

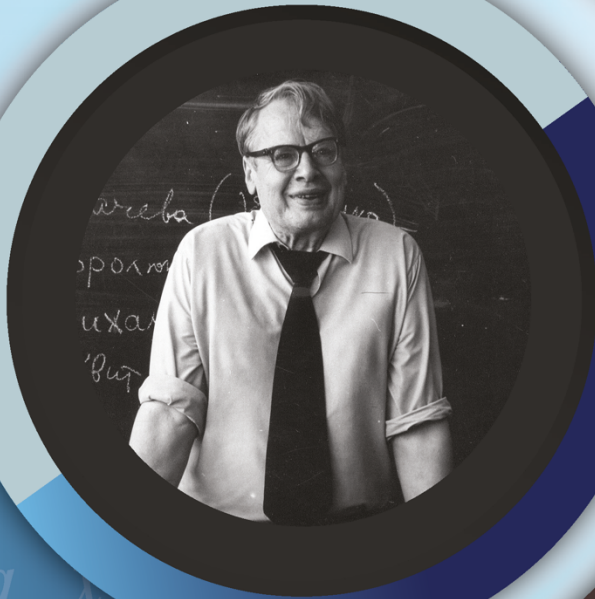
RTA

ISSN 1932-2321

JOURNAL IS REGISTERED
IN THE LIBRARY OF THE
U.S. CONGRESS

RELIABILITY:
THEORY & APPLICATIONS

INTERNATIONAL
GROUP ON
RELIABILITY



GNEDENKO FORUM PUBLICATIONS

#3

(79) VOL.19
SEPTEMBER
2024

SAN DIEGO

RELIABILITY

RISK ANALYSIS

MAINTENANCE

SAFETY

ISSN 1932-2321

© "Reliability: Theory & Applications", 2006, 2007, 2009-2024

© " Reliability & Risk Analysis: Theory & Applications", 2008

© I.A. Ushakov

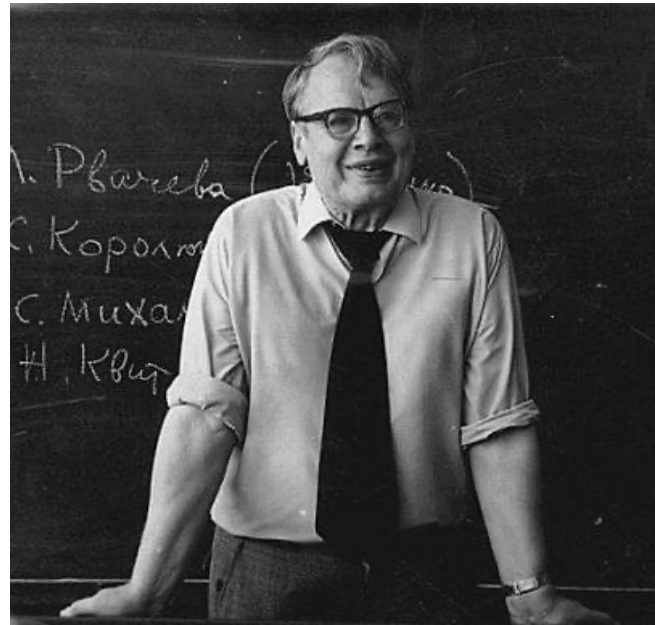
© A.V. Bochkov, 2006-2024

© Kristina Ushakov, Cover Design, 2024

<http://www.gnedenko.net/Journal/index.htm>

All rights are reserved

The reference to the magazine "Reliability: Theory & Applications"
at partial use of materials is obligatory.



RELIABILITY: THEORY & APPLICATIONS

Vol.19 No.3 (79),
September 2024

San Diego
2024

Editorial Board

Editor-in-Chief

Rykov, Vladimir (Russia)

Doctor of Sci, Professor, Department of Applied Mathematics & Computer Modeling, Gubkin Russian State Oil & Gas University, Leninsky Prospect, 65, 119991 Moscow, Russia.
e-mail: vladimir_rykov@mail.ru

Managing Editors

Bochkov, Alexander (Russia)

Doctor of Technical Sciences, Scientific Secretary JSC NIIAS, Scientific-Research and Design Institute Informatization, Automation and Communication in Railway Transport, Moscow, Russia, 107078, Orlikov pereulok, 5, building 1
e-mail: a.bochkov@gmail.com

Gnedenko, Ekaterina (USA)

PhD, Lecturer Department of Economics Boston University, Boston 02215, USA
e-mail: gnedenko@bu.edu

Bushinskaya, Anna (Russia)

Candidate of Tech. Sci. Leading Research Fellow of the Sci & Engng Center of the Russian Academy of Sciences, Ekaterinburg
e-mail: bushinskaya@gmail.com

Sazonov, Aleksey (Russia)

Leading Specialist of the Standardization Department, JSC NIIAS (Joint Stock Company "Design & Research Institute for Information Technology, Signaling and Telecommunications on Railway Transport), Bild 1, 5 Orlikov Pereulok, Moscow, Russia, 107078
e-mail: sazono2007@gmail.com

Deputy Editors

Dimitrov, Boyan (USA)

Ph.D., Dr. of Math. Sci., Professor of Probability and Statistics, Associate Professor of Mathematics (Probability and Statistics), GMI Engineering and Management Inst. (now Kettering)
e-mail: bdimitro@kettering.edu

Gnedenko, Dmitry (Russia)

Doctor of Sci., Assos. Professor, Department of Probability, Faculty of Mechanics and Mathematics, Moscow State University, Moscow, 119899, Russia
e-mail: dmitry@gnedenko.com

Kashtanov, Victor A. (Russia)

PhD, M. Sc (Physics and Mathematics), Professor of Moscow Institute of Applied Mathematics, National Research University "Higher School of Economics" (Moscow, Russia)
e-mail: VAKashtan@yandex.ru

Krishnamoorthy, Achyutha (India)

M.Sc. (Mathematics), PhD (Probability, Stochastic Processes & Operations Research), Professor Emeritus, Department of Mathematics, Cochin University of Science & Technology, Kochi-682022, INDIA.
e-mail: achyuthacusat@gmail.com

Recchia, Charles H. (USA)

PhD, Senior Member IEEE Chair, Boston IEEE Reliability Chapter A Joint Chapter with New Hampshire and Providence, Advisory Committee, IEEE Reliability Society
e-mail: charles.recchia@macom.com

Shybinsky Igor (Russia)

Doctor of Sci., Professor, Division manager, VNIAS (Russian Scientific and Research Institute of Informatics, Automatics and Communications), expert of the Scientific Council under Security Council of the Russia
e-mail: igor-shubinsky@yandex.ru

Yastrebenetsky, Mikhail (Ukraine)

Doctor of Sci., Professor. State Scientific and Technical Center for Nuclear and Radiation Safety
e-mail: ma.yastreb2013@gmail.com

Associate Editors

Aliyev, Vugar (Azerbaijan)

Doctor of Sci., Professor, Chief Researcher of the Institute of Physics of the National Academy of Sciences of Azerbaijan, Director of the AMIR Technical Services Company
e-mail: prof.vugar.aliyev@gmail.com

Balakrishnan, Narayanaswamy (Canada)

Professor of Statistics, Department of Mathematics and Statistics, McMaster University
e-mail: bala@mcmaster.ca

Carrión García, Andrés (Spain)

Professor Titular de Universidad, Director of the Center for Quality and Change Management, Universidad Politécnica de Valencia, Spain
e-mail: acarrion@eio.upv.es

Chakravarthy, Srinivas (USA)

Ph.D., Professor of Industrial Engineering & Statistics, Departments of Industrial and Manufacturing Engineering & Mathematics, Kettering University (formerly GMI-EMI) 1700, University Avenue, Flint, MI48504
e-mail: schakrav@kettering.edu

Cui, Lirong (China)

PhD, Professor, School of Management & Economics, Beijing Institute of Technology, Beijing, P. R. China (Zip:100081)
e-mail: lirongcui@bit.edu.cn

Finkelstein, Maxim (SAR)

Doctor of Sci., Distinguished Professor in Statistics/Mathematical Statistics at the UFS. Visiting researcher at Max Planck Institute for Demographic Research, Rostock, Germany and Visiting research professor (from 2014) at the ITMO University, St Petersburg, Russia
e-mail: FinkelM@ufs.ac.za

Kaminsky, Mark (USA)

PhD, principal reliability engineer at the NASA Goddard Space Flight Center
e-mail: mkaminskiy@hotmail.com

Krivtsov, Vasiliy (USA)

PhD. Director of Reliability Analytics at the Ford Motor Company. Associate Professor of Reliability Engineering at the University of Maryland (USA)
e-mail: VKrivtso@Ford.com_krivtsov@umd.edu

Lemeshko Boris (Russia)

Doctor of Sci., Professor, Novosibirsk State Technical University, Professor of Theoretical and Applied Informatics Department
e-mail: Lemeshko@ami.nstu.ru

Lesnykh, Valery (Russia)

Professor, Doctor of Sci., Adviser to Director General, LLC Gazprom gaznadzor, Novocheryomushkinskaya Street, 65, Moscow, 117418, Russia
e-mail: vvlesnykh@gmail.com

Levitin, Gregory (Israel)

PhD, The Israel Electric Corporation Ltd. Planning, Development & Technology Division. Reliability & Equipment Department, Engineer-Expert; OR and Artificial Intelligence applications in Power Engineering, Reliability.
e-mail: levitin@iec.co.il

Limnios, Nikolaos (France)

Professor, Université de Technologie de Compiègne, Laboratoire de Mathématiques, Appliquées Centre de Recherches de Royallieu, BP 20529, 60205 COMPIEGNE CEDEX, France
e-mail: Nikolaos.Limnios@utc.fr

Papic, Ljubisha (Serbia)

PhD, Professor, Head of the Department of Industrial and Systems Engineering Faculty of Technical Sciences Cacak, University of Kragujevac, Director and Founder the Research Center of Dependability and Quality Management (DQM Research Center), Prijedor, Serbia
e-mail: dqmcenter@mts.rs

Ram, Mangey (India)

Professor, Department of Mathematics, Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun, India. Visiting Professor, Institute of Advanced Manufacturing Technologies, Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia.
e-mail: mangeyram@gmail.comq

Timashev, Sviatoslav (Russia)

Doctor of Sci., Professor, Director and principal scientist the Sci & Engng Center of the Russian Academy of Sciences, Ekaterinburg
e-mail: timashevs@cox.net

Zio, Enrico (Italy)

PhD, Full Professor, Direttore della Scuola di Dottorato del Politecnico di Milano, Italy.
e-mail: Enrico.Zio@polimi.it

e-Journal *Reliability: Theory & Applications* publishes papers, reviews, memoirs, and bibliographical materials on Reliability, Quality Control, Safety, Survivability and Maintenance.

Theoretical papers must contain new problems, finger practical applications and should not be overloaded with clumsy formal solutions.

Priority is given to descriptions of case studies.
General requirements for presented papers.

1. Papers must be presented in English in MS Word or LaTeX format.
2. The total volume of the paper (with illustrations) can be up to 15 pages.
3. A presented paper must be spell-checked.
4. For those whose language is not English, we kindly recommend using professional linguistic proofs before sending a paper to the journal.

The manuscripts complying with the scope of journal and accepted by the Editor are registered and sent for external review. The reviewed articles are emailed back to the authors for revision and improvement.

The decision to accept or reject a manuscript is made by the Editor considering the referees' opinion and considering scientific importance and novelty of the presented materials. Manuscripts are published in the author's edition. The Editorial Board are not responsible for possible typos in the original text. The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Publication in this e-Journal is equal to publication in other International scientific journals.

Papers directed by Members of the Editorial Boards are accepted without referring. The Editor has the right to change the paper title and make editorial corrections.

The authors keep all rights and after the publication can use their materials (re-publish it or present at conferences).

Send your papers to Alexander Bochkov, e-mail: a.bochkov@gmail.com

Table of Contents

MY WAY TO RELIABILITY AND SAFETY 27

M. Yastrebenetsky

I was born in Moscow on September 10, 1934. I wrote about my family and previous generations of Yastrebenetsky (a soldier under Emperor Nicholas I, a businessman, a doctor, a chemical engineer) in my book "Generations of Yastrebenetsky" [1]. After finishing the institute and postgraduate studies, 2/3 years of my life were connected with activity in the fields of reliability and safety of technical systems. This activity is the subject of the article.

ALGORITHMS FOR APPROXIMATING A FUNCTION BASED ON INACCURATE OBSERVATIONS..... 41

G.Sh. Tsitsiashvili, M.A. Osipova

This paper is devoted to the approximation of a function by a trigonometric polynomial based on its inaccurate values at selected points. Two methods of observation are considered. The first method is to make observations at points evenly distributed on the segment where the function is specified. The second method is to take observations at the points of division into a finite number of equal parts of the neighbourhoods of the selected points. Upper estimates of the standard deviation of the function from trigonometric polynomials are constructed and the rate of their convergence is estimated. Differences were found in the computational complexity of these approximations and in the number of observations of the function values at the selected points. Thus, the problem of approximating a function from inaccurate observations of their values at selected points is a multi-criteria one and its solution depends on the choice of observation points.

AVAILABILITY AND PROFIT OPTIMIZATION OF CONTINUOUS CASTING SYSTEM OF THE STEEL INDUSTRY USING ARTIFICIAL NEURAL NETWORK TECHNIQUE 47

Shikha Bansal, Sohan Lal Tyagi, Urvashi

The main purpose of this study was to optimize the performance parameters of the casting process using a neural network approach. The casting process is molten or liquefied metal is poured into the mould cavity, after solidification it takes the near net shape of the cavity. The entire manufacturing process goes through six stations viz Pouring turret "ladle", "Tundish", "Mould", "Water spray chamber", "Support roller" and "Torch cutter". The Artificial Neural Network (ANN) technique is used in this paper to analyze the casting system's availability, profitability, and state probability variation. An effort has been made to identify the most critical component of the system. The outcomes of the analysis will help the practitioners in deciding effective maintenance strategies.

TOPP-LEONE EXPONENTIATED GOMPERTZ INVERSE RAYLEIGH DISTRIBUTION: PROPERTIES AND APPLICATIONS 59

Sule Omeiza Bashiru, Alaa Abdulrahman Khalaf, Alhaji Modu Isa

This paper focused on deriving a new lifetime distribution having five parameters by compounding the Gompertz inverse Rayleigh model and the Topp-Leone exponentiated-G family of distributions. The new model is called Topp-Leone exponentiated Gompertz inverse Rayleigh (TLEGoIRa) distribution. The new model is very flexible and the shape of its pdf can be positively or negatively skewed and symmetric. Some statistical characteristics of the new model, such as the moments, incomplete moments, quantile function, rényi entropy and order statistics are derived and investigated. The pdf of the minimum and maximum order statistics of the new model were

derived and studied. The model's parameters are estimated using the maximum likelihood approach. A simulation study was conducted to investigate the consistency of the newly proposed model, using the average bias and root mean square error (RMSE) as metrics. The outcome of the simulation suggested that as sample sizes increase, both the average bias and root mean square error (RMSE) decrease, indicating that the distribution is consistent. Finally, two real-life datasets were used to explore the new model's importance and adaptability in comparison to other competing models. The results of the application revealed that the new distribution outperforms its competitors.

A PROBABILITY MODEL FOR SURVIVAL ANALYSIS OF CANCER PATIENTS 78

Mousumi Ray, Rama Shanker

It has been observed by statistician that to find a suitable model for the survival analysis of cancer patients is really challenging. The main reasons for that is the highly positively skewed nature of datasets. During recent decades several statistician tried to propose one parameter, two-parameter, three-parameter, four-parameter and five-parameter probability models but due to either theoretical or applied point of view the goodness of fit provided by these distributions are not very satisfactory. In this paper a compound probability model called gamma-Sujatha distribution, which is a compound of gamma and Sujatha distribution, has been proposed for the modeling of survival times of cancer patients. Many important properties of the suggested distribution including its shape, moments (negative), hazard function, reversed hazard function, quantile function have been discussed. Method of maximum likelihood has been used to estimate its parameters. A simulation study has been conducted to know the consistency of maximum likelihood estimators. Two real datasets, one relating to acute bone cancer and the other relating to head and neck cancer, has been considered to examine the applicability, suitability and flexibility of the proposed distribution. The goodness of fit of the proposed distribution shows quite satisfactory fit over other considered distributions.

ON A DISCRETE TIME SHOCK MODEL IN CRITICAL SITUATION..... 95

Reza Farhadian, Habib Jafari

In this paper, we study a discrete time shock model which is defined based on the length of the time between successive shocks. For a system that is exposed to a sequence of random shocks over time under this model, if the interarrival time between two successive shocks is equal to a prefixed critical time point such as δ , the system fails, and the system is not damaged otherwise. We have considered two situations for the system, which are regular situation and critical situation, then we investigate the statistical behavior of the system's lifetime under these situations. More precisely, we obtain the probability generating function of the system's lifetime, the mean time to failure, the variance of the system's lifetime, the Laplace transform of the system's lifetime, and some other related results. We end the paper with an example including numerical comparisons of the results.

ENHANCED METHODS UNDER EXPONENTIAL DISTRIBUTION CONCERN WITH EWMA AND DEWMA METHODS..... 107

Dr. V. Kaviyarasu and M. Inbarasi

Statistical Process Control plays a crucial part in improving the quality and lowering the fluctuation in the production process environment. In SPC, the most popularly used methods are Shewhart control chart techniques and EWMA techniques which distinguish itself for its quick identification of minute process deviations, which makes it an essential tool for guaranteeing product. EWMA methods detect variances in quality of the product as well as services, measure process mean shifts with control charts, and track manufacturing process parameters to find deviations and make necessary adjustments. The exponential distribution was employed in this study because it may reflect vast and bulk production in everyday life. Exponentially distributed data, evaluate it alongside the EWMA function. This paper's objective is to study the impact of EWMA & DEWMA parameters within the EWMA control chart's performance using exponential distribution. Further, A few tables are provided with suitable illustrations that can be available with parameters with the help of these findings. The study also examines how the EWMA parameter affects the shape of the distribution.

RELIABILITY SAMPLING PLAN FOR GENERALIZED INVERTED EXPONENTIALLY DISTRIBUTED UNDER PROGRESSIVE TYPE-II CENSORED DATA 120

S. Singh, A. Kaushik

This article aims to explore a sampling strategy designed to assess the reliability of products that exhibit lifetimes following a GIED. Considered sampling approach has been specially constructed for a Type-II progressive censoring scheme, which includes binomial removals as part of its methodology. Its core objectives is to find out acceptance constant and the optimum sample size. To facilitate practical implementation, the article presents a tabulated form of the sampling plan for the selected specification, as per the considered censoring scheme. To validate the dependability and precision of the suggested sampling approach, we perform a Monte Carlo experiment under various scenarios.

ROLE OF AGEING METRICS TO ANALYSE THE SURVIVAL DATA OF TONGUE CANCER PATIENTS 132

B. Elina, Pulak Swain, Satya Kr. Misra, Subarna Bhattacharjee

The paper vividly describes the non-parametric estimation of basic quantities for right censored data of times to death for patients with tongue cancer. Here we compare patients with two different sets of DNA profile using several parameters like reliability function, cumulative hazard rate function, smoothed hazard rate function and ageing intensity function. With the help of graphical representations of these functions, we analyse which DNA profile patients have better prognosis.

STATISTICAL PROPERTIES AND APPLICATIONS OF TRANSMUTED SKEW STUDENT t DISTRIBUTION 144

David Ikwuoche John, Mathew Stephen

In this study, a modified 2-parameter skew t distribution called the transmuted skew student t distribution (TSSStD) was presented. Some statistical and reliability properties of TSSStD such as the quantile function, the raw moments, and the moment generating function (among others), were derived. Through the method of maximum likelihood, the two parameters of the model were estimated. The stability of the model was studied via Montecarlo simulations utilizing bias, mean square error, and root mean square error as metrics. The results from the stability study revealed that the TSSStD was well-behaved. Four datasets were modeled with the transmuted skewed student t distribution and four other probability density models. On the basis of information criteria, the results revealed that the transmuted skew student t distribution provides a better fit for all the datasets compared to the other competing models.

SINE-TOPP-LEONE EXPONENTIATED G FAMILY OF DISTRIBUTIONS: PROPERTIES, SURVIVAL REGRESSION AND APPLICATION 157

A. M. Isa, S. I. Doguwa, B. B. Alhaji, H. G. Dikko

In this research article, we have introduced a new class of continuous probability distributions known as the Sine Topp-Leone Exponentiated-G family of distributions. This newly proposed family exhibits a higher degree of flexibility compared to some of the established distribution families. The various models within this family find wide-ranging applications in fields such as physics, engineering, and medicine. Some statistical properties of the Sine Topp-Leone Exponentiated-G family of distributions such as moments, moment generating function, quantile function and order statistics are derived. Two special models were also presented and studies. Maximum likelihood estimation method was used to estimate parameters of the models. The consistency of the proposed family was determine using simulation studies. Two real life datasets were analyzed to show the flexibility of the proposed model and the results of the analysis showed that, the proposed model was more efficient and best fit the data sets than its competitors.

ANALYSIS OF RELIABILITY OF TYPICAL POWER SUPPLY CIRCUITS..... 173

S.V. Rzayeva, N.M. Piriyeva, I.A. Guseynova

This article is devoted to the analysis of the reliability of typical power supply circuits. The question of what factors have the greatest impact on the reliability of power supply circuits, as well as what methods and tools are used to analyze and improve their reliability is considered. Particular attention is paid to the comparative analysis of various types of power supply schemes and the determination of their advantages and disadvantages in terms of reliability. The article also discusses current trends and developments in the field of increasing the reliability of power supply and possible ways to optimize existing circuits. The results obtained can be useful for specialists in the field of power engineering and electrical engineering in the design, maintenance and modernization of power supply systems.

FUZZY LOGIC RELIABILITY BLOCK DIAGRAM APPROACH FOR PATIENT HEALTH MONITORING USING R PROGRAMMING 179

Liji Sebastian, Rita S, Vennila J

In this research, a new approach using fuzzy logic and reliability block diagram (RBD) techniques is used to ensure the reliability of patient health monitoring systems. This technique handles uncertainties in health information, while RBD assesses system reliability by displaying factor relations. The RBD model construct for system components and measure reliability using probabilistic models. Fuzzy logic identifies the effect of uncertainties on overall reliability. Using this approach in a simulated health monitoring scenario, using R, we demonstrate its effectiveness and potential to increase reliable health monitoring for improved patient outcomes and healthcare efficiency. Furthermore, the awareness gained from this study can be directed beyond healthcare such as modern process control and environmental sensing.

RELIABILITY ANALYSIS OF AN ANTI-DRONE SYSTEM BY CONSIDERING RANDOM ENVIRONMENTAL FACTORS 186

Dharmaraja Selvamuthua, Harshita Badiyasara, Smrati Tripathia, Priyanka Kalitab, Raina Rajc

In today's security landscape, the proliferation of unauthorized drones in restricted airspace has emerged as a significant threat. These drones pose various risks, from potential surveillance and espionage to more sinister possibilities such as physical attacks. Consequently, the development of effective anti-drone laser systems has become increasingly vital. Our study focuses on three main objectives: modeling internal reliability, identifying critical components, and studying the factors affecting the reliability of anti-drone systems. We aim to enhance the overall performance and effectiveness of anti-drone laser systems by analyzing the reliability of critical components and understanding how system parameters influence system reliability. To this end, reliability block diagram (RBD) methodology has been employed to compute the reliability of the laser subsystem in the anti-drone system. Additionally, we conduct a comprehensive review of component-wise reliability to identify vulnerable points within the system, thus enabling targeted improvements and optimizations. To capture the realistic scenario of system failure behavior, different distributions have been used to compute the reliability of the system, ensuring a thorough understanding of its operational reliability in diverse conditions. Finally, the energy values and probability of hitting are obtained for the anti-drone laser system to effectively mitigate environmental challenges.

STRATEGIES FOR REPLACEMENT IN WORKFORCE SCHEDULING RELIABILITY MODELS 206

Iyappan. M, Balaji. M, R. Saranraj, G. Sathya Priyanka

It is a typical occurrence to replace some industrial equipment or components, such as electronic chips, bulbs, etc. It deals with the ideas of dependability theory, in which the likelihood of an equipment malfunctioning instantly is calculated assuming that it has operated normally for a given amount of time, 't'. In reliability, it's

known as the hazard rate. The rate of hazard may be rising, falling, or staying the same. However, replacement tactics are employed to maintain output. Basically, two distinct approaches are employed. 1. Replacing a broken item; 2. Replacing items on a regular basis. Since it is a preventive measure to maintain production, the effective administration of maintaining system functionality depends on the application of both reliability concepts in addition to replacement theory. One may envision a similar issue with the personnel system as well. To determine the best manpower policies in this situation, replacement methods and dependability theory can also be coupled. This theory's application to labor systems is examined, and appropriate methods for employee replacement and advancement are covered in order to ensure the system's successful upkeep. In order to obtain the best fit, manpower planning is a dynamic process that controls the movement of workers into, through, and out of the organization. In order to determine appropriate manpower replacement policies for promotion and replacement of personnel for the successful maintenance of the system, this study discusses the application of dependability theory and the renewal process.

PERFORMANCE ANALYSIS OF M[X]/GB/1 FEEDBACK RETRIAL QUEUE WITH VARIABLE SERVER MODEL 215

N. Micheal Mathavavisakan, K. Indhira

In this article, a working vacation policy-based on bulk arrival feedback retrial queueing system with variable server capacity has been analyzed. The server can serve a minimum of one customer and a maximum of B customers in a batch in accordance with the variable server capacity bulk service rule. As soon as the orbit becomes empty at the time of service completion, the server goes for a working vacation. The server works at a lower speed during a working vacation period. In addition, the steady state probability generating function for system size and orbit size is generated by incorporating the supplementary variables technique (SVT). Further, the conditional decomposition law is shown for this retrial queueing system. Moreover, system performance metrics, and significant special instances are discussed. Finally, the effects of various parameters on the system performance are analyzed numerically.

STOCHASTIC MODELING AND PERFORMABILITY ANALYSIS OF REPAIRABLE SYSTEM OF A PLYWOOD INDUSTRY 230

Mr. Amit Kumar Singh, Dr. P. C. Tewari

The current paper analyzes the performance behavior concerning the performability of the Veneer layup system in a plywood industry. A Markovian Approach is utilized to develop a process model for the system and enhance to evaluate system performability i.e. the function of system availability. The study investigates the impact of varying failure and repair rates on the availability of system, variation in the availability is also determined by varying available repair facilities, using a licensed software package. Particle Swarm Optimization (PSO) method has been employed to optimize the results. Additionally, a Decision Support System (DSS) has been proposed for making strategic decisions regarding financial investments and maintenance order priorities. The findings of the paper will aid the practitioners in deciding the maintenance order priorities among various subsystems.

E-BAYESIAN ESTIMATION FOR BATHTUB-SHAPED LIFETIME DISTRIBUTION BASED ON UPPER RECORD VALUES 242

Sana, M. Faizan

In this research paper, we presents the expected Bayesian (E-Bayesian) estimation of bathtub-shaped lifetime (BSL) distribution for scale parameter based on upper record values (URV) using a conjugate prior distribution. Also, we are considered different prior distributions for the E-Bayesian estimators. Some properties of the E-Bayesian estimators are discussed. A simulation study is given to compare the performance of the E-Bayesian estimators with Bayesian estimator. we notice that the E-Bayesian estimators are perform better than the Bayesian estimators. Moreover, the performance of the Bayesian estimators and E-Bayesian estimators for Prior II are better than Prior I. Also, we observe that if we increase the sample size n then the estimators are showing lesser mean square error (MSE).

A CLASS OF LOGARITHMIC-CUM-EXPONENTIAL ESTIMATORS FOR POPULATION MEAN WITH RISK ANALYSIS USING DOUBLE SAMPLING..... 248

Diwakar Shukla, Astha Jain

In order to improve upon the efficiency of an estimate in double sampling for estimating population mean of character under study using an auxiliary variable, a part of survey resources are used to collect the information on auxiliary variable. Some authors have suggested exponential-type estimators and some others advocated for log-type estimators. But combination of such is required for specific situation. This paper presents a class of logarithmic-cum-exponential ratio estimators in double sampling setup. The expressions for the mean squared error and bias of the proposed class of estimators are derived for two different cases (sub-sample and independent sample). Sometimes the persons involved in the sample survey have to undergo for risk on life. For example, data collection in naxalites area, working in intense forest, interview during spread of epidemic or data collection in politically disturbed region. Such risk may affect the accuracy, efficiency of estimation. A linear Risk function is used for the proposed class of estimators. Two cases of double sampling are compared in terms of relative efficiency in view to risk aspect. It is found that the proposed class of estimators has a lower mean squared error than the simple mean estimator, usual ratio, usual exponential, usual log estimators in the double sampling setup. In addition, these theoretical results are supported by a numerical example. Risk function based simulated study is performed for the support of findings of the content. Optimal sample sizes under risk are derived and compared under two cases.

STOCHASTIC ANALYSIS OF A GAS TURBINE SYSTEM WITH PRIORITY AND RANDOM INSPECTION BY SINGLE SERVER UNDER DIFFERENT HUMID CONDITIONS..... 263

Pinki, Vijeta Kumari, Dalip Singh

In this study, we investigated the impact of two different humid levels on the reliability measures of a stochastic model for a gas turbine system composed of a gas turbine and a steam turbine. To enhance the system's overall performance, we prioritize gas turbine repair over steam turbine repair in addition to a combined inspection and preventative maintenance approach. To find some reliability measures, such as the mean time to system failure, availability, etc., semi-Markov process and regenerating point technique are utilized. These measures are analysed graphically based on the data obtained from a gas turbine power plant in Delhi, India.

PYTHON IMPLEMENTATION OF FUZZY LOGIC FOR ZERO-INFLATED POISSON SINGLE SAMPLING PLANS..... 275

Kavithanjali S, Sheik Abdullah A

Acceptance sampling is used in Statistical Quality Control (SQC) to conduct lot quality evaluations through sample inspections which involve probability theory and fuzzy sets. It aims to optimize quality, costs, and productivity, frequently applying linguistic variables when accurate parameter values are not good enough which is handled using fuzzy set theory. This research analyses single sampling plans (SSP) in the presence of fuzzy number non-conformities, modelling them with the Zero-inflated Poisson (ZIP) distribution structure. This study presents a unique method to single sampling plans (SSP) inside the Zero-inflated Poisson (ZIP) distribution framework that makes use of fuzzy logic approaches. In addition, we show how to apply this method using a Python programme, providing practical suggestions for real-world quality control complications.

ANALYSIS OF NON-MARKOVIAN BATCH ARRIVAL RETRIAL QUEUE WITH PRIORITY SERVICES, IMMEDIATE FEEDBACK, PUSH OUT, DIFFERENTIATED BREAKDOWNS, DELAYED REPAIR, RANDOMIZED VACATION..... 282

G. Ayyappan, S. Nithya

Priority and ordinary customers arrive according to Poisson processes, and their service time based on the general distribution. The server constantly offers a single service for both priority and ordinary customers. We compute the Laplace transforms of the time-dependent probabilities of system states using the probability generating function and supplementary variable technique. Numerical results are obtained which are also examined to facilitate the sensitivity analysis of system descriptions.

PERSONALIZED FEATURES-BASED STRESS DETECTION WITH HYPERPARAMETER TUNING USING GENETIC ALGORITHM..... 298

Jigna Jadav, Uttam Chauhan

In recent years, there have been considerable improvements in how we keep track of mental health, especially with devices you can wear, which give us a better chance of spotting and dealing with problems like stress before they become serious. This research paper presents an innovative approach. Experimental validation uses a comprehensive dataset of 15 subjects working as multinational company employees. Heart Rate Variability(HRV) was obtained from wearable sensors using Apple Watch during working hours. We have calculated time, frequency and non-linear domains as well and added personalized features like a person's age, height, weight, etc. Recurrent Neural Network(RNN)and Long Short-Term Memory (LSTM)models are applied and get an accuracy of 87% and 90%, respectively. To enhance stress detection accuracy by optimizing hyperparameters using a genetic algorithm (GA) explicitly targeting the configuration of LSTM models. Key hyperparameters, including the number of units in the LSTM layer and the number of training epochs, are optimized to maximize stress detection accuracy. Model Through 5 generations of evolution, the GA identifies optimal hyperparameter settings of 45 units in the LSTM layer 49 epochs, significantly improving stress detection accuracy compared to baseline configurations. It gives 92 % accuracy with optimized hyperparameters. Analyzing recorded data, we observe that the time per training step decreases gradually, indicating efficient convergence during optimization. Simultaneously, stress detection accuracy steadily improves over epochs, showcasing the model's effectiveness in learning patterns from physiological data. So, This study provides insights into the practical application of genetic algorithms for hyperparameter optimization in healthcare contexts, contributing to advancements in personalized monitoring and intervention strategies for mental well-being.

ANALYSIS OF SINGLE SERVER FEEDBACK RETRIAL QUEUE WITH BERNOULLI WORKING VACATION AND STARTING FAILURE..... 310

Keerthiga S, Indhira K

The suggested queueing model describes a single-server feedback retrial queueing system with starting failure, Bernoulli working vacation and vacation interruptions. The server departs on a working vacation as soon as orbit is empty. During the working vacation period, the server provides a slower level of service. The supplementary variable method was utilized to determine the steady-state probability-generating functions for the system and its orbit. If there are consumers in the system at the end of each vacation, the server becomes idle and ready to serve new customers. The average busy time and the average busy cycle are presented as important system performance indicators. Additionally, the adaptive neuro-fuzzy interface system has compared the numerical results with the neuro-fuzzy results. Finally, particle swarm optimization (PSO) were utilized to obtain the best (optimal) cost for the system in this study. We have examined the convergence of these optimization strategies.

IMPROVING THE SPECTRAL EFFICIENCY IN DOWNLINK MULTIPLE USER MULTIPLE INPUT MULTIPLE OUTPUT TRANSMISSION FOR FIFTH GENERATION AND BEYOND WIRELESS COMMUNICATIONS..... 327

Abdulmujeeb Akajewole Masud, Donatus Uchechukwu Onyishi

This research presents a solution in Multiple Input Multiple Output (MIMO) wireless systems to meet the growing demand for high data rates in cellular networks. Although MIMO systems offer greater capacity, the higher frequencies used have caused interference problems, especially for mobile User Equipment (UE). This research aims to reduce interference problems in the downlink of Multi-user MIMO (MU-MIMO) systems, with a specific focus on improving Quality of Service (QoS) metrics, such as outage probability and Signal-to-Interference plus Noise Ratio (SINR). Existing solutions to these challenges are complex due to the dynamic nature of the factors involved in modelling real-world scenarios. As such, an Improved Downlink MU-MIMO (ID-MU-MIMO) algorithm is developed as a solution to these problems. The ID-MU-MIMO method employs both single antenna users and multiple transmitter antennas. The performance of the suggested algorithm is compared to the IEEE 802.11ax standard specification and a previous research work for validation and evaluation. Performance measures considered to aid validation included outage probability, spectrum efficiency, and communication connection reliability. On this premise, the outcomes showed that the proposed ID-MU-MIMO scheme outperforms both the IEEE 802.11ax standard and current MD-MU-MIMO systems. In particular, compared to IEEE 802.11ax, the ID- MU-MIMO technique achieved a 7.71% reduction in interference. When compared to the performance of the random and uniform MD-MU-MIMO algorithms, the proposed ID-MU- MIMO scheme showed a reduction in interference in percentages of 8.90% and 2.28%, respectively. The ID-MU-MIMO scheme outperformed the random and uniform MD-MU-MIMO algorithms in terms of Signal-to-Interference Noise Ratio (SINR), outperforming them by 4.27% and 2.75%, respectively, and resource block use, outperforming them by 20.05% and 3.89%, respectively.

ENHANCING PROCESS CAPABILITY ANALYSIS FOR LOGNORMAL DATA UTILIZING BOX COX TRANSFORMATION AND GOODNESS OF FIT TESTS 343

J. Krishnan, R. Vijayaraghavan

Process capability analysis is a valuable tool in quality assurance, but deviations from normal distribution necessitate adjustments to basic process capability indices. Process control literature offers solutions for non-normality, with data transformation being a common approach. The Box- Cox transformation (BCT) is often used to normalize non-normal data, relying on maximum likelihood estimation (MLE) to determine the transformation parameter, lambda. Alternative methods exist for estimating the single transformation parameter lamda, employing goodness-of-fit tests instead of the MLE method. This study explores two expressions within the Box-Cox transformation (BCT), encompassing both optimal and rounded values of lambda. The primary goal is to identify an effective method for transforming non-normal data into a distribution closer to normality through goodness-of-fit tests, aiming to obtain accurate estimates for process capability analysis in alignment with six sigma standards. Furthermore, this study focuses on the influence of utilizing both optimal and rounded values of lambda when transforming non-normal data to normal, and how these lambda values impact the estimates of process capability analysis. The findings reveal that methods such as Shapiro-Wilk's (SW) and Artificial Covariate (AC) outperform the MLE method. Moreover, employing the optimal lambda value during data transformation leads to improved estimates of process capability. Data simulation and analysis were conducted using Minitab software and the R programming language.

INFERENCE ON THE INVERSE POWER BURR-HATKE DISTRIBUTION UNDER TYPE II CENSORING 356

Pavitra Kumari, Vinay Kumar

There are many real-life situations, where data require probability distribution function which have decreasing or upside-down bathtub (UBT) shaped failure rate function. The inverse power burr hatke distribution consists both decreasing and UBT shaped failure rate functions. Here, we address the different estimation methods of the parameter and reliability characteristics of the inverse Pareto distribution from both classical and Bayesian approaches. We consider classical estimation procedures to estimate the unknown parameter of inverse power burr-hatke distribution, such as maximum likelihood. Also, we consider Bayesian estimation using squared error loss function based joint priors. The Monte Carlo simulations are performed to compare the performances of the obtained estimators in mean square error sense. Finally, the flexibility of the proposed distribution is illustrated empirically using one real-life datasets. The analyzed data shows that the introduced distribution provides a superior fit than some important competing distributions such as the Weibull, inverse Pareto and Burr-Hatke distributions.

ANALYSIS OF TWO NON-IDENTICAL UNIT SYSTEM HAVING SAFE AND UNSAFE FAILURES WITH REBOOTING AND PARAMETRIC ESTIMATION IN CLASSICAL AND BAYESIAN PARADIGMS 364

Poonam Sharma, Pawan Kumar

The present paper aims at the study of a two non-identical system model having safe and unsafe failures and rebooting. The focus centers on the analysis w.r.t important reliability measures and estimation of parameters in Classical and Bayesian paradigms. At first one of the units is operational whereas other one is confined to standby mode. Any unit may suffer safe or unsafe failure. A safe failure is immediately taken up for remedial action by a repairman available with the system all the time, while the case of unsafe failure cannot be dealt directly but first rebooting is performed to convert the unsafe failure to safe failure mode so as to start repair normally. A switching device is used to make the repaired and standby units operational. The lifetime of both the units and switching device are taken to be exponentially distributed random variables whereas the distribution of repair times are assumed to be general. Regenerative point technique is employed to derive associated measures of effectiveness. To make the study more elaborative and visually attractive, some of the derived characteristics have been studied graphically too. A simulation study has also been undertaken to exhibit the behaviour of obtained characteristics in Classical and Bayesian setup. Valuable inferences about MLE and Bayes estimates have been drawn from the tables and graphs for varying values of failure and repair parameters.

RELIABILITY ESTIMATION OF STRESS-STRENGTH MODEL USING FUZZY DISTORTION FUNCTION UNDER UNCERTAINTY IN ENVIRONMENTAL FACTORS 380

K Sruthi, M Kumar

In the reliability estimation of stress-strength models, external factors such as temperature, humidity, etc. may influence the distribution of stress and strength random variables. In traditional reliability analysis, these external factors are accounted for by introducing a real-valued distortion function, which replaces the original distribution with a distorted one. However, it's important to note that the effect of these external factors is not always adequately represented by a single real-valued function. To address this issue, we propose the use of fuzzy numbers within the distortion function. In this paper, we introduce the concept of a "fuzzy distortion function" to incorporate the uncertainty stemming from external factors when estimating the reliability of stress-strength relationships. We present a methodology for estimating fuzzy reliability by employing this fuzzy distortion function. Through an illustrative example, we demonstrate how this approach to estimating fuzzy reliability offers a wider range of possibilities for system reliability and provides more comprehensive insights into the system's behaviour. Throughout our exploration, we have delved into the diverse properties inherent in fuzzy distortion functions. These properties highlight the versatility and adaptability of such functions in capturing

uncertainty within data sets. Moreover, we have scrutinized several methods for constructing fuzzy distortion functions from pre-existing ones. By examining these methods, we gain valuable insights into how fuzzy distortion functions can be tailored to specific contexts and applications, thereby enhancing the accuracy and robustness of reliability analysis in complex systems. Additionally, in the conventional stress-strength model, reliability is determined without considering the uncertainty in the parameters of the distribution function. The drawback of existing methods in the literature is that they do not consider the uncertainty or fuzziness in the parameters of the distribution. Therefore, we estimate the system reliability in the presence of fuzzy parameters in the distribution function of corresponding random variables. The method we discuss in this paper provides a reliability estimate of the given system under realistic situations. A sensitivity analysis study is carried out to examine the behaviour of mean square errors (MSE) of estimated system reliability under various scenarios. It is observed that MSE can be significantly reduced by a suitable choice of parameters in the membership function of fuzzy parameters.

ANALYTICAL AND COMPUTATIONAL ASPECTS OF A MULTI-SERVER QUEUE WITH IMPATIENCE UNDER DIFFERENTIATED WORKING VACATIONS POLICY 393

Aimen Dehimi, Mohamed Boualem, Amina Angelika Bouchentouf, Sofiane Ziani, Louiza Berdjoudj

A multi-server queueing system with synchronous differentiated working vacation policy, Bernoulli schedule vacation interruption, and customer impatience (balking and reneging) is studied. The system consists of c servers and a finite capacity N , where customers arrive according to a Poisson process and are served in the chronological order of their arrival. When the system becomes empty, servers wait for a random duration before entering a type-1 working vacation, during which service is provided at a reduced rate. If customers are present in the system at the moment of service achievement during this period, the vacation is interrupted. With a certain probability, servers return to the regular busy period; otherwise, they continue the working vacation. Upon completion of the working vacation, if the system is still empty, servers can take another working vacation of shorter duration, named type-2 working vacation; otherwise, they switch to the regular busy period. Customer impatience is considered during both the normal busy period and working vacations. A recursive analysis method is used to find the steady-state probabilities of the system. Then, some important performance measures are obtained. Furthermore, an optimal operational policy for the model is developed to minimize the total expected cost. The Grey Wolf Optimization (GWO) meta-heuristic approach is employed to determine the optimal service rates for both working vacations and normal busy periods. Finally, several numerical examples are provided to validate and support the theoretical findings.

COSINE MARSHAL-OLKIN-G FAMILY OF DISTRIBUTION: PROPERTIES AND APPLICATIONS..... 408

Akeem Ajibola Adepoju, Alhaji Modu Isa, Olalekan Akanji Bello

Trigonometric distributions have recently been emphasized due to its applicability and relevance for modeling different phenomena. This article contributes to the existing literature on trigonometric family by introducing and investigating new trigonometric family of distribution which is developed by compounding the cosine family of distribution with Marshall-olkin family of distribution to form a new Cosine Marshall-Olkin family of distribution (CMO). Graphical, numerical and analytical approach was explored to study the properties and applicability of the new CMO family of distribution. Special representations and important reliability properties and other statistical properties were defined. Simulation study was conducted in order to have an insight on the estimates of the three parameters model using maximum products of spacing (MPS). Emphases on the greater flexibility of the new CMO family of distribution beyond the cosine-G family and other top models of the Cosine related family was made through Weibull distribution. The results revealed the superiority of the Cosine Marshall-Olkin Weibull model (CMO-W) over others via two data sets.

**ANALYSIS OF MMAP/PH1,PH2/1 PREEMPTIVE PRIORITY INVENTORY RETRIAL
QUEUEING SYSTEM WITH SINGLE VACATION, WORKING BREAKDOWN, REPAIR AND
CLOSEDOWN 423**

G. Ayyappan, S. Meena

This paper analyzes preemptive priority inventory retrial queueing system with a single vacation, working breakdown, repair, and closedown. We assume that an arrival follows the Marked Markovian arrival process and that the server will provide them with phase-type services. The (s, S) policy to replenish the items and the replenishing duration follow an exponential distribution. In this paper, we consider two types of customers: high-priority(HP) customers and low-priority(LP) customers. Arriving HP customers should get the service if the server is idle and has a positive inventory level; otherwise, they should wait in front of the service station. Arriving LP customers get service only if there is a positive inventory level and there are no high-priority customers in the system; otherwise, go for the finite capacity size of the orbit. After the completion of service, if no one is present in the high-priority queue and orbit, the server will close down the system and then go on a single vacation. The server is idle when the vacation period ends. When the server breaks down, it only serves the present customer and operates in slow mode while it is being repaired. The number of high-priority customers in the system, the number of low-priority customers in the orbit, the inventory level, and server status may all be determined in a steady state. Numerous key performance indicators are defined, and a cost analysis is obtained. To make our mathematical concept clearer, a few numerical examples are provided.

**OPTIMIZATION OF AN INVENTORY MODEL FOR DETERIORATING ITEMS ASSUMING
DETERIORATION DURING CARRYING WITH TWO-WAREHOUSE FACILITY 442**

Krishan Kumar Yadav, Ajay Singh Yadav, Shikha Bansal

A common topic in the context of its application in today's business contexts is inventory modelling and management. It is well-known that deterioration has a big impact on inventory management. One of the most frequent supply chain concerns is the deterioration of items during transit from a supplier's storehouse to a retailer's storehouse. In light of this, a two-level supply chain inventory model for decaying goods is developed with two warehouse (storehouse) facilities for retailers, namely Owned Warehouse (OW) and Rented Warehouse (RW), assuming deterioration both during carrying from a supplier's storehouse to a retailer's storehouses and in the retailer's storehouses themselves. Also, we are assuming the selling price and time sensitive demand. We are developed this model under inflation. Shortages are not allowed. The main objective of this study is to determine the optimal ordering policy in order to maximizes the retailer's profit per unit of time. The applicability of our suggested model is investigated using a numerical example and with the support of MATLAB programming software (version: R2021b). Sensitivity analysis is used to examine the effects of changing the values of system parameters. Graphical representations are also shown in this paper.

**ON MODELING OF BIOMEDICAL DATA WITH EXPONENTIATED GOMPERTZ INVERSE
RAYLEIGH DISTRIBUTION 460**

Sule Omeiza Bashiru, Alaa Abdulrahman Khalaf, Alhaji Modu Isa, Aishatu Kaigama

This paper introduces and thoroughly examines the Exponentiated Gompertz Inverse Rayleigh (EtGoIr) Distribution, a four-parameter extension of the Gompertz Inverse Rayleigh distribution. The primary focus is on its application to biomedical datasets, shedding light on its mathematical and statistical properties. Some properties of the distribution that were derived include the quantile function, median, moments, incomplete moments, Rényi entropy, and probability weighted moments. The model parameters were estimated using the method of maximum likelihood. A simulation study was conducted to investigate the consistency of the proposed model. The outcome of the investigation revealed that the model demonstrates consistency, as evidenced by the reduction in both root mean square error (RMSE) and bias as sample sizes increase. To showcase the practical relevance of the EtGoIr distribution, the paper applies the model to three distinct biomedical datasets. The results highlight its enhanced flexibility, demonstrating superior fit compared to its counterpart.

BEHAVIOR ANALYSIS PRESENTED SYSTEM WITH FAILURE AND MAINTENANCE RATE WITH USING DEEP LEARNING ALGORITHMS 476

Shakuntla Singla, Shilpa Rani, Diksha Mangla, Umar Muhammad Modibbo

The paper discusses the behavioral analysis and dependability of a three-unit system utilizing RPGT for system parameters. Since all three units P, Q and R include parallel subcomponents, in the event that one of them fails, the system continues to operate although at a reduced capacity, but it is not profitable to run the system when two units are in reduced state hence considered failed state. The rates of failures are exponentially distributed, but the rates of repair are generalized, independent, and differ based on the operational unit. Fuzzy concept is used to declare/ determine whether the system is in failed/ reduced/ failed state. Graphs and tables are drawn to compare failure/repair effect on the parameters values. The system parameters are modelled using Regenerative Point graphical Technique (RPGT) and optimized using Deep learning methods such as Adam, SGD, RMS prop. The results of the optimization may be used to validate and challenge existing models and assumptions about the systems.

OPTIMIZING INVENTORY CONTROL THROUGH A GRADIENT-BASED MULTILEVEL APPROACH IN THE FACE OF DEMAND AND LEAD TIME UNCERTAINTIES 486

Muragesh Math, D.Gopinath, B. S.Biradar

Systems of two-level assembly with unknown timing of leads are taken into consideration while arranging supplies. Probably, the final product's demand and its deadline are known. When all required parts are on hand, each level's assembly process gets underway. To address these problems, we have developed a model for the control of inventories for an uncapitated warehousing space in a manufacturing plant with unpredictable demand and lead times. The goal is to choose orders in a way that minimizes the overall system's cost. We present a multilevel optimization model including a rotating horizon that utilizes gradients to handle unknown lead time and demand, irrespective of the distributions at the core of them. Furthermore, a precise algorithm is created to solve the model. In a case study, we compare our approach with the current model. Our computational results indicate that while the new gradient-based multi-level optimization model nearly continuously yields the least expensive overall across all parameter settings. These models' performances are either systematically worse or extremely sensitive to cost parameters (holding cost, shortfall cost, etc.).

THE EXPONENTIATED SKEW LAPLACE DISTRIBUTION: PROPERTIES AND APPLICATIONS..... 497

Timothy Kayode Samson, Christian Elendu Onwukwe, Ekaette Inyang Enang

In this paper, a 4-parameter Exponentiated Skew Laplace distribution is defined and studied. Various statistical properties including its moment generating function, characteristics function, hazard function, and reliability function of the proposed ESLD were derived. The estimation of its parameters was carried out using the maximum likelihood method of estimation. The performance of the proposed ESLD compared with other similar distributions was demonstrated empirically with daily returns of S & P 500 between 2/02/24 and 28/03/2024 and daily returns of Bitcoin between 2/02/24 and 1/04/24 as obtained from Yahoo Finance. The fitness performance of the proposed distribution was evaluated based on log-likelihood, AIC, and BIC. Results obtained show that the proposed ESLD reported the highest log likelihood as well as the lowest AIC and BIC in the two data sets. This study therefore underscores the superiority of the proposed distribution over the some of the similar existing distributions.

A NOVEL ASYMMETRIC COMPOUND CLASS OF DISTRIBUTIONS WITH ESTIMATION AND APPLICATION 510

A.G. Al-Kilany, Amal S. Hassan, L.S. Diab, E.S. El-Atfy

This paper introduces and discusses the novel asymmetric class of distributions that have the name inverse power Lomax power series (IPLPS). This class of distributions is produced by combining the inverse power Lomax with the power series distributions. This combined approach provides an opportunity for the creation of flexible distributions with significant physical implications in many fields, like biology and engineering. The IPLPS distributions encompass several new compound distributions as sub-models along with a new class of compound distributions. Many statistical features, including moments, quantile function, conditional moments, inverse moments, uncertainty measures, and probability-weighted moments, are obtained. As a special model of the generated class, the parameters of the inverse power Lomax Poisson distribution are estimated by different methods, including least squares, Cramér von Mises, maximum likelihood, and weighted least squares. Through an extensive simulation analysis, the execution of different parameter estimation techniques for the inverse power Lomax Poisson model is performed to show its validity based on its mean squared error and absolute bias. Two real datasets are utilized to show the practicality of the newly generated model. Results show that the inverse power Lomax Poisson distribution provides the most fitted model for these datasets in comparison to other distributions such as power Lomax, Marshall- Olkin power Lomax, power Lomax Poisson, and Topp-Leone Lomax distributions.

EXPLORING LENGTH BIASED QUASI SUJA DISTRIBUTION: PROPERTIES AND APPLICATIONS..... 530

Vidya Yerneni, Aafaq A. Rather

This paper introduces a new statistical distribution called length biased quasi suja distribution (LBQS). It explores its properties, including moments, moment generating function(MGF), characteristic function(CF), harmonic mean, reliability, hazard rate and reverse hazard rate. Order statistics of the above distribution is obtained. Furthermore, the paper also examines various entropy which measures the randomness of system, like Renyi entropy and Tsalli's entropy. It also evaluates Bonferroni and Lorenz curves which are useful in measuring the inequality. It also discusses parameter estimation techniques specifically maximum likelihood estimation and likelihood ratio testing. Moreover, a simulation study has been conducted to demonstrate how well the distribution would perform in real-life situation. The validity of the distribution is also demonstrated with real-world data example of failure data, highlighting its potential for practical applications in data analysis.

A NEW GENERALIZATION OF SABUR DISTRIBUTION 545

Suvarna Ranade, Aafaq A. Rather

When the weight function depends on the lengths of the units of interest, the resulting distribution is called length biased. Length biased distribution is thus a special case of the more general form, known as weighted distribution. In this study, we introduce a novel probability distribution named the Length- Biased Sabur distribution (LBSD). This new distribution enhances the traditional Sabur distribution by incorporating a weighted transformation approach. The paper investigates the probability density function (pdf) and the cumulative distribution function (cdf) associated with the LBSD. A thorough examination of the distinctive structural properties of the proposed model is conducted, covering the survival function, conditional survival function, hazard function, cumulative hazard function, mean residual life, moments, moment generating function, characteristic function, likelihood ratio test, ordered statistics, entropy measures, and Bonferroni and Lorenz curve.

MODELING THE INTERCONNECTED OPERATION OF ENERGY SYSTEMS FOR ENERGY SECURITY STUDY IN TODAY'S CONTEXT..... 554

Dmitry Krupenev, Natalia Pyatkova

The paper shows the need for comprehensive research into energy security problems to assess the possibilities of interconnected operation of all energy industries with the view to identifying the implications for consumers of energy resources in the event of emergencies in one or several industries at the same time. The paper presents a methodological framework and features of modeling the interrelated operation of the industries in current context and a model developed for these studies. The results of experimental studies using the developed methodology are shown through the analysis of several critical situations (threats to energy security) of various nature.

ESTIMATION OF PARAMETERS FOR KUMARASWAMY EXPONENTIAL DISTRIBUTION BASED ON PROGRESSIVE TYPE-I INTERVAL CENSORED SAMPLE..... 567

Manoj Chacko, Shilpa S Dev

In this paper, we consider the problem of estimation of parameters of the Kumaraswamy exponential distribution using progressive type-I interval censored data. The maximum likelihood estimators (MLEs) of the parameters are obtained. As it is observed that there is no closed-form solutions for the MLEs, we implement the Expectation-Maximization (EM) algorithm for the computation of MLEs. Bayes estimators are also obtained using different loss functions such as the squared error loss function and the LINEX loss function. For the Bayesian estimation, Lindley's approximation method has been applied. To evaluate the performance of the various estimators developed, we conduct an extensive simulation study. The different estimators and censoring schemes are compared based on average bias and mean squared error. A real data set is also taken into consideration for illustration.

ANALYSIS OF MX/G/1 QUEUE WITH OPTIONAL SECOND SERVICE, FEEDBACK AND BERNOULLI VACATION..... 583

S. Karpagam, B. Somasundaram, A. Kavin Sagana Mary, R. Lokesh,

In this article the single-server queue situation described with batch arrivals, a mandatory first service and a choice of second service are provided to the customers. A general distribution governs the service times, whereas a compound Poisson distribution follows customer arrivals. Although each new customer requests the first mandatory service, only some of them choose the optional second service. Customers who are dissatisfied with mandatory service are more likely to get the required services later on. After every service is finished, the server might choose to go on Bernoulli vacation. Time dependent probability generating functions are constructed in terms of Laplace transforms using the supplementary variable approach, and explicit results are obtained for the steady state. Additionally, mean waiting time and mean queue length expressions are examined. The graphical and numerical representations improve comprehension of the results even further.

INVERTED DAGUM DISTRIBUTION: PROPERTIES AND APPLICATION TO LIFETIME DATASET 595.

Abdulhameed A. Osi, Shamsuddeen A. Sabo, Ibrahim Z. Musa

This article presents the introduction of a novel univariate probability distribution termed the inverted Dagum distribution. Extensive analysis of the statistical properties of this distribution, including the hazard function, survival function, Renyi's entropy, quantile function, and the distribution of the order statistics, was conducted. Parameter estimation of the model was performed utilizing the maximum likelihood method, with the consistency of the estimates validated through Monte Carlo simulation. Furthermore, the applicability of the proposed distribution was demonstrated through the analysis of two real datasets.

A NOVEL HYBRID DISTRIBUTED INNOVATION EGARCH MODEL FOR INVESTIGATING THE VOLATILITY OF THE STOCK MARKET 605

Mubarak M.T., Adubisi O.D., Abbas U.F.

When calculating risk and making decisions, investors and financial institutions heavily rely on the modeling of asset return volatility. For the exponentiated generalized autoregressive conditional heteroscedasticity (EGARCH) model, we created a unique innovation distribution in this study called the type-II-Topp-Leone-exponentiated-Gumbel (TIITLEGU) distribution. The key mathematical characteristics of the distribution were determined, and Monte Carlo experiments were used to estimate the parameters of the novel distribution using maximum likelihood estimation (MLE) procedure. The performance of the EGARCH (1,1) model with TIITLEGU distributed innovation density in relation to other innovation densities in terms of volatility modeling is examined through applications using two Nigerian shock returns. The results of the diagnostic tests indicated that, with the exception of the EGARCH (1,1)-Johnson (SU) reparametrized (JSU) innovation density, the fitted models have been sufficiently specified. The parameters for the EGARCH (1,1) model with different innovation densities are significant at various levels. Furthermore, in out-of-sample prediction, the fitted EGARCH (1,1)-TIITLEGU innovation density performed better than the EGARCH (1,1)- existing innovation densities. As a result, it is decided that the EGARCH-TIITLEGU model is the most effective for analyzing Nigerian stock market volatility.

SINGLE AND DOUBLE ACCEPTANCE SAMPLING PLAN FOR TRUNCATED LIFE TESTS BASED ON GAMMA LINDLEY DISTRIBUTION..... 619

Sriramachandran G. V.

For time-truncated life tests, this work defines single acceptance and double acceptance sampling plans assuming that the product's lifespan follows the Gamma Lindley distribution. The minimum sample size needed in a single acceptance sampling plan for lot approval is calculated for a range of parameter combinations and a fixed test termination time. This ensures the given average product life and the corresponding number of failures. Operational characteristic and producer risk values are also tabulated for these parameter values. Using a double acceptance sampling plan, the best first and second samples are obtained to ensure that the products specified average with a certain level of customer trust. Finally, under the same conditions, the minimum sample size obtained using these strategies are compared with other acceptance sampling plans.

FUZZY VARIABLE LINEAR PROGRAMMING PROBLEMS USING A FUZZY DUAL SIMPLEX ALGORITHM 630

Srinivasa Rao Kolli, U.V. Adinarayana Rao, Taviti Naidu Gongada

In modern research, several brilliant minds investigate linear programming problems involving fuzzy variable quantities. Many researchers have turned to linear programming by fuzzy variables to address this problem. Various fuzzy simplex approaches have been developed, using ranking functions to handle fuzzy numbers. Results from this research suggest that linear ranking functions can provide a straightforward interpretation of problems involving linear programming by fuzzy variable quantities. To solve these types of problems, the Fuzzy Dual Simplex Tableau method is often applied, which proves useful for sensitivity analysis when modifications are made to the activity vectors of the fundamental columns. In this study, a numerical case is presented to demonstrate the potential benefits of this approach for future technologies.

A MODIFIED AILAMUJIA DISTRIBUTION: PROPERTIES AND APPLICATION 638

David Ikwuoche John, Okeke Evelyn Nkiru, Franklin Lilian

This study presents a modified one-parameter Ailamujia distribution called the Entropy Transformed Ailamujia distribution (ETAD) is introduced to handle both symmetric and asymmetric lifetime data sets. The ETAD properties like order and reliability statistics, entropy, moment and moment generating function, quantile function, and its variability measures were derived. The maximum likelihood estimation (MLE) method was used in estimating the parameter of ETAD and through simulation at different sample sizes, the MLE was found to be consistent, efficient, and unbiased for estimating the ETAD parameter. The flexibility of ETAD was shown by fitting it to six different real lifetime data sets and compared it alongside seven competing one-parameter distributions. The goodness of fit (GOF) results from Akaike information criteria, Bayesian information criteria, corrected Akaike information criteria, and Hannan-Quinn information criteria show that the ETAD was the best fit amongst all the seven competing distributions across all the six data sets.

MODIFIED GROUP RUNS CONTROL CHART FOR MONITORING PROCESS DISPERSION..... 653

Chandrakant G. Gardi, Vikas B. Ghute

Due to a rise in competitiveness, it has become an intense concern to the manufacturers to monitor process dispersion to avoid low quality production. To ensure quality production, the control chart that gives early detection of change in the dispersion is always encouraged. Researchers have suggested various control charts based on different estimators of process dispersion. Recently, many synthetic control charts based on such estimators are put forth by researchers to effectively monitor the dispersion in the process. Modified Group Runs (MGR) control chart is an extension of synthetic charts with further enhancement in the detection ability. In this paper, we propose a MGR control chart based on Downton's estimator (D). Comparison of MGR control chart with synthetic chart based on estimator D reveals the enhanced performance of MGR-D chart.

ON THE FLEXIBILITY OF TYPE I HALF LOGISTIC EXPONENTIATED FRECHET DISTRIBUTION 660

Olalekan Akanji Bello, Sani Ibrahim Doguwa, Abukakar Yahaya, Haruna Mohammed Jibril,

In this article, we delve into the modeling and analysis of lifetimes, which hold substantial importance across various scientific and industrial fields. Our focus is on introducing a novel distribution termed the Type I Half-Logistic Exponentiated Frechet (TIHLEtF) Distribution, which is an extension of the Frechet distribution. We have derived a crucial representation of the density function for this distribution. Furthermore, we explore several statistical properties associated with the TIHLEtF distribution. These properties encompass explicit expressions for the quantile function, probability-weighted moments, moments, moments generating function, reliability function, hazard function, and order statistics. To estimate the model parameters, we employ the maximum likelihood estimation technique and present the results of a simulation study. To emphasize the superiority of our newly introduced distribution, we apply it to two real datasets. The outcomes of our analysis reveal that the TIHLEtF distribution outperforms the other considered distributions in terms of fitting the data in these real-world cases.

BAYESIAN ANALYSIS OF EXTENDED MAXWELL-BOLTZMANN DISTRIBUTION USING SIMULATED AND REAL-LIFE DATA SETS 675

Nuzhat Ahad, S.P.Ahmad, J.A.R eshi

The objective of the study is to use Bayesian techniques to estimate the scale parameter of the 2Kth order weighted Maxwell-Boltzmann distribution(KWMBD). This involved using various prior assumptions such as extended Jeffrey's , Hartigan's , Inverse-gamma and Inverse-exponential, as well as different loss functions including squared error loss function (SELF), precautionary loss function (PLF), Al Bayyati's loss function (ALBF), and Stein's Loss Function (SLF).The maximum likelihood estimation (MLE) is also obtained. We compared the performances of MLE and bayesian estimation under each prior and its associated loss functions. And demonstrated the effectiveness of Bayesian estimation through simulation studies and analyzing real-life datasets.

BAYESIAN NON-INFERIORITY TEST BETWEEN TWO BINOMIAL PROPORTIONS..... 689

W. B. Yahya, C. P. Ezenweke, O. R. Olaniran, I. A. Adeniyi, K. Jimoh, R. B. Afolayan, M. K. Garba, I. Ahmed

The paper aimed to propose a new Bayesian test method for establishing a non-inferiority measure between an active treatment (drug) and a new (cheaper) treatment using two independent binomial samples. A Bayesian test statistic was developed for testing non-inferiority between two independent binomial proportions. Conjugate Beta prior was assumed for the binomial proportions to elicit posterior from the same Beta family of distributions. The efficiency of this test method was established via power analysis and its ability to yield the nominal Type I error rate (alpha) in a detailed Monte-Carlo study. Results from this study showed that the proposed test method yielded higher powers and good estimates of the Type I error rate at the chosen sample sizes and varying non-inferiority margins (effect sizes). Thus, the new Bayesian test method is very efficient at detecting the significance of the non-inferiority margin between two independent binomial proportions when such is not negligible at all sample sizes. Further results showed that the size of the two population proportions being tested influences the power and the estimated nominal Type I error rate with an increase in power and a good estimate of Type I error rate achieved when both population proportions being tested are less than 0.5. It is therefore concluded that the new Bayesian test method can be employed whenever it is desirable to establish the existence of non- inferiority or otherwise between a pair of (clinical) treatments (drugs). All the simulations and analyses were performed with the R statistical package.

BAYESIAN AND E-BAYESIAN ESTIMATION OF EXPONENTIATED INVERSE RAYLEIGH DISTRIBUTION USING CONJUGATE PRIOR 704

Ramesh Kumar, Hemani Sharma, Rahul Gupta, Ableen Kaur

This study explores the application of Bayesian and E-Bayesian techniques to estimate the scale parameter of the Exponentiated Inverse Rayleigh distribution. Bayesian estimates for the parameter are derived using an informative Gamma prior and evaluated under three distinct loss functions: De- Groot, Squared Error, and Al-Bayyati loss functions. Various Properties of the E-Bayesian estimators under different loss functions have also been studied. To compare the effectiveness of E-Bayesian estimates against the Bayesian counterpart, a simulation study is conducted using MatLab. The various derived estimators were compared in terms of their Mean Squared Error. The results of a simulation study reveal that E-Bayesian estimates exhibit a smaller Mean Squared Error in comparison to Bayesian estimates, thereby demonstrating their enhanced efficiency. Among the E- Bayesian estimates, the third one stands out as the most effective. Moreover, the analysis highlights that the Squared Error loss function outperforms the Al-Bayyati and De-Groot loss functions, exhibiting a smaller MSE. Furthermore, the efficacy of these estimators is demonstrated through an analysis of a real-life dataset.

AVAILABILITY ANALYSIS FOR IDENTIFICATION OF CRITICAL FACTOR OF A THERMAL POWER PLANT 717

Pardeep Kumar, Vipin Kumar Sharma, Dinesh Kumar

In the present stimulated business environment, power sector is playing a major role in the economic growth of India. During the last 20 years, the country had been facing a poor supply of energy and this supply-demand gap is increasing continuously. So, it is important for power plants to improve its power generation capacity drastically by reducing the failure rate. In the present paper, to analyze the causes of poor availability, thermal power plant has divided into six different systems and a system comprising of waste gases heating system has been considered. With the help of transition diagram, mathematical equations have been used to find out the availability. After analyzing, it was found that the value of availability is very low and boiler tube failure is one of the most critical factors for this low availability of system. Economizer zone has identified having long existence time of failures and frequency of occurrence is very high. So, minimizing the failure rate with the help of a proper maintenance schedule will result in decreasing the shutdown period of the plant and increasing the system availability.

WEIGHTED R-NORM ENTROPY FOR LIFETIME DISTRIBUTIONS: PROPERTIES AND APPLICATION 725

Bilal Ahmad Bhat, M.A.K Baig

In the field of information theory, different uncertainty measures have been introduced by various researchers. These measures are widely used in reliability and survival studies. In this article, we introduce two new weighted uncertainty measures which are known as weighted R-Norm entropy (WRNE) and weighted R-Norm residual entropy (WRNRE). WRNE and WRNRE are "length-biased" shift-dependent uncertainty measures in which higher weight is assigned to large values of the observed random variable. Several important properties of these measures are studied. Some significant characterization results and the relationships of WRNRE with other reliability measures are presented. We also show that the survival function is uniquely determined by the WRNRE. Finally, based on a real life data set of bladder cancer patients, we illustrate the importance of WRNE and WRNRE.

AN IMPROVED ESTIMATOR OF FINITE POPULATION MEAN UNDER RANKED SET SAMPLING 736

Francis Delali Baeta, Diogban Jakperik, Michael Jackson Adjabui

To obtain reliable estimates of population parameters, data that is sampled for estimation must accurately represent the underlying population. Sampled data that is representative of the underlying population depends also on the sampling technique that was used in obtaining them. This is very important since sampling bias could lead to over or under estimation of parameters. Ranked Set Sampling is considered to be a better alternative to the classical sampling designs in obtaining such data. Ranked Set Sampling is designed to minimize the number of measured observations required to achieve a desired precision in making inferences, and thus it is more economical to use for the purposes of estimation, compared to the classical sampling designs. This is also an added advantage in cases where it is difficult to obtain data. Many estimators have been developed recently for the estimation of finite population mean under ranked set sampling. This paper aims to improve estimation by modifying an existing estimator using a simple linear combination of the known population mean, square root of the known coefficient of variation, and the known median of an auxiliary variable. The theoretical properties of the proposed estimator, such as the bias and mean squared error were derived up to the first order of approximation, using Taylor's expansion. The bias, mean squared error, absolute relative bias, and the relative efficiency were used as means of evaluation and comparison between the proposed modified estimator and its competitors. The R software was used to aid computations. Empirical applications to real data showed that the

proposed modified estimator is superior to the competing estimators that were compared since it has least bias, the least mean squared error, the least absolute relative bias, and the highest relative efficiency in all sample sizes that were considered. The bias and mean squared error of the modified estimator under Ranked Set Sampling was found to be smaller than those of the existing estimators that were compared. Hence it is more efficient and capable of providing reliable estimates than the existing estimators that were compared and so we recommend that it should be used in survey estimations.

BAYESIAN ESTIMATION OF PARAMETERS AND RELIABILITY CHARACTERISTICS IN THE INVERSE GOMPERTZ DISTRIBUTION..... 744

Taiwo. M. Adegoke, Latifat A. Abimbola, Oladapo M. Oladoja, Oyindamola. R. Oyebanjo, K.O. Obisesan

In this study, we derive Bayes' estimators for the unknown parameters of the Inverse Gompertz Distribution (IGD) using three alternative loss functions: the Squared Error Loss Function (SELF), the Entropy Loss Function (ELF), and the Linex Loss Function. Closed-form formulas for Bayes estimators are not possible when both parameters are unknown, hence Lindley's approximation (L-Approximation) is used for computation. We examine the performance of these estimators using their simulated hazards and assess their effectiveness in parameter estimation. It was discovered that as the sample size increases, parameter estimations became more precise and accurate across all functions. However, ELF consistently has lower MSE values than SELF and LINEX, indicating better parameter estimation. This pattern was also seen in the estimation of the hazard function, where ELF regularly beat SELF and LINEX, implying more efficient parameter estimation overall.

A TWO-PARAMETER ARADHANA DISTRIBUTION WITH APPLICATIONS TO RELIABILITY ENGINEERING 757

Ravi Shanker, Nitesh Kumar Soni, Rama Shanker, Mousumi Ray, Hosenuur Rahman Prodhani

The search for a statistical distribution for modelling the reliability data from reliability engineering is challenging and the main cause is the stochastic nature of the data and the presence of skewness, kurtosis and over-dispersion. During recent decades several one and two-parameter statistical distributions have been proposed in statistics literature, but all these distributions were unable to capture the nature of data due to the presence of skewness, kurtosis and over-dispersion in the data. In the present paper, two-parameter Aradhana distribution, which includes one parameter Aradhana distribution as a particular case, has been proposed. Using convex combination approach of deriving a new statistical distribution, a two-parameter Aradhana distribution has been proposed. Various interesting and useful statistical properties including survival function, hazard function, reverse hazard function, mean residual life function, stochastic ordering, deviation from mean and median, stress-strength reliability, Bonferroni and Lorenz curve and their indices have been discussed. The raw moments, central moments and descriptive measures based on moments of the proposed distribution have been obtained. The estimation of parameters using the maximum likelihood method has been explained. The simulation study has been presented to know the performance in terms of consistency of maximum likelihood estimators as the sample size increases and. The goodness of test of the proposed distributions has been tested using the values of Akaike Information criterion and Kolmogorov-Smirnov statistics. Finally, two examples of real lifetime datasets from reliability engineering have been presented to demonstrate its applications and the goodness of fit, and it shows a better fit over two-parameter generalized Aradhana distribution, quasi Aradhana distribution, new quasi Aradhana distribution, Power Aradhana distribution, weighted Aradhana distribution, gamma distribution and Weibull distribution. The flexibility, tractability and usefulness of the proposed distribution show that it is very much useful for modelling reliability data from reliability engineering. As this is a new distribution and it has wide applications, it will draw the attention of researchers in reliability engineering and biomedical sciences to search many more applications in the future.

**STATISTICAL DESIGN OF CONDITIONAL REPETITIVE GROUP SAMPLING PLAN
BASED ON TRUNCATED LIFE TEST FOR PERCENTILE LIFETIME USING
EXPONENTIATED GENERALIZED FRECHET DISTRIBUTION 775**

S. Jayalakshmi, S. Vijilamery

Reliability Acceptance sampling plan is used to assess whether to accept or reject a product depending on its lifetime. An inspection carried out for the purpose of determining if lifetime inspections are performing properly can be tested by submitting a truncated lifetime test. In this paper describes a new approach on Conditional Repetitive Group Sampling Plan based on Truncated life test is proposed and the lifetime follows an Exponentiated Generalized Frechet Distribution. For each consumer risk, it is determined whether minimum sample sizes are required to assert a percentile life. It is calculated that the operating characteristic function values of the sampling plans as well as the producer's risk ratio corresponding to the sampling plans. The results are illustrated with numerical examples and a real-world data set is used to demonstrate the impact and performance of the suggested acceptance sampling plans.

**NEW EXTENSION OF INVERTED MODIFIED LINDLEY DISTRIBUTION WITH
APPLICATIONS..... 788**

Devendra Kumara, Anju Goyalb, P. Pareekc, M. Sahaa

In this article we, proposed a new two parameter distribution called inverted power modified Lindley distribution. The main objective is to introduce an extension to inverted modified Lindley distribution as an alternative to the inverted exponential, inverted gamma and inverted modified Lindley distributions, respectively. The proposed distribution is more flexible than the above mentioned distributions in terms of its hazard rate function. In the part of estimation of the proposed model, we first utilize the maximum likelihood (ML) estimator and parametric bootstrap confidence intervals, viz., standard bootstrap, percentile bootstrap, bias-corrected percentile (BCPB), bias-corrected accelerated bootstrap (BCAB) from the classical point of view as well the Bayesian estimation under different loss functions, squared error loss function, modified squared error loss function, and Bayes credible interval as to obtain the model parameter based on order statistics. A simulation study is carried out to check the efficiency of the classical and the Bayes estimators in terms of mean squared errors and posterior risks, respectively. Two real life data sets, have been analyzed for order statistics to demonstrate how the proposed methods may work in practice.

**A MAP/PH1, PH2/2 INVENTORY QUEUEING SYSTEM WITH TWO COMMODITY,
MULTIPLE VACATION, SERVER FEEDBACK, WORKING BREAKDOWN, REPAIR AND
EMERGENCY REPLENISHMENT 805**

G. Ayyappan, N. Arulmozhi

We investigate a continuous review inventory queuing system in the present study that has two heterogeneous servers: Server-2, which is reliable, and Server-1, which is unreliable. An exponentially distributed random time is used to describe the repair process when server-1 has an interruption. On the other hand, server-2 is completely dependable, but it goes on vacation when the system is empty. These two goods can be reordered under ordering regulations. To ensure customer satisfaction, an emergency replenishment of one item with no lead time occurs when the on-hand inventory level falls to zero. We use the matrix analytic approach for the QBD process under a steady-state probability vector. We also take into account the overall cost and the busy time. Furthermore, numerical data shows the benefits of the suggested approach in a range of random circumstances.

MY WAY TO RELIABILITY AND SAFETY

M. Yastrebenetsky

•
Dr. Sc. (Techn), Prof., Honor Doctor of Science and Technique
of Ukraine, Gnedenko-Forum Past President

Abstract

I was born in Moscow on September 10, 1934. I wrote about my family and previous generations of Yastrebenetsky (a soldier under Emperor Nicholas I, a businessman, a doctor, a chemical engineer) in my book "Generations of Yastrebenetsky" [1]. After finishing the institute and postgraduate studies, 2/3 years of my life were connected with activity in the fields of reliability and safety of technical systems. This activity is the subject of the article.

Keywords: Reliability, Safety, Nuclear Power Plant, Instrumentation and Control system.

1. Reliability

I worked in the Central Scientific Research Institute of Complex Automatization (CSRICA) in Moscow and was engaged in the analysis of the dynamic characteristics of industrial automatic regulators. At the international exhibition in Moscow, I saw a big poster with the inscription "We have high reliability: $\lambda = \dots 1/\text{hours}$ ". I didn't know, what its reliability? What its λ ? And a question naturally happened: what is the reliability of the regulators with dynamic characteristics I researched? Neither we nor the creators and manufacturers of the regulators knew this. We did not have stands for reliability tests, manufacturers did not fulfill reliability tests at that time, and an analytical assessment of reliability could not be done due to the lack of data on the component's reliability. I got "carte blanche" in CSRICA for reliability work and created in CSRICA the group of reliability (what later became the Reliability Laboratory and the Reliability Department). The idea arose- the determination of reliability measures from collecting and processing statistical data on regulators' failures and repairs during their operation. At that time, it was not clear whether operational data could be used to obtain objective information about the reliability of industrial automation. We began with the development of methods for collecting and processing information, ways to increase its validity. We chose the nearest thermal power plant with many new automatic regulators, trained and monitored the plant personnel who collected information. The first quantitative estimates of automatic regulator reliability measures were published in 1965 [2]. In the future, the analysis of statistical data about operational reliability, in particular point stochastic processes (flows) of failures became one of the directions of my future work.

My activities in the reliability area included:

- Assessing and ensuring the reliability of information computer systems (ICS) developed by CSRICA in different branches of industries (thermal power plants with units 300 MWt and 800 MWt , nuclear power plants (NPP) in the USSR and Bulgaria with VVER-1000 reactors, including 6 units of Zaporizhzhya NPP (ZNPP)- the biggest in Europe, and unit RBMK-1500 Ignalina NPP in Lithuania, chemical-technological enterprises, metallurgy etc.). Even before the first fuel loading at unit 1 ZNPP I was taken around the locations inside the reactor and touched the reactor equipment with my hands. After the start-up of unit 1 ZNPP, we collected and analyzed data on the reliability of the ICS and all Instrumentation and Control systems (I&C) of this unit. It turned out that the weak point of these systems was the Universal Complex of Technical Means (UCTS). I remember what the adjusters said to me about the UKTS at unit 1: "We'll start the unit with it, but will we be able to stop it later?". We

analyzed information about the reliability of this complex - as a result, it was soon replaced. It's hard to count the time I spent on business trips to ZNPP. I often remember the resolution of the USSR Deputy Minister of Energy and Electrification A. Maioret on the letter: "Efforts to improve the reliability of NPP instrumentation and control systems (I&CS) can never be considered sufficient".

- Collecting and processing statistical data on the reliability of I&C components -sensors, actuators, meters, various devices, and first, new computer systems. The aims of these works were to create feedback with information about reliability between the users of automation equipment and its manufacturers, development of recommendations on reliability improvement (e.g., Moscow plant "Manometer", plant "Lvivpribor"), create reliability databases, collection and processing of information for maintenance optimization (overhaul intervals, nomenclature and several spare parts, etc.). Proposed norms for overhauls and spare parts for automation equipment were developed together with the enterprises of the USSR Ministry of Energy and Electrification and introduced at all USSR thermal power plants.

- Development of methods to assess and ensure the reliability of automatic I&C systems for technological processes, including methods of assessing these systems functional reliability, analysis of the impact of reliability on efficiency, analysis of metrological failures, etc.

- The analysis of the operating reliability of the I&CS in different branches of industry resulted in the necessity to develop new mathematical models of reliability which were published in Russian ("Proceedings of the USSR Academy of Science. Technical Cybernetics", "Automatic and Remote Control"), then translated and reprinted in English:

- new class of regenerative random processes [2],
- models of unrenewable and renewed elements reliability under external disturbances ([3], [4]),
- models of failures flow, when external disturbance are stochastic processes ([5]),
- a rarefaction of the Markov renewal process ([6])
- flow of failures caused by the crossing of non-constant levels''''''00/ by regenerating random process [7].

I can add the model of regenerative processes with some types of regenerative points (together with V. Rykov), published in Ukrainian journal [8].

- Elaboration of USSR standards ([9], [10]), and sets of departmental standards (example- [11]) which contain requirements for the reliability of ICSTP and components, their reliability assessment and testing, organization of the maintenance.

- My first books on reliability [12], [13] (together with B. Solyanik) and [14] (author of foreword – acad. B.V. Gnedenko) were published in 1968, 1978 and in 1982 in Moscow. These books contained reliability models and information on the reliability of I&C systems and their components under operating conditions. Those were the times of rapid development of reliability research when such books were sold out in 14000, 10000 and 7000 copies -unthinkable quantities for the present time. The books were based on the data of the operational reliability of automatics collected by us.

I participated in many scientific and technical conferences, meetings, and reliability tests in different parts of the huge Soviet Union: - from the northeast (Bilibinskaya NPP in Chukotka) to the southwest (Azerbaijan, Armenia). From the southeast (Vladivostok, Irkutsk) to the northwest (Riga, Tallinn). The map of the USSR with the places of my official trips is shown in Fig. 1.

These places included different parts of USSR from Nord-West (Riga- collaboration with schools of prof. Kordonsky and Scliarevich) to South-East (Vladivostok, Branch of USSR Academy of Science- O.V. Abramov), from Nord-East (Bilibino NPP in Chukotka- verification of reliability and safety) to South-West (Baku, Sumgait- common work with Reliability Department of Oil and Chemical Avtomatization Institute- Sh.A. Kiasbeily) Safety of Nuclear Power Plants (NPP's) Problems of NPP's safety have always been urgent since NPP's began to operate. In Ukraine the work on NPP safety was especially vital after the accident at Chernobyl NPP - the most severe and widely known in the world of all technogenic accidents. In all countries, where NPP's are operated, there are state organizations that regulate the state policy in the field of NPP safety. Currently, such an organization in Ukraine is called the State Nuclear Regulatory Inspectorate of Ukraine (SNRIU). It includes Technical Support

Organization - State Scientific and Technical Center for Nuclear and Radiation Safety (SSTC NRS), also subordinate to the National Academy of Sciences of Ukraine.



Figure 1. The map of USSR with marked places of trips

The impetus for the work on the reliability of Information and Control Systems was the Board of Directors of the enterprises of the Ministry of Instrument Engineering in 1984 under the leadership of the Director of the Institute of Control Problems Acad. V.A. Trapeznikov and the Minister of Instrument Engineering M.S. Shkabardni. I made a report on the state of work on the reliability of NPP automation. After my report, I was regularly invited to the Industrial Department of the Central Committee of the CPSU (Yu.I. Abramov) – Moscow, Old Square to discuss the situation with work on the reliability of NPP automation.

My university activity in reliability:

- lectures on reliability as a professor of Kharkov Polytechnical University (Departments “Technical Cybernetics” and “Systems Analysis”) and the other universities in USSR and the other countries (the most distant university was Havana, Cuba),

- issue of textbook for universities on the reliability of automatic control systems, agreed upon USSR State Educational Committee (together with G. Ivanova) [15].

A considerable help to me, as to many my colleagues in reliability, was rendered by the outstanding specialists in the mathematical reliability theory, grouped around the of the Moscow State University Department of Probability Theory, headed by acad. B.V. Gnedenko. Guidance of this Department had handed him by famous mathematic – acad. A. Kolmogorov. The book B.V. Gnedenko, A.D. Soloviev, Yu.V. Beliaev [16] became as classical for reliability specialists and was issued in many languages (fig. 2).

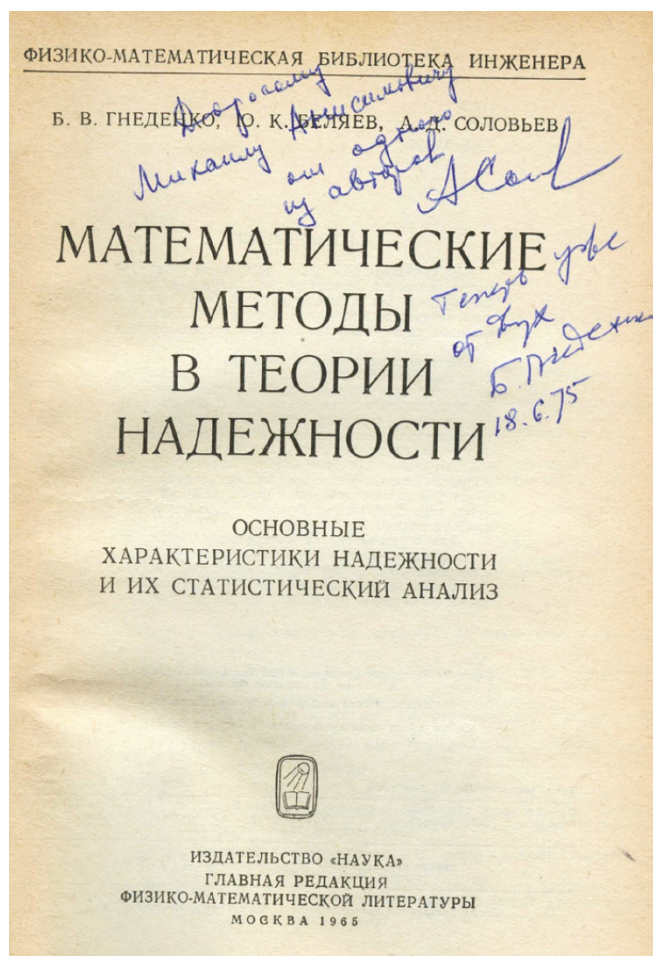


Figure 2. The book "Mathematical methods in reliability theory" with autographs of two authors

I gave reports on the seminars and school organized by this Department. Prof. A.D. Soloviev (Fig. 3) and prof. I.A. Ushakov performed as opponents in defense of my Dr. Sc. thesis. The most memorable meeting for me was the school in Dilijan (Armenia) in 1970. The Department of Probability Theory at Moscow University brought together reliability specialists from all over the Soviet Union. Many of them became my lifelong friends.

Subsequently in 2006 Gnedenko was named international non-commercial association "Gnedenko-Forum" created by I.A. Ushakov and A. V. Bochkov. "Gnedenko-Forum" joined the specialists working in the fields of reliability and mathematic statistics applications in safety, risk analysis, survivability. "Gnedenko-Forum" aims are establishing contacts between specialists in the world and exchange of professional information (about new publications, international conferences, meetings, participation in organization of conferences, etc.).

"Gnedenko-Forum" publishes in the USA the quarterly electronic journal in English "Reliability: Theory & Applications". Since 2006 about 1000 articles have been published that contained both theoretical and methodological problems and practical papers related to reliability. The journal is registered with the U.S. Library of Congress. Prof. V. Rykov is the Editor-in-Chief of the journal now. I was honored to be as President of the "Gnedenko-Forum" in 2014-2020. Before me, President of the "Gnedenko-Forum" in 2012-2015 was prof. Way Kuo, one of the world's most famous specialists in safety and the reliability of electronic systems. After me President of "Gnedenko-Forum" became prof. B. Dimitrov (USA). I am very glad that thanks to "Gnedenko-Forum" I have renewed regular contacts with my long-time reliability colleagues – prof. V. Kashtanov, prof. V. Rykov, prof. I. Shubinsky, who continue to be active.

The word "regulation" has gone through my life. My first business trip was to test regulators for a steam boiler. Then was followed by work on the analysis of dynamic characteristics of technological processes automatic regulators, the results of which are given in the book "Industrial Automatic Regulators" [17]. And then, for many years, safety regulation of NPP's and other critical facilities was the area of my activity.

SNRIU and SSTC NRS were based on the specialists who worked at the Chernobyl NPP during the accident and later connected with the restoration of the Chernobyl NPP after the accident. The first SNRIU head and creator was N.A. Steinberg, who was appointed to position chief engineer Chernobyl NPP immediately after the accident.

The former employees of the Chernobyl NPP who joined SNRIU and SSTC NRS felt on their shoulders what it is NPP safety. I had a lot to learn from these people because after the USSR collapse my activity related to SSTC NRS, which included Kharkov Reliability Department as Kharkov subsidiary. Main objects of SNRIU and SSTC NRS activities were 4 operating NPPs of Ukraine with 13 (after commissioning of two power units in 2004 - 15) power units. Besides, the objects of activity are NPP spent fuel storage facilities, "Shelter" facility above the unit 4 of Chernobyl NPP, research nuclear reactors, various sources of ionized radiation, etc.

Starting from 1993, for long time I have been supervising a set of activities on nuclear and radiation safety regulation, defined by instrumentation and control systems (I&CS), for all these objects. The main directions of my work on safety were:

- Development of SNRIU regulatory requirements to NPP I&CS and their components (hardware, software, software -hardware complexes) according to nuclear and radiation safety criteria. The document [18] containing such requirements was issued in 2000. The I&CS creation for two new VVER-1000 power units and I&CS modernization of many Ukrainian units for their life extension were carried out in accordance with the requirements of document [18]. In 2015, this document was revised [19] taking into account the experience of Ukrainian NPPs operation, the lessons of the accident at Fukushima NPP and changes in the international requirements (first of all, IAEA and In addition, the requirements were divided into regulatory requirements established by SNRIU, and technical requirements established by the operating organization – National Energy Company "Energoatom". [20]. According to [18] and [19], Ukrainian companies developed I&CS for NPPs in Bulgaria, the Czech Republic, Armenia and other countries.
- Development of SNRIU regulatory requirements:
 - methods of assessment of compliance of safety important ICS NPP with the requirements for nuclear and radiation safety;
 - requirements for the procedure and content of work for the life extension of I&CS included in NPP safety important systems, etc.
- More than 1000 state nuclear and radiation safety expertise's (safety reviews) of I&CS for NPP's , research reactors, NPP spent fuel storage facilities, the Object "Shelter" above unit 4 Chernobyl NPP, etc. It was performed ICS safety reviews, designed for Ukrainian NPP not only by Ukrainian companies, but also by the largest foreign companies - "Westinghouse" (USA), "Siemens" (Germany), "Skoda-Controls" (Czech Republic).
- Development of methods of safety analysis:
 - Fukushima NPP accident lessons related to ICS and the participation in the Ukrainian NPP activity following from these lessons (stress- tests, the implementation of post-accident monitoring systems, reserve diesel generators control systems, black boxes systems, etc.)
 - I&CS for NPP units life extension,
 - development of NPP I&CS hardware ageing investigations, including analysis of the drift of their characteristics in time,
 - impact of I&CS failures on NPP violations,
 - new NPP I&CS safety important digital systems reliability in operation conditions.

- determination of ICS NPP functional safety measures.

My international activities in NPP I&CS safety included:

- the activity in International Atomic Energy Agency (IAEA) as a member of Technical Working Group on NPP Instrumentation and Control (1999-2011). Co-author of IAEA documents, the main of them - IAEA Safety standard SSG-39 [21] and book "Core knowledge of I&CS in NPP" [22],
- the activity in International Electrotechnical Commission (IEC) as an expert in Subcommittee SC-45 A "Instrumentation, control and electric power systems of nuclear facilities" of Technical committee TC-45 "Nuclear Instrumentation" (from 2002 to now).
- Participation in elaboration of different IEC standards, in IEC meetings, collaboration with regulatory bodies and organizations for technical support USA, Germany, France, Bulgaria and other countries.
- the organization in Kharkov by Kharkov Department SSTC NRS 1-5 International Scientific Technical Conferences "NPP Instrumentation and control systems: safety aspects" with participants of countries from Argentina to Hong Kong, including USA, Russia, Germany, France, etc.
- Speaker at 7 (2000- 2015) American Nuclear Society conferences "Nuclear power instrumentation, control and human-machine interface technologies (NPIC &HMIT)".

My activity in NPP Safety related to a lot of business trips in the world. The map of the world with marked places of business trips is shown in fig.3.

As the result of the work were published the books:

- Two books on I&CS NPP safety published in Russian in Ukraine in 2004 [23] and 2011 [24], where I was an editor and one of the authors. Book [23] was translated into English by the US Nuclear Regulatory Commission.
- Two books in English published by IGI Global (USA) on the safety and cybersecurity of NPP I&CS in 2014 [25] and in 2020 [26] edited by myself and prof. V. Kharchenko. The authors of most chapters of these books were my colleagues from SSTC NRS and V. Kharchenko's colleagues from the National Airspace University KhaI (Ukraine).
- In 2022 SSTC NRS began to publish the books series to 30th anniversary of SSTC NRS (in Ukrainian). Among them there is the book [27] prepared by SSTC NRS Kharkiv subsidiary with the results of our works devoted to NPP I&CS safety. New direction of work related to NPP safety- safety of small modular reactors [28].

2. Safety of Nuclear Power Plants (NPP's)

Problems of NPP's safety have always been urgent since NPP's began to operate. In Ukraine the work on NPP safety was especially vital after the accident at Chernobyl NPP- the most severe and widely known in the world of all technogenic accidents. In all countries, where NPP's are operated, there are state organizations that regulate the state policy in the field of NPP safety. Currently, such an organization in Ukraine is called the State Nuclear Regulatory Inspectorate of Ukraine (SNRIU). It includes Technical Support Organization - State Scientific and Technical Center for Nuclear and Radiation Safety (SSTC NRS), also subordinate to the National Academy of Sciences of Ukraine.

The word "regulation" has gone through my life. My first business trip was to test regulators for a steam boiler. This was followed by work on the analysis of dynamic characteristics of technological processes automatic regulators, the results of which are given in the book "Industrial Automatic Regulators" [17]. And then, for many years, safety regulation of NPP's and other critical facilities was the area of my activity.

SNRIU and SSTC NRS were based on the specialists who worked at the Chornobyl NPP during the accident and later connected with the restoration of the Chornobyl NPP after the accident. The first SNRIU head and creator was N.A. Steinberg, who was appointed to the position of chief engineer Chornobyl NPP immediately after the accident. The former employees of the Chornobyl NPP who joined SNRIU and SSTC NRS felt on their shoulders what it is NPP safety. I had a lot to learn from these people because after the USSR collapse my activity related to SSTC NRS, which included Kharkov Reliability Department as Kharkov subsidiary. Main objects of SNRIU and SSTC NRS activities were 4 operating NPPs of Ukraine with 13 (after commissioning of two power units in 2004 - 15) power units. Besides, the objects of activity are NPP spent fuel storage facilities, "Shelter" facility above unit 4 of Chornobyl NPP, research nuclear reactors, various sources of ionized radiation, etc. Starting from 1993, for long time I have been supervising a set of activities on nuclear and radiation safety regulation, defined by instrumentation and control systems (I&CS) for all these objects.

The main directions of my work on safety were:

- Development of SNRIU regulatory requirements to NPP I&CS and their components (hardware, software, software-hardware complexes) according to nuclear and radiation safety criteria. The document [18] containing such requirements was issued in 2000. The I&CS creation of two new VVER-1000 power units and I&CS modernization of many Ukrainian units for their life extension was carried out in accordance with the requirements of the document [18]. In 2015, this document was revised [19] considering the experience of Ukrainian NPPs operation, the lessons of Fukushima NPP accident, and changes in the international requirements (first, the International Atomic Energy Agency (IAEA) and International Electrotechnical Commission (IEC). In addition, the requirements were divided into regulatory requirements established by SNRIU, and technical requirements established by the operating organization – National Energy Company "Energoatom" [20]. According to [18] and [19], Ukrainian companies developed I&CS for NPPs in Bulgaria, the Czech Republic, Armenia, and other countries. The regulatory requirements contain methods of assessment of compliance with of safety important ICS NPP with the requirements for nuclear and radiation safety, requirements for the procedure and content of work for the I&CS life extension, etc.

- More than 1000 state nuclear and radiation safety expertise (safety reviews) of I&CS were fulfilled, designed for Ukrainian NPP not only by Ukrainian companies but also by the largest foreign companies - "Westinghouse" (USA), "Siemens" (Germany), "Skoda-Controls" (Czech Republic).

Development of methods of safety analysis:

- Fukushima NPP accident lessons related to I@CS lessons (stress- tests, the implementation of post-accident monitoring systems, reserve diesel generators, black boxes systems, etc.)

- I&CS for NPP unit's life extension, NPP I&CS hardware aging investigations, including analysis of the drift of their characteristics in time, impact of I&CS failures on NPP violations,

- Comparison of NPP and other systems safety development and safety assurance principles (launch vehicles control systems for nuclear warheads -example most powerful intercontinental ballistic missile SS-18 "Satan" [21] together with the Chief Constructor of this system Y. Aizenberg; control systems for rocket-cosmic complexes [22] together with Prof. V. Kharchenko and V. Scliar; hydraulic power plant-example emergency on Saiano-Sushenskaia plant [23] together with Prof. S. Artuh)

- Analysis of new NPP I&CS safety important digital systems reliability in operation conditions.

My international activities in NPP I&CS safety included:

- the activity in IAEA as a member of the IAEA Technical Working Group on NPP Instrumentation and Control (1999-2011). Co-author of IAEA documents, the main of them - IAEA Safety standard SSG-39 [24] and book "Core knowledge of I&CS in NPP" [25],
- the activity in IEC as an expert in Subcommittee SC-45 A "Instrumentation, control and electric power systems of nuclear facilities" of Technical Committee TC-45 "Nuclear Instrumentation" (from 2002 to now). Participation in the elaboration of different IEC standards, in IEC meetings,
- collaboration with USA, Germany, France, Bulgaria and other countries regulatory bodies and organizations for technical support,
- the organization in Kharkov-by-Kharkov Department SSTC NRS 1-5 International Scientific Technical Conferences "NPP Instrumentation and control systems: safety aspects" with participants from different countries from Argentina to Hong Kong, including USA, Russia, Germany, France, etc.
- Speaker at 7 American Nuclear Society conferences "Nuclear power instrumentation, control and human-machine interface technologies (NPIC & HMIT)".

My activity in NPP Safety related to a lot of business trips around the world. The map of the world with marked places for business trips is shown in Fig.3.

As a result of the work, the books were published:

- Two books on I&CS NPP safety were published in the Russian language in Ukraine in 2004 [26] and 2011 [27], where I was an editor and one of the authors. Book [26] was translated into English by the US Nuclear Regulatory Commission.

- Two books in English were published by IGI Global (USA) on the safety and cybersecurity of NPP I&CS in 2014 [28] and in 2020 [29] edited by myself and Prof. V. Kharchenko.

SSTC NRS published the book series on the 30th anniversary of SSTC NRS (in Ukrainian). Among them, was the book [30] prepared by SSTC NRS Kharkiv subsidiary with the results of our works devoted to NPP I&CS safety.



Figure 3. My activity in NPP Safety

The new direction of my work is related to NPP safety - small modular reactors (SMR) ([31]). In this definition: small- physically a fraction of the size of a conventional nuclear power reactor, modular-making it possible for systems and components to be factory-assembled and transported as a unit to location for installation, reactor- harnessing nuclear fission to generate heat to produce energy.

My colleagues from SSTC NRS S.A. Trubchaninov and I.I. Chervonenko are now prepared "M.A. Yastrebenetsky. Biographical indexes. Aladin Print. Kharkov. 2024 (In Ukrainian, Russian, English)" [34]. (Author of preface - Academic A.V. Nosovsky, Director of "Nuclear Power Plant Safety Institute" of Ukrainian National Academy of Science). The common number of my publications in this document included 350 names.

3. Big Safety

The definition of "Big Safety" for safety-critical technical and organization-technical systems was proposed and introduced in our articles with Prof. V. Kharchenko in 2020-2021 ([32], [33]). Big Safety is a result of a crossing of the following 10 attributes:

- big/complex systems,
- big safety properties (types of Big Safety)- nuclear, radiation, fire, information, transport, functional, ecological, infection safety, protectability from natural hazards, from military hazards,
- big/complex environment,
- big consequences from fatal failures,
- big number of fatal failures reasons,
- big data,
- big requirements for safety,
- big time system development,
- big toolbox for safety assessment and assurance,
- big resources for safety assurance.

A typical example of Big Safety is NPP safety.

Unfortunately, time leads to the originating of new types of Big Safety. Not so far, the problem of cybersecurity raised. Infection safety was included in some Big Safety types after the COVID-19 appearance. Military actions in Ukraine made it relevant to calculate as protectability of non-military objects (NPP, hydraulic and thermal power plants, offshore gas pipelines for military hazards. What may be the next?

4. My nearest teachers and colleagues

In this anniversary article I would like to thank my teachers and colleagues in reliability and safety, who have provided me with great assistance throughout my life



Academician Boris Vladimirovich Gnedenko (B.V.) (1912-1995)

B.V.'s works on probability theory are widely known throughout the world - they were the pride of Soviet science, and many generations of specialists were educated in them. My acquaintance with B.V.'s textbook "Probability Theory Course" dates to 1961. B.V. quickly appreciated that reliability theory is based on works on probability theory, and in 1965 a classic appeared - the book by B.V., A.D. Soloviev, Yu.K. Belyaev [34]. I'd like to say about signatures on this book: A.D. Soloviev's "To dear Mikhail Anisimovich from one of the authors" and B.V.'s "Now from two." Personal acquaintance with B.V. took place at the All-Union school - meeting "Theory of queueing systems", headed by B.V. in Dilijan (Armenia) in 1970, where I met many of my future

colleagues for life. It was "a holiday that is always with me." B.V. invited me to report on my works, and one day he even asked me to chair a meeting. Then I participated in the following schools in Pushchino-on-Oka and Zagulba (Azerbaijan). The role of B.V. in my fate is extremely high. B.V. invited me to participate in the seminars he led at Moscow University, supported my doctoral dissertation and suggested opponents for it, and wrote a preface to my book [15]. He visited me in Kharkov and spoke there at my seminar. I am glad to have friendly relations with B.V.'s descendants - his son Dima, who works at the Department of Probability Theory at Moscow State University, and in Boston (USA) - with B.V. granddaughter Katya.



Professor Alexander Dmitrievich Soloviev (A.D.) (1927-2001)

The next gift that the school in Dilijan gave me was meeting A.D. A.D.'s whole life related to Moscow State University - he graduated from it and taught there, and for many years - at the department headed by B.V. In Dilijan we tried not to leave A.D. - we listened to his stories and tales. A.D. reviewed my articles in the journal "Izvestiya of the USSR Academy of Sciences. Technical Cybernetics" and it is not surprising that B.V. suggested him as my opponent for my doctoral dissertation. We often met at his home, discussing the results. And when A.D. came to Kharkov for my defense, I did not recognize him - he sacrificed his luxurious long artistic hair for me - "So that there would be no negative reaction from conservative members of the Kharkov Academic Council." After the main defense, he again defended me in the Academic Council of the Moscow Bauman Higher Technical School, where the Higher Attestation Commission sent my work. A.D.'s authority was indisputable. A.D. twice (in 1974 and 1977) came to Kharkov for my seminar, where his reports attracted a full hall of listeners, visited my home and his favorite place - the Kharkov bazaar, charmed my entire family and, to the horror of the women, he himself went to show off his culinary skills in our kitchen.



Professor Igor Alekseevich Ushakov (I.A.) (1933-2015) (Fig.),

I.A. - a pupil of B.V. Gnedenko, was the leader of my generation of reliability specialists. Much of what he did is associated with the word "first" - the first reference book on reliability calculations, repeatedly reprinted in different languages, the first works on optimal redundancy, on the analysis of complex networks, etc. I got acquainted with I.A. before Dilijan and met him many times in Moscow, in Kharkov, in Irkutsk, in Leningrad, (where I.A. arrived one summer in a light shirt with a tennis racket, and there it snowed,) and after that in San Diego. His talents were diverse and energy colossal - he worked at the research institute on a closed topic. headed departments at universities, supervised countless dissertations, supervised reliability journals, wrote a series of children's books "History of Science through the Prism of Insights", wrote wonderful memoirs, drew and played football. I.A.'s results to domestic science should have made him a member of the USSR Academy of Sciences, but this did not happen because of envy and lack of objectivity. I.A. moved to the USA, taught, and worked in well-known companies. He came up with the idea of uniting reliability specialists who remained in the USSR, those who left the USSR abroad, and foreign specialists in an association under the name of his teacher - "Gnedenko-Forum". I was proud that I.A. offered me to become one of the next vice presidents (and then president)) of "Gnedenko-Forum".



Professor Volodymir Vladimirovich Rykov (V.V.)

Of all my colleagues, the duration of my contacts with V.V. is the longest - we came to work at CSRiCA almost simultaneously - he to the computing center after Moscow Lomonosov University (The head of his diploma was Acad. A.N. Kolmogorov), me- I to the dynamics department. V.V. was the editor of my book [14] and a co-author of an article [9] on regenerative processes with several types of regeneration points. And he also gave me a bed in his apartment in Praslin during difficult times. I continued to work at CSRiCA until the collapse of the USSR. V.V. went to universities-at the Gubmint Institute of Petrochemical and Gas Industry and at the University of Peoples Friendship. He is a professor at both to this day. V.V. is a world-famous specialist in reliability and queueing systems". He retains great energy. His books and articles have been published in Russia and abroad. He implements the interface between specialists from Russia and the rest of the world through various conferences and meetings. In addition, VV is the editor-in-chief of the journal "[Reliability: Theory & Applications](#)". Of all my colleagues, V.V. is the absolute record man of the number of grandchildren () and great-grandchildren!



Professor Boian Dimitrov (B.D.)

B.D. was a postgraduate student at Moscow Lomonosov University- he was a pupil of Acad. Gnedenko and he had contact with this school all his life. My acquaint with him began from Dilijan. B.D. long time lived in the USA; he was a professor at Kettering University. Thank him, V.V. Rykov was invited to this university for 2 years. B.D. change me to the position of Gnedenko -Forum President. I think that he is more energetic than me.



Professor Victor Alekseevich Kashtanov (V.A.)

V.A. graduated Department "Probability Theory" in Moscow Lomonosov State University and after that he was the first post-graduate student of Prof. A.D. Soloviev. During long time V.A. was head of Department and Dean of the Moscow Institute of Electronics and Mathematics, named after A. N. Tikhonov. Now he is an ordinary Professor of Department of Applied Mathematics of National Research University's "Higher School of Economics". He is an Honors Scientist of Russia. V.A. scientific results related to controlled stochastic models, controlled random processes, theory of complex systems service, methods of assurance complex systems reliability. Now V.A. is Vice-President (and future President) of Gnedenko- Forum. I had contact with V.A. for many years.



Professor Igor Borisovich Shubinsky (I.B.)

I.B. graduated in 1961 from Kyiv High Engineer Radio Technic School (KVIRTU). This was a very non-standard school - leading specialists in probability theory and reliability were the teachers, and my near colleague in reliability and friend General Nikolai Shishonok was the deputy of the chief. I first time met I.B. in the Leningrad area in 1970 in the military form- he served in the Pushkin military school of radio-electronic. In 1989 colonel I.B. ended military activity and began to work/ at Leningrad University instead of famous specialist Prof. A.M.Polovko. Then I.B. created the power organization "Information Security in Railway Transport ", I.B. is the author of an - interesting book about the analysis of information system's functional reliability. I didn't see I.B. for a long time. After that, I met him at the Second International Symposium on Stochastic Models in Reliability Engineering, Life Science, and Operations Management, 2016, in Beer Sheva, Israel. Now we have regular contacts during "Gnedenko Forim" meetings. I add, that I.B. is the chief redactor of the scientific-technical journal "Dependability", I am his deputy.



Professor Viacheslav Sergeevich Kharchenko (V.S.)

The number of common works with co-authors in my bibliographic indexes [34] is the maximum together with V.S., more than 20. There are books in Russian and English [26, 27, 28, 29] and a lot of articles. V.S. graduated from Kharkiv High Command Engineering Military College of Rocket Troops where his scientific supervisor was my friend Prof. Alexander Larin- an honored scientist of Ukraine and Russia. V.S. was a “real colonel”- his military activity ended as head of the department of the Kharkiv Military University. V.S. is the author of more than 700 inventions (!) and an Honored Inventor of Ukraine. Now V.S. is the head of the department “Computer Systems, Networks, and Cybersecurity” At National Aerospace University’s Kharkiv Aviation Institute”. The energy of V.S. is fantastic- he prepared more than 60 PhD and D.Sc. and organizes every month seminar “Critical Computer Systems and Technology” (2001 - at present time), he is a founder and general chair of international conference “Dependable Systems, Services and Technologies” (IEEE DESSERT, 2006 - at present time), he was the President of Ukrainian Scientific-Educational IT Society (2018-2023). I fulfilled all my work in the Big Safety area together with V.S. Today we live in different countries, but we can have regular contact. We continue our common work!



Dr. Sc. Alexander Vladimirovich Bochkov (A.V.)

A.V., as I.A. Ushakov, graduated from the Moscow Aviation Institute. I wrote higher about Gnedenko-Forum. Together with I. A. Ushakov , A.V. was the so-founder Gnedenko-Forum and for many years he fulfilled a huge work devoted to its development, he is the Scientific Secretary of the e-Journal “[Reliability: Theory & Applications](#)”. Area of his scientific interests - safety very big systems, and now he is Scientific Secretary of the Scientific and Technical Council Research and Design Institute for Information Technology, Signaling and Telecommunications in Railway Transportation – NIIAS, JSC. I became acquainted with A.V. thanks to our invitations to Hong Kong City University President Prof. Way Kyo. One of his colleagues at this university was a famous specialist in reliability prof. Singpurvalla. Many talented people are very rare events. But A.V. is included in this class. He is lyrical poet, and I have his 7(!) books on my shelf. Every poem is accompanied by the corresponding picture of different artists. I’m waiting music and song on A.V. poems!

References

- [1] Yastrebenetsky M. Generations of Yastrebenetsky. Part 1. Kharkov, Aladin Print. 2014 (in Russian).
- [2] Yastrebenetsky M., Solyanik B., Komarov G. Determining the reliability of automatics controllers in a thermal power station. Thermal Engineering. 1965. No4
- [3] Yastrebenetsky M. A class of regenerative random processes. Engineering Cybernetics. 1969, No5
- [4] Yastrebenetsky M. Reliability of unrenewable elements under external disturbances. Engineering Cybernetics. 1970, No3
- [5] Yastrebenetsky M. Reliability of renewable elements under external disturbances. - Engineering Cybernetics. 1970. No4
- [6] Yastrebenetsky M., Solyanik B. Model of failure flow of automatic equipment in case of random stationary external disturbances. Automation and remote control. 1969. No7
- [7] Yastrebenetsky M. On a rarefied Markov renewal process connected with some reliability problems. Engineering Cybernetics. 1972. No4.
- [8] Yastrebenetsky M. A flow of failures caused by crossing of nonconstant level by a regenerating random process. Engineering Cybernetics. 1974. No3.
- [9] Yastrebenetsky M., Rykov V. About the regenerative process with some types of regeneration points. Cybernetics. Kyev. 1971. No3.
- [10] USSR GOST 25804.3-83. Apparatus, instruments, for nuclear power plant devices, equipment of nuclear power plants technological process control systems. Requirements to reliability. 1983. M. Standard publishing house. (in Russian).
- [11] USSR GOST 24701-86. Unified system of standards of automatic control systems. Reliability of automatic control systems. Main positions. 1986. M. Standard publishing house (in Russian).
- [12] RTM 25 850-87. Automatic control systems for NPP technological process. Reliability. Common requirements. M. Minpribor. 1987. (in Russian).
- [13] Yastrebenetsky M., Solianik B. Evaluation of industrial automatic apparatus in operation conditions. M. Energy. 1968 (in Russian).
- [14] Yastrebenetsky M., Solianik B. Reliability of industrial automatic systems in operation conditions. Failures flow and methods of their statistical processing. M. Energy. 1978. (in Russian).
- [15] Yastrebenetsky M. Reliability of instrumentation for technological processes automatic control systems. M. Energyizdat. 1982 (in Russian).
- [16] Yastrebenetsky M., Ivanova G. Reliability of technological processes automatic control systems. (For university students). M. Energyatomizdat. 1989. (in Russian).
- [17] Steinberg S., Hvilivitsky L., Yastrebenetsky M. Industrial automatic regulators. M. Energy. 1973. (in Russian).
- [18] NP 306.5.02/3.035-2000. Nuclear and radiation safety requirements to instrumentation and control systems important to NPP safety. Kiev (K.). State administration of nuclear regulation. 2000. (in Russian).
- [19] NP 306.5.02/3.035-2015. Nuclear and radiation safety requirements to instrumentation and control systems important to NPP safety. Kiev (K.). SNRIU. 2015 (in Russian).
- [20] SOU NAEK 100:2016. Engineering, scientific and technical support. Instrumentation and control systems, safety important to NPP. NAEK "Energoatom". 2016. (in Russian).
- [21] Aizenberg Y., Yastrebenetsky M. The comparison of safety assurance principles of control systems for launch vehicles and nuclear power plants. Cosmic science and technology. 2002, №1 (In Russian)
- [22] Sclar V., Kharchenko Yastrebenetsky M. Digital information and control systems NPP and rocket-cosmic complexes: comparison analysis, tendential grow, safety assurance. Nuclear and radiation safety. 2004, №2 (In Russian)
- [23] Artuh S., Yastrebenetsky M. Accident at Saiano-Sushenskoi Hydraulic power plant. Lessons for Ukraine/ "2000", №49 (In Russian)

- [24] SSG-39. Design of Instrumentation and Control Systems for Nuclear Power Plants EA Safety Standards. Specific Safety Guide. 2016. IAEA. Vienna
- [25] NP-T-3.12. Core knowledge on instrumentation and control systems in nuclear power plants. IAEA Nuclear Energy Serial. IAEA. 2011. IAEA. Vienna
- [26] Yastrebenetsky M., Goldrin V., Kharchenko V., Rozen Y., etc. (Editor – Yastrebenetsky M), Nuclear power plants safety. Instrumentation and control systems. K. Technika. 2004.
- [27] Yastrebenetsky M., Jonhson G. Eliseev V., Kharchenko V., etc. (Editor -Yastrebenetsky M.), Nuclear power plants safety. Nuclear reactors control systems and protection systems..K. Osnova-Print. 2011 (in Russian).
- [28] Yastrebenetsky M., Kharchenko V. (Editors). Nuclear power plant instrumentation and control systems for safety and security. USA. Hershey. IGI Global. 2014.
- [29] Yastrebenetsky M., Kharchenko V. (Editors). Cyber security and safety of nuclear power plant instrumentation and control systems. USA. Hershey. IGI Global. 2020.
- [30] Trubchaninov S., Klevtsov O., Yastrebetsky M., Rozen Y., Simonov A., Chervonenko I., etc. Safety aspects of Ukrainian NPP information and control systems. SSTC NRS. K. 2022 (in Ukrainian)
- [31] Bregnev E., Fesenko G., Kharchenko V., Yastrebenetsky M. Safety of small modular reactors digital infrastructure: new requirements and challenges. Electronic modeling, 2023, No 4 (in Ukrainian).
- [32] Kharchenko V., Yastrebenetsky M. NPP Safety and Big Safety in Coronavirus time. Nuclear and Radiation Safety. 2020, No3 (in Russian).
- [33] Kharchenko V., Yastrebenetsky M. About the concept of big safety. Reliability: Theory and Application 2021, No.1
- [34] M.A.Yastrebenetsky. Biographical indexes. Aladin Print. Kharkov. 2024 (In Ukrainian, Russian, English)". -350. Preface - Academic A.V. Nosovsky, Director of "Nuclear Power Plant Safety Institute" of Ukrainian National Academy of Science. Compilers- S.A.Trubchaninov and I.I.Chervonenko.
- [35] Gnedenko B., Soloviev A., Beliaev Ya. Mathematical methods in reliability theory. - M. Nauka. 1965.

ALGORITHMS FOR APPROXIMATING A FUNCTION BASED ON INACCURATE OBSERVATIONS

G.SH. TSITSIASHVILI¹, M.A. OSIPOVA^{1,2}



¹Institute for Applied Mathematics, Far Eastern Branch of Russian Academy Sciences, Russia

²Far Eastern Federal University, Russia

guram@iam.dvo.ru, mao1975@list.ru

Abstract

This paper is devoted to the approximation of a function by a trigonometric polynomial based on its inaccurate values at selected points. Two methods of observation are considered. The first method is to make observations at points evenly distributed on the segment where the function is specified. The second method is to take observations at the points of division into a finite number of equal parts of the neighbourhoods of the selected points. Upper estimates of the standard deviation of the function from trigonometric polynomials are constructed and the rate of their convergence is estimated. Differences were found in the computational complexity of these approximations and in the number of observations of the function values at the selected points. Thus, the problem of approximating a function from inaccurate observations of their values at selected points is a multi-criteria one and its solution depends on the choice of observation points.

Keywords: trigonometric functions, inaccurate observations, error in function estimation, experimental plan

1. INTRODUCTION

This paper is devoted to the approximation of the periodic function from inaccurate observations. To solve this problem, Chebyshev, Hermite, Jacobi, Laguerre polynomials, trigonometric polynomials are used for the exact values of the function at the selected points (see, for example, [1], [2]). However, the task becomes significantly more complicated if it is necessary to evaluate the function based on inaccurate observations at selected points. In this case, there are many different solutions that need to be compared by various indicators (solution error, computational complexity, number of observation points). In this paper, two solutions to this problem are proposed and compared.

Trigonometric polynomials were used to approximate the function from inaccurate observations. The first method of approximation consists in observing the function at points evenly distributed over the segment of its assignment. In the second method, observations are considered at the points of division into a finite number of equal parts of the neighbourhoods of the selected points. In both cases, upper estimates of the standard deviation of the approximation of the function from its exact value are constructed. Despite the proximity of the upper estimates obtained, differences in the computational complexity of these approximations are found in the number of observations of the function values at the selected points.

The proposed algorithm for estimating the value of a function by a trigonometric polynomial using inaccurate deterministic or stochastic observations, unlike classical algorithms, allows us to estimate the rate of convergence of the estimates obtained to the estimated parameters.

And considering a small interval of time observation makes it possible to build an experiment planning procedure. The authors previously used this idea to solve the problem of estimating the parameters of a number of ordinary differential equations and their systems, partial differential equations [3], [4].

The paper considers a function $f(x)$ that is continuously differentiable on the segment $[0, 2\pi]$. This function decomposes into a Fourier series in the space $L_2[0, 2\pi]$. Denote

$$f_n(x) = \frac{\bar{a}_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)), \quad (1)$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx, \quad k = 0, 1, \dots, n.$$

It is known (see, for example, [5]) that under given conditions for the function $f(x)$ there exists a number C such that $|a_k| \leq C/k$, $|b_k| \leq C/k$, $k = 1, \dots$, and, therefore,

$$\pi \sum_{k=n+1}^{\infty} (a_k^2 + b_k^2) = \int_0^{2\pi} (f(x) - f_n(x))^2 dx = D(n) = O(n^{-1}). \quad (2)$$

In this paper, we will consider two different estimates of the function $f(x)$ based on inaccurate observations. The first estimate of $\hat{f}_n(x)$ of the function $f(x)$ is constructed as follows. Let $t_{p,k}^a$, $t_{p,k}^b$, $\varepsilon_{p,k}^a$, $\varepsilon_{p,k}^b$, $p = 1, \dots$, $k = 0, \dots$, independent random variables. Moreover, the random variables $t_{p,k}^a$, $t_{p,k}^b$ have a uniform distribution on the segment $[0, 2\pi]$, and random variables $\varepsilon_{p,k}^a$, $\varepsilon_{p,k}^b$ characterizing measurement errors have zero mean and variance σ^2 . Then we define random variables

$$\hat{a}_k = \frac{2}{m} \sum_{p=0}^{m-1} (f(t_{p,k}^a) + \varepsilon_{p,k}^a) \cos(kt_{p,k}^a), \quad \hat{b}_k = \frac{2}{m} \sum_{p=0}^{m-1} (f(t_{p,k}^b) + \varepsilon_{p,k}^b) \sin(kt_{p,k}^b) \quad (3)$$

and we will make the first assessment

$$\hat{f}_n(x) = \frac{\hat{a}_0}{2} + \sum_{k=1}^n (\hat{a}_k \cos(kx) + \hat{b}_k \sin(kx)). \quad (4)$$

The second estimate $\bar{f}(x)$ of the function $f(x)$ is based on inaccurate observations in the following way. Let $x_p = 2\pi p/m$, $p = 0, \dots, m-1$, and random variables $\varepsilon_{p,j}$, $p = 0, \dots, m-1$, $j = -(2N+1), \dots, 2N+1$, characterizing measurement errors are independent, and have zero mean and variance σ^2 . Let's assume random measurements of quantities $f(x_p)$ equal

$$\bar{f}(x_p) = \frac{1}{2N+1} \sum_{j=-N}^N (f(x_p + jh) + \varepsilon_{p,j}), \quad h = N^{-\alpha}.$$

Let's construct a second estimate of the function $f(x)$

$$\bar{f}_n(x) = \frac{\bar{a}_0}{2} + \sum_{k=1}^n (\bar{a}_k \cos(kx) + \bar{b}_k \sin(kx)). \quad (5)$$

$$\bar{a}_k = \frac{2}{m} \sum_{p=0}^{m-1} \bar{f}_p \cos(kx_p), \quad \bar{b}_k = \frac{2}{m} \sum_{p=0}^{m-1} \bar{f}_p \sin(kx_p), \quad k = 0, 1, \dots, n.$$

Our task is to build upper bounds

$$M \int_0^{2\pi} (f(x) - \hat{f}_n(x))^2 dx, \quad M \int_0^{2\pi} (f(x) - \bar{f}_n(x))^2 dx$$

and compare them.

2. THE ERROR OF THE FIRST ESTIMATE

Obviously, the inequality is fair

$$\begin{aligned} M \int_0^{2\pi} (f(x) - \widehat{f}_n(x))^2 dx &= \int_0^{2\pi} M[(f(x) - f_n(x)) + (f_n(x) - \widehat{f}_n(x))]^2 dx \leq \\ &\leq 2 \left[\int_0^{2\pi} (f(x) - f_n(x))^2 dx + M \int_0^{2\pi} (f_n(x) - \widehat{f}_n(x))^2 dx \right]. \end{aligned}$$

In turn,

$$M \int_0^{2\pi} (f_n(x) - \widehat{f}_n(x))^2 dx = \pi M \left[\left(\frac{a_0 - \widehat{a}_0}{2} \right)^2 + \sum_{k=1}^n (a_k - \widehat{a}_k)^2 + \sum_{k=1}^n (b_k - \widehat{b}_k)^2 \right].$$

Then it is not difficult to get from the formulas (4)

$$\begin{aligned} M\widehat{a}_0 = a_0 &\Rightarrow M \left(\frac{a_0 - \widehat{a}_0}{2} \right)^2 = \frac{1}{m} \left(\sigma^2 + \frac{1}{2\pi} \int_0^{2\pi} f^2(x) dx - a_0^2 \right), \\ M\widehat{a}_k = a_k &\Rightarrow M(a_k - \widehat{a}_k)^2 \leq \frac{4}{m} \left(\sigma^2 + \frac{1}{2\pi} \int_0^{2\pi} f^2(x) dx - a_k^2 \right), \quad k = 1, 2, \dots, \\ M\widehat{b}_k = b_k &\Rightarrow M(b_k - \widehat{b}_k)^2 \leq \frac{4}{m} \left(\sigma^2 + \frac{1}{2\pi} \int_0^{2\pi} f^2(x) dx - b_k^2 \right), \quad k = 1, 2, \dots \end{aligned}$$

It follows from this and from the formula (2) that

$$M \int_0^{2\pi} (f(x) - \widehat{f}_n(x))^2 dx \leq 2 \left[D(n) + \frac{\pi(8n+1)}{m} \left(\sigma^2 + \frac{1}{2\pi} \int_0^{2\pi} f^2(x) dx \right) \right] \quad (6)$$

and, therefore, due to the formulas (2), (6) we have

$$M \int_0^{2\pi} (f(x) - \widehat{f}_n(x))^2 dx = O(n^{-1}) + O(nm^{-1}). \quad (7)$$

In particular, for $n = [m^{1/2}]$ we have (here $[a]$ is the integer part of the real number a)

$$M \int_0^{2\pi} (f(x) - \widehat{f}_n(x))^2 dx = O(m^{-1/2}). \quad (8)$$

3. THE ERROR OF THE SECOND ESTIMATE

Let $\Delta f_p = \widehat{f}(x_p) - f(x_p)$, first evaluate $M(\Delta f_p)^2$. From the continuous differentiability of the function $f(x)$ by $[0, 2\pi]$ we have

$$\max \left(\sup_{0 \leq x \leq 2\pi} |f(x)|, \sup_{0 \leq x \leq 2\pi} |f'(x)| \right) = C < \infty.$$

Then

$$\begin{aligned} M(\Delta f_p)^2 &= M \left[\frac{1}{2N+1} \sum_{j=-N}^N (f(x_p + jh) + \varepsilon_{p,j}) - f(x_p) \right]^2 = \\ &= M \left[\frac{1}{2N+1} \sum_{j=-N}^N (f(x_p + jh) + \varepsilon_{p,j} - f(x_p)) \right]^2 = \end{aligned}$$

$$\begin{aligned}
 &= M \left[\frac{1}{2N+1} \sum_{j=-N}^N (f(x_p + jh) - f(x_p)) + \frac{1}{2N+1} \sum_{j=-N}^N \varepsilon_{p,j} \right]^2 = \\
 &= \left[\frac{1}{2N+1} \sum_{j=-N}^N (f(x_p + jh) - f(x_p)) \right]^2 + \frac{\sigma^2}{2N+1} \leq \\
 &\leq \left[\frac{1}{2N+1} \sum_{j=-N}^N C|j|h \right]^2 + \frac{\sigma^2}{2N+1} = O(h^2N^2) + O(N^{-1}).
 \end{aligned}$$

Therefore, the relation is fulfilled

$$\sup_{0 \leq p \leq m-1} M(\Delta f_p)^2 = O(h^2N^2) + O(N^{-1}).$$

In particular, for $h = N^{-3/2}$ we get

$$\sup_{0 \leq p \leq m-1} M(\Delta f_p)^2 = O(N^{-1}). \quad (9)$$

Using this definition of the estimate \bar{f}_p , we construct an estimate of the error of the function $\bar{f}_n(x)$.

Denote

$$f_n^*(x) = \frac{a_0^*}{2} + \sum_{k=1}^n (a_k^* \cos(kx) + b_k^* \sin(kx)), \quad (10)$$

$$a_k^* = \frac{2}{m} \sum_{p=0}^{m-1} f_p \cos(kx_p), \quad b_k^* = \frac{2}{m} \sum_{p=0}^{m-1} f_p \sin(kx_p), \quad k = 0, 1, \dots, n.$$

Consider

$$\begin{aligned}
 M \int_0^{2\pi} (f(x) - \hat{f}_n(x))^2 dx &= \int_0^{2\pi} M[(f(x) - f_n(x)) + (f_n(x) - f_n^*(x)) + (f_n^*(x) - \hat{f}_n(x))]^2 dx \leq \\
 &\leq 3 \left[\int_0^{2\pi} (f(x) - f_n(x))^2 dx + \int_0^{2\pi} (f_n(x) - f_n^*(x))^2 dx + \int_0^{2\pi} M(f_n^*(x) - \hat{f}_n(x))^2 dx \right]. \quad (11)
 \end{aligned}$$

Let's focus first on the assessment $\int_0^{2\pi} (f(x) - f_n^*(x))^2 dx$.

$$\begin{aligned}
 \int_0^{2\pi} (f_n(x) - f_n^*(x))^2 dx &= \int_0^{2\pi} dx \left[\frac{a_0 - a_0^*}{2} + \sum_{k=1}^n (a_k - a_k^*) \cos(kx) + \sum_{k=1}^n (b_k - b_k^*) \sin(kx) \right]^2 = \\
 &= \frac{\pi}{2} \left[(a_0 - a_0^*)^2 + 2 \sum_{k=1}^n [(a_k - a_k^*)^2 + (b_k - b_k^*)^2] \right].
 \end{aligned}$$

We have

$$\begin{aligned}
 (a_0 - a_0^*)^2 &= \left(\frac{1}{\pi} \int_0^{2\pi} f(x) dx - \frac{2}{m} \sum_{p=0}^{m-1} f(x_p) \right)^2 = \left(\frac{1}{\pi} \sum_{p=0}^{m-1} \int_{x_p}^{x_{p+1}} f(x) dx - \frac{2}{m} \sum_{p=0}^{m-1} f(x_p) \right)^2 = \\
 &= \left(\frac{1}{\pi} \sum_{p=0}^{m-1} \int_{x_p}^{x_{p+1}} (f(x) - f(x_p)) dx \right)^2 \leq \left(\frac{1}{\pi} \sum_{p=0}^{m-1} \int_{x_p}^{x_{p+1}} \frac{2\pi C}{m} dx \right)^2 \leq \frac{16\pi^2 C^2}{m^2}.
 \end{aligned}$$

It is not difficult to get when $x_p \leq x \leq x_{p+1}$, $k = 0, 1, \dots, n-1$,

$$\begin{aligned}
 |f(x) \cos(kx) - f(x_p) \cos(kx_p)| &\leq |f(x)| \cdot |\cos(kx) - \cos(kx_p)| + \\
 &+ |\cos(kx_p)| \cdot |f(x) - f(x_p)| \leq \frac{2\pi C(k+1)}{m}.
 \end{aligned}$$

It follows that

$$\begin{aligned} |a_k - a_k^*| &= \left| \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx - \frac{2}{m} \sum_{p=0}^{m-1} f(x_p) \cos(kx_p) \right| = \\ &= \left| \frac{1}{\pi} \sum_{p=0}^{m-1} \int_{x_p}^{x_{p+1}} (f(x) \cos(kx) - f(x_p) \cos(kx_p)) dx \right| \leq \frac{4\pi C(k+1)}{m}. \end{aligned}$$

Then

$$\sum_{k=0}^n (a_k - a_k^*)^2 \leq \frac{16C^2\pi^2(n+1)(n+2)(2n+3)}{6m^2} = O(n^3m^{-2}),$$

By analogy, we have

$$\sum_{k=0}^n (b_k - b_k^*)^2 \leq \frac{16C^2\pi^2(n+1)(n+1)(2n+3)}{6m^2} = O(n^3m^{-2}).$$

Therefore,

$$\int_0^{2\pi} (f_n(x) - f_n^*(x))^2 dx = O(n^3m^{-2}). \quad (12)$$

Let's now move on to the evaluation of $M \int_0^{2\pi} (f_n^*(x) - \bar{f}_n(x))^2 dx$, by calculating first

$$M \int_0^{2\pi} (f_n^*(x) - \bar{f}_n(x))^2 dx = \pi M \left(\frac{(a_0^* - \bar{a}_0)^2}{4} + \sum_{k=1}^n M(a_k^* - \bar{a}_k)^2 + \sum_{k=1}^n M(b_k^* - \bar{b}_k)^2 \right). \quad (13)$$

And then let's use the equalities

$$\begin{aligned} A &= \sum_{k=1}^n M(a_k^* - \bar{a}_k)^2 = \sum_{k=1}^n M \left(\frac{2}{m} \sum_{p=0}^{m-1} \Delta f_p \cos(kx_p) \right)^2 = \\ &= \frac{4}{m^2} \sum_{k=1}^n M \left(\sum_{p=0}^{m-1} \Delta f_p \cos(kx_p) \sum_{q=0}^{m-1} \Delta f_q \cos(kx_q) \right) = \\ &= \frac{4}{m^2} \sum_{p=0}^{m-1} \sum_{q=0}^{m-1} M \Delta f_p \Delta f_q \sum_{k=1}^n \cos(kx_p) \cos(kx_q). \end{aligned}$$

Therefore, we have

$$|A| \leq \frac{4n}{m^2} \sum_{p=0}^{m-1} \sum_{q=0}^{m-1} |M \Delta f_p \Delta f_q| \leq \frac{2n}{m^2} \sum_{p=0}^{m-1} \sum_{q=0}^{m-1} M((\Delta f_p)^2 + (\Delta f_q)^2) \leq 4n \sup_{0 \leq p \leq m-1} M(\Delta f_p)^2.$$

From these relations and the formula (9) we obtain $|A| = O(nN^{-1})$.

In turn,

$$M(a_0^* - \bar{a}_0)^2 = \frac{4}{m^2} M \left(\sum_{p=0}^{m-1} \Delta f_p \right)^2 = O(N^{-1}).$$

Then

$$\sum_{k=0}^n M(a_k^* - \bar{a}_k)^2 = O(nN^{-1}). \quad (14)$$

Similarly, it is not difficult to obtain equality

$$\sum_{k=0}^n M(b_k^* - \bar{b}_k)^2 = O(nN^{-1}) \quad (15)$$

Combining the formulas (13) - (15), we get

$$M \int_0^{2\pi} (f_n^*(x) - \bar{f}_n(x))^2 dx = O(nN^{-1}). \quad (16)$$

Finally from the formulas (2), (11), (12), (16) we come to the ratio

$$M \int_0^{2\pi} (f(x) - \bar{f}_n(x))^2 dx = O(n^{-1}) + O(n^3 m^{-2}) + O(nN^{-1}). \quad (17)$$

In particular, for $n = m^{1/2}$ and $N = m$ we get

$$M \int_0^{2\pi} (f(x) - \bar{f}_n(x))^2 dx = O(m^{-1/2}).$$

4. CONCLUSION

From formulas (8), (17) it follows that the estimation error $\bar{f}_n(x)$, as well as the estimation error $\hat{f}_n(x)$ are equal to $O(m^{-1/2})$. In turn, the number of observations in the first case is equal to $O(nm) = O(m^{3/2})$, and in the second case is equal to $O(Nm) = O(m^2)$. However, it should be noted that the formula for calculating the Fourier coefficients (5) can be made more economical using the fast Fourier transform method (see, for example, [6]).

This work was supported by the Ministry of Science and Higher Education of Russian Federation (Agreement No 075-02-2024-1440).

REFERENCES

- [1] Laurent P. J. Approximation et optimisation [Approximation and optimization], Editions Hermann, Paris, 1972.
- [2] Suetin P. K. Classical orthogonal polynomials, Nauka, Moscow, 1979.
- [3] Tsitsiashvili, G. Sh., Osipova, M. A., Kharchenko, Yu. N. (2022). Estimating the Coefficients of a System of Ordinary Differential Equations Based on Inaccurate Observations. *Mathematics*, 10(3), 502.
- [4] Tsitsiashvili, G. Sh., Osipova, M. A., Gudimenko, A. I. (2023) and Statistical Aspects of Estimating Small Oscillations Parameters in a Conservative Mechanical System Using Inaccurate Observations. *Mathematics*, 11(12), 2643.
- [5] Ivanov G.E. Lectures on mathematical analysis: in 2 parts: textbook, MIPT, Moscow, 2011.
- [6] Nussbaumer G. Fast Fourier transform and convolution calculation algorithms, Radio and Communications, Moscow, 1985.

AVAILABILITY AND PROFIT OPTIMIZATION OF CONTINUOUS CASTING SYSTEM OF THE STEEL INDUSTRY USING ARTIFICIAL NEURAL NETWORK TECHNIQUE

SHIKHA BANSAL¹, SOHAN LAL TYAGI^{2*}, URVASHI³

•

1,2*. Assistant Professor, Department of Mathematics, SRM Institute of Science
and Technology, Delhi-NCR Campus, Ghaziabad, 201204, India
srbansal2008@gmail.com, drsohanttyagi@gmail.com *

3. Research Scholar, Department of Mathematics, SRM Institute of Science
and Technology, Delhi-NCR Campus, Ghaziabad, 201204, India
urvashigodara8@gmail.com

* Corresponding author

Abstract

The main purpose of this study was to optimize the performance parameters of the casting process using a neural network approach. The casting process is molten or liquefied metal is poured into the mould cavity, after solidification it takes the near net shape of the cavity. The entire manufacturing process goes through six stations viz Pouring turret "ladle", "Tundish", "Mould", "Water spray chamber", "Support roller" and "Torch cutter". The Artificial Neural Network (ANN) technique is used in this paper to analyze the casting system's availability, profitability, and state probability variation. An effort has been made to identify the most critical component of the system. The outcomes of the analysis will help the practitioners in deciding effective maintenance strategies.

Keywords: Neural Network Approach, Neural weights, Profit, Availability

1. INTRODUCTION

The business sector must adopt the paths of quick technological progress to survive the current business climate. The absolute necessity for a corporation to survive is now massively automated systems. Many conventional industrial materials are finding it difficult to fulfill the demands of today's industries as industrial manufacturing technologies continue to advance. The manufacturers' primary focus is on the development of systems with lower development costs so that maximum profit may be gained, and they have succeeded in doing so to a significant extent by incorporating some performance improvement approaches into their systems. Even if such massive systems will eventually fail, the damage can be reduced by increasing their availability. Because of this, the availability of these systems has emerged as the most important factor in process industries. They can better comprehend the impact of variations in failure and repair rates of different subcomponents and units on the system by analyzing performance characteristics. Tzong [1] emphasized improving cast parts availability and dependability while attempting to

minimize them as much as feasible. Agarwal and Bansal [2] use high-level repair specialties susceptible to shifting environmental circumstances to boost system dependability. Agarwal et al. [3] applied for cold standby redundancy in order to determine different reliability metrics. Bansal et al. [4] illustrate how to solve milk powder production facilities using the Boolean function technique and how mission phases affect the system's reliability. Amit Kumar [5] employed a probabilistic technique that helps the foundry plant's quality control departments increase production by minimizing cast component rejections, which raises dependability and lowers casting process costs. Agarwal and Bansal [6] examined the costs and dependability of a multi-component system, such as a thermal power plant, utilizing the supplemental variable technique to ensure the system's maximum dependability. Kumar [7] gives a thorough grasp of the casting industry to explore the several factors influencing the foundry industry's casting process and maximize casting quality, assuming adequate repair facilities are constantly accessible. Bansal [8] enhanced the availability of a repairable system by introducing the preemptive resume Repair Discipline. Bansal et al. [9] suggested an orthogonal matrix approach and the reliability was calculated when the exponential time distribution and Weibull were used to express the failure rate. Bansal and Tyagi [10] recognized the system's weak points so that suitable maintenance techniques can be implemented to strengthen the system's maintenance policy. Bansal et al. [11] Utilised the Markov Birth-Death Process, focusing on the leaf spring plant's performance to ascertain the system's availability. Chaudhari and Vasudevan [12] concentrated on creating a casting failure Markov chain model. It demonstrates how this predictive modeling technique may be used to produce the chance of casting failure through different casting faults and casting process parameters. Chaudhary and Bansal [13] offer improved standards for upcoming proposals and execution that will act as a basis for planning the expansion of hydroelectric power plant generation. Godara and Bansal [14] discussed the use of neural networks and Boolean functions as two separate methods for estimating the reliability of a steam turbine power facility. Tyagi and Bansal [15] resolved a water treatment plant using the Runge-Kutta technique, which aims to enhance the plant's operational performance by putting in place efficient maintenance practices. Godara and Bansal [16] examined the neural network approach's potential to improve the computer network system's dependability and profitability.

Recently, models based on artificial neural networks (ANN) have gained popularity for estimating system reliability. In this study, we develop a mathematical model of the casting industry using the Artificial Neural Network (ANN) technique. It has broad applicability in many areas, including dependability. The ANN is composed of microscopic units known as neurons, each of which has a distinct weight (synaptic strength) and is linked by synapses. We create neuron equations employing the neural weight in the mathematical model, and we assess the plant's dependability and performance with the aid of these equations. The applicability of different types of neural networks for probabilistic analysis needs to be thoroughly investigated these days. When using a neural network approach for analysis, all failure kinds and repair rates are treated as weights, and these weights are determined by the exponential distribution.

2. SYSTEM DESCRIPTION

Continuous casting technology is a major advancement in the history of steelmaking and is influencing the worldwide steel industry. Globally, about 750 million tonnes of steel are produced annually, with continuous casting accounting for the majority of this solidification. The continuous casting method, seen in Figure 1, involves the flow of molten steel via a tundish and into the mould from a ladle. To protect each vessel from air exposure, a slag cover and ceramic nozzles should be positioned between each vessel. After entering the mould, the molten steel solidifies into a shell when it freezes up against the water-cooled copper walls. In order to maintain optimal operation in a steady state, the machine's driving rolls descend continuously, removing the shell from the mould at a rate known as the "casting speed" that corresponds to the flow of incoming metal. The remaining liquid is held in a container by the steel shell that is solidifying beneath the mould exit. Rolls are used to support steel to lessen ferrostatic pressure-induced bulging.

The strand's surface temperature is kept constant by air and water sprays until the molten core solidifies.

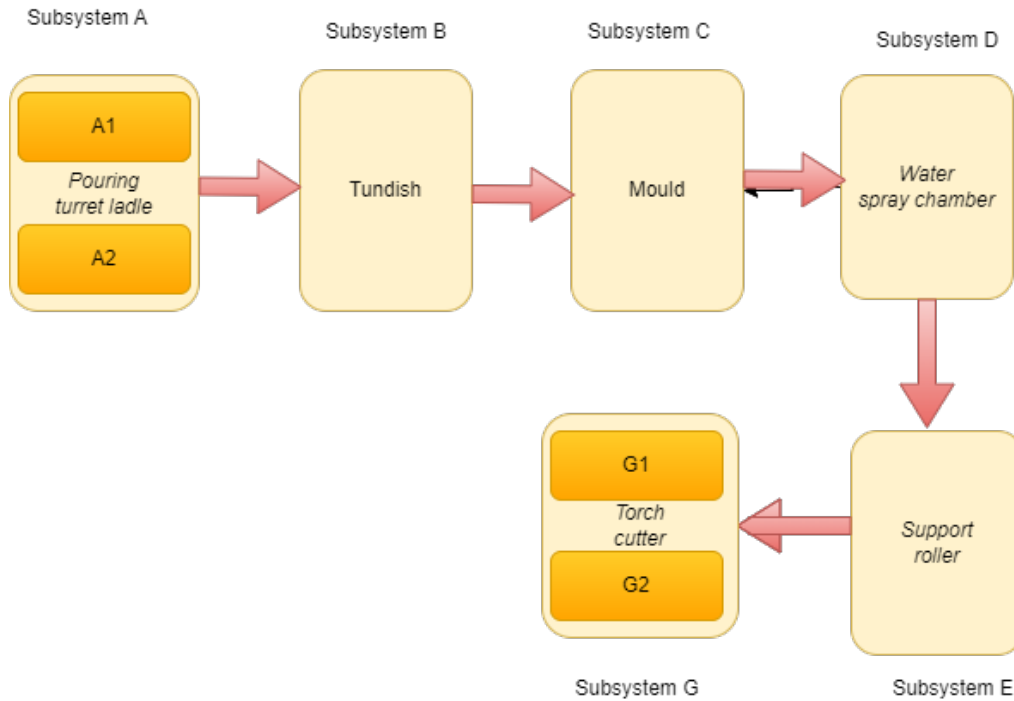


Figure 1: Block diagram

2.1. Assumptions and Notations

- First, we know the probabilities of the states.
- Every subsystem is initially in a fully operational state.
- Every subsystem has a repair facility in case it fails or deteriorates.
- There is a facility for repairs and the Repair rates are constant.
- Each subsystem's failure rate is both time-dependent and time-independent.

Moreover, the following notations are applied:

- f_{a1}/r_{a1} failure and repair rate of the 1st unit of turret ladle subsystem.
- f_{a2}/r_{a2} failure and repair rate of the standby unit of turret ladle subsystem.
- f_b/r_b failure and repair rate of the Tundish subsystem.
- f_c/r_c failure and repair rate of the Mould subsystem.
- f_d/r_d failure and repair rate of the Water spray chamber subsystem.
- f_e/r_e failure and repair rate of the Support roller subsystem.
- f_{g1}/r_{g1} failure and repair rate of the 1st unit of Torch cutter subsystem.
- f_{g2}/r_{g2} failure and repair rate of the standby unit of the Torch cutter subsystem.
- $w_{i,j}$ Neural weight of the i th state to the j th state.
- $P_i(t)$ Probability of the i th subsystem.
- A,B,C,D,E,G - Representation of the system is in a fully working state.
- a b c d e g Representation of the system is in the Failed state
- A^-, G^- Representation of the system is in the partially failed state.
- P_i is the probability that the system is in S_i state.

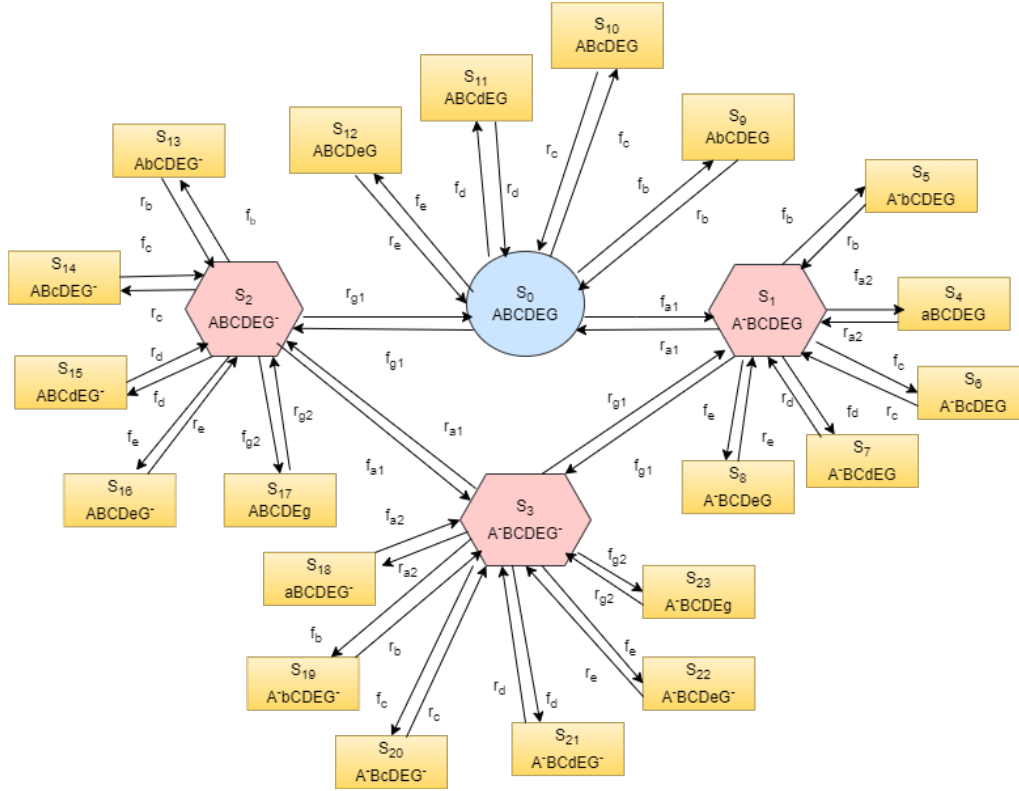


Figure 2: State transition diagram

2.2. Formulation of Artificial Neural Network Mathematical Model of the Continuous Casting System of the Steel Industry

The following set of difference differential equations governing the current mathematical model can be obtained using probability considerations and continuity arguments. There are three layers in the suggested ANN model: input, hidden, and output.

Input Layer. Input is defined by

$$P_i(t) = Z_i(t)$$

The following equations define the hidden layer:

$$P_0(t + \Delta t) = w_{0,0}Z_0 + w_{1,0}Z_1 + w_{2,0}Z_2 + w_{9,0}Z_9 + w_{10,0}Z_{10} + w_{11,0}Z_{11} + w_{12,0}Z_{12} \quad (1)$$

$$P_1(t + \Delta t) = w_{0,1}Z_0 + w_{1,1}Z_1 + w_{3,1}Z_3 + w_{4,1}Z_4 + w_{5,1}Z_5 + w_{6,1}Z_6 + w_{7,1}Z_7 + w_{8,1}Z_8 \quad (2)$$

$$P_2(t + \Delta t) = w_{0,2}Z_0 + w_{2,2}Z_2 + w_{3,2}Z_3 + w_{13,2}Z_{13} + w_{14,2}Z_{14} + w_{15,2}Z_{15} + w_{16,2}Z_{16} + w_{17,2}Z_{17} \quad (3)$$

$$P_3(t + \Delta t) = w_{1,3}Z_1 + w_{2,3}Z_2 + w_{3,3}Z_3 + w_{18,3}Z_{18} + w_{19,3}Z_{19} + w_{20,3}Z_{20} + w_{21,3}Z_{21} + w_{22,3}Z_{22} + w_{23,3}Z_{23} \quad (4)$$

$$P_4(t + \Delta t) = w_{1,4}Z_1 + w_{4,4}Z_4 \quad (5)$$

$$P_5(t + \Delta t) = w_{1,5}Z_1 + w_{5,5}Z_5 \quad (6)$$

$$P_6(t + \Delta t) = w_{1,6}Z_1 + w_{6,6}Z_6 \quad (7)$$

$$P_7(t + \Delta t) = w_{1,7}Z_1 + w_{7,7}Z_7 \quad (8)$$

$$P_8(t + \Delta t) = w_{1,8}Z_1 + w_{8,8}Z_8 \quad (9)$$

$$P_9(t + \Delta t) = w_{0,9}Z_0 + w_{9,9}Z_9 \quad (10)$$

$$P_{10}(t + \Delta t) = w_{0,10}Z_0 + w_{10,10}Z_{10} \quad (11)$$

$$P_{11}(t + \Delta t) = w_{0,11}Z_0 + w_{11,11}Z_{11} \quad (12)$$

$$P_{12}(t + \Delta t) = w_{0,12}Z_0 + w_{12,12}Z_{12} \quad (13)$$

$$P_{13}(t + \Delta t) = w_{2,13}Z_2 + w_{13,13}Z_{13} \quad (14)$$

$$P_{14}(t + \Delta t) = w_{2,14}Z_2 + w_{14,14}Z_{14} \quad (15)$$

$$P_{15}(t + \Delta t) = w_{2,15}Z_2 + w_{15,15}Z_{15} \quad (16)$$

$$P_{16}(t + \Delta t) = w_{2,16}Z_2 + w_{16,16}Z_{16} \quad (17)$$

$$P_{17}(t + \Delta t) = w_{2,17}Z_2 + w_{17,17}Z_{17} \quad (18)$$

$$P_{18}(t + \Delta t) = w_{2,18}Z_2 + w_{18,18}Z_{18} \quad (19)$$

$$P_{19}(t + \Delta t) = w_{3,19}Z_3 + w_{19,19}Z_{19} \quad (20)$$

$$P_{20}(t + \Delta t) = w_{3,20}Z_3 + w_{20,20}Z_{20} \quad (21)$$

$$P_{21}(t + \Delta t) = w_{3,21}Z_3 + w_{21,21}Z_{21} \quad (22)$$

$$P_{22}(t + \Delta t) = w_{3,22}Z_3 + w_{22,22}Z_{22} \quad (23)$$

$$P_{23}(t + \Delta t) = w_{3,23}Z_3 + w_{23,23}Z_{23} \quad (24)$$

Where

$$w_{0,0} = 1 - w_{0,1} - w_{0,2} - w_{0,9} - w_{0,10} - w_{0,11} - w_{0,12} \quad (25)$$

$$w_{1,1} = 1 - w_{1,0} - w_{1,5} - w_{1,3} - w_{1,4} - w_{1,6} - w_{1,7} - w_{1,8} \quad (26)$$

$$w_{2,2} = 1 - w_{2,0} - w_{2,3} - w_{2,13} - w_{2,14} - w_{2,15} - w_{2,16} - w_{2,17} \quad (27)$$

$$w_{3,3} = 1 - w_{3,2} - w_{3,1} - w_{3,18} - w_{3,19} - w_{3,18} - w_{3,19} - w_{3,20} - w_{3,21} - w_{3,22} - w_{3,23} \quad (28)$$

The output layer is defined by the following equation with the aid of the equation above.

$$Availability = P_0(t + \Delta t) + P_1(t + \Delta t) + P_2(t + \Delta t) + P_3(t + \Delta t) \quad (29)$$

3. RESULTS AND DISCUSSIONS

In this section, we explore the estimated time-varying state probabilities as revealed by our analysis applied to the steel industry's continuous casting system. The utilization of the Artificial Neural Network (ANN) methodology has yielded valuable insights regarding the optimization of profit and availability of the system. The system's initial state probabilities are given in Table 1, and it is assumed that the system has been operating continuously.

Table 1: Initial Probability of the System

Probability	P_0	P_1	P_2	P_3	$P_i, i= 4 \text{ to } 23$
Value	0.45	0.1	0.1	0.1	0.0125

Based on an analysis of Table 1, it can be observed that a good state has a higher probability than a failing or degraded state. Additionally, the initial probability of all failed states is 0.0125. When a system malfunctions, a repair facility is accessible because it is in three states: good, degraded, and failed.

Table 2 lists the repair rates for each subsystem.

Table 2: Constant Repair Rate of the System

Repair Rate	r_{a1}	r_{a2}	r_b	r_c	r_d	r_e	r_{g1}	r_{g2}
Value	0.2	0.2	0.1	0.1	0.1	0.1	0.2	0.2

The subsystem's failure rate is determined by the exponential and Weibull time distribution basis; there are two possible scenarios.

(1) Useful Life Phase: during this phase, the system's failure rate is constant and is distributed exponentially.

(2) Wear-out Phase: this phase is characterized by a Weibull time distribution that follows a time-dependent failure rate of the system.

In the Useful life Phase, the system's failure rate is constant and is distributed exponentially, and the wear-out phase is characterized by a Weibull time distribution that follows a time-dependent failure rate of the system. In the case of the wear-out phase, the failure rate of the system is calculated by equation.

$$f(t) = \beta t^{(\beta-1)} \eta^\beta \quad (30)$$

Where (β) and (η) represent the wear-out phase failure rate's shape and scale parameters, respectively. In order to match the starting failure rates of each subsystem and assume constant failure rates, the shape parameters (β) and scale parameters (η) of the time-dependent failure rates of each subsystem at time $t = 0$ have been taken as shown in Table 3.

Table 3: Constant and Time-Dependent Failure Rate

Failure rate	f_{a1}	f_{a2}	f_b	f_c	f_d	f_e	f_{g1}	f_{g2}
Shape Parameter(β)	2	2	2	2	2	2	2	2
Scale Parameter (η)	1.65	2.36	0.75	0.65	0.85	0.9	3	5
Time-dependent Failure rate	.0103	.050	.0498	0.066	0.033	0.034	0.031	0.011
Constant Failure rate	0.01	0.05	0.05	0.07	0.04	0.04	0.026	0.03

According to the ANN approach, the linear combination of the initial probability with the weight function and bias function is added to determine the state probability of any subsystem over time. To bring the estimated values of state probabilities closer to the measured state probabilities, bias is a constant that is added to the activation function in the described ANN model. Therefore, assuming $b_j = 0$ because measured state probabilities have not been taken into account in this numerical computation.

The ANN input layer of the casting system is determined by the initial probability of the system, which is given in Table 1. The hidden layer of the network is calculated with the help of equations 1 to 24, in which the neural weight of the system is calculated with the help of a linear combination of the failure and repair rate of the system, as shown in equations 25 to 28. The hidden layer of the ANN model shows the state probability of the system. The variation in the state probability of the system is calculated in both the cases namely useful life period phase and wear-out phase, these variations are calculated with a time increment of 10 hours. Here in this model, for the calculation purpose time, the factor is taken in month 10 hours= 0.014 month. The state probability fluctuates during the wear-out phase and the useful-life period, as Tables 4 to 9 demonstrate.

Table 4: Change in State Probabilities With Time In Useful-Life Phase)

T	P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
10	.4491	.0998	.0999	.0992	.0125	.0126	.0126	.0125	.0125
20	.4483	.0996	.0999	.0986	.0126	.0126	.0127	.0126	.0126
30	.4474	.0993	.0998	.0977	.0126	.0127	.0127	.0126	.0126
40	.4466	.0991	.0997	.0970	.0126	.0127	.0128	.0127	.0127
50	.4457	.0989	.0997	.0962	.0127	.0128	.0129	.0127	.0127
60	.4449	.0987	.0996	.0954	.0127	.0128	.0130	.0127	.0127
70	.4446	.0985	.0995	.0947	.0127	.0129	.0131	.0128	.0128
80	.4431	.0983	.0995	.0939	.0128	.0129	.0131	.0128	.0128
90	.4423	.0980	.0994	.0931	.0128	.0130	.0132	.0128	.0128
100	.4414	.0978	.0993	.0924	.0129	.0130	.0133	.0129	.0129

Table 5: Change in State Probabilities With Time In Useful-Life Phase

T	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}
10	.0126	.0129	.0127	.0127	.0126	.0126	.0125	.0125	.0125
20	.0126	.0133	.0130	.0130	.0126	.0127	.0126	.0126	.0125
30	.0127	.0138	.0132	.0132	.0127	.0127	.0126	.0126	.0125
40	.0127	.0142	.0134	.0134	.0127	.0128	.0127	.0127	.0125
50	.0128	.0146	.0137	.0137	.0128	.0129	.0127	.0127	.0125
60	.0128	.0150	.0139	.0139	.0128	.0130	.0127	.0127	.0125
70	.0129	.0155	.0141	.0141	.0129	.0131	.0128	.0128	.0125
80	.0129	.0159	.0144	.0144	.0129	.0131	.0128	.0128	.0126
90	.0130	.0163	.0146	.0146	.0130	.0132	.0128	.0128	.0126
100	.0130	.0167	.0148	.0148	.0130b	.0133	.0129	.0129	.0126

Table 6: Change in State Probabilities With Time In Useful-Life Phase

T	P_{18}	P_{19}	P_{20}	P_{21}	P_{22}	P_{23}
10	.0125	.0126	.0126	.0125	.0125	.0125
20	.0126	.0126	.0127	.0126	.0126	.0125
30	.0126	.0127	.0127	.0126	.0126	.0125
40	.0126	.0127	.0128	.0127	.0127	.0125
50	.0127	.0128	.0129	.0127	.0127	.0125
60	.0127	.0128	.0130	.0127	.0127	.0125
70	.0127	.0129	.0131	.0128	.0128	.0125
80	.0128	.0129	.0131	.0128	.0128	.0126
90	.0128	.0130	.0132	.0128	.0128	.0126
100	.0129	.0130	.0133	.0129	.0129	.0126

Table 7: Change in State Probabilities With Time In Wear-Out Phase

T	P_0	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
10	.4492	.0998	.0100	.0993	.0125	.0126	.0126	.0125	.0125
20	.4462	.0994	.0992	.0981	.0125	.0127	.0128	.0127	.0127
30	.4404	.0984	.0980	.0964	.0125	.0131	.0133	.0129	.0129
40	.4321	.0970	.0962	.0942	.0125	.0135	.0139	.0133	.0132
50	.4212	.0952	.0940	.0915	.0125	.0142	.0147	.0138	.0136
60	.4078	.0930	.0912	.0883	.0125	.0149	.0157	.0143	.0141
70	.3918	.0904	.0879	.0846	.0126	.0158	.0169	.0150	.0147
80	.3733	.0873	.0841	.0803	.0127	.0168	.0183	.0158	.0155
90	.3522	.0838	.0798	.0756	.0128	.0180	.0199	.0167	.0163
100	.3285	.0799	.0749	.0703	.0129	.0193	.0216	.0178	.0172

Table 8: Change in State Probabilities With Time In Wear-Out Phase

T	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	P_{17}
10	.0126	.0129	.0127	.0127	.0126	.0126	.0125	.0125	.0125
20	.0127	.0141	.0134	.0133	.0127	.0128	.0127	.0127	.0124
30	.0131	.0162	.0146	.0144	.0131	.0133	.0129	.0129	.0124
40	.0135	.0191	.0163	.0159	.0135	.0139	.0133	.0132	.0124
50	.0142	.0229	.0185	.0178	.0142	.0147	.0138	.0136	.0124
60	.0149	.0274	.0212	.0202	.0149	.0157	.0143	.0141	.0123
70	.0158	.0328	.0243	.0230	.0158	.0169	.0150	.0147	.0123
80	.0168	.0391	.0280	.0263	.0168	.0183	.0158	.0155	.0123
90	.0180	.0462	.0321	.0300	.0180	.0199	.0167	.0163	.0123
100	.0193	.0541	.0367	.0341	.0193	.0216	.0178	.0172	.0123

Table 9: Change in State Probabilities With Time In Wear-Out Phase

T	P_{18}	P_{19}	P_{20}	P_{21}	P_{22}	P_{23}
10	.0125	.0126	.0126	.0126	.0125	.0125
20	.0125	.0127	.0128	.0127	.0127	.0124
30	.0125	.0131	.0133	.0129	.0129	.0124
40	.0125	.0135	.0139	.0133	.0132	.0124
50	.0125	.0142	.0147	.0138	.0136	.0124
60	.0125	.0149	.0157	.0143	.0141	.0123
70	.0126	.0158	.0169	.0150	.0147	.0123
80	.0127	.0168	.0183	.0158	.0155	.0123
90	.0128	.0180	.0199	.0167	.0163	.0123
100	.0129	.0193	.0216	.0178	.0172	.0123

Tables 4 to 9 analysis show that variations in the wear-out phase decrease by 4.5% to 3.2%, while variations in the useful life phase decrease by 4.5% to 4.4% in the good state probability. The variation in the good state probability occurred because of an increase in the failure rate with respect to time during the wear-out phase. The similarity of state probability fluctuates Because of the two facts: (1) The values of their individual weights dominate this summation, making the cumulative effect of the other weights and states insignificant, and (2) The state probability values at time $t=0$ are presumptively similar, P_4 versus P_{18} and P_{17} versus P_{23} are in both the useful-life and wear-out phases.

Here, the upstate (availability) of the system is indicated in the current ANN model output layer. With the aid of equation (29), the system’s time-related availability is computed, and the results are displayed in Table (10) and Figure (3).

Table 10: *Availability of the system*

Time	Availability (Wear-out)	Availability (useful-life)
10	0.7488	0.7481
20	0.7428	0.7462
30	0.7332	0.7443
40	0.7196	0.7424
50	0.7020	0.7405
60	0.6803	0.7386
70	0.6547	0.7367
80	0.6250	0.7348
90	0.5913	0.7329
100	0.5536	0.7310

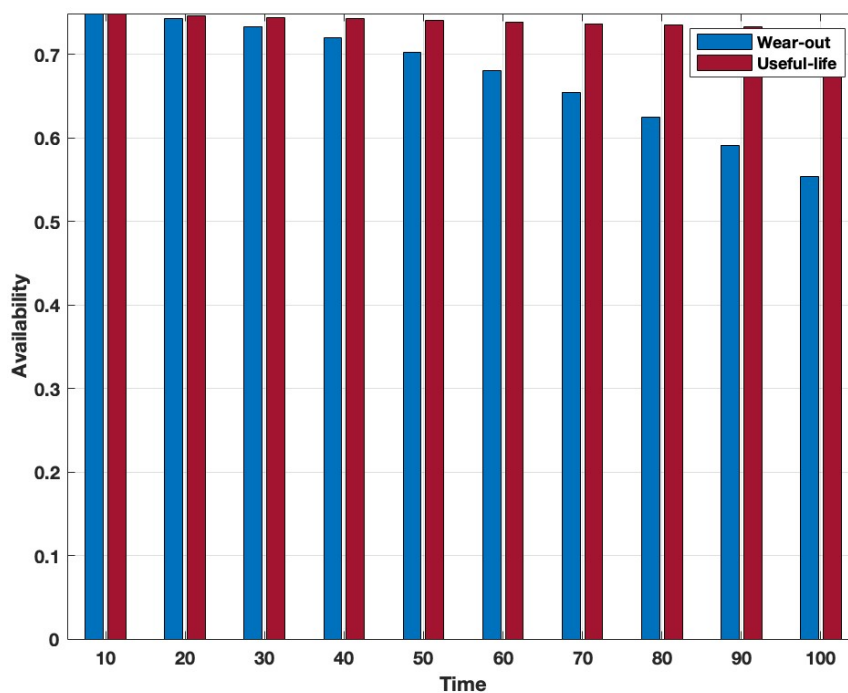


Figure 3: *Availability with Time*

Table 10 analysis shows that the system’s availability decreases over time in both the wear-out and useful life phases. The system availability drops from 7.4% to 5.5% during the wear-out phase. On the other hand, availability slightly decreases from 7.4% to 7.3% during the useful-life phase. When compared to the useful-life phase, the availability decrement varies more during the wear-out phase. This variation is caused by the wear-out phase failure rate increasing over time, whereas it remains constant during the useful-life phase.

With the aid of revenue costs and preventive maintenance costs, which fix revenue costs and variable preventive costs, the profit of the continuous casting system is computed using the given

equation.

$$Profit = R_e A(t) - C_1 - C_2 \tag{31}$$

Where R_e , C_1 and C_2 is the revenue cost per unit time, the total cost per unit, and the preventive maintenance cost of time, here we take $C_1 = 2400$, $R_e = 19000$, $C_2 = 56$

Table 11: Profit of the System

Time	Profit (Wear-out)	Profit (useful-life)
10	11761	11758
20	23371	23499
30	34537	35168
40	45033	46765
50	54632	58290
60	63100	69742
70	70215	81121
80	75745	92429
90	79461	10366
100	81126	11483

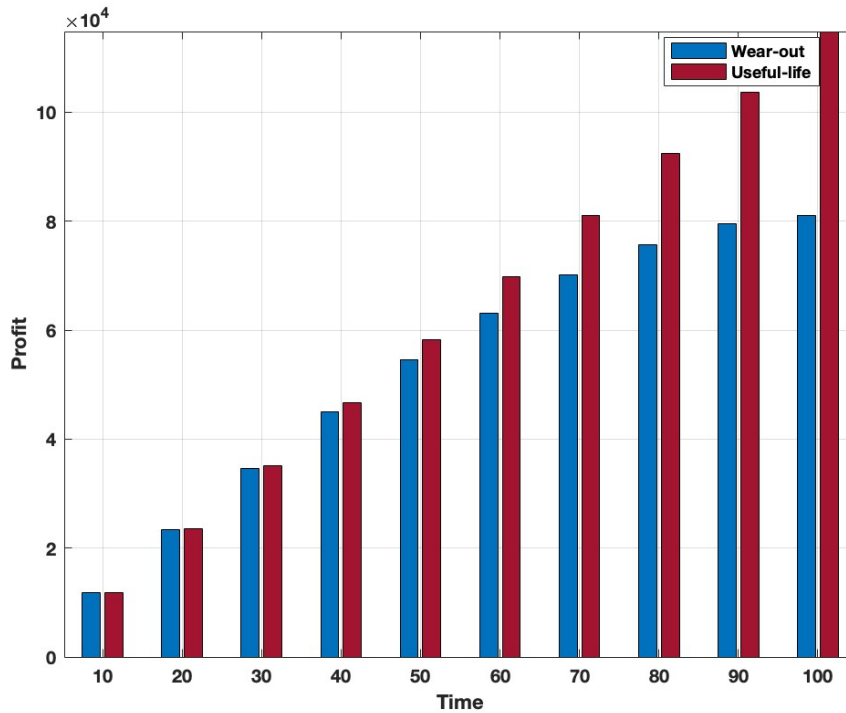


Figure 4: Profit with Time

Based on the data presented in Table 11, it can be inferred that the steel industry’s casting system’s profit increases over time as it wears out and the useful-life phase. The system’s wear-out phase turns out to be more profitable than its useful-life phase for the main reason that the wear-out phase’s availability exceeds the useful-life phases. Wear-out phase is a strategic focal point for optimizing financial returns because of its direct correlation with increased profitability due to its increased availability.

4. CONCLUSION

The implementation of Artificial Neural Network (ANN) techniques in the metals industry's continuous casting system has been shown to be beneficial in terms of enhancing productivity parameters. This study uses an artificial neural network (ANN) approach to optimize casting system availability in the steel industry during both the useful life and wear-out phases. Consequently, an ANN model has been used to compare the changes in multi-state probabilities with respect to time for the system's (i) wear-out phase and (ii) useful life phase. Therefore, it can be concluded that while the probability of the good state P_0 decreases dramatically from 44.92% to 32.85% during the wear-out phase, it decreases smoothly from 44.91% to 44.14% during the useful-life segment.

The probability behavior of the degraded states (P_1 , P_2 , and P_3) is also analyzed. It is found that the wear-out phase has a greater impact on the variation in the degraded state's state probability than the useful-life phase. In the wear-out phase, the probability of degraded states decreases by approximately 9.98% to 7%, whereas in the useful-life phase, it decreases by approximately 9.98% to 9.24%. In both the useful-life and wear-out phases, the state probability of the failed state increases with time.

The availability of the casting system is examined in both the useful life and wear-out phases. It is found that, in comparison to the useful-life phase, the wear-out phase majorly impacts the system's availability. The system's availability drops from 74.9% to 55.4% during the wear-out phase and from 74.8% to 73.1% during the useful life.

The profit margin of the system has also been examined using revenue and preventive maintenance costs, which adjust for revenue and other expenses. It is concluded that, in comparison to a useful-life case, the system yields higher profit in terms of time during the wear-out phase.

REFERENCES

- [1] Tzong, R. Y., & Lee, S. L. (1992). Solidification of arbitrarily shaped casting in mold casting system. *International Journal of Heat and Mass Transfer*, 35(11), 2795-2803.
- [2] Agarwal, S. C., & Bansal, S. (2009). Reliability analysis of a standby redundant complex system with changing environment under the head of the line repair discipline. *Bulletin of Pure and Applied Sciences*, 28(1), 165-173.
- [3] Agarwal, S. C., Sahani, M., & Bansal, S. (2010). Reliability characteristic of a cold-standby redundant system *International Journal of Research and Review in Applied Sciences (IJRRAS)*, 3(2), 193-199.
- [4] Bansal, S., Agarwal, S.C., Sahani, M., Sharma K. (2018). Evaluation of reliability factors using B.F technique in milk powder manufacturing plant. *International Journal of Research and Review in Applied Science*, vol.4, no.4, pp. 416-424.
- [5] Amit Kumara, Aman Kumar Varshneyb and Mangey Rama (2015). Sensitivity analysis for casting process under stochastic modelling, *International Journal of Industrial Engineering Computations* 6, 419 -432.
- [6] Agarwal, S. C., & Bansal, S. (2015) Cost analysis of solar thermal electric power plant. *International Journal of Advanced Technology in Engineering and Science*, 3(10), 12-22.
- [7] Kumar, A., Ram, M. and Rawat, R.S. (2017). Optimization of casting process through reliability approach, *International Journal of Quality & Reliability Management*, Vol. 34 No. 6, pp. 833-848. <https://doi.org/10.1108/IJQRM-07-2016-0103>
- [8] Bansal, S., (2018). Availability Analysis of a Repairable Redundant System under Preemptive-Resume Repair Discipline. *International Journal of Mathematics And its Applications*, 6(1"D), 665-671.
- [9] Bansal, Shikha, and Sohan Tyagi(2018). Reliability analysis of screw manufacturing plant using orthogonal matrix method. *Pertanika Journal of Science & Technology* 26(4),1789-1800.

- [10] Bansal S., Tyagi, S. (2021). Mathematical modeling and availability analysis of leaf spring manufacturing plant. *Pertanika Journal of Science & Technology*, Vol.29, No.2, pp. 1041-1051.
- [11] Bansal, S., Tyagi, S., Verma, V.K (2022). Performance Modeling and Availability Analysis of Screw Manufacturing Plant. *Materials Today: Proceedings*, Vol.57, No.5, pp.1985-1988.
- [12] Amit Chaudhari, Hari Vasudevan (2022). Reliability-based design optimization of casting process parameters using Markov chain model, *Materials Today: Proceedings*, Volume 63, Pages 602-606.
- [13] Chaudhary, P., and Bansal, S.,(2023). Assessment of the Reliability Performance of Hydro-Electric Power Station. 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, pp. 148-153.
- [14] Godara, U., and Bansal, S., (2023). Performance of Reliability Factors in Steam Turbine Generator Power Plant Using Boolean Function Technique and Neural Network Approach. International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), Bangalore, India, pp. 507-512.
- [15] Tyagi, S.L., and Bansal, S., (2023). Optimization Model for Wastewater Treatment Process 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 165-168.
- [16] Godara, U., and Bansal, S., (2023). Prediction of Reliability Factor For Multi-State Computer System With Neural Network Approach. IEEE 2nd International Conference on Industrial Electronics: Developments & Applications (ICIDeA), Imphal, India, pp. 132-135.

TOPP-LEONE EXPONENTIATED GOMPERTZ INVERSE RAYLEIGH DISTRIBUTION: PROPERTIES AND APPLICATIONS

Sule Omeiza Bashiru¹, Alaa Abdulrahman Khalaf² and Alhaji Modu Isa³

¹Department of Mathematics and Statistics, Confluence University of Science and Technology,
Osara, Kogi State, Nigeria.

²Diyala Education Directorate, Diyala, Iraq.

³Department of Mathematics and Computer Science, Borno State University, Nigeria.

Email: ¹bash0140@gmail.com; ²alaa.a.khalaf35510@st.tu.edu.iq; ³alhajimoduisa@bosu.edu.ng

Abstract

This paper focused on deriving a new lifetime distribution having five parameters by compounding the Gompertz inverse Rayleigh model and the Topp-Leone exponentiated-G family of distributions. The new model is called Topp-Leone exponentiated Gompertz inverse Rayleigh (TLEGoIRa) distribution. The new model is very flexible and the shape of its pdf can be positively or negatively skewed and symmetric. Some statistical characteristics of the new model, such as the moments, incomplete moments, quantile function, rényi entropy and order statistics are derived and investigated. The pdf of the minimum and maximum order statistics of the new model were derived and studied. The model's parameters are estimated using the maximum likelihood approach. A simulation study was conducted to investigate the consistency of the newly proposed model, using the average bias and root mean square error (RMSE) as metrics. The outcome of the simulation suggested that as sample sizes increase, both the average bias and root mean square error (RMSE) decrease, indicating that the distribution is consistent. Finally, two real-life datasets were used to explore the new model's importance and adaptability in comparison to other competing models. The results of the application revealed that the new distribution outperforms its competitors.

Keywords: Topp-Leone Exponentiated G., Gompertz Inverse Rayleigh, Quantile Function, Order Statistics, MLE.

I. Introduction

In the realm of distribution theory, the pursuit of developing models that accurately reflect the prevailing trends across various disciplines has proven challenging. Classical distributions, which form the foundation, often lack the required flexibility and robustness. This inherent limitation has spurred researchers within the distribution theory field to undertake the task of extending or generalizing existing distributions. The overarching goal is to imbue these distributions with greater flexibility and resilience, enabling them to effectively capture the evolving patterns present in datasets originating from diverse fields like engineering, environmental sciences, biological sciences, medical sciences, and beyond.

This process of extension or generalization entails the introduction of one or more additional parameters to the existing distributions. Contemporary approaches to distribution generalization frequently involve the utilization of distribution families. Examples of these families include the Topp-Leone exponentiated-G distribution by [1], Topp-Leone Kumaraswamy-G distribution by [2], Topp-Leone-G distribution by [3], type II half logistic-G distribution by [4], type I half logistic exponentiated-G distribution by [5], type II half logistic exponentiated-G distribution by [6], transmuted exponentiated generalized G distribution by [7], Topp-Leone odd Lindley G distribution by [8], and Topp-Leone Gompertz-G distribution by [9], among others. These families are constructed by introducing supplementary shape parameter(s) to the foundational distribution, thereby augmenting the efficacy and practicality of data modeling.

One significant continuous probability distribution, referred to as the inverse Rayleigh (IRa) distribution, was initially introduced by [10] and has since found extensive application in modeling system failure times. Notably, the IRa distribution is a specialized form of the broader inverse Weibull (IW) distribution. The statistical literature offers various adaptations and extensions of the IRa distribution, which can be explored further through references such as [11-13]. An extension of particular interest is the Gompertz inverse Rayleigh (GoIRa) distribution, developed by [14], which is considered as the baseline distribution in this study. Through the extension of the GoIRa distribution, we aim to develop a more adaptable compound distribution.

Reference [1] introduced the TLE-G, a distinct family of continuous distributions with the cumulative distribution function (cdf) and probability density function (pdf) given as:

$$F(x, \theta, \alpha) = \left(1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right)^\theta \quad (1)$$

$$f(x, \theta, \alpha) = 2\theta\alpha g(x; \xi) G(x; \xi)^{\alpha-1} \left(1 - G(x; \xi)^\alpha \right) \left(1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right)^{\theta-1} \quad (2)$$

The cdf and pdf corresponding to the baseline Gompertz inverse Rayleigh (GoIRa) distribution are given as:

$$G(x, \gamma, \beta, \lambda) = 1 - e^{-\frac{\gamma}{\beta} \left(1 - \left(1 - e^{-\left(\frac{\lambda}{x}\right)^2} \right)^{-\beta} \right)} \quad (3)$$

and

$$g(x, \gamma, \beta, \lambda) = 2\gamma\lambda^2 x^{-3} e^{-\left(\frac{\lambda}{x}\right)^2} \left(1 - e^{-\left(\frac{\lambda}{x}\right)^2} \right)^{-\beta-1} e^{-\frac{\gamma}{\beta} \left(1 - \left(1 - e^{-\left(\frac{\lambda}{x}\right)^2} \right)^{-\beta} \right)} \quad (4)$$

Where $\lambda > 0$ is the scale parameter and $\gamma, \beta > 0$ are the shape parameters respectively.

The primary objective of this research is to utilize the GoIRa distribution as a foundational model within the TLE-G framework, aiming to develop a novel extension known as the TLEGoIRa distribution. This extension seeks to enhance the flexibility and applicability of the GoIRa distribution in capturing complex data patterns across various fields.

The remaining content of this article is organized as follows: Section 2 presents the development of the TLEGoIRa distribution, including the derivation of its properties and the method for estimating its parameters. Section 3 discusses simulation studies conducted to investigate the consistency of the estimates and the application of the model to two real datasets to demonstrate the practical potential of the new distribution. Finally, Section 4 provides concluding remarks.

II. Methods

2.1 Development of Topp-Leone Exponentiated Gompertz Inverse Rayleigh (TLEGoIRa) Distribution

To derive the cdf of the new model, equation (3) is inserted into equation (1) as:

$$F\left(x, \theta, \alpha, \gamma, \beta, \lambda\right) = \left(1 - \left(1 - \left(1 - e^{-\left(\frac{\gamma}{\beta}\right)\left(1 - e^{-\left(\frac{\lambda}{x}\right)^2}\right)^{-\beta}}\right)^{\alpha}\right)^2\right)^{\theta} \quad (5)$$

To derive the PDF of the new model, equations (3) and (4) are inserted into equation (2) as follows:

$$f\left(x, \theta, \alpha, \gamma, \beta, \lambda\right) = 4\gamma\theta\alpha\lambda^2 x^{-3} e^{-\left(\frac{\lambda}{x}\right)^2} \left(1 - e^{-\left(\frac{\lambda}{x}\right)^2}\right)^{-\beta-1} e^{\frac{\gamma}{\beta}\left(1 - e^{-\left(\frac{\lambda}{x}\right)^2}\right)^{-\beta}} \left(1 - e^{-\left(\frac{\gamma}{\beta}\left(1 - e^{-\left(\frac{\lambda}{x}\right)^2}\right)^{-\beta}\right)}\right)^{\alpha-1} \left(1 - \left(1 - \left(1 - e^{-\left(\frac{\gamma}{\beta}\left(1 - e^{-\left(\frac{\lambda}{x}\right)^2}\right)^{-\beta}\right)}\right)^{\alpha}\right)^2\right)^{\theta-1} \quad (6)$$

where $x \geq 0$, and $\theta, \alpha, \lambda, \beta, \gamma > 0$.

The hazard function for the TLEGoIRa distribution can be obtained using this expression:

$$h\left(x, \theta, \alpha, \gamma, \beta, \lambda\right) = \frac{f\left(x, \theta, \alpha, \gamma, \beta, \lambda\right)}{1 - F\left(x, \theta, \alpha, \gamma, \beta, \lambda\right)}$$

The pdf and hazard function plots of the TLEGoIRa distribution are given figures 1 and 2 below:

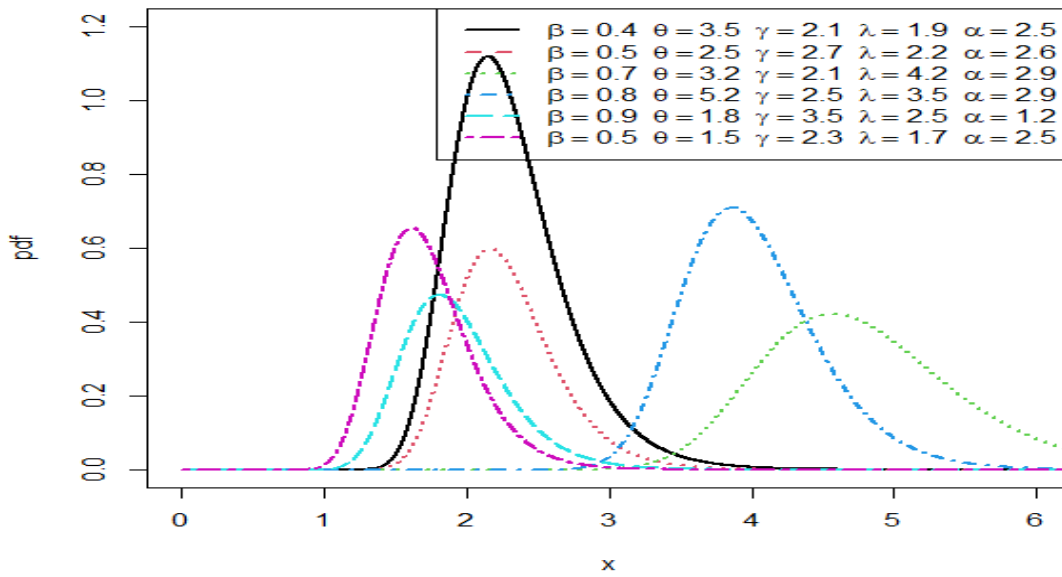


Figure 1: pdf plots of the TLEGoIRa distribution with different parameter values

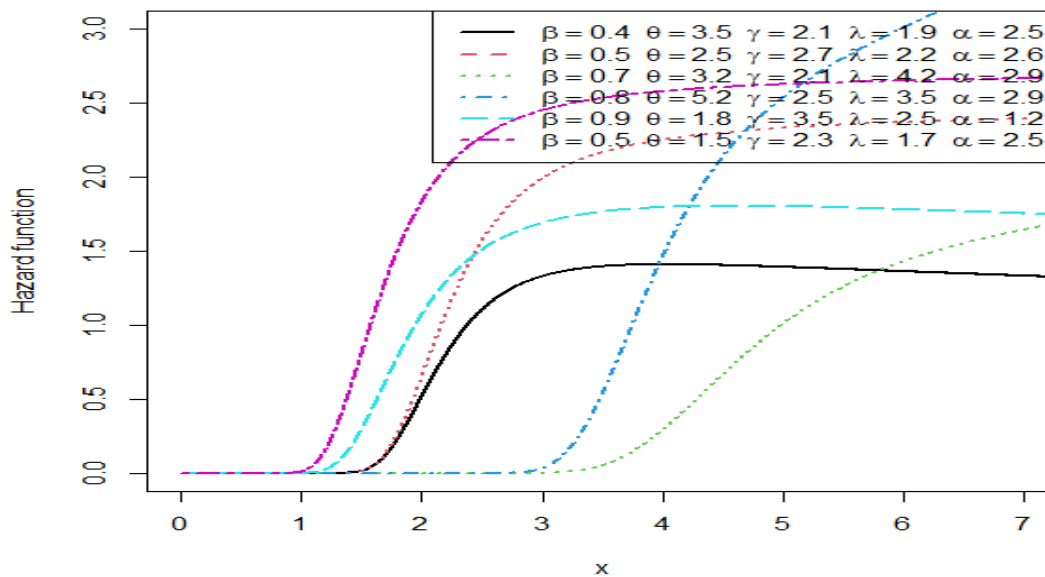


Figure 2: hazard function plots of the TLEGoIRa distribution with different parameter values

2.2 Statistical Properties of the TLEGoIRa Distribution

This section derives some statistical properties of the TLEGoIRa distribution including moments, survival function, hazard function, quantile functions, and order statistics.

2.2.1 Moment

The r^{th} moment of the TLEGoIRa model is computed using the following expression:

$$\dot{\mu}_r = E(X^r) = \int_0^\infty x^r f \left(x, \theta, \alpha, \lambda, \beta, \gamma \right) dx$$

First expanding equation (6) by using generalized binomial expansion $[1+u]^{-q} = \sum_{z=0}^{\infty} (-1)^z \binom{q}{z} u^z$,

$$[1+u]^{-q} = \sum_{z=0}^{\infty} (-1)^z \binom{q}{z} u^z, [1-u]^{-q} = \sum_{w=0}^{\infty} \frac{\Gamma(q+w)}{w! \Gamma(q)} u^w : |u| < 1, q > 0. [15-16]$$

$$\left(1 - \left(1 - \left(1 - e^{-\frac{\gamma}{\beta} \left(1 - \left(\frac{\lambda}{x} \right)^2 \right)^{-\beta}} \right) \right)^\alpha \right)^{2i-1} = \sum_{i=0}^{\infty} (-1)^i \binom{\theta-1}{i} \left(1 - \left(1 - e^{-\frac{\gamma}{\beta} \left(1 - \left(\frac{\lambda}{x} \right)^2 \right)^{-\beta}} \right) \right)^{2i}$$

And

$$\left(1 - \left(1 - e^{-\frac{\gamma}{\beta} \left(1 - \left(\frac{\lambda}{x} \right)^2 \right)^{-\beta}} \right) \right)^{2i+1} = \sum_{j=0}^{\infty} (-1)^j \binom{2i+1}{j} \left(1 - e^{-\frac{\gamma}{\beta} \left(1 - \left(\frac{\lambda}{x} \right)^2 \right)^{-\beta}} \right)^{\alpha j}$$

Again, using the generalized binomial expansion and the exponential expansion formula

$$e^{-c} = \sum_{c=0}^{\infty} \frac{(-1)^c}{c!} u^c [17]$$

Then we get expansion pdf of TLEGoIRa distribution:

$$f \left(x, \theta, \alpha, \gamma, \beta, \lambda \right) = \psi_p x^{-3} e^{-\left(\frac{\lambda}{x} \right)^2} \tag{7}$$

Where

$$\psi_p = 4\theta\alpha\gamma\lambda^2 \sum_{i=j=k=m=w=q=0}^{\infty} \frac{(-1)^{i+j+k+w+q}}{m!} \binom{\theta}{i} \binom{2i+1}{j} \binom{\alpha(j+1)-1}{k} \binom{m}{w} \binom{-\beta(1+w)-1}{q}$$

Hence,

$$\dot{\mu}_r = E(X^r) = \psi_p \int_0^\infty x^{r-3} e^{-\left(\frac{\lambda}{x} \right)^2} dx$$

On solving the integral part in the equation above, then the $\dot{\mu}_r$ is:

$$\dot{\mu}_r = E\left(X^r\right) = \psi_p \frac{\lambda^2 \Gamma\left(1 - \frac{r}{2}\right)}{\left(q+1\right)^{1-\frac{r}{2}}} \quad (8)$$

2.2.2 Incomplete Moments

Equation (7) yields the incomplete moments for the TLEGoIRa distribution with r^{th} ($r > 0$). [18-19]

$$\dot{\mu}_r(u) = \psi_p \int_0^u x^{r-3} e^{-\left(q+1\right)\left(\frac{\lambda}{x}\right)^2} dx$$

$$\text{Let } t = \left(q+1\right)\left(\frac{\lambda}{x}\right)^2 \Rightarrow x = \left(\frac{\left(q+1\right)\lambda^2}{t}\right)^{\frac{1}{2}}$$

When $x=0 \Rightarrow t=0$, and if $x=u \Rightarrow t = \left(q+1\right)\left(\frac{\lambda}{u}\right)^2$, then

$$\dot{\mu}_r(u) = \frac{\psi}{2\left(\left(q+1\right)\lambda^2\right)^{1-\frac{r}{2}}} \gamma\left(1 - \frac{r}{2}, \left(q+1\right)\lambda^2\right) \quad (9)$$

2.2.3 Quantile Function

The quantile function of TLEGoIRa distribution is given as

$$x = Q(u) = \lambda \left\{ -\log \left[1 - \left[1 - \left[\frac{\beta}{\gamma} \log \left(1 - \left(1 - u^{\frac{1}{\theta}} \right)^{\frac{1}{2}} \right)^{\frac{1}{\alpha}} \right] \right]^{\frac{1}{\beta}} \right] \right\}^{\frac{1}{2}} \quad (10)$$

The median of TLEGoIRa distribution is obtained by setting $u = 0.5$ in equation (10)

$$x_{\text{median}} = Q(0.5) = \lambda \left\{ -\log \left[1 - \left[1 - \left[\frac{\beta}{\gamma} \log \left(1 - \left(1 - (0.5)^{\frac{1}{\theta}} \right)^{\frac{1}{2}} \right)^{\frac{1}{\alpha}} \right] \right]^{\frac{1}{\beta}} \right] \right\}^{\frac{1}{2}}$$

Table 1: Quantiles for given parameter values of the TLEGoIRa distribution.

U	$(\theta, \lambda, \alpha, \beta, \gamma)$		
	(0.4, 0.8, 1.2, 1.3, 2.2)	(3, 1.3, 0.9, 0.7, 3.3)	(2.3, 3, 1.7, 1.4, 1.2)
0.1	0.45276	0.59975	1.33661
0.2	0.52743	0.63546	1.41855
0.3	0.59314	0.66110	1.48066
0.4	0.65832	0.68289	1.53586
0.5	0.72712	0.70315	1.58933
0.6	0.80314	0.72329	1.64475
0.7	0.89111	0.74472	1.70628
0.8	0.99973	0.76965	1.78139
0.9	1.15324	0.80387	1.89139

2.2.4 Rényi Entropy

Define the Rényi entropy of the TLEGoIRa distribution using the following formula. [20]

$$T_R(\tau) = \frac{1}{1-\tau} \log \int_0^\infty f^\tau(x) dx, \quad \tau > 0, \tau \neq 1$$

By substitution equation (7) into the equation above:

$$T_R(\tau) = \frac{1}{1-\tau} \log \left(\psi_p \int_0^\infty x^{-3\tau} e^{-\frac{(q+1)\tau}{x}} dx \right)$$

The last integral, we get

$$T_R(\tau) = \frac{1}{1-\tau} \log \left(\frac{\psi_p \Gamma\left(\frac{3}{2}(\tau-1)+1\right)}{2 \left(\left(q+1 \right) \tau \lambda^2 \right)^{\frac{1}{2}(3\tau+1)}} \right) \quad (11)$$

2.2.5 Order Statistic

The pdf of the order statistics for the TLEGoIRa distribution is obtained as follows: [21-24]

$$g_{t:n}(x) = \frac{n!}{(t-1)!(n-t)!} [F(X)]^{t-1} [1-F(X)]^{n-t} f(x) \quad (12)$$

Substituting equations (5) and (6) into equation (12), we have:

$$\begin{aligned} \frac{\partial(I)}{\partial\alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \log \left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right) - \sum_{i=1}^n \frac{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha \log \left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)}{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha} \\ &+ 2(\theta-1) \sum_{i=1}^n \frac{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha \left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha \log \left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)}{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha} \right)^2 \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{\partial(I)}{\partial\gamma} &= \frac{n}{\gamma} + \frac{1}{\beta} \sum_{i=1}^n \log \left(1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right) - (\alpha-1) \sum_{i=1}^n \frac{e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \frac{1}{\beta} \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta}}{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha} + \alpha \sum_{i=1}^n \frac{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^{\alpha-1} e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \frac{1}{\beta} \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta}}{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha} \\ &+ 2\alpha(\theta-1) \sum_{i=1}^n \frac{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha \left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \frac{1}{\beta} \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta}}{\left(1 - e^{-\left(\frac{\gamma}{\beta} \left[1 - \left(1 - e^{-\left(\frac{\lambda}{x_i} \right)^2} \right)^{-\beta} \right] \right)} \right)^\alpha} \right)^2 \end{aligned} \quad (17)$$

$$\begin{aligned}
 \frac{\partial(J)}{\partial\lambda} &= \frac{2n}{\lambda} - 2\lambda \sum_{i=1}^n \left(\frac{1}{x_i}\right)^2 - 2\lambda(\beta+1) \sum_{i=1}^n \frac{e^{-\left(\frac{\lambda}{x_i}\right)^2}}{x_i^2 \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)} - 2\lambda\gamma \sum_{i=1}^n \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-(\beta+1)} e^{-\left(\frac{\lambda}{x_i}\right)^2} \left(\frac{1}{x_i}\right)^2 \\
 &+ \frac{2\lambda\beta(\alpha-1) \sum_{i=1}^n e^{-\left(\frac{\lambda}{x_i}\right)^2} \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2}}{1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}} + \frac{2\lambda\beta\alpha \sum_{i=1}^n \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2}}{x^2 \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{\alpha}} \\
 &+ \frac{4\lambda\beta\alpha(\theta-1) \sum_{i=1}^n \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{\alpha} \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2} e^{-\left(\frac{\lambda}{x_i}\right)^2}}{1 - \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{\alpha}} \tag{18}
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial(J)}{\partial\beta} &= -\sum_{i=1}^n \log \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right) + \frac{\gamma}{\beta^2} \sum_{i=1}^n \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} \log \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right) - (\alpha-1) \sum_{i=1}^n \frac{e^{-\left(\frac{\lambda}{x_i}\right)^2} \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} e^{-\left(\frac{\lambda}{x_i}\right)^2} \log \left(e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)}{1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}} \\
 &+ \alpha \sum_{i=1}^n \frac{\left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} e^{-\left(\frac{\lambda}{x_i}\right)^2} \log \left(e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)}{1 - \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{\alpha}} \\
 &- 2\alpha(\theta-1) \sum_{i=1}^n \frac{\left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{\alpha} \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{-\beta} e^{-\left(\frac{\lambda}{x_i}\right)^2} \log \left(e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)}{1 - \left(1 - e^{-\left(\frac{\lambda}{x_i}\right)^2}\right)^{\alpha}} \tag{19}
 \end{aligned}$$

Since equations (15), (16), (17), (18), and (19) are non-linear in parameters, an iterative technique is resorted to using Newton-Raphson iterative algorithm to obtain the estimate of the parameters.

III. Results

3.1 Simulation

This section describes the conclusions of a simulation research of the TLEGoIRa distribution. The study investigates five distinct sets of parameter values: $(\theta = 0.9, \alpha = 1.2, \gamma = 0.3, \beta = 0.3, \lambda = 0.5)$, $(\theta = 1.2, \alpha = 0.9, \gamma = 0.7, \beta = 0.3, \lambda = 0.5)$, $(\theta = 1.3, \alpha = 1.3, \gamma = 0.8, \beta = 0.6, \lambda = 0.4)$, and $(\theta = 2, \alpha = 2, \gamma = 0.6, \beta = 0.75, \lambda = 0.2)$. Each parameter set yields 1000 samples, with $n = 50, 100, 150,$ and 300 . We utilize these samples to compute the mean, average bias, and root mean square error (RMSE). To calculate bias and RMSE for the calculated parameters, use the formulas below:

$$Abias(\hat{\alpha}) = \frac{\sum_{i=1}^N \hat{\alpha}_i}{N} - \alpha, \text{ and } RMSE(\hat{\alpha}) = \sqrt{\frac{\sum_{i=1}^N (\hat{\alpha}_i - \alpha)^2}{N}}.$$

Tables 2 and **3** illustrate the results, which demonstrate a clear pattern: as the sample size increases, the mean parameter estimates get more precise and closer to the true values. Simultaneously, the corresponding RMSEs and Abias approach zero, proving MLEs' reliability and consistency.

Table 2: Results from Monte Carlo simulations of the TLEGoIRa distribution

$\theta = 0.9, \alpha = 1.2, \gamma = 0.3, \beta = 0.3, \lambda = 0.5$					$\theta = 1.2, \alpha = 0.9, \gamma = 0.7, \beta = 0.3, \lambda = 0.5$		
Parameter	N	Mean	RMSE	Abias	Mean	RMSE	Abias
θ	50	3.3963	4.9674	2.4963	3.7293	3.6574	2.5293
	100	2.2823	2.9373	1.3823	3.7169	3.3882	2.5169
	150	1.9920	2.1471	1.0920	3.1464	2.3203	1.9464
	300	1.5072	1.4519	0.6072	2.2718	2.2284	1.0718
α	50	1.1618	1.3257	0.0981	1.1671	1.6299	0.2671
	100	1.2951	1.1776	0.0951	1.1055	1.3425	0.2055
	150	1.1993	1.1490	0.0644	1.1023	1.3182	0.2060
	300	1.3244	1.0653	0.0455	1.0609	1.0041	0.1609
γ	50	0.2144	0.1879	0.0855	0.5295	0.6465	0.1704
	100	0.2459	0.1461	0.0540	0.5168	0.4134	0.1531
	150	0.2632	0.1382	0.0367	0.5023	0.3071	0.1276
	300	0.2599	0.1155	0.0200	0.4994	0.2535	0.1005
β	50	0.3311	0.1891	0.0396	0.3246	0.2079	0.0246
	100	0.3274	0.1864	0.0374	0.3136	0.1747	0.0136
	150	0.3233	0.1581	0.0240	0.2956	0.1504	0.0043
	300	0.3209	0.1324	0.0233	0.3020	0.1368	0.0020
λ	50	0.9800	0.9588	0.4800	1.9086	1.6432	1.4086
	100	0.8343	0.6752	0.3343	1.2474	1.4809	0.7474
	150	0.7453	0.5092	0.2453	0.9105	1.0670	0.4105
	300	0.6237	0.3144	0.1237	0.7456	0.5470	0.2456

Table 3: Results from Monte Carlo simulations of the TLEGoIRa distribution

$\theta = 1.3, \alpha = 1.3, \gamma = 0.8, \beta = 0.6, \lambda = 0.4$					$\theta = 2, \alpha = 2, \gamma = 0.6, \beta = 0.75, \lambda = 0.2$		
Parameter	N	Mean	RMSE	Abias	Mean	RMSE	Abias
θ	50	3.9020	3.2589	2.6020	4.6867	4.6171	2.6867
	100	3.1574	2.2916	1.8574	3.9532	3.5984	1.9532
	150	2.7310	2.6574	1.4310	3.4859	2.7741	1.4859
	300	2.9938	1.1012	0.6938	3.0303	2.0658	1.0303
α	50	1.5031	2.1812	0.2031	2.2077	3.9702	0.2077
	100	1.4930	2.1254	0.3930	2.6847	3.9583	0.0847
	150	1.4756	2.0739	0.2756	2.0827	3.9374	0.0827
	300	1.3273	1.3373	0.2273	2.4920	3.0661	0.0720
γ	50	0.7726	1.0620	0.9709	0.4879	0.3450	0.1920
	100	0.7423	0.5469	0.1073	0.4565	0.2704	0.1434
	150	0.7259	0.4181	0.0976	0.4367	0.2280	0.1132
	300	0.7016	0.3897	0.0594	0.5271	0.1969	0.0728
β	50	0.7850	0.5824	0.1850	1.3345	1.2143	0.5845
	100	0.7450	0.5133	0.1450	1.2569	1.1121	0.5069
	150	0.7342	0.4793	0.1342	1.3577	1.1337	0.4877
	300	0.6869	0.4226	0.0869	1.1989	0.9787	0.4489
λ	50	2.0712	2.7215	1.6719	1.0376	1.9009	0.8376
	100	1.3579	2.1418	0.9579	0.6940	1.0649	0.4940
	150	0.9976	1.3302	0.5976	0.6385	0.9570	0.4385
	300	0.7083	0.7030	0.3083	0.4223	0.4678	0.2223

3.2 Applications

In this section, the practical use of the TLEGoIRa distribution is explored via two real-life data sets. Table 4 displays the cdf of the models, which will be compared to the TLEGoIRa distribution.

Table 4: CDF for the Comparative distributions

Distribution	CDF
Truncated Exponentiated Exponential Gompertz inverse Rayleigh (TEGoIRa) [15]	$\frac{1 - \exp\left(-\theta \left(1 - e^{-\left(\frac{\gamma}{\beta} \left(1 - e^{-\left(\frac{\lambda}{x}\right)^{-\beta}}\right)}\right)\right)\right)}{\left(1 - \exp(-\theta)\right)^\alpha}$
Beta Gompertz inverse Rayleigh (BeGoIRa) (New)	$pbeta\left(1 - e^{-\left(\frac{\gamma}{\beta} \left(1 - e^{-\left(\frac{\lambda}{x}\right)^{-\beta}}\right)}\right), \theta, \alpha\right)$

Kumaraswamy Gompertz inverse Rayleigh (KuGoIRa) (New)	$1 - \left(1 - \left(1 - e^{-\left(\frac{\gamma}{\beta} \left(1 - \left(1 - e^{-\left(\frac{\lambda}{x} \right)^{-\beta}} \right) \right)^{\theta}} \right)^{\alpha} \right)$
Exponential Generalized Gompertz inverse Rayleigh (EGGoIRa) (New)	$\left(1 - \left(1 - \left(1 - e^{-\left(\frac{\gamma}{\beta} \left(1 - \left(1 - e^{-\left(\frac{\lambda}{x} \right)^{-\beta}} \right) \right)^{\theta}} \right)^{\alpha} \right)$
Weibull Gompertz inverse Rayleigh (WeGoIRa) (New)	$1 - \exp \left(-\alpha^{-\theta} \left(-\log \left(1 - \left(1 - e^{-\left(\frac{\gamma}{\beta} \left(1 - \left(1 - e^{-\left(\frac{\lambda}{x} \right)^{-\beta}} \right) \right)^{\theta}} \right)^{\alpha} \right) \right) \right)$

The first dataset (I), shows the tensile strength in GPa of 69 carbon fibers evaluated at 20mm gauge lengths. It was utilized by Bader and Priest [25]

(1.312, 1.314, 1.479 ,1.552,1.700 ,1.803, 1.861 ,1.865 ,1.944, 1.958 ,1.966, 1.997 ,2.006, 2.021 ,2.027, 2.055, 2.063 ,2.098, 2.140, 2.179 ,2.224 ,2.240, 2.253 ,2.270, 2.272, 2.274, 2.301, 2.301 ,2.359 ,2.382 ,2.382 ,2.426 ,2.434, 2.435, 2.478 ,2.490, 2.511, 2.514, 2.535 ,2.554, 2.566, 2.570, 2.586, 2.629 ,2.633, 2.642, 2.648, 2.684 ,2.697, 2.726, 2.770, 2.773, 2.800, 2.809, 2.818, 2.821, 2.848 ,2.880, 2.954, 3.012, 3.067 ,3.084, 3.090, 3.096, 3.128, 3.233, 3.433 ,3.585, 3.585).

The second dataset (II), shown here represents COVID-19 mortality rate data for Mexico over a 108-day period from March 4th to July 20, 2020. It was utilized by Alongy et al. [26]

(8.826, 6.105 ,10.383, 7.267 ,13.220, 6.015 ,10.855, 6.122 ,10.685, 10.035, 5.242 ,7.630 ,14.604 ,7.903 ,6.327 ,9.391 ,14.962 ,4.730 , 3.215 ,16.498, 11.665 ,9.284, 12.878, 6.656,3.440 ,5.854, 8.813 , 10.043, 7.260, 5.985, 4.424 ,4.344 ,5.143 ,9.935 ,7.840 ,9.550 , 6.968 ,6.370 ,3.537 ,3.286 ,10.158, 8.108 ,6.697 ,7.151 ,6.560 , 2.988 ,3.336 ,6.814 ,8.325 ,7.854 ,8.551 ,3.228, 3.499 ,3.751, 7.486 ,6.625 ,6.140 ,4.909 ,4.661 ,1.867 ,2.838 ,5.392, 12.042, 8.696 ,6.412 ,3.395 ,1.815 ,3.327 ,5.406 ,6.182 ,4.949 ,4.089 , 3.359 ,2.070, 3.298 ,5.317 ,5.442 ,4.557 ,4.292 ,2.500 ,6.535 , 4.648 ,4.697 ,5.459 ,4.120, 3.922 ,3.219, 1.402 ,2.438, 3.257 , 3.632, 3.233 ,3.027, 2.352 ,1.205 ,2.077, 3.778, 3.218, 2.926, 2.601, 2.065, 1.041, 1.800, 3.029, 2.058, 2.326, 2.506, 1.923).

Tables 5, and 6 for data (I), and (II) show that the TLEGoIRa distribution beats Comparative distributions in several key criteria, including Akaike information criterion (AIC) , Consistent AIC (CAIC),Bayesian information criterion (BIC), Hannan-Quinn information (HQIC),Kolmogorov-Smimov (KS) ststistic, Anderson-Darling (A), and Cramer-von Mises (W) values. The lower values of these measures for the TLEGoIRa distribution are preferable for comparative distributions.

Figures 3, and **5** show the Fitted densities for Data I, and II, respectively, and **Figures 4**, and **5** show the empirical cdf plots for Data I and II. These visualizations enable us to evaluate the goodness of fit and see how well the model fits the data.

Table 5: Goodness-of-Fit Statistics for Data I

Dist.	MLEs	-2L	AIC	CAIC	BIC	HQIC	W	A	K-S	p-value
TLEGoIRa	$\hat{\theta}$:1.4173 $\hat{\alpha}$:1.4214 $\hat{\lambda}$:0.1896 $\hat{\beta}$:1.6041 $\hat{\gamma}$:2.0224	48.81	107.62	108.58	118.79	112.05	0.0178	0.1600	0.0414	0.9997
TEEGoIRa	$\hat{\theta}$:0.0443 $\hat{\alpha}$:0.8522 $\hat{\lambda}$:3.3049 $\hat{\beta}$:3.5479 $\hat{\gamma}$:4.1427	49.36	108.74	109.69	119.91	113.17	0.0356	0.2930	0.0530	0.9900
BeGoIRa	$\hat{\theta}$:1.7430 $\hat{\alpha}$:1.0087 $\hat{\lambda}$:0.0292 $\hat{\beta}$:0.7674 $\hat{\gamma}$:1.8205	50.50	111.03	111.99	122.21	115.47	0.0165	0.1487	0.0982	0.5177
KuGoIRa	$\hat{\theta}$:1.7872 $\hat{\alpha}$:1.0734 $\hat{\lambda}$:0.0238 $\hat{\beta}$:0.7657 $\hat{\gamma}$:1.8912	49.63	109.29	110.24	120.46	113.72	0.0164	0.1481	0.0750	0.8313
EGGoIRa	$\hat{\theta}$:0.9743 $\hat{\alpha}$:1.7609 $\hat{\lambda}$:0.0223 $\hat{\beta}$:0.7207 $\hat{\gamma}$:1.8454	50.12	110.28	111.23	121.45	114.71	0.0165	0.1483	0.0826	0.7330
WeGoIRa	$\hat{\theta}$:1.8695 $\hat{\alpha}$:1.0161 $\hat{\lambda}$:0.0322 $\hat{\beta}$:0.7167 $\hat{\gamma}$:1.4519	49.64	109.29	110.24	120.46	113.72	0.0308	0.2515	0.0598	0.9659

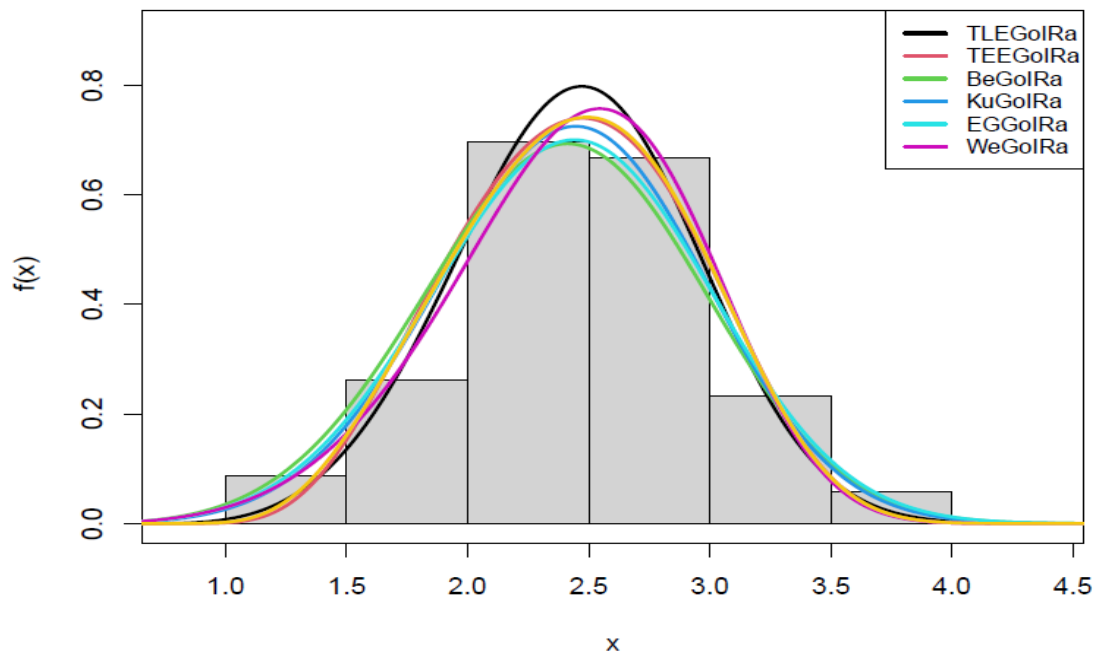


Figure 3: Fitted densities for Data I

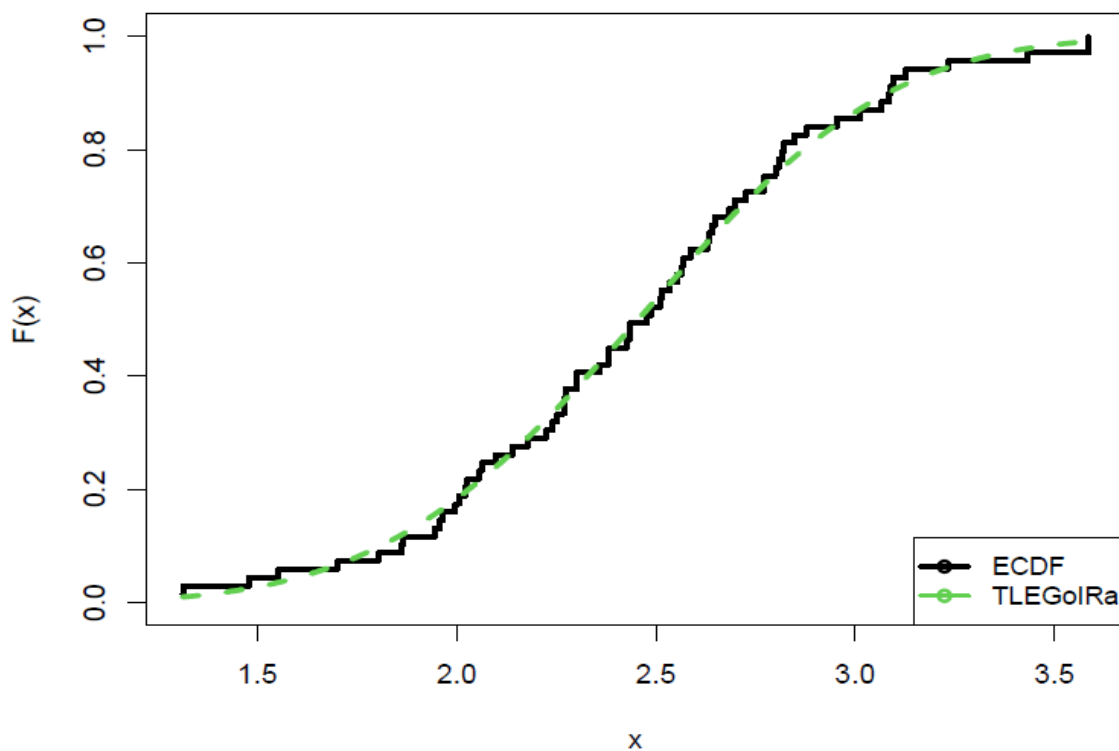


Figure 4: Empirical cdf plot for Data I

Table 6: Goodness-of-Fit Statistics for Data II

Dist.	MLEs	-2L	AIC	CAIC	BIC	HQIC	W	A	K-S	p-value
TLEGoIRa	$\hat{\theta}$:0.6769 $\hat{\alpha}$:1.4676 $\hat{\lambda}$:0.1046 $\hat{\beta}$:2.1082 $\hat{\gamma}$:0.7180	265.12	540.24	540.83	553.65	545.68	0.0357	0.2127	0.0565	0.8805
TEEGoIRa	$\hat{\theta}$:0.9677 $\hat{\alpha}$:1.1598 $\hat{\lambda}$:0.0450 $\hat{\beta}$:1.2991 $\hat{\gamma}$:0.8986	265.62	541.25	541.84	554.66	546.69	0.0517	0.3001	0.0646	0.7577
BeGoIRa	$\hat{\theta}$:1.2413 $\hat{\alpha}$:0.9903 $\hat{\lambda}$:0.0079 $\hat{\beta}$:0.3841 $\hat{\gamma}$:0.8618	267.77	545.54	546.13	558.95	550.98	0.0909	0.5702	0.0833	0.4415
KuGoIRa	$\hat{\theta}$:1.1977 $\hat{\alpha}$:0.9940 $\hat{\lambda}$:0.0108 $\hat{\beta}$:0.3686 $\hat{\gamma}$:0.7788	268.42	546.85	547.44	560.26	552.29	0.0802	0.4945	0.0683	0.6936
EGGoIRa	$\hat{\theta}$:0.9884 $\hat{\alpha}$:1.2186 $\hat{\lambda}$:0.0095 $\hat{\beta}$:0.3768 $\hat{\gamma}$:0.8146	267.83	545.66	546.25	559.07	551.10	0.0846	0.5260	0.0605	0.8236
WeGoIRa	$\hat{\theta}$:1.4253 $\hat{\alpha}$:1.0311 $\hat{\lambda}$:0.1373 $\hat{\beta}$:1.3555 $\hat{\gamma}$:0.51167	265.26	540.56	541.15	553.97	546.00	0.0433	0.2484	0.0607	0.8208

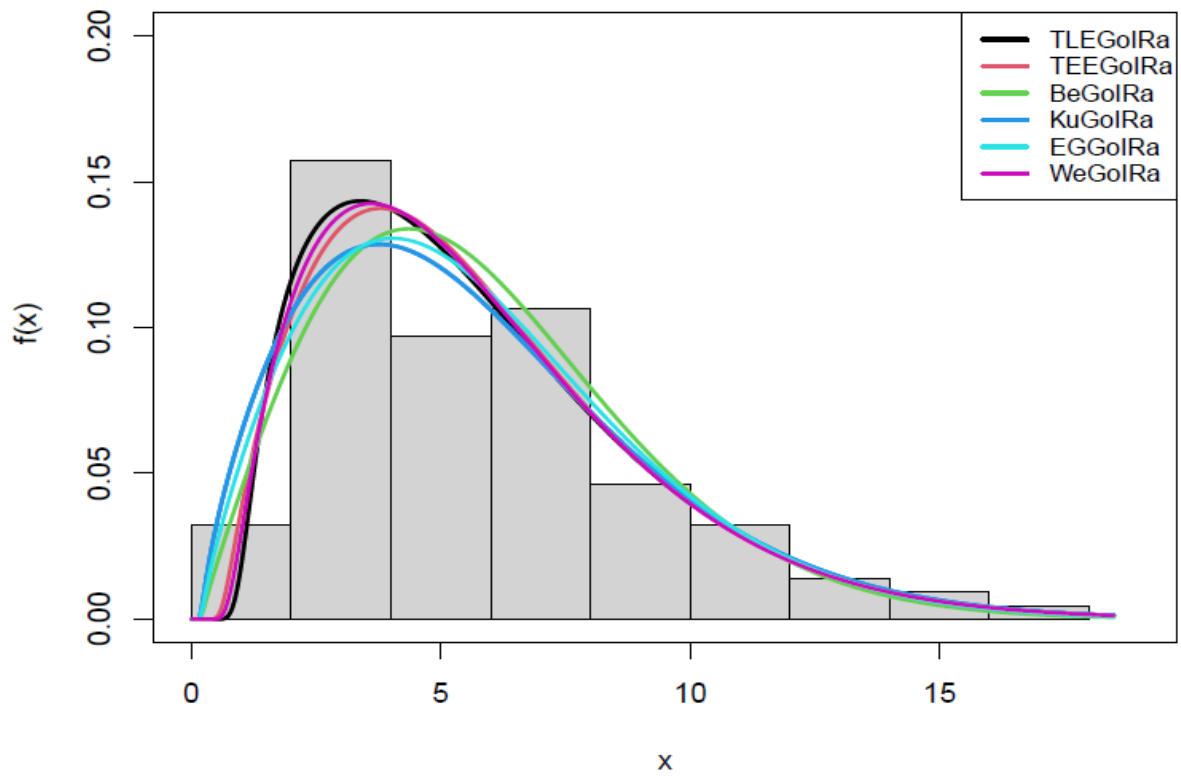


Figure 5: Fitted densities for Data II

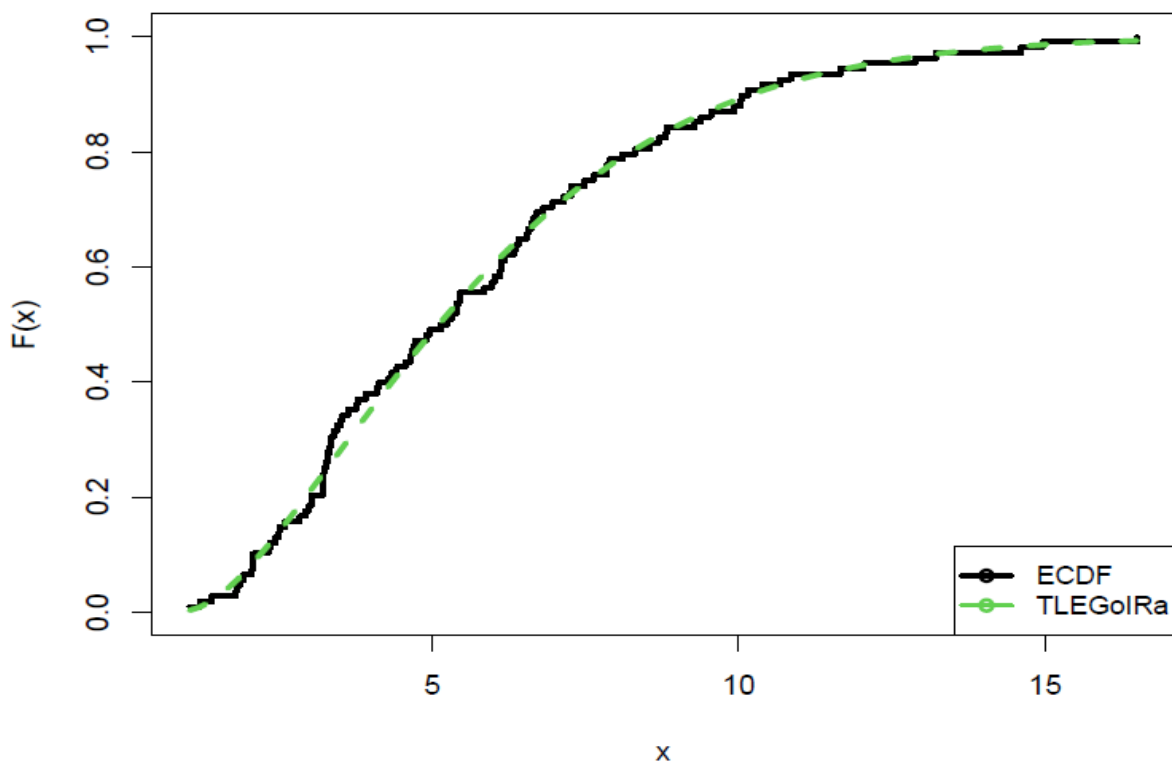


Figure 6: Empirical cdf plot for Data II

IV. Discussion

In this paper, we introduce a novel extension of the GoIRa model, referred to as the TLEGoIRa distribution. We provide explicit expressions for various statistical properties, including ordinary moments, incomplete moments, quantile function, rényi entropy and order statistics of the TLEGoIRa distribution. To estimate the unknown parameters, we employ the method of maximum likelihood estimation and undertake a simulation research to investigate the average bias and root mean square error (RMSE) as sample sizes rise. The results show a consistent model performance, with diminishing Abias and Rmse as the sample size grows. Furthermore, we validate the effectiveness of the proposed model through two real-life applications. Through these applications, we demonstrate that the proposed model outperforms several other competitive models in terms of goodness of fit. This empirical evidence underscores the enhanced flexibility and robustness of our novel model in accurately representing and modeling the characteristics of the given datasets when compared to the other competitive models under consideration.

References

- [1] Sule I., Sani I.D., Audu I., and Jibril H.M. (2020). On the Topp Leone exponentiated-G family of distributions: Properties and applications. *Asian Journal of Probability and Statistics*, 7, 1-15.
- [2] Sule I., Doguwa S.I., Isah A., and Jibril H.M. (2020). The Topp Leone Kumaraswamy-G Family of Distributions with Applications to Cancer Disease Data. *Journal of Biostatistics and Epidemiology*, 6, 37-48.
- [3] Al-Shomrani A., Arif O., Shawky A., Hanif S., and Shahbaz M.Q. (2016). Topp-Leone Family of Distributions: Some Properties and Application. *Pakistan Journal of Statistics and Operation Research*, 12(3), 443-451.
- [4] Soliman A.H., Elgarhy M.A.E., and Shakil M. (2017). Type II half logistic family of distributions with applications. *Pakistan Journal of Statistics and Operation Research*, 13, 245-264.
- [5] Bello O.A., Doguwa S.I., Yahaya A., and Jibril H.M. (2021). A Type I Half Logistic Exponentiated-G Family of Distributions: Properties and Application. *Communication in Physical Sciences*, 7(3), 147-163.
- [6] Bello O.A., Doguwa S.I., Yahaya A., and Jibril H.M. (2021). A Type II Half Logistic Exponentiated-G Family Of Distributions with Applications to Survival Analysis. *FUDMA Journal of Sciences*, 5(3), 177-190.
- [7] Yousof H.M., Afify A.Z., Alizadeh M., Butt N.S., Hamedani G., and Ali M.M. (2015). The transmuted exponentiated generalized-G family of distributions. *Pakistan Journal of Statistics and Operation Research*, 11(4), 441-464.
- [8] Reyad H., Alizadeh M., Jamal F., and Othman S. (2018). The Topp Leone odd Lindley-G family of distributions: Properties and applications. *Journal of Statistics and Management Systems*, 21(7), 1273-1297.
- [9] Oluyede B., Chamunorwa S., Chipepa F., and Alizadeh M. (2022). The Topp-Leone Gompertz-G family of distributions with applications. *Journal of Statistics and Management Systems*, 25(6), 1399-1423.
- [10] Voda, V. G. (1972). On the inverse Rayleigh distributed random variable, *Rep. Statist. App. Res.*, JUSE 19: 13-21.
- [11] Sule O. B., Adegoke T. M and Kayafat T.U (2021). Bayes Estimators of Exponentiated Inverse Rayleigh Distribution using Lindleys Approximation. *Asian Research Journal of Mathematics*, 17(2): 60 – 71.

- [12] Malik A., and Ahmad S. (2018). A New inverse Rayleigh distribution: Properties and application. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 5(5), 92-96.
- [13] Sule O.B., and Halid O.Y. (2023). On the properties and applications of Toppleone Gompertz Inverse Rayleigh Distribution. *Reliability: Theory and Applications*, 18(4), 1032-1045.
- [14] Halid O.Y., and Sule O.B. (2022). A classical and Bayesian techniques for Gompertz Inverse Rayleigh Distribution. *Pakistan Journal of Statistics*, 38(1), 49-76.
- [15] Khalaf A.A., Yusur K., and Khaleel M.A. (2023). [0, 1] Truncated Exponentiated Exponential Inverse Weibull Distribution with Applications of Carbon Fiber and COVID-19 Data. *Journal of Al-Rafidain University College For Sciences*, 1, 387-399.
- [16] Khalaf A.A., and Khaleel M.A. (2022). [0, 1] Truncated exponentiated exponential Gompertz distribution: Properties and applications. *AIP Conference Proceedings*, 2394(1), 1-12.
- [17] Khalaf A.A., and Khaleel M.A. (2020). Truncated Exponential Marshall-Olkin-Gompertz Distribution Properties and Applications. *Tikrit Journal of Administration and Economics Sciences*, 16, 483-497.
- [18] Noori N.A., Khalaf A.A., and Khaleel M.A. (2023). A New Generalized Family of Odd Lomax-G Distributions: Properties and Applications. *Advances in the Theory of Nonlinear Analysis and its Applications*, 7(4), 01-16.
- [19] Ibrahim M.Q., Khalaf A.A., Noori N.A., and Khaleel M.A. (2023). Exploring the Properties, Simulation, and Applications of the Odd Burr XII Gompertz Distribution. *Advances in the Theory of Nonlinear Analysis and its Applications*, 7(4), 60-75.
- [20] Khalaf A.A., Ibrahim M.Q., Noori N.A., and Khaleel M.A. (2024). [0, 1] Truncated Exponentiated Exponential Burr Type X Distribution with Applications. *Iraq Journal of Science*, 65(8), 15.
- [21] Sule O. B. (2021). A New Extended Generalized Inverse Exponential Distribution: Properties and Applications. *Asian Journal of Probability and Statistics*. 11(2): 30 – 46.
- [22] Adegoke T.M., Oladoja O.M., Sule O.B., Mustapha A.A., Aderupatan D.E and Nzei L.C. (2023). Topp-Leone Inverse Gompertz Distribution: Properties and different estimations techniques and Applications. *Pakistan Journal of Statistics*. 39(4): 433 – 456.
- [23] Sule O. B. and Ibrahim I.I (2023). Modeling of Reliability and Survival data with Exponentiated Generalized Inverse Lomax Distribution. *Reliability: theory & applications*. 18(4): 493 – 501.
- [24] Sule O. B. (2023). A study on the properties of a New Exponentiated Extended Inverse Exponential Distribution with applications. *Reliability: theory & applications*. 18(3): 59 – 72.
- [25] Bader M.G., and Priest A.M. (1982). Statistical aspects of fibre and bundle strength in hybrid composites. *Progress in Science and Engineering of Composites*, 1129-1136.
- [26] Almongy H.M., Almetwally E.M., Aljohani H.M., Alghamdi A.S., and Hafez E.H. (2021). A new extended Rayleigh distribution with applications of COVID-19 data. *Results in Physics*, 23, 1-9.

A PROBABILITY MODEL FOR SURVIVAL ANALYSIS OF CANCER PATIENTS

Mousumi Ray* , Rama Shanker

•
Department of Statistics, Assam University, Silchar, India
mousumiray616@gmail.com , shankerrama2009@gmail.com

*Corresponding Author

Abstract

It has been observed by statistician that to find a suitable model for the survival analysis of cancer patients is really challenging. The main reasons for that is the highly positively skewed nature of datasets. During recent decades several statistician tried to propose one parameter, two-parameter, three-parameter, four-parameter and five-parameter probability models but due to either theoretical or applied point of view the goodness of fit provided by these distributions are not very satisfactory. In this paper a compound probability model called gamma-Sujatha distribution, which is a compound of gamma and Sujatha distribution, has been proposed for the modeling of survival times of cancer patients. Many important properties of the suggested distribution including its shape, moments (negative), hazard function, reversed hazard function, quantile function have been discussed. Method of maximum likelihood has been used to estimate its parameters. A simulation study has been conducted to know the consistency of maximum likelihood estimators. Two real datasets, one relating to acute bone cancer and the other relating to head and neck cancer, has been considered to examine the applicability, suitability and flexibility of the proposed distribution. The goodness of fit of the proposed distribution shows quite satisfactory fit over other considered distributions.

Keywords: Survival analysis, compounding, hazard function, reversed hazard rate function, stress-strength parameter, maximum likelihood estimation, applications.

I. Introduction

Several statistical distributions have been extensively used for the modeling and analysis of survival times (time to event) data, also known as reliability data in biomedical sciences. On comparative studies on gamma and Weibull [1] distribution done by Shanker et al [2] shows that on some datasets relating to head and neck cancer these two classical two-parameter lifetime distributions does not provide good fit and on some datasets they perform diversely. During recent decades researchers were trying to modify Weibull distribution which would provide better fit to survival times of cancer patients. We know that the Weibull distribution is the most popular distribution for modeling survival data that properly explain the mortality and failure. Several authors have extended the Weibull distribution by adding one or more additional shape parameters to bring more flexibility in the shape of the distribution to accommodate the nature of

the data. For example, exponentiated generalized Weibull (EGW) distribution by Cordeiro et al [3], Beta-Weibull (BW) distribution by Famoye et al [4], Kumaraswamy Weibull (Kum-W) distribution by Cordeiro et al [5], exponentiated Kumaraswamy Weibull (EKumW) distribution by Eissa [6], Alpha power Weibull (APW) distribution by Nassar et al [7], are some among others. Although, these two, three and four parameters extended Weibull distribution provide good fit to survival times of cancer patients, but are not quite satisfactory because, in general, cancer data are highly positively skewed.

During recent decades several researchers have been trying to derive a suitable lifetime distribution to model data which are highly positively skewed, especially survival times of cancer patients. The search for highly positively skewed continuous distribution (mean is much less than the variance) has been studied by several researchers using compounding technique as the compounding always provides a highly positively skewed distributions. For instance, gamma distribution is a positively skewed distribution and its compounding with other positively skewed distribution provides highly positively skewed distribution. A compound gamma distribution arises when a random variable say X , follows gamma distribution with a shape parameter φ and scale parameter λ and the parameter λ itself behaves as a random variable with some distribution which is known as mixing distribution. There are four important one parameter positively skewed lifetime distributions namely, exponential distribution, Lindley distribution by Lindley [8], Shanker distribution by Shanker [9] and Sujatha distribution by Shanker [10] for modeling and analysis of survival time of cancer patients and out of these four distributions, Sujatha distribution provides much better fit as compared to the other distributions. The gamma-Lindley distribution (G-LD) proposed by Abdi et al [11] which is a compound of gamma distribution with Lindley distribution of Lindley [8] is highly positively skewed distribution. The gamma – Shanker distribution (G-SD) introduced by Ray and Shanker [12], which is a compound of gamma distribution with Shanker distribution of Shanker [9] is also highly positively skewed distribution. Further exponential-Shanker distribution (E-SD) suggested by Ray and Shanker [13] which is the compound of exponential distribution with Shanker distribution is also positively skewed distribution. The G-LD and the G-SD for $x > 0, \varphi > 0, \omega > 0$ are defined by its probability density function (pdf) and cumulative density function (cdf) as follows

$$f_{G-LD}(x; \varphi, \omega) = \frac{\varphi \omega^2 (1 + \varphi + \omega + x) x^{\varphi-1}}{(\omega + 1)(\omega + x)^{\varphi+2}} \tag{1}$$

$$F_{G-LD}(x; \varphi, \omega) = \frac{x^\varphi [(\omega + 1)x + (1 + \varphi + \omega)\omega]}{(\omega + 1)(\omega + x)^{\varphi+1}} \tag{2}$$

$$f_{G-SD}(x; \varphi, \omega) = \frac{\varphi \omega^2 (1 + \varphi + \omega x + \omega^2) x^{\varphi-1}}{(1 + \omega^2)(\omega + x)^{2+\varphi}} \tag{3}$$

$$F_{G-SD}(x; \varphi, \omega) = \frac{x^\varphi [x(1 + \omega^2) + (1 + \varphi + \omega^2)\omega]}{(1 + \omega^2)(\omega + x)^{1+\varphi}} \tag{4}$$

Sujatha distribution is defined by its pdf and cdf

$$f_{SUD}(x; \omega) = \frac{\omega^3 (1 + x + x^2) e^{-\omega x}}{(\omega^2 + \omega + 2)} \tag{5}$$

$$F_{SUD}(x; \omega) = 1 - \left[1 + \frac{\omega x (\omega x + \omega + 2)}{\omega^2 + \omega + 2} \right] e^{-\omega x} \tag{6}$$

The motivations for considering the gamma-Sujatha distribution (G-SUD), the compound of gamma and Sujatha distribution are as follows:

(i). Suppose X is the lifetime of component following gamma distribution with shape parameter φ and scale parameter λ . If the sample is drawn from the population having variability in the scale parameter λ , then the variability can be well explained by assuming the distribution of λ to be Sujatha distribution.

(ii). In real life situation, the sustainability of the components of population differs from each other in terms of heterogeneity. The analysis of data from such populations, heterogeneity can easily be taken into consideration using compound distributions. G-LD and G-SD are the two compound distributions proposed for the analysis of such variation in the components of populations. As Sujatha distribution provides better fit over Lindley and Shanker distributions, it is the expectation that the G-SUD would provide better fit over existing compound distributions.

(iii). In general, compound distribution is the most suited distributions for the datasets having long right tail, which have been observed in some real lifetime datasets relating to cancer datasets.

(iv). As Sujatha distribution performs well compared to exponential and Lindley distribution so it is hoped that G-SUD would performs better over the classical gamma and Weibull distributions as well as other two-parameters distributions.

The whole paper is divided into eleven sections. The section one is introductory in nature. The gamma-Sujatha probability model and some of its results are given in section two. The hazard function and the reversed hazard function of the proposed probability model are given in section 3. Section four contains the quantile and the moments of the distribution. The extreme order statistics and the stochastic ordering of the distribution are given in sections 5 and 6 respectively. The maximum likelihood estimation of parameters and the estimation of stress-strength parameter of the distribution are discussed in sections seven and eighth. The simulation study to know the consistency of maximum likelihood estimators and applications of the distribution are provided in sections nine and ten respectively. The conclusion of the whole paper is given in section eleven.

II. Gamma-Sujatha Distribution

The pdf and the cdf of gamma-Sujatha distribution (G-SUD) are obtained as

$$f_{G-SUD}(x; \varphi, \omega) = \frac{\varphi \omega^3 \left[(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right] x^{\varphi-1}}{(\omega^2 + \omega + 2)(\omega + x)^{\varphi+3}}; x > 0, \varphi > 0, \omega > 0 \tag{7}$$

$$F_{G-SUD}(x; \varphi, \omega) = \frac{x^\varphi \left[(\omega + x)^2 \{ \omega + (\varphi + 1)(\varphi + \omega + 2) \} - 2x(\omega + x) \{ \omega + \varphi(\varphi + 2) \} + \varphi(\varphi + 1)x^2 \right]}{(\omega^2 + \omega + 2)(\omega + x)^{2+\varphi}}; x > 0, \varphi > 0, \omega > 0 \tag{8}$$

Figure 1 and 2 shows the pdf and cdf of G-SUD for selected values of parameters. The G-SUD shows the tendency to accommodate right tail and for particular values of parameters, the tail approach to zero at a faster rate. This means that G-SUD would provide better fit appropriately to those datasets where there is an extended right tail or the right tail approaches to zero at a faster rate. Such datasets are quite prevalent in the biomedical sciences relating to survival times of cancer patients.

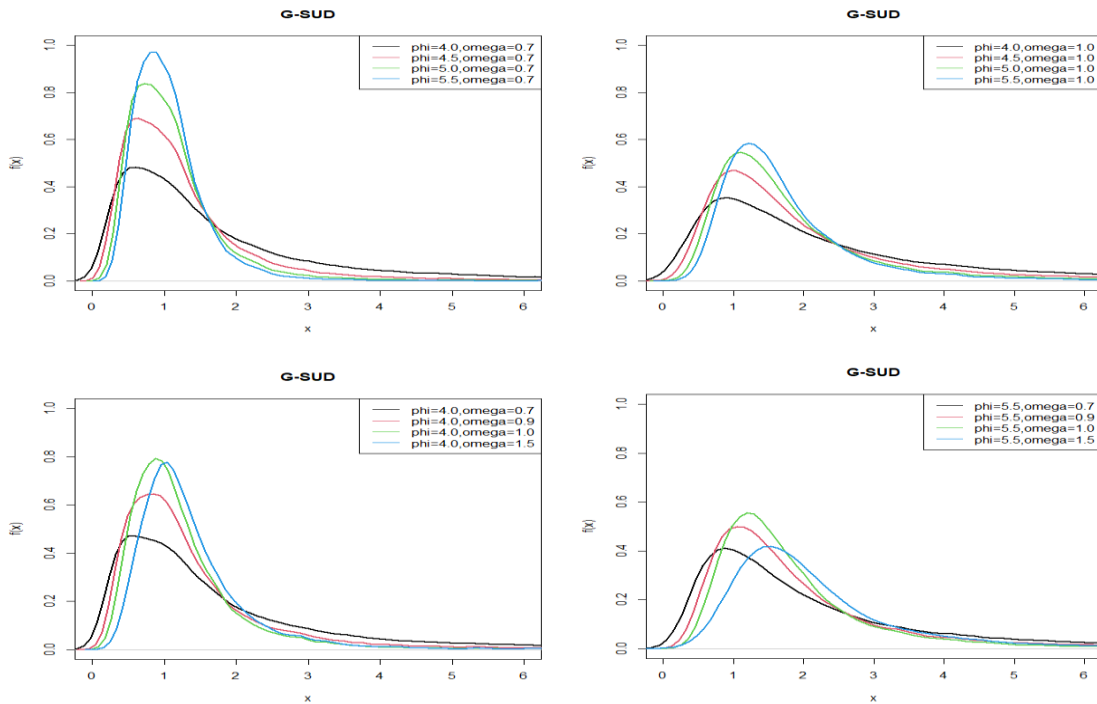


Fig. 1: pdf plots of G-SUD

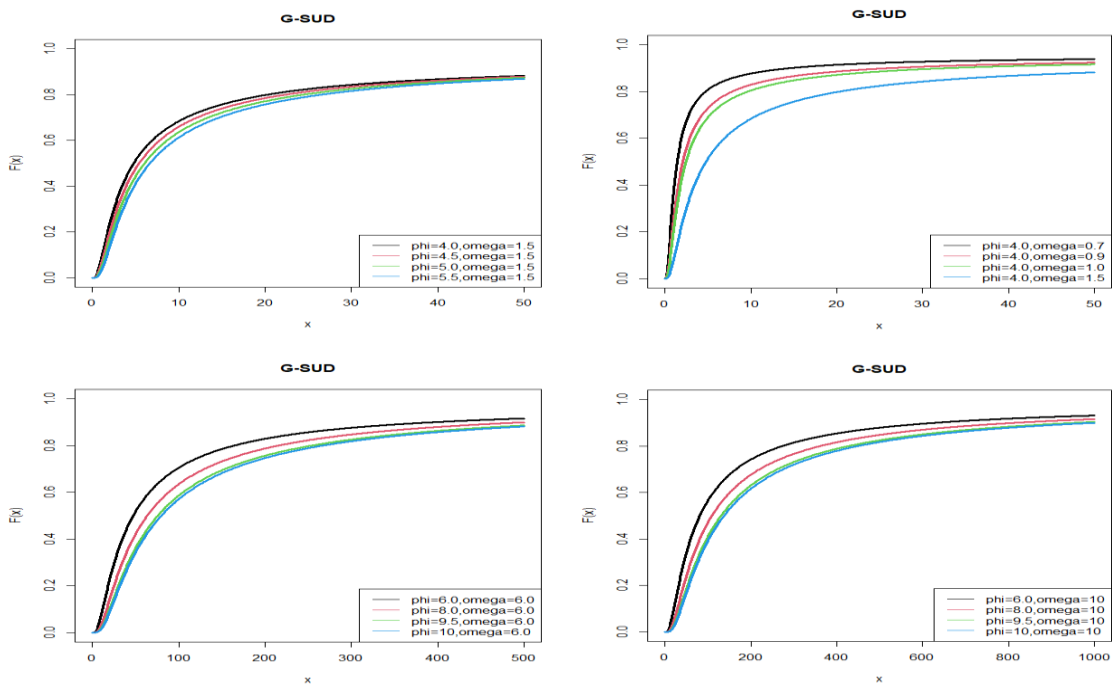


Fig. 2: cdf plots of G-SUD

Theorem 1: The G-SUD is decreasing for $\varphi \leq 1$.

Proof: We have,

$$f(x; \varphi, \omega) = \frac{\varphi \omega^3 \left[(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right] x^{\varphi-1}}{(\omega^2 + \omega + 2)(\omega + x)^{\varphi+3}}; x > 0, \varphi > 0, \omega > 0$$

$$\log f(x; \varphi, \omega) = \log \left[(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right] + (\varphi - 1) \log(x) - (\varphi + 3) \log(\omega + x) + C.$$

where C is a constant. We have

$$\frac{d}{dx} \log f(x; \varphi, \omega) = \frac{\varphi - 1}{x} - \left[\frac{(\varphi + 1) \{ (\omega + x)^2 + (\varphi + 2)(\varphi + \omega + x + 3) \}}{(\omega + x) \{ (\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \}} \right]$$

For $\varphi \leq 1$, $\frac{d}{dx} \log f(x; \varphi, \omega) < 0$ and this means that $f(x)$ is decreasing for all x

III. Hazard function and Reversed hazard function

The hazard function and the reverse hazard function are two important functions of a distribution. The reliability (survival) function of G-SUD is given by

$$R(x; \varphi, \omega) = \frac{(\omega^2 + \omega + 2)(\omega + x)^{\varphi+2} - x^\varphi \left[(\omega + x)^2 \{ \omega + (\varphi + 1)(\varphi + \omega + 2) \} - 2x(\omega + x) \{ \omega + \varphi(\varphi + 2) \} + \varphi(\varphi + 1)x^2 \right]}{(\omega^2 + \omega + 2)(\omega + x)^{\varphi+2}} \tag{9}$$

The corresponding hazard and reversed Hazard function of G-SUD are given by

$$h(x; \varphi, \omega) = \frac{f(x; \varphi, \omega)}{R(x; \varphi, \omega)} = \frac{\varphi \omega^3 x^{\varphi-1} \left[(\omega + x)^2 + (\varphi + 1)(\omega + x) + (\varphi + 1)(\varphi + 2) \right]}{(\omega^2 + \omega + 2)(\omega + x)^{\varphi+3} - x^\varphi \left[(\omega + x)^2 \{ \omega + (\varphi + 1)(\varphi + \omega + 2) \} - 2x(\omega + x) \{ \omega + \varphi(\varphi + 2) \} + \varphi(\varphi + 1)x^2 \right]} \tag{10}$$

$$r(x; \varphi, \omega) = \frac{f(x; \varphi, \omega)}{F(x; \varphi, \omega)} = \frac{\varphi \omega^3 \left[(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right]}{x(\omega + x) \left[(\omega + x)^2 \{ \omega + (\varphi + 1)(\varphi + \omega + 2) \} - 2x(\omega + x) \{ \omega + \varphi(\varphi + 2) \} + \varphi(\varphi + 1)x^2 \right]} \tag{11}$$

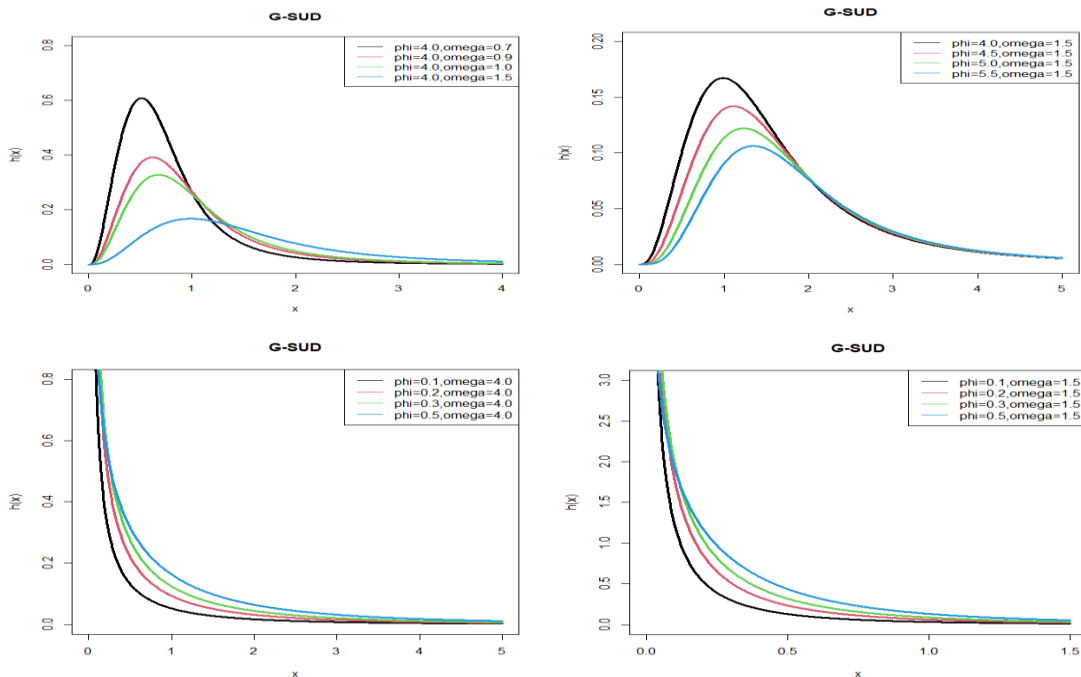


Fig.3: Hazard function of G-SUD

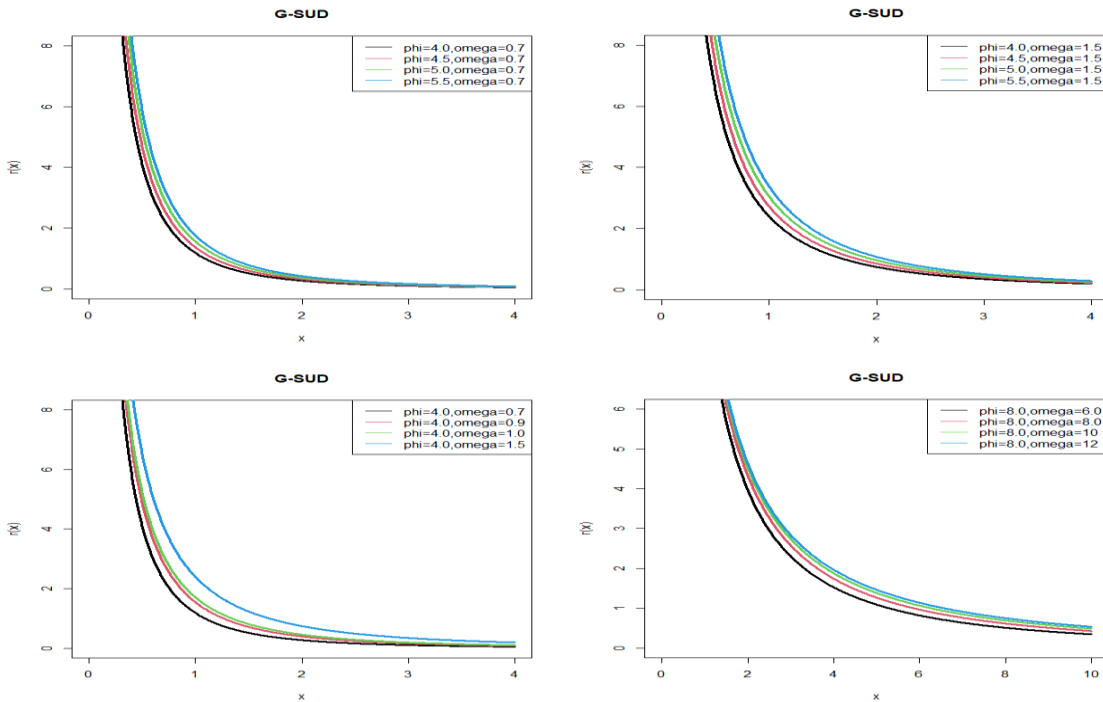


Fig.4: Reverse hazard function of G-SUD

Theorem 2: For $\varphi \leq 1$, the hazard function of the G-SUD is decreasing and for $\varphi > 1$ it is unimodal.

Proof: We have

$$f(x; \varphi, \omega) = \frac{\varphi \omega^3 \left[(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right] x^{\varphi-1}}{(\omega^2 + \omega + 2)(\omega + x)^{\varphi+3}}; x > 0, \varphi > 0, \omega > 0, \text{ and}$$

$$f'(x; \varphi, \omega) = \frac{\varphi \omega^3 \left[(\varphi \omega - \omega - 4x) \left\{ (\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right\} + x(\omega + x)(\varphi + 2\omega + 2x + 1) \right] x^{\varphi-1}}{x(\omega^2 + \omega + 2)(\omega + x)^{\varphi+4}}; x > 0, \varphi > 0, \omega > 0$$

Now, suppose that

$$\xi(x) = -\frac{f'(x; \varphi, \omega)}{f(x; \varphi, \omega)}$$

$$= -\frac{(\varphi - 1)}{x} + \frac{\left\{ (\omega + x)(\varphi^2 + \varphi(\omega + x + 3) + (\omega + x + 2)) + (\varphi + 1)(\varphi + 2)(\varphi + 3) \right\}}{(\omega + x) \left\{ (\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right\}}$$

This gives

$$\xi'(x) = \frac{(\varphi - 1)}{x^2} + \frac{(\omega + x) \left\{ (\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right\} \left\{ \varphi^2 + \varphi(2\omega + 2x + 3) + (2\omega + 2x + 2) \right\} - \left\{ (\omega + x)(\varphi^2 + \varphi(\omega + x + 3) + (\omega + x + 2)) + (\varphi + 1)(\varphi + 2)(\varphi + 3) \right\} \times \left\{ (\omega + x)(\varphi + 3\omega + 3x + 1) + (\varphi + 2)(\varphi + \omega + x + 2) \right\}}{(\omega + x)^2 \left\{ (\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2) \right\}}$$

It is quite obvious that for $\varphi \leq 1$, $\xi'(x) < 0$ and for $\varphi > 1$, $\xi'(x) < 0$ has a global maximum at mode (say x_0).

Theorem 3: The G-SUD has decreasing reverse hazard function.

Proof: We have,

$$r(x) = \frac{\varphi\omega^3 \left[(\omega+x)^2 + (\varphi+1)(\varphi+\omega+x+2) \right]}{x(\omega+x) \left[(\omega+x)^2 \{ \omega + (\varphi+1)(\varphi+\omega+2) \} - 2x(\omega+x) \{ \omega + \varphi(\varphi+2) \} + \varphi(\varphi+1)x^2 \right]}$$

This gives

$$\frac{d}{dx} \log r(x) = \frac{\varphi\omega \left\{ \varphi + \omega + 2(2x+2-x^2) \right\} - \varphi\omega \left[\begin{matrix} -\omega \{ \omega(3+4x) - 2x(2-3x) - 7 \} \\ -x \{ 5x - \varphi(4-3x) - 2 \} \end{matrix} \right] - 2 \left[\omega \{ \omega(\omega^2 + x^2 + x) + 2x \} + x^2 \right]}{\left\{ (\omega+x)^2 + (\varphi+1)(\varphi+\omega+x+2) \right\} \left[\begin{matrix} (\omega+x)^2 \{ \omega + (\varphi+1)(\varphi+\omega+2) \} \\ -2x(\omega+x) \{ \omega + \varphi(\varphi+2) \} + \varphi(\varphi+1)x^2 \end{matrix} \right]} - \frac{1}{x} - \frac{1}{(\omega+x)} < 0$$

This proves the theorem for all φ, ω .

IV. Quantiles and Moments

The p th quantiles x_p of G-SUD is defined by $F(x_p) = p$, is the root of the equation

$$\frac{x_p^\varphi \left[(\omega+x_p)^2 \{ \omega + (\varphi+1)(\varphi+\omega+2) \} - 2x(\omega+x_p) \{ \omega + \varphi(\varphi+2) \} + \varphi(\varphi+1)x_p^2 \right]}{(\omega^2 + \omega + 2)(\omega+x_p)^{2+\varphi}} = p \tag{12}$$

This gives

$$x_p = \frac{(\omega+x_p)^2 \{ \omega + (\varphi+1)(\varphi+\omega+2) \} - 2x(\omega+x_p) \{ \omega + \varphi(\varphi+2) \} + \varphi(\varphi+1)x_p^2}{p(\omega^2 + \omega + 2) \left(1 + \frac{\omega}{x_p} \right)^{\varphi+1}} \tag{13}$$

It should be noted that this x_p may be used to generate G-SUD random variates. Further, the median of G-SUD can be obtained from above equation by taking $p = \frac{1}{2}$.

The moments of G-SUD can be obtained as follows:

If $X \sim \text{G-SUD}(\varphi, \omega)$ then,

$$E(X) = E(E(X | \lambda)) = E\left(\frac{\varphi}{\lambda}\right) = \varphi E\left(\frac{1}{\lambda}\right) = \infty$$

Thus, in general, $E(X^r) = \infty$ for $r \geq 1$. This means that all moments of G-SUD are infinite and hence G-SUD has no mean. As G-SUD has no mean, if we take a sample (X_1, X_2, \dots, X_n) from G-SUD, then mean \bar{X} does not tend to a particular value. Since G-SUD has no raw and central moments, we have to derive inverse moments. Negative moments are useful in several life applications, such as life testing problems and estimation purpose. The negative moments for G-SUD can be obtained as follows:

The r^{th} negative moment (about origin) $\mu_{(-r)}'$, of the G-SUD is given by,

$$\mu_{(-r)}' = \frac{\Gamma(\varphi-r)}{\Gamma(\varphi)} \cdot \frac{r! \left[\omega^2 + \omega(r+2) + (r+1)(r+2) \right]}{\omega^r (\omega^2 + \omega + 2)}; r = 1, 2, 3, \dots \tag{14}$$

Thus, for $r = 1, 2, 3$ and 4 , we have

$$\mu_{(-1)}' = \frac{(\omega^2 + 2\omega + 6)}{\omega(\omega^2 + \omega + 2)(\varphi - 1)}, \varphi > 1 \tag{15}$$

$$\mu_{(-2)}' = \frac{2(\omega^2 + 3\omega + 12)}{\omega^2(\omega^2 + \omega + 2)(\varphi - 1)(\varphi - 2)}, \varphi > 2 \tag{16}$$

$$\mu_{(-3)}' = \frac{6(\omega^2 + 4\omega + 20)}{\omega^3(\omega^2 + \omega + 2)(\varphi - 1)(\varphi - 2)(\varphi - 3)}, \varphi > 3 \tag{17}$$

$$\mu_{(-4)}' = \frac{24(\omega^2 + 5\omega + 30)}{\omega^4(\omega^2 + \omega + 2)(\varphi - 1)(\varphi - 2)(\varphi - 3)(\varphi - 4)}, \varphi > 4 \tag{18}$$

It is obvious from the above expressions for negative moments that negative moments are not defined for $\varphi \leq 1$.

V. Extreme Order Statistics

Let, $X_{1:n}, \dots, X_{n:n}$ be the order statistics of a random sample of size n from the G-SUD(φ, ω) distribution with distribution function $F(x)$. The cdf of the minimum order statistic $X_{1:n}$ is given by

$$F_{X_{1:n}}(x) = 1 - \left[\frac{(\omega^2 + \omega + 2)(\omega + x)^{2+\varphi} - x^\varphi \left[(\omega + x)^2 \{ \omega + (\varphi + 1)(\varphi + \omega + 2) \} - 2x(\omega + x) \{ \omega + \varphi(\varphi + 2) \} + \varphi(\varphi + 1)x^2 \right]}{(\omega^2 + \omega + 2)(\omega + x)^{2+\varphi}} \right]^n$$

The cdf of the maximum order statistic $X_{n:n}$ is given by

$$F_{X_{n:n}}(x) = \left\{ \frac{x^\varphi \left[(\omega + x)^2 \{ \omega + (\varphi + 1)(\varphi + \omega + 2) \} - 2x(\omega + x) \{ \omega + \varphi(\varphi + 2) \} + \varphi(\varphi + 1)x^2 \right]}{(\omega^2 + \omega + 2)(\omega + x)^{2+\varphi}} \right\}^n$$

VI. Stochastic Orderings

In probability theory and Statistics, a stochastic order quantifies the concept of one random variable being “bigger” than other. In many problems, it becomes necessary to compare two lifetime distributions with reference to some of their characteristics. Stochastic orders provide the necessary tools in such case.

A random variable X is said to be smaller than a random variable Y in the

- i. Stochastic order ($X \prec_{st} Y$) if $F_X(x) \geq F_Y(y)$ for all x
- ii. Hazard rate order ($X \prec_{hr} Y$) if $h_X(x) \geq h_Y(y)$ for all x
- iii. Mean residual life order ($X \prec_{mrl} Y$) if $m_X(x) \geq m_Y(y)$ for all x
- iv. Likelihood ratio order ($X \prec_{lr} Y$) if $\frac{f_X(x)}{f_Y(y)}$ decrease in x
- iv. Likelihood ratio order ($X \prec_{lr} Y$) if $\frac{f_X(x)}{f_Y(y)}$ decrease in x

The following results due to Shaked and Shantikumar [14] are well known for establishing

stochastic ordering of distributions:

$$\begin{aligned} X \prec_{lr} Y &\Rightarrow X \prec_{hr} Y \Rightarrow X \prec_{mrl} Y \\ &\Downarrow \\ X &\prec_{st} Y \end{aligned}$$

Theorem 4: Let $X_1 \sim \text{G-SUD}(\varphi_1, \omega_1)$ and $X_2 \sim \text{G-SUD}(\varphi_2, \omega_2)$. If $\varphi_1 = \varphi_2 = \varphi$ and $\omega_1 \leq \omega_2$ if $\omega_1 = \omega_2 = \omega \geq 1$ with $\varphi_1 \leq \varphi_2$, then $X_1 \prec_{lr} X_2 \Rightarrow X_1 \prec_{hr} X_2 \Rightarrow X_1 \prec_{st} X_2$.

Proof: We have

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} = \frac{\varphi_1 \omega_1^3 \left[(\omega_1 + x)^2 + (\varphi_1 + 1)(\varphi_1 + \omega_1 + x + 2) \right] (\omega_2^2 + \omega_2 + 2) (\omega_2 + x)^{\varphi_2 + 3}}{\varphi_2 \omega_2^3 \left[(\omega_2 + x)^2 + (\varphi_2 + 1)(\varphi_2 + \omega_2 + x + 2) \right] (\omega_1^2 + \omega_1 + 2) (\omega_1 + x)^{\varphi_1 + 3}} x^{\varphi_1 - \varphi_2} \quad (19)$$

Case I: For $\varphi_1 = \varphi_2 = \varphi$, we get

$$\begin{aligned} G_1(x) &= \frac{\omega_1^3 \left[(\omega_1 + x)^2 + (\varphi + 1)(\varphi + \omega_1 + x + 1) \right] (\omega_2^2 + \omega_2 + 2) \left(\frac{\omega_2 + x}{\omega_1 + x} \right)^{\varphi + 3}}{\omega_2^3 \left[(\omega_2 + x)^2 + (\varphi + 1)(\varphi + \omega_2 + x + 1) \right] (\omega_1^2 + \omega_1 + 2) \left(\frac{\omega_2 + x}{\omega_1 + x} \right)^{\varphi + 3}} \\ \frac{d \log G_1(x)}{dx} &= \left(\frac{\varphi + 3}{\omega_2 + x} - \frac{2(\omega_2 + x) + (\varphi + 1)}{(\omega_2 + x)^2 + (\varphi + 1)(\varphi + \omega_2 + x + 2)} \right) - \left(\frac{\varphi + 3}{\omega_1 + x} - \frac{2(\omega_1 + x) + (\varphi + 1)}{(\omega_1 + x)^2 + (\varphi + 1)(\varphi + \omega_1 + x + 2)} \right) \\ &= Q(\omega_2) - Q(\omega_1) \end{aligned} \quad (20)$$

Where

$$\begin{aligned} Q(\omega) &= \left(\frac{\varphi + 3}{\omega + x} - \frac{2(\omega + x) + (\varphi + 1)}{(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2)} \right) \\ \frac{d}{d\omega} Q(\omega) &= \frac{-(\varphi + 3)}{(\omega + x)^2} - \frac{2(\omega + x)(\varphi - \omega - x + 1) + (\varphi + 1)(\varphi + 2)}{\{(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2)\}^2} < 0 \end{aligned} \quad (21)$$

The X_1 is stochastically smaller than X_2 with respect to the likelihood ratio for $\varphi_1 = \varphi_2 = \varphi$ provided $\omega_1 \leq \omega_2$.

Case II: For $\omega_1 = \omega_2 = \omega \geq 1$, we get

$$G_2(x) = \frac{\varphi_1 \left[(\omega + x)^2 + (\varphi_1 + 1)(\varphi_1 + \omega + x + 2) \right] \left(\frac{x}{\omega + x} \right)^{\varphi_1 - \varphi_2}}{\varphi_2 \left[(\omega + x)^2 + (\varphi_2 + 1)(\varphi_2 + \omega + x + 2) \right] \left(\frac{x}{\omega + x} \right)^{\varphi_1 - \varphi_2}} \quad (22)$$

$$\begin{aligned} \frac{d \log G_2(x)}{dx} &= \left(\frac{2(\omega + x) + (\varphi_1 + 1)}{(\omega + x)^2 + (\varphi_1 + 1)(\varphi_1 + \omega + x + 2)} + \frac{\varphi_1}{x} - \frac{\varphi_1}{\omega + x} \right) - \left(\frac{2(\omega + x) + (\varphi_2 + 1)}{(\omega + x)^2 + (\varphi_2 + 1)(\varphi_2 + \omega + x + 2)} + \frac{\varphi_2}{x} - \frac{\varphi_2}{\omega + x} \right) \\ &= S(\varphi_1) - S(\varphi_2) \end{aligned} \quad (23)$$

Where

$$\begin{aligned} S(\varphi) &= \left(\frac{2(\omega + x) + (\varphi + 1)}{(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2)} + \frac{\varphi}{x} - \frac{\varphi}{\omega + x} \right) \\ \frac{d}{d\varphi} S(\varphi) &= \frac{-\{\omega(\omega + 4\varphi + 6) + x(x + 4\varphi + 2\omega + 6) + 2\varphi\}}{\{(\omega + x)^2 + (\varphi + 1)(\varphi + \omega + x + 2)\}^2} + \frac{1}{x} - \frac{1}{\omega + x} > 0 \text{ for } \omega \geq 1 \end{aligned}$$

Thus, for $\varphi_1 \leq \varphi_2$, $\frac{d \log G_2(x)}{dx} < 0$. The X_1 is stochastically smaller than X_2 with respect to the likelihood ratio for $\omega_1 = \omega_2 = \omega \geq 1$ provided $\varphi_1 \leq \varphi_2$.

VII. Estimation of parameters

Let (x_1, x_2, \dots, x_n) be the observed values of a random sample (X_1, X_2, \dots, X_n) from the G-SUD. Then the log-likelihood function is given by

$$L(\varphi, \omega) = \left(\frac{\varphi \omega^3}{\omega^2 + \omega + 2} \right)^n \frac{\prod_{i=1}^n \left[(\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) + (\varphi + 1)(\varphi + 2) \right] \left(\prod_{i=1}^n x_i \right)^{\varphi - 1}}{\prod_{i=1}^n (\omega + x_i)^{\varphi + 3}}$$

The log-likelihood function of G-SUD is thus obtained as

$$\begin{aligned} \ln L(\varphi, \omega) = n \ln \varphi + 3n \ln \omega - n \ln (\omega^2 + \omega + 2) + \sum_{i=1}^n \ln \left[(\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) + (\varphi + 1)(\varphi + 2) \right] \\ + (\varphi - 1) \sum_{i=1}^n \ln(x_i) - (\varphi + 3) \sum_{i=1}^n \ln(\omega + x_i) \end{aligned}$$

The maximum likelihood estimators of φ and ω , say $\hat{\varphi}$ and $\hat{\omega}$ are the simultaneous solutions of the following log likelihood

$$\begin{aligned} \frac{\partial \ln L(\varphi, \omega)}{\partial \varphi} = \frac{n}{\varphi} + \sum_{i=1}^n \frac{(\omega + x_i) + (2\varphi + 3)}{(\omega + x_i)^2 + (\varphi + 1)(\omega + x_i)} + \sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n \ln(\omega + x_i) = 0 \\ \frac{\partial \ln L(\varphi, \omega)}{\partial \omega} = \frac{3n}{\omega} - \frac{n(2\omega + 1)}{(\omega^2 + \omega + 2)} + \sum_{i=1}^n \frac{2(\omega + x_i) + (\varphi + 1)}{\left\{ (\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) + (\varphi + 1)(\varphi + 2) \right\}} - (\varphi + 3) \sum_{i=1}^n \frac{1}{(\omega + x_i)} = 0 \end{aligned}$$

It is very difficult to solve these two log-likelihood equations directly, so we will use Fisher's scoring method. We have

$$\begin{aligned} \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \varphi^2} = \frac{-n}{\varphi^2} + \sum_{i=1}^n \frac{2 \left[(\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) \right] - (\omega + x_i) \left[(\omega + x_i) + (2\varphi + 3) \right]}{\left\{ (\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) \right\}^2} \\ \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \varphi \partial \omega} = \sum_{i=1}^n \frac{- \left[(\omega + x_i) + (2\varphi + 3) \right] \left[2(\omega + x_i) + (\varphi + 1) \right]}{\left\{ (\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) \right\}^2} - \sum_{i=1}^n \frac{1}{\omega + x_i} = \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \omega \partial \varphi} \\ \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \omega^2} = \frac{-3n}{\omega^2} - \frac{2n \left[(\omega^2 + \omega + 2) - n(2\omega + 1)^2 \right]}{(\omega^2 + \omega + 2)^2} - \sum_{i=1}^n \frac{2\omega \left[(\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) + (\varphi + 1)(\varphi + 2) \right] - \left[2(\omega + x_i) + (\varphi + 1) \right]^2}{\left\{ (\omega + x_i)^2 + (\varphi + 1)(\omega + x_i) + (\varphi + 1)(\varphi + 2) \right\}^2} \end{aligned}$$

The following equation can be solved for MLE's of $\hat{\varphi}$ and $\hat{\omega}$ of G-SUD

$$\begin{pmatrix} \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \varphi^2} & \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \varphi \partial \omega} \\ \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \omega \partial \varphi} & \frac{\partial^2 \ln L(\varphi, \omega)}{\partial \omega^2} \end{pmatrix}_{\substack{\hat{\varphi} = \varphi_0 \\ \hat{\omega} = \omega_0}} \begin{pmatrix} \hat{\varphi} - \varphi_0 \\ \hat{\omega} - \omega_0 \end{pmatrix} = \begin{pmatrix} \frac{\partial \ln L(\varphi, \omega)}{\partial \varphi} \\ \frac{\partial \ln L(\varphi, \omega)}{\partial \omega} \end{pmatrix}_{\substack{\hat{\varphi} = \varphi_0 \\ \hat{\omega} = \omega_0}}$$

where φ_0 and ω_0 are initial value of φ and ω respectively. The initial values of the parameters taken in this paper for estimating parameters are $\varphi_0 = 0.5$ and $\omega_0 = 0.5$.

VIII. Estimation of the Stress-Strength parameter $R = P(X > Y)$

In reliability, the stress-strength model describes the life of a component which has a random strength X subjected to a random stress Y . The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactory Whenever $X > Y$. In this section our objective is to estimate $R = P(X > Y)$ when

$X \sim \text{G-SUD}(\varphi_1, \omega_1)$ and $Y \sim \text{G-SUD}(\varphi_2, \omega_2)$, X and Y are independently distributed. The, the Stress- Strength Parameter is given by

$$\begin{aligned} R &= P(X > Y) = \int_0^\infty P(X > Y | Y = y) f_Y(y) dy \\ &= \int_0^\infty [1 - F_X(y)] f_Y(y) dy \\ &= 1 - \int_0^\infty \frac{\varphi_2 \omega_2^3 \left[\begin{array}{l} (\omega_1 + y)^2 \{ \omega_1 + (\varphi_1 + 1)(\omega_1 + \varphi_1 + 2) \} \\ - 2y(\omega_1 + y) \{ \omega_1 + \varphi_1(\varphi_1 + 2) \} \\ + \varphi_1(\varphi_1 + 1)y^2 \end{array} \right] \times \left[\begin{array}{l} (\omega_2 + y)^2 + (\varphi_2 + 1)(\omega_2 + y) \\ + (\varphi_2 + 1)(\varphi_2 + 2) \end{array} \right]}{(\omega_1^2 + \omega_1 + 2)(\omega_2^2 + \omega_2 + 2)(\omega_1 + y)^{\varphi_1 + 2} (\omega_1 + y)^{\varphi_2 + 3}} y^{\varphi_1 + \varphi_2 - 1} dy \\ &= G(\varphi_1, \varphi_2, \omega_1, \omega_2) \end{aligned}$$

Let, (x_1, x_2, \dots, x_n) be the observed value of a random sample of size n from $\text{G-SUD}(\varphi_1, \omega_1)$ and (y_1, y_2, \dots, y_m) be the observed value of a random sample of size m from $\text{G-SUD}(\varphi_2, \omega_2)$.

The log-likelihood function of $\varphi_1, \varphi_2, \omega_1$ and ω_2 is given by

$$\begin{aligned} \ln L(\varphi_1, \varphi_2, \omega_1, \omega_2) &= n \ln(\varphi_1) + 3n \ln(\omega_1) - n \ln(\omega_1^2 + \omega_1 + 2) + \sum_{i=1}^n \ln \left[(\omega_1 + x_i)^2 + (\varphi_1 + 1)(\omega_1 + x_i) + (\varphi_1 + 1)(\varphi_1 + 2) \right] \\ &+ (\varphi_1 - 1) \sum_{i=1}^n \ln(x_i) - (\varphi_1 + 3) \sum_{i=1}^n \ln(\omega_1 + x_i) + m \ln(\varphi_2) + 3m \ln(\omega_2) - m \ln(\omega_2^2 + \omega_2 + 2) \\ &+ \sum_{i=1}^m \ln \left[(\omega_2 + y_i)^2 + (\varphi_2 + 1)(\omega_2 + y_i) + (\varphi_2 + 1)(\varphi_2 + 2) \right] + (\varphi_2 - 1) \sum_{i=1}^m \ln(y_i) - (\varphi_2 + 3) \sum_{i=1}^m \ln(\omega_2 + y_i) \end{aligned}$$

Now,

$$\begin{aligned} \hat{\varphi}_1 &= \frac{n}{\varphi_1} + \sum_{i=1}^n \frac{(\omega_1 + x_i) + (2\varphi_1 + 3)}{(\omega_1 + x_i)^2 + (\varphi_1 + 1)(\omega_1 + x_i)} + \sum_{i=1}^n \ln(x_i) - \sum_{i=1}^n \ln(\omega_1 + x_i) = 0 \\ \hat{\varphi}_2 &= \frac{m}{\varphi_2} + \sum_{i=1}^m \frac{(\omega_2 + y_i) + (2\varphi_2 + 3)}{(\omega_2 + y_i)^2 + (\varphi_2 + 1)(\omega_2 + y_i)} + \sum_{i=1}^m \ln(y_i) - \sum_{i=1}^m \ln(\omega_2 + y_i) = 0 \\ \hat{\omega}_1 &= \frac{3n}{\omega_1} - \frac{n(2\omega_1 + 1)}{(\omega_1^2 + \omega_1 + 2)} + \sum_{i=1}^n \frac{2(\omega_1 + x_i) + (\varphi_1 + 1)}{(\omega_1 + x_i)^2 + (\varphi_1 + 1)(\omega_1 + x_i) + (\varphi_1 + 1)(\varphi_1 + 2)} - (\varphi_1 + 3) \sum_{i=1}^n \frac{1}{(\omega_1 + x_i)} = 0 \\ \hat{\omega}_2 &= \frac{3m}{\omega_2} - \frac{m(2\omega_2 + 1)}{(\omega_2^2 + \omega_2 + 2)} + \sum_{i=1}^m \frac{2(\omega_2 + y_i) + (\varphi_2 + 1)}{(\omega_2 + y_i)^2 + (\varphi_2 + 1)(\omega_2 + y_i) + (\varphi_2 + 1)(\varphi_2 + 2)} - (\varphi_2 + 3) \sum_{i=1}^m \frac{1}{(\omega_2 + y_i)} = 0 \end{aligned}$$

Solving these non-linear equations using any iterative methods available in R packages we can obtain the MLEs of the parameters as $(\hat{\varphi}_1, \hat{\varphi}_2, \hat{\omega}_1, \hat{\omega}_2)$ and hence the MLE of R can thus be obtained as

$$\hat{S} = G(\hat{\phi}_1, \hat{\phi}_2, \hat{\omega}_1, \hat{\omega}_2)$$

IX. A Simulation Study

This section contains a simulation study to examine the consistency of maximum likelihood estimators of the G-SUD. The mean, bias (B), MSE and variance of the MLE's are computed using the formulae

$$Mean = \frac{1}{n} \sum_{i=1}^n \hat{H}_i, B = \frac{1}{n} \sum_{i=1}^n (\hat{H}_i - H), MSE = \frac{1}{n} \sum_{i=1}^n (\hat{H}_i - H)^2, Variance = MSE - B^2$$

Where, $H = (\omega, \varphi)$ and $\hat{H} = (\hat{\omega}_i, \hat{\phi}_i)$.

The simulation results for different parameter values of G-SUD have been presented in tables 1 and 2 respectively using acceptance-rejection method:

a. Acceptance -rejection method for generating random samples from the G-SUD consists of following steps.

- i. Generate a random variable Y from exponential (ω) and U from Uniform (0,1)
- ii. If $U \leq \frac{f(y)}{M g(y)}$, then set $X = Y$ ("accept the sample"); otherwise ("reject the sample")

and if reject then repeat the whole process until we get the required samples, where M is a constant.

b. The sample sizes $n = 25, 50, 100, 150, 200$ are taken

c. The parameter values are considered as $\varphi = 5.5, \omega = 0.6$ and $\varphi = 6, \omega = 10$

d. Each sample size is replicated 10000 times

Tables 1 and 2 reveal that for increasing sample size, the value of the biases, MSE and variances of the MLE of the parameters of G-SUD becoming smaller and certify the first-order asymptotic theory of maximum likelihood estimators.

Table 1: The mean, Biases, MSE and Variances of G-SUD for $\varphi = 5.0, \omega = 0.6$

Parameters	Sample Size	Mean	Bias	MSE	Variance
$\hat{\phi}$	25	5.105803	0.1058031	0.01352763	0.002333342
	50	5.097851	0.0978509	0.01195673	0.00238192
	100	5.093918	0.0939184	0.0109286	0.00210792
	150	5.092278	0.0922778	0.01075683	0.00224162
	200	5.089048	0.0890482	0.00983284	0.00190325
$\hat{\omega}$	25	0.595456	-0.0045436	0.00004471	0.00002407
	50	0.595628	-0.0043716	0.00004846	0.00002935
	100	0.596119	-0.0038801	0.00004259	0.00002753
	150	0.596454	-0.0035456	0.00003761	0.00002504
	200	0.596588	-0.0034117	0.00003651	0.00002487

Table 2: The mean, Biases, MSE and Variances of G-SUD for $\varphi = 6.0, \omega = 10$

Parameters	Sample Size	Mean	Bias	MSE	Variance
$\hat{\phi}$	25	5.945172	-0.05482844	0.0042597	0.00125354
	50	5.961664	-0.03833594	0.0027186	0.00271866
	100	5.980525	-0.01947528	0.0025010	0.00212172
	150	5.985068	-0.01893228	0.0023664	0.00200800
	200	5.987536	-0.01246439	0.0022490	0.00209365
$\hat{\omega}$	25	10.08853	0.08852744	0.0120358	0.00419872
	50	10.06313	0.06317850	0.0088113	0.00481984

	100	10.03813	0.03813436	0.0064691	0.00501485
	150	10.02125	0.02124674	0.0064981	0.00604670
	200	10.00821	0.00820760	0.0055774	0.00550691

X. Applications

This section deals with the goodness of fit of G-SUD over G-LD, G-SD, Weibull and gamma distributions to illustrate its applications and using two real datasets relating to survival time of acute bone cancer and head and neck cancer patients. The summary of the two datasets are presented in tables 3 and 4 respectively. The total time to test (TTT) plots of the two datasets are given in figures 5 and 6 respectively. The goodness of fit of the considered distributions for two datasets is provided in tables 5 and 6 respectively. The fitted plots of the considered distributions for the two datasets are given in figure 7. The p-p plots of the considered distributions for the two datasets are finally presented in figures 8 and 9 respectively. The datasets are as follows:

Dataset 1: Acute bone cancer

This dataset represents the survival times (in days) of 73 patients who diagnosed with acute bone cancer available in Mansour et al [15] and are as follows:

0.09, 0.76, 1.81, 1.10, 3.72, 0.72, 2.49, 1.00, 0.53,0.66, 31.61, 0.60, 0.20, 1.61, 1.88, 0.70, 1.36, 0.43, 3.16, 1.57, 4.93, 11.07, 1.63, 1.39, 4.54, 3.12,86.01, 1.92, 0.92, 4.04, 1.16, 2.26, 0.20, 0.94, 1.82, 3.99, 1.46, 2.75, 1.38, 2.76, 1.86, 2.68, 1.76,0.67, 1.29, 1.56, 2.83, 0.71, 1.48, 2.41, 0.66, 0.65, 2.36, 1.29, 13.75, 0.67, 3.70, 0.76, 3.63, 0.68,2.65, 0.95, 2.30, 2.57, 0.61, 3.93, 1.56, 1.29, 9.94, 1.67, 1.42, 4.18, 1.37.

Table 3: The summary of acute bone cancer dataset

Min.	1st Qu.	Median	Mean	Variance	3rd Qu.	Max.
0.090	0.920	1.570	3.755	112.33	2.750	86.010

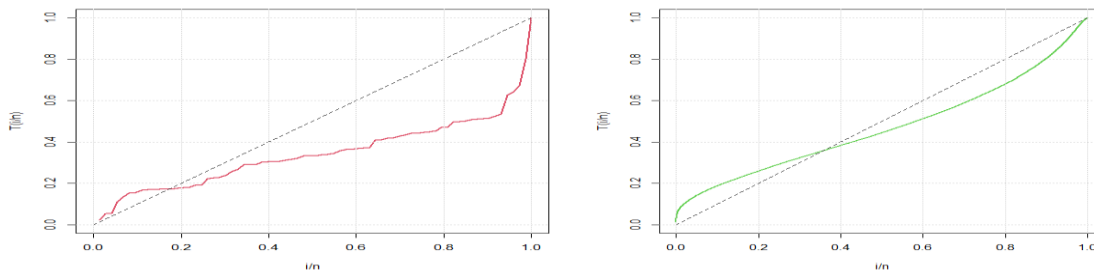


Fig.5: TTT-plot of the acute bone cancer dataset and simulated data of G-SUD respectively.

Dataset 2: Head and Neck cancer

This dataset is the survival time of 44 patients diagnosed by Head and Neck cancer disease are available in Efron [16] and are given by

12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194,195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776

Table 4: The summary of head and neck cancer dataset

Min.	1st Qu.	Median	Mean	Variance	3rd Qu.	Max.
12.20	67.21	128.50	223.48	93286.41	219.00	1776.00

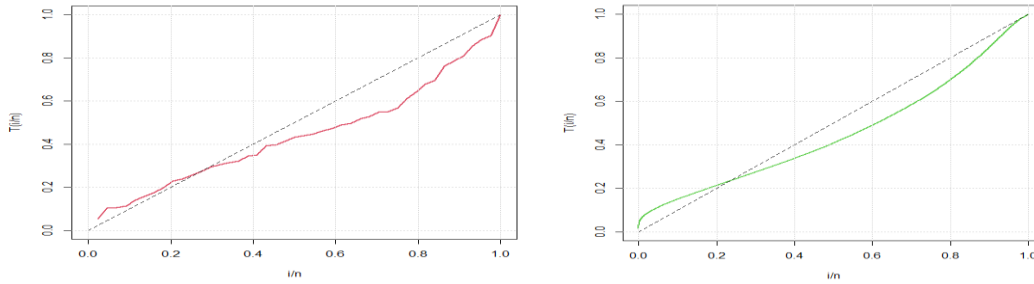


Fig.6: TTT-plot of the head and neck cancer dataset and simulated data of G-SUD respectively.

Table5: ML estimates, $-2\log L$, AIC , BIC and K-S statistics with their P-values of the distributions for acute bone cancer data set

Distributions	ML estimates $\hat{\phi}(S.E\ of\ \hat{\phi})$ $\hat{\omega}(S.E\ of\ \hat{\omega})$	$-2\log L$	AIC	BIC	K-S	p- value
G-SUD	4.4567 (1.1253) 0.7646 (0.1776)	281.7757	285.7757	300.0857	0.09	0.86
G-SD	4.8969(1.3904) 0.4967(0.1360)	282.8051	286.8051	301.1151	0.10	0.39
G-LD	5.1600(1.8468) 0.4375(0.1602)	284.315	288.315	302.625	0.11	0.33
Gamma	0.1985(0.0389) 0.7456(0.1057)	334.5311	338.5311	352.8411	0.56	0.00
Weibull	0.4395(0.0687) 0.7655(0.0567)	322.8033	326.8033	341.1133	0.25	0.00

Table 6: ML estimates, $-2\log L$, AIC,BIC and K-S statistics with their P-values of the distributions for head and neck cancer dataset.

Distributions	ML estimates $\hat{\phi}(S.E\ of\ \hat{\phi})$ $\hat{\omega}(S.E\ of\ \hat{\omega})$	$-2\log L$	AIC	BIC	K-S	p- value
G-SUD	8.6223 (11.3202) 11.1699(14.5932)	558.4763	562.4763	576.7863	0.08	0.90
G-SD	8.6787(11.7435) 10.0923(14.8515)	558.4641	562.4641	576.7741	0.09	0.81
G-LD	8.4483(10.4902) 11.1557(14.3688)	558.4555	562.4555	576.7655	0.09	0.70
Gamma	0.0047(0.0010) 1.0522(0.1886)	564.0254	568.0254	582.3354	1.00	0.00
Weibull	0.0070(0.0034) 0.9234(0.0809)	563.7155	567.7155	582.0255	0.5	0.04

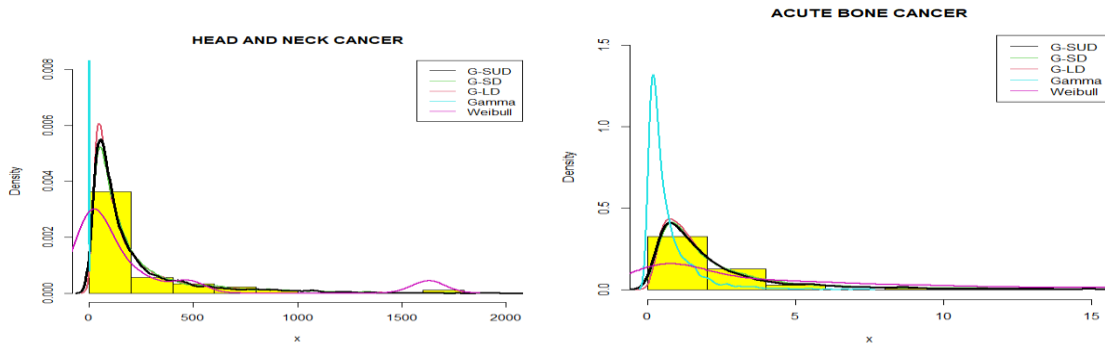


Fig. 7: Fitted plots of distributions for acute bone cancer and neck cancer datasets

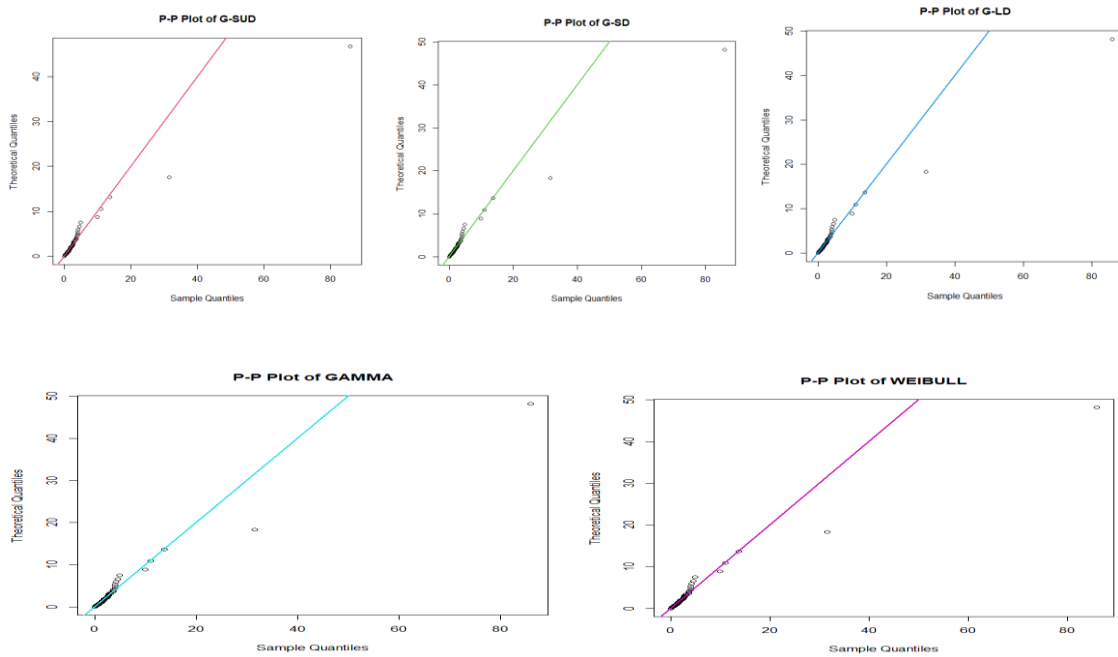


Fig. 8: P-P plots for considered distributions of acute bone cancer dataset

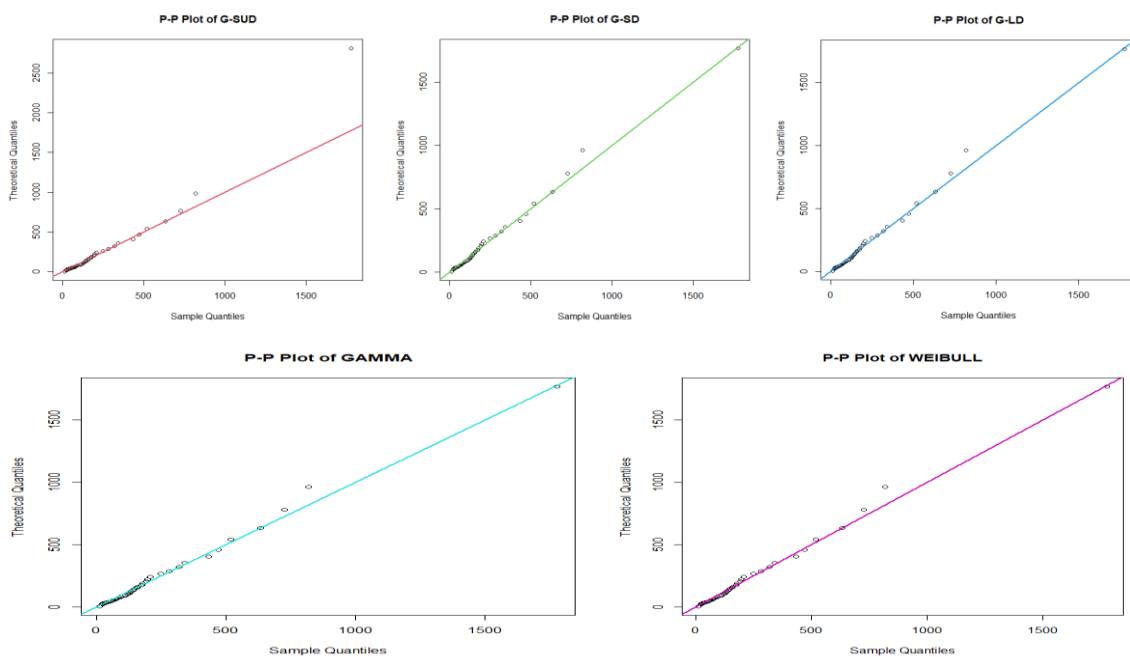


Fig. 9: P-P plots for considered distributions of head and neck cancer dataset

From the summary of the two datasets in tables 3 and 4, it is quite obvious that the considered datasets are highly positively skewed and highly over-dispersed. Based on the values of $-2\log L$, AIC (Akaike information criterion), Kolmogorov – Smirnov (K-S) statistic and the fitted plots of two parameter lifetime distributions, it is crystal clear from the goodness of fit that two parameters G-SUD is the best for modelling survival times of patients suffering from acute bone cancer and head and neck cancer. It can be recalled that recently Klakattawi [17] proposed a new extended Weibull distribution with five parameters and used it for analysing survival time of cancer patients and found that it gave much better fit than several two-parameter, three parameter, four parameter and five parameter lifetime distribution including Weibull distribution, alpha power Weibull (APW) distribution by Nassar et al [7], Beta-Weibull (BW) distribution by Famoye et al [4], Kumararaswamy-Weibull (Kum-W) distribution by Cordeiro et al [5], exponentiated generalized Weibull (EGW) distribution by Cordeiro et al [3], a new Kumaraswamy family of generalized Weibull distribution by Ahmed et al [18] and exponentiated Kumaraswamy Weibull distribution by Eissa [6], some among others. Here we would like to emphasize that the proposed gamma-Sujatha distribution (G-SUD) provides much closure fit than all these two-parameter, three-parameter, four-parameter and five-parameter lifetime distributions as it can be seen from the test of goodness of fit given by Klakattawi [17]. The most interesting feature of G-SUD is that being two-parameter distribution is much easier to characterize and handle the distribution as compared to three-parameter, four-parameter and five parameter distributions and hence it can be considered an important probability model for modeling survival time of cancer patients.

XI. Concluding Remarks

In this paper, we propose a gamma-Sujatha probability model, a compound of gamma and Sujatha distribution to model data of long tails. Some important statistical and reliability properties have been discussed. Maximum likelihood estimation has been discussed for estimating parameters and simulation studies to know the consistency of ML estimators are presented. The goodness of fit of the G-SUD has been compared with several well-known two-parameter distributions and observed that it provides much better fit and hence it can be considered as an important probability models for survival time of patients suffering from acute bone cancer and head and neck cancer in biomedical science. As the proposed distribution is the new probability model, a lot of works can be done in the future and definitely it will draw the attention of research workers in biomedical sciences and biomedical engineering.

Conflict of Interest

The Authors declare that there is no conflict of Interest.

References

- [1] Weibull, W. (1951). A Statistical distribution function of wide applicability. *Journal of applied mechanics*, 18:293-297
- [2] Shanker, R., Shukla, K.K., Shanker, R. and Tekie, A.L. (2016). On modelling of Lifetime data using two-parameter gamma and Weibull distributions . *Biometrics & Biostatistics International journal*, 4: 201 – 206.

- [3]Cordeiro,G.M.,Ortega,E.M.M and Da-Cunha, D.C.C. (2013). The exponentiated Generalized Class of Distributions. *Journal of Data Science*, 11:1-27.
- [4] Famoye,F., Lee,C. and Olumolade,O.(2005). The Beta-Weibull Distribution. *Journal of Statistical Theory and Applications*, 4:121-136.
- [5] Cordeiro,G.M.,Ortega,E.M.M and Nadarajah,S. (2010). The Kumaraswamy Weibull Distribution with Application to failure data. *Journal of Franklin Institute* ,347:1399-1429.
- [6] Eissa.F.H.(2017). The Exponentiated Kumaraswamy-Weibull Distribution with Application to Real Data. *International Journal of Statistics and Probability*, 6:167-182.
- [7] Nassar,M. Alzaatreh,A., Mead M. and Abo-Kasem,O. (2017). Alpha Power Weibull Distribution: Properties and Applications. *Communications in Statistics-theory and Methods*, 46:10236-10252.
- [8] Lindley, D.V. (1958). Fiducial Distribution and Bayes' Theorem. *Journal of The Royal Statistical Society*, 20: 102-107.
- [9] Shanker, R.(2015). Shanker Distribution and its Applications. *International Journal of Statistics and Applications*, 5:338-348.
- [10] Shanker ,R. (2016). Sujatha distribution and its Applications. *Statistics in Transition New Series*, 17:391-410.
- [11] Abdi, M., Asgharzadeh, A., Bakouch, H.S. and Alipour, Z. (2019). A new compound Gamma and Lindley distribution with Application to failure data. *Austrian Journal of Statistics*, 48:54-75.
- [12] Ray, M. and Shanker, R. (2023a). A Compound of Gamma and Shanker Distribution. *Reliability Theory & Applications*, 18: 87-99.
- [13] Ray, M. & Shanker, R. (2023b).A Compound of Exponential and Shanker Distribution With an Application. *Journal of Scientific Research of the Banaras Hindu University*, 67:39-46.
- [14] Shaked, M. and Shanthikumar, J.G. (1994). Stochastic Orders and Their Applications. Academic Press New Work.
- [15] Mansour M., Yousof H.M., Shehata W.A. and Ibrahim M.(2020). A new two parameter Burr XII distribution: properties, copula, different estimation methods and modeling acute bone cancer data. *Journal of Nonlinear Science and Applications*, 13:223–238.
- [16] Efron B.(1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association* ,83:414–425.
- [17] Klakattawi.H.S. (2022). Survival Analysis of Cancer Patients using a New Extended Weibull Distribution. *PLOS ONE*, 17:1-20.
- [18] Ahmed, M.A., Mahmoud M.R.and Elsherbini E.A. (2015). The New Kumaraswamy Family of Generalized Distributions with Application. *Pakistan Journal of Statistics and Operation Research*, 11:159-180.

ON A DISCRETE TIME SHOCK MODEL IN CRITICAL SITUATION

REZA FARHADIAN

•

Department of Statistics, Razi University, Kermanshah, Iran
farhadian.reza@yahoo.com

HABIB JAFARI

•

Department of Statistics, Razi University, Kermanshah, Iran
h.jafari@razi.ac.ir

Abstract

In this paper, we study a discrete time shock model which is defined based on the length of the time between successive shocks. For a system that is exposed to a sequence of random shocks over time under this model, if the interarrival time between two successive shocks is equal to a prefixed critical time point such as δ , the system fails, and the system is not damaged otherwise. We have considered two situations for the system, which are regular situation and critical situation, then we investigate the statistical behavior of the system's lifetime under these situations. More precisely, we obtain the probability generating function of the system's lifetime, the mean time to failure, the variance of the system's lifetime, the Laplace transform of the system's lifetime, and some other related results. We end the paper with an example including numerical comparisons of the results.

Keywords: discrete time shock model, intershock time, critical time point

1. INTRODUCTION

Studying systems in various sciences is very important in maintaining, improving and reducing their errors. Shock models are useful models for studying the systems which are subject to random shocks at random times. Since the probabilistic behavior of a system under the influence of factors such as shocks is important to prevent system failure, therefore, shock models have attracted a great deal of interest in applied probability, reliability theory, and engineering. There are three basic types of shock models in the literature, which are cumulative shock models, extreme shock models, and run shock models. In a cumulative shock model, a system break down because of a cumulative effect; in an extreme shock model, a system break down because of one single large shock; in a run shock model, the range of a certain number of consecutive shocks is considered a failure criterion. For more on these, see, e.g., [1], [8], [18], and [23]. Moreover, models that can be produced by combining these traditional models can be found in [9].

In addition to the above traditional shock models, there are other shock models that have been introduced and developed in recent years. The so-called δ -shock model is among these new models that have received more attention. According to the δ -shock model, the system fails when the interarrival time between two successive shocks (intershock time) falls below a

prefixed threshold $\delta > 0$. Therefore, the lifetime of the system is defined by $T = \sum_{i=1}^N X_i$, where X_1, X_2, \dots represent the intershock times, and the random variable N is the waiting time for the first intershock time which is less than δ , i.e., $\{N = n\}$ iff $\{X_1 > \delta, \dots, X_{n-1} > \delta, X_n \leq \delta\}$. The δ -shock model was first introduced by Li et al. in [13], after which it was widely studied by many scientists and researchers, for example, [14], [27], [15], [16], and [26]. Some extensions and generalizations were provided for the δ -shock model, see, e.g., [3], [19], and [20]. Moreover, a mixed shock model is defined by Wang and Zhang in [26] in which the system fails when an extreme shock occurs or a δ -shock. Ma and Li [17] have introduced and studied a censored δ -shock model. Eryilmaz [4] studied the lifetime behavior of a discrete time δ -shock model. Tuncel and Eryilmaz [24] investigate the survival function and the mean lifetime of the system failure considering the proportional hazard rate model. Goyal et al. [6] studied a general class of shock models with dependent intershock times.

In this paper, we aim to study a new discrete time shock model, which can be reduced to the classical δ -shock model in a special case. In general, in the new model, we do not have a critical threshold as was the case for δ -shock models, however, we have a critical time point so that if the intershock time is equal to the critical time point, then the system fails. To our knowledge, there is no such a shock model as this in the literature. Furthermore, we investigate the system under the assumption that the intershock times X_1, X_2, \dots are inflated at a particular point of time. This means that the frequency of some points of the times between successive shocks may occur more than expected in regular models. We will study the lifetime of the system under the new model and new assumptions. To this end, we obtain the probability generating function of the system's lifetime, the mean time to failure of the system, the variance of the system's lifetime, the Laplace transform of the system's lifetime, and some other related results. Furthermore, we provide an illustrative example for the new model.

The paper organized as follows. Section 2 introduces the model and assumptions. The statistical properties of the intershock times under the inflation property are investigate in Section 3. The life properties of the system under the new model are derived in Section 4. The paper ends with an example in Section 5.

2. MODEL DESCRIPTION

2.1. Regular Situation

We consider a system that is subject to a sequence of external shocks occur randomly over time. Let X_i denote the time between i th and $(i + 1)$ th shocks for $i = 1, 2, \dots$. We assume that the intershock times X_1, X_2, \dots take positive integer values and are independent identically distributed (i.i.d.) by an arbitrary discrete distribution with probability mass function (pmf) $P(X = x)$ (with $P(X = 0) = 0$). The performance of the system is such that when the intershock time is not equal to a prefixed positive integer δ (i.e. $X_n \neq \delta$), the system does not fail, and if $X_n = \delta$, the system fails. Under these assumptions the lifetime of the system is defined as follows:

$$T_\delta = \sum_{i=1}^N X_i, \tag{1}$$

where the stopping random variable N is defined as

$$\{N = n\} \Leftrightarrow \{X_1 \neq \delta, X_2 \neq \delta, \dots, X_{n-1} \neq \delta, X_n = \delta\}. \tag{2}$$

Hence, the pmf of N is obtained as follows:

$$P(N = n) = (1 - P(X_1 = \delta))^{n-1} P(X_1 = \delta), \quad n = 1, 2, \dots,$$

that is, the random variable N has the geometric distribution with mean $\frac{1}{P(X_1 = \delta)}$.



Figure 1: Visual understanding of the system under the model.

Note that in the case where $\delta = \min\{X_i : i = 1, 2, \dots\}$ we have

$$\begin{aligned} \{N = n\} &= \{X_1 \neq \delta, X_2 \neq \delta, \dots, X_{n-1} \neq \delta, X_n = \delta\} \\ &= \{X_1 > \delta, X_2 > \delta, \dots, X_{n-1} > \delta, X_n = \delta\}, \\ &= \{X_1 > \delta, X_2 > \delta, \dots, X_{n-1} > \delta, X_n \leq \delta\}, \end{aligned}$$

thus, the model reduces to the classical δ -shock model.

Note also that when the intershock times X_i 's follow a regular distribution, we say that the system is in a regular situation.

2.2. Critical Situation

We are interested in studying the conditions in which the system is in a critical situation. Since the system fails whenever the intershock time equals the critical time point δ , so one of the conditions that makes the system to go from a regular situation to a critical situation is that the intershock times have an overdispersion at the point δ . This means that the frequency of intershock times at point δ may occur more than expected in a regular model. In this case, the intershock times X_i 's follow an inflated distribution (with inflation point δ), and we denote the lifetime of the system by T_δ^+ .

3. PROPERTIES OF INTERSHOCK TIMES UNDER INFLATION

According to the descriptions in Section 2, the intershock time X takes positive integer values with a pmf $P(X = x)$. Based on the assumptions, the distribution of intershock times X is either regular (not inflated) or inflated. Thus, as useful notations, henceforth, instead of $P(X = x)$ we will use $P_{reg}(X = x)$ and $P_+(X = x)$, respectively, if X is distributed by a regular distribution or if X is distributed by an inflated distribution.

In the context of inflated distributions, a discrete random variable X with support \mathcal{X} is said to be inflated at a particular point k ($k \in \mathcal{X}$), if its pmf is given by

$$P_+(X = x) = \begin{cases} \alpha + (1 - \alpha)P_{reg}(X = x) & \text{if } x = k, \\ (1 - \alpha)P_{reg}(X = x) & \text{if } x \in \mathcal{X} - \{k\}, \end{cases} \quad (3)$$

where $\alpha \in [0, 1]$ is an inflation parameter.

For example, if X_1, X_2, \dots are i.i.d. 3-inflated geometric distribution, we have $P_{reg}(X = x) = p(1 - p)^{x-1}$ for $x = 1, 2, \dots$, and $0 < p \leq 1$, thus

$$P_+(X = x) = \begin{cases} \alpha + (1 - \alpha)p(1 - p)^2 & \text{if } x = 3, \\ (1 - \alpha)p(1 - p)^{x-1} & \text{if } x \in \{1, 2, 4, 5, 6, \dots\}. \end{cases}$$

For some references on inflated distributions, see, for example, [5], [10], [12], and [21].

In following, we investigate some distributional properties of a general k -inflated distribution. First, the main property of an inflated distribution is proved in the following theorem.

Theorem 1. Let the intershock time X follows the k -inflated distribution in (3). Then $P_+(X = k) \geq P_{reg}(X = k)$.

Proof. We have $0 \leq P_{reg}(X = x) \leq 1$ for any x in its support. On the other hand, for $\alpha = 0$ we have $P_+(X = k) = P_{reg}(X = k)$, and for $\alpha = 1$, $P_+(X = k)$ is degenerated pmf. Now, let us consider $\alpha \in (0, 1)$. Therefore multiplying both sides of $P_{reg}(X = k) < 1$ by α and adding $-\alpha P_{reg}(X = k)$ to both sides gives

$$\alpha - \alpha P_{reg}(X = k) > 0.$$

Finally, adding both sides by $P_{reg}(X = k)$, we get $\alpha + (1 - \alpha)P_{reg}(X = k) > P_{reg}(X = k)$, that is, $P_+(X = k) > P_{reg}(X = k)$. This completes the proof. ■

Therefore, by Theorem 1, in a k -inflated distribution the probability of occurrence of k is higher than in a distribution with regular pmf $P_{reg}(X = x)$.

In the next theorem, we obtain the cumulative distribution function (cdf) of a k -inflated random variable distributed by (3).

Theorem 2. Let the intershock time X follows the k -inflated distribution in (3). If the cdf of X in regular mode is denoted by $F_{reg}(x)$, then the cdf of X in inflated mode is given by

$$F_+(x) = \begin{cases} (1 - \alpha)F_{reg}(x) & \text{if } x < k, \\ \alpha + (1 - \alpha)F_{reg}(x) & \text{if } x \geq k. \end{cases}$$

Proof. Using (3), we have for $x < k$,

$$F_+(x) = \sum_{j=0}^x P_+(X = j) = (1 - \alpha) \sum_{j=0}^x P_{reg}(X = j) = (1 - \alpha)F_{reg}(x),$$

and if $x \geq k$, we have

$$\begin{aligned} F_+(x) = \sum_{j=0}^x P_+(X = j) &= \left(\left((1 - \alpha) \sum_{j=0}^x P_{reg}(X = j) \right) - (1 - \alpha)P_{reg}(X = k) \right) \\ &+ \alpha + (1 - \alpha)P_{reg}(X = k) \\ &= \alpha + (1 - \alpha)F_{reg}(x). \end{aligned}$$

This completes the proof. ■

Theorem 3. Let the intershock time X follows the k -inflated distribution in (3). If the reliability function of X in regular mode is denoted by $\bar{F}_{reg}(x)$, then the reliability function of X in inflated mode is given by

$$\bar{F}_+(x) = \begin{cases} \alpha + (1 - \alpha)\bar{F}_{reg}(x) & \text{if } x < k, \\ (1 - \alpha)\bar{F}_{reg}(x) & \text{if } x \geq k. \end{cases}$$

Proof. By using the definition of reliability function ($\bar{F}_+(x) = P_+(X > x)$) and using Theorem 2 the proof is straightforward. ■

In the following theorem, we obtain the moments related to interarrival times between successive shocks under k -inflated distribution in (3).

Theorem 4. Let $E_{reg}[X^r]$ be the r th moment of the intershock time X in its regular mode. The r th moment of the k -inflated version of X distributed by (3) is

$$E_+[X^r] = \alpha k^r + (1 - \alpha)E_{reg}[X^r].$$

In particular, for $r = 1$ and $r = 2$, we have

$$E_+[X] = \alpha k + (1 - \alpha)E_{reg}[X], \tag{4}$$

$$E_+[X^2] = \alpha k^2 + (1 - \alpha)E_{reg}[X^2]. \tag{5}$$

Proof. We have

$$\begin{aligned} E_+[X^r] = \sum_x x^r P_+(X = x) &= \left(\left((1 - \alpha) \sum_x x^r P_{reg}(X = x) \right) - (1 - \alpha)k^r P_{reg}(X = k) \right) \\ &+ k^r (\alpha + (1 - \alpha)P_{reg}(X = k)) \\ &= \alpha k^r + (1 - \alpha)E_{reg}[X^r]. \end{aligned}$$

The theorem is proved. ■

Corollary 1. The variance of a k -inflated intershock time X distributed by (3) is given by

$$\begin{aligned} Var_+(X) = Var_{reg}(X) &+ \alpha^2 \left(2kE_{reg}[X] - E_{reg}^2[X] - k^2 \right) \\ &+ \alpha \left(2E_{reg}^2[X] - E_{reg}[X^2] - 2kE_{reg}[X] + k^2 \right), \end{aligned}$$

where $Var_{reg}(X)$ is the variance of X in its regular mode.

Proof. Use Equations (4) and (5) and the definition of variance. This completes the proof. ■

In following, we calculate the probability generating function (pgf) of the inflated distribution in (3).

Theorem 5. If intershock time X follows the k -inflated distribution in (3), then its pgf is

$$G_X^+(z) = \alpha z^k + (1 - \alpha)G_X^{reg}(z),$$

where $G_X^{reg}(z)$ is the pgf of X in its regular mode.

Proof. We have

$$\begin{aligned} G_X^+(z) = E_+[z^X] = \sum_x z^x P_+(X = x) &= \alpha z^k + (1 - \alpha) \sum_x z^x P_{reg}(X = x) \\ &= \alpha z^k + (1 - \alpha)E_{reg}(z^X) \\ &= \alpha z^k + (1 - \alpha)G_X^{reg}(z). \end{aligned}$$

The theorem is proved. ■

Note that if $\alpha = 0$, then Equation (3) gives $P_+(X = x) = P_{reg}(X = x)$, $F_+(x) = F_{reg}(x)$, $\bar{F}_+(x) = \bar{F}_{reg}(x)$, $E_+[X^r] = E_{reg}[X^r]$, and $G_X^+(z) = G_X^{reg}(z)$. Therefore, when the inflation parameter α tends to zero, then the inflation property for the random variable X reduces the regular state.

4. STATISTICAL PROPERTIES OF THE SYSTEM'S LIFETIME

Following Section 2, when the intershock times are not equal to δ , the system continues to work safely. Indeed, if our observations of intershock times have overdispersion at point δ , then the system is in an critical situation and is expected to fail soon. Hence, it is important to study the lifetime of the system when the intershock times follow an δ -inflated distribution. This is done in the following.

The reliability function (or survival function) of the system's lifetime can be expressed by

$$P(T_\delta > t) = P\left(\sum_{i=0}^N X_i > t\right) = \sum_{n=1}^{\infty} P\left(\sum_{i=1}^n X_i > t\right) P(N = n). \quad (6)$$

Since the distribution of intershock times is arbitrary, therefore, deriving a general explicit representation of the reliability function from Equation (6) is difficult, or very complex if obtained. Therefore, the probability generating function (pgf) can be useful for calculation of the probability mass function of the system's lifetime. Following theorem gives us the pgf of the system's lifetime.

Theorem 6. Let X_1, X_2, \dots be the intershock times in the model described in Section 2, and let also X denotes a generic random variable of X_i 's. When the system is in the critical situation, the pgf of the system's lifetime becomes

$$G_{T_\delta}(z) = \frac{(\alpha z^\delta + (1 - \alpha)G_X^{reg}(z)) (\alpha + (1 - \alpha)P_{reg}(X = \delta))}{1 - (1 - \alpha)P_{reg}(X \neq \delta) (\alpha z^\delta + (1 - \alpha)G_X^{reg}(z))}. \quad (7)$$

Proof. Since the system is in the critical situation, so X is δ -inflated distributed. The pgf of T_δ^+ can be written as (by using Equation (1))

$$\begin{aligned} G_{T_\delta^+}(z) &= E[z^{T_\delta^+}] = E\left[z^{\sum_{i=1}^N X_i}\right] = E\left[E_+\left[z^{\sum_{i=1}^N X_i} | N\right]\right] \\ &= E\left[\left(G_X^+(z)\right)^N\right] \\ &= G_N(G_X^+(z)), \end{aligned} \quad (8)$$

where $G_N(z)$ is the pgf of random variable N .

Since N has geometric distribution with parameter $P_+(X = \delta)$, therefore, the pgf of N is obtained as (use also (3))

$$G_N(z) = \frac{z (\alpha + (1 - \alpha)P_{reg}(X = \delta))}{1 - (1 - (\alpha + (1 - \alpha)P_{reg}(X = \delta)))z}. \quad (9)$$

Hence, by using Equation (9) in Equation (8), we obtain

$$G_{T_\delta^+}(z) = \frac{G_X^+(z) (\alpha + (1 - \alpha)P_{reg}(X = \delta))}{1 - (1 - (\alpha + (1 - \alpha)P_{reg}(X = \delta)))G_X^+(z)}. \quad (10)$$

Consequently, by applying Theorem 5 to Equation (10) we get the desired result. This completes the proof. ■

In the next theorem, we obtain an explicit formula for the mean lifetime of the system, which defines the system's mean time to failure. The second moment is also provided.

Theorem 7. Let X_1, X_2, \dots be intershock times in the model described in Section 2, and let also X denotes a generic random variable of X_i 's. The mean lifetime of the system in the critical situation is

$$E[T_\delta^+] = \frac{\alpha\delta + (1 - \alpha)E_{reg}[X]}{\alpha + (1 - \alpha)P_{reg}(X = \delta)}, \quad (11)$$

and the second moment is

$$\begin{aligned} E[T_\delta^{+2}] &= \left(\frac{(1 - \alpha)P_{reg}(X \neq \delta)}{\alpha + (1 - \alpha)P_{reg}(X = \delta)}\right) (1 - \alpha) (E_{reg}[X^2] - \delta^2 P_{reg}(X = \delta)) \\ &+ \delta^2 (\alpha + (1 - \alpha)P_{reg}(X = \delta)) \\ &+ \frac{2((1 - \alpha)P_{reg}(X \neq \delta))^2}{(\alpha + (1 - \alpha)P_{reg}(X = \delta))^2} ((1 - \alpha) (E_{reg}[X] - \delta P_{reg}(X = \delta)))^2 \\ &+ 2\left(\frac{(1 - \alpha)P_{reg}(X \neq \delta)}{\alpha + (1 - \alpha)P_{reg}(X = \delta)}\right) ((1 - \alpha) (E_{reg}[X] - \delta P_{reg}(X = \delta))) \\ &\times \delta (\alpha + (1 - \alpha)P_{reg}(X = \delta)). \end{aligned} \quad (12)$$

Proof. Since N is a stopping time for X_i 's ($i = 1, 2, \dots$), so the mean lifetime of the system can be computed as

$$E[T_\delta^+] = E\left[E\left[\sum_{i=1}^N X_i \mid N\right]\right] = \sum_{n=1}^{\infty} E_+\left[\sum_{i=1}^n X_i\right] P(N = n) = E_+[X]E[N]. \quad (13)$$

Since N follows the geometric distribution with parameter $P_+(X = \delta)$, therefore (use also (3))

$$E[N] = \frac{1}{P_+(X = \delta)} = \frac{1}{\alpha + (1 - \alpha)P_{reg}(X = \delta)}, \quad (14)$$

and by using Equation (14) in Equation (13), we obtain

$$E[T_\delta^+] = \frac{E_+[X]}{\alpha + (1 - \alpha)P_{reg}(X = \delta)}. \quad (15)$$

Using Theorem 4 we have $E_+[X] = \alpha\delta + (1 - \alpha)E_{reg}[X]$, and by putting this in Equation (15), we get to Equation (11).

For the second moment, by conditioning on N , we can write

$$E[T_\delta^{+2}] = \sum_{n=1}^{\infty} E_+ \left(\left(\sum_{i=1}^n X_i \right)^2 \mid N = n \right) P_+(N = n). \quad (16)$$

A simple calculation show that

$$E_+ \left(\left(\sum_{i=1}^n X_i \right)^2 \mid N = n \right) = \sum_{i=1}^n E_+[X_i^2 \mid N = n] + 2 \sum_{1 \leq i < j \leq n} E_+[X_i X_j \mid N = n]. \quad (17)$$

From the definition of random variable N (see (2)), we have

$$E_+[X_i^2 \mid N = n] = \begin{cases} E_+[X^2 \mid X \neq \delta], & i = 1, 2, \dots, n-1, \\ E_+[X^2 \mid X = \delta], & i = n, \end{cases}$$

thus,

$$\sum_{n=1}^{\infty} E_+[X_i^2 \mid N = n] = (n-1)E_+[X^2 \mid X \neq \delta] + E_+[X^2 \mid X = \delta]. \quad (18)$$

Similarly,

$$\sum_{1 \leq i < j \leq n} E_+[X_i X_j \mid N = n] = \binom{n-1}{2} (E_+[X \mid X \neq \delta])^2 + (n-1)E_+[X \mid X \neq \delta]E_+[X \mid X = \delta]. \quad (19)$$

By using Equations (19) and (18) in (17) and then via Equation (16), we obtain

$$\begin{aligned} E[T_\delta^{+2}] &= (E[N] - 1)E_+[X^2 \mid X \neq \delta] + E_+[X^2 \mid X = \delta] \\ &+ E[(N-1)(N-2)](E_+[X \mid X \neq \delta])^2 \\ &+ 2(E[N] - 1) \left(E_+[X \mid X \neq \delta]E_+[X \mid X = \delta] \right). \end{aligned}$$

Since $E[N] = \frac{1}{P_+(X = \delta)}$ and $E[(N-1)(N-2)] = \frac{2(P_+(X = \delta) - 1)^2}{(P_+(X = \delta))^2}$, therefore,

$$\begin{aligned} E[T_\delta^{+2}] &= \left(\frac{P_+(X \neq \delta)}{P_+(X = \delta)} \right) E_+[X^2 \mid X \neq \delta] + E_+[X^2 \mid X = \delta] \\ &+ \frac{2(P_+(X \neq \delta))^2}{(P_+(X = \delta))^2} (E_+[X \mid X \neq \delta])^2 \\ &+ 2 \left(\frac{P_+(X \neq \delta)}{P_+(X = \delta)} \right) E_+[X \mid X \neq \delta]E_+[X \mid X = \delta]. \end{aligned} \quad (20)$$

On the other hand, using Equation (3) and Theorem 4, we have

$$P_+(X = \delta) = \alpha + (1 - \alpha)P_{reg}(X = \delta), \tag{21}$$

$$P_+(X \neq \delta) = 1 - P_+(X = \delta) = (1 - \alpha)P_{reg}(X \neq \delta), \tag{22}$$

$$E_+[X|X = \delta] = \delta (\alpha + (1 - \alpha)P_{reg}(X = \delta)), \tag{23}$$

$$E_+[X|X \neq \delta] = E_+[X] - E_+[X|X = \delta] = (1 - \alpha) (E_{reg}[X] - \delta P_{reg}(X = \delta)), \tag{24}$$

$$E_+[X^2|X = \delta] = \delta^2 (\alpha + (1 - \alpha)P_{reg}(X = \delta)), \tag{25}$$

$$E_+[X^2|X \neq \delta] = E_+[X^2] - E_+[X^2|X = \delta] = (1 - \alpha) (E_{reg}[X^2] - \delta^2 P_{reg}(X = \delta)). \tag{26}$$

The desired result (12) is obtained by putting Equations (21)–(26) in Equation (20). The proof is complete. ■

Next, an explicit expression for the variance of lifetime of the system is given for the defined shock model.

Theorem 8. Let X_1, X_2, \dots be the intershock times in the model described in Section 2, and let also X denotes a generic random variable of X_i 's. When the system is in the critical situation, the variance of the system's lifetime is

$$\begin{aligned} Var(T_\delta^+) &= \frac{1}{\alpha + (1 - \alpha)P_{reg}(X = \delta)} \left(Var_{reg}(X) + \alpha^2 (2\delta E_{reg}[X] - E_{reg}^2[X] - \delta^2) \right. \\ &+ \left. \alpha (2E_{reg}^2[X] - E_{reg}[X^2] - 2\delta E_{reg}[X] + \delta^2) \right) \\ &+ \frac{(\alpha\delta + (1 - \alpha)E_{reg}[X]) ((1 - \alpha)P_{reg}(X \neq \delta))}{(\alpha + (1 - \alpha)P_{reg}(X = \delta))^2}. \end{aligned}$$

Proof. By using the second Wald's identity (see page 30 from [11]), the variance of the system's lifetime can be written as

$$Var(T_\delta^+) = Var\left(\sum_{i=1}^N X_i\right) = Var_+(X)E[N] + (E_+(X))^2 Var(N). \tag{27}$$

Since N has geometric distribution with parameter $P_+(X = \delta)$, therefore

$$Var(N) = \frac{1 - P_+(X = \delta)}{(P_+(X = \delta))^2} = \frac{P_+(X \neq \delta)}{(P_+(X = \delta))^2}. \tag{28}$$

By using Equations (28) and (14) (for $E[N]$) in Equation (27), we obtain

$$Var(T_\delta^+) = \frac{Var_+(X)}{P_+(X = \delta)} + \frac{(E_+[X])^2 P_+(X \neq \delta)}{(P_+(X = \delta))^2}. \tag{29}$$

From Equation (3) we have $P_+(X = \delta) = \alpha + (1 - \alpha)P_{reg}(X = \delta)$, and thus, $P_+(X \neq \delta) = (1 - \alpha)P_{reg}(X \neq \delta)$. Besides, by Theorem 4 and Corollary 8 we have, respectively,

$$\begin{aligned} E_+[X] &= \alpha\delta + (1 - \alpha)E_{reg}[X], \\ Var_+(X) &= Var_{reg}(X) + \alpha^2 (2\delta E_{reg}[X] - E_{reg}^2[X] - \delta^2) + \alpha (2E_{reg}^2[X] - E_{reg}[X^2] - 2\delta E_{reg}[X] + \delta^2). \end{aligned}$$

Finally, by putting these identities in Equation (29) we obtain the desired result. The theorem is proved. ■

The Laplace transform (or the Laplace–Stieltjes transform) of T_δ^+ is derived in the following theorem.

Theorem 9. Let X_1, X_2, \dots be the intershock times in the model described in Section 2, and let also X denotes a generic random variable of X_i 's. Assuming that the system is in the critical situation, the Laplace transform of the system's lifetime is

$$\mathcal{L}_{T_\delta^+}(s) = \frac{e^{-s\delta}(\alpha + (1 - \alpha)P_{reg}(X = \delta))}{1 - \sum_{x \neq \delta} [e^{-sx}(\alpha + (1 - \alpha)P_{reg}(X = x))]}.$$

Proof. By properties of conditional expectation and using the fact that the random variable N is independent of X_i 's ($i = 1, 2, \dots$), we have

$$\begin{aligned} \mathcal{L}_{T_\delta^+}(s) &= E[e^{-sT}] = E[E[e^{-s\sum_{i=1}^N X_i} | N]] = \sum_{n=1}^{\infty} E[e^{-s\sum_{i=1}^n X_i} | N = n] P(N = n) \\ &= \sum_{n=1}^{\infty} E_+ [e^{-s\sum_{i=1}^n X_i} I_{(N=n)}]. \end{aligned}$$

Using the definition of N (see Equation (2)) and the fact that X is δ -inflated distributed, we have

$$\mathcal{L}_X(s) = \sum_{n=1}^{\infty} E_+ [e^{-s\sum_{i=1}^n X_i} I_{(X_1 \neq \delta, X_2 \neq \delta, \dots, X_{n-1} \neq \delta, X_n = \delta)}] = \sum_{n=1}^{\infty} (E_+ [e^{-sX} I_{(X \neq \delta)}])^{n-1} E_+ [e^{-sX} I_{(X = \delta)}].$$

Hence (by geometric series),

$$\mathcal{L}_{T_\delta^+}(s) = \frac{E_+ [e^{-sX} I_{(X = \delta)}]}{1 - E_+ [e^{-sX} I_{(X \neq \delta)}]}, \tag{30}$$

that is

$$\mathcal{L}_{T_\delta^+}(s) = \frac{e^{-s\delta} P_+(X = \delta)}{1 - (E_+ [e^{-sX}] - e^{-s\delta} P_+(X = \delta))} = \frac{e^{-s\delta} (\alpha + (1 - \alpha)P_{reg}(X = \delta))}{1 - \sum_{x \neq \delta} [e^{-sx} (\alpha + (1 - \alpha)P_{reg}(X = x))]}.$$

This completes the proof. ■

Remark 1. Note that if we consider $\alpha = 0$ in the above theorems, the distribution of intershock times reduces to a regular distribution, consequently, the above results investigate the system's lifetime in its regular situation.

5. ILLUSTRATIVE EXAMPLE

Transistors are one of the crucial components that are almost always present in any electronic device. It is common knowledge that transistors tend to heat up during operation and that this temperature increase can significantly affect their performance and dependability. Understanding the thermal challenges that transistors confront is crucial for engineers and designers. One of the reasons for the heat increase of transistors is high voltage, and high voltages usually damage the transistor as shocks. Transistors are usually cooled by using cooling fans (often with an passive heat exchanger such as a heat sink) so that they don't get damaged. For more information, we refer the reader to [22] and [2].

Here, we consider a specific transistor as a system which is exposed to a sequence of external shocks in the form of high voltages. The transistor must operate within its specified voltage and current limits to prevent overheating and shocks occur randomly in any period of time $x = 1, 2, \dots$, the values of x 's are as minutes. If the interarrival time between two successive high voltages is more than 5 minutes, the transistor has enough time to cool itself. If the interarrival time between two successive high voltages is less than 5 minutes, the cooling fan will automatically turn on to cool the transistor. In the case when the interarrival time between two

successive high voltages is equal to 5 minutes, the cooling fan will not turn on due to incorrect detection at the time length 5 minutes, so the transistor will be damaged. Observations show that the number of high voltages with interarrival times greater than 5 minutes and less than 5 minutes are approximately equal.

Indeed, the above example follows the shock model described in Section 2. Now, we present some illustrative computational results. Let us consider the case when the intershock times X_1, X_2, \dots are i.i.d. distributed by the geometric distribution with mean $\frac{1}{p}$, that is, the pmf of X_i ($i = 1, 2, \dots$) is $P(X_i = x_i) = p(1 - p)^{x_i - 1}$ for $x_i = 1, 2, \dots$. Assuming that the critical intershock time $x = \delta = 5$ is the median of X_i 's (based on observations), some different values can be chosen for the parameter p . The median of a geometric distribution with parameter p is given by

$$\text{Median} = \left\lceil \frac{-1}{\log_2(1 - p)} \right\rceil.$$

By solving the equation $\left\lceil \frac{-1}{\log_2(1 - p)} \right\rceil = 5$, we find that $p \in [\frac{2 - 2^{4/5}}{2}, \frac{2 - 2^{3/4}}{2})$. Thus, some different choices for the parameter p can be $p = 0.135$, $p = 0.145$, and $p = 0.155$. Assuming that $T_\delta = \sum_{i=1}^N X_i$ is the lifetime of the above system, in Table 1 we present the pmf $P(T_\delta = t)$ and the reliability function $P(T_\delta > t)$ for $p = 0.135, 0.145, 0.155$, $\delta = 5$, and some values of t . Also, we have considered the lifetime of the system in its critical situation, that is, $T_\delta^+ = \sum_{i=1}^N X_i$, in which the distribution of X_i 's is inflated at the critical time point $x = \delta = 5$.

From Table 1 we can see that the probability of system's lifetime in critical situation ($P(T_\delta^+ = t)$) at the critical time point $t = \delta = 5$ is greater than the probability of system's lifetime at other time points, and also the reliability function ($P(T_\delta^+ > t)$) at the critical point $t = \delta = 5$ is smaller than the reliability function at other time points. In general, the probability $P(T_\delta^+ = \delta)$ increases when the inflation parameter α increases. Also, as expected, an increase in inflation parameter α leads to a decrease in reliability function. By comparing $P(T_\delta > t)$ and $P(T_\delta^+ > t)$, it is clear that the system is more stable in the regular situation. This is consistent with what was expected from the theory, results, and purpose of the paper.

Table 1: The pmf and reliability function of the system's lifetime

p	δ	t	$P(T_\delta = t)$	$P(T_\delta > t)$	α	$P(T_\delta^+ = t)$	$P(T_\delta^+ > t)$
0.135	5	1	0.00571	0.99429	0.3	0.0333	0.9667
		2	0.00567	0.98869		0.0308	0.9359
		3	0.00564	0.98289		0.0286	0.9073
		4	0.00561	0.97737		0.0264	0.8809
		5	0.00558	0.97179		0.1304	0.7505
		6	0.00555	0.96624		0.0356	0.7149
		7	0.00551	0.96073		0.0334	0.6815
0.145	5	1	0.00600	0.99400	0.6	0.0365	0.9635
		2	0.00596	0.98804		0.0320	0.9315
		3	0.00593	0.98211		0.0281	0.9034
		4	0.00589	0.97622		0.0246	0.8788
		5	0.00586	0.97036		0.4001	0.4787
		6	0.00582	0.96454		0.0351	0.4436
		7	0.00579	0.95875		0.0309	0.4127
0.155	5	1	0.00624	0.99376	0.9	0.0140	0.9860
		2	0.00620	0.98756		0.0119	0.9741
		3	0.00616	0.98140		0.0100	0.9641
		4	0.00612	0.97528		0.0085	0.9556
		5	0.00609	0.96919		0.8243	0.1313
		6	0.00605	0.96314		0.0084	0.1229
		7	0.00601	0.95713		0.0071	0.1158

The following figure shows the mean lifetime of the above system versus p for critical time point $t = \delta = 5$ and some different values of the inflation parameter α . For $\alpha = 0$, we have the mean lifetime of the system in regular situation ($E[T_\delta]$). Figure 2 shows that the mean lifetime of the system in the regular state is higher than the mean lifetime of the system in the critical situation ($E[T_\delta^+]$). Also, we see that as the inflation parameter increases, the mean lifetime of the system decreases.

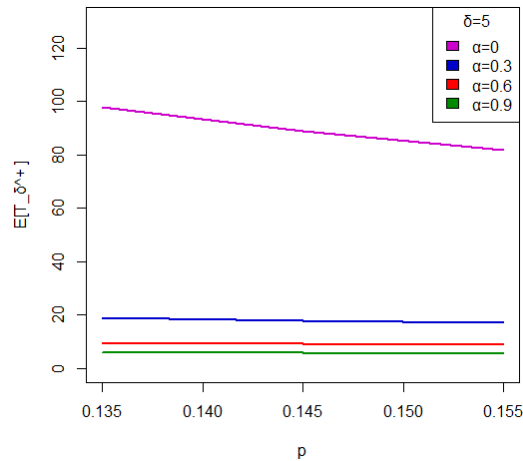


Figure 2: Plot of $E[T_\delta^+]$ versus p for $\delta = 5$ and some different values of inflation parameter α .

REFERENCES

- [1] Anderson, K.K., (1988). A note on cumulative shock models. *J. Appl. Probab*, **25**, 220–223.
- [2] Balasubbareddy, M., Sivasankaran, K., Atamuratov, A. E. and Khalilloev, M. M. (2023). Optimization of vertically stacked nanosheet FET immune to self-heating. *Micro and Nanostructures*, **182**, Article 207633.
- [3] Eryilmaz, S. (2013). On the lifetime behavior of a discrete time shock model. *Journal of Computational and Applied Mathematics*, **237**, 384–388.
- [4] Eryilmaz, S. (2012). Generalized δ -shock model via runs. *Statistics and Probability Letters*, **82**, 326–331.
- [5] Feng, C., Wang, H., Han, Y., Xia, Y, Lu, N. and Tu, X.M. (2015). Some Theoretical Comparisons of Negative Binomial and Zero-Inflated Poisson Distributions, *Commun Stat Theory Methods*. **44**, 3266–3277.
- [6] Goyal, D., Hazra, N.K., and Finkelstein, M. (2023). A general class of shock models with dependent inter-arrival times. *TEST*, **32**, 1079–1105.
- [7] Graham, R., Knuth, D. and Patashnik, O. (1994). *Concrete mathematics: a foundation for computer science*. New York, Addison-Wesley Publishing Company.
- [8] Gut, A. (1990). Cumulative shock models. *Advances in Applied Probability*, **22**, 504–507.
- [9] Gut, A. (2001). Mixed shock models. *Bernoulli*, **7**, 541–555.
- [10] Jornsatian, C., W. Bodhisuwan. (2022). Zero-one inflated negative binomial-beta exponential distribution for count data with many zeros and ones. *Commun Stat Theory Methods*. **51**, 8517–8531.
- [11] Janssen, J. and ?R. Manca. (2007). *Semi-Markov Risk Models for Finance, Insurance and Reliability*. New York, Springer.
- [12] Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, **34** (1):1–14.
- [13] Li, Z. H., Chan, L.Y. and Yuan, Z. X. (1999). Failure time distribution under a δ -shock model and its application to economic design of system. *International Journal of Reliability, Quality and Safety Engineering*, **6**, 237–247.

- [14] Li, Z.H., Huang, B.S. and Wang, G.J. (1999). Life distribution and its properties of shock models under random shocks. *Journal of Lanzhou University*, **35**, 1–7.
- [15] Li, Z.H. and Kong, X.B. (2007). Life behavior of δ -shock model. *Statistics and Probability Letters*, **77**, 577–587.
- [16] Li, Z.H. and Zhao, P. (2007). Reliability analysis on the δ -shock model of complex systems. *IEEE Transactions on Reliability*, **56**, 340–348.
- [17] Ma, M. and Li, Z. (2010). Life behavior of censored δ -shock model. *Indian J Pure Appl Math*, **41**, 401–420.
- [18] Mallor, F. and Omey E. (2001). Shocks, runs and random sums. *Journal of Applied Probability*, **38**, 438–448.
- [19] Parvardeh, A. and Balakrishnan, N. (2015). On mixed δ -shock models. *Statist. Probab. Lett*, **102**, 51–60.
- [20] Poursaeed, M. H. (2019). A generalization of δ -shock model. *Journal of Advanced in Mathematical Modeling*, **9**, 58–70.
- [21] Rivas, L. and Compos, F. (2023). Zero inflated Waring distribution, *Commun Stat Simulation Computation*, 52 (8): 3676–3691.
- [22] Rumyantsev, S.V., Novoselov, A.S. and Masalsky, N.V. (2022). Study of the Effect of Self-Heating in High-Voltage SOI Transistors with a Large Drift Region. *Russ Microelectron*, **51**, 325–333.
- [23] Sumita, U. and Shanthikumar, J.G. (1985). A class of correlated cumulative shock models. *Advances in Applied Probability*, **17**, 347–366.
- [24] Tuncel, A. and Eryilmaz, S. (2018): System reliability under δ -shock model, *Commun Stat Theory Methods*, **47**, 4872–4880.
- [25] Wang, G.J., and Zhang, Y.L. (2001). δ -shock model and its optimal replacement policy. *Journal of Southeast University*, **31**, 121–124.
- [26] Wang, G.J. and Zhang, Y.L. (2007). A shock model with two-type failures and optimal replacement policy. *International Journal of Systems Science*, **36**, 209–214.
- [27] Xu, Z.Y. and Li, Z.H. (2004). Statistical inference on d shock model with censored date. *Chinese J. Appl. Probab. Statist.* **20**, 147–153.

ENHANCED METHODS UNDER EXPONENTIAL DISTRIBUTION CONCERN WITH EWMA AND DEWMA METHODS

DR. V. KAVIYARASU AND M. INBARASI



- (1).Associate Professor, Department of Statistics, Bharathiar University
(2).Research Scholar, Department of Statistics, Bharathiar University
kaviyarasu@buc.edu.in, inbarasi.statistics@buc.edu.in

Abstract

Statistical Process Control plays a crucial part in improving the quality and lowering the fluctuation in the production process environment. In SPC, the most popularly used methods are Shewhart control chart techniques and EWMA techniques which distinguish itself for its quick identification of minute process deviations, which makes it an essential tool for guaranteeing product. EWMA methods detect variances in quality of the product as well as services, measure process mean shifts with control charts, and track manufacturing process parameters to find deviations and make necessary adjustments. The exponential distribution was employed in this study because it may reflect vast and bulk production in everyday life. Exponentially distributed data, evaluate it alongside the EWMA function. This paper's objective is to study the impact of EWMA & DEWMA parameters within the EWMA control chart's performance using exponential distribution. Further, A few tables are provided with suitable illustrations that can be available with parameters with the help of these findings. The study also examines how the EWMA parameter affects the shape of the distribution.

Keywords: Average run length, Exponential distribution, EWMA control Chart, Statistical Process Control.

1. INTRODUCTION

Walter A. Shewhart created Statistical process control in the (SPC) early 1920s at Bell Laboratories. To keep an eye on industrial activities, SPC uses technology that evaluates and regulates quality. SPC is a technique that is commonly used to discover production-line flaws and verify that the final product meets defined quality requirements. SPC is commonly used in manufacturing or production processes to assess how reliably a product functions under its design specification parameters. SPC helps to improve product quality, eliminate process variation, maintain regulatory and customer requirements, and reduce scrap, waste, defects, and reworks. Shewhart pioneered the control chart and the theory of a statistical control condition in 1920. A control chart can be used to observe how a process develops over time. Control charts are utilized to ensure quality regularly. The basic limitation of the memory-less charting system is its inability to track minute changes in process parameters. The exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) are the two finest frequent types of memory-based control charts. They integrate historical and present data to discover minor and reduce changes in process parameters. The name Geometric weighted moving average was adapted to reflect the fact that EWMA charts are constructed using exponential smoothing, for additional coverage for the EWMA, see [6] and [10] developed the double exponential moving average (DEMA) to

reduce the temporal delay caused by traditional moving averages. Moving averages, on the other hand, are prominent for their long lag. The double exponential moving average (DEMA) reacts by producing a more efficient averaging approach.

2. EXPONENTIAL WEIGHTED MOVING AVERAGES (EWMA)

The EWMA chart is an instance of a control chart used in Statistical Quality Control (SQC) for tracking variables or attribute-type data through an entire period of outcome from the tracked enterprise or industrial process. In contrast to other control charts, the EWMA chart estimates the EWMA of all historical sample averages. Although the EWMA chart relies on the normal distribution, it is also fairly robust in considering the existence of non-normally distributed quality parameters. There is a chart shift which lets quality attributes that are more accurately reflected through the Poisson distribution. The graphic just monitors the process mean; assessing the process variability implies an additional approach. Unlike a conventional moving average, which allocates equal weight to all data points within a specific period, EWMA offers higher importance to newly acquired observations while dropping the weight for previous observations exponentially. EWMA differs from a simple moving average because every data point in the time interval is given the same significance. EWMA can be especially beneficial in finance and time series analysis. Compared to a normal moving average, it gives a better representation of the data. The responsiveness might be transformed by improving the λ value. Calculating the EWMA over a shorter time period (smaller in size λ) causes it to be more responsive to recent changes. It is also utilized in quality control and process improvement, as recent observations provide a better indication of a process's present status.

Enhanced the Shewhart control chart's power by integrating the EWMA statistic [14]. The exponentially weighted moving average concepts which is develops [8]. The properties and enhancements of EWMA was developed by [9]. The EWMA control charts are utilized by [4] to monitor an analytical procedure. The features of the exponential EWMA chart are developed with parameter estimates in [12]. Improving EWMA chart performance [1]. Using an exponential type estimator of mean, [13] functioned in a hybrid exponentially weighted moving average (HEWMA) control chart. A nonparametric HEWMA-p control chart is developed by [2] to adjust for variance in monitoring procedures. The maximum EWMA and DEWMA charts based on auxiliary information are compared with sampling intervals for process mean and variance in [7]. The EWMA charts are more sensitized to tiny fluctuations into the process mean, exceeding ordinary \bar{X} charts to detect tiny shifts. The main difference is the smoothing effect resulting from exponential weighting, which reduces the influence of outliers or random fluctuations and emphasizes the underlying trend in the process. What distinguishes EWMA is its dynamic mobility in adapting to process scenarios, particularly adapting promptly to fluctuations in the mean by prioritizing recent data. This versatility is crucial in dynamic corporate environments where procedures undergo continual alteration. The rapid detection of errors in EWMA charts is a game changer, allowing for prompt intervention and correction before processes fall out of control or create out-of-spec goods. EWMA charts excel at managing auto-correlated data, outperforming traditional control charts in accurately illustrating process states by understanding autocorrelation trends via their weighted averaging strategy. Continuous monitoring capabilities boost EWMA's value by allowing real-time assessment of process performance for long-term stability. The convenience of perception adds the last feather to its gap, with expanding patterns or departures from the midline indicating probable alterations into the process mean.

Before designing the EWMA control chart, a researcher must pick two parameters:

- The 1st parameter act as the amount of weight given to the most current rational subgroup mean. The criteria $0 < \lambda \leq 1$ could be met, although determining the "proper" value is subjective depending on specific events.
- The 2nd parameter L act as a multiple of a rational subgroup standard deviation and specifies the control limitations. During alignment with other control charts, L is often set at 3. However, for smaller amounts of λ , L may need to be reduced significantly.

Instead of directly charting rational subgroup averages, the EWMA chart estimates consecutive observations z_i by analyzing the rational subgroup average, using the running average of all past observations, z_{i-1} , using the carefully determined weight, λ . It states that the EWMA is

$$z_i = \lambda x_i + (1 - \lambda)z_{(i-1)}; i = 1, 2, 3, \dots, n. \quad (1)$$

In which X_i represents present measure value, λ represents the smoothing constant that controls the depth of the memory EWMA, λ must fulfill $0 < \lambda \leq 1$, z_i signifies the present EWMA represents a EWMA statistic observed in the past measurement. (Needed the first sample at $i=1$) is the procedure aim, therefore $z_0 = \mu_0$. In some cases the average of preliminary information is applied as the EWMA's initial value, therefore $z_0 = \bar{x}$. To show that the EWMA z_i is a weighted average of the early sample means, we may add $z_{(i-1)}$ to the right side of equation (1) to get

$$\begin{aligned} z_i &= \lambda x_i + (1 - \lambda)[1 - \lambda x_{(i-1)} + (1 - \lambda)z_{(i-2)}] \\ z_i &= \lambda x_i + \lambda(1 - \lambda)x_{(i-1)} + (1 - \lambda)^2 z_{(i-2)} \end{aligned}$$

Trying to displace reclusively for $z_{(i-j)}$, $j=1,2,3,\dots t$, we get

$$z_i = \lambda \sum_{j=0}^{i-1} (1 - \lambda)^j x_{(i-j)} + (1 - \lambda)^i z_0 \quad (2)$$

The weights $\lambda(1 - \lambda)^j$ diminishes geometrically with the age of the sample mean. Moreover, the weights add to unity, given that

$$\lambda \sum_{j=0}^{i-1} (1 - \lambda)^j = \lambda \frac{(1 - (1 - \lambda)^i)}{(1 - (1 - \lambda))} = 1 - (1 - \lambda)^i$$

When the observations x_i are random variables that are independent at variance (σ^2), the variance of z_i

$$\sigma_{z_i}^2 = \sigma^2 \left(\frac{\lambda}{2 - \lambda} \right) [1 - (1 - \lambda)^{2i}] \quad (3)$$

As a consequence, the EWMA control chart might be generated by arranging z_i against the sample number i (or time). The EWMA control chart's centerline and control limits are displayed below: Control chart for EWMA:

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)} [1 - (1 - \lambda)^{2i}]} \quad (4)$$

$$CL = \mu_0 \quad (5)$$

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)} [1 - (1 - \lambda)^{2i}]} \quad (6)$$

In equations (4) and (6), the factor L represents the breadth of the control limit and it explore the possibility of the values L and λ subsequently. Approach unity i as gets larger, the $(1 - \lambda)^{2i}$ gets very close to zero and the equation (4), (5) and (6) preformed as the value of $(1 - \lambda)^{2i}$ approaches 0, the equation is rearranged as follows:

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)}} \quad (7)$$

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2 - \lambda)}} \quad (8)$$

Nonetheless, for tiny values of i and highly advise applying specific control limits in equations (4) and (6). By doing this, the control chart's ability to identify an off-target process as soon as the EWMA is activated will be greatly increased.

3. DOUBLE EXPONENTIAL WEIGHTED MOVING AVERAGES (DEWMA)

In 1994 Patrick G. Mulloy introduced the DEMA, often known as the DEWMA scheme, as an extension of classic EWMA concepts. They demonstrated that the DEWMA scheme beats the Shewhart scheme in tiny to moderate changes and has similar qualities for anticipating variation into the process mean to the EWMA control scheme. The DEWMA, likewise referred to by the Holt-Winters exponential smoothing, is widely used in forecasting, particularly in scenarios that require adapting to changing trends. The DEWMA has evolved to accommodate many variations and is now widely used in time series forecasting. Its applications range from banking to demand estimation, handling inventory, and environmental monitoring, all of which require accurate projections based on past data for decision-making. The DEWMA's adaptability and historical performance make it an invaluable resource for analysts & practitioners looking for reliable forecasting tools in a variety of fields.

The DEWMA is an indicator of trend designed to decrease noise in price charts employed by technical traders. It also tries to eliminate the lag time inherent in classic moving averages. A DEWMA variation was presented by [16], who also showed that it worked better than the EWMA system in detecting small mean shifts. Several researchers concluded that the DEWMA scheme is superior to the old EWMA system. The examination of the DEWMA control chart is presented in [11]. Using repetitive sampling, [3] develops the new DEWMA control chart. A Comparative Analysis of the EWMA and DEWMA is developed by [5]. The New Neutrosophic Double and Triple Exponentially Weighted Moving Average Control Charts are developed by [15]. The DEWMA is useful in financial and time series research because it may capture trends and produce smoother forecasts than regular moving averages. It is especially beneficial when working with data that has non-constant volatility. The DEWMA emphasizes recent observations by giving them larger weights, while simultaneously taking into account the trend using a second smoothing parameter. This dual-weighting strategy increases its responsiveness to changes in the underlying data pattern.

4. EXPONENTIAL DISTRIBUTION

In the Poisson point process, the exponential distribution describes the probability distribution of the time among events. The exponential distribution is considered a version of the exponential distribution. Furthermore, the exponential distribution is the continuous equivalent of the geometric distribution. The exponential distribution was

$$f(x) = \lambda e^{-\lambda x}, x \geq 0 \quad (9)$$

Here $\lambda > 0$ is a constant.

$$F(a) = \int_0^a \lambda e^{-\lambda x}, a \geq 0 \quad (10)$$

The exponential distribution is a commonly used time-to-failure model in reliability engineering. An exponential distribution is a continuous probability distribution that is frequently used in statistics and probability to represent the amount of time that will pass before a particular event occurs. Events happen continuously, independently, and at a set average pace during this process. One important property of the exponential distribution is that it requires less memory. It is possible for the exponential random variable to have fewer large values or more small ones. As a result, a customer's total grocery shop spending on a single visit follows an exponential curve.

The exponential distribution is a probability distribution that defines the time between occurrences in a process that happens at a constant pace regardless of how long it has been since the last event. It's frequently utilized in process control and reliability testing. The exponential distribution is frequently used in process control to describe the duration between successive occurrences or failures, such as system breakdowns, manufacturing problems, or client arrivals in a queue. A key feature of this distribution is its lack of memory, which means that the chance

of an event occurring in the following time interval remains constant regardless of the amount of time since the last occurrence. This Memorylessness property is especially useful in process management since it suggests a consistent chance of failure or event occurrence in the next instant, whatever the amount of time has passed since the previous event. The simulation settings are shown in Table 1.

Table 1: Simulation Setting

Simulation setting	Value	Simulation setting	Value
The size of the Sample	1000	Replication	100
Distribution	Exponential	Alpha level	0.95
Rate parameter (?)	0.2	Statistical Software (execution)	Excel
Confidence Level	0.05	Statistical software (validation)	R

The following Figure 1 depicts some shape of the exponential distribution

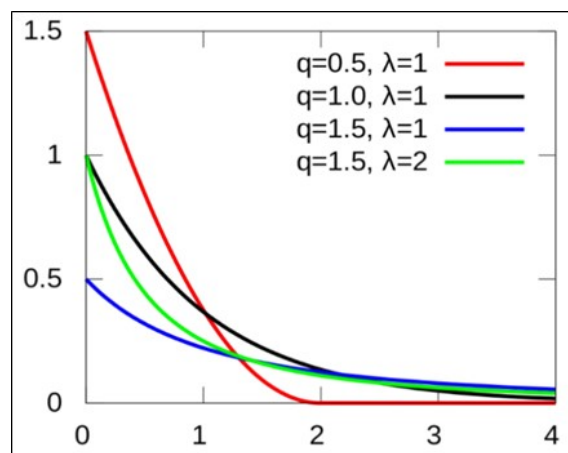


Figure 1: Simulation Setting

Figure 2 shows that the simulation has three steps

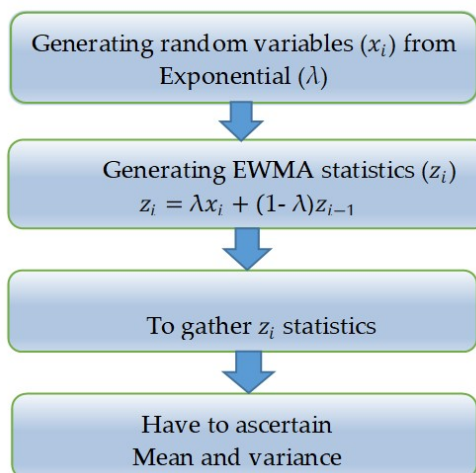


Figure 2: Simulation Steps

The following Figure 3 illustrates the Monte Carlo simulation using MS-Excel in detail

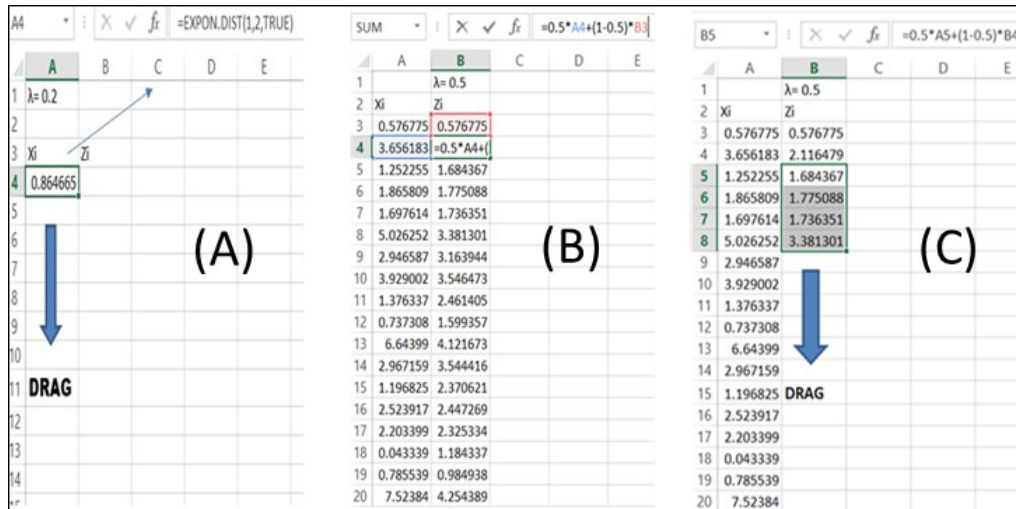


Figure 3: Monte Carlo simulation with MS-Excel in detail

5. EXPERIMENTAL APPROACH

In 1946 During World War II, John von Neumann and Stainlaw Ulam pioneered the Monte Carlo simulation to aid in decision-making in unpredictable situations. Given that randomness lies at the heart of modeling, comparable approaches, and roulette games, the name of the Monte Carlo comes from the famous city with casinos. All the three phases of the simulation are as follows, as seen in Figure 2. To create a sequence of EWMA statistic, a first set of random variates is used. When creating the EWMA statistic (i.e., z_1), a random number variates is pulled, and this variate has the similar value of the variates that come from the exponential distribution (i.e., x_1) z_2 is then calculated using z_1 and the second produced random variates, x_2 , as indicated by equation (1). Upto until the last random variates are formed, the procedure keeps going. This graphic is important since it explains the simulation stages and the mathematical portion of the EWMA estimations. Table 1 explains the simulated setup. This study applies Monte Carlo simulation using software. The exponential distribution was used to create random variates using software. The inverse value that is cumulative density of the exponential distribution was found using the Excel code "Exponential. INV (probability; λ)". As long as random probabilities are fed through the code, "Exponential. INV (probability; λ)" will produce random variates based on the given Exponential distribution. The probability distributed random probabilities must be uniform in order to provide non-disordered results because of the Nature of probabilities, that they're scattered over the field [0, 1]. Thus, these are evenly distributed randomly probabilities across the domain [0, 1] are produced using the Excel code `Rand ()`. And adding "`Rand ()`" in the place of "Probability," it is possible to archive the process of feeding randomized probabilities into the Exponential inverse code. "Exponential. INV (`Rand ()`; 1)" for instance, if $\lambda = 1$. Refer to Figure 5a. It is quite helpful to assign numbers in the first column between 1 and 1 million. The first cell at the top of that column is then used to produce a random variates from the Exponential distribution. The next step is to Double-click on the cell's owner the right corner or move the cell to produce one million variations in a column with two million cells containing random variates fitted to an exponential distribution is the end result (1). The next step after creating a column of random variables need to create EWMA statistics in the subsequent column. The first cell in this column is set to the exact same value (0.576775) as the top cell in the random variates columns, and this value only occurs once. Assuming λ to be 0.2, the EWMA equation is then applied to the second cell, as depicted in Figure.3 B, which is represented by cells x_2 (the second cell in this

column) and z_1 (the first cell in the column z_i). You can drag and click on this cell twice in the lower right corner. Doing a million EWMA statistical studies on EWMA data is the third and last step. The authors checked and validated the model. The mean and variance of the output variables were computed for each simulation run and compared to theoretical values.

$$Error\ fraction = \left| \frac{Theoretical\ value - Simulation\ value}{Theoretical\ value} \right| \tag{11}$$

In addition, the error fraction was calculated for each simulation run, as shown in equation (11). The error fraction for the variance and mean for all 150 simulation runs were less than 0.01. The error fraction was only 0.02. This validation of the random variable applied for the EWMA functions was excellent. Because of this, the EWMA equation is straightforward and this application has a lower error rate. Figure 4 illustrates how λ affects the structure of EWMA statistics with exponential distribution

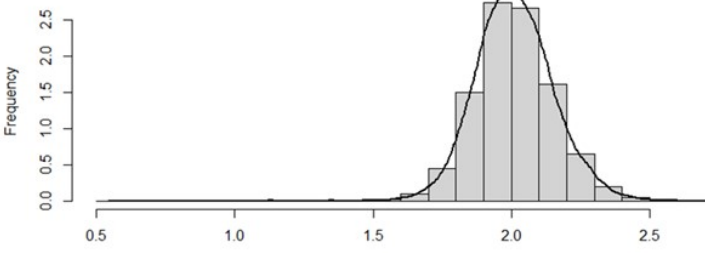
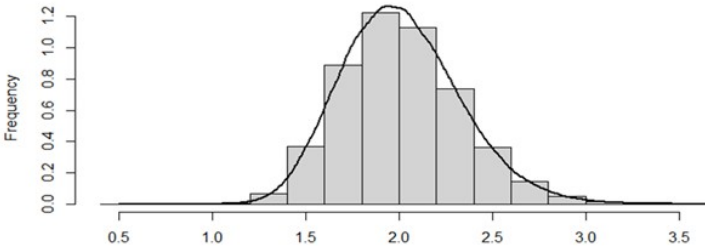
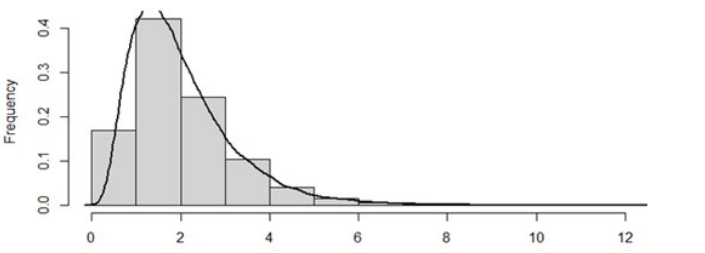
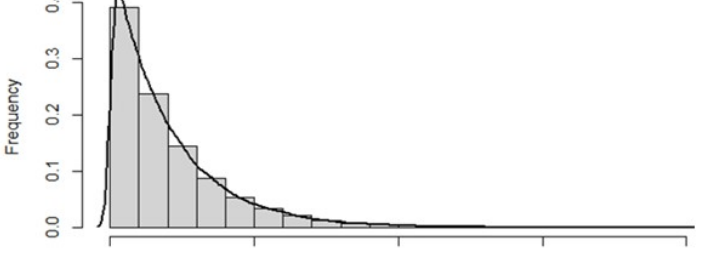
λ	Shape	Notes
0.01		Mean = 0.2814417 Standard deviation = 0.1614081 Which is very close to normal distribution
0.05		Mean = 2.998192 Standard deviation = 0.1615081
0.5		Mean = 2.997354 Standard deviation = 0.9961142
0.1		Mean = 2.997308 Standard deviation = 1.72673

Figure 4: The influence of λ in the structure of EWMA statistics with exponential distribution

6. CASE RESEARCH WITH SENSITIVE ANALYSIS

The store chosen for this inquiry is a prominent business around a crowded place metropolis of Maruthamalai, Coimbatore. The ABC store, located in the retail district serves a diverse customer base that includes residents as well as tourists. Recognized for its products and beneficial customer service, the company sees a continuous stream of customers throughout the day. A systematic data collection technique it was executed for this study's client arrival time data. Experienced investigators was discreetly stationed across the store for a few weeks to record customer's exact arrival times. Data collection entailed capturing customer arrival trends. Using real-time arrival data, the purpose is to investigate and understand how temporal dynamics flow is happening.

Understanding customer arrival patterns is critical for managing retail operations and providing better customer service. The case study focuses on the arrival time of customers, the busy capital city. Using empirical observation and data analysis, it was discovered that customer arrival times at this specific firm have an exponential distribution. The retail business at ABC Shop displays market dynamics shaped by city-specific cultural norms and consumer behavior. Customers' arrival timings at a store were investigated and documented throughout a certain time period. That was the case determined to the distribution of arrival timing follows established patterns that correspond to the exponential distribution, which is commonly used to monitor linear arrival times. The exponential distribution is an adaptable framework to describing the variability as well as mutual dependency of customer arrival time intervals, allowing for in-depth analysis and comprehension of the variability and interdependence of customer arrival times, as well as the identification of customer behavior and arrival processes at the retail shop. The first component of arrival times is established during the store's operating hours of 7:00 a.m. to 12:00 p.m. Table 2 displays the period between customer arrivals. The goal of the case study was to understanding the consumer behavior and patterns of arrival in an ABC retail atmosphere. In developing the adherence of client arrival times with the exponential distribution, retail managers and decision-makers with ABC retail shop gained significant insights. These findings have the potential to influence programs for workforce efficiency and queuing management enhancements, and standard customer service enhancement, all of which are based on a better knowledge of the elements that determine arrival times. The efficiency of EWMA and DEWMA may be assessed by measuring conduction sensitivity and varying λ values. Please see the Tables. 3 and 4. The case study's customer arrival time data showed significant changes and improvements after using the EWMA and DEWMA functions. The combination of EWMA and DEWMA resulted in smoothed data that provided a clearer picture of underlying patterns and behaviors associated to client arrival, enabling for a more in-depth study of the relevant dynamics involved. Furthermore, by assigning larger weight to recent data sets, the EWMA function allows for real-time study of arrival trends. Customer arrival time trends were quickly discovered, allowing for proactive monitoring and response. Overall, the EWMA and DEWMA functions improved the analysis and interpretation of arriving time data at ABC retail, resulting in better resource utilization, operational scheduling, and outstanding customer management. Table 2 displays the period between customer arrivals.

Table 2: Customers' arrival & inter-arrival times at the ABC shop

Arrival no	Arrival time	Inter-arrival time(min)	Arrival no	Arrival time	Inter-arrival time(min)
1	7.06 a.m.	6.2	16	10.04 a.m.	3.6
2	7.31 a.m.	25.2	17	10.17 a.m.	13.61
3	7.36 a.m.	4.98	18	10.23 a.m.	5.87
4	7.51 a.m.	14.55	19	10.26 a.m.	3.01
5	8.05 a.m.	14.3	20	10.39 a.m.	12.72
6	8.08 a.m.	2.99	21	10.54 a.m.	14.47
7	8.1 a.m.	2.39	22	10.55 a.m.	1.11
8	8.35 a.m.	24.9	23	10.56 a.m.	0.63
9	9.26 a.m.	51.39	24	10.57 a.m.	0.19
10	9.28 a.m.	1.74	25	11.12 a.m.	15.15
11	9.3 a.m.	2.09	26	11.16 a.m.	4.17
12	9.4 a.m.	9.57	27	11.19 a.m.	12.89
13	9.44 a.m.	3.98	28	11.31 a.m.	12.09
14	9.54 a.m.	10.39	29	11.55 a.m.	24.25
15	9.59 a.m.	4.59	30	11.58 a.m.	2.58

Table 3 illustrates the Inter-arrival time after using EWMA with varied λ

Table 3: *Inter-arrival time after using EWMA with varied λ*

λ	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.01
<i>Inter-arrival time</i>											
6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25	6.25
25.23	8.148	10.046	11.945	13.843	15.741	17.639	19.538	21.436	23.334	24.283	25.042
4.99	7.832	9.034	9.857	10.3	10.364	10.047	9.351	8.276	6.821	5.951	5.186
14.55	8.504	10.138	11.266	12.002	12.459	12.751	12.993	13.298	13.781	14.124	14.46
14.3	9.084	10.971	12.176	12.921	13.379	13.681	13.908	14.1	14.248	14.291	14.302
2.99	8.474	9.375	9.421	8.949	8.185	7.267	6.266	5.213	4.117	3.556	3.104
2.4	7.867	7.98	7.314	6.329	5.292	4.346	3.559	2.962	2.571	2.457	2.407
24.99	9.58	11.383	12.619	13.795	15.144	16.736	18.564	20.588	22.753	23.868	24.769
51.39	13.761	19.385	24.251	28.835	33.269	37.531	41.545	45.233	48.53	50.018	51.128
1.74	12.559	15.856	17.498	17.998	17.505	16.057	13.683	10.44	6.421	4.156	2.236
2.09	11.513	13.104	12.877	11.636	9.799	7.679	5.57	3.762	2.526	2.196	2.094
9.57	11.319	12.398	11.886	10.811	9.687	8.816	8.373	8.412	8.87	9.206	9.5
3.98	10.585	10.715	9.515	8.079	6.834	5.915	5.299	4.868	4.47	4.243	4.037
10.4	10.566	10.651	9.78	9.007	8.616	8.605	8.869	9.292	9.806	10.091	10.335
4.59	9.969	9.44	8.224	7.241	6.605	6.198	5.876	5.533	5.114	4.868	4.65
3.61	9.333	8.273	6.839	5.788	5.106	4.643	4.288	3.992	3.758	3.67	3.618
13.62	9.761	9.342	8.872	8.92	9.362	4.643	10.818	11.692	12.631	13.12	13.517
5.87	9.372	8.648	7.972	7.701	7.617	4.643	7.356	7.036	6.548	6.235	5.949
3.02	8.737	7.522	6.486	5.827	5.317	4.643	4.319	3.821	3.371	3.178	3.047
12.73	9.136	8.564	8.359	8.588	9.023	4.643	10.206	10.948	11.794	12.252	12.633
14.47	9.67	9.746	10.194	10.943	11.749	4.643	13.194	13.769	14.206	14.363	14.456
1.12	8.815	8.02	7.471	7.012	6.433	4.643	4.74	3.647	2.426	1.779	1.251
0.63	7.996	6.542	5.419	4.46	3.532	4.643	1.863	1.234	0.81	0.688	0.636
0.19	7.216	5.272	3.851	2.753	1.862	4.643	0.694	0.401	0.255	0.218	0.197
15.16	8.01	7.249	7.242	7.714	8.509	4.643	10.817	12.204	13.665	14.408	15.006
4.17	7.626	6.634	6.322	6.298	6.342	4.643	6.167	5.781	5.124	4.686	4.283
2.9	7.153	5.886	5.294	4.937	4.618	4.643	3.877	3.472	3.118	2.985	2.909
12.1	7.648	7.129	7.335	7.802	8.359	4.643	9.632	10.374	11.201	11.643	12.007
24.26	9.309	10.555	12.412	14.384	16.308	4.643	19.87	21.481	22.952	23.627	24.135
2.59	8.637	8.961	9.465	9.665	9.448	4.643	7.772	6.366	4.624	3.639	2.803

Table 4 illustrates the Inter-arrival time after using DEWMA with varied λ .

Table 4: *Inter-arrival time after using DEWMA with varied λ*

	0.9	0.8	0.7	0.6	0.5	0.4
	<i>Inter-arrival time</i>					
6.25	6.3	6.3	6.3	6.3	6.3	6.3
25.23	8.3	10.8	13.7	16.9	20.5	24.5
4.99	7.8	9	9.7	10.1	10	9.6
14.55	8.6	10.3	11.7	12.6	13.3	14.1
14.3	9.1	11.2	12.6	13.5	14.3	15
2.99	8.4	9.1	8.7	7.6	6	4.2
2.4	7.8	7.6	6.5	4.7	2.8	0.8
24.99	9.7	12	14	16.1	18.8	22
51.39	14.2	21.1	28.1	35.8	44.2	53.2
1.74	12.5	15.5	16.6	16.4	15.1	12.6
2.09	11.4	12.5	11.2	8.5	4.7	0.6
9.57	11.3	12.1	11.1	9.2	7.1	5.2
3.98	10.5	10.3	8.6	6.3	4.1	2
10.4	10.6	10.6	9.6	8.7	8.1	7.9
4.59	9.9	9.2	7.7	6.4	5.4	4.3
3.61	9.3	8	6.3	4.9	3.7	2.6
13.62	9.8	9.5	9.3	9.8	10.8	3.4
5.87	9.3	8.5	7.8	7.6	7.5	3.9
3.02	8.7	7.3	6	5	4.1	4.2
12.73	9.2	8.7	8.8	9.4	10.3	4.4
14.47	9.7	10	10.9	12.2	13.7	4.5
1.12	8.7	7.7	6.9	5.9	4.8	4.5
0.63	7.9	6.2	4.6	3	1.2	4.6
0.19	7.1	4.9	3.1	1.5	0.1	4.6
15.16	8.1	7.6	8	9.2	10.8	4.6
4.17	7.6	6.6	6.3	6.3	6.4	4.6
2.9	7.1	5.7	5	4.4	3.8	4.6
12.1	7.7	7.3	7.9	8.7	9.8	4.6
24.26	9.5	11.3	14.1	17.4	21	4.6
2.59	8.6	8.8	9.1	9	8.4	4.6

Figures 5 depict the histograms for inter-arrival time, after applying EWMA and DEWMA, respectively.

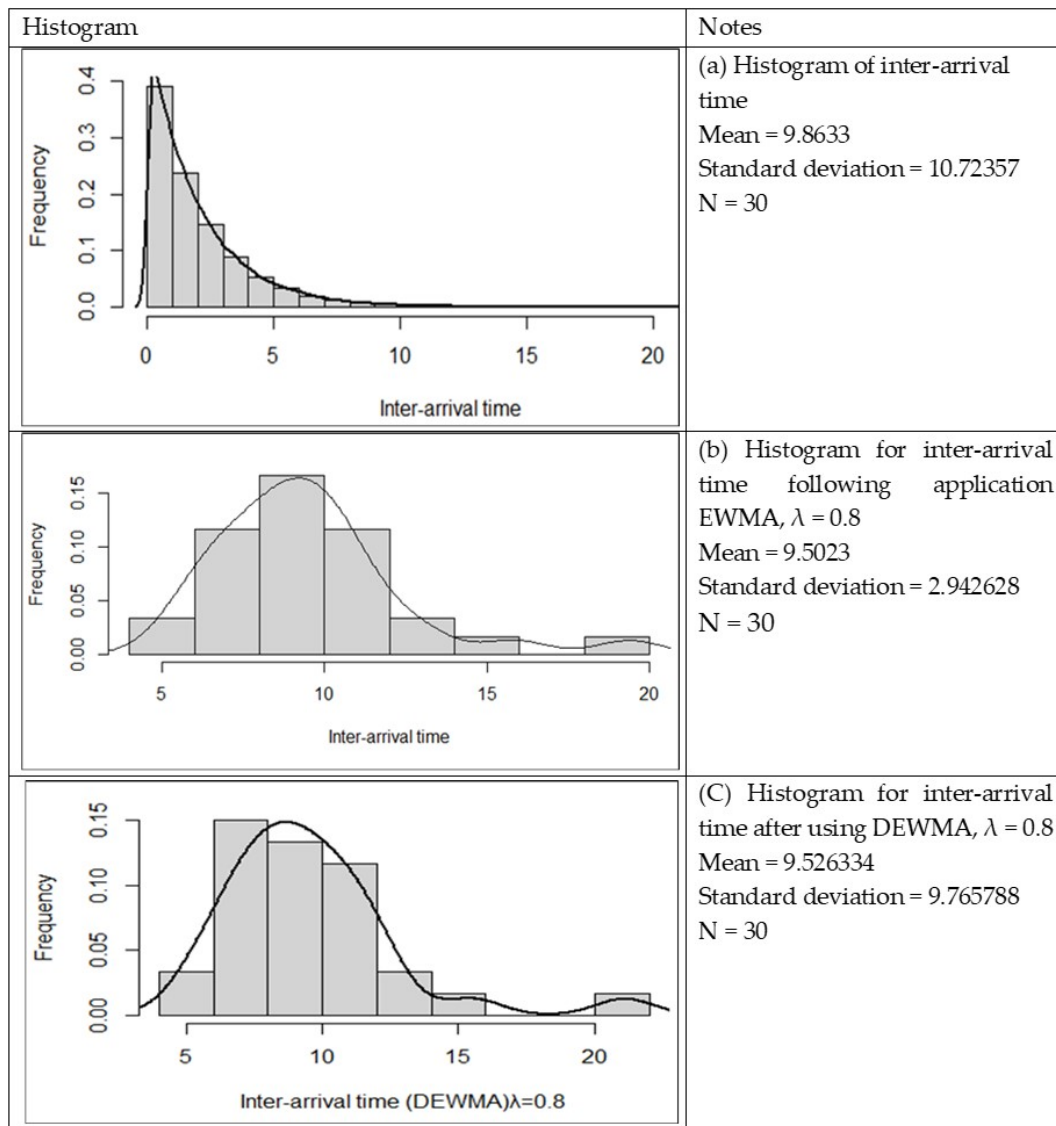


Figure 5: Histogram for inter-arrival time, after applying EWMA and DEWMA

7. CONCLUSION

Monte Carlo simulation methods were utilized to create exponential distribution data that was evaluated using EWMA control chart functions to determine the value. The overall goal of this study is to assess the impact of distribution parameters in the operation of the EWMA control chart. Furthermore, the investigation is predicated on many key assumptions. The study's findings give an essential new knowledge of how distribution characteristics influence the Effectiveness of the EWMA control charts. Overall, the EWMA and DEWMA functions enhanced the analysis and interpretation of customer's arrival times, leading to superior management and better use of resources. There are a few limitations to consider while interpreting the findings and applying them to real-world situations. More research is required to get overcome these limitations and improve understand the situation. Additional study is needed to solve these limits

and obtain an improved better grasp of the situation. Future research may include the creation of new statistical methodologies, namely for improved visualization methods and machine learning anomalies identification. In future research, this approach may be expanded to estimate the variance parameter and used to real data with different distributions.

REFERENCES

- [1] Abbas, N., Riaz, M., & Does, R. J. (2011). Enhancing the performance of EWMA charts. *Quality and Reliability Engineering International*, 27(6), 821-833.
- [2] Aslam, M., Rao, G. S., Al-Marshadi, A. H., & Jun, C. H. (2019). A Nonparametric HEWMA-p Control chart for variance in monitoring processes. *Symmetry*, 11(3), 356.
- [3] Adeoti, O. A. (2018). A new double exponentially weighted moving average control chart using repetitive sampling. *International Journal of Quality & Reliability Management*, 35(2), 387-404.
- [4] Carson, P. K., & Yeh, A. B. (2008). Exponentially weighted moving average (EWMA) control charts for monitoring an analytical process. *QIndustrial & engineering chemistry research*, 47(2), 405-411.
- [5] Chan, K. M., Chong, Z. L., Khoo, M. B. C., Khaw, K. W., & Teoh, W. L. (2021). A Comparative Study of the EWMA and Double EWMA Control Schemes. In *Journal of Physics: Conference Series*, Vol. 2051, No. 1, p. 012067.
- [6] Crowder, S.V. (1989). Design of Exponentially Weighted Moving Average Schemes. *Journal of Quality Technology*, 21, 155-162.
- [7] Haq, A., & Akhtar, S. (2022). Auxiliary information based maximum EWMA and DEWMA charts with variable sampling intervals for process mean and variance. *Communications in Statistics-Theory and Methods*, 51(12), 3985-4005.
- [8] Hunter, J. S. (1986). The exponentially weighted moving average. *Journal of quality technology*, 18(4), 203-210.
- [9] Lucas, J. M., & Saccucci, M. S. (1987). Exponentially weighted moving average control schemes: Properties and enhancements. *Drexel University, College of Business and Administration.*, 32(1), 1-12
- [10] Lucas, J.M. and Saccucci, M.S. (1990). Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics*, 32, 1-12.
- [11] Mahmoud, M. A., & Woodall, W. H. (2010). An evaluation of the double exponentially weighted moving average control chart. *Communications in Statistics—Simulation and Computation*, 39(5), 933-949.
- [12] Ozsan, G., Testik, M. C., & Wei?, C. H. (2010). Properties of the exponential EWMA chart with parameter estimation. *Quality and Reliability Engineering International*, 26(6), 555-569..
- [13] Raza, S. M. M., Sial, M. H., Haider, M., & Butt, M. M. (2019). Hybrid exponentially weighted moving average (HEWMA) control chart based on exponential type estimator of mean. *Journal of Reliability and Statistical Studies*, 187-198.
- [14] Roberts, S. W. (1959). Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, 1, 239-250.
- [15] Shafqat, A., Aslam, M., Saleem, M., & Abbas, Z. (2021). The New Neutrosophic Double and Triple Exponentially Weighted Moving Average Control Charts. *CMES-Computer Modeling in Engineering & Sciences*, 129(1).
- [16] Zhang, L., & Chen, G. (2005). An extended EWMA mean chart. *Quality Technology & Quantitative Management.*, 2(120), 39-52.

RELIABILITY SAMPLING PLAN FOR GENERALIZED INVERTED EXPONENTIALLY DISTRIBUTED UNDER PROGRESSIVE TYPE-II CENSORED DATA

S. SINGH¹, A. KAUSHIK²

•

^{1,2}Department of Statistics, Banaras Hindu University, Varanasi, India.

arundevkaushik@gmail.com

Corresponding Author Email: ¹statsshubham@bhu.ac.in

Abstract

This article aims to explore a sampling strategy designed to assess the reliability of products that exhibit lifetimes following a GIED. Considered sampling approach has been specially constructed for a Type-II progressive censoring scheme, which includes binomial removals as part of its methodology. Its core objectives is to find out acceptance constant and the optimum sample size. To facilitate practical implementation, the article presents a tabulated form of the sampling plan for the selected specification, as per the considered censoring scheme. To validate the dependability and precision of the suggested sampling approach, we perform a Monte Carlo experiment under various scenarios.

Keyword : Generalized Inverted Exponential Distribution (GIED); OC-Curve; Reliability Sampling Plan; Simulation; Progressive Censoring.

1. INTRODUCTION

In life testing and reliability studies, direct observation of the exact lifetime of a specific event of interest for all tested units is often impractical. This situation arises in various scenarios, such as clinical trials and in engineering where individuals may remain alive or disease-free beyond the study period. To streamline costs and time, some units may be randomly withdrawn from the experiment, resulting in censored data. It is essential to assess the effect of censoring on reliability and determine whether it provides meaningful information or not.

In the field of statistics, a variation of the exponential distribution, termed the one-parameter inverse exponential or inverted exponential distribution (IED), has been advanced. This distribution exhibits an inverted bathtub hazard rate. The utilization of IED in survival analysis has been advocated by several researchers, as exemplified by [22] and [23]. A two-parameter extension of the inverted exponential distribution (IED), called the generalized inverted exponential distribution (GIED), was proposed by [24] and demonstrated that GIED fits real datasets better than IED, based on K-S statistics and likelihood ratio tests. Furthermore, [25] conducted a study on reliability estimation based on progressive Type-II censored samples under the classical paradigm. A common way to evaluate the quality of a product is to check if it meets certain specifications related to its reliability and lifetime. Some standard sampling plans, such as MIL – STD – 414 and MIL – STD – 105 as discussed by [1], can be used to compare the results with predefined criteria. However, these plans may not be suitable for situations where observing all failures is too expensive or time-consuming, especially for products with high reliability. In such cases,

censored tests are often used. Censoring is one of the main feature of lifetime study or reliability study. Censoring desirably or undesirably occurred in the experiment. There are several type of censoring schemes discussed by [2] and [9]. Now a days practitioners and researchers have advocated for a versatile censoring scheme known as progressive Type-II censoring. Progressive Type-II censoring is a method of reliability sampling that involves removing a certain number of units that have not failed at each failure time. This method can reduce the cost and time of testing, but it also introduces some challenges in the analysis.

The progressive Type-II censoring scheme is an extension of Type-II censoring, which incorporates the removal of units from a life-test at predetermined or random inspection times. In this scheme, out of the initial total of (n) units simultaneously placed on a life test, only (m) units are fully observed, while the remaining $(n - m)$ units are withdrawn from the experiment at different time points. Some of the researchers who have developed reliability sampling plans with progressive Type-II censoring are [6], [8] and [7]. A reliability sampling plan by [8] focused on the exponential distribution, while [7] considered the Log-normal and Weibull distributions. Also, [10], [12], and [11] also studied the exponential, Weibull, and Log-normal distributions, respectively, but with different assumptions on the number of units removed at each failure. A comprehensive review of progressive Type-II censoring and its applications provided by [9]. Progressive Type-II censoring is a complex process that requires careful planning and analysis.

An example of the application of progressive censoring in evaluating the performance of electronic components provided by [13]. In such cases, certain test units may require removal due to factors like excessive heat, resulting in situations that fall under the purview of Type-II PCR. Furthermore, [16] conducted extensive investigations into issues related to parameter estimation and the expected duration of experiments under Type-II PCR censoring. In numerous practical scenarios, managing removals presents a formidable challenge, rendering the assumption of fixed and known removals impractical. In acknowledgment of this constraint, [3] advocates the adoption of random removals, emphasizing its practical viability. Therefore, the implementation of Type-II censoring with random removals becomes a more pragmatic choice. In this approach, a suitable distribution, such as the binomial distribution, can be employed to model the removal pattern. To the best of our knowledge, the utilization of reliability sampling plans for Type-II censoring with binomial removals has not been previously documented. In current study, our primary focus lies in the development of a reliability sampling plan for the GIED under Type-II progressive censoring with random removals. This entails that the number of removals at each failure is subject to a binomial distribution. In Section 2, we will introduce our proposed model and establish the maximum likelihood estimators (MLEs) for the model parameters. In Section 3, we present the Operating Characteristic (O.C.) curve, providing insights into the performance of our sampling plan. Section 4 delves into an in-depth examination of the sampling plan's design. Finally, in Section 5, we offer our concluding remarks and provide a succinct summary of the key findings derived from this study.

2. METHODS

A generalisation of the one parameter IED is a two parameter GIED having PDF and CDF as follows:

$$\begin{aligned} \xi(t) &= \frac{\nu\eta}{t^2} \exp\left(-\frac{\eta}{t}\right) \left(1 - \exp\left(-\frac{\eta}{t}\right)\right)^{\nu-1}; & t > 0, \nu > 0, \eta > 0. \\ \Xi(t) &= 1 - \left(1 - \exp\left(-\frac{\eta}{t}\right)\right)^\nu & ; & t > 0, \nu > 0, \eta > 0. \end{aligned} \tag{1}$$

Where, ν is the shape parameter and η is the scale parameter. Let t_p be the p^{th} percentile of the GIED, it is given by,

$$p = 1 - \left(1 - \exp\left(-\frac{\eta}{t_p}\right)\right)^\nu. \tag{2}$$

On simplification, we get

$$t_p = \frac{-\eta}{\ln(1 - (1 - p)^{1/\nu})}$$

and median of the distribution is given by

$$m_d = \frac{-\eta}{\ln(1 - (0.5)^{1/\nu})}$$

the reliability function is given by

$$S(t) = \left(1 - \exp\left(-\frac{\eta}{t}\right)\right)^\nu; \quad t \geq 0, (\eta, \nu) > 0.$$

The failure rate function of the GIED(η, ν) is given by

$$h(t) = \frac{f(t)}{S(t)} = \frac{\nu\eta}{t^2} \frac{e^{-\frac{\eta}{t}}}{1 - e^{-\frac{\eta}{t}}}; \quad t \geq 0, (\eta, \nu) > 0.$$

For simplicity point of view, let us make a transformation $Z = \ln(t)$.

$$\psi(z) = \nu\eta \exp(-z - \eta \exp(-z)) (1 - \exp(-\eta \exp(-\eta \exp(-z))))^{\nu-1} \quad (3)$$

$$; z > 0, \nu > 0, \eta > 0.$$

and its distribution function is given by

$$\Psi(z) = 1 - (1 - \exp(-\eta \exp(-z)))^\nu; z > 0, \nu > 0, \eta > 0. \quad (4)$$

Let's consider the following transformations: $\mu = \ln \nu$ and $\sigma = \frac{1}{\eta}$. It simplifies our analysis to work with the model represented by equation 3. Now, we have a set of m ordered log-failure times, denoted as $Z_1 < Z_2 < \dots < Z_m$, selected from a pool of n items. The value of m is pre-determined, indicating the number of failures that occur before the testing concludes.

At the i^{th} failure event, a random removal of r_i items takes place from the testing pool. The number of items removed, r_i , follows a binomial distribution characterized by parameters $(n - m)$ and removal probability (p_r). In the context of a Type II progressive censoring (Type II PCR), we define the likelihood function as follows:

$$L(t; \mu, \sigma) = L_1(t; \mu, \sigma) P_R.$$

where,

$$L_1(t; \mu, \sigma) = \prod_{i=1}^m \psi(z_i) (1 - \Psi(z_i))^{r_i}.$$

and

$$P_R = P(R_{m-1} = r_{m-1} | R_{m-2} = r_{m-2}, R_{m-3} = r_{m-3} \dots R_1 = r_1)$$

$$\times P(R_{m-2} = r_{m-2} | R_{m-3} = r_{m-3}, R_{m-4} = r_{m-4} \dots R_1 = r_1) \dots P(R_2 = r_2 | R_1 = r_1).$$

$$= \frac{(n - m)!}{\prod_{i=1}^m r_i! (n - m - \sum_{j=1}^{m-1} r_j)!} p_r^{\sum_{j=1}^{m-1} r_j} (1 - p_r)^{(m-1)(n-1) - \sum_{j=1}^{m-1} (m-j)r_j}.$$

Where, $C_i = n - \sum_{j=1}^{i-1} (r_j + 1)$. To obtain the maximum likelihood estimators (MLEs) of μ and σ , the likelihood have been maximized at MLEs of the parameters, for more details see [28] and [27]. In this type of censoring scheme, the number of failures is predetermined before the experiment begins. The experiment is terminated once the desired number of failures is observed. Assuming that the number of failures is fixed as m , we denote t_i as the time at the i^{th} removal, and r_i as the number of the random removals of the i^{th} component.

3. OC CURVE

The OC curve is a tool used to evaluate the effectiveness of a sampling plan. It does so by charting the likelihood of accepting a lot against the proportion of non-conforming items within that lot. This evaluation relies on the principles of asymptotic distribution theory.

In the context of this evaluation, we utilize the following equations:

First, we have:

$$\frac{T' - (\mu - k\sigma)}{\sqrt{AsVar[T']}} \sim N(0, 1)$$

The standardized variate is expressed as:

$$W = \frac{(T' - (\mu - k\sigma))\sqrt{n}}{\sqrt{V}}$$

To construct the OC curve, which represents the probability of accepting a lot, denoted as $L(p)$, we use the following equation:

$$L(p) = Pr[T' \geq L'] = 1 - \Phi \left[\frac{\sigma(u_p + k_1)\sqrt{n}}{\sqrt{V}} \right]$$

In this equation, u_p stands for the quantile of the standard logistic distribution that corresponds to the given proportion of non-conforming items, denoted as p . $\Phi(\cdot)$ represents the standard normal distribution function.

Therefore, to determine an optimal sampling plan for specific points on the OC curve, denoted as $(p_\alpha, 1 - \alpha)$ and (p_β, β) , the following equations need to be solved for the variables k and n :

$$\begin{aligned} z_\alpha - \frac{\sigma(u_{p_\alpha} + k_1)\sqrt{n}}{\sqrt{V}} &= 0 \\ z_{1-\beta} - \frac{\sigma(u_{p_\beta} + k_1)\sqrt{n}}{\sqrt{V}} &= 0 \end{aligned} \tag{5}$$

where, u_{p_α} and u_{p_β} denote the quantiles of the standard normal distribution and z_α and z_β denotes the quantiles of the log-life distribution. Thus on solving equation (5), we get

$$k = \frac{z_{p_\alpha}u_{1-\beta} - z_{p_\beta}u_\alpha}{u_\alpha - u_{1-\beta}} \tag{6}$$

and

$$n = \left(\frac{u_\alpha - u_{1-\beta}}{z_{p_\alpha} - z_{p_\beta}} \right)^2 \left(\frac{\sigma^2}{n} (\gamma_{11}(n, p_r, f_c) - 2k\gamma_{12}(n, p_r, f_c) + \gamma_{22}(n, p_r, f_c)) \right) \tag{7}$$

4. LAYOUT OF SAMPLING PLAN

4.1. Sampling Plan

In our study, we adopt the methodology originally proposed by [17] to evaluate the acceptability of a batch. Specifically, we concentrate on variable sampling plans with one-sided specification limits. Let's consider a lot of size n randomly drawn from a larger population. The log-lifetimes of the items in this lot follow a distribution characterized by Equation (3). This distribution is defined by a set of unknown parameters, denoted as ν and η . We seek to obtain the maximum likelihood estimators for these parameters, denoted as $\hat{\nu}$ and $\hat{\eta}$.

In this context, we have a lot with a proportion of non-conforming items, denoted as p_0 (where $p_0 \leq p_\alpha$), which is considered acceptable and should be approved with a probability of

at least $(1 - \alpha)$. Here, p_α represents the proportion of non-conforming items that corresponds to the desired probability of acceptance, denoted as $(1 - \alpha)$ on the operating characteristic (OC) curve. We define L' as the quantile of the $\Psi(\cdot)$ given in the equation (4) that corresponds to the proportion of non-conforming items for the chosen probability of acceptance $(1 - \alpha)$. This is calculated as $L' = \Psi^{-1}(p_\alpha)$. The decision to accept or reject the lot hinges on the comparison of the estimate $(\hat{\mu} - k\hat{\sigma})$ with the value of L' . If $(\hat{\mu} - k\hat{\sigma})$ is greater than or equal to L' , the lot is accepted; otherwise, it is rejected. The acceptance constant, denoted as k , is a pivotal factor in making this decision.

The key focus is on specifying the optimal sample size (n) and the pertinent acceptance constant (k) within the framework of the proposed censoring scheme. One can note that, the distribution of the variable $(\hat{\mu} - k\hat{\sigma})$ will be $AN\left((\mu - k\sigma), \frac{\sigma^2}{n}(\gamma_{11}(n, p_r, f_c) - 2k\gamma_{12}(n, p_r, f_c) + \gamma_{22}(n, p_r, f_c))\right)$. Here, $\gamma_{11}(n, p_r, f_c)$, $\gamma_{12}(n, p_r, f_c)$, and $\gamma_{22}(n, p_r, f_c)$ are elements of the asymptotic dispersion matrix. You can find detailed expressions for these in the Appendix provided by [26]. Here, p_r represents the removal probability, and f_c denotes the censoring fraction. Choose two points, $(p_\alpha, 1 - \alpha)$ and (p_β, β) , on the OC curve suggested by [7]. To calculate these points, we use the formulas: $y_\tau = \Psi^{-1}(\tau) = -\ln\left(-\frac{1}{\eta} \ln\left(1 - (1 - U)^{\frac{1}{\nu}}\right)\right)$ and $u_\alpha = \Phi^{-1}(\alpha)$, where $\Psi(\cdot)$ and $\Phi(\cdot)$ are the cumulative distribution functions (CDF) of the log-GIED and the standard normal distribution, respectively. To determine the acceptance constant (k) and the sample size (n) for a given pair of points, $(p_\alpha, 1 - \alpha)$ and (p_β, β) , on the OC curve, along with specified censoring fraction (f_c) and removal probability (p_r), we solve equations (5).

In Tables 1 and 2, we present the results for various removal probabilities ($p_r = 0.1, 0.3, \text{ and } 0.5$) and censoring fractions ($f_c = 0.2, 0.3, 0.4, 0.5, 0.6, \text{ and } 0.7$). The selection of values for p_α and p_β aligns with the criteria set by MIL-STD-105D. Additionally, we include results for a limit case of standard Type-II censoring, where $p_r = 0.00001$, allowing for a comparison with the findings of [7]. For the computation of the terms $\gamma_{11}(n, p_r, f_c)$, $\gamma_{12}(n, p_r, f_c)$, and $\gamma_{22}(n, p_r, f_c)$, we employ a Monte Carlo simulation, generating progressive Type-II censored samples initially.

Specifically, we calculate the moments based on 2000 simulations, assessing the average values of the terms $\gamma_{11}(n, p_r, f_c)$, $\gamma_{12}(n, p_r, f_c)$, and $\gamma_{22}(n, p_r, f_c)$ for various values of n . The outcomes are detailed in Table 1 and Table 2 for $\beta = 0.05$ and $\beta = 0.10$, respectively. The results reveal that, when maintaining a constant p_r , the optimum value of n decreases as f_c declines, irrespective of the acceptance constant (k). A lower f_c implies lesser dropouts, resulting in fewer accurate lifetime observations. Consequently, a larger sample size is necessary to compensate for the loss of information when assessing lot acceptability.

On the other hand, when the censoring fraction f_c is held constant, and the same acceptance constant k is used, the sample size does not exhibit a consistent pattern with respect to the removal probability. This discrepancy arises because removal shifts the observations toward the tail of the lifetime distribution, improving the accuracy of lifetime parameter estimation but leading to a loss of information due to dropouts. However, an excessive number of dropouts early in the process diminishes this advantage. Hence, for higher values of p_r , the sample size increases as p_r , such as $p_r = 0.5$, increases. Conversely, for low to moderate values of p_r , the sample size decreases as p_r increases due to the impact of a significant number of dropouts.

Table 1: Type-II PCR reliability sampling plan for p_α and p_β to match with MIL – STD – 105D for $1 - \alpha = 0.95, \beta = 0.10$.

p_α	p_β	$f_r \rightarrow$	n						k
			0.7	0.6	0.5	0.4	0.3	0.2	
$p_r = 0.1$									
0.00041	0.01840		14	10	9	9	8	8	1.1352
0.00284	0.03110		24	19	18	17	16	16	0.9834
0.00654	0.04260		32	27	26	25	24	23	0.8981
0.01090	0.05350		40	34	33	31	30	29	0.8372
0.02090	0.07420		53	47	45	43	41	39	0.7482
0.03190	0.09420		65	58	55	53	50	48	0.6816
$p_r = 0.3$									
0.00041	0.01840		14	12	11	10	9	8	1.1352
0.00284	0.03110		26	23	21	19	18	17	0.9834
0.00654	0.04260		36	33	30	28	25	24	0.8981
0.01090	0.05350		45	41	37	35	32	30	0.8372
0.02090	0.07420		61	56	51	47	44	41	0.7482
0.03190	0.09420		74	68	63	58	54	50	0.6816
$p_r = 0.5$									
0.00041	0.01840		16	13	12	10	9	9	1.1352
0.00284	0.03110		29	26	23	20	18	17	0.9834
0.00654	0.04260		40	36	32	29	26	24	0.8981
0.01090	0.05350		50	45	40	36	33	31	0.8372
0.02090	0.07420		68	61	55	50	45	42	0.7482
0.03190	0.09420		83	75	67	61	56	51	0.6816
$p_r = 0.0001$									
0.00041	0.01840		50	15	9	8	8	7	1.1352
0.00284	0.03110		66	23	17	16	15	15	0.9834
0.00654	0.04260		75	30	24	23	23	22	0.8981
0.01090	0.05350		81	37	31	30	29	28	0.8372
0.02090	0.07420		91	49	44	42	41	39	0.7482
0.03190	0.09420		99	61	56	54	51	49	0.6816

4.2. Simulated sampling plan

It is worth noting that in the discussion of the distribution of $(\hat{\mu} - k\hat{\sigma})$, asymptotic distribution theory is applied, and the derived sampling plans are based on this approximation. However, it is essential to investigate the finite sample behavior of these sampling plans by conducting a Monte Carlo simulations to assess the true probability of acceptance.

In this research, we employ a Monte Carlo simulation to compare the expected probability of acceptance with the actual probability when a designed sampling plan is put into practice within a specific censoring framework. We investigate various scenarios, incorporating removal probabilities (p_r) of 0.1, 0.3, and 0.5, as well as censoring fractions (f_c) of 0.3, 0.5, and 0.7. Additionally, we consider fixed producer's and consumer's risk (v, β) settings at (5%, 10%) and (5%, 5%). For each combination of these parameters, we conduct 2000 Monte Carlo simulations to provide precise estimates of the probability of acceptance, denoted as $\widehat{L(p)}$.

To obtain estimates for the parameters k and $\widehat{L(p)}$, we employ the bias-corrected maximum likelihood estimators (MLEs), as detailed by [9]. The obtained results are presented in Table 3 through Table 11, covering various removal probabilities p_r and diverse levels of censoring proportions.

Table 2: Progressive Type-II reliability sampling plan with random removals aligned with the requirements of MIL – STD – 105D for $1 - \alpha = 0.95$ and $\beta = 0.05$.

p_α	p_β	$f_r \rightarrow$	n						k
			0.7	0.6	0.5	0.4	0.3	0.2	
$p_r = 0.1$									
0.00041	0.01840		19	31	41	50	67	82	1.1664
0.00284	0.03110		13	25	35	43	60	74	1.0066
0.00654	0.04260		12	23	33	41	57	70	0.9180
0.01090	0.05350		11	22	31	40	54	67	0.8553
0.02090	0.07420		11	21	30	37	51	63	0.7641
0.03190	0.09420		10	20	28	36	49	61	0.6962
$p_r = 0.3$									
0.00041	0.01840		19	33	46	57	76	93	1.1664
0.00284	0.03110		16	29	41	52	70	86	1.0066
0.00654	0.04260		14	26	37	47	64	78	0.9180
0.01090	0.05350		13	25	35	44	60	74	0.8553
0.02090	0.07420		11	22	32	40	54	67	0.7641
0.03190	0.09420		11	21	29	37	51	63	0.6962
$p_r = 0.5$									
0.00041	0.01840		20	37	52	64	86	106	1.1664
0.00284	0.03110		17	32	45	56	76	93	1.0066
0.00654	0.04260		15	29	40	51	69	85	0.9180
0.01090	0.05350		13	26	37	46	63	77	0.8553
0.02090	0.07420		12	23	33	42	57	70	0.7641
0.03190	0.09420		11	21	30	38	52	64	0.6962
$p_r = 0.0001$									
0.00041	0.01840		79	101	112	118	128	135	1.1664
0.00284	0.03110		21	32	41	49	64	79	1.0066
0.00654	0.04260		11	21	30	39	55	71	0.9180
0.01090	0.05350		10	19	28	36	52	66	0.8553
0.02090	0.07420		10	19	28	36	51	64	0.7641
0.03190	0.09420		9	19	27	35	48	60	0.6962

Table 3: Simulated probabilities of acceptance for GIED with $p_r = 0.1$ and 70% censoring.

$p_r = 0.1$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L}(p)$
$\alpha = 0.05, \beta = 0.10$	15	4	1.1352	0.5278
	24	7	0.9834	0.3812
	32	9	0.8981	0.3025
	39	11	0.8372	0.2403
	52	15	0.7482	0.1518
	64	19	0.6815	0.0833
$\alpha = 0.05, \beta = 0.05$	18	5	1.1664	0.6235
	30	9	1.0066	0.4398
	40	12	0.9180	0.3424
	49	14	0.8553	0.2751
	66	19	0.7641	0.1797
	81	24	0.6962	0.1083

Table 4: Simulated probabilities of acceptance for GIED with $p_r = 0.1$ and 50% censoring. .

$p_r = 0.1$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L}(p)$
$\alpha = 0.05, \beta = 0.10$	10	5	1.1352	0.5284
	19	9	0.9834	0.3869
	26	13	0.8981	0.2988
	33	16	0.8372	0.2393
	46	23	0.7482	0.1520
	57	28	0.6815	0.0842
$\alpha = 0.05, \beta = 0.05$	12	6	1.1664	0.6150
	23	11	1.0066	0.4357
	33	16	0.9180	0.3416
	42	21	0.8553	0.2755
	58	29	0.7641	0.1803
	71	35	0.6962	0.1084

In Table 3, with a constant removal probability p_r , varying consumer's risk (β) and fixed producer's risk (ν), and fixed censoring proportion f_c , as the sample size n rises, the corresponding value of the number of the failures m also rises. However, the probability of acceptance $\widehat{L}(p)$ diminishes. Similar pattern have been observed from Table 3 to Table 11 for the different values of the censoring proportion f_c . From the Table 3 and Table 4, one can study the effect of the change of censoring proportion f_c . Here, with a constant removal probability p_r , consumer's risk (β) and fixed producer's risk (ν), size of the sample n decreases as the censoring proportion f_c decreases. A similar patterns has been observed for the rest of the tables for different values of the censoring fraction f_c and removal probability p_r , so one can conclude the same in general.

5. CONCLUSION

Our study has delved into the challenges and intricacies of Type-II Progressive Censoring (Type-II PCR), a common practical scenario where the number of removals is uncertain. Our primary focus has been the development of optimum reliability sampling plans for the GIED lifetime distribution within the framework of Type-II PCR. We have rigorously examined a range of scenarios involving removal probabilities and censoring fractions, shedding light on their influence on these sampling plans.

Table 5: Simulated probabilities of acceptance for GIED with $p_r = 0.1$ and 30% censoring. .

$p_r = 0.1$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L(p)}$
$\alpha = 0.05, \beta = 0.10$	8	5	1.1352	0.4955
	17	11	0.9834	0.3853
	24	16	0.8981	0.3014
	30	21	0.8372	0.2399
	41	28	0.7482	0.1508
	51	35	0.6815	0.0838
$\alpha = 0.05, \beta = 0.05$	11	7	1.1664	0.6245
	21	14	0.9745	0.4376
	30	21	0.9180	0.3421
	38	26	0.8553	0.2755
	52	36	0.7641	0.1798
	64	44	0.6962	0.1086

Table 6: Simulated probabilities of acceptance for GIED with $p_r = 0.3$ and 70% censoring..

$p_r = 0.3$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L(p)}$
$\alpha = 0.05, \beta = 0.10$	14	4	1.1352	0.5104
	26	7	0.9834	0.3833
	36	10	0.8981	0.3012
	45	13	0.8372	0.2415
	60	18	0.7482	0.1513
	74	22	0.6815	0.0840
$\alpha = 0.05, \beta = 0.05$	18	5	1.1664	0.6163
	33	9	1.0066	0.4396
	46	13	0.9180	0.3444
	56	16	0.8553	0.2750
	76	22	0.7641	0.1800
	93	27	0.6962	0.1089

Table 7: Simulated probabilities of acceptance for GIED with $p_r = 0.3$ and 50% censoring. .

$p_r = 0.3$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L(p)}$
$\alpha = 0.05, \beta = 0.10$	11	5	1.1352	0.5438
	21	10	0.9834	0.3973
	30	15	0.8981	0.3130
	37	18	0.8372	0.2485
	51	25	0.7482	0.1590
	63	31	0.6815	0.0907
$\alpha = 0.05, \beta = 0.05$	14	7	1.1664	0.6225
	27	13	1.0065	0.4410
	38	19	0.9180	0.3436
	47	23	0.8553	0.2746
	65	32	0.7641	0.1803
	80	40	0.6962	0.1092

Table 8: Simulated probabilities of acceptance for GIED with $p_r = 0.3$ and 30% censoring.

$p_r = 0.3$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L(p)}$
$\alpha = 0.05, \beta = 0.10$	9	6	1.1352	0.5106
	18	12	0.9834	0.3838
	25	17	0.8981	0.2983
	32	22	0.8372	0.2400
	44	30	0.7482	0.1518
	54	37	0.6815	0.0837
$\alpha = 0.05, \beta = 0.05$	12	8	1.1664	0.6311
	23	16	1.0066	0.4418
	32	22	0.9180	0.3421
	40	28	0.8553	0.2742
	55	38	0.7641	0.1794
	68	47	0.6962	0.1087

Table 9: Simulated probabilities of acceptance for GIED with $p_r = 0.5$ and 70% censoring..

$p_r = 0.5$	Sampling plan			Probabilities of acceptance
	n	m	k	$\widehat{L(p)}$
$\alpha = 0.05, \beta = 0.10$	16	4	1.1352	0.5240
	29	8	0.9834	0.3841
	40	12	0.8981	0.3006
	50	15	0.8372	0.2407
	67	20	0.7482	0.1510
	83	24	0.6815	0.0842
$\alpha = 0.05, \beta = 0.05$	20	6	1.1664	0.6224
	37	11	1.0066	0.4414
	51	15	0.9180	0.3436
	63	18	0.8553	0.2756
	85	25	0.7641	0.1799
	104	31	0.6962	0.1088

Table 10: Simulated probabilities of acceptance for GIED with $p_r = 0.5$ and 50% censoring.

$p_r = 0.5$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L(p)}$
$\alpha = 0.05, \beta = 0.10$	12	6	1.1352	0.5217
	23	11	0.9834	0.3845
	32	16	0.8981	0.2995
	41	20	0.8372	0.2415
	55	27	0.7482	0.1508
	68	34	0.6815	0.0836
$\alpha = 0.05, \beta = 0.05$	15	7	1.1664	0.6213
	29	14	1.0066	0.4402
	41	20	0.9180	0.3434
	51	25	0.8553	0.2750
	70	35	0.7641	0.1800
	86	43	0.6962	0.1088

Table 11: Simulated probabilities of acceptance for GIED with $p_r = 0.5$ and 30% censoring.

$p_r = 0.5$	Sampling plan			Probability of acceptance
	n	m	k	$\widehat{L}(p)$
$\alpha = 0.05, \beta = 0.10$	10	7	1.1352	0.5259
	19	13	0.9833	0.3834
	27	18	0.8981	0.3010
	34	23	0.8372	0.2405
	46	32	0.7482	0.1507
	57	39	0.6815	0.0836
$\alpha = 0.05, \beta = 0.05$	12	8	1.1664	0.6128
	24	16	1.0066	0.4396
	34	23	0.9180	0.3427
	43	30	0.8553	0.2759
	58	40	0.7641	0.1791
	72	50	0.6962	0.1088

Our key findings, as evident in the parameters of sample size (n) and the acceptance constant (k), underscore the crucial role of an increasing censoring fraction (f_c) in necessitating a larger sample size. Generally, the optimal sample size (n) exhibits stability across varying removal probabilities (p_r). Nonetheless, it is of paramount importance to highlight the pivotal role played by the removal probability (p_r) in shaping the overall test duration. [16] has convincingly demonstrated that an escalation in the removal probability (p_r) leads to a significant extension of the test duration. In such cases, an increased sample size becomes imperative to effectively mitigate the extended testing period. These insights emphasize the practical significance of our research in addressing real-world challenges related to reliability testing under Type-II PCR.

Declaration of conflicting interest: All authors have no conflict of interests in the publication of the manuscript.

Funding : This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data: No such data are provided in this manuscript. Only simulation study has been performed.

REFERENCES

- [1] Koyama, T., Suga, R., Yokoh, T., Ohmae, Y., Yamamoto, T., Pabst Jr, W. R. (1970). MIL-STD-105D and the Japanese modified standard. *Journal of Quality Technology*, 2(2), 99–108.
- [2] Pandey, A., Kaushik, A., Singh, S. K., Singh, U. (2021). On the estimation problems for exponentiated exponential distribution under generalized progressive hybrid censoring: On the generalised progressive hybrid censoring. *Austrian Journal of Statistics*, 50(1), 24-40.
- [3] Kaushik, A., Singh, U., Singh, S. K. (2017). Bayesian inference for the parameters of Weibull distribution under progressive Type-I interval censored data with beta-binomial removals. *Communications in Statistics-Simulation and Computation*, 46(4), 3140-3158.
- [4] Fertig, K. W., Mann, N. R. (1980). Life-test sampling plans for two-parameter Weibull populations. *Technometrics*, 22(2), 165-177.
- [5] Hosono, Y., Ohta, H., Kase, S. (1981). Design of single sampling plans for doubly exponential characteristics. *Frontiers in Quality Control*, 94-112.

- [6] Kocherlakota, S., Balakrishnan, N. (1986). One-and two-sided sampling plans based on the exponential distribution. *Naval research logistics quarterly*, 33(3), 513-522.
- [7] Schneider, H. (1989). Failure-censored variables-sampling plans for lognormal and Weibull distributions. *Technometrics*, 31(2), 199-206.
- [8] Balasooriya, U. (1995). Failure-censored reliability sampling plans for the exponential distribution. *Journal of Statistical Computation and Simulation*, 52(4), 337-349.
- [9] Balakrishnan, N., Aggarwala, R. (2000). *Progressive censoring: theory, methods, and applications*. Springer Science Business Media.
- [10] Balasooriya, U., Saw, S. L. (1998). Reliability sampling plans for the two-parameter exponential distribution under progressive censoring. *Journal of Applied Statistics*, 25(5), 707-714.
- [11] Balasooriya, U., Balakrishnan, N. (2000). Reliability sampling plans for lognormal distribution, based on progressively-censored samples. *IEEE Transactions on Reliability*, 49(2), 199-203.
- [12] Balasooriya, U., Saw, S. L., Gadag, V. (2000). Progressively censored reliability sampling plans for the Weibull distribution. *Technometrics*, 42(2), 160-167.
- [13] Montanari, G. C., Cacciari, M. (1988). Progressively-censored aging tests on XLPE-insulated cable models. *IEEE Transactions on Electrical Insulation*, 23(3), 365-372.
- [14] Yuen, H. K., Tse, S. K. (1996). Parameters estimation for Weibull distributed lifetimes under progressive censoring with random removals. *Journal of Statistical Computation and Simulation*, 55(1-2), 57-71.
- [15] Tse, S. K., Yuen, H. K. (1998). Expected experiment times for the Weibull distribution under progressive censoring with random removals. *Journal of Applied Statistics*, 25(1), 75-83.
- [16] Tse, S. K., Yang, C., Yuen, H. K. (2000). Statistical analysis of Weibull distributed lifetime data under Type II progressive censoring with binomial removals. *Journal of Applied Statistics*, 27(8), 1033-1043.
- [17] Lieberman, G. J., Resnikoff, G. J. (1955). Sampling plans for inspection by variables. *Journal of the American Statistical Association*, 50(270), 457-516.
- [18] Balasooriya, U., Saw, S. L. (1999). A note on approximate moments of progressively censored order statistics. *Metron*, 57(1-2), 117-130.
- [19] Balakrishnan, N., Sandhu, R. A. (1995). A simple simulational algorithm for generating progressive Type-II censored samples. *The American Statistician*, 49(2), 229-230.
- [20] Wu, S. J., Huang, S. R. (2012). Progressively first-failure censored reliability sampling plans with cost constraint. *Computational Statistics Data Analysis*, 56(6), 2018-2030.
- [21] Barlow, R. E., Proschan, F. (1974). *Statistical theory of reliability and life testing: probability models*.
- [22] Lin, C. T., Duran, B. S., Lewis, T. O. (1989). Inverted gamma as a life distribution. *Microelectronics Reliability*, 29(4), 619-626.
- [23] Singh, S. K., Singh, U., Kumar, D. (2013). Bayes estimators of the reliability function and parameter of inverted exponential distribution using informative and non-informative priors. *Journal of Statistical computation and simulation*, 83(12), 2258-2269.
- [24] Abouammoh, A. M., Alshingiti, A. M. (2009). Reliability estimation of generalized inverted exponential distribution. *Journal of statistical computation and simulation*, 79(11), 1301-1315.
- [25] Krishna, H., Kumar, K. (2013). Reliability estimation in generalized inverted exponential distribution with progressively type II censored sample. *Journal of Statistical Computation and Simulation*, 83(6), 1007-1019.
- [26] Srivastava, P. W., Sharma, D. (2014). Optimum Time?Censored Constant?Stress PALTSP for the Burr Type XII Distribution Using Tampered Failure Rate Model. *Journal of Quality and Reliability Engineering*, 2014(1), 564049.
- [27] Kaushik, A., Pandey, A., Maurya, S. K., Singh, U., Singh, S. K. (2017). Estimations of the parameters of generalised exponential distribution under progressive interval type-I censoring scheme with random removals. *Austrian Journal of Statistics*, 46(2), 33-47.
- [28] Kaushik, A., Singh, U., Singh, S. K. (2017). Bayesian inference for the parameters of Weibull distribution under progressive Type-I interval censored data with beta-binomial removals. *Communications in Statistics-Simulation and Computation*, 46(4), 3140-3158.

ROLE OF AGEING METRICS TO ANALYSE THE SURVIVAL DATA OF TONGUE CANCER PATIENTS

B. ELINA^{1*}, PULAK SWAIN^{2†}, SATYA KR. MISRA³, SUBARNA BHATTACHARJEE⁴

^{1,4} Department of Mathematics, Ravenshaw University, Cuttack-753003, Odisha, India

² Department of Mathematics, ITER (SOA University), Bhubaneswar-751030, Odisha, India

³ Department of Mathematics, KIIT University, Bhubaneswar-751024, Odisha, India

¹elina.2294@gmail.com, ²pulakswain1994@gmail.com,

³satyamisra05@gmail.com, ⁴subarna.bhatt@gmail.com

Abstract

The paper vividly describes the non-parametric estimation of basic quantities for right censored data of times to death for patients with tongue cancer. Here we compare patients with two different sets of DNA profile using several parameters like reliability function, cumulative hazard rate function, smoothed hazard rate function and ageing intensity function. With the help of graphical representations of these functions, we analyse which DNA profile patients have better prognosis.

Keywords: survival function, hazard rate, cumulative hazard rate, ageing intensity function.

AMS 2020 Subject Classification: Primary 60E15, Secondary 62N05, 60E05.

1. INTRODUCTION

A quantitative analysis of failure data through various reliability functions such as, survival function, hazard (failure) rate, reversed hazard rate is known to researchers since a long time. These failure data are usually related to mechanical or a biological systems. In recent literature [4], [11], [5], [7], we also find the use of ageing intensity function along with other ageing metrics as discussed above to know about the ageing phenomena underlying a given failure data.

In many biomedical applications the primary endpoint of interest is time to a certain event. Example include: time of deaths, time it takes for a patient to respond to a therapy, time from response until disease relapse (that is, disease returns), etc.

Two important issues arise when studying time-to-event data (we will assume the "event" to be death):

(i) Some individuals are still alive till the end of the study or at the time of analysis. So the event of interest, namely death, has not occurred. Therefore we have right censored data.

(ii) Length of follow-of varies due to staggered entry. So we cannot observe the event for those individuals with insufficient times.

Suppose the events occur at D distinct times $t_1 < t_2 < \dots < t_D$, and at time t_i there are d_i events. Let Y_i be the number of individuals who are at risk at time t_i . Y_i is a count of the number of individuals with a time on study of t_i or more. The quantity $\frac{d_i}{Y_i}$ provides an estimate of the conditional probability that an individual who survives to just prior to time t_i experiences the

*The work was jointly done with the first author when she was in Ravenshaw University, Cuttack-753003, Odisha, India.

†Corresponding author : E-mail: pulakswain1994@gmail.com

event at time t_i .

The ageing intensity function $L(t)$ of any system at time $t > 0$, with probability density function $f(t)$, survival function $\bar{F}(t)$, hazard rate $h(t) = \frac{f(t)}{\bar{F}(t)}$, and cumulative hazard rate

$$H(t) = \frac{\int_0^t h(u)du}{t}$$

is given by [8]

$$\begin{aligned} L(t) &= \frac{-tf(t)}{\bar{F}(t) \ln \bar{F}(t)}, \text{ where defined,} \\ &= \frac{th(t)}{\int_0^t h(u)du} \\ &= \frac{h(t)}{H(t)}. \end{aligned}$$

Works on aforementioned functions can be found in [8], [17], [2], [6], [16], [12, 13], [14], [18], [5], [19].

In the present work, we first make a review on the notion of commonly known Kaplan-Meier and Nelson-Aalen Estimator used in survival analysis. Further, we take up a right censored data of times to death for patients with tongue cancer to illustrate the significance of the ageing metrics. In particular, we examine through different ageing metrics for drawing an inference about the distribution of the time to some event X , based on sample of right censored survival data of "Times to Death for Patients with Tongue Cancer".

The rest of the paper is organized as follows. A brief literature on Kaplan-Meier and Nelson-Aalen estimation is given in Section 2. Consequently, Section 3 presents the kernel based estimation for the ageing intensity function. Further, the survival analysis of the tongue cancer patients is done in Section 4 with the help of several ageing metrics. Finally, the concluding remarks are provided in Section 5.

2. VARIANCE OF KAPLAN-MEIER AND NELSON-AALEN ESTIMATORS

The standard estimator of the survival function proposed by [9] called Product limit estimators defined as

$$\hat{S}(t) = \begin{cases} 1 & \text{if } t < t_1 \\ \prod_{t_i \leq t} \frac{d_i}{Y_i} & \text{if } t > t_1 \end{cases} \quad (1)$$

The variance of the Product-Limit estimator is estimated by Greenwood's formula

$$\hat{V}[\hat{S}(t)] = (\hat{S}(t))^2 \sum_{t_i \leq t} \frac{d_i}{Y_i(Y_i - d_i)}. \quad (2)$$

The standard error of the product-Limit estimator is given by $\{\hat{V}[\hat{S}(t)]\}^{\frac{1}{2}}$. An estimator of the cumulative hazard rate, was first suggested by [15] and then rediscovered by [1] which is referred as Nelson-Aalen estimator of the cumulative hazard, defined as

$$\tilde{H}(t) = \begin{cases} 0 & \text{if } t \leq t_1 \\ \sum_{t_i \leq t} \frac{d_i}{Y_i} & \text{if } t \geq t_1 \end{cases} \quad (3)$$

The estimated variance of the Nelson-Aalen estimator is given by

$$\sigma_H^2(t) = \sum_{t_i < t} \frac{d_i}{(Y_i)^2}. \quad (4)$$

The standard error of the Nelson-Aalen estimator is given by $(\sigma_H^2(t))^{\frac{1}{2}}$.

3. KERNEL BASED ESTIMATION OF THE AGEING INTENSITY

Kernel-smoothed estimators of $h(t)$ are based on the Nelson-Aalen estimator $\tilde{H}(t)$ and its variance $\hat{V}[\tilde{H}(t)]$.

Let $\Delta\tilde{H}(t_i) = \tilde{H}(t_i) - \tilde{H}(t_{i-1})$ and $\Delta\hat{V}[\tilde{H}(t)] = \hat{V}[\tilde{H}(t_i)] - \hat{V}[\tilde{H}(t_{i-1})]$ denote the magnitude of the jumps in $\tilde{H}(t_i)$ and $\hat{V}[\tilde{H}(t_{i-1})]$ at time t_i . $\Delta\tilde{H}(t_i)$ provides a crude estimator of $h(t)$ at the death times. The Kernel smoothed estimator of $h(t)$ is a weighted average of these crude estimates over event times close to t . Closeness is determined by a bandwidth b , so that event times in the range $t - b$ to $t + b$ are included in the weighted average which estimate $h(t)$. The weights are controlled by the choice of a kernel function $K(\cdot)$, defined on the interval $[-1, 1]$, which determines how much weight is given to points at a distance from t .

The kernel used in following estimation of hazard rate is uniform kernel with

$$K(x) = \frac{1}{2} \text{ for } -1 \leq x \leq 1, \quad \text{if } b \leq t \leq t_D - b \quad (5)$$

$$K_q(x) = \frac{4(1+q^3)}{(1+q)^4} + \frac{6(1-q)}{(1+q)^3} \text{ for } -1 \leq x \leq q, \quad \text{if } t \leq b \text{ given } q = \frac{t}{b} \quad (6)$$

$$K_q(x) = \frac{4(1+q^3)}{(1+q)^4} - \frac{6(1-q)}{(1+q)^3} \text{ for } -1 \leq x \leq q, \quad \text{if } t_D - b \leq t \leq t_D \text{ given } q = \frac{(t_D - t)}{b} \quad (7)$$

The kernel smoothed estimator of $h(t)$ based on the kernel $K(\cdot)$ is given by

$$\hat{h}(t) = b^{-1} \sum_{i=1}^D K\left(\frac{t-t_i}{b}\right) \Delta\tilde{H}(t_i). \quad (8)$$

The variance of $\hat{h}(t)$ is estimated by the quantity

$$\sigma^2[\hat{h}(t)] = b^{-2} \sum_{i=1}^D K\left(\frac{t-t_i}{b}\right)^2 \Delta\hat{V}[\tilde{H}(t_i)]. \quad (9)$$

Section 4 is based on the study of ageing phenomenon on the patients with cancer of the tongue.

4. STUDY ON THE EFFECTS OF PLOIDY ON THE PROGNOSIS OF PATIENTS WITH MOUTH CANCER

4.1. Background

Patients were selected who had a paraffin-embedded sample of the cancerous tissue taken at the time of surgery. The tissue samples were examined using a flow cytometer to determine if the tumor had an aneuploid or diploid DNA profile. The data in the Table 1 is on the patients with tongue cancer (c.f. [10]).

Data on 79 Patients with Cancer of the Tongue:

g : Tumor DNA profile – 1: Aneuploid, 2: Diploid

T : Time (in weeks) to death or on study time

δ : Death indicator – 1: Dead, 0: Alive

4.2. Results

From the survival function graph given in Figure 1, it can be observed that the curve end at different points as the times on study are different for two DNA group patients (i.e., 400 weeks for aneuploid patients and 231 weeks for diploid patients). Secondly the figure suggests the aneuploid patients have the best and diploid patients the least favourable prognosis. The disease free survival probability are 0.2286 ($SE = 0.0954$) for aneuploid patients and 0.0833 ($SE = 0.0716$)

Table 1: Data on 79 patients with cancer of the tongue

g	T	δ	g	T	δ
1	1	1	1	93	0
1	3	1	1	93	0
1	3	1	1	101	0
1	4	1	1	104	0
1	10	1	1	108	0
1	13	1	1	109	0
1	16	1	1	131	0
1	16	1	1	150	0
1	24	1	1	231	0
1	26	1	1	240	0
1	27	1	1	400	0
1	28	1	2	1	1
1	30	1	2	3	1
1	30	1	2	4	1
1	32	1	2	5	1
1	41	1	2	5	1
1	51	1	2	8	1
1	65	1	2	12	1
1	67	1	2	13	1
1	70	1	2	18	1
1	73	1	2	26	1
1	77	1	2	27	1
1	91	1	2	30	1
1	93	1	2	42	1
1	96	1	2	56	1
1	100	1	2	62	1
1	104	1	2	69	1
1	157	1	2	104	1
1	167	1	2	104	1
1	61	0	2	112	1
1	74	0	2	129	1
1	79	0	2	181	1
1	80	0	2	8	0
1	81	0	2	67	0
1	87	0	2	76	0
1	87	0	2	104	0
1	88	0	2	176	0
1	89	0	2	231	0

for diploid patients.

We also observe from Table 2 and Table 3 that the estimated survival function at 12 months after the transplant for aneuploid group is 0.6731 and diploid group is 0.4863 (that is, at 1 year (12 months), 67.31% of aneuploid patients were alive, whereas 48.63% of diploid patients were alive). The extended final plateau of the graph indicates that people are being cured.

From the cumulative hazard rate graph (Figure 2), we interpreted that the estimate of the cumulative hazard rate function is steeper for first 100-110 weeks (i.e., in first 110 weeks the hazard rate is approximately constant). And the plot shows that the aneuploid group patients

Table 2: Product limit estimator and its estimated variance for aneuploid group patients

t_i	d_i	Y_i	$\hat{S}(t)$	$\hat{V}[\hat{S}(t)]$	$\{\hat{V}[\hat{S}(t)]\}^{\frac{1}{2}}$
1	1	52	0.980769	0.000362711	0.01904496
3	2	51	0.942308	0.00104546	0.03233357
4	1	49	0.923077	0.001365498	0.03695265
10	1	48	0.903846	0.001671313	0.0408817
13	2	47	0.865385	0.002240271	0.0473315
16	2	45	0.826923	0.002752333	0.05246268
24	1	43	0.807692	0.002987028	0.05465371
26	1	42	0.788462	0.003207499	0.05663478
27	1	41	0.769231	0.003413746	0.05842727
28	1	40	0.75	0.003605769	0.06004806
30	2	39	0.711538	0.003947144	0.0628263
32	1	37	0.692308	0.004096495	0.06400387
41	1	36	0.673077	0.004231623	0.06505092
51	1	35	0.653846	0.004352526	0.06597368
65	1	33	0.634033	0.004473413	0.06688358
67	1	32	0.614219	0.004578501	0.06766462
70	1	31	0.594406	0.00466779	0.00466779
72	1	30	0.574592	0.00474128	0.06885695
73	1	29	0.554779	0.004798971	0.06927461
77	1	27	0.534231	0.004856632	0.06968954
91	1	19	0.506114	0.005107841	0.07146916
93	1	18	0.477996	0.005302736	0.07281989
96	1	16	0.448122	0.005497328	0.07414397
100	1	14	0.416113	0.005691417	0.07544148
104	1	12	0.381437	0.005884598	0.07671114
157	1	5	0.305149	0.008421952	0.0917712
167	1	4	0.228862	0.009102169	0.09540529

have the smallest death rate and the diploid group patients have the highest death rate. We also observe from Table 4a and Table 4b respectively that the estimated cumulative hazard function at 12 months after the transplant for aneuploid group is 0.3892 and diploid group is 0.6999 (i.e., at 1 year (12 months), 38.92% of aneuploid patients were dead whereas 69.99% of diploid patients were dead). Table 5a, Table 6a give a record of hazard rate and ageing intensity of aneuploid and diploid patients respectively. One can note that Table 5b to Table 5d depict the method to calculate hazard rate of aneuploid group at $t = 4, 30, 167$ respectively. On a similar line, the values of $\hat{h}(t_i)$ at different t_i s for aneuploid group are obtained in Table 5a. Table 6b to Table 6d reflect the computation of hazard rate of diploid group at $t = 8, 27, 181$ respectively. The required values of $\hat{h}(t_i)$ at other t_i 's for diploid group are shown in Table 6a. Thus we get a crude estimate of hazard function.

Since Figure 2 shows a crude estimate of the hazard rate so as to provide a smoothed estimated hazard rate we used uniform kernel estimation which is shown in Figure 3. The figure indicates the risk of death or hazard rate decreases slowly but the initial peak is high for diploid group patients.

From the smoothed hazard rate graphs of two DNA group patients, we compare the hazard rate of two graphs on various subintervals. These comparisons are given in Table 7.

Let $h_1(t)$ be the hazard rate function of the aneuploid patients and $h_2(t)$ be the hazard rate

Table 3: Product limit estimator and its estimated variance for diploid group patients

t_i	d_i	Y_i	$\hat{S}(t)$	$\hat{V}[\hat{S}(t)]$	$\{\hat{V}[\hat{S}(t)]\}^{\frac{1}{2}}$
1	1	28	0.964286	0.001229956	0.035070732
3	1	27	0.928571	0.002368805	0.048670367
4	1	26	0.892857	0.003416545	0.058451221
5	2	25	0.821429	0.005238703	0.072378882
8	1	23	0.785714	0.00601312	0.077544307
12	1	21	0.748299	0.006787295	0.082385044
13	1	20	0.710884	0.00745542	0.086344773
18	1	19	0.673469	0.008017493	0.089540453
23	1	18	0.636054	0.008473514	0.092051693
26	1	17	0.598639	0.008823484	0.093933402
27	1	16	0.561224	0.009067402	0.095222909
30	1	15	0.52381	0.009205269	0.095944095
42	1	14	0.486395	0.009237085	0.096109754
56	1	13	0.44898	0.009162849	0.095722771
62	1	12	0.411565	0.008982561	0.094776375
69	1	10	0.370408	0.008800344	0.093810147
104	2	8	0.277806	0.00816587	0.090365204
112	1	5	0.222245	0.007695797	0.087725689
129	1	4	0.166684	0.006644173	0.081511795
181	1	2	0.083342	0.005133974	0.071651756

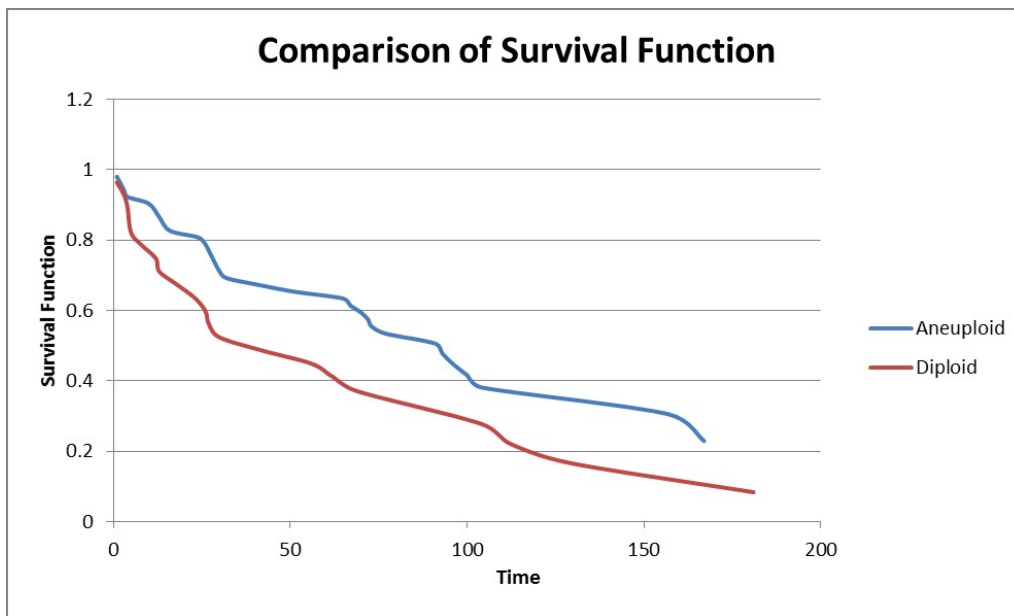


Figure 1: Comparison of survival function between aneuploid and diploid patients.

function of diploid patients.

On the basis of the above analysis we can clearly observe at interval $[1, 5)$, $[5, 13)$, $[13, 23)$, $[23, 41)$ and $[41, 42)$, the hazard rate for aneuploid patients is less than the diploid patients where as at the subintervals $[42, 51)$, the hazard rate for diploid patients is less than the aneuploid patients. Similarly we can compare the hazard rates on remaining subintervals and on an overall we

Table 4: Construction of the Nelson-Aalen estimator and its estimated variance for aneuploid and diploid tumor

t_i	$\tilde{H}(t_i)$	$\sigma_{\tilde{H}}^2(t)$	Standard Error
1	0.019230769	0.00036982	0.01923077
3	0.058446456	0.00113876	0.03374548
4	0.078854619	0.00155525	0.03943667
10	0.099687952	0.00198928	0.04460133
13	0.142241144	0.00289467	0.0538021
16	0.186685588	0.00388232	0.06230826
24	0.209941402	0.00442315	0.06650679
26	0.233750926	0.00499005	0.07064026
27	0.25814117	0.00558493	0.07473239
28	0.28314117	0.00620993	0.07880311
30	0.334423221	0.00752485	0.08674592
32	0.361450248	0.00825531	0.09085876
41	0.389228026	0.00902692	0.0950101
51	0.417799454	0.00984325	0.09921313
65	0.448102485	0.01076152	0.10373775
67	0.479352485	0.01173808	0.10834243
70	0.511610549	0.01277867	0.11304276
72	0.544943883	0.01388978	0.11785489
73	0.579426641	0.01507884	0.12279592
77	0.616463678	0.01645058	0.12825981
91	0.669095257	0.01922066	0.1386386
93	0.724650813	0.02230708	0.14935555
96	0.787150813	0.02621333	0.16190532
100	0.858579384	0.03131537	0.1769615
104	0.941912717	0.03825982	0.19560117
157	1.141912717	0.07825982	0.27974956
167	1.391912717	0.14075982	0.37517971

(a) Aneuploid

t_i	$\tilde{H}(t_i)$	$\sigma_{\tilde{H}}^2(t)$	Standard Error
1	0.035714286	0.00127551	0.03571429
3	0.072751323	0.002647252	0.05145146
4	0.111212861	0.004126542	0.06423817
5	0.191212861	0.007326542	0.08559522
8	0.234691122	0.009216901	0.09600469
12	0.28231017	0.011484475	0.10716564
13	0.33231017	0.013984475	0.11825597
18	0.384941749	0.016754558	0.1294394
23	0.440497304	0.019840978	0.14085801
26	0.499320834	0.023301186	0.15264726
27	0.561820834	0.027207436	0.16494677
30	0.6284875	0.03165188	0.17790975
42	0.699916072	0.036753921	0.19171312
56	0.776839149	0.042671081	0.2065698
62	0.860172482	0.049615525	0.22274543
69	0.960172482	0.059615525	0.24416291
104	1.210172482	0.090865525	0.30143909
112	1.410172482	0.130865525	0.3617534
129	1.660172482	0.193365525	0.43973347
181	2.160172482	0.443365525	0.66585699

(b) Diploid

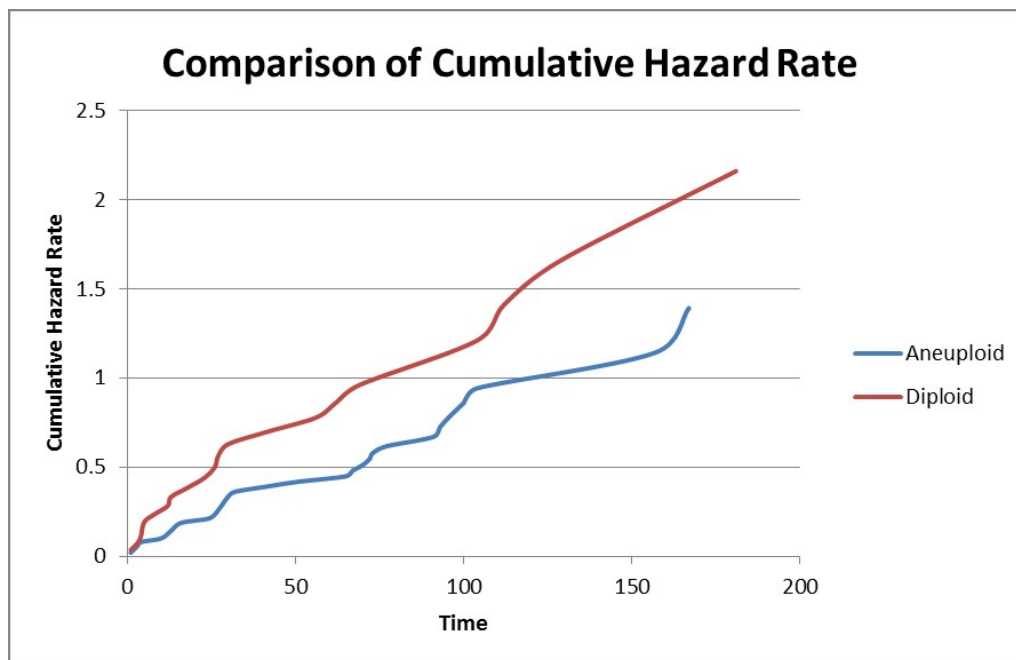


Figure 2: Comparison of cumulative hazard rate between aneuploid and diploid patients.

can again confirm the impression that aneuploid group patients have the lowest rate of death. From the ageing intensity graphs of two DNA group patients Figure 4, we compare the ageing intensities of two graphs at different intervals of time. These comparisons are given in Table 8.

Table 5: Analysis of uniform smoothed hazard rate for aneuploid group patients

t_i	d_i	Y_i	$\tilde{H}(t_i)$	$\hat{h}(t_i)$	$\hat{L}(t_i)$
1	1	52	0.019230769	0.013989484	0.727453187
3	2	51	0.058446456	0.009881253	0.169065048
4	1	49	0.078854619	0.010365443	0.131450046
10	1	48	0.099687952	0.009334279	0.093634980
13	2	47	0.142241144	0.008372741	0.058863004
16	2	45	0.186685588	0.007744815	0.041485877
24	1	43	0.209941402	0.010960455	0.052207212
26	1	42	0.233750926	0.010960455	0.046889462
27	1	41	0.258141170	0.008738233	0.033850598
28	1	40	0.283141170	0.008738233	0.030861754
30	2	39	0.334423221	0.008738233	0.026129265
32	1	37	0.361450248	0.010127122	0.028018024
41	1	36	0.389228026	0.004168812	0.010710461
51	1	35	0.417799454	0.002817460	0.006743571
65	1	33	0.448102485	0.008081359	0.018034623
67	1	32	0.479352485	0.009933211	0.020722144
70	1	31	0.511610549	0.009933211	0.019415571
72	1	30	0.544943883	0.009933211	0.018227952
73	1	29	0.579426641	0.009933211	0.017143173
77	1	27	0.616463678	0.009933211	0.016113214
91	1	19	0.669095257	0.012105785	0.018092768
93	1	18	0.724650813	0.012105785	0.016705681
96	1	16	0.787150813	0.016272452	0.020672598
100	1	14	0.858579384	0.085857938	0.100000000
104	1	12	0.941912717	0.010863095	0.011533017
157	1	5	1.141912717	0.022500000	0.019703783
167	1	4	1.391912717	0.100000000	0.071843585

(a)

t_i	$\Delta H(t_i)$	x	$K_q(x)$	$K_q(x)\Delta H(t_i)$
1	0.019231	0.3	1.501458	0.02887419
3	0.039216	0.1	1.239067	0.048590875
4	0.020408	0	1.107872	0.022609633
10	0.020833	-0.6	0.3207	0.006681254
13	0.042553	-0.9	-0.07289	-0.003101519
16	0.044444	-1.2	0	0
24	0.023256	-2	0	0
26	0.02381	-2.2	0	0
27	0.02439	-2.3	0	0
28	0.025	-2.4	0	0
30	0.051282	-2.6	0	0
32	0.027027	-2.8	0	0
41	0.027778	-3.7	0	0
51	0.028571	-4.7	0	0
65	0.030303	-6.1	0	0
67	0.03125	-6.3	0	0
70	0.032258	-6.6	0	0
72	0.033333	-6.8	0	0
73	0.034483	-6.9	0	0
77	0.037037	-7.3	0	0
91	0.052632	-8.7	0	0
93	0.055556	-8.9	0	0
96	0.0625	-9.2	0	0
100	0.071429	-9.6	0	0
104	0.083333	-10	0	0
157	0.2	-15.3	0	0
167	0.25	-16.3	0	0
				$h(4) = 0.010365443$

(b) At $t = 4 < b (= 10)$, $K_q(x) = \frac{4(1+q^3)}{(1+q)^4} + \frac{6(1-q)}{(1+q)^3}$

$$x = \frac{4-t_i}{10}, \hat{h}(8) = (10)^{-1} \sum_i K_q(x) \Delta \tilde{H}(t_i)$$

t_i	$\Delta H(t_i)$	x	$K_q(x)$	$K_q(x)\Delta H(t_i)$
1	0.019231	2.9	0	0
3	0.039216	2.7	0	0
4	0.020408	2.6	0	0
10	0.020833	2	0	0
13	0.042553	1.7	0	0
16	0.044444	1.4	0	0
24	0.023256	0.6	0.5	0.011627907
26	0.02381	0.4	0.5	0.011904762
27	0.02439	0.3	0.5	0.012195122
28	0.025	0.2	0.5	0.0125
30	0.051282	0	0.5	0.025641026
32	0.027027	-0.2	0.5	0.013513514
41	0.027778	-1.1	0	0
51	0.028571	-2.1	0	0
65	0.030303	-3.5	0	0
67	0.03125	-3.7	0	0
70	0.032258	-4	0	0
72	0.033333	-4.2	0	0
73	0.034483	-4.3	0	0
77	0.037037	-4.7	0	0
91	0.052632	-6.1	0	0
93	0.055556	-6.3	0	0
96	0.0625	-6.6	0	0
100	0.071429	-7	0	0
104	0.083333	-7.4	0	0
157	0.2	-12.7	0	0
167	0.25	-13.7	0	0
				$h(30) = 0.008738233$

(c) At $t = 30, (10 =) b < t < t_D - b (= 157)$, $K_q(x) = \frac{1}{2}$
 $x = \frac{30-t_i}{10}, \hat{h}(27) = (10)^{-1} \sum_i K_q(x) \Delta \tilde{H}(t_i)$

t_i	$\Delta H(t_i)$	x	$K_q(x)$	$K_q(x)\Delta H(t_i)$
1	0.019231	16.6	0	0
3	0.039216	16.4	0	0
4	0.020408	16.3	0	0
10	0.020833	15.7	0	0
13	0.042553	15.4	0	0
16	0.044444	15.1	0	0
24	0.023256	14.3	0	0
26	0.02381	14.1	0	0
27	0.02439	14	0	0
28	0.025	13.9	0	0
30	0.051282	13.7	0	0
32	0.027027	13.5	0	0
41	0.027778	12.6	0	0
51	0.028571	11.6	0	0
65	0.030303	10.2	0	0
67	0.03125	10	0	0
70	0.032258	9.7	0	0
72	0.033333	9.5	0	0
73	0.034483	9.4	0	0
77	0.037037	9	0	0
91	0.052632	7.6	0	0
93	0.055556	7.4	0	0
96	0.0625	7.1	0	0
100	0.071429	6.7	0	0
104	0.083333	6.3	0	0
157	0.2	1	0	0
167	0.25	0	4	1
				$h(167) = 0.1$

(d) At $t = 167, (157 =) t_D - b < t < t_D (= 167)$, $K_q(x) = \frac{4(1+q^3)}{(1+q)^4} - \frac{6(1-q)}{(1+q)^3}$, $x = \frac{167-t_i}{10}, \hat{h}(167) = (10)^{-1} \sum_i K_q(x) \Delta \tilde{H}(t_i)$

Table 6: Analysis of uniform smoothed hazard rate for diploid group patients

t_i	d_i	Y_i	$\tilde{H}(t_i)$	$\hat{h}(t_i)$	$\hat{L}(t_i)$
1	1	28	0.035714286	0.031166063	0.872649778
3	1	27	0.072751323	0.023461182	0.322484609
4	1	26	0.111212861	0.024474361	0.220067724
5	2	25	0.191212861	0.024130904	0.126199167
8	1	23	0.234691122	0.021894230	0.093289553
12	1	21	0.282310170	0.017461373	0.061851733
13	1	20	0.332310170	0.020239151	0.060904398
18	1	19	0.384941749	0.018530399	0.048138189
23	1	18	0.440497304	0.017308867	0.039293922
26	1	17	0.499320834	0.014808867	0.029658019
27	1	16	0.561820834	0.014808867	0.026358700
30	1	15	0.628487500	0.012177288	0.019375545
42	1	14	0.699916072	0.003571429	0.005102653
56	1	13	0.776839149	0.008012821	0.010314646
62	1	12	0.860172482	0.008012821	0.009315365
69	1	10	0.960172482	0.009166667	0.009546896
104	2	8	1.210172482	0.022500000	0.018592391
112	1	5	1.410172482	0.022500000	0.015955495
129	1	4	1.660172482	0.012500000	0.007529338
181	1	2	2.160172482	0.200000000	0.092585199

(a)

t_i	$\Delta H(t_i)$	x	$K_q(x)$	$K_q(x)\Delta H(t_i)$
1	0.03571429	0.7	0.720164	0.025720132
3	0.03703704	0.5	0.679012	0.025148574
4	0.03846154	0.4	0.658435	0.025324438
5	0.08000000	0.3	0.637859	0.051028744
8	0.04347826	0	0.576131	0.025049174
12	0.04761905	-0.4	0.493827	0.023515552
13	0.05000000	-0.5	0.473251	0.023662525
18	0.05263158	-1	0.37037	0.019493158
23	0.05555556	-1.5	0	0
26	0.05882353	-1.8	0	0
27	0.06250000	-1.9	0	0
30	0.06666667	-2.2	0	0
42	0.07142857	-3.4	0	0
56	0.07692308	-4.8	0	0
62	0.08333333	-5.4	0	0
69	0.10000000	-6.1	0	0
104	0.25000000	-9.6	0	0
112	0.20000000	-10.4	0	0
129	0.25000000	-12.1	0	0
181	0.50000000	-17.3	0	0

(b) At $t = 8 < b (= 10)$, $K_q(x) = \frac{4(1+q^3)}{(1+q)^4} + \frac{6(1-q)}{(1+q)^3}$,

$x = \frac{8-t_i}{10}$, $\hat{h}(8) = (10)^{-1} \sum_i K_q(x) \Delta \tilde{H}(t_i)$

t_i	$\Delta H(t_i)$	x	$K_q(x)$	$K_q(x)\Delta H(t_i)$
1	0.03571429	2.6	0	0
3	0.03703704	2.4	0	0
4	0.03846154	2.3	0	0
5	0.08000000	2.2	0	0
8	0.04347826	1.9	0	0
12	0.04761905	1.5	0	0
13	0.05000000	1.4	0	0
18	0.05263158	0.9	0.5	0.026315789
23	0.05555556	0.4	0.5	0.027777778
26	0.05882353	0.1	0.5	0.029411765
27	0.06250000	0	0.5	0.03125
30	0.06666667	-0.3	0.5	0.033333333
42	0.07142857	-1.5	0	0
56	0.07692308	-2.9	0	0
62	0.08333333	-3.5	0	0
69	0.10000000	-4.2	0	0
104	0.25000000	-7.7	0	0
112	0.20000000	-8.5	0	0
129	0.25000000	-10.2	0	0
181	0.50000000	-15.4	0	0

(c) At $t = 27$, $(10 =) b < t < t_D - b (= 171)$, $K_q(x) = \frac{1}{2}$,

$x = \frac{27-t_i}{10}$, $\hat{h}(27) = (10)^{-1} \sum_i K_q(x) \Delta \tilde{H}(t_i)$

t_i	$\Delta H(t_i)$	x	$K_q(x)$	$K_q(x)\Delta H(t_i)$
1	0.03571429	1.8	0	0
3	0.03703704	1.78	0	0
4	0.03846154	1.77	0	0
5	0.08000000	1.76	0	0
8	0.04347826	1.73	0	0
12	0.04761905	1.69	0	0
13	0.05000000	1.68	0	0
18	0.05263158	1.63	0	0
23	0.05555556	1.58	0	0
26	0.05882353	1.55	0	0
27	0.06250000	1.54	0	0
30	0.06666667	1.51	0	0
42	0.07142857	1.39	0	0
56	0.07692308	1.25	0	0
62	0.08333333	1.19	0	0
69	0.10000000	1.12	0	0
104	0.25000000	-7.7	0.77	0
112	0.20000000	0.69	0	0
129	0.25000000	0.52	0	0
181	0.50000000	0	4	2

(d) At $t = 181$, $(171 =) t_D - b < t < t_D (= 181)$, $K_q(x) = \frac{4(1+q^3)}{(1+q)^4} - \frac{6(1-q)}{(1+q)^3}$,

$x = \frac{181-t_i}{10}$, $\hat{h}(181) = (10)^{-1} \sum_i K_q(x) \Delta \tilde{H}(t_i) = 0.2$

Let $L_1(t)$ be the ageing intensity of the aneuploid patients and $L_2(t)$ be the ageing intensity of diploid patients.

From the Table 8 we get, at interval $[1, 5)$, the ageing intensity for aneuploid patients is less than the diploid patients where as at the subintervals $[5, 13)$, the ageing intensity for aneuploid patients is more than the diploid patients. But as we proceed, we observe that there is an alternate sign for the rest of the subintervals. Thus we cannot get a concluding remark. To obtain the desired result, we now calculate the total time for which aneuploid patients have less ageing intensity than diploid patients and vice versa.

After calculations, we get to know that the aneuploid patients have low ageing intensity than the diploid patients for a time of 101 weeks, whereas the diploid patients have less ageing intensity than aneuploid patients for a time of 79 weeks.

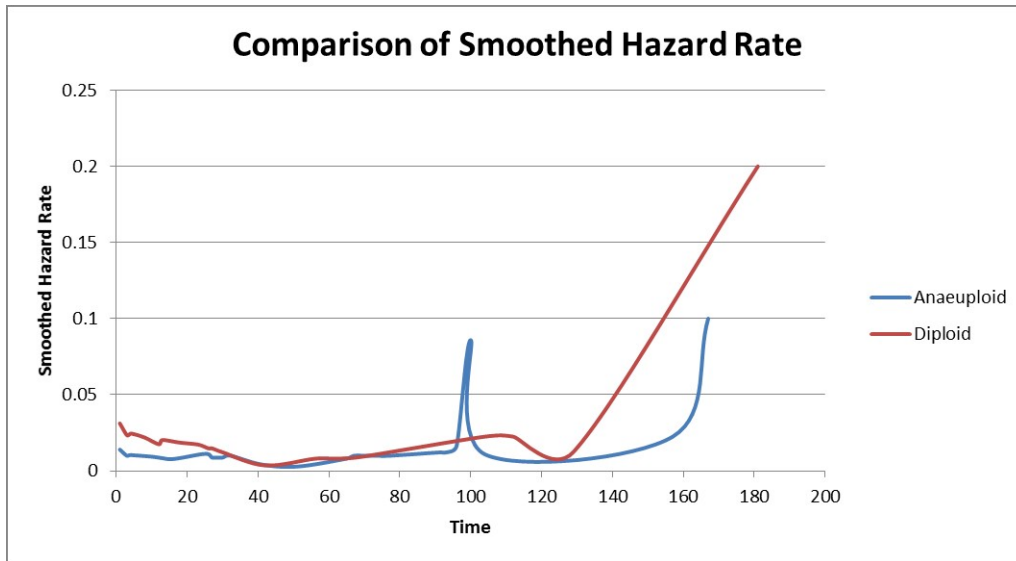


Figure 3: Comparison of smoothed hazard rate between aneuploid and diploid patients.

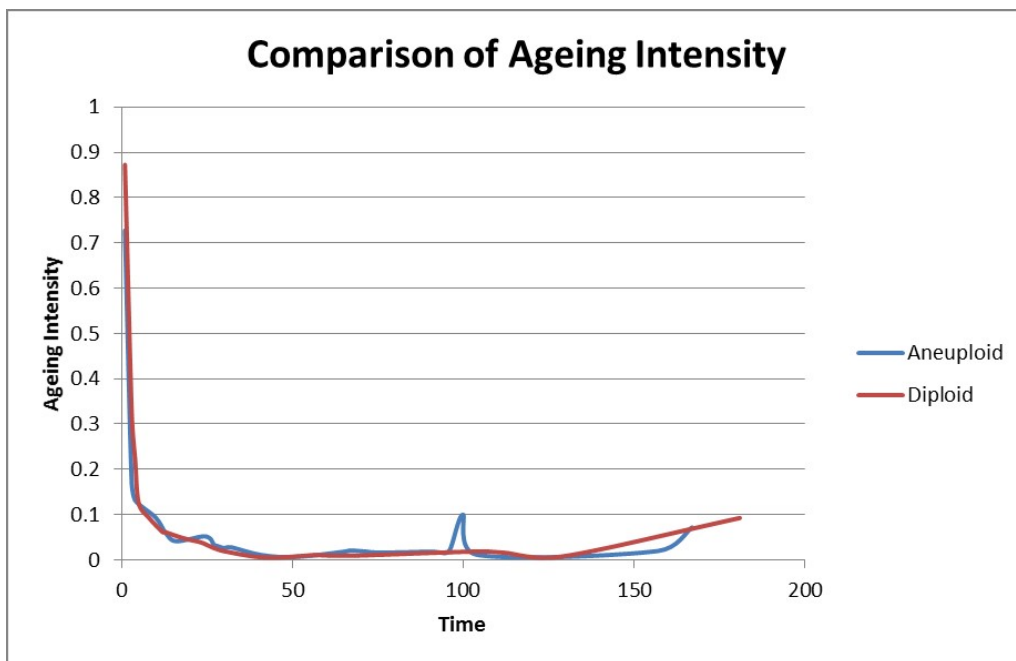


Figure 4: Comparison of ageing intensity between aneuploid and diploid patients.

On the basis of the above analysis we can clearly observe that the cancer patients with aneuploid DNA have less ageing intensity than the patients with diploid DNA profile.

5. CONCLUSION

There are several factors which are the majors of prognosis for the death (failure) of a man in case of the transplant among the patients with tongue cancer. Nonetheless, the study is confined to the effect of ploidy on the survival of patients with tongue cancer. Based on the failure data available for the patients with tongue cancer with aneuploid and diploid DNA profile, statistical analyses were made and graphical interpretation was studied from the curves obtained with several parameters like reliability function cumulative hazard rate function, smoothed hazard

Table 7: Interval-wise comparison of hazard rates between aneuploid and diploid patients

Interval	Comparison of $h(t)$
[1, 5)	$h_1(t) < h_2(t)$
[5, 13)	$h_1(t) < h_2(t)$
[13, 23)	$h_1(t) < h_2(t)$
[23, 41)	$h_1(t) < h_2(t)$
[41, 42)	$h_1(t) < h_2(t)$
[42, 51)	$h_1(t) > h_2(t)$
[51, 65)	$h_1(t) < h_2(t)$
[65, 104)	$h_1(t) > h_2(t)$
[104, 157)	$h_1(t) < h_2(t)$
[157, 181)	$h_1(t) > h_2(t)$

Table 8: Interval-wise comparison of ageing intensities between aneuploid and diploid patients

Interval	Comparison of $L(t)$
[1, 5)	$L_1(t) < L_2(t)$
[5, 13)	$L_1(t) > L_2(t)$
[13, 23)	$L_1(t) < L_2(t)$
[23, 41)	$L_1(t) > L_2(t)$
[41, 42)	$L_1(t) < L_2(t)$
[42, 51)	$L_1(t) > L_2(t)$
[51, 65)	$L_1(t) < L_2(t)$
[65, 104)	$L_1(t) > L_2(t)$
[104, 157)	$L_1(t) < L_2(t)$
[157, 181)	$L_1(t) < L_2(t)$

rate function and ageing intensity. It is analysed from the graph in all the cases that patients with aneuploid DNA profile ought to be one of the basis for prognosis for the patients with tongue cancer. Hence it is inferred that patients with aneuploid tumors may get benefit significantly from a prolonged tumor free period. Here we discuss about the analysis of the censored and uncensored failure data through various measures of ageing phenomenon. Moreover, we summarize various ageing concepts of the lifetimes that have been widely studied in the field of reliability.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

ACKNOWLEDGEMENTS

Subarna Bhattacharjee would like to thank Odisha State Higher Education Council for providing support to carry out the research project under OURIIP, Odisha, India (Grant No. 22-SF-MT-073).

REFERENCES

- [1] Aalen, O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics*, 701-726.
- [2] Barlow, R. E., and Proschan, F. (1975). Statistical theory of reliability and life testing: probability models. Florida State Univ Tallahassee.
- [3] Bhattacharjee, S., Nanda, A. K., and Misra, S. K. (2013). Reliability analysis using ageing intensity function. *Statistics & Probability Letters*, 83(5): 1364-1371, DOI: 10.1016/j.spl.2013.01.016.
- [4] Bhattacharjee, S., and Misra, S. K. (2016). Some aging properties of Weibull models. *Electronic Journal of Applied Statistical Analysis*, 9(2): 297-307, DOI: 10.1285/i20705948v9n2p297.
- [5] Bhattacharjee, S., Mohanty, I., Szymkowiak, M., and Nanda, A. K. (2022). Properties of aging functions and their means. *Communications in Statistics-Simulation and Computation*, 1-20, DOI: 10.1080/03610918.2022.2141257.
- [6] Ebeling, C. E. (2004). An introduction to reliability and maintainability engineering. Tata McGraw-Hill Education.
- [7] Giri, R. L., Nanda, A. K., Dasgupta, M., Misra, S. K., and Bhattacharjee, S. (2023). On ageing intensity function of some Weibull models. *Communications in Statistics-Theory and Methods*, 52(1): 227-262, DOI: 10.1080/03610926.2021.1910845.
- [8] Jiang, R., Ji, P., and Xiao, X. (2003). Aging property of unimodal failure rate models. *Reliability Engineering & System Safety*, 79(1): 113-116, DOI: 10.1016/S0951-8320(02)00175-8.
- [9] Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282): 457-481, DOI: 10.1080/01621459.1958.10501452.
- [10] Klein, J. P., and Moeschberger, M. L. (2003). Survival analysis: techniques for censored and truncated data, 1230, New York: Springer, DOI: 10.1007/b97377.
- [11] Misra, S. K., and Bhattacharjee, S. (2018). A case study of aging intensity function on censored data. *Alexandria Engineering Journal*, 57(4): 3931-3952, DOI: 10.1016/j.aej.2018.03.009.
- [12] Nanda, A. K., Bhattacharjee, S., and Alam, S. S. (2006). On upshifted reversed mean residual life order. *Communications in Statistics-Theory and Methods*, 35(8): 1513-1523, DOI: 10.1080/03610920600637271.
- [13] Nanda, A. K., Bhattacharjee, S., and Alam, S. S. (2007). Properties of aging intensity function. *Statistics & Probability Letters*, 77(4): 365-373, DOI: 10.1016/j.spl.2006.08.002.
- [14] Nanda, A. K., Bhattacharjee, S., and Balakrishnan, N. (2009). Mean residual life function, associated orderings and properties. *IEEE Transactions on Reliability*, 59(1): 55-65, DOI: 10.1109/TR.2009.2035791.
- [15] Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14(4): 945-966, DOI: 10.1080/00401706.1972.10488991.
- [16] Shaked, M., and Shanthikumar, J. G. (Eds.). (2007). Stochastic orders. New York, NY: Springer New York, DOI: 10.1007/978-0-387-34675-5.
- [17] Shooman, M. L. (1968). Probabilistic reliability: an engineering approach.
- [18] Swain, P., Bhattacharjee, S., and Misra, S. K. (2021). A Case Study to Analyze Ageing Phenomenon in Reliability Theory. *Reliability: Theory & Applications*, 16(4 (65)): 275-285, DOI: 10.24412/1932-2321-2021-465-275-285.
- [19] Szymkowiak, M., Roychowdhury, A., Misra, S. K., Giri, R. L., and Bhattacharjee, S. (2023). A study of a survival data using kernel estimates of hazard rate and aging intensity functions. *Statistics in Transition new series*, 24(5): 109-127, DOI: 10.59170/stattrans-2023-066.

STATISTICAL PROPERTIES AND APPLICATIONS OF TRANSMUTED SKEW STUDENT t DISTRIBUTION

DAVID Ikwuoche John, MATHEW Stephen

•

Department of Mathematics and Statistics, Federal University Wukari, Nigeria

davidij@fuwukari.edu.ng, matsteve231@gmail.com

<https://orcid.org/0000-0002-7100-5357>

Abstract

In this study, a modified 2-parameter skew t distribution called the transmuted skew student t distribution (TSS t D) was presented. Some statistical and reliability properties of TSS t D such as the quantile function, the raw moments, and the moment generating function (among others), were derived. Through the method of maximum likelihood, the two parameters of the model were estimated. The stability of the model was studied via Montecarlo simulations utilizing bias, mean square error, and root mean square error as metrics. The results from the stability study revealed that the TSS t D was well-behaved. Four datasets were modeled with the transmuted skewed student t distribution and four other probability density models. On the basis of information criteria, the results revealed that the transmuted skew student t distribution provides a better fit for all the datasets compared to the other competing models.

Keywords: Transmuted, Skew, Raw moments, Quantile, Reliability function, Hazard function

I. Introduction

Empirical probability models, or probability distributions, are an essential aspect of parametric statistical investigation. Classical distributions are more susceptible to an anomaly when characterizing various data and data generating processes, according to a prevalent reality across a variety of sectors, including finance, environmental science, biological sciences, engineering, and others. This emphasizes the need to hybridize classical probability models in order to meet this complex task [1]. Owing to the applicability of distribution theory and the availability of diverse data in today's world, the thirst for improved statistical distributions that might be used to describe and model these events have increased spontaneously [2], [3], [4], [5], and [6]. Many scholars have over the years help modified simple statistical methodologies in relation to distribution theory and these methods have been found immensely useful in statistical modeling. Several methods of modifying probability distributions have been proposed over the years by scholars to improve statistical methodology of distribution theory. Among the methods are Exponentiated Exponential Distribution [7], the Sine-G family [8], the New Sine-G Family [9], the G-families using the transformed-transformer [10], Transmuted G family by [11], and amongst others. Several distributions have been modified using the Transmuted G family of distributions over the years and some of them include Transmuted Lomax and Transmuted Exponentiated Lomax distributions [12], Transmuted Frechet distribution [13], Transmuted Exponentiated Gamma distribution [14], Transmuted additive Weibull distribution [15], Transmuted generalized Gompertz distribution [16], Kumaraswamy Transmuted Exponentiated modified Weibull

distribution [17], Transmuted Exponential power distribution [18], etc. Using the method of transmutation by [19], this paper proposes a new probability distribution called the Transmuted Skew Student t Distribution (TSS t D).

II. Methods

I. The Transmuted Family of Distribution

This section presents the TSS t D, some of its statistical properties, simulation study as well as application to real life data. The cumulative density function (CDF) and probability density function (PDF) of the Transmuted family of distribution generator is given by;

$$F(v) = (1 + \alpha)G(v) - \alpha G(v)^2 \quad (1)$$

$$f(v) = g(v)\{(1 + \alpha) - 2\alpha G(v)\} \quad (2)$$

where, α is the transmuted parameter (shape), $F(v)$ is the CDF and $f(v)$ is the PDF of the baseline distribution.

II. The Skew Student t Distribution

Using the simplified version of the Skew Student t Distribution (SS t D) with 2 degrees of freedom introduced by Jonson *et al.*, [20] whose CDF and PDF are expressed as;

$$\left. \begin{aligned} G(v) &= \frac{1}{2} \left(1 + \frac{v}{\sqrt{\Lambda + v^2}} \right) \\ g(v) &= \frac{\Lambda}{2(\Lambda + v^2)^{3/2}} \end{aligned} \right\} ; -\infty < v < \infty \quad (3)$$

where, Λ is the shape parameter.

III. Transmuted Skew Student t Distribution (TSS t D)

On substituting $G(v)$ and $g(v)$ in equation (3) into (1) and (2) the PDF and CDF of TSS t D are obtained as;

$$f(v; \Lambda, \alpha) = \frac{\Lambda}{2(\Lambda + v^2)^{3/2}} \left[(1 + \alpha) - 2\alpha \left(\frac{\sqrt{\Lambda + v^2} + v}{2\sqrt{\Lambda + v^2}} \right) \right] \quad (4)$$

and

$$F(v; \Lambda, \alpha) = (1 + \alpha) \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right) - \alpha \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right)^2 \quad (5)$$

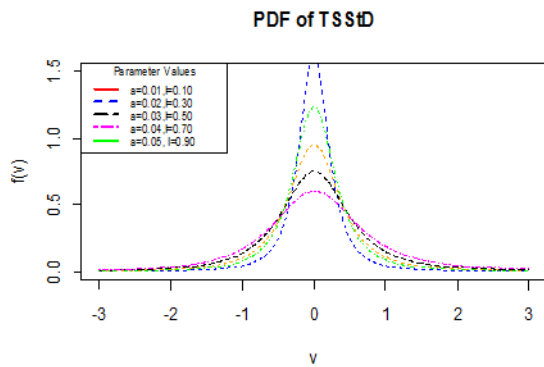


Figure 1: The PDF Plot of TSStD

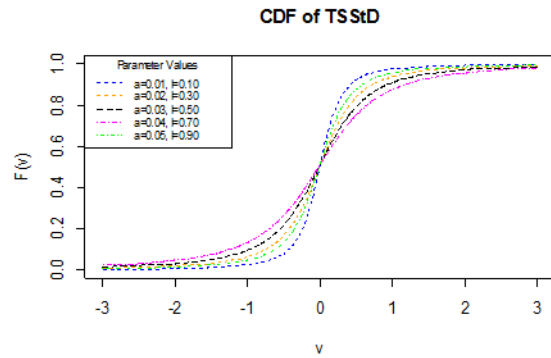


Figure 2: The CDF Plot of TSStD

IV. Properties of Transmuted Skew Student t Distribution

The survival $[S(x)]$, hazard $[h(x)]$, odd $[O(x)]$, and quantile $[Q(v)]$ functions, skewness, and kurtosis statistics are presented as well as the $S(x)$ plot as follows

$$S(x) = 1 - F(V \leq v) = 1 - \left((1 + \alpha) \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right) - \alpha \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right)^2 \right) \quad (6)$$

$$h(x) = \frac{f(x)}{S(x)} = \frac{\frac{\Lambda}{2(\Lambda + v^2)^{\frac{3}{2}}} \left[(1 + \alpha) - 2\alpha \left(\frac{\sqrt{\Lambda + v^2} + v}{2\sqrt{\Lambda + v^2}} \right) \right]}{1 - \left((1 + \alpha) \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right) - \alpha \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right)^2 \right)} \quad (7)$$

$$O(x) = \frac{F(x)}{S(x)} = \frac{(1 + \alpha) \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right) - \alpha \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right)^2}{1 - \left((1 + \alpha) \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right) - \alpha \left(\frac{1}{2} + \frac{v}{2\sqrt{\Lambda + v^2}} \right)^2 \right)} \quad (8)$$

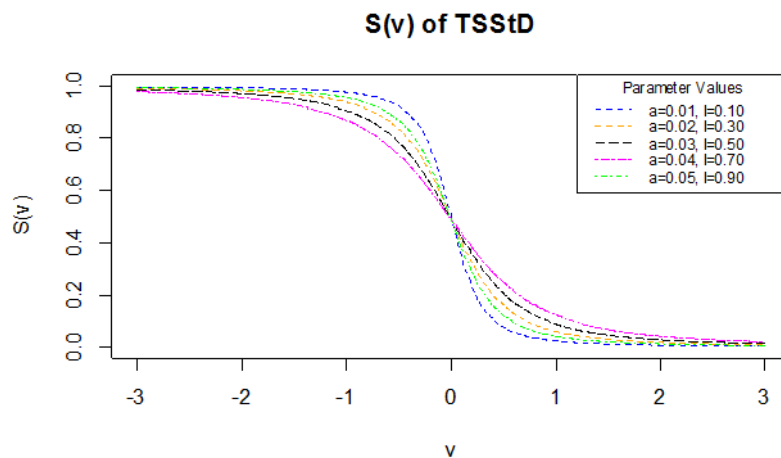


Figure 3: The $S(v)$ Plot of TSStD for varied parameter value

Quantile Function: the inverse of the CDF of TSStD gives the quantile function of TSStD and after numerous algebraic simplifications, this is expressed as;

$$Q(v) = \pm \left(-\frac{\left(\alpha^2\Lambda^2 - 4\alpha\Lambda^2u^* + 2\alpha\Lambda^2 + \Lambda^2\right)^{\frac{1}{2}}}{8(u^{*2} - u^*)} + \frac{\alpha\Lambda u^*}{4(u^{*2} - u^*)} - \frac{\alpha\Lambda}{8(u^{*2} - u^*)} - \frac{\Lambda u^{*2}}{u^{*2} - u^*} + \frac{\Lambda u^*}{u^{*2} - u^*} - \frac{\Lambda}{8(u^{*2} - u^*)} \right)^{\frac{1}{2}} \quad (9)$$

Skewness: the Galton measure of skewness [GSK] which measures the presence and lack of symmetry of a probability distribution is presented for $TSStD$ using $Q(v)$ as follows;

$$GSK = \frac{Q\left(\frac{1}{8}\right) + Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right) - Q\left(\frac{1}{2}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)} \quad (10)$$

Kurtosis: the kurtosis measure whether or not the probability distribution is heavy-tailed. For $TSStD$ the Moor's kurtosis measure is derived using $Q(v)$ as follows;

$$MKT = \frac{Q\left(\frac{1}{8}\right) + Q\left(\frac{3}{8}\right) - Q\left(\frac{5}{8}\right) - Q\left(\frac{1}{4}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)} \quad (11)$$

IV. Moments, Moment Generating Function, and Characteristic Function

Moments: A random variable v with $TSStD$ has its moment in the form;

$$u_r = \frac{\lambda}{2} \left[G\beta \left[\frac{r^2 - r + 2}{2}, \frac{r^2}{2} \right] - H\beta \left[\frac{r+2}{2}, \frac{r}{2} \right] \right] \quad (12)$$

Proof

The r^{th} moment of a random variable v with a valid probability distribution is given by;

$$u_r = E[v^r] = \int_{-\infty}^{\infty} v^r f(v) dv \quad (13)$$

Let v be a random variable following the $TSStD$. Then the moment is derived as follows;

$$u_r = E[v^r] = \int_{-\infty}^{\infty} v^r \frac{\Lambda}{2(\Lambda + v^2)^{\frac{3}{2}}} \left[1 - \frac{\alpha v}{(\Lambda + v^2)^{\frac{1}{2}}} \right] dv \quad (14)$$

$$\begin{aligned} u_r &= \int_{-\infty}^{\infty} \frac{\Lambda v^r}{2} \left[\frac{1}{(\Lambda + v^2)^{\frac{3}{2}}} - \frac{\alpha v}{(\Lambda + v^2)^{\frac{1}{2} \cdot \frac{3}{2}}} \right] dv = \frac{2\Lambda}{2} \int_0^{\infty} \left[\frac{v^r}{(\Lambda + v^2)^{\frac{3}{2}}} - \frac{\alpha v v^r}{(\Lambda + v^2)^2} \right] dv \\ u_r &= \Lambda \int_0^{\infty} \left[\frac{v^r}{(\Lambda + v^2)^{\frac{3}{2}}} - \frac{\alpha v^{r+1}}{(\Lambda + v^2)^2} \right] dv \end{aligned} \quad (15)$$

By transformation, let

$$u = v^r; v = u^{\frac{1}{r}}; \frac{du}{dv} = r v^{r-1}; dv = \frac{1}{r v^{r-1}} du. \text{ Substituting the transformations into (15) to obtain;}$$

$$u_r = \Lambda \int_0^\infty \left[\frac{u^{\frac{r}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^{\frac{3}{2}}} - \frac{\alpha u^{\frac{1}{r(r+1)}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^2} \right] \frac{1}{ru^{\frac{1}{r}(r-1)}} du \quad (16)$$

$$u_r = \frac{\Lambda}{r} \left[\int_0^\infty \frac{u^{\frac{r-1}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^{\frac{3}{2}}} du - \int_0^\infty \frac{\alpha u^{\frac{r+1}{r} \left(\frac{r-1}{r}\right)} }{\left(\Lambda + u^{\frac{2}{r}}\right)^2} du \right] = \frac{\lambda}{r} \left[\int_0^\infty \frac{u^{\frac{r-r+1}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^{\frac{3}{2}}} du - \int_0^\infty \frac{\alpha u^{\frac{r+1-r+1}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^2} du \right]$$

$$u_r = \frac{\Lambda}{r} \left[\Lambda^{\frac{3}{2}} \int_0^\infty \frac{u^{\frac{1}{r}}}{\left(1 + \frac{u^r}{\Lambda}\right)^{\frac{3}{2}}} du - \frac{\alpha}{\Lambda^2} \int_0^\infty \frac{u^{\frac{2}{r}}}{\left(1 + \frac{u^r}{\Lambda}\right)^2} du \right] \quad (17)$$

Also, letting $k = \frac{u^r}{\Lambda}$, $\Rightarrow u = (\Lambda k)^{\frac{r}{2}}$, $\Rightarrow \frac{dk}{du} = \frac{2}{\Lambda r} u^{\frac{2-r}{r}}$, $\Rightarrow du = \frac{\Lambda r}{2u^{\frac{2-r}{r}}} dk$ then by substituting for u and du in

(17), we obtain the following;

$$u_r = \frac{\Lambda}{r} \left[\Lambda^{\frac{3}{2}} \int_0^\infty \frac{(\Lambda k)^{\frac{r}{2} \left(\frac{1}{r}\right)}}{(1+k)^{\frac{3}{2}}} \frac{\Lambda r}{2(\Lambda k)^{\frac{r}{2} \left(\frac{2-r}{r}\right)}} dk - \frac{\alpha}{\Lambda^2} \int_0^\infty \frac{(\Lambda k)^{\frac{r}{2} \left(\frac{2}{r}\right)}}{(1+k)^2} \frac{\Lambda r}{2(\Lambda k)^{\frac{r}{2} \left(\frac{2-r}{r}\right)}} dk \right]$$

$$u_r = \frac{\Lambda}{2} \left[\frac{\Lambda^{\frac{r^2-r}{2}}}{\Lambda^{\frac{1}{2}}} \int_0^\infty \frac{k^{\frac{r^2-r}{2}}}{(1+k)^{\frac{3}{2}}} dk - \frac{\alpha \Lambda^{\frac{r}{2}}}{\Lambda} \int_0^\infty \frac{k^{\frac{r}{2}}}{(1+k)^2} dk \right] = \frac{\Lambda}{2} \left[G \int_0^\infty \frac{k^{\frac{r^2-r}{2}}}{(1+k)^{\frac{3}{2}}} dk - H \int_0^\infty \frac{k^{\frac{r}{2}}}{(1+k)^2} dk \right] \quad (18)$$

where, $G = \frac{\Lambda^{\frac{r^2-r}{2}}}{\Lambda^{\frac{1}{2}}}$; $H = \frac{\alpha \Lambda^{\frac{r}{2}}}{\Lambda}$. Therefore, the Beta function is given as;

$$\beta[p, q] = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}} dt \quad (19)$$

Therefore, on transforming (18) into the form presented in (19), the following is obtained;

$$u_r = \frac{\Lambda}{2} \left[G \int_0^\infty \frac{k^{\frac{r^2-r}{2}+1}}{(1+k)^{\frac{3}{2}+\frac{r^2}{2}}} dk - H \int_0^\infty \frac{k^{\frac{r}{2}+1}}{(1+k)^{2+\frac{r^2}{2}+\frac{r^2}{2}}} dk \right] = \frac{\Lambda}{2} \left[G\beta\left[\frac{r^2-r+2}{2}, \frac{r^2}{2}\right] - H\beta\left[\frac{r+2}{2}, \frac{r}{2}\right] \right] \quad (20)$$

which completes the proof.

The mean of TSS t D, that is, $E[v]$ based on equation (20) by substituting r to be 1 is given as;

$$E[v] = u_1 = \frac{\Lambda}{2} \left[G\beta\left[\frac{1^2-1+2}{2}, \frac{1^2}{2}\right] - H\beta\left[\frac{1+2}{2}, \frac{1}{2}\right] \right] = \frac{\Lambda}{2} \left[G\beta\left[1, \frac{1}{2}\right] - H\beta\left[\frac{3}{2}, \frac{1}{2}\right] \right] \quad (21)$$

Moment Generating Function (MGF): The MGF of the random variable v with TSS t D is given by:

$$M_v(t) = \sum_{m=0}^\infty \frac{t^m}{m!} \cdot \frac{\Lambda}{2} \left[G\beta\left[\frac{m^2-m+2}{2}, \frac{m^2}{2}\right] - H\beta\left[\frac{m+2}{2}, \frac{m}{2}\right] \right] \quad (22)$$

Proof

The MGF is given by:

$$M_v(t) = E[e^{tv}] = \int_{-\infty}^{\infty} e^{tv} f(v) dv \quad (23)$$

By McLaurin's series expansion, e^{tv} is expressed as $e^{tv} = \sum_{m=0}^{\infty} \frac{(tv)^m}{m!}$ then equation (23) becomes

$$M_v(t) = E[e^{tv}] = E\left[\sum_{m=0}^{\infty} \frac{(tv)^m}{m!}\right] = \sum_{m=0}^{\infty} \frac{t^m}{m!} E[v^m] \quad (24)$$

As obtained in equation (20), $E[v^m] = u_m$, therefore,

$$M_v(t) = \sum_{m=0}^{\infty} \frac{t^m}{m!} \cdot \frac{\Lambda}{2} \left[G\beta \left[\frac{m^2 - m + 2}{2}, \frac{m^2}{2} \right] - H\beta \left[\frac{m+2}{2}, \frac{m}{2} \right] \right] \quad (25)$$

Characteristic Function [CF]: for a random variable v with TSS t D, the CF is given as;

$$f_v(t) = \sum_{c=0}^{\infty} \frac{(it)^c}{c!} \cdot \frac{\Lambda}{2} \left[G\beta \left[\frac{r^2 - r + 2}{2}, \frac{r^2}{2} \right] - H\beta \left[\frac{r+2}{2}, \frac{r}{2} \right] \right] \quad (26)$$

Proof

The CF of a valid probability distribution is expressed as;

$$f_v(t) = E[e^{itv}] = \int_{-\infty}^{\infty} e^{itv} f(v) dv \quad (27)$$

By Taylor's series expansion of $e^{itv} = \sum_{c=0}^{\infty} \frac{(itv)^c}{c!}$, therefore,

$$f_v(t) = \sum_{c=0}^{\infty} \frac{(itv)^c}{c!} \cdot \frac{\Lambda}{2} \left[G\beta \left[\frac{r^2 - r + 2}{2}, \frac{r^2}{2} \right] - H\beta \left[\frac{r+2}{2}, \frac{r}{2} \right] \right] \quad (28)$$

V. Order Statistics for TSS t D

Sample values such as the smallest, largest, or middle observation from a random sample provide important information. Order Statistics could be used to determine the distribution of the smallest

(minimum) order statistic and the largest (maximum) order statistic of a given distribution. Let V_1, V_2, \dots, V_n denote n -independent random sample from a distribution function $F(v)$ and probability density function, $f(v)$, then v_1, v_2, \dots, v_n represent the order sample arrangement and the pdf of $v_{(n)}$ is given by:

$$f_{i:n}(v) = \frac{n!}{(i-1)!(n-i)!} f(v) F(v)^{i-1} [1-F(v)]^{n-i} ; \text{ for } i=1, 2, \dots, n. \quad (29)$$

For simplicity, $[1-F(v)]^{n-i}$ in (29) can be expressed using the sum of a binomial series as

$\sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m [F(v)]^m$. By substitution the following is obtained:

$$f_{i:n}(v) = \frac{n!}{(i-1)!(n-i)!} f(v) \sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m F(v)^{m+i-1} \quad (30)$$

Now, making the substitution of $f(v)$ and $F(v)$ into (30) will yield the following;

$$f_{in}(v) = \frac{n!}{(i-1)!(n-i)!} \left[\frac{\Lambda}{2(\Lambda+v^2)^{\frac{3}{2}}} \left[1 - \alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right] \right] \times \sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m \left[(1+\alpha) \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) - \alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right)^2 \right]^{m+i-1} \quad (31)$$

Applying binomial expansion to the term $\left[(1+\alpha) \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) - \alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right)^2 \right]^{m+i-1}$ to get

$$\sum_{q=0}^{m+i-1} \binom{m+i-1}{q} \left[(1+\alpha) \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right]^q \left[-\alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right)^2 \right]^{m+i-1-q} \text{ and when substituted into (31), the } f_{in}(v) \text{ becomes;}$$

$$f_{in}(v) = \frac{n!}{(i-1)!(n-i)!} \left[\frac{\Lambda}{2(\Lambda+v^2)^{\frac{3}{2}}} \left[1 - \alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right] \right] \times \sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m \sum_{q=0}^{m+i-1} \binom{m+i-1}{q} \left[(1+\alpha) \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right]^q \left[-\alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right)^2 \right]^{m+i-1-q} \quad (32)$$

Expanding the term $\left[(1+\alpha) \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right]^q$ binomially yields $\sum_{r=0}^q \binom{q}{r} \alpha^{q-r} \left[\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right]^q$ and on substitution into (32), $f_{in}(v)$ becomes,

$$f_{in}(v) = \frac{n!}{(i-1)!(n-i)!} \left[\frac{\Lambda}{2(\Lambda+v^2)^{\frac{3}{2}}} \left[1 - \alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right] \right] \times \sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m \sum_{q=0}^{m+i-1} \binom{m+i-1}{q} \sum_{r=0}^q \binom{q}{r} (-\alpha)^{m+i-1-q+r} \left[\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right)^2 \right]^{m+i-1-q+r} \quad (33)$$

$$f_{in}(v) = \frac{n!}{(i-1)!(n-i)!} \left[\frac{\Lambda}{2(\Lambda+v^2)^{\frac{3}{2}}} \left[1 - \alpha \left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right) \right] \right] H_{in}(-\alpha)^{m+i-1} \left[\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}} \right)^2 \right]^{m+i-1} \quad (34)$$

where, $H_{in} = \sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m \sum_{q=0}^{m+i-1} \binom{m+i-1}{q} \sum_{r=0}^q \binom{q}{r}$.

VI. Parameter Estimation Using Maximum Likelihood Estimation Technique

Let $l(\theta)$ be a parameter vector for the transmuted family of distributions. Consider a random variable $v \sim TSSStD(\Lambda, \alpha)$, then by definition, the likelihood function of v with PDF $f(v)$ is given as;

$$l(\theta) = l(f(v; \Lambda)) = \prod_{i=1}^n \left[\Lambda \left[\frac{\sqrt{(\Lambda+v^2)} - \alpha v}{\sqrt{(\Lambda+v^2)}} \right] 2(\Lambda+v^2)^{\frac{3}{2}} \right] \quad (35)$$

$$= \Lambda^n 2^{-n} \left[1 - \frac{\alpha \sum_{i=1}^n v}{\sum_{i=1}^n \sqrt{(\Lambda + v^2)}} \right] \sum_{i=1}^n (\Lambda + v^2)^{\frac{3}{2}} \quad (36)$$

The log of $l(\theta)$ is as follows;

$$\log(l(\theta)) = \log \left(\Lambda^n \left[\frac{\sum_{i=1}^n \sqrt{(\Lambda + v^2)} - \alpha^n \sum_{i=1}^n v}{\sum_{i=1}^n \sqrt{(\Lambda + v^2)}} \right] \right) - \log \left(2^n \sum_{i=1}^n (\Lambda + v^2)^{\frac{3}{2}} \right) \quad (37)$$

Differentiating equation (37) with respect to the parameters Λ and α will yield the estimates of the parameters. The differentiation w.r.t. Λ is obtained as

$$\frac{d[\log(l(\theta))]}{d\Lambda} = \frac{1 + n\alpha^n \sum_{i=1}^n v \left(\sum_{i=1}^n (\Lambda + v^2)^{\frac{3}{2}} \right)}{n} \left[\frac{\sum_{i=1}^n \sqrt{(\Lambda + v^2)} - \alpha^n \sum_{i=1}^n v}{\sum_{i=1}^n \sqrt{(\Lambda + v^2)}} \right]^{-1} - \left(3 \times 2^{n-1} \sum_{i=1}^n (\Lambda + v^2)^{\frac{3}{2}} \right)^{-1} \sum_{i=1}^n (\Lambda + v^2)^{\frac{5}{2}} = 0.$$

The result obtained shows that the parameter Λ does not exist in a closed form. A numerical estimate for the parameter will be obtained using R-software. Differentiation w.r.t. α gives

$$\frac{d[\log(l(\theta))]}{d\alpha} = - \frac{n\alpha^{n-1} \sum_{i=1}^n v}{\sum_{i=1}^n \sqrt{(\Lambda + v^2)}} \left[\frac{\sum_{i=1}^n \sqrt{(\Lambda + v^2)} - \alpha^n \sum_{i=1}^n v}{\sum_{i=1}^n \sqrt{(\Lambda + v^2)}} \right]^{-1} = 0$$

$$\hat{\alpha} = \sqrt{\frac{\sum_{i=1}^n (\Lambda + v^2)^{\frac{1}{2}}}{\sum_{i=1}^n v}} \quad (38)$$

VII. Measures of Goodness-of-Fit Adopted

This section presents the measures used in model selection. They include the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Corrected AIC (CAIC), and Hannan-Quinn Information Criterion (HQIC).

$$AIC = 2k - 2(l) \quad (39)$$

$$BIC = k(\log_{10}(n)) - (2l) \quad (40)$$

$$CAIC = \frac{AIC(n-k-1) + 2k(k+1)}{(n-k-1)} \quad (41)$$

$$HQIC = 2k \log_{10}(\ln(n)) - 2(l) \quad (42)$$

where, l is the log-likelihood value, n is the in-sample size, and k is the parameter.

III. Results

I. Stability Study via Monte Carlo Simulations

In this section, we examined the stability of the probability model with an increase in sample size using simulation. The parameters were fixed at (0.6, 0.5) and (0.3, 0.7). Utilizing the quantile function presented in Equation (8), a random sample of sizes—30, 75, 300, 500, and 1000—was generated. The measures used for assessing the models were the Average Absolute Bias (AAB),

Mean Square Error (MSE), and Root Mean Square Error (RMSE), respectively. These measures were calculated using the expressions given below:

$$AAB = \sum_{q=1}^Q \frac{|\hat{u}_q - u|}{q} \tag{43}$$

$$MSE = \sum_{q=1}^Q \frac{(\hat{u}_q - u)^2}{q} \tag{44}$$

$$RMSE = \sqrt{MSE} \tag{45}$$

In accordance with the central limit theorem, an increase in sample sizes is expected to result in the reduction of estimation errors, approaching zero. Analyzing the results of the simulation as presented in Table 1, it becomes evident that this holds true for the new model. An increase in sample sizes corresponds to a decrease in bias and mean square error, as demonstrated.

Table 1: The AAB, MSE and RMSE

Size (n)	Parameter	Parameter Value	MLE	AAB	MSE	RMSE
30	λ	0.21	-2.929E-06	0.21000293	0.04410123	0.21000293
30	α	0.72	-0.3688692	1.08886921	1.18563616	1.08886921
100	λ	0.21	0.00062679	0.20937321	0.04383714	0.20937321
100	α	0.72	-0.378557	1.09855699	1.20682746	1.09855699
300	λ	0.21	0.08369549	0.12630451	0.01595283	0.12630451
300	α	0.72	0.74825993	0.02825993	0.00079862	0.02825993
500	λ	0.21	0.2257876	0.0157876	0.00024925	0.0157876
500	α	0.72	0.57707861	0.14292139	0.02042652	0.14292139

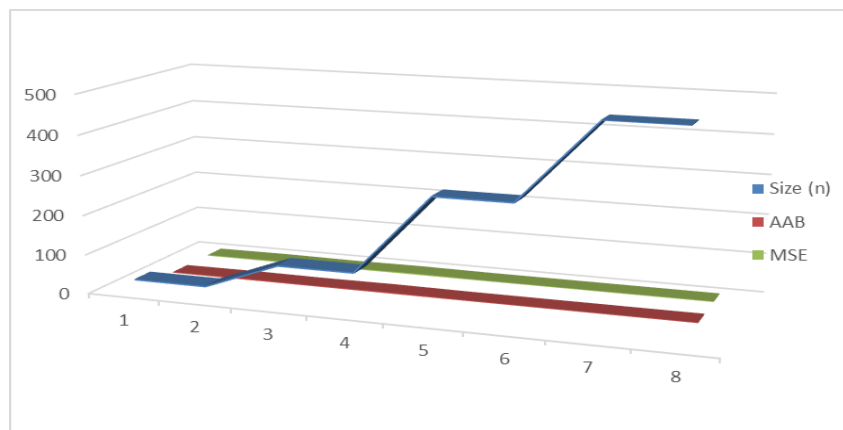


Figure 4: The AAB and MSE

As illustrated in the plot of the AAB and MSE presented in Figure 4, it can be observed that the probability model is well-behaved. This was due to the decay in the value of the AAB and MSE. The larger the sizes of the sample, the better the estimates are, the smaller the error and the more consistent the parameters are.

II. Application to Real-Life Data

This section provides an application of the model to real data. Other competing models were fitted to the data, and the goodness of fit of these models was assessed using various statistical information criteria. This illustration involves four data sets. The first and second datasets were sourced from [21]. The first dataset pertains to the duration of symptom decrease or disappearance

in patients with bladder cancer, measured in months for one hundred and twenty-eight patients. The second dataset focuses on the response time of patients to treatments, measured in minutes from the moment the treatment was administered.

First data set: 0.08, 2.09, 2.73, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.22, 3.52, 4.98, 6.99, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 15.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.93, 8.65, 12.63, 22.69

Second data set: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0

The third and fourth dataset were both sourced from [22]. The third dataset consisted of three hundred and forty-six measures of nicotine taken as obtained from different cigarette product category. The fourth data was on the windshield of an aircraft. The data comprises of one hundred and fifty-three measurements, of which eighty-eight were categorized as failed windshields and the remaining sixty-five were service times of windshields that were in good condition at the time of the observation. 1000h is the measuring unit.

Third data set: 1.3, 1.0, 1.2, 0.9, 1.1, 0.8, 0.5, 1.0, 0.7, 0.5, 1.7, 1.1, 0.8, 0.5, 1.2, 0.8, 1.1, 0.9, 1.2, 0.9, 0.8, 0.6, 0.3, 0.8, 0.6, 0.4, 1.1, 1.1, 0.2, 0.8, 0.5, 1.1, 0.1, 0.8, 1.7, 1.0, 0.8, 1.0, 0.8, 1.0, 0.2, 0.8, 0.4, 1.0, 0.2, 0.8, 1.4, 0.8, 0.5, 1.1, 0.9, 1.3, 0.9, 0.4, 1.4, 0.9, 0.5, 1.7, 0.9, 0.8, 0.8, 1.2, 0.9, 0.8, 0.5, 1.0, 0.6, 0.1, 0.2, 0.5, 0.1, 0.1, 0.9, 0.6, 0.9, 0.6, 1.2, 1.5, 1.1, 1.4, 1.2, 1.7, 1.4, 1.0, 0.7, 0.4, 0.9, 0.7, 0.8, 0.7, 0.4, 0.9, 0.6, 0.4, 1.2, 2.0, 0.7, 0.5, 0.9, 0.5, 0.9, 0.7, 0.9, 0.7, 0.4, 1.0, 0.7, 0.9, 0.7, 0.5, 1.3, 0.9, 0.8, 1.0, 0.7, 0.7, 0.6, 0.8, 1.1, 0.9, 0.9, 0.8, 0.8, 0.7, 0.7, 0.4, 0.5, 0.4, 0.9, 0.9, 0.7, 1.0, 1.0, 0.7, 1.3, 1.0, 1.1, 1.1, 0.9, 1.1, 0.8, 1.0, 0.7, 1.6, 0.8, 0.6, 0.8, 0.6, 1.2, 0.9, 0.6, 0.8, 1.0, 0.5, 0.8, 1.0, 1.1, 0.8, 0.8, 0.5, 1.1, 0.8, 0.9, 1.1, 0.8, 1.2, 1.1, 1.2, 1.1, 1.2, 0.2, 0.5, 0.7, 0.2, 0.5, 0.6, 0.1, 0.4, 0.6, 0.2, 0.5, 1.1, 0.8, 0.6, 1.1, 0.9, 0.6, 0.3, 0.9, 0.8, 0.8, 0.6, 0.4, 1.2, 1.3, 1.0, 0.6, 1.2, 0.9, 1.2, 0.9, 0.5, 0.8, 1.0, 0.7, 0.9, 1.0, 0.1, 0.2, 0.1, 0.1, 1.1, 1.0, 1.1, 0.7, 1.1, 0.7, 1.8, 1.2, 0.9, 1.7, 1.2, 1.3, 1.2, 0.9, 0.7, 0.7, 1.2, 1.0, 0.9, 1.6, 0.8, 0.8, 1.1, 1.1, 0.8, 0.6, 1.0, 0.8, 1.1, 0.8, 0.5, 1.5, 1.1, 0.8, 0.6, 1.1, 0.8, 1.1, 0.8, 1.5, 1.1, 0.8, 0.4, 1.0, 0.8, 1.4, 0.9, 0.9, 1.0, 0.9, 1.3, 0.8, 1.0, 0.5, 1.0, 0.7, 0.5, 1.4, 1.2, 0.9, 1.1, 0.9, 1.1, 1.0, 0.9, 1.2, 0.9, 1.2, 0.9, 0.5, 0.9, 0.7, 0.3, 1.0, 0.6, 1.0, 0.9, 1.0, 1.1, 0.8, 0.5, 1.1, 0.8, 1.2, 0.8, 0.5, 1.5, 1.5, 1.0, 0.8, 1.0, 0.5, 1.7, 0.3, 0.6, 0.6, 0.4, 0.5, 0.5, 0.7, 0.4, 0.5, 0.8, 0.5, 1.3, 0.9, 1.3, 0.9, 0.5, 1.2, 0.9, 1.1, 0.9, 0.5, 0.7, 0.5, 1.1, 1.1, 0.5, 0.8, 0.6, 1.2, 0.8, 0.4, 1.3, 0.8, 0.5, 1.2, 0.7, 0.5, 0.9, 1.3, 0.8, 1.2, 0.9

Fourth data set: 0.040, 1.866, 2.385, 3.443, 0.301, 1.876, 2.481, 3.467, 0.309, 1.899, 2.610, 3.478, 0.557, 1.911, 2.625, 3.578, 0.943, 1.912, 2.632, 3.595, 1.070, 1.914, 2.646, 3.699, 1.124, 1.981, 2.661, 3.779, 1.248, 2.010, 2.688, 3.924, 1.281, 2.038, 2.823, 4.035, 1.281, 2.085, 2.890, 4.121, 1.303, 2.089, 2.902, 4.167, 1.432, 2.097, 2.934, 4.240, 1.480, 2.135, 2.962, 4.255, 1.505, 2.154, 2.964, 4.278, 1.506, 2.190, 3.000, 4.305, 1.568, 2.194, 3.103, 4.376, 1.615, 2.223, 3.114, 4.449, 1.619, 2.224, 3.117, 4.485, 1.652, 2.229, 3.166, 4.570, 1.652, 2.300, 3.344, 4.602, 1.757, 2.324, 3.376, 4.663.

Five models were fitted to the above four datasets using the *Adequacy Model* package in R [23], these models include; Exponentiated Ailamujia distribution (EAD), Exponentiated Exponential distribution (EED), Exponentiated Weibull (EWD), the Logistic distribution (LD) and the transmuted skew student t distribution (TSS t D) respectively. The resulting fitted models selected on the basis of Akaike Information Criterion, (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (AIC) and Hannan-Quinn information criterion, (HQIC).

Table 2: The AIC, CAIC, BIC, HQIC and MLE of the First data

Models	AIC	BIC	CAIC	HQIC	MLE	Rank
TSS <i>t</i> D	-642.6124	-636.9719	-634.9719	-627.3313	0.0006 1.1348	1
EED	799.7295	805.3700	799.8286	802.0208	1.2682 0.1041 0.4493	3
EWD	275.1219	282.4499	275.4182	278.0694	1.6494 1.7404	2
EAD	7285.8160	7291.4570	7285.9150	7288.1080	1.4819 1.8290	5
LD	1080.3360	1085.9770	1080.4350	1082.6270	1.9978 1.8904	4

Table 2 to Table 5 presents the model estimates for each of the datasets. The results revealed that the model with the smallest measure of the entire information criterion was the TSS*t*D. The ranks for the performance of the models were based on the information criteria of each of the models. From the results obtained, for the five models estimated, the TSS*t*D was the models with the best fit.

Table 3: The AIC, CAIC, BIC, and HQIC of the second data

Models	AIC	BIC	CAIC	HQIC	MLE	Rank
TSS <i>t</i> D	-642.6124	-640.621	-638.621	-634.6295	9.3568e-03 -7.6898e-16	1
EED	36.3450	38.3365	37.0509	36.7338	54.366966 2.172273 0.6234	2
EWD	46.9561	49.9433	48.4561	47.5392	1.7261 1.7733	3
EAD	48.3887	50.3802	49.0946	48.7775	1.8667 0.7558	4
LD	1204.3680	1210.0080	1204.4670	1206.6590	1.7387 1.7264	5

Table 4: The AIC, CAIC, BIC, and HQIC of the Third data

Models	AIC	BIC	CAIC	HQIC	MLE	Ranks
TSS <i>t</i> D	-642.6124	-634.9196	-632.9196	-623.2267	-0.1381 1.4114	1
EED	325.3694	333.0623	325.4044	328.4327	6.0022 2.3477 1.3488	4
EWD	261.0285	272.5679	261.0987	265.6235	1.7973 1.5733	2
EAD	300.5637	308.2566	300.5987	303.6271	1.9331 1.5668	3
LD	293.6612	301.3540	293.6961	296.7245	0.9789 0.2148	5

Table 5: The AIC, CAIC, BIC, and HQIC of the fourth data

Models	AIC	BIC	CAIC	HQIC	MLE	Ranks
TSS t D	-347.3415	-342.4562	-340.4562	-333.5709	-0.1613 1.1483	1
EED	292.3541	297.2394	292.5004	294.3191	3.6582 0.6240 0.4305	4
EWD	275.2268	282.5548	275.5231	278.1743	1.7537 1.3413	2
EAD	283.6236	288.5089	283.7699	285.5886	1.7293 0.5065	3
LD	412.6227	417.5080	412.7690	414.5877	0.8074 1.3006	5

IV. Conclusion

This research paper presented a novel two-parameter distribution known as the Transmuted Skew Student t distribution. Some of the statistical and reliability properties for the TSS t D were derived and they included the survival function, the r^{th} moment, the hazard function, the mean, the quantile function, the moment generating function, the characteristic function and the order statistics. Before application to real dataset, a Monte-Carlo simulation study was conducted to assess the stability of the model with more sample sizes. The results revealed that the model was consistent with increase in the number of samples. The new PDF was applied to four different real datasets. Using information criterions, it was found that TSS t D performs better than other competing models.

References

- [1] David, I. J., Mathew, S., & Falgore, J. Y. (2024). Reliability Analysis with New Sine Inverted Exponential Distribution: Properties, Simulation and Application. *European Journal of Statistics*, 4, 1-14.
- [2] David, I. J., Mathew, S., & Eghwerido, J. T. (2023). Reliability Analysis with New Sine Inverse Rayleigh Distribution. *Journal of Reliability and Statistical Studies*, 16(2), 255-268.
- [3] David, I. J., Asiribo, O. E., Dikko, H. G., & Ikwuoche, P. O. (2023). Johnson-Schumacher Split-Plot Design Modelling of Rice Yield. *Biometrical Letters*, 60(1), 37-52.
- [4] David, I. J., Asiribo, O. E., & Dikko, H. G. (2023). Nonlinear Split-Plot Design Modeling and Analysis of Rice Varieties Yield. *Scientific African*, 19(9):e01444.
- [5] David, I. J., Asiribo, O. E., & Dikko, H. G. (2022). A Weibull Split-Plot Design Model and Analysis. *Thailand Statistician*, 20(2), 420-434.
- [6] David, I. J., Asiribo, O. E., & Dikko, H. G. (2022). A Bertalanffy-Richards Split-Plot Design Model and Analysis. *Journal of Statistical Modeling and Analysis*, 4(1), 56-71.
- [7] Gupta, R. D., & Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1), 117-130.
- [8] Kumar, D., Singh, U., & Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients data. *Journal of Statistics Applications & Probability*, 4(3), 417.
- [9] Mahmood, Z., Chesneau, C., & Tahir, M. H. (2019). A new sine-G family of distributions: properties and applications. *Bull. Comput. Appl. Math.*, 7(1), 53-81.
- [10] Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1), 63-79.
- [11] Ashour S. K. & Eltehiwy M. A., (2013). Transmuted exponentiated Lomax distribution. *Australian Journal of Basic and Applied Sciences*, 7(7), 658-667.
- [12] Ashour, S. K., & Eltehiwy, M. A. (2013). Transmuted lomax distribution. *American Journal of Applied Mathematics and Statistics*, 1(6), 121-127.
- [13] Mahmoud, M. R., & Mandouh, R. M. (2013). On the transmuted Fréchet distribution. *Journal of Applied*

- Sciences Research*, 9(10), 5553-5561.
- [14] Hussian, M. A. (2014). Transmuted exponentiated gamma distribution: A generalization of the exponentiated gamma probability distribution. *Applied Mathematical Sciences*, 8(27), 1297-1310.
- [15] Elbatal, I., & Aryal, G. (2013). On the transmuted additive weibull distribution. *Austrian Journal of statistics*, 42(2), 117-132.
- [16] Khan, M. S., King, R., & Hudson, I. L. (2017). Transmuted Generalized Gompertz distribution with application. *Journal of Statistical Theory and Applications*, 16(1), 65-80.
- [17] Al-Babtain, A., Fattah, A. A., Ahmed, A. H. N., & Merovci, F. (2017). The Kumaraswamy-transmuted exponentiated modified Weibull distribution. *Communications in Statistics-Simulation and Computation*, 46(5), 3812-3832.
- [18] Saraçoğlu, B., & Taniş, C. (2021). A new lifetime distribution: transmuted exponential power distribution. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 70(1), 1-14.
- [19] Shaw, W. T., & Buckley, I. R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*.
- [20] Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Volume 2, Second Edition. New York: Wiley.
- [21] Jones M. C., (2001). A skew t distribution. In *Probability and Statistical Models with Applications* (eds C. A. Charalambides, M. V. Koutras and N. Balakrishnan), pp. 269–277. London: Chapman and Hall.
- [22] Jones M. C. & Faddy M., (2003). A skew extension of the t -distribution, with applications.," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(1)159-174.
- [23] Rather, A. A., Subramanian, C., Al-Omari, A. I., & Alanzi, A. R. (2022). Exponentiated Ailamujia distribution with statistical inference and applications of medical data. *Journal of Statistics and Management Systems*, 25(4), 907-925.

SINE-TOPP-LEONE EXPONENTIATED G FAMILY OF DISTRIBUTIONS: PROPERTIES, SURVIVAL REGRESSION AND APPLICATION

A. M. Isa¹, S. I. Doguwa², B. B. Alhaji³ and H. G. Dikko⁴

¹Department of Mathematics and Computer Science, Borno State University, Maiduguri, Nigeria
alhajimoduisa@bosu.edu.ng

^{2,4}Department of Statistics, Ahmadu Bello University Zaria, Kaduna State, Nigeria
sidoguwa@gmail.com
hgdkko@gmail.co

³Nigerian Defense Academy, Kaduna State, Nigeria
bbukar@nda.edu.ng

Abstract

In this research article, we have introduced a new class of continuous probability distributions known as the Sine Topp-Leone Exponentiated-G family of distributions. This newly proposed family exhibits a higher degree of flexibility compared to some of the established distribution families. The various models within this family find wide-ranging applications in fields such as physics, engineering, and medicine. Some statistical properties of the Sine Topp-Leone Exponentiated-G family of distributions such as moments, moment generating function, quantile function and order statistics are derived. Two special models were also presented and studies. Maximum likelihood estimation method was used to estimate parameters of the models. The consistency of the proposed family was determined using simulation studies. Two real life datasets were analyzed to show the flexibility of the proposed model and the results of the analysis showed that, the proposed model was more efficient and best fit the data sets than its competitors.

Keywords: Sine-G Family, Topp-Leone Exponentiated G, Survival Analysis, Survival Regression, Maximum Likelihood Estimate

1. Introduction

The Topp-Leone Distribution is a statistical concept that finds its roots in probability theory and data analysis. It was developed by [1] in 2015 and it is a powerful tool in the field of statistics for modeling and understanding random variables with various applications across different domains. It is a relatively recent addition to the family of probability distributions in statistics and has gained prominence for its adaptability in modeling various types of data with flexibility and precision. This distribution offers a valuable tool for statisticians, data scientists, and researchers in diverse fields, enabling them to capture the underlying characteristics of data sets that may not conform to traditional distribution assumptions. The Topp Leone Distribution is particularly well-suited for modeling data with heavy tails, which means it can effectively describe observations that exhibit extreme values or outliers. This characteristic is especially important in fields like finance, where extreme events can have significant consequences, and in environmental science, where rare but impactful events need to be accounted for in risk assessment.

Recently, some researchers have developed numerous families of Topp-Leone Distribution which include: Topp-Leone G Family of Distribution by [2], Topp-leone Marshal Olkin G by [3], Transmuted Topp-Leone G by [4], Topp-Leone Exponentiated G Family by [5], Topp-Leone Odd Lindley G by [6], Odd Log Logistic Topp-Leone by [7], Frechet Topp-Leone G Family by [8], Topp-Leone Odd Log Logistic Family by [9], Type II generalized Topp-Leone G Family by [10], Type II Exponentiated Half-Logistic Topp-Leone G by [11], Topp-Leone Marshal Olkin G by [12], Type II Topp-Leone G by [13], the Weibull Topp-Leone G by [14], Odd Weibull Topp-Leone G by [15], Topp Leone Odd Burr III G by [16], The Burr III Topp-Leone G by [17], Topp-Leone Gompertz G by [18], Topp-Leone Exponential G by [19] and Topp-Leone Generalized Half-Log Logistic G by [20].

In order hand, the Sine-G family of probability distributions is a class of continuous probability distributions that is often used in statistical modeling and data analysis. This family is characterized by its flexibility and ability to capture a wide range of data patterns, making it a valuable tool for statisticians and data scientists. The PDF of a Sine-G distribution is defined in terms of the sine function, which introduces oscillatory behavior into the distribution. This oscillatory behavior can be adjusted by varying the distribution's parameters, allowing it to fit data with different shapes and characteristics. One of the notable features of the Sine-G family is its ability to model data with heavy tails, which means it can effectively describe extreme or outlier values in a dataset. This makes it useful in fields such as finance, where extreme events can have a significant impact on investment portfolios and risk assessment. The Sine-G family is also capable of modeling data with various degrees of skewness and kurtosis, providing a versatile tool for capturing complex data patterns that may not conform to traditional distribution assumptions like the normal distribution.

Some of the recent development of the Sine G family include: Sine Topp-Leone G by Al-[21], the New Sine G Family by [22], the Sine Kumaraswamy G by [23], Exponentiated Sine G by [24], Transmuted Sine G by [25], Sine Marshall–Olkin G by [26] and Sine Inverse Lomax G by [27]. These developments of flexible families of distributions through innovative transformations, as mentioned in this research article, reflects the dynamic nature of statistical research. Such advancements hold promise for improving the accuracy and applicability of statistical models in diverse domains and addressing the complexities of real-world data.

2. Methods

2.1 The Sine-G Family of Probability Distribution

Let $h(x; \xi)$ and $H(x; \xi)$ be the pdf and cdf of a Univariate continuous distribution, then, the Sine-G family of probability distribution according to [28] is defined by:

$$F(x; \xi) = \int_0^{\frac{\pi}{2} H(x; \xi)} \cos t dt = \sin \left\{ \frac{\pi}{2} H(x; \xi) \right\} \quad (1)$$

with corresponding pdf given by:

$$f(x; \xi) = \frac{\pi}{2} h(x; \xi) \cos \left\{ \frac{\pi}{2} H(x; \xi) \right\} \quad (2)$$

where $H(x; \xi)$ and $h(x; \xi)$ are the cdf and the pdf of any baseline distribution with vector parameter ξ .

2.2 Topp-Leone Exponentiated-G Family of Distributions

The cdf of the Topp-Leone Exponentiated G family of distribution according to [4] is given by:

$$F(x; \alpha, \theta, \xi) = \left\{ 1 - \left[1 - G(x; \xi)^\alpha \right]^2 \right\}^\theta \quad (3)$$

with corresponding pdf defined by:

$$f(x; \alpha, \theta, \xi) = 2\alpha\theta g(x; \xi)G(x; \xi)^{\alpha-1} \left[1 - G(x; \xi)^\alpha \right] \left\{ 1 - \left[1 - G(x; \xi)^\alpha \right]^2 \right\}^{\theta-1} \quad (4)$$

where α is a shape parameter, $g(x; \xi)$ and $G(x; \xi)$ are pdf and cdf of any baseline distribution respectively and ξ is a vector parameter of the baseline distribution.

2.3 The Proposed Sine Topp-Leone Exponentiated G Family of Distributions

The cdf of the new Sine Topp-Leone Exponentiated G Family is given by:

$$F(x; \alpha, \theta, \xi) = \sin \left\{ \frac{\pi}{2} \left[1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right]^\theta \right\} \quad (5)$$

with corresponding pdf given by:

$$f(x; \alpha, \theta, \xi) = \frac{\pi}{2} 2\alpha\theta g(x; \xi)G(x; \xi)^{\alpha-1} \left[1 - G(x; \xi)^\alpha \right] \left\{ 1 - \left[1 - G(x; \xi)^\alpha \right]^2 \right\}^{\theta-1} \times \cos \left\{ \frac{\pi}{2} \left[1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right]^\theta \right\} \quad (6)$$

The survival function $S(x)$, hazard function $h(x)$, reversed hazard function $r(x)$ and the quantile functions $Q(x)$ of the STLE-G are presented in equation (7) to (10).

$$S(x) = 1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left(1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^\alpha \right)^2 \right]^\theta \right\} \quad (7)$$

$$h(x) = \frac{\frac{\pi}{2} 2\alpha\theta g(x; \xi)G(x; \xi)^{\alpha-1} \left[1 - G(x; \xi)^\alpha \right] \left\{ 1 - \left[1 - G(x; \xi)^\alpha \right]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} \left[1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right]^\theta \right\}}{1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right]^\theta \right\}} \quad (8)$$

$$r(x) = \frac{\frac{\pi}{2} 2\alpha\theta g(x; \xi)G(x; \xi)^{\alpha-1} \left[1 - G(x; \xi)^\alpha \right] \left\{ 1 - \left[1 - G(x; \xi)^\alpha \right]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} \left[1 - \left(1 - G(x; \xi)^\alpha \right)^2 \right]^\theta \right\}}{\sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - G(x; \xi)^\alpha \right]^2 \right]^\theta \right\}} \quad (9)$$

$$x = \Phi(u) = G^{-1} \left\{ 1 - \left[1 - \left(1 - \frac{\sin^{-1}(u)}{\pi/2} \right)^{\frac{1}{\theta}} \right]^{\frac{1}{2}} \right\}^{\frac{1}{\alpha}} \quad (10)$$

where $G^{-1}(x, \xi)$ is the quantile function of the baseline distribution $G(x; \xi)$.

2.4 Expansion of Density

The pdf and the cdf of the Sine Topp-Leone Exponentiated G family can be expanded using power series expansion as follows:

$$f(x; \alpha, \theta, \xi) = \frac{\pi}{2} 2\alpha\theta g(x; \xi) G(x; \xi)^{\alpha-1} [1 - G(x; \xi)^\alpha] \left\{ 1 - [1 - G(x; \xi)^\alpha]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} [1 - (1 - G(x; \xi)^\alpha)^2]^\theta \right\}$$

$$\cos \left\{ \frac{\pi}{2} [1 - [1 - G(x)^\alpha]^2]^\theta \right\} = \sum_{i=0}^{\infty} \frac{(-1)^i \pi^{2i}}{(2i)! 2^{2i}} [1 - [1 - G(x)^\alpha]^2]^{\theta i}$$

$$[1 - [1 - G(x)^\alpha]^2]^{\theta i + \theta - 1} = \sum_{j=0}^{\infty} (-1)^j \binom{\theta i + \theta - 1}{j} [1 - G(x)^\alpha]^{2j}$$

$$[1 - G(x)^\alpha]^{2j+1} = \sum_{k=0}^{\infty} (-1)^k \binom{2j+1}{k} G(x)^\alpha$$

$$f(x) = \sum_{i,j,k=0}^{\infty} \frac{2\alpha\theta(-1)^{i+j+k}}{(2i)!} \frac{\pi^{2i+1}}{2^{2i+1}} \binom{\theta i + \theta - 1}{j} \binom{2j+1}{k} g(x) G(x)^{\alpha(k+1)-1}$$

Let $\Psi = \frac{2\alpha\theta(-1)^{i+j+k}}{(2i)!} \frac{\pi^{2i+1}}{2^{2i+1}} \binom{\theta i + \theta - 1}{j} \binom{2j+1}{k}$

Therefore, the reduced form of the pdf is given by:

$$f(x) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi g(x) G(x)^{\alpha(k+1)-1} \tag{11}$$

The cdf can also be expanded as follows:

$$\sin \left\{ \frac{\pi}{2} [1 - [1 - G(x)^\alpha]^2]^\theta \right\} = \sum_{l=0}^{\infty} \frac{(-1)^l \pi^{2l+1}}{(2l+1)! 2^{2l+1}} [1 - [1 - G(x)^\alpha]^2]^{2l+1}$$

$$[1 - [1 - G(x)^\alpha]^2]^{2l+1} = \sum_{m=0}^{\infty} (-1)^m \binom{2l+1}{m} [1 - G(x)^\alpha]^{2m}$$

$$[1 - G(x)^\alpha]^{2m} = \sum_{n=0}^{\infty} (-1)^n \binom{2m}{n} G(x)^{\alpha n}$$

$$F(x) = \sum_{l,m,n=0}^{\infty} \frac{(-1)^{l+m+n} \pi^{2l+1}}{(2l+1)! 2^{2l+1}} \binom{2l+1}{m} \binom{2m}{n} G(x)^{\alpha n}$$

Let $\Phi = \frac{(-1)^{l+m+n} \pi^{2l+1}}{(2l+1)! 2^{2l+1}} \binom{2l+1}{m} \binom{2m}{n}$

Therefore, the reduced form of the cdf is given as:

$$F(x) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \Phi G(x)^{\alpha n} \tag{12}$$

2.5 Mathematical Properties

2.5.1 The Moment

Moments is used to study many important properties of distribution such as dispersion, tendency, skewness and kurtosis. The r^{th} moments of the Sine Type II Topp Leone G family of distribution is obtained as follow:

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x) dx$$

Therefore, the moment of the Sine Topp-Leone Exponentiated G is obtained as follows:

$$\mu_r' = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi \int_0^{\infty} g(x) G(x)^{\alpha(k+1)-1} dx$$

$$\text{Let } \Theta = \int_0^{\infty} g(x) G(x)^{\alpha(k+1)-1} dx,$$

Therefore, the r th moment of the Sine Topp-Leone Exponentiated is given by:

$$\mu_r' = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi \Theta \tag{13}$$

2.5.2 Moment Generating Function

The moment generating function of a random variable X is defined as $E(e^{tx})$.

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x; \alpha, \xi) dx$$

We say that mgf of X exists, if there exists a positive constant a such that $M_t(x)$ is finite for all $s \in [-a, a]$. The moment generating function of the Sine Topp-Leone Exponentiated G family of Distributions is given by:

$$M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} g(x) G(x)^{\alpha(k+1)-1} dx$$

$$M_x(t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi \int_0^{\infty} e^{tx} g(x) G(x)^{\alpha(k+1)-1} dx$$

$$\text{Let } \Upsilon = \int_0^{\infty} e^{tx} g(x) G(x)^{\alpha(k+1)-1} dx$$

Therefore, the moment generating function of the Sine Topp-Leone Exponentiated G is given by:

$$M_x(t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi \Upsilon \tag{14}$$

2.5.3 Entropy

Entropy is used as a measure of information or uncertainty, which present in a random observation of its actual population. There will be the greater uncertainty in the data if the value of entropy is large. For some probability distributions expression, the differential entropy is considered mostly effective. It can be derived using the formula:

$$I_{\theta}(x) = \frac{1}{1-\theta} \log \int_{-\infty}^{\infty} f(x)^{\theta} dx$$

The entropy for the Sine Topp-Leone Exponentiated G family of distributions is given by:

$$f(x)^{\theta} = \left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi g(x) G(x)^{\alpha(k+1)-1} \right)^{\theta}$$

$$f(x)^{\theta} = \left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi \right)^{\theta} \left(g(x) G(x)^{\alpha(k+1)-1} \right)^{\theta}$$

Let $\Omega = g(x)G(x)^{\alpha(k+1)-1}$

Therefore, the entropy is given as:

$$I_{\theta}(x) = \frac{1}{1-\theta} \left[\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi \right)^{\theta} \int_0^{\infty} \Omega^{\theta} dx \right] \quad (15)$$

2.5.6 Order Statistics

Let x_1, x_2, \dots, x_n be a random sample of size n from a continuous population having a pdf $f(x)$ and cdf $F(x)$, Let $X_{1:n} \leq X \leq X_{n:n}$ be the corresponding order statistics (OS). [29] defined the pdf of $X_{1:n}$ that is the i^{th} order statistics by:

$$f_{p,q}(x) = \frac{f(x)}{B(p, q-p+1)} (F(x))^{p-1} (1-F(x))^{q-p}$$

The order statistics of the Sine Topp-Leone Exponentiated G is given by:

$$f_{p,q}(x) = \frac{\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \Psi g(x) G(x)^{\alpha(k+1)-1}}{B(p, q-p+1)} \left(\sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \Phi G(x)^{\alpha n} \right)^{p-1} \left(1 - \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \Phi G(x)^{\alpha n} \right)^{q-p} \quad (16)$$

2.6 Parameter Estimation

2.6.1 Maximum Likelihood Estimate

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample of size n from the Sine Type II Topp-Leone G family of distribution. Then, the likelihood function of the Sine Type II Topp-Leone G family is derived as follows:

$$\ell = n \log(\pi) + n \log(\alpha) + n \log(\theta) + \sum_{i=1}^n \log g(x; \alpha, \theta, \xi) + (\alpha - 1) \sum_{i=1}^n \log G(x; \alpha, \theta, \xi) + \sum_{i=1}^n \log(1 - G(x; \alpha, \theta, \xi)^\alpha) \\
 (\theta - 1) \sum_{i=0}^n \log \left[1 - (1 - G(x; \alpha, \theta, \xi)^\alpha)^2 \right] + \sum_{i=0}^n \log \cos \left\{ \frac{\pi}{2} \left[1 - (1 - G(x; \alpha, \theta, \xi)^\alpha)^2 \right]^\theta \right\} \quad (17)$$

Differentiating the likelihood function in equation (17) with respect to α gives the following expression:

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \log G(x) - \sum_{i=1}^n \frac{G(x)^\alpha \ln G(x)}{[1 - G(x)^\alpha]} + \sum_{i=1}^n \pi \theta (1 - G(x)^\alpha) \left[1 - [1 - G(x)^\alpha]^2 \right]^{\theta-1} \\
 \times G(x)^\alpha \ln G(x) \tan \left\{ \frac{\pi}{2} \left[1 - [1 - G(x)^\alpha]^2 \right]^\theta \right\} \quad (18)$$

Differentiating the likelihood function in equation (17) with respect to θ gives:

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + \sum_{i=0}^n \log \left[1 - (1 - G(x)^\alpha)^2 \right] - \sum_{i=0}^n \frac{\pi}{2} \tan \left\{ \frac{\pi}{2} \left[1 - (1 - G(x)^\alpha)^2 \right]^\theta \right\} \left[1 - (1 - G(x)^\alpha)^2 \right]^\theta \ln \left[1 - (1 - G(x)^\alpha)^2 \right] \quad (19)$$

Differentiating the likelihood function in equation (17) with respect to ξ gives:

$$\frac{\partial \ell}{\partial \xi} = \sum_{i=0}^n \frac{g'(x)}{g(x)} + (\alpha - 1) \sum_{i=0}^n \frac{g(x)}{G(x)} - \sum_{i=0}^n \frac{\alpha g(x) G(x)^{\alpha-1}}{(1 - G(x)^\alpha)} + (\theta - 1) \sum_{i=0}^n \frac{2\alpha g(x) G(x)^{\alpha-1} (1 - G(x)^\alpha)}{[1 - (1 - G(x)^\alpha)^2]} \\
 - \sum_{i=0}^n \pi \alpha \theta g(x) G(x)^{\alpha-1} (1 - G(x)^\alpha) \left[1 - (1 - G(x)^\alpha)^2 \right]^{\theta-1} \tan \left\{ \frac{\pi}{2} \left[1 - (1 - G(x)^\alpha)^2 \right]^\theta \right\} \quad (20)$$

The expression in equation (18), (19) and (20) are the maximum likelihood estimates of the parameters α, θ and the vector parameter ξ .

2.7 Special Models of STLE-G Family

Here, we consider two special models of the STLE-G family along with the plots of their density and hazard rate function.

2.7.1 Sine Topp-Leone Exponentiated Lomax (STLE-L) Distribution

Let the Lomax distribution be the baseline distribution with cdf and pdf defined by:

$$F(x) = 1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \quad (21)$$

And

$$f(x) = \frac{\gamma}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\gamma+1)} \tag{22}$$

where γ is a shape parameter and λ is a scale parameter, the cdf and pdf of the STLE-L distribution are respectively given by:

$$F(x; \alpha, \theta, \gamma, \lambda) = \sin \left\{ \frac{\pi}{2} \left[1 - \left(1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^\alpha \right)^2 \right]^\theta \right\} \tag{23}$$

$$f(x; \alpha, \theta, \gamma, \lambda) = \frac{\pi}{2} 2\alpha\theta \left(\frac{\gamma}{\lambda} \left(1 + \frac{x}{\lambda} \right)^{-(\gamma+1)} \right) \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha-1} \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^\alpha \right] \times \left\{ 1 - \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^\alpha \right]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} \left[1 - \left(1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^\alpha \right)^2 \right]^\theta \right\} \tag{24}$$

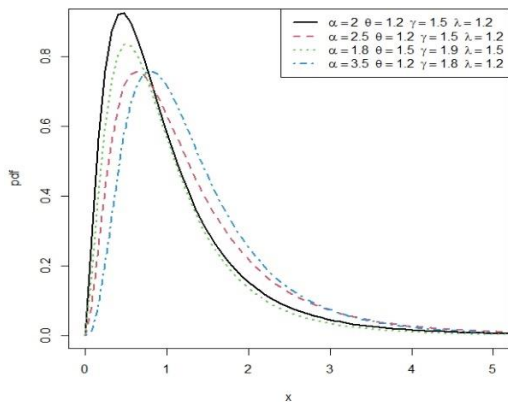


Figure 1: pdf plot of STLE-Lomax distribution

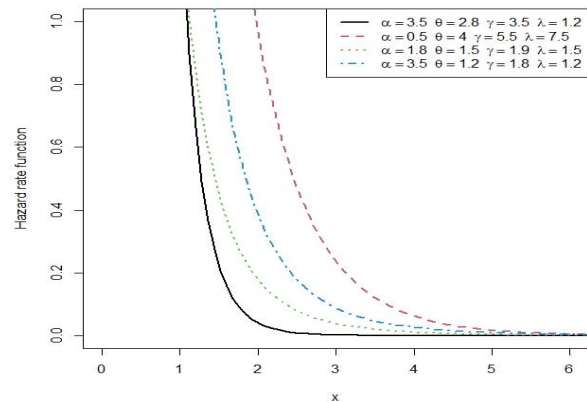


Figure 2: plot of the hrf of the STLE-L distribution

The survival function, hazard function, reverse hazard function and quantile function of the proposed STLE-Lomax distribution is presented below:

$$S(x) = 1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - \left(1 - \frac{x}{\lambda} \right)^\gamma \right)^\alpha \right]^2 \right]^\theta \right\}$$

$$h(x) = \frac{\frac{\pi}{2} 2\alpha\theta \left(\frac{\gamma}{\lambda} \left(1 + \frac{x}{\lambda} \right)^{-(\gamma+1)} \right) \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha-1} \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right] \left\{ 1 - \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right]^2 \right]^{\theta} \right\}}{1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right]^2 \right\}^{\theta}}$$

$$r(x) = \frac{\pi}{2} 2\alpha\theta \left(\frac{\gamma}{\lambda} \left(1 + \frac{x}{\lambda} \right)^{-(\gamma+1)} \right) \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha-1} \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right] \left\{ 1 - \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right]^2 \right\}^{\theta-1} \cot \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left[1 - \left(1 + \frac{x}{\lambda} \right)^{-\gamma} \right]^{\alpha} \right]^2 \right]^{\theta} \right\}$$

$$Q(u) = \lambda \left\{ \left[\left[1 - \left[1 - \left(1 - \left(\frac{\sin^{-1}(u)}{\pi/2} \right)^{\frac{1}{\theta}} \right)^{\frac{1}{2}} \right]^{\frac{1}{\alpha}} \right]^{\frac{1}{\gamma}} - 1 \right] \right\}$$

2.7.2 Sine Topp-Leone Exponentiated Weibull (STLE-W) Distribution

Let the Weibull distribution be the baseline distribution with cdf and pdf defined by:

$$f(x) = \frac{\gamma}{\lambda} \left(\frac{x}{\lambda} \right)^{\gamma-1} e^{-\left(\frac{x}{\lambda}\right)^{\gamma}} \tag{25}$$

$$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^{\gamma}} \tag{26}$$

where γ is a shape parameter and λ is a scale parameter, the cdf and pdf of the STLE-L distribution are respectively given by:

$$F(x; \alpha, \theta, \gamma, \lambda) = \sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^{\gamma}} \right)^{\alpha} \right]^2 \right]^{\theta} \right\} \tag{27}$$

$$f(x; \alpha, \theta, \gamma, \lambda) = \frac{\pi}{2} 2\alpha\theta \left[\frac{\gamma}{\lambda} \left(\frac{x}{\lambda} \right)^{\gamma-1} e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right] \left[1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right]^{\alpha-1} \left\{ 1 - \left[1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right]^\alpha \right\} \left\{ 1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right]^\theta \right\} \quad (28)$$

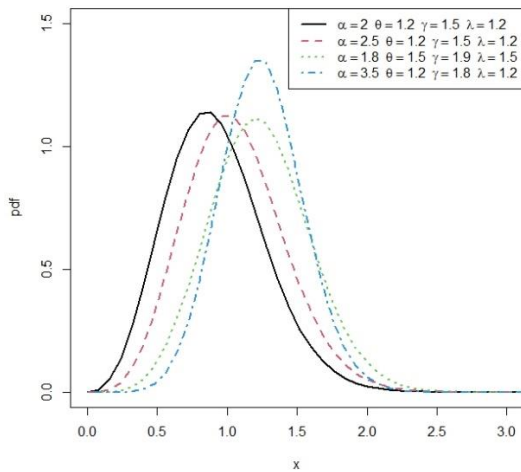


Figure 3: pdf plot of STLE-W distribution

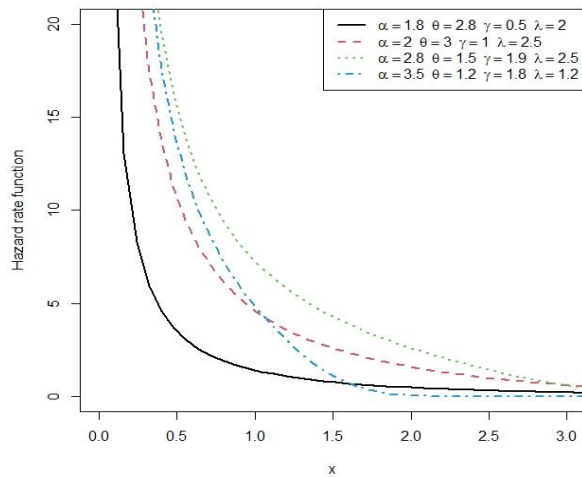


Figure 4: Plot of the hrf of the STLE-W Distribution

The survival function, hazard function, reverse hazard function and quantile function of the proposed STLE-Weibull distribution is presented below.

$$S(x) = 1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right]^\theta \right\}$$

$$h(x) = \frac{\frac{\pi}{2} 2\alpha\theta \left[\frac{\gamma}{\lambda} \left(\frac{x}{\lambda} \right)^{\gamma-1} e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right] \left[1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right]^{\alpha-1} \left\{ 1 - \left[1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right]^\alpha \right\} \left\{ 1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right\}^{\theta-1} \cos \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right]^\theta \right\}}{1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right]^\theta \right\}}$$

$$r(x) = \frac{\frac{\pi}{2} 2\alpha\theta \left[\frac{\gamma}{\lambda} \left(\frac{x}{\lambda} \right)^{\gamma-1} e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right] \left[1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right]^{\alpha-1} \left\{ 1 - \left[1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right]^\alpha \right\} \left\{ 1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right\}^{\theta-1} \cot \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right]^\theta \right\}}{1 - \sin \left\{ \frac{\pi}{2} \left[1 - \left[1 - \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\gamma} \right)^\alpha \right]^2 \right]^\theta \right\}}$$

$$Q(u) = \lambda \left\{ -\log \left[1 - \left[1 - \left[1 - \left(\frac{\sin^{-1}(u)}{\pi/2} \right)^{\frac{1}{\theta}} \right]^{\frac{1}{2}} \right]^{\frac{1}{\alpha}} \right]^{\frac{1}{\gamma}} \right\}$$

3. Results

3.1 Assessing the Consistency of the Parameter Estimates of the New Family

To evaluate the performance of the recently introduced Sine Tope Leone Exponentiated Lomax distribution, we conducted a Monte Carlo Simulation method. The aim of the simulation is to compute the mean, bias, and root mean square error of the estimated parameters obtained through maximum likelihood estimation. The simulated data was generated using the quantile function of STLE-Lomax distribution for various sample sizes, specifically: n=20, 50, 100, 150, 200, and 250, with 1000 replications for each sample size. Throughout these simulations, we set the parameters to $\alpha = 1.72$, $\theta=1.2$, $\theta = 0.05$ and $\gamma = 0.99$. The results of the parameter estimates, bias, and root mean square error from the new distribution are summarized in Table 1

Table 1: Estimate, Bias and RMSE of the new STLE-Lomax Distribution

N	Properties	$\alpha = 1.72$	$\lambda = 1.2$	$\theta = 0.05$	$\gamma = 0.99$
20	Est.	1.8431	1.3277	0.0562	1.1256
	Bias	0.1231	0.1277	0.0062	0.1356
	RMSE	0.3805	0.4730	0.0164	0.5804
50	Est.	1.7846	1.2416	0.0525	1.0996
	Bias	0.0646	0.0416	0.0025	0.1096
	RMSE	0.2661	0.2967	0.0093	0.5034
100	Est.	1.7541	1.2241	0.0512	1.0692
	Bias	0.0341	0.0241	0.0012	0.0792
	RMSE	0.2148	0.2022	0.0063	0.4446
150	Est.	1.7488	1.2086	0.0507	1.0646
	Bias	0.0288	0.0086	0.0007	0.0746
	RMSE	0.1836	0.1693	0.0051	0.4603
200	Est.	1.7480	1.2024	0.0506	1.0535
	Bias	0.0280	0.0024	0.0006	0.0635
	RMSE	0.1561	0.1397	0.0044	0.4229
250	Est.	1.7488	1.1957	0.0504	1.0575
	Bias	0.0288	-0.0043	0.0004	0.0675
	RMSE	0.1521	0.1250	0.0040	0.4147

The results of the Monte Carlo Simulations are shown in table 1 above. The values of biases and RMSEs tend to zero as shown in Table 1 and the estimates tend to the true parameter values as the sample size increases, indicating that the estimates are efficient and consistent.

3.2 Application

Two datasets were considered for illustrative purposes and comparison with the baseline distribution [30] and other extensions of the Lomax distribution such as: Tope-Leone Exponentiated Lomax distribution by [31], Type II Topp-Leone Lomax by [32], Half Logistic-Lomax distribution developed by [33], Weibull Lomax distribution by [34] and Gompertz Lomax distribution by [35]. For each data set, we estimated the unknown parameters of each distribution by the maximum-likelihood method and also obtained the values of the Akaike information criterion (AIC) for the proposed model and the competitors.

The pdf of Lomax developed by [30] is given by:

$$f(x) = \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} \quad (34)$$

The pdf of TLE-Lm as developed by [31] has pdf defined by:

$$f(x) = 2\alpha\beta\gamma\lambda(1 + \lambda x)^{-(\alpha+1)}(1 - (1 + \lambda x)^{-\beta})^{\gamma-1} [1 - (1 - (1 + \lambda x)^{-\alpha})^\beta] \{1 - [1 - (1 - (1 + \lambda x)^{-\alpha})^\beta]\}^{\gamma-1} \quad (33)$$

The pdf of TIITL-Lm developed by [32] is defined by:

$$f(x) = \frac{2\alpha\gamma}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} \left[1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}\right] \left\{1 - \left[1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}\right]^2\right\}^{\gamma-1} \quad (35)$$

The pdf of HL-Lm developed by [33] is defined by:

$$f(x) = \frac{2\alpha\lambda(1 + \lambda x)^{-(\alpha+1)}}{[1 + (1 + \lambda x)^{-\alpha}]^2} \quad (36)$$

The pdf of W-Lm distribution by [34] is defined by:

$$f(x) = \frac{2\gamma\beta\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{\beta\alpha-1} \left[1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}\right]^{\beta-1} \exp\left\{-\gamma \left\{\left(1 + \frac{x}{\lambda}\right)^{-\alpha} - 1\right\}^\beta\right\} \quad (37)$$

The pdf of GoLm distribution by [35] is defined by:

$$f(x) = \theta\gamma\lambda(1 + \lambda x)^{\gamma\alpha-1} \exp\left\{\frac{\theta}{\alpha} [1 - (1 + \lambda x)^{-\gamma\alpha}]\right\} \quad (38)$$

3.2.1 First Data Set

The first data set as listed below represents the COVID-19 positive cases record in Pakistan from March 24 to April 28, 2020, previously used by [36] and [37]: 2, 2, 3, 4, 26, 24, 25, 19, 4, 40, 87, 172, 38, 105, 155, 35, 264, 69, 283, 68, 199, 120, 67, 36, 102, 96, 90, 181, 190, 228, 111, 163, 204, 192, 627, 263.

3.2.2 Second Dataset

The second data set represents the failure times of the air conditioning system of an airplane. The data set was given by [38], it has been used by [39], and also by [4]. The data set is presented below:

23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95.

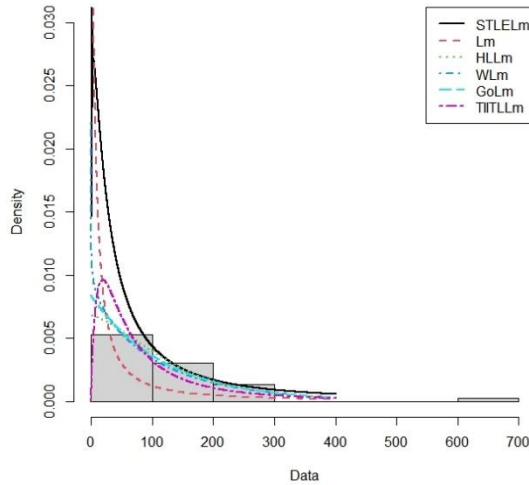


Figure 5: Density plot of the STLE-Lomax distribution for the first data sets

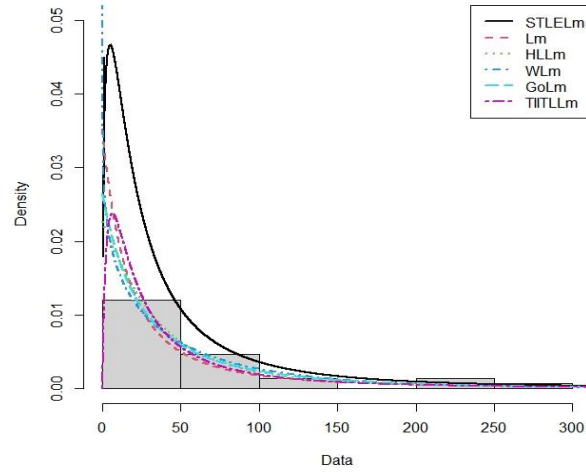


Figure 6: Density plot for the STLE-Lomax distribution for the second data set

Table 2 and table 3 gives the summary statistics of the two data sets such as the mean, the median, the first and third quartile, the minimum and the maximum values.

Table 2: Summary Statistics of the two data sets

Data	Minimum	Q_1	Median	Mean	Q_3	Maximum
Dataset I	2.00	32.75	93.00	119.28	183.25	627.0
Dataset II	1.00	12.50	22.00	59.60	83.00	261.0

Table 3: Estimate, Bias and RMSE of the new STLE-Lomax Distribution

Data set I	α	λ	θ	γ	LL	AIC
STLE-Lm	1.03323	0.11602	0.40035	2.56160	-189.4279	386.8558
TIITL-Lm	0.18430	24.5957	14.4732	-	-213.8879	433.7758
Lm	0.41228	6.73500	-	-	-224.5627	453.1253
HL-Lm	7.14437	0.00191	-	-	-208.3690	420.7379
W-Lm	0.01813	0.84166	1.12087	1.67046	-207.9697	423.9394
Go-Lm	7.96169	2.14265	47.014355	-	-209.7675	425.5349
Dataset II	α	λ	θ	γ	LL	AIC
STLE-Lm	1.77323	0.13691	0.64458	1.70984	-131.7712	271.5425
TIITL-Lm	34.2579	2.67719	0.06294	-	-167.6184	341.2368
Lm	17.1694	27.7386	-	-	-153.0699	310.1398
HL-Lm	0.05118	0.00178	-	-	-134.2354	272.4707
W-Lm	5.54118	5.62194	0.23070	0.18368	-183.2427	274.4856
Go-Lm	0.03744	0.01936	0.00819	-	-186.4186	378.8372

Table 3 presents the results of the two datasets. The analysis compared the performance of the Sine-Topp-Leone Exponentiated Lomax distribution against several other distributions, namely the Type II Topp Leone Lomax distribution, Lomax distribution, Half-Logistic-Lomax distribution, Weibull-Lomax distribution, and Gompertz-Lomax distribution. The results indicated that the proposed Sine Topp-Leone Exponentiated Lomax distribution outperformed some competing distributions, as it exhibits the lowest AIC value.

The visual assessment of the goodness of fit, as depicted in Figures 5 and 6, further validates the superiority of the proposed distribution when compared to other competing distributions. Therefore, it can be concluded that the proposed family of distributions is the most suitable choice for modeling both the COVID-19 and failure times of the air conditioning system of an airplane datasets.

4. Discussion

In this research, we introduced a new family of lifetime distributions by applying a Sine Transformation. Two special distributions were derived from the family by considering Lomax and Weibull distributions. Numerical analysis of fitting two real live data sets was presented using a maximum likelihood technique and density plots were provided to visually assess the outcomes.

References

- [1] Topp, C. W., & Leone, F. C. (1955). A family of J-shaped frequency functions. *Journal of the American Statistical Association*, 50(269), 209-219.
- [2] Sangsanit, Y., & Bodhisuwan, W. (2016). The Topp-Leone generator of distributions: properties and inferences. *Songklanakarinn Journal of Science & Technology*, 38(5), 537-548.
- [3] Chipepa, F., Oluyede, B., and Makubate, B. (2020). The Topp-Leone-Marshall-Olkin-G family of distributions with applications. *International Journal of Statistics and Probability*, 9(4), 15-32.
- [4] Yousof, H. M., Alizadeh, M., Jahanshahi, S. M. A., Ghosh, T. G. R. I., & Hamedani, G. G. (2017). The transmuted Topp-Leone G family of distributions: theory, characterizations and applications. *Journal of Data Science*, 15(4), 723-740.
- [5] Sule, I., Doguwa, S. I., Audu, I., & Jibril, H. M. (2020). On the Topp Leone exponentiated-G family of distributions: Properties and applications. *Asian Journal of Probability and Statistics* 7(1), 1-15.
- [6] Reyad, H., Alizadeh, M., Jamal, F., & Othman, S. (2018). The Topp Leone odd Lindley-G family of distributions: Properties and applications. *Journal of Statistics and Management Systems*, 21(7), 1273-1297.
- [7] Alizadeh, M., Lak, F., Rasekhi, M., Ramires, T. G., Yousof, H. M., and Altun, E. (2018). The odd log-logistic Topp-Leone G family of distributions: heteroscedastic regression models and applications. *Computational Statistics*, 33, 1217-1244.
- [8] Reyad, H., Korkmaz, M. Ç., Afify, A. Z., Hamedani, G. G., & Othman, S. (2021). The Fréchet Topp Leone-G family of distributions: Properties, characterizations and applications. *Annals of Data Science*, 8, 345-366.
- [9] Brito, E., Cordeiro, G. M., Yousof, H. M., Alizadeh, M., and Silva, A. O. (2017). The Topp-Leone odd log-logistic family of distributions. *Journal of Statistical Computation and Simulation*, 87(15), 3040-3058.
- [10] Hassan, A. S., Elgarhy, M., & Ahmad, Z. (2019). Type II Generalized Topp-Leone Family of Distributions: Properties and Applications. *Journal of data science*, 17(4), 638-659.
- [11] Moakofi, T., Oluyede, B., & Chipepa, F. (2021). Type II exponentiated half-logistic Topp-Leone Marshall-Olkin-G family of distributions with applications. *Heliyon*, 7(12), e08590.

- [12] Chipepa, F., Oluyede, B., & Makubate, B. (2020). The Topp-Leone-Marshall-Olkin-G family of distributions with applications. *International Journal of Statistics and Probability*, 9(4), 15-32.
- [13] Elgarhy, M., Arslan Nasir, M., Jamal, F., & Ozel, G. (2018). The type II Topp-Leone generated family of distributions: Properties and applications. *Journal of Statistics and Management Systems*, 21(8), 1529-1551.
- [14] Karamikabir, H., Afshari, M., Yousof, H. M., Alizadeh, M., & Hamedani, G. (2020). The Weibull Topp-Leone generated family of distributions: statistical properties and applications. *Journal of the Iranian Statistical Society*, 19(1), 121-161.
- [15] Vasileva, M., Rahneva, O., Malinova, A., & Arnaudova, V. (2021). The odd Weibull-Topp-Leone-G power series family of distributions. *International Journal of Differential Equations and Applications*, 20, 43-58.
- [16] Moakofi, T., Oluyede, B., & Gabanakgosi, M. (2022). The Topp-Leone Odd Burr III-G Family of distributions: Model, properties and applications. *Statistics, Optimization & Information Computing*, 10(1), 236-262.
- [17] Chipepa, F., Oluyede, B., & Peter, P. O. (2021). The Burr III-Topp-Leone-G family of distributions with applications. *Heliyon*, 7(4).
- [18] Oluyede, B., Chamunorwa, S., Chipepa, F., & Alizadeh, M. (2022). The Topp-Leone Gompertz-G family of distributions with applications. *Journal of Statistics and Management Systems*, 25(6), 1399-1423.
- [19] Sanusi, A. A., Doguwa, S. I. S., Audu, I., & Baraya, Y. M. (2020). Topp Leone Exponential-G Family of Distributions: Properties and Application. *Journal of Science and Technology Research*, 2(4).
- [20] Korkmaz, M. Ç., Yousof, H. M., Alizadeh, M., & Hamedani, G. G. (2019). The Topp-Leone generalized odd log-logistic family of distributions: properties, characterizations and applications. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 68(2), 1506-1527.
- [21] Al-Babtain, A. A., Elbatal, I., Chesneau, C., and Elgarhy, M. (2020). Sine Topp-Leone-G family of distributions: Theory and applications. *Open Physics*, 18(1), 574-593.
- [22] Mahmood, Z., Chesneau, C., & Tahir, M. H. (2019). A new sine-G family of distributions: properties and applications. *Bull. Comput. Appl. Math.*, 7(1), 53-81.
- [23] Chesneau, C., and Jamal, F. (2020). The sine Kumaraswamy-G family of distributions. *Journal of Mathematical Extension*, 15.
- [24] Muhammad, M., Alshanbari, H. M., Alanzi, A. R., Liu, L., Sami, W., Chesneau, C., & Jamal, F. (2021). A new generator of probability models: the exponentiated sine-G family for lifetime studies. *Entropy*, 23(11), 1394.
- [25] Sakthivel, K. M., & Rajkumar, J. (2021). Transmuted sine-G family of distributions: theory and applications. *Statistics and Applications*, (Accepted: 10 August 2021).
- [26] Rajkumar, J., & Sakthivel, K. M. (2022). A New Method of Generating Marshall-Olkin Sine-G Family and Its Applications in Survival Analysis. *Lobachevskii Journal of Mathematics*, 43(2), 463-472.
- [27] Fayomi, A., Algarni, A., & Almarashi, A. M. (2021). Sine Inverse Lomax Generated Family of Distributions with Applications. *Mathematical Problems in Engineering*, 2021, 1-11.
- [28] Kumar, D., Singh, U., & Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients' data. *Journal of Statistics Applications & Probability*, 4(3), 417-427.
- [29] David, H. A. (1970). Order statistics, Second edition. Wiley, New York.
- [30] Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American statistical association*, 49(268), 847-852.
- [31] Sule, I., Doguwa, S. I., Audu, I., & Jibril, H. M. (2020). On the Topp Leone exponentiated-G family of distributions: Properties and applications. *Asian Journal of Probability and Statistics* 7(1), 1-

15.

[32] Elgarhy, M., Arslan Nasir, M., Jamal, F., & Ozel, G. (2018). The type II Topp-Leone generated family of distributions: Properties and applications. *Journal of Statistics and Management Systems*, 21(8), 1529-1551.

[33] Anwar, M., and Zahoor, J. (2018). The half-logistic Lomax distribution for lifetime modeling. *Journal of probability and Statistics*, 1-12.

[34] Tahir, M. H., Cordeiro, G. M., Mansoor, M., & Zubair, M. (2015). The Weibull-Lomax distribution: properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 44(2), 455-474.

[35] Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., Adejumo, A. O., & Odetunmbi, O. A. (2017). A new generalization of the Lomax distribution with increasing, decreasing, and constant failure rate. *Modelling and Simulation in Engineering*,

[36] Al-Marzouki, S., Jamal, F., Chesneau, C., and Elgarhy, M. (2020). Topp-Leone odd Fréchet generated family of distributions with applications to COVID-19 data sets. *Computer Modeling in Engineering & Sciences*, 125(1), 437-458.

[37] Bello, O. A., Doguwa, S. I., Yahaya, A., and Jibril, H. M. (2021). A Type I Half Logistic Exponentiated-G Family of Distributions: Properties and Application. *Communication in Physical Sciences*, 7(3), 147-163.

[38] Linhart H, Zucchini W. (1986). *Model selection*. John Wiley, New York, USA; 1986.

[39] Shanker, R., Hagos, F., & Sujatha, S. (2015). On modeling of Lifetimes data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*, 2(5), 1-9.

ANALYSIS OF RELIABILITY OF TYPICAL POWER SUPPLY CIRCUITS

S.V. Rzayeva¹, N.M. Piriyeva², I.A. Guseynova³

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan

¹sona.rzayeva@asoiu.edu.az ; ²naciba.piriyeva@asoiu.edu.az ; ³huseynova.ilduze@asoiu.edu.az

Abstract

This article is devoted to the analysis of the reliability of typical power supply circuits. The question of what factors have the greatest impact on the reliability of power supply circuits, as well as what methods and tools are used to analyze and improve their reliability is considered. Particular attention is paid to the comparative analysis of various types of power supply schemes and the determination of their advantages and disadvantages in terms of reliability. The article also discusses current trends and developments in the field of increasing the reliability of power supply and possible ways to optimize existing circuits. The results obtained can be useful for specialists in the field of power engineering and electrical engineering in the design, maintenance and modernization of power supply systems.

Keywords: electrical power systems, reliability indicators, cross-section classes, typical circuits.

I. Introduction

In the modern world, electricity supply is a key aspect of ensuring the life of society and the functioning of its economy. The performance of industrial enterprises, the safety of residential complexes, the efficiency of vehicles and much more depend on the reliability of power supply systems. In this regard, analysis and improvement of the reliability of standard power supply circuits is a relevant and important task for specialists in the field of electrical engineering and power engineering.

The purpose of this article is to study the factors affecting the reliability of power supply circuits, as well as to develop methods and tools for their analysis and improvement. It is proposed to consider various types of power supply schemes, analyze their advantages and disadvantages from a reliability point of view, and also consider current trends and developments in the field of increasing the reliability of power supply.

It is important to note that the efficient operation of modern power supply systems requires not only technical competence, but also consideration of various factors such as climatic conditions, technological changes and energy efficiency requirements. Therefore, analysis of the reliability of typical power supply schemes has many practical applications and can become the basis for optimizing existing power supply systems and developing new, more reliable solutions.

II. Formulation of the problem

Determining reliability indicators for modern electric power systems is impossible without the use of appropriate software systems. In our country and abroad today, the following software

systems are most widely used, allowing one to model and calculate probabilistic reliability indicators of electric power systems:

1. Software systems "RISK SPECTRUM" (Sweden); "SAPHIRE" (USA), using "fault trees" and "event trees" as initial data.
2. Software package "WINDCHILL RBD" (USA), using a special block diagram of system performance.
3. Domestic software complex "ARBITR" ("ASM SZMA"), using the logical-probabilistic method.
4. Software systems for modeling energy systems: "MATLAB", a software environment that allows you to simulate energy facilities and develop control systems; "ETAR SYSTEMS" (USA), software for electrical power systems, allowing for the design, analysis, and maintenance of electrical power systems; "PSCAD" (Canada), a software package that allows you to simulate the operation of power systems.

The main disadvantage of foreign-made software systems is the high cost and complexity of training personnel to study specialized software systems.

The use of complex and expensive specialized software systems is justified only in those industries where equipment failure can cause catastrophic consequences, for example, in nuclear energy. For projects in which equipment failure does not entail such serious consequences, it is possible to use proprietary software products, for example [1, 6, 8], the cost of which is not comparable with specialized ones, and the limited set of functionality is compensated by the ease of development.

III. Problem solution

Most of the main step-down substations currently being built have a simplified circuit with separators and short circuiters on the high voltage side. Refusal to install an oil or air circuit breaker saves capital and operating costs and reduces the construction time of a substation (Figure 1). Jumpers on the high voltage side increase the maneuverability of switching of dead-end substations, especially if they are equipped with a separator with a double-acting drive. At dead-end substations made in the form of a "radial line – transformer" block, separators need not be installed.

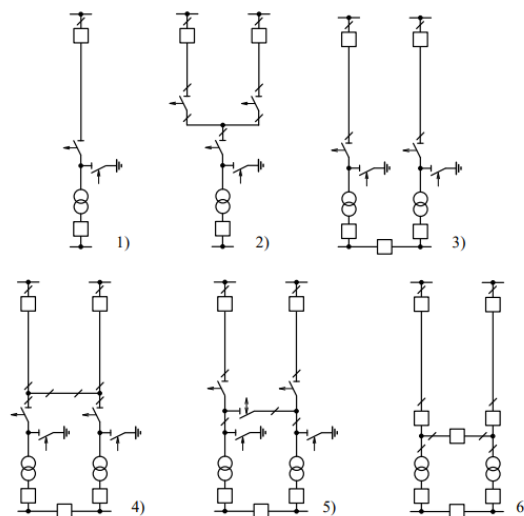


Figure 1: Comparison of typical external power supply schemes

During the operation of simplified substations, significant shortcomings were identified in the operation of open-type separators and short circuiters. The high response time of these devices

makes it difficult to automatically reclose the main switch and contributes to the development of damage that occurs in the transformer. In addition, turning on the short circuit causes a sharp decrease in voltage at the supply substation. If an air circuit breaker with a voltage of 110–220 kV is used as a main switch, installing a short circuit in a zone of 0.5–6 km is unacceptable due to the kilometer effect. In this zone, the short-circuiter is replaced by various teletrip pulse systems while maintaining the backup function of the short-circuiter. The use of a tele-breaking pulse also avoids the reduction in voltage caused by the activation of the short circuit.

Table 1 shows reliability indicators for the circuits presented in Figure 1.

Table 1: reliability indicators for the circuits presented in Figure 1.

Scheme	ω_{Σ} , year ⁻¹	n, year ⁻¹	T _{rec.} , hour	K _{rec.} , p.u.	Conclusion
1	0.401	0.3048	12.5	5.7·10 ⁻⁴	Best by ω_{Σ}
2	0.793	0.1160	14.4	1.8·10 ⁻⁴	-
3	0.802	0.0154	12.0	0.21·10 ⁻⁴	Best by n
4	0.804	0.0164	10.3	0.19·10 ⁻⁴	-
5	0.823	0.1660	11.6	0.22·10 ⁻⁴	-
6	0.758	0.0250	5.6	0.16·10 ⁻⁴	Best by T _{rec.} and K _{rec.}

Analysis of Table 1 allows us to draw the following conclusions. As the number of equipment in the circuit increases, the total failure flow parameter ω_{Σ} increases. From the point of view of uninterrupted power supply (number of outages n), the most effective is scheme 3. Scheme 6 turns out to be the best in terms of recovery time and emergency downtime coefficient, however, it is approximately 10 times more expensive than the most complex of the simplified schemes - scheme 5. Dead-end substations made according to schemes 3, 4 and 5, are close to the optimal solution both from the point of view of uninterrupted and efficiency.

Taking reliability into account when planning the development of power supply systems (PSS) of industrial enterprises (IEs) and when designing its individual links, as well as in operating conditions, is to ensure optimal reliability of power supply to consumers, taking into account the known reliability of system elements.

The main directions for increasing the reliability of power supply to consumers include:

- Creation of rational power supply schemes with an increased degree of reliability.
- Increasing the reliability of PSS elements.
- Improvement or implementation of electrical automation devices and the use of telemechanics.
- Improving the operational maintenance of solar power plants.
- Application of monitoring the technical condition of solar power plant equipment.

In addition to the listed areas, the reliability of PSS is influenced by the following factors:

- Quality of electrical energy.
- Correct choice of neutral mode of electrical networks.
- Optimization of short circuit currents.

Development of rational power supply schemes

The development of efficient power supply schemes is a critical aspect in the planning and operation of energy systems. Optimization of these schemes is aimed at ensuring an uninterrupted and reliable supply of electrical energy to consumers while minimizing economic costs and maximizing operational efficiency. The main parameters that are taken into account when developing circuits are the reliability of system elements, their interaction, as well as possible operating modes in emergency situations.

A critical step in developing sustainable power supply designs is risk assessment and failure analysis of system components. The use of modern modeling and simulation methods makes it

possible to predict the behavior of the system in various operational scenarios and minimize the likelihood of failures. This takes into account statistical data on the reliability of components, test and operation results of similar systems, as well as requirements for the quality of power supply.

One of the key areas for optimizing power supply schemes is increasing the degree of redundancy and introducing fault-tolerant technologies. This includes the use of backup lines, automatic transfer switches and intelligent control systems that ensure rapid restoration of power in the event of a major equipment failure. In addition, the use of telemechanical systems and electrical automation devices makes it possible to quickly identify and eliminate faults, which significantly reduces downtime and increases the overall reliability of the system.

Sustainable electricity supply schemes must also take into account economic aspects. This involves analyzing the costs of installing and maintaining equipment, as well as assessing the cost-effectiveness of various circuit options. An important factor is the selection of equipment with an optimal balance of cost and reliability, which allows for a high level of power supply at an acceptable cost.

Increasing the reliability of PSS elements

Increasing the reliability of elements of power supply systems (PSS) is one of the key tasks in the planning and operation of energy systems. The high reliability of individual components can significantly reduce the frequency of emergency situations and ensure a stable power supply to consumers.

One of the main approaches to increasing the reliability of SES elements is the use of high-quality equipment with high performance characteristics. This includes selecting equipment from trusted manufacturers, undergoing rigorous testing and certification, and regularly updating and upgrading obsolete components. The use of innovative materials and technologies, such as intelligent control and monitoring systems, also helps improve reliability indicators.

Regular maintenance and preventive measures play an important role in maintaining the reliability of PSS elements. Routine inspections, diagnostics and replacement of worn parts allow potential problems to be identified and corrected before they develop into serious accidents. The introduction of systems for monitoring the technical condition of equipment, such as vibration sensors, temperature sensors and other parameters, allows you to quickly obtain data on the condition of elements and take measures to maintain them.

Another important aspect is the advanced training and training of personnel responsible for the operation and maintenance of solar power plants. Qualified specialists are able to quickly respond to emerging problems, carry out necessary repairs and ensure correct operation of the equipment. Organizing regular trainings, seminars and advanced training courses helps improve the professional skills and knowledge of employees.

The development of normative and methodological documents regulating the operation and maintenance of solar power plant elements is also important. Standardization of procedures, compliance with technical regulations and recommendations for the operation of equipment make it possible to ensure a uniform level of quality and reliability in all parts of the power supply system.

Application of relay protection and automation devices

The use of relay protection and automation devices plays a key role in modern power supply systems, ensuring the reliability, safety and efficiency of their operation. These technologies are designed to quickly identify and isolate faults in electrical networks, minimize downtime and prevent possible equipment damage. Relay protection provides protection against overloads, short circuits and other anomalies, automatically disconnecting damaged sections of the network to prevent the spread of emergency situations.

Automation devices, in turn, carry out automatic control of power supply, including automatic restart after short-term outages (AR) and automatic switching to backup power supplies

(APS) in the event of failure of the main equipment. This significantly reduces downtime and ensures continuity of power supply for consumers.

Integration of relay protection and automation into power supply control systems allows not only to increase its reliability, but also to significantly improve controllability and operational safety. The use of modern technologies in this area is a prerequisite for the efficient functioning of modern energy systems, where every minute of downtime can have significant economic and social consequences.

Factors influencing the reliability of power supply systems (PSS)

The reliability of power supply systems (PSS) is determined by many factors that affect their ability to provide uninterrupted power supply to consumers. The main aspects affecting the reliability of solar power systems include technical, operational, organizational and economic factors.

Among the technical aspects, the quality and condition of the equipment plays a key role. The choice of modern and reliable technologies, the correct power supply scheme and the use of modern protective devices and automation significantly influence the degree of protection of the system from emergency situations. Regular maintenance and condition checks of equipment also play an important role in preventing possible failures.

Operational aspects include operating mode management, technical condition monitoring and personnel qualifications. Trained and qualified personnel are able to effectively manage the system and quickly respond to possible problems.

Organizational aspects include compliance with standards, risk management and optimization of management and operational processes. Economic factors also play a role: investments in modernization and renewal of equipment help to increase the reliability of solar power plants while optimizing operating costs.

All these aspects are interconnected and require an integrated approach to achieve high reliability indicators. Only systematic management of all factors can ensure stable and efficient operation of power supply systems in the context of modern technological and economic challenges.

IV. Conclusions

Electrical power systems are vital infrastructure that ensures the continued functioning of industrial plants, commercial properties and residential areas. The reliability of these systems plays a critical role in ensuring the stability and security of energy supply for millions of consumers.

An optimal power supply scheme, the use of modern relay protection and automation technologies, as well as an integrated approach to the management and operation of equipment significantly affect the reliability of the system. Effective operating mode management, regular maintenance and qualified personnel are key to minimizing risks and preventing downtime.

Uninterrupted power supply is necessary not only for economic stability, but also to ensure the comfort and safety of people's lives. Investments in modern technologies and continuous improvement of electricity supply systems are strategic priorities that contribute to improving the quality of life and sustainable development of society.

Only an integrated approach to the management, maintenance and modernization of power supply systems allows us to achieve high standards of reliability and ensure sustainable operation in the conditions of modern dynamic economy and technological challenges..

References

- [1] E.Safiyev, N.Pirieveva, G.Bagirov //Analysis of the application of active lightning rods in lightning protection objects// Interscience: electron. scientific magazine 2023. No. 6(276). pp 14-17
- [2] N.M.Pirieveva, S.V.Rzaeva, S.N.Talibov Analysis of surge protection devices for electrical networks

- "Internauka": scientific journal – No. 43 (266). Part 3. Moscow, Publishing house. "Internauka", 2022. pp. 14-17
- [3] N.M.Piriyeva, S.V.Rzayeva, E.M.Mustafazadeh /Evaluation of the application of various methods and equipment for protection from emergency voltage in 6-10 kV electric networks of oil production facilities/. Interscience: electron. Scientific magazine 2022. No. 39(262). p.40-44
- [4] GV Mamedova, GS Kerimzade, NM Piriyeva /Electromagnetic calculation of tension devices for winding wires of small cross sections/ IJ TPE Journal, ISSUE 53. Volume 14. Number 4. Decembe, 2022,
- [5] Piriyeva N.M., Tagizade L.N. "Surge suppressors and transformer surge protection" International scientific journal "BULLETIN OF SCIENCE. No. 1 (70) Volume 3. 2024. pp. 772-778.
- [6] Piriyeva N.M., Veliev Q.A., Abbasov A.I., Suleymanov E.E. "Switching processes in electric networks 10-35 kV." Energy problem no. 2, Baku, 2021 pp. 100-106.12.
- [7] Piriyeva N.M., Makhmudov U.I.//Analysis of substance grounding and surge protection system// Flaqman nauki: naucniy jurnal. Yanvar 2024. - St. Petersburg, Izd. State Research Institute "Naprazvitie" - 2024. No. 1(12).
- [8] Rahimli I.N., Rzayeva S.V., Choluev M.E. Technology for designing alternative energy supply systems. interscience, Volume 5, 2023, pp.23-25
- [9] Rahimli I. N., Rzayeva S. V., Zairov T. N. Applicability of wind-solar energy as an alternative source of power supply to oil facilities // Bulletin of Science. – 2023. – T. 4. – No. 4 (61). – pp. 289-293.

FUZZY LOGIC RELIABILITY BLOCK DIAGRAM APPROACH FOR PATIENT HEALTH MONITORING USING R PROGRAMMING

Liji Sebastian¹, Rita S², Vennila J³

¹Department of Statistics, Periyar University, Salem-636011

lijigeorge2000@gmail.com

²Department of Statistics, Periyar University, Salem-636011

ritasmikannu@gmail.com

³Manipal College of Health Professions

dr.j.vennila@gmail.com

Abstract

In this research, a new approach using fuzzy logic and reliability block diagram (RBD) techniques is used to ensure the reliability of patient health monitoring systems. This technique handles uncertainties in health information, while RBD assesses system reliability by displaying factor relations. The RBD model construct for system components and measure reliability using probabilistic models. Fuzzy logic identifies the effect of uncertainties on overall reliability. Using this approach in a simulated health monitoring scenario, using R, we demonstrate its effectiveness and potential to increase reliable health monitoring for improved patient outcomes and healthcare efficiency. Furthermore, the awareness gained from this study can be directed beyond healthcare such as modern process control and environmental sensing.

Keywords: Fuzzy logic, reliability block diagram, health care, efficiency, R programming, membership function.

I. Introduction

The engineering of reliability is essential to the effectiveness, safety, and economic viability of technologies across various sectors. However, reliability models tend to minimize the complexity and unpredictability that characterize modern technology. Fuzzy logic highlights the potential for this new model with a framework to deal with the fundamental inaccuracy and uncertainty observed in the real-world process of decision-making. By providing the interpretation of qualitative and fuzzy models, fuzzy logic simulates human logic and allows challenging decision-making in uncertain situations. This adaptability and quickness provide a perfect model for improving conventional reliability engineering techniques.

Fuzzy logic and artificial intelligence together indicate the most significant development in analysis of reliability. The application of AI methods like artificial neural networks and algorithm development permits applications using fuzzy logic to keep evolving. Since the outcome, by gradually understanding using data, clients will improve their methods for making decisions. The ability of fuzzy inference systems to evolve and prosper in difficult and unexpectedly shifting

environments could be increased through this combination of factors. Reliability block diagrams (RBDs) are needed for evaluating the system's reliability to highlight the relationships between components and their influence on the general effectiveness of the entire system. Even though systems sometimes fail to prepare for the inherent uncertainty of real-world systems, conventional RBDs are effective at predicting dependent scenarios.

This paper provides an approach for integrating fuzzy logic with health information from patients. To provide an understanding of a person's medical scenario, the approach evaluates the possibility related to each risk factor and takes into factor the person's family history, lifestyles, prior medical history, and factors in the environment. Fuzzy logic, FALCON, and BP in combination might improve the diagnostic rate of MSSA patients, according to study findings of Modai, I et al. [1]. The ecological relevance of Asian tiger mosquitoes for infectious diseases has been evaluated by Proestos, Y et al. [2] using fuzzy logic. Fuzzy logic plays an essential function in evaluating human resources performance, as Sadegh Amalnick, M et al. [3] indicate. In their idea, Davoodi, R et al. [4] had higher accuracy than conventional algorithms for forecasting ICU patient mortality using the Deep Rule-Based Fuzzy System (DRBFS). A fuzzy-based Bayesian model has been developed by Rallapalli, S. et al. [5] in the campaign over COVID-19 to help quickly isolate SARS-CoV-2 RNA by determining the most effective sets for wastewater sampling.

Li, S et al. [6] improved performance in the context of interference challenges by providing a network of sensors health forecasting framework using a system base model with attribute reliability (BRB-r). Reliability is essential for the implementation of blockchain technology in medical care, based on Du, X et al. [7]. Dovic, K et al. [8] studied the performance of AHCL and HCL methods for controlling glucose levels to an intermittent baseline of 150 mg/dL as implementing fuzzy logic in the FLAIR learning. Abd Rahman, N H et al. [9] addressed an observation on the centrality of the healthcare system's accuracy. The medical information research studies done by Sebastian, L et al. [10], Vijayan, K et al. [11], and Vennila, J et al. [12], [13] mainly using R programming.

II. Methods

This study's primary objective is to show how the fuzzy logic idea and reliability block design can be used effectively in a simulated health observation situation to increase reliability, which in turn can lead to better patient outcomes and more efficient healthcare delivery. The following steps were used in methodological parts.

- Step 1: Construction of RBD model
- Step 2: Integration with fuzzy logic concept
- Step 3: Simulation in R
- Step 4: Evaluating the results.

III. Case Scenarios

Based on the above methodological concept, we consider few examples and demonstrate the effectiveness of using fuzzy logic concept and the results obtained are discussed below:

I. Case Scenario One

In this scenario, fuzzy logic and reliability are incorporated to evaluate the overall reliability level of a system composed of two components based on their individual reliabilities and predefined thresholds. Let the predefined thresholds be as follows: high is 0.65, average is 0.39 and low is 0.21; and let the individual reliability for two components follow U (0,1). The system accuracy can be classified to be high, average, or low according to this inference outcome.

II. Case Scenario Two

This scenario includes a health assessment system that evaluates the health status of patients as per their blood pressure, cholesterol, and Body Mass Index (BMI). Applying predefined membership functions, the system classifies “healthy” or “unhealthy” based on the vital statistics of health status of patients. The case studies reveal that, how the system assesses each patient data to measure their overall health status, providing valuable awareness for healthcare filed to constitute notified findings and interventions.

II. Case Scenario Three

In this scenario, a patient's health evaluation status, based on variables such as age, body mass index (BMI), heart rate, cholesterol and blood pressure, can be obtained using the code in concern. The information has been organized into various health status groups for each evidence.

IV. Results and Discussion

I. Output for Case Scenario One

For case scenario one, on executing the R coding a bar diagram is obtained, note that here the programme has been executed thrice to depictive 3 different types of output, also to ensure the randomness of the reliability for the components A and B, they are randomly assigned a value from Uniform (0,1) distribution. The outputs are as shown in figure 1, figure 2 and figure 3.

From figure1, observe that when the when the reliability of Component A= 0.0884 and of component B = 0.2487 the overall system reliability fuzzy logic is Low. It can be observed from figure 2 that when the reliability of Component A= 0.4799 and of component B= 0.3713 the overall system reliability based on fuzzy logic is Medium. From figure 3 we notice that when the reliability of Component A= 0.7317 and of component B= 0.5334, the overall system reliability based on fuzzy logic is High.

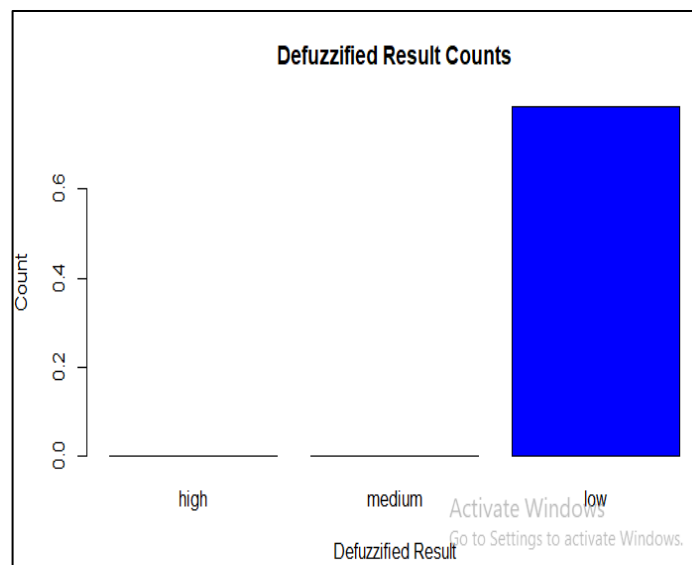


Figure 1: Low system reliability when A=0.0884; B=0.2487

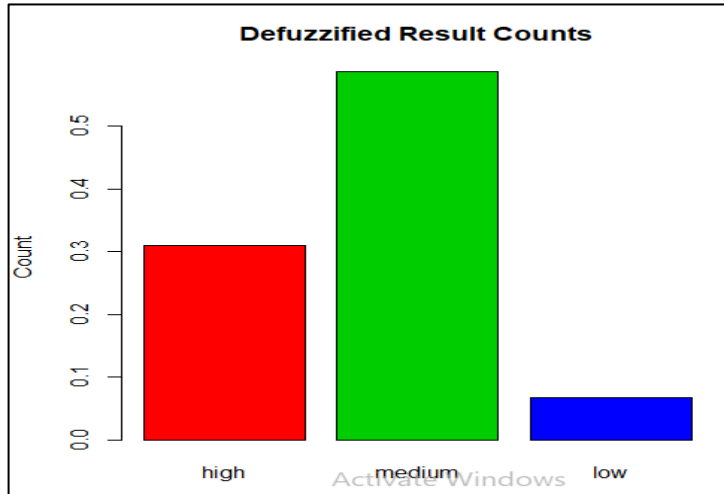


Figure 2: Medium system reliability when $A=0.4799$; $B=0.3713$

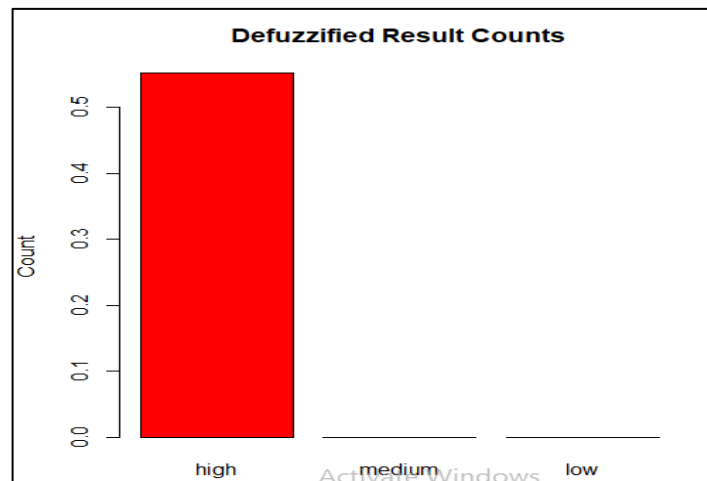


Figure 3: High system reliability when $A=0.7317$; $B=0.5334$

Thus, by considering the individual reliability of component, A and B, the code in R determines the overall reliability of the system. It identifies high, medium, and low reliability levels and gives results based on fuzzy. This approach permits screening variables based on changes in component reliability. The findings obtained are useful for decision-making on efficiency, baseline, and improvements to system designs. Based on predetermined requirements, the fuzzy inference evaluates components A and B and demonstrates its validity. Furthermore, in practical applications, reliability evaluation would need to consider additional variables such as factor mutuality and the active using environment.

II. Output for Case Scenario Two

For case scenario two, the R coding ensures that based on the varying random inputs regarding the vital statistics of blood pressure (BP), Cholesterol and BMI for the patient list provided, the output gives an accurate prediction on the health status of these individuals using fuzzy membership function. A patient is termed “healthy” if either one of the two conditions are met.

- Condition 1: BP is between 120 and 140, Cholesterol is between 100 and 240 and BMI is between 18.5 and 25.
- Condition 2: BP is less than or equal to 120, Cholesterol is less than or equal to 200

and BMI is between 18.5 and 25. The output for case scenario 2 is shown in Table 1.

Table 1: Output table for case scenario 2

Patient "ABC". with BP 92 & Chol 245 & BMI 23 Health Status: unhealthy
Patient "PQR" with BP 127 & Chol 185 & BMI 21 Health Status: healthy
Patient "XYZ" with BP 171 & Chol 300 & BMI 33 Health Status: unhealthy

From the above table we draw the following inference, for patient "ABC" although BMI of 23 is within the specified range (18.5, 25), BP of 92 is less than the specified range, and cholesterol level of 245 is above specified range with respect to condition 1 and hence the patient is labeled "unhealthy"; on verifying these vital statistics as per condition 2, we notice that although BMI of 23 is within the specified range (18.5, 25) and BP of 92 is less than 120 but the cholesterol level of 245 is above specified range and hence the patient is labeled "unhealthy". Since neither of the two conditions are met the overall health status is "unhealthy"

For patient "PQR", BP of 127 is within the specified range (120, 140), cholesterol level of 185 within the specified range (100, 240) and BMI of 21 is within the specified range (18.5, 25) with respect to condition 1 and hence the patient's overall health status is "healthy".

For patient "XYZ", BP is 171, Cholesterol is 300 and BMI is 33; all these vital statistics are above the permissible range with respect to condition 1 and hence the patient's overall health status is "unhealthy".

Since the vital statistics of patients are unique to every individual, the coding has been framed in a manner so as to incorporate the fuzzy nature of vital statistics namely Blood pressure, Cholesterol and BMI and gives an assured reliability of the health status of individuals. The fuzzy membership functions for Blood pressure, Cholesterol and BMI are depicted in figure 4.

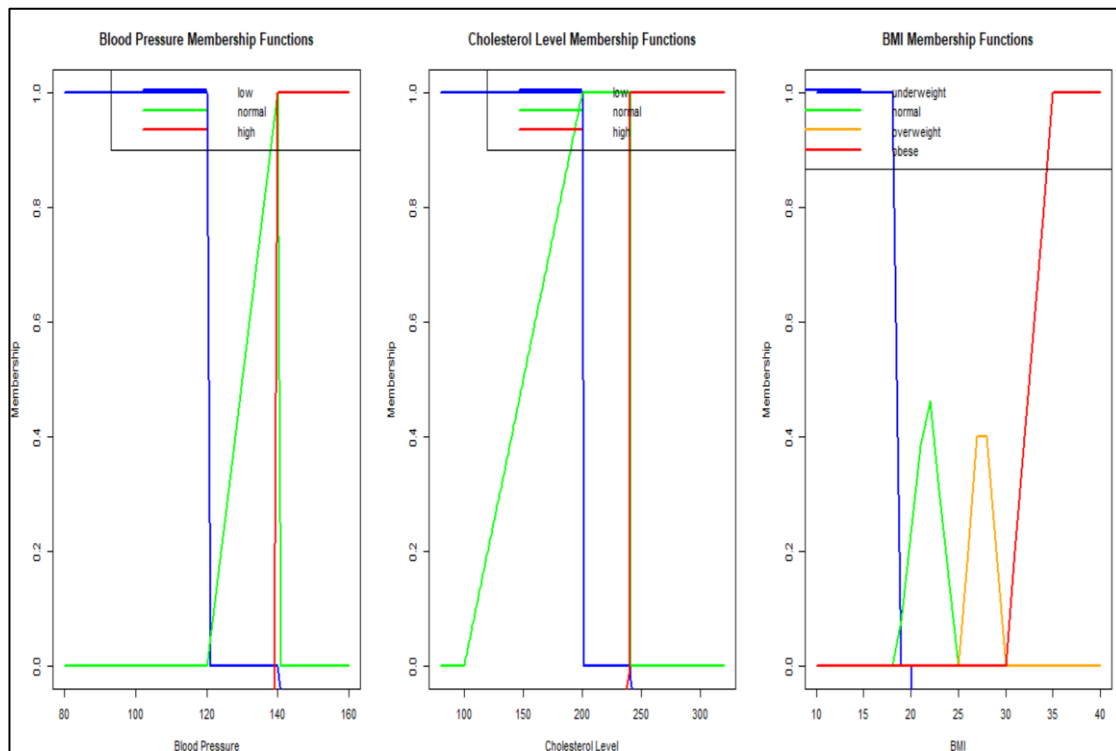


Figure 4: Fuzzy membership functions for Blood pressure, Cholesterol and BMI generate using R coding

From the above result for case scenario 2, observe that the health assessment system shown is a significant tool in healthcare practice, supporting clinicians in managing notified results and

upgrading patient outcomes through directed interventions and monitoring.

III. Output for Case Scenario Three

For case scenario three, the R coding is done in such a manner that each time we run the code the parameters of the vital statistics taken on random uniform nos. within a specific range as mentioned in the code and accurately determines the inference with respect to each of the varying vital statistics of a patient. Given below in table 2 are three randomly generated outputs along with their accurate inference

Table 2: Output table for case scenario 3

Patient 1:				
Age	BMI	Heart Rate	Cholesterol	BP
40	25	97	179	95
age_inference	bmi_inference	hr_inference	chol_inference	bp_inference
Medium	Medium	Medium	Low	Low
Patient 2:				
Age	BMI	Heart Rate	Cholesterol	BP
19	26	88	297	132
age_inference	bmi_inference	hr_inference	chol_inference	bp_inference
Low	Medium	Medium	High	Medium
Patient 3:				
Age	BMI	Heart Rate	Cholesterol	BP
60	27	70	222	67
age_inference	bmi_inference	hr_inference	chol_inference	bp_inference
High	Medium	Low	High	Low

Observe in Table 2; based on the predefined thresholds; for patient 1, Age (40), BMI (25), Heart Rate (97) are categorized as 'Medium'; Cholesterol (179), BP (95) are categorized as 'Low'. Similarly, for Patient 2, Age (19) is categorized as 'Low'; BMI (26), Heart Rate (88), BP (132) are categorized as 'Medium'; and Cholesterol (297) is categorized as 'High' and for Patient 3, Age (60), Cholesterol (179) are categorized as 'High'; BMI (27) is categorized as 'Medium'; Heart Rate (70), BP (67) are categorized as 'Low'.

The given R code generates random values for a patients Age, BMI, Heart Rate, Cholesterol and BP, and then classify each of these values into categories (High, Medium and Low) based on predefined thresholds. Thus, a thorough assessment of the health of an individual along a number of factors can be executed, permitting focus to be given to medication prescribed along with specific diseases.

V. Conclusion

This study gives the combination of fuzzy logic and machine learning with reliability block diagrams. It represents a capable boundary in reliability engineering. This method gives a complete procedure for measuring a patient's health status over observing many health parameters. It gives a meaningful valuation of the patient's complete health status. But it is worthwhile to notice that while fuzzy logic provides a useful framework for health evaluation, clinical decision and skill must also be measured when interpreting the findings and making healthcare assessments. Fuzzy logic with reliability block diagram recommends a great structure for evaluating the reliability of health

systems. This gives to improvements in diagnostic accuracy, patient care, and healthcare outcomes. The health evaluation gives actionable awareness for patients and their healthcare providers, highlighting the significance of active health controlling and aimed interventions to maintain ideal health and well-being.

References

- [1] Modai, I. Kurs, R. Ritsner, M. Oklander, S. Silver, H. Segal, A. Goldberg, I. and Mendel, S. (2002). Neural network identification of high-risk suicide patients. *Medical Informatics and the Internet in Medicine*, 27(1), 39-47.
- [2] Proestos, Y. Christophides, G. K. Ergüler, K. Tanarhte, M. Waldock, J. and Lelieveld, J. (2015). Present and future projections of habitat suitability of the Asian tiger mosquito, a vector of viral pathogens, from global climate simulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1665), 20130554
- [3] Sadegh Amalnick, M. and Zarrin, M. (2017). Performance assessment of human resource by integration of HSE and ergonomics and EFQM management system. *International Journal of Health Care Quality Assurance*, 30(2), 160-174
- [4] Davoodi, R. and Moradi, M. H. (2018). Mortality prediction in intensive care units (ICUs) using a deep rule-based fuzzy classifier. *Journal of Biomedical Informatics*, 79, 48-59.
- [5] Rallapalli, S. Aggarwal, S. and Singh, A. P. (2021). Detecting SARS-CoV-2 RNA prone clusters in a municipal wastewater network using fuzzy-Bayesian optimization model to facilitate wastewater-based epidemiology. *Science of The Total Environment*, 778, 146294.
- [6] Li, S. Feng, J. He, W. Qi, R. and Guo, H. (2021). A new health prediction model for a sensor network based on belief rule base with attribute reliability. *Scientific Reports*, 11(1), 2806.
- [7] Du, X. Chen, B. Ma, M. and Zhang, Y. (2021). Research on the Application of Blockchain in Smart Healthcare: Constructing a Hierarchical Framework. *Journal of Healthcare Engineering*, 2021, 6698122.
- [8] Dovic, K. Battelino, T. Beck, R. W. Sibayan, J. Bailey, R. J. Calhoun, P. Turcotte, C., Weinzimer, S. Smigoc Schweiger, D. Nimri, R. and Bergenstal, R. M. (2022). Impact of Temporary Glycemic Target Use in the Hybrid and Advanced Hybrid Closed-Loop Systems. *Diabetes Technology & Therapeutics*, 24(11), 848-852
- [9] Abd Rahman, N. H. Ibrahim, A. K. Hasikin, K. and Abd Razak, N. A. (2023). Critical Device Reliability Assessment in Healthcare Services. *Journal of Healthcare Engineering*, 2023, 3136511.
- [10] Sebastian, L. et al. (2022). Analyzing the Knowledge, Attitude and Practice about Obesity among College Students using Python Programming. *Indian Journal of Natural Sciences*, 13(71).
- [11] Vijayan, K. et al. (2022). Productive modeling for Coffee production using R programming. In *Proceedings of 2022 3rd International Conference on Communications, Computing, and Industry 4.0* (C214).
- [12] Vennila, J. et al. (2022). Analyzing the educational challenges during Covid-19 by using R-Programming. *International Journal of Ecological Economics and Statistics*, 43(3), 12-22.
- [13] Vennila, J. et al. (2022). Analyzing the Impact of Inflammatory Bowel Disease (IBD) by using R-Programming. *JP Journal of Biostatistics*, 19, 123-144.

RELIABILITY ANALYSIS OF AN ANTI-DRONE SYSTEM BY CONSIDERING RANDOM ENVIRONMENTAL FACTORS

DHARMARAJA SELVAMUTHU^a, HARSHITA BADIYASAR^a, SMRATI TRIPATHI^a,
PRIYANKA KALITA^b, RAINA RAJ^c

Department of Mathematics, Indian Institute of Technology Delhi, New Delhi, India^a

dharmar@maths.iitd.ac.in

mt1200807@maths.iitd.ac.in

mt1200855@maths.iitd.ac.in

Department of Statistics, Bhattadev University, Assam, India^b

priya.24723@gmail.com

Bharti School of Telecommunication Technology & Management,

Indian Institute of Technology Delhi, New Delhi, India^c

rainaraj.curaj@gmail.com

Abstract

In today's security landscape, the proliferation of unauthorized drones in restricted airspace has emerged as a significant threat. These drones pose various risks, from potential surveillance and espionage to more sinister possibilities such as physical attacks. Consequently, the development of effective anti-drone laser systems has become increasingly vital. Our study focuses on three main objectives: modeling internal reliability, identifying critical components, and studying the factors affecting the reliability of anti-drone systems. We aim to enhance the overall performance and effectiveness of anti-drone laser systems by analyzing the reliability of critical components and understanding how system parameters influence system reliability. To this end, reliability block diagram (RBD) methodology has been employed to compute the reliability of the laser subsystem in the anti-drone system. Additionally, we conduct a comprehensive review of component-wise reliability to identify vulnerable points within the system, thus enabling targeted improvements and optimizations. To capture the realistic scenario of system failure behavior, different distributions have been used to compute the reliability of the system, ensuring a thorough understanding of its operational reliability in diverse conditions. Finally, the energy values and probability of hitting are obtained for the anti-drone laser system to effectively mitigate environmental challenges.

Keywords: Reliability Block Diagram, Laser Source Subsystem, Weibull Distribution, Mean Time to Failure (MTTF), Rayleigh Distribution, Exponential Distribution, Environmental Factors.

1. INTRODUCTION

Drones have swiftly become an integral part of modern life, finding widespread use across various sectors. While initially linked mainly to military operations, drones now play vital roles in civilian domains. Their applications are diverse, spanning entertainment (aerial photography, videography), geology (mapping, surveying), transportation (traffic monitoring), security (search and rescue, crowd monitoring, disaster relief), shipping (parcel delivery), agriculture (crop monitoring, spraying), and communication (emergency infrastructure). These innovative uses mark a significant shift towards a more autonomous society, with drones poised to revolutionize various aspects of daily life.

In today's security landscape, the threat posed by unauthorized drones operating in restricted airspace is a growing concern. These drones can be utilized for various malicious activities, including surveillance, espionage, and even physical attacks. To address this threat effectively, the development of robust anti-drone laser systems has become imperative. Figure 1 represents the laser implemented anti-drone systems.

An anti-drone laser system serves as a critical security technology designed to detect, track, and neutralize unauthorized drones operating in restricted airspace. By employing advanced detection mechanisms, precise tracking capabilities, and effective neutralization methods, these systems aim to safeguard sensitive areas from the potential risks posed by rogue drones. Laser weapons are emerging as a potent solution for countering the escalating threat posed by drones, leveraging their rapid light-speed engagement, pinpoint accuracy in beam targeting, and cost-effectiveness per shot [19]. In order to analyze the strike capability of laser to drone engine, a comprehensive assessment method for the vulnerability of target to laser is studied in [16]. In [24], Ball suggested that assessing a target's vulnerability to lasers parallels evaluating how non-explosive penetrating objects cause damage when striking a target, although a detailed method was not explicitly outlined.

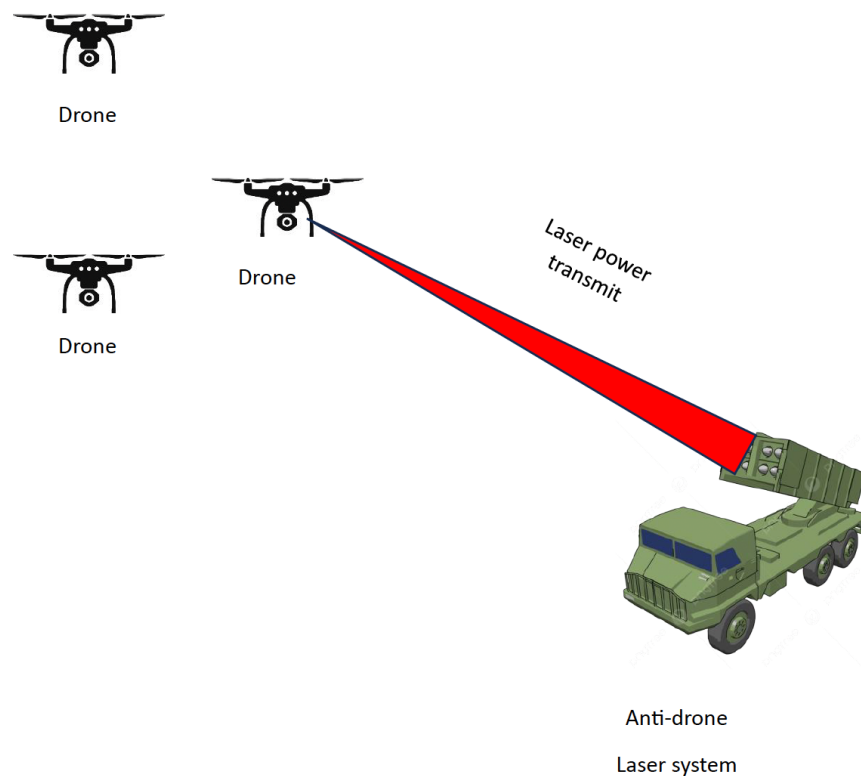


Figure 1: *Anti-Drone Laser System*

1.1. Literature Review

Drones are poised to become a significant factor in future warfare, driven by advancements in artificial intelligence (AI) and information technology. Simultaneously, drones pose a significant challenge to conventional air defense systems. Concerning this practical point of view anti-drone systems are designed and developed. Presently, the majority of anti-drone systems utilize military-grade components to ensure the definitive destruction of hostile drones. Military anti-drone measures commonly employ jamming systems to disrupt the control channel of the target drone [36]. However, in non-military contexts, using RF jamming to thwart fast-moving drones poses

the risk of temporarily disabling existing wireless network systems, such as mobile access or wireless sensor networks [1], [2]. Therefore, the majority of national regulations prohibit the non-military deployment of jamming systems, compelling civilian anti-drone systems to explore alternative methods to halt illegal or unauthorized drones. To destroy or neutralize the illegal drone, different destructive technology is studied. Out of all destructive technology, the most useful technologies are laser, killer drone and anti-aircraft weapons.

Killer drones are referred to as legally operated drone designed to track and neutralize target drones by inflicting damage upon them [18]. Killer drones are remotely operated aircraft equipped with weaponry designed to engage and neutralize targets. These drones have gained significant attention due to their role in modern warfare, intelligence gathering, and targeted strikes. The development of killer drones represents a paradigm shift in military tactics, providing armed forces with unprecedented capabilities in surveillance, reconnaissance, and precision strikes. The legal and ethical considerations surrounding the use of killer drones, also known as armed drones, are complex and multifaceted [23].

Another useful destructive neutralization technology for an anti-drone system is a laser power transmitter. Laser power in anti-drone systems plays a pivotal role in countering the proliferation of drones and protecting critical infrastructure, military installations, and public events from potential threats posed by unauthorized drones. Laser based anti-drone systems utilize directed energy technology to disable or destroy hostile drones through the focused emission of high-energy laser beams. In [15], Huang et al. designed and developed laser integrated anti-drone system. Laser integrated anti-drone system combines cutting-edge laser technology with advanced sensors, tracking systems, and command-and-control interfaces to provide a comprehensive and effective defense against hostile drones.

The key components and capabilities of a laser integrated anti-drone system are:

High-Energy Laser Weapons: Compared to traditional projectile weapons, high-energy laser weapons are particularly well-suited for countering such threats due to their precise and scalable effects, with minimal collateral damage [38]. In [17], Lyu and Zhan made a comprehensive overview of the current state of high-energy laser weapons on a global scale. High energy laser weapons nowadays are most useful in military application to protect the country from the evil. The overview of different technologies in high-energy laser systems encompasses both strategic and tactical roles for high-energy laser weapons on the modern battlefield. It delves into the current performance limitations of weapon system components, including various types of laser devices, beam control systems, atmospheric propagation, and issues related to targetting the killing power [21].

Detection system: Laser integrated anti-drone system incorporates a sophisticated sensor system comprising radar systems, electro-optical/infrared cameras, and radio frequency (RF) detectors. These detectors provide comprehensive situational awareness, enabling operators to detect, track, and classify incoming drones with precision. By integrating data from multiple sensors, laser integrated anti-drone system enhances its ability to identify threats and mitigate false alarms. In [3], Abunada et al. discussed a drone detection mechanism utilizing the RF control signal exchanged between the drone and its remote controller. Wang et al. [37] studied the problems and difficulties of existing radar detection technology for small drone detection and provided better outlook on the development of detection technology. The laser can serve as a supplementary sensor, complementing others such as radar to detect, recognize, and track the drone. Additionally it can dazzle and destroy the drone's optical sensors, enhancing its defensive capabilities. In [32], Steinvall examined diverse laser functionalities and their significance in detecting, identifying, tracking, and countering a drone.

Tracking System: A robust tracking and targeting system forms an integral part of laser integrated anti-drone system, allowing operators to accurately track the movement of hostile drones and maintain a precise lock on the target throughout engagement. This system utilizes advanced algorithms and predictive modeling to compensate for factors such as target motion, atmospheric conditions, and platform dynamics, ensuring optimal laser beam placement for effective engagement. In [33], Steinvall explored the impact of atmospheric conditions and beam jitter resulting

from tracking and platform pointing errors on the effectiveness of the laser, whether employed as a sensor, countermeasure, or weapon.

Control Unit: Laser integrated anti-drone system features a user-friendly control unit that enables operators to monitor system status, analyze threat data, and execute engagement protocols with ease. The unit provides real-time feedback on target tracking, laser engagement, and system performance, allowing operators to make informed decisions and adjust tactics as necessary. In [8], Chen et al. presented the design, simulation, control scheme, and implementation of a capture mechanism. A high-power laser is chosen to control the motor motion, ensuring synchronous operation of the surrounding six launch mechanisms. Abunada et al. [3] have proposed a study aimed at devising a systematic design for a drone detection mechanism utilizing the RF control signal exchanged between the drone and its remote controller. The proposed system entails the generation of a high-power jamming signal transmitted over the identical carrier frequency and band of the detected drone. This jamming signal is then directed toward the drone's location with the intent of disconnecting it from its controller, thereby facilitating a safe landing or activating a mechanism to prompt its return to home. However To destroy the illegal drones, Shi et al. [29] conducted a comprehensive review of the technologies employed in drone surveillance as well as the current anti-drone systems.

Communication system: The communication system facilitates real-time data exchange and coordination between different subsystems of the laser integrated anti-drone system, including sensors, tracking systems, laser emitters, command centers, and operator interfaces. This enables synchronized operation and response to detected threats. Laser beam steering plays a critical role in a wide array of applications, including military targeting and surveillance, space communication, optical data storage, and diverse medical procedures. In [6], Chaudhay et al. provided an extensive literature review on the multifaceted aspects of laser beam pointing and stabilization through the utilization of a fast steering mirror.

In the previous research work, very few authors developed the mathematical model of an anti-drone system. Among them, Garcia et al. [12] studied a simulation model for visual based anti-drone system to detect a UAV. The network's accuracy is 93.40% and it has successfully detected the UAVs. In [39], Zheng et al. developed a simulation model to find the accuracy of visual detection anti-drone system. Chen et al. [9] studied a dynamical modelling and simulation for a capture technology based anti-drone system. To get a brief idea about the comparison of our work with some existing works, Table 1 is provided.

In the literature, no work has been recorded for the reliability analysis of an anti-drone laser system. This motivates us to study the reliability analysis of an anti-drone laser system. The proposed work is the first one to analyse the reliability of an anti-drone laser system. To assess the reliability of an anti-drone laser system, a stochastic model has been constructed. For the model analysis, the reliability block diagram (RBD) is proposed. In [11], Fesenko et al. presented an RBD model to find the reliability of drone systems. The analytical results of reliability are verified using the simulation approach.

Significant research has been dedicated to assessing the reliability of diverse communication systems, including drones and high-altitude platforms. Vishnevsky et al. [35] provided a comprehensive overview of recent advancements in k -out-of- n system theory, particularly applicable to reliability evaluations of high-altitude unmanned platforms. Similarly, Selvamuthu et al. ([27] developed a Markov model to analyze tether reliability in high-altitude platforms. Chen et al. [7] focused on reliability modeling of the NASA Remote Exploration and Experimentation system, employing fault trees and stochastic reward nets. Vishnevsky et al. [34] examined the reliability of tethered high-altitude telecommunication platform modules using k -out-of- $n:F$ models. Gautam and Dharmaraja [13] proposed hierarchical models for LTE-A networks, while the studies [26, 28] explored reliability in UMTS and VANET, respectively. However Feng et al. [10] discussed the optimization model to maximize the mission reliability by changing anti-drone number.

Table 1: Comparisons with the related works

Related works	Anti-Drone	Detection technology	Neutralization technology	Purpose	Mathematical Models
Korsoveczki et al. [14]	✓	Radar	-	Detection in hovering and maneuvering circumstances	Simulation
Multerer et al. [20] [15]	✓	Radar	Jamming	Detect and jamming the signal In long rang confirmatory destruction	-
	✓	-	Laser		-
Pisa et al. [22]	✓	Radar	-	Evaluation of mono-static Radar cross section	Simulation
Shin et al. [30]	✓	Position tracking	-	To detect and track drones	Simulation
Zhou et al. [40]	✓	Fast steering mirror optical focusing system	Laser	To detect and track drones	Simulation
Mohamed and Somaya[5]	✓	optical focusing system	Laser	To detect and destroy the rogue drones	Simulation
Steinvall [31]	✓	Video Sensor	Laser	To detect and destroy drones	-

1.2. Contribution

Our study encompasses three primary objectives.

1. Firstly, we aim to delve into the internal reliability of anti-drone Laser systems through a comprehensive component-wise analysis. This involves a meticulous examination of each component to identify and assess its reliability. By doing so, we strive to enhance the overall performance and effectiveness of the anti-drone laser system.
2. Secondly, we endeavor to identify the critical components within the anti-drone Laser system. This will be achieved by utilizing statistical distributions for reliability analysis. Understanding the reliability of these critical components is paramount as it enables us to prioritize maintenance efforts and ensure uninterrupted operation of the system.
3. Lastly, we seek to determine the optimal number of lasers required in the anti-drone Laser system and analyze the contributing factors influencing this decision. By leveraging statistical analysis and considering various operational scenarios, we aim to optimize the system configuration to achieve maximum efficiency and effectiveness in neutralizing unauthorized drones.

This work is arranged into five sections. In Section 2, a model for the calculation of reliability has been introduced. Section 3 describes the reliability analysis of the laser source subsystem with the numerical illustration. Further, by considering environmental factors, the energy values and probability of hitting of the laser anti-drone system are obtained in Section 4. Finally, the underlying model is concluded with insight for future works in Section 5.

2. RELIABILITY BLOCK DIAGRAMS

RBDs depict system reliability by illustrating the interconnection of components and their potential failure modes. In constructing RBDs, three fundamental patterns of component connections are

employed: (i) Series connection, where components are arranged linearly, rendering the system susceptible to failure if any component fails; (ii) Active Redundancy, involving the simultaneous activation of identical components to ensure system functionality despite failures; and (iii) Standby Redundancy, exemplified by the k -out-of- n configuration, where only a subset of components is active, with others serving as backups, ready to be deployed if necessary. These methods offer diverse approaches to enhancing system reliability and resilience within the framework of RBD analysis.

2.1. Components of an Anti-Drone Laser System

1. **Power Supply:** Provides electrical power to the entire anti-drone system, ensuring continuous and stable power for all the components.
2. **Detection System:** Utilizes sensors technology and triggering mechanism to detect the presence of drones.
3. **Tracking System:** Tracks the movements of detected drones and provides precise targeting information to the laser anti-drone system.
4. **Laser System:** Emits a laser beam to disable and neutralize the targeted drone. Combination of laser diode, focusing lenses, adjustable lenses and, water cooling system.
5. **Control Unit:** Manages and coordinates the overall operation of the anti-drone laser system. Controls the activation and deactivation of the anti-drone laser system based on tracking information.
6. **Communication System:** Facilitates communication between the anti-drone laser system and external command/control centers. Enables the system to receive commands and transmit status updates.
7. **Environment:** Environmental factors significantly impact the performance and reliability of the anti-drone laser system. Variables such as temperature, humidity, wind, and visibility can influence the effectiveness of the laser beam in neutralizing drones.

In Figure 2, all seven components are arranged in series, implying that if one component fails, the failure will cascade throughout the entire mission, ultimately leading to mission failure. Our focus will be on addressing the two most critical components: the laser system and the environment. For simplicity, we assume a reliability value of 1 for all other components, emphasizing the critical importance of addressing issues related to the laser system and environmental factors.

2.2. Components of a Laser System

1. **DC-DC Step Down Converter Voltage Regulator:** Regulates the voltage supplied to the laser diode for optimal performance. This device converts one DC voltage level to another. In the case of a step-down converter, it reduces the input DC voltage level to a lower output DC voltage level. This conversion process is achieved through electronic components such as transistors, diodes, and capacitors.
2. **Laser Diode:** Emits the laser beam for targeting and disabling drones.
3. **Focusing Lenses and Adjustable Lenses:** Focusing Lenses concentrate the laser beam for precise targeting. Adjustable Lenses fine-tune the focus and direction of the laser beam.
4. **Water Cooling System:** Prevents overheating of the laser diode during prolonged use.

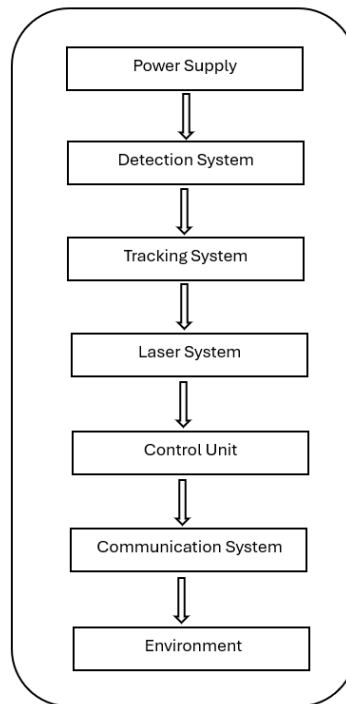


Figure 2: Block Diagram of Anti-Drone Laser System

In Figure 3, all components are indeed interconnected in series, with the exception of both lenses, as they constitute integral parts of a single component to adjust and focus the laser beam.

3. RELIABILITY ANALYSIS OF LASER SOURCE SUBSYSTEM

3.1. Assumption

1. We assume the reliability of power supply, detection system, tracking, control unit, communication unit to be 1 that is, they are completely reliable.
2. Under the laser system inside the anti drone system, we specifically consider the most critical subsystem which is the laser source subsystem.
3. The other subsystems- focusing lenses, cooling system and voltage regulator are expected to withstand for longer intervals.

3.2. Reliability Analysis

In this section, the reliability analysis of the laser source subsystem will be performed using different distributions.

3.2.1 Weibull Distribution

To create a mathematical model for reliability analysis of the laser source, we'll consider the failure behavior of the laser source over time. Since failure rates may vary over time and the failure behavior of electronic components often follows the bathtub curve (initial high failure rate, followed by a period of low failure rate, and then an increase in failure rate over time), the Weibull distribution is used for reliability modeling.

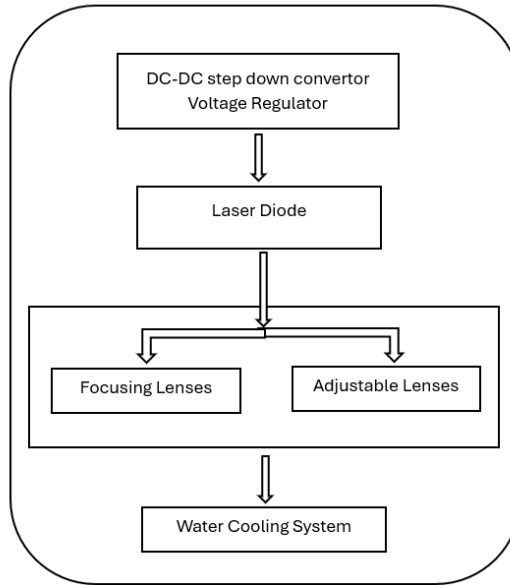


Figure 3: Block Diagram of Laser System

1. Mathematical model

The Weibull distribution has the probability density function (PDF) given by:

$$f(t) = \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1} e^{-(t/\lambda)^k}, t \geq 0$$

where:

- t is the time variable.
- k is the shape parameter (reflects the failure behavior: $k < 1$ for decreasing failure rate, $k = 1$ for constant failure rate, $k > 1$ for increasing failure rate).
- λ is the scale parameter (reflects the characteristic life or scale of the distribution).

The reliability function $R(t)$, which represents the probability that the laser source will function without failure up to time t , is the complement of the cumulative distribution function (CDF) of the Weibull distribution:

$$R(t) = 1 - F(t)$$

where the CDF $F(t)$ is given by:

$$F(t) = 1 - e^{-(t/\lambda)^k}.$$

The reliability function $R(t)$ for the Weibull distribution is given by [4]:

$$R(t) = e^{-(\frac{t}{\lambda})^k}.$$

2. Analysis:

- (a) **Scale Parameter (λ) Estimation:** The estimated scale parameter λ represents the characteristic life of the laser source. The range of mean time to failure (MTTF) [25] values have been taken and corresponding lambda values as listed in table 1 have been calculated using the stated equation

$$MTTF = \lambda \Gamma \left(1 + \frac{1}{k}\right). \tag{1}$$

Table 2: MTTF for Different Values of λ and k

MTTF	λ ($k = 0.8$)	λ ($k = 1$)	λ ($k = 1.2$)
500	442	500	532
1500	1327	1500	1595
2500	2212	2500	2658
3500	3097	3500	3721
4500	3982	4500	4784

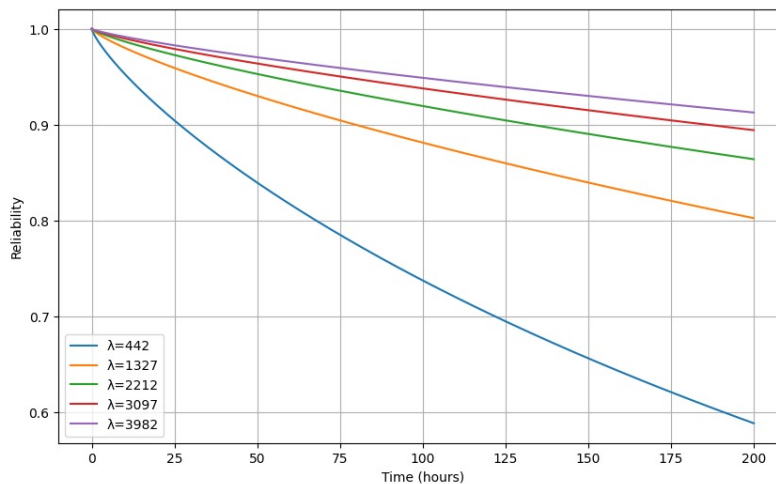


Figure 4: Reliability of laser subsystem for Weibull distribution ($k = 0.8$)

(b) **Shape Parameter (k) Estimation:** The estimated shape parameter k reflects the failure behavior of the laser source.

- A shape parameter $k < 1$ suggests that the laser source is prone to early-life failures. This means the failure rate decreases over time, indicating a “burn-in” period where defective units fail early.
- If $k = 1$, it indicates that the failure rate is constant over time, implying a constant hazard rate or “random failures” occurring independent of time.
- When $k > 1$, it suggests that the laser source experiences wear-out failures, where the failure rate increases over time due to aging or degradation of components. This is the most relevant case for the laser source subsystem as more failures are observed in later stages.

In Table 2 values of λ have been calculated for listed MTTF values and assumed k values. The corresponding λ values for $k = 0.8$ have been then used to plot the reliability time graph (Figure 4) for the first 200 working hours of laser subsystem assumed to follow a Weibull distribution. Figure 5 represents the reliability time plots for $k = 1$ that is constant failure rates and λ values as calculated in Table 2 for $k = 1$. Figure 6 shows the reliability time plots for $k = 1.2$ and corresponding λ values as obtained from Table 2. The $k = 1.2$ that is the increasing failure rate case is the relevant case for laser subsystem. For the laser subsystem with MTTF values above 2500 hours the reliability is greater than 0.95 for the observed period of 200 hours.

CDF $F(t)$ when we have n number of lasers in the system is given by:

$$F(t) = \left(1 - e^{-(t/\lambda)^k}\right)^n.$$

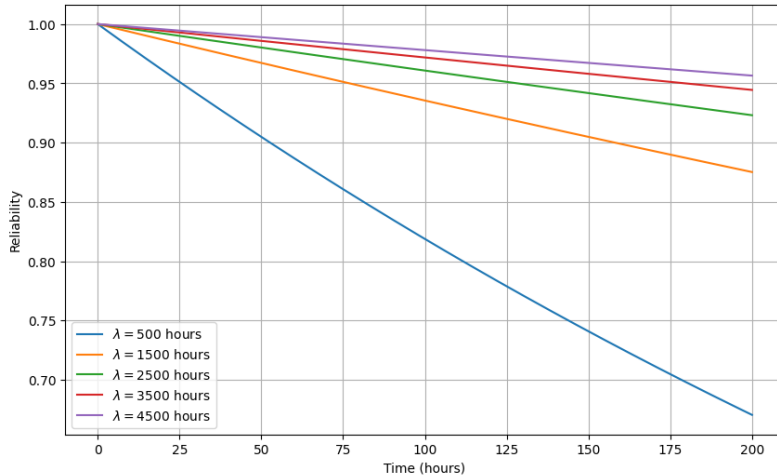


Figure 5: Reliability of laser subsystem for Weibull distribution ($k = 1$)

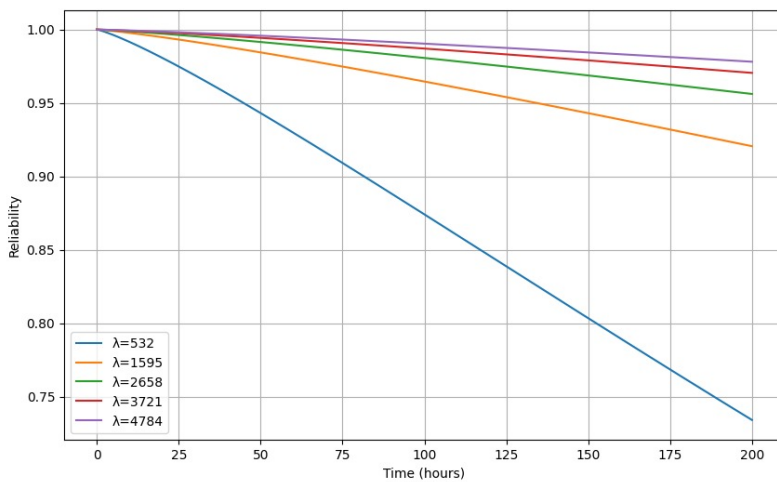


Figure 6: Reliability of laser subsystem for Weibull distribution ($k = 1.2$)

The reliability function $R(t)$ for the Weibull distribution is given by:

$$R(t) = 1 - \left(1 - e^{-(t/\lambda)^k}\right)^n$$

Figure 7 represents the reliability time plots for system with number of lasers = 2, 4, 6 and 8. The failure rate of each of these lasers is assumed to follow the Weibull distribution ($k = 1.2$) with mean time to failure taken as 4000 hours.

3.2.2 Rayleigh Distribution

The Rayleigh distribution is often used to model the time-to-failure of systems where failure events are influenced by the accumulation of random factors. It can model a wide range of failure patterns, from early-life failures to wear-out failures.

1. Mathematical Model

Notations:

- T : Time-to-failure of the laser source subsystem, which follows a Rayleigh distribution.
- σ : Scale parameter of the Rayleigh distribution, representing the characteristic life of the laser source subsystem.

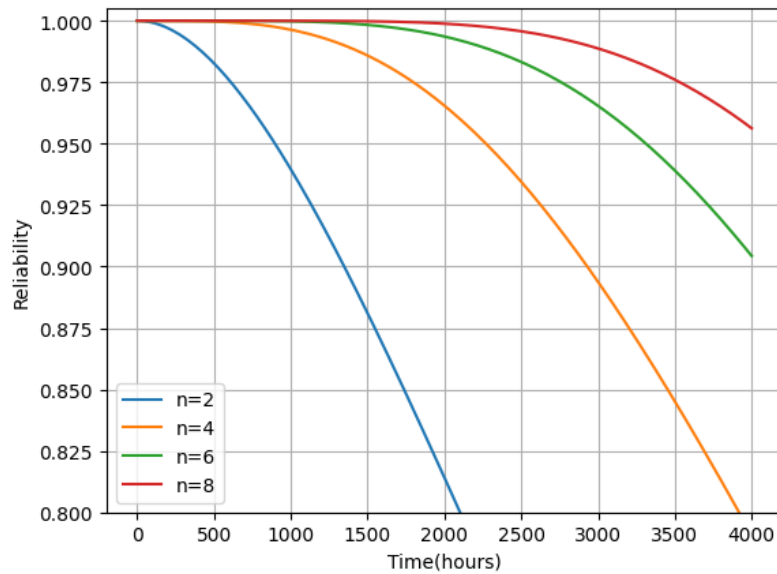


Figure 7: Reliability of laser subsystem (Weibull distribution) for different number of lasers

PDF

$$f(t; \sigma) = \frac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}}, \quad t \geq 0.$$

Reliability Function

$$R(t; \sigma) = 1 - F(t; \sigma)$$

where CDF $F(t; \sigma)$ is given by:

$$F(t; \sigma) = 1 - e^{-\frac{t^2}{2\sigma^2}}.$$

2. Analysis

A larger value of σ indicates a longer characteristic life of the laser source subsystem, while a smaller value suggests a shorter characteristic life. In Table 3 corresponding σ values are calculated for the MTTF values. Figure 8 represents reliability time plots for these σ values obtained.

Table 3: MTTF vs σ

MTTF	σ
500	399
1500	1197
2500	1995
3500	2793
4500	3591

CDF $F(t)$ when we have n number of lasers in the system is given by:

$$F(t; \sigma) = \left(1 - e^{-\frac{t^2}{2\sigma^2}} \right)^n.$$

The reliability function $R(t)$ for the rayleigh distribution is given by:

$$R(t) = 1 - \left(1 - e^{-\frac{t^2}{2\sigma^2}} \right)^n.$$

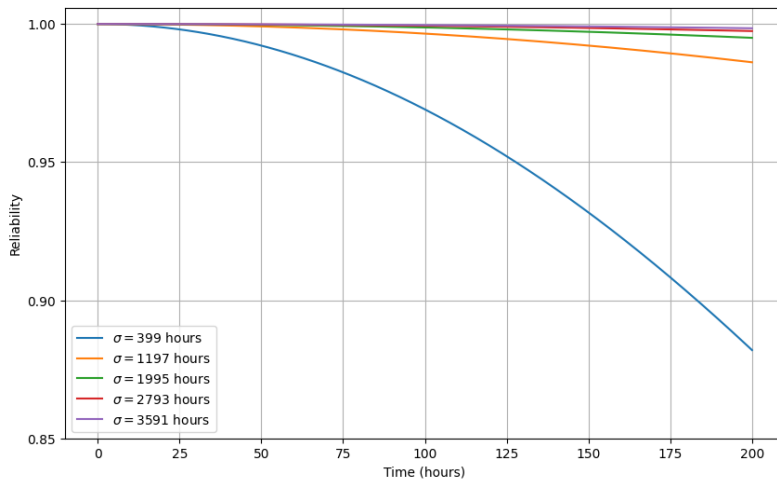


Figure 8: Reliability of laser subsystem for Rayleigh distribution

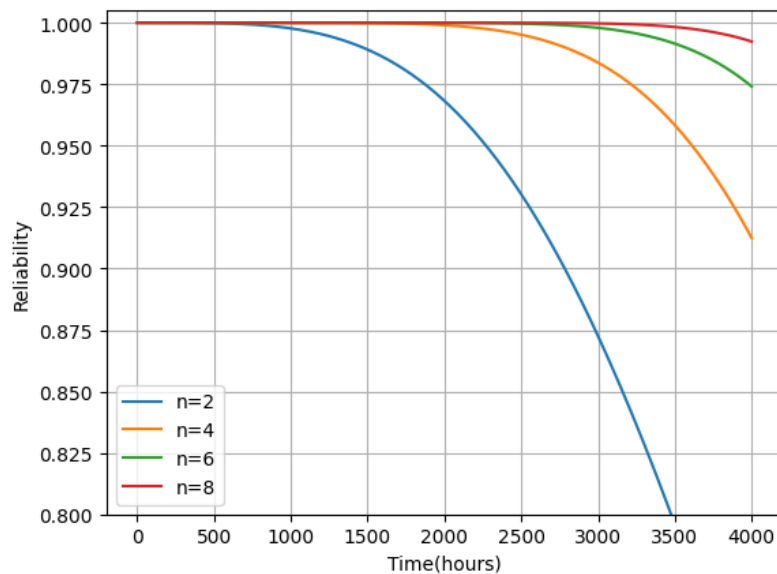


Figure 9: Reliability of laser subsystem (Rayleigh distribution) for different number of lasers

Figure 9 represents the reliability time plots for system with number of lasers= 2,4,6 and 8. The failure rate of each of these lasers is assumed to follow rayleigh distribution with mean time to failure taken as 4000 hours.

3.2.3 Exponential Distribution

The exponential distribution is commonly used to model the time between events in a Poisson process, where events occur continuously and independently at a constant average rate [4]. It is characterized by a single parameter, the failure rate λ .

1. Mathematical Model

Notations:

- T : Time-to-failure of the laser source subsystem, which follows an exponential distribution.
- λ : Failure rate of the exponential distribution, representing the average number of failures per unit time.

The PDF is given as follows

$$f(t; \lambda) = \lambda e^{-\lambda t}, t \geq 0.$$

The Reliability function is given as follows.

$$R(t; \lambda) = e^{-\lambda t}.$$

2. Analysis

The MTTF for an exponential distribution is the reciprocal of the failure rate λ , given by:

$$\text{MTTF} = \frac{1}{\lambda}.$$

Table 4: MTTF for Different Values of λ

MTTF	λ
500	0.00200
1500	0.00067
2500	0.00040
3500	0.00029
4500	0.00022

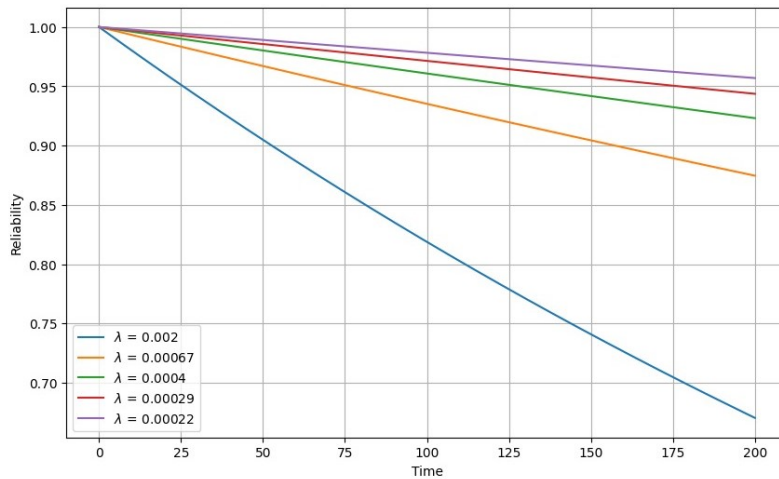


Figure 10: Reliability of laser subsystem for Exponential distribution

Once λ is estimated, it can be interpreted in the context of the laser source subsystem's reliability. A larger value of λ indicates a higher failure rate, meaning shorter time-to-failure, while a smaller value suggests a lower failure rate and longer time-to-failure. The Table 4 maps the corresponding λ values according to the MTTF values obtained using the MTTF formula. Figure 10 represents the plots of laser subsystems with obtained λ values assumed to follow exponential distribution.

CDF $F(t)$ when we have n number of lasers in the system is given by:

$$F(t) = (1 - e^{-\lambda t})^n.$$

The reliability function $R(t)$ for the exponential distribution is given by:

$$R(t) = 1 - (1 - e^{-\lambda t})^n.$$

Figure 11 represents the reliability time plots for system with number of lasers= 2,4,6 and 8. The failure rate of each of these lasers is assumed to follow exponential distribution with mean time to failure taken as 4000 hours.

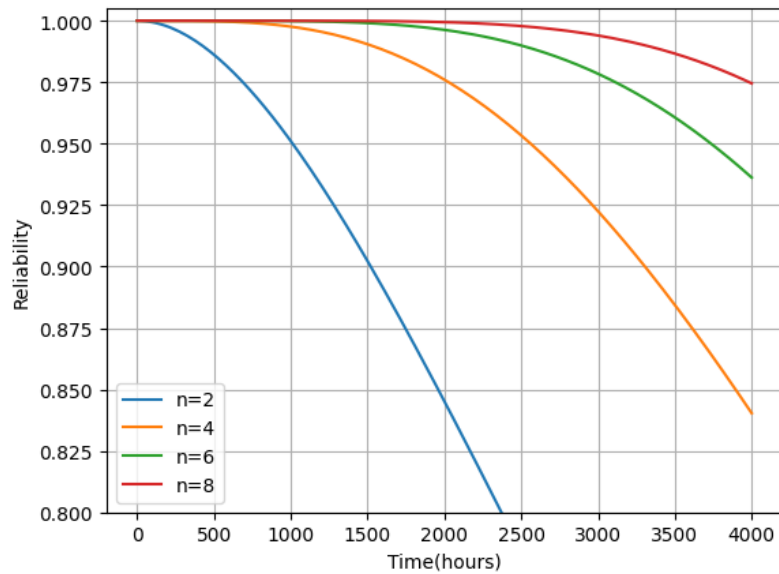


Figure 11: Reliability of laser subsystem (Exponential distribution) for different number of lasers

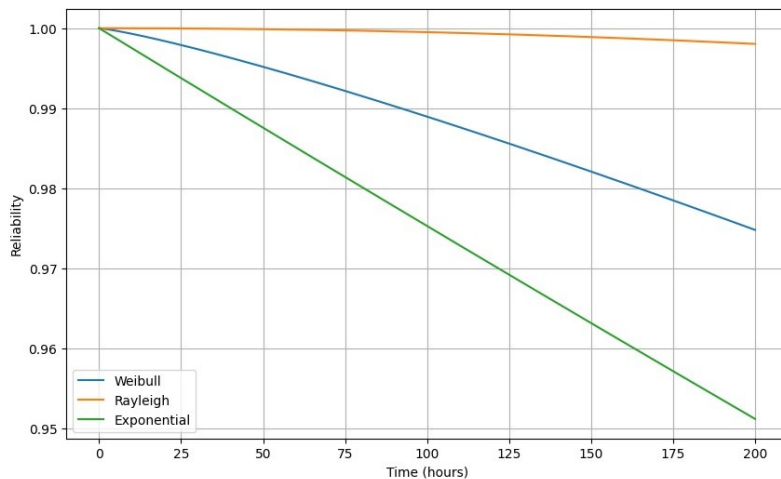


Figure 12: Reliability comparison for different distributions (MTTF = 4000 hours)

Figure 12 represents the reliability time curves assuming the laser subsystem follows the three distributions: Weibull ($k = 1.2$), Rayleigh and exponential with MTTF of 4000 hours. This depicts the reliability values comparison over time with Rayleigh distribution depicting the highest reliability.

Figure 13 represents the reliability time curves assuming the laser subsystem follows the three distributions: Weibull ($k = 1.2$), Rayleigh and exponential with MTTF of 4000 hours for six number of lasers. This depicts the reliability values comparison over time with Rayleigh distribution depicting the highest reliability. The reliability is better than 0.995 (the desirable range) for the first 2000 hours of operation for all three distributions for $n = 6$.

4. ENVIRONMENTAL FACTORS

The effectiveness of laser-based anti-drone systems is significantly influenced by environmental conditions encountered during operation. Environmental factors such as wind, fog, snow, and atmospheric turbulence vary spatially and temporally, posing challenges to the performance and

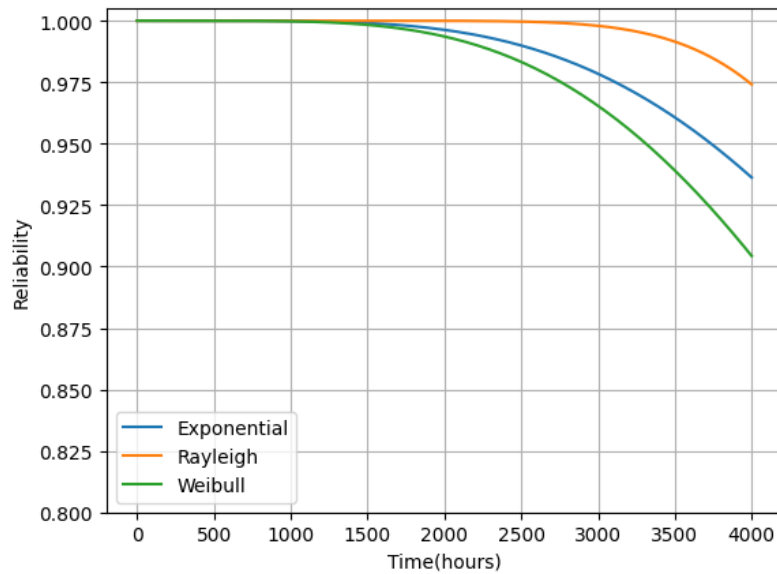


Figure 13: Reliability comparison for different distributions (MTTF = 4000 hours, n=6)

accuracy of the system. These conditions directly impact parameters crucial for laser propagation, such as diffraction and turbulence, ultimately affecting the beam quality (effectiveness in damaging drones) and targeting precision of the anti-drone system.

Contributing factors: The propagation of laser beams in the atmosphere is subject to the effects of diffraction and turbulence, which can significantly impact their trajectory and intensity. These effects can be quantified using mathematical formulations that consider various environmental parameters.

The spreading angle of laser beam due to diffraction and beam quality ($\theta_{\text{diff\&qual}}$) are determined by the wavelength of the laser beam (λ), the diameter of its emitting aperture (D), and beam quality factor (M^2):

$$\theta_{\text{diff\&qual}} = M^2 \frac{C\lambda}{D} \quad (2)$$

where, C is a constant, usually set to be 1.22.

The spreading angle resulting from atmospheric turbulence (θ_{turb}) is calculated based on the Fried parameter (r_0), which is influenced by the wavelength (λ), path length of laser beam (target distance, L) and the refractive index structure constant (C_n^2):

$$\theta_{\text{turb}} = \frac{1.6\lambda}{\pi r_0} \quad (3)$$

where,

$$r_0 = 0.184 \left(\frac{\lambda^2}{C_n^2 L} \right)^{3/5} \quad (4)$$

The dependency of these parameters on environmental conditions such as humidity and temperature is crucial for understanding their impact on laser beam propagation. Specifically, the refractive index structure constant (C_n^2) is essential for characterizing atmospheric turbulence and is influenced by the potential refractive index gradient, denoted by M :

$$C_n^2 = a^2 A L_0^{4/3} M^2 \quad (5)$$

where a^2 is a dimensionless constant, most commonly used at a value of 2.8, equal to unity. L_0 is the outer scale of turbulence, which can be set equal to the resolution of the radiosonde data (e.g.,

10 m).

$$M = -\frac{77.6 \times 10^{-6}}{T} \cdot p \cdot \frac{\partial \ln \theta}{\partial z} \cdot \left[1 + \frac{15500 \cdot q}{T} - \frac{15500}{2 \cdot T} \cdot \frac{\left(\frac{dq}{dz}\right)}{\left(\frac{\partial \ln \theta}{\partial z}\right)} \right]$$

where p is average pressure for a layer. ∂z is thickness of layer. T is absolute average temperature. θ is potential temperature. q is specific humidity.

We can vary the refractive index structure constant (C_n^2) with time (i.e., changing temperature, pressure, and humidity with time). To obtain an overall measure of the spreading angle (θ_{total}) of the laser beam considering both diffraction and turbulence effects, we can compute the root mean square (RMS) of the individual spreading angles:

$$\theta_{total} = \sqrt{\theta_{diff}^2 + \theta_{turb}^2} \quad (6)$$

In the context of our mission, the success criterion is met when the energy per unit area reaching the target drone surpasses a predefined threshold energy value. This threshold energy level, denoted as e_{th} , represents the minimum energy required to effectively neutralize the drone. Upon meeting this condition, the drone is effectively destroyed.

To quantify this condition, we employ the concept of the brightness (B) of our laser weapon system. This brightness must exceed the ratio of the square of the target distance (L^2) to the threshold energy (e_{th}) multiplied by the dwell time of the laser beam (τ).

Mathematically, this relationship is represented as:

$$B = \frac{P_0 \cdot (1 - \alpha)^L}{\pi \theta_{total}^2} > \frac{L^2 \cdot e_{th}}{\tau} \quad (7)$$

Here, P_0 denotes the total output beam power, α represents the atmospheric attenuation ratio of the laser beam, L signifies the target distance measured in kilometers, θ represents the root mean square of both the diffraction and quality of the laser beam, as well as the atmospheric turbulence, and τ denotes the dwell time of the laser beam irradiated on the target.

$$\text{Energy_on_target} = \frac{P_0 \cdot \tau \cdot (1 - e^{-\eta \cdot L})^L}{\pi \cdot \theta_{total}^2 \cdot L^2} \quad (8)$$

We are utilizing the fixed laser system with the following parameter values and is shown in Table 5.

Table 5: Parameters Values

Parameters	Values
Beam quality (M)	1.5
Power Output	150KW
Aperture Diameter	50cm
Wavelength	1.045 micro-meter
Target Range	17km
Alpha	0.2
Threshold energy	50KJ

By assuming the parameter values mentioned in Table 5, we have obtained the energy values and probability of hitting which is mentioned in Table 6.

$$\text{Probability of hitting} = 1 - e^{-\frac{\text{Energy_on_target}}{e_{th}}} \quad (9)$$

From Figure 14, it is noticed that as the turbulence increases the reliability of the system decreases. This depicts that environmental turbulence can affect the system reliability. This happens because higher levels of turbulence may introduce more variability or unpredictability into

the system, leading to disruptions or failures. For instance, increased turbulence might cause misalignment or disturbances in the laser beam, impacting its effectiveness or stability.

Therefore, the relationship depicted in Figure 14 suggests that environmental turbulence has a significant impact on the reliability of the system. This underscores the importance of considering environmental factors when designing or operating the system to ensure its consistent and dependable performance.

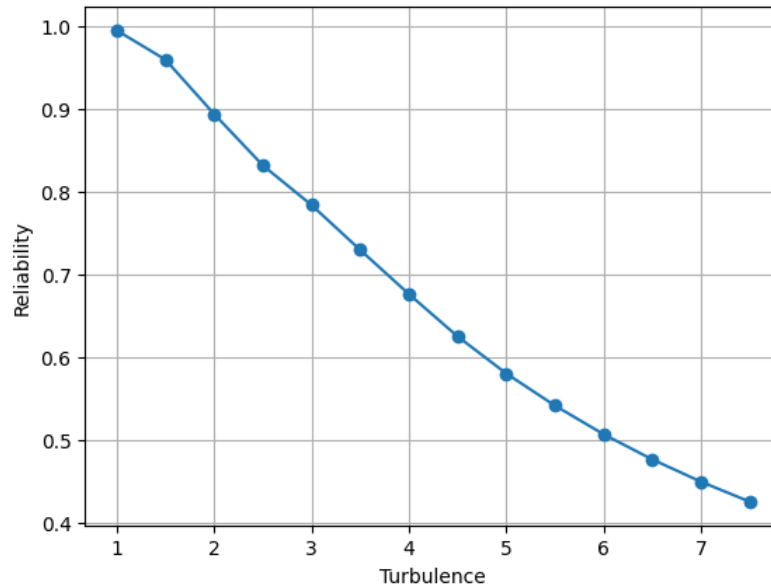


Figure 14: Reliability values at different turbulence levels

Table 6: Energy Values and Probability of Hitting

Turbulence (in $10^{-14} \text{ m}^{-\frac{2}{3}}$)	Energy (KJ)	Probability of hitting
1	257.1968	0.9946
1.5	158.66	0.9593
2	112.53	0.8937
2.5	86	0.8317
3	69.28	0.7838
3.5	57	0.7296
4	49.09	0.6763
4.5	42	0.6252
5	37.57	0.5804
5.5	33.51	0.5416
6	30.1995	0.5070
6.5	27.43	0.4766
7	25.1039	0.4497
7.5	23.11	0.4256

5. CONCLUSIONS AND FUTURE DIRECTION

The proposed study present reliability analysis of laser subsystem in anti-drone system using RBD methodology. To represent the failure behaviour of the reliability blocks, Weibull and Rayleigh distribution have been used. Numerical results are presented to demonstrate the behaviour of

reliability of the system with respect to several parameters of the system. Further, the study focus on investigating and modeling the environmental factors that impact the reliability of anti-drone laser systems. Factors such as temperature, humidity, wind, and visibility can significantly affect the system's performance. By gaining a comprehensive understanding of these environmental variables, we obtain the energy values and probability of hitting of the anti-drone laser system to effectively mitigate environmental challenges.

The parameters involved in the reliability analysis of the anti-drone system will be estimated in future work. Furthermore, various stochastic models, such as Markov model, semi-Markov model, etc., will be constructed to evaluate the reliability of the anti-drone system. To verify the validity of the obtained results, simulation analysis will be conducted.

REFERENCES

- [1] Uk public general acts, wireless telegraphy act. *U Legislation, London, U.K.*, 2006.
- [2] Fcc enforcement advisory, cell jammers, gps jammers, and other jamming devices. *document FCC RCD, 1329*, 2017.
- [3] Abdularahman Hasan Abunada, Ahmed Yousif Osman, Amith Khandakar, Muhammad Enamul Hoque Chowdhury, Tamer Khattab, and Farid Touati. Design and implementation of a rf based anti-drone system. pages 35–42. 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIOT), 2020.
- [4] Liliana Blanco Castañeda, Viswanathan Arunachalam, and Dharmaraja Selvamuthu. *Introduction to probability and stochastic processes with applications*. John Wiley & Sons, 2012.
- [5] Mohamed Zied Chaari and Al-Maadeed Somaya. Testing the efficiency of laser technology to destroy the rogue drones. *Security and Defence Quarterly*, 32(5):31–38, 2020.
- [6] Himanshu Chaudhary, Shahida Khatoun, Ravindra Singh, and Ashish Pandey. Fast steering mirror for optical fine pointing applications: A review paper applications: A review paper. 2018 3rd International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH), pages 1–5, 2018.
- [7] Dong Chen, Dharmaraja Selvamuthu, Dongyan Chen, Lei Li, Kishor S Trivedi, Raphael R Some, and Allen P Nikora. Reliability and availability analysis for the jpl remote exploration and experimentation system. pages 337–342. Proceedings International Conference on Dependable Systems and Networks, 2002.
- [8] Yingzi Chen, Zhiqing Li, Longchuan Li, Shugen Ma, Fuchun Zhang, and Chao Fan. An anti-drone device based on capture technology. *Biomimetic Intelligence and Robotics*, 2(3):100060, 2022.
- [9] Yingzi Chen, Zhiqing Li, Longchuan Li, Shugen Ma, Fuchun Zhang, and Chao Fan. An anti-drone device based on capture technology. *Biomimetic Intelligence and Robotics*, 2(3):100060, 2022.
- [10] Qiang Feng, Meng Liu, Hongyan Dui, Yi Ren, Bo Sun, and Dezhen Yang. Importance measure-based phased mission reliability and uav number optimization for swarm. *Reliability Engineering & System Safety*, 223:108478, 2022.
- [11] Herman Fesenko, Oleg Illiashenko, Vyacheslav Kharchenko, Ihor Kliushnikov, Olga Morozova, and Anatoliy Sachenko. Flying sensor and edge network-based advanced air mobility systems: reliability analysis and applications for urban monitoring. *Drones*, 7(7):409, 2023.
- [12] A J Garcia, J Min Lee, and D S Kim. Anti-drone system: A visual-based drone detection using neural networks. 2020 International Conference on Information and Communication Technology Convergence (ICTC), 2:559–561, 2020.
- [13] Anupam Gautam and Dharmaraja Selvamuthu. Reliability and survivability assessment of lte-a architecture and networks. *OPSEARCH*, 60(1):370–392, 2023.
- [14] Gyula Korsoveczki, Miklos Kende Orosz, and Istvan Balajti. Performance analysis of quadcopter drones to radar detection and track initialization characteristics optimization. pages 1–10. 2023 24th International Radar Symposium (IRS), 2023.

- [15] Longwen Liao, Xiaoliang Huang, and Fengyu Xie. Development status and operation analysis of laser weapon in anti-drone warfare. pages 305–310. 2023 IEEE International Conference on Unmanned Systems (ICUS), 2023.
- [16] Le Liu, Chengyang Xu, Changbin Zheng, Sheng Cai, Chunrui Wang, and Jin Guo. Vulnerability assessment of uav engine to laser based on improved shotline method. *Defence Technology*, 33:588–600, 2023.
- [17] C. Lyu and R. Zhan. Global analysis of active defense technologies for unmanned aerial vehicle. *IEEE Aerospace and Electronic Systems Magazine*, 30(1):6–31, 2022.
- [18] Michael. Mayer. The new killer drones: Understanding the strategic implications of next-generation unmanned combat aerial vehicles. *International Affairs*, 91(4):765–780, 2015.
- [19] Zhu Mengzhen, Chen Xia, Liu Xu, Tan Chaoyong, and Li Wei. Situation and key technology of tactical laser anti-uav. *Infrared Laser Eng*, 50(7):20200230–1, 2021.
- [20] Thomas Multerer, Alexander Ganis, Ulrich Prechtel, Enric Miralles, Askold Meusling, Jan Mietzner, Martin Vossiek, Mirko Loghi, and Volker Ziegler. Low-cost jamming system against small drones using a 3d mimo radar based tracking. pages 299–302. 2017 European radar conference (EURAD), 2017.
- [21] Glen P Perram, Michael A Marciniak, and Matthew Goda. High-energy laser weapons: technology overview. *Laser Technologies for Defense and Security*, 5414:1–25, 2004.
- [22] Stefano Pisa, Emanuele Piuze, Erika Pittella, Pierfrancesco Lombardo, and Nertjana Ustalli. Evaluating the radar cross section of the commercial iris drone for anti-drone passive radar source selection. pages 699–703. 2018 22nd International Microwave and Radar Conference (MIKON), 2018.
- [23] V. Ramazani. Killer drones, legal ethics, and the inconvenient referent. *Lateral*, 7(2), 2018.
- [24] Ball RE. The fundamentals of aircraft combat survivability: analysis and design. 2th ed. Reston: American Institute of Aeronautics and Astronautics, 1985.
- [25] Dharmaraja Selvamuthu and Dipayan Das. *Introduction to probability, statistical methods, design of experiments and statistical quality control*. Springer, 2024.
- [26] Dharmaraja Selvamuthu, Vaneeta Jindal, and Upkar Varshney. Reliability and survivability analysis for umts networks: An analytical approach. *IEEE Transactions on Network and Service Management*, 5(3):132–142, 2008.
- [27] Dharmaraja Selvamuthu, Adwaith H Sivama, Raina Raj, and Vladimir Vishnevsky. Study of reliability of the on-tether subsystem of a tethered high-altitude unmanned telecommunication platform. *Reliability: Theory & Applications*, 18(1 (72)):172–178, 2023.
- [28] Dharmaraja Selvamuthu, Resham Vinayak, and Kishor S Trivedi. Reliability and survivability of vehicular ad hoc networks: An analytical approach. *Reliability Engineering & System Safety*, 153:28–38, 2016.
- [29] Xiufang Shi, Chaoqun Yang, Weige Xie, Chao Liang, Zhiguo Shi, and Jiming Chen. Anti-drone system with multiple surveillance technologies: Architecture, implementation, and challenges. *IEEE Communications Magazine*, 56(4):68–74, 2018.
- [30] Jae-Min Shin, Yu-Sin Kim, Tae-Won Ban, Suna Choi, Kyu-Min Kang, and Jong-Yeol Ryu. Position tracking techniques using multiple receivers for anti-drone systems. *Sensors*, 21(1):35, 2020.
- [31] O Steinvall. The potential role of laser in combating UAVs: Part 2; laser as a countermeasure and weapon. pages 14–30. Technologies for Optical Countermeasures XVIII and High-Power Lasers: Technology and Systems, Platforms, Effects V, 2021.
- [32] Ove Steinvall. The potential role of lasers in combating uavs, part 1: detection, tracking, and recognition of uavs. In *Electro-Optical and Infrared Systems: Technology and Applications XVIII and Electro-Optical Remote Sensing XV*, 11866:174–185, 2021.
- [33] Ove Steinvall. Beam tracking and atmospheric influence on laser performance in defeating uav: s. pages 100–116. SPIE Electro-Optical Remote Sensing XVI, 2022.
- [34] Vladimir Vishnevsky, Dharmaraja Selvamuthu, Vladimir Rykov, Dmitry Kozyrev, and Nika Ivanova. Reliability modeling of a flight module of a tethered high-altitude telecommuni-

- cation platform. pages 1–6. 2022 International Conference on Information, Control, and Communication Technologies (ICCT), 2022.
- [35] Vladimir M. Vishnevsky, Dharmaraja Selvamuthu, Rykov Vladimir, V. Kozyrev Dmitry, Ivanova Nika, and Krishnamoorthy Achyutha. *Reliability Assessment of Tethered High-altitude Unmanned Telecommunication Platforms: k-out-of-n Reliability Models and Applications*. Springer Nature, 2023.
- [36] B. R. Van Voorst. Counter drone system. *U.S. Patent 15 443 143*, 14, 2017.
- [37] Ping Wang, Jing Luo, Jiwei Tang, Yi Qin, Jie Cui, and Liangjun Zhang. Progress and prospects of anti-drone radar detection technology. pages 2331–2334. 2021 CIE International Conference on Radar (Radar), 2021.
- [38] Shubin Zhao, Renwei Xie, and Jun Wan. Design of anti-drone laser weapon systems. pages 24–30. SPIE High-Power Lasers and Applications XI, 2020.
- [39] Ye Zheng, Zhang Chen, Dailin Lv, Zhixing Li, Zhenzhong Lan, and Shiyu Zhao. Air-to-air visual detection of micro-UAVs: An experimental evaluation of deep learning. *IEEE Robotics and automation letters*, 6(2):1020–1027, 2021.
- [40] Qingkun Zhou, Pinhas Ben-Tzvi, Dapeng Fan, and Andrew A Goldenberg. Design of fast steering mirror systems for precision laser beams steering. pages 144–149. 2008 International workshop on robotic and sensors environments, 2008.

STRATEGIES FOR REPLACEMENT IN WORKFORCE SCHEDULING RELIABILITY MODELS

Iyappan. M^{1*}, Balaji. M², R. Saranraj³, G. Sathya Priyanka⁴

^{1*}Assistant Professor, Department of Statistics, St. Francis College, Bengaluru-34, India

²Associate Professor, Department of Management, Seshadripuram College, Bengaluru-20, India

³Lecturer in Statistics, Kunthavai Naachiyar Government Arts College for Women, Thanjavur, India

⁴Ph.D Research Scholar, Department of Statistics, Periyar University, Salem-11, India

^{1*}iyappastat@gmail.com, ²balajim0405@gmail.com, ³saranraj316@gmail.com,

⁴sathyapriyankastat@gmail.com

Abstract

It is a typical occurrence to replace some industrial equipment or components, such as electronic chips, bulbs, etc. It deals with the ideas of dependability theory, in which the likelihood of an equipment malfunctioning instantly is calculated assuming that it has operated normally for a given amount of time, 't'. In reliability, it's known as the hazard rate. The rate of hazard may be rising, falling, or staying the same. However, replacement tactics are employed to maintain output. Basically, two distinct approaches are employed. 1. Replacing a broken item; 2. Replacing items on a regular basis. Since it is a preventive measure to maintain production, the effective administration of maintaining system functionality depends on the application of both reliability concepts in addition to replacement theory. One may envision a similar issue with the personnel system as well. To determine the best manpower policies in this situation, replacement methods and dependability theory can also be coupled. This theory's application to labor systems is examined, and appropriate methods for employee replacement and advancement are covered in order to ensure the system's successful upkeep. In order to obtain the best fit, manpower planning is a dynamic process that controls the movement of workers into, through, and out of the organization. In order to determine appropriate manpower replacement policies for promotion and replacement of personnel for the successful maintenance of the system, this study discusses the application of dependability theory and the renewal process.

Keywords: Degradation problem, Life time, Replacing, Reliability model, Normal Distribution, Hazard Rate.

I. Introduction

Manpower planning evaluates the organization's present staffing and skill levels, connects these factors to the market demand for its goods, and offers options to align manpower resources with projected demand. It is a dynamic process that controls the movement of workers into, through, and out of the company in order to find the best fit. The methodical process of using available labour in accordance with the demands of the nation's many industries is known as manpower planning. As a result, manpower models are created while accounting for the diverse real-world scenarios. Manpower refers to a group of individuals who have obtained a specific ability or

expertise to perform a specific kind of work. Developing appropriate economic policies and applying scientific approaches to economic problems are crucial components of effective administration. To determine the best workforce policy, replacement strategies and reliability theory can be coupled. This theory's applicability to labour systems is examined, and appropriate methods for employee replacement and advancement are covered in order to ensure the system's continued viability. Robinson [18] has examined the application of replacement techniques to workforce planning.

The grade I, also known as the training grade, and the grade II, also known as the grade in the organization, are conceptualized as the two divisions that are taken into consideration in this study. In stage I, the cost of an individual replacement is higher than in step II. When an individual in stage II departs the organization, the void is filled by someone in stage I. If any wastage in stage I as a result of exits and transfers from stage I to stage II, after which training grade recruiting is completed. Planning for manpower must take into account the overall business strategy as well as factors that affect employment, such as technology, competition, government laws and regulations, and shifting social norms and expectations. Plans that guarantee important concerns are addressed, suitable measures are performed, and the process is maintained will be effective; this will give the plan credibility and instill a feeling of pride in those who carry it out. When analyzing the planning of human resources, statistical techniques are a vital resource. Predicting demand may involve looking at productivity changes, technological changes, market forces and trends and the corporate strategies. Predicting supply involves knowledge of the current manpower stocks and looking at future recruitment, wastage, working conditions, promotion policies and labour market trends. Closing the gap means examining training, remuneration career planning, redundancies and further consideration of all the factors under the other headings.

II. Methods

I. Model

The current model is designed with two compartments: Grade I, sometimes known as the training grade, and Grade II, which denotes a specific working location or formal hierarchy position. In stage I, the cost of an individual replacement is higher than in step II. When an individual in stage II departs the organization, the void is filled by someone in stage I. Recruitment for training grade is carried out if there is any waste in stage I as a result of exit and transfer from stage I to stage II. It can be seen that the force of separation, sometimes referred to as the hazard rate, is dependent on the total period of service. It may be observed that the force of separation or the propensity of the individual to leave the organization depends upon completed length of service and it is defined as the probability that a person leaves in a small time interval $(\xi, \xi + \delta\xi)$ having served a period of t and $\lambda(\xi) = \frac{f(\xi)}{G(\xi)}$. Under these assumptions it is proposed to determine the size of grade I or the number of persons in the training grade. The mean time to promotion is also to be determined.

In any organization there is loss of manpower, which arises due to the leaving of the personnel from the organization. The rate at which the loss of manpower arises due to the leaving process is very important. If the rate of leaving can be accurately measured then the appropriate policies for the number of persons to be recruited at various time points in the future can be decided. In the study of leaving process the completed length of service (CLS) distribution plays an important role. Hence an appropriate statistical model assuming different probability distributions of the random variable should be fitted to the data obtained from any organization. Several distributions have been suggested for this purpose. Lawless [15] have suggested a model for the

CLS distribution Chien-Yu Peng [11] has suggested a model for the CLS distribution. Initially exponential distribution has been suggested for the CLS distribution. The density function is $f(\xi) = \lambda e - \lambda \xi$ where λ is called the loss intensity. Tweedie [17] has suggested that the parameter of the distribution λ be treated as a random variable whose distribution is $H(\lambda)$. Therefore the CLS distribution in the Silcock form is given by

$$f(\xi) = \int \lambda e^{-\lambda \xi} dH(\lambda), \quad (\xi \geq 0)$$

The mixed exponential distribution has been used by Bartholomew and Forbes [2]. The concept of change of distribution is discussed in Chien-Yu Peng [11]. Meeker [16] has used this concept in shock model and cumulative damage process, to estimate the expected time to cross the threshold, where the threshold random variable Y undergoes a change in the distribution itself. In inventory theory, the probabilistic demand can undergo change in the very distribution itself after a change point. This concept is used in manpower planning problems. Assuming that the wastages in successive decision epochs are correlated random variables, the expected time to recruitment is derived and in doing so the results of Gurland [8] are used. In addition to this assumption it is also assumed that the random variable which denotes the threshold level for wastages is one which satisfies the so called Setting the Clock Back to Zero (SCBZ) property, due to Jerry Lawless [14]. The expression for mean and variance of the expected time to recruitment are derived using the shock model approach by Esary Marshall and Proschan [12].

The term "Manpower planning" defined by Walker [9], "Manpower planning refers to the rather complex task of forecasting and planning for the right numbers and at the right kinds of people at the right places and the right times to perform activities that will benefit both the organization and the individuals in it". According to Bartholomew [1], "Manpower planning is concerned, in aggregate terms, with matching jobs to people. In the broadest sense, manpower planning is concerned with matching the supply of people available for employment with the jobs available". According to Grinold and Marshall [8 & 12], Manpower planning must be an ancient art, since manpower problems have existed for centuries. People, jobs, time and money are the basic ingredients of a manpower system. A decision-maker must be aware of the interactions among these four ingredients in order to formulate and evaluate manpower policy. Manpower planning within an organization has the basic purpose of producing the correct numbers of correct type of people in the correct jobs at appropriate times. System constraints do not often allow for perfect matching of the people to jobs. A more realistic view of manpower planning is that it avoids having too many of the wrong types of people in the wrong jobs too frequently. For a detailed study of the subject in this direction can be seen in Bartholomew [4], Bartholomew and Forbes [6], Nikulin et al. [10], Butler [7].

The reliability $R(\xi)$ of the device is given by

$$R(\xi) = \sum_{k=0}^{\infty} P_k V_k(\xi)$$

where $V_k(\xi)$ is the probability that k damages are caused during $(0, t]$. The above model has been considered by Esary, Marshall and Proschan [12] with the underlying process generating the shocks as Poisson. Gurland [8] has shown that the characteristic function $\phi(\lambda_1, \lambda_2, \dots, \lambda_n)$ of the joint distribution of any n random variables from a sequence $\{X_n\}$ of constantly correlated, exchangeable random variables each following the exponential distribution with p.d.f $f(x) = \left(\frac{1}{a}\right)e^{-\frac{x}{a}}$, $a > 0$, $0 < x < \infty$ such that the correlation coefficient R between any X_i and X_j , $i \neq j$ (independent of i and j) is given by

We suppose that all individuals have the same completed length of service (CLS) distribution, with density function $f(\xi)$, where t denotes the time which has elapsed since the

person in question joined the organization. Thus $f(\xi) \delta t$ is the probability that a man leaves with t period of service in $(\xi, \xi + \delta \xi)$. The survivor function $G(\xi)$ gives the probability that a person remains in the organization for at least time t .

Hence

$$G(x) = \int_{\xi}^{\infty} f(\omega) d\omega \quad (1)$$

We define the further quantity $\lambda(\xi)$, known as the 'force of separation at length of service ξ ', as follows. $\lambda(\xi) \delta \xi$ is the probability that an individual who has been in the organization for a length of time ' ξ ', leaves in the interval $(\xi, \xi + \delta \xi)$. It is easily shown that

$$\lambda(\xi) = \frac{f(\xi)}{G(\xi)}, \text{ for all } \xi \geq 0 \quad (2)$$

Since the input and output of each grade must balance, we have for Grade II

$$N_1 P = N_2 W W_2 \quad (3)$$

Moreover, in equilibrium the expected input to the system over unit time is N/μ , where μ the mean length of completed service. Thus for grade I, we have,

$$\frac{N}{\mu} = N_1(P + W_1) = N_1 W_1 + N_2 W_2 \quad (4)$$

In general W_1, W_2 and P are functions of time. It shows that P, W_1, W_2 tend to equilibrium values which are independent of the age of the system. Let μ_1 be the average time spent in Grade I. Then in equilibrium the expected number of vacancies occurring per unit time in this Grade is N_1/μ_1 . These can be caused either by promotion or losses, and hence

$$\frac{N_1}{\mu_1} = N_1(P + W_1) \quad (5)$$

$$\frac{1}{\mu_1} = P + W_1 \quad (6)$$

It follows that equation (4) and equation (6) that

$$\frac{N}{\mu_1} = \frac{N_1}{\mu_1} \quad (7)$$

II. Promotion by Seniority

Let $a(\xi/T)$ denote the age distribution of the system at time T given that the system was established at $(T = 0)$. Thus $a(\xi/T) \delta T$ is the probability that an individual chosen at random at time T has length of service in $(\xi, \xi + \delta \xi)$. Thus

$$\begin{aligned} a(\xi/T) \delta T &= P\{\text{individual joined in } (T - \xi, T - \xi + \delta \xi) \text{ and remained for time } t\} \\ &= h(\xi - T) \delta T G(\xi) \end{aligned}$$

Where $h(\xi)$ is the renewal density for the whole system, i.e. $h(\xi) \delta \xi$ is the probability of loss in $(\xi, \xi + \delta \xi)$. Hence $a(\xi/T) = h(T - \xi) G(\xi)$. This is true for $\xi < T$, when $\xi = T$, We have

$$\begin{aligned} a(T/T) &= P\{\text{Original member of organization is still there at time } T\} \\ &= G(T) \end{aligned}$$

Now,

$$\lim_{T \rightarrow \infty} h(T) = 1/\mu \quad (8)$$

Whatever the form of the CLS distribution.

Also

$$\lim_{T \rightarrow \infty} G(T) = 0 \quad (9)$$

Thus

$$\lim_{T \rightarrow \infty} a\left(\frac{T}{\mu}\right) = G(\xi)/\mu \quad (10)$$

Now, since a loss from Grade II is replaced by the most senior member of Grade I, it follows that at any time every individual in Grade II has length of service at least as long as any individual in Grade I and hence that there exists some threshold value t_1 such that all individuals with length of service less than t_1 are in Grade I, where t_1 is a random variable. But if the grade sizes are large their expected value can be found from the approximate formula,

$$\int_{t_1}^{\infty} a(\xi) d\xi = N_2 / (N_1 + N_2) \quad (11)$$

The expected number of promotions per unit time will be the proportion of new recruits whose services to the threshold length of service t_1 is

$$N_1 P = G(\xi_1) N / \mu \quad (12)$$

III. Promotion at Random

Let $F_1(\xi)$ be the probability that an individual remains in the system for a time t without being promoted. Let $n\xi$ be the number of promotions in $(0, \xi)$, then

$F_1(\xi) = \text{Prob}\{\text{Individual not promoted in } (0, \xi) / \text{ doesn't leave in } (0, \xi)\} \cdot \text{Prob}\{\text{doesn't leave in } (0, \xi)\}$

$$= \left(1 - \frac{1}{N_1}\right)^{n\xi} G(\xi)$$

Now expected value of $n\xi$ is $N_1 P_1$. Thus, as a first approximation we can take

$$F_1(\xi) = G(\xi) \left(1 - \frac{1}{N_1}\right)^{N_1 P_1 \xi} \text{ as } N_1 \rightarrow \infty, \text{ we have}$$

$$F_1(\xi) = G(\xi) e^{-P\xi}$$

It follows that the average time spent in Grade I

$$\mu_1 = \int_0^{\infty} F_1(\xi) d\xi = \int_0^{\infty} G(\xi) e^{-P\xi} d\xi$$

But from (7), $\mu_1 = N_1 \mu / N$, Thus

$$1/\mu \int_0^{\infty} G(\xi) e^{-P\xi} d\xi = N_1 / N \quad (13)$$

IV. The Mean Time to Promotion

Let μ_L be the average length of time spent in Grade I, and let μ_p be the average length of time spent in Grade I by those who are eventually promoted to Grade II. As before, let μ_1 be the average sojourn time in Grade I. (This can be terminated by promotion on leaving). Let us consider the problem of choosing N_1 so that μ_p has some predetermined value. The two promotion rules are considered separately.

V. Promotion by Seniority

In this case μ_p is equivalent to the average value of ξ_1 introduced earlier, and hence equations (11) and (12) hold.

Thus we have

$$\int_{\mu_p}^{\infty} a(\xi) d\xi = N_2/N \quad (14)$$

$$1/\mu \int_{\mu_p}^{\infty} G(\xi) d\xi = N_2/N \quad (15)$$

Also $N_1 P = G(\mu_p) \frac{N}{\mu}$ Equation (15) gives $R = N_1/N_2$ as a function of μ_p and hence knowing the size N_2 of the organization, we can determine N_1 for any specified value of μ_p .

III. Results

I. Numerical Study

Case (i)

CLS is taken to be a exponential distribution,

$$f(\xi) = \lambda e^{-\lambda\xi}; \quad \lambda > 0, \xi > 0$$

Then $G(\xi) = e^{-\lambda\xi}$ and $\mu = 1/\lambda$

From equation (15) implies

$$\frac{1}{\lambda(1+R)} = \int_{\mu_p}^{\infty} e^{-\lambda\xi} dt, \text{ then}$$

$$R = (e^{-\lambda\mu_p} - 1) \quad (16)$$

The values of R for various values of μ_p and λ are given in Table 1.

Case (ii)

CLS is taken to be mixed exponential distribution

$$f(\xi) = x\lambda e^{-\lambda_1\xi} + (1-x)\lambda_2 e^{-\lambda_2\xi}; \quad 0 < x < 1; \lambda_1, \lambda_2 > 0; \xi \geq 0$$

Then

$$G(t) = x e^{-\lambda_1 t} + (1-x) e^{-\lambda_2 t}$$

$$\mu = \frac{x}{\lambda_1} + (1-x)/\lambda_2 \quad (17)$$

From equation (15) implies

$$R = \frac{\frac{x(1-e^{-\lambda_1\mu_p})}{\lambda_1} + \frac{(1-x)(1-e^{-\lambda_2\mu_p})}{\lambda_2}}{\frac{x(e^{-\lambda_1\mu_p})}{\lambda_1} + \frac{(1-x)(1-e^{-\lambda_2\mu_p})}{\lambda_2}} \quad (18)$$

The values of R corresponding to various values of μ_p for the mixed Exponential CLS distribution with parameters $x = 0.4, \lambda_1 = 0.2$ and $\lambda_2 = 2.0$ so that $\mu = 2.3$ are given in table .2.

Case (iii)

Now the CLS taken to be a Pearsonian Type XI distribution (Silcock's form) and the behavior of R with respect to μ_p, γ and c is studied.

$$f(t) = \frac{\gamma}{c} \left(1 + \frac{\xi}{c}\right)^{-(\gamma-1)} \quad (19)$$

$$G(t) = (1 + \frac{\xi}{c})^{-\gamma}; \quad \xi \geq 0, \gamma > 1, c > 0 \quad (20)$$

$$\mu = \frac{c}{\gamma - 1}$$

From equation (15) implies

$$\frac{N_2}{N} = \frac{1}{\mu} \xi \quad (21)$$

$$\frac{N_2}{N} = \frac{1}{\mu} \int_{\mu_0}^{\infty} (1 + \xi/c)^{-\gamma} d\xi \quad (22)$$

$$R = \left[\left(\frac{c}{\mu_p + c} \right)^{1-\gamma} - 1 \right] \quad (23)$$

Then, the value of R for the various values of μ_p, γ and $c = 3$ is given in Table 3.

Table 1: CLS as Exponential Distribution

μ_p λ	0.5	1.0	1.5	2.0	2.5	3.0
0.2	0.111	0.212	0.315	0.492	0.651	0.821
0.4	0.221	0.491	0.822	1.232	1.722	2.322
0.8	0.351	0.823	1.453	2.331	3.483	5.053
1.0	0.451	1.231	2.322	3.952	6.381	10.023
1.2	0.652	1.723	3.482	6.391	11.182	19.092
1.4	0.823	2.322	5.051	10.022	19.081	35.601
1.6	1.011	3.064	7.172	15.442	32.122	65.693
1.8	1.232	3.951	10.021	23.533	53.592	120.511
1.9	1.461	5.052	13.882	35.601	89.021	220.412
2.0	1.721	6.392	19.092	53.602	147.412	402.432

Table 2: CLS as Mixed Exponential Distribution

μ_p	0.5	1.0	1.5	2.0	2.5	3.0
R	0.1196	0.3919	0.6109	0.8128	1.0519	1.3016

Table 3: CLS as Pearsonian Type XI Distribution

γ μ	1.5	2.0	2.5	3.0	3.5
2	0.29109	0.66667	1.15117	1.77718	2.58611
4	0.52175	1.33333	2.56412	4.44444	7.31615
6	0.71321	2.00010	4.19612	7.99999	14.58815
8	0.91148	2.66677	6.02111	12.44444	24.74412
10	1.08116	3.33333	8.02015	17.77718	38.08911

It is discovered that in the case of the mixed exponential distribution, it is preferable to have fewer recruits at the training grade in order to get promoted sooner. Since a tiny percentage of trainees would survive to be promoted in the case of CLS as a Pearsonian Type XI distribution, a high training grade would be necessary. The analysis of data pertaining to completed length of service will be helpful in large organizations with different hierarchical grades to predict the size or number of individuals to be trained at each grade, ensuring that the organization's regular operations are not disrupted by the waste of personnel in the various grades. This is because it is feasible to anticipate the amount of leavers for varying CLS lengths.

IV. Discussion

Table 1 shows values of R calculated for various values of λ and μ_p . The values of R required becomes large when $\lambda\mu_p > 1$, and then increases rapidly as $(\lambda\mu_p)$ increases. This is what one would intuitively expect, since $\lambda\mu_p > 1$ implies that $\mu_p > 1/\lambda$, i.e. the average time to promotion is greater than the overall mean length of completed service. Thus only a very small proportion of trainees would survive to promotion and hence a large training grade would be required. Table 2 shows the values of R corresponding to various values of μ_p and the particular distribution is fitted to manpower data, it is normally found that x takes a value near $1/2$ and that λ_2 is about ten times the value of λ_1 , and hence producing a fairly skew distribution. Then the mixed exponential CLS distribution with parameters $\lambda_1 = 0.2$, $\lambda_2 = 2.0$, $x = 0.4$, so that $\mu = 2.3$.

Table 3 shows values of R calculated for various values of γ and μ_p . The values of R becomes large when γ, μ_p increase. The average time to promotion is greater than the overall mean length of completed service. Thus only a very small proportion of trainees would survive to promotion and hence a large training grade would be required. The mean duration of finished services is less than the average time to promotion. Because of the extremely low rate of trainee survival and promotion, a high training grade would be necessary.

References

- [1] Bartholomew, D.J. (1963). A Multi-Stage Renewal Process, *Journal of the Royal Statistical Society, Series B*, vol. 25, pp. 150 -168.
- [2] Bartholomew, D.J. (1969). Renewal Theory Models for Manpower Systems, *In.N.A.B. Wilson* (1969), pp.120 - 128.
- [3] Bartholomew, D.J. (1971). The Statistical Approach to Manpower Planning, *The Statistician*, Vol. 20, pp. 3-26.

- [4] Bartholomew, D.J. (1973). *Stochastic Model for Social Process*, 2nd ed., *John Wiley and Sons*, New York.
- [5] Bartholomew, D.J. A.F. Forbes and S.I. McClean (1991). *Statistical Techniques for Manpower Planning*, *John Wiley and Sons*, New York.
- [6] Bartholomew, D.J. and A.F. Forbes (1979). *Statistical Techniques for Manpower Planning*, *John Wiley*, Chichester.
- [7] Butlers, A.D. (1971). An Analysis of Flows in a Manpower System. *The Statistician*, Vol. 20, No.1, 69-84.
- [8] Grinold, R.C., and K.T. Marshall (1977). *Manpower Planning Models*, North-Holland, New York.
- [9] Walker.J (1968). Trends in Manpower Management Research. *Business Horizons*, pp.36-46.
- [10] Bagdonavicius, V., & Nikulin, M. S. (2001). Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, 7(1): 85-103.
- [11] Chien-Yu Peng (2015): Inverse Gaussian Processes with Random Effects and Explanatory Variables for Degradation Data, *Technometrics*, 57(1): 100-111.
- [12] Esary, J.D., A.W. Marshall and F. Prochan (1973). Shock models and wear process. *Annals of probability*, Vol. I, No. 4, pp. 627-650.
- [13] Grinold, R.C., and K.T. Marshall (1977). *Manpower Planning Models*, North-Holland, New York.
- [14] Jerry Lawless and Martin Crowder. (2004). Covariates and Random Effects in a Gamma Process Model with Application to Degradation and Failure, *Lifetime Data Analysis*, 10:213–227.
- [15] Lawless. J and Martin. C. (2004). Covariates and random effects in a Gamma process model with application to degradation and failure. *Lifetime Data Analysis*, 10(3):213-219.
- [16] Meeker, W.Q. and Hamada, M. (1995). Statistical tools for the rapid development and evaluation of high-reliability products. *IEEE Transactions on Reliability*, 44(2):187– 198.
- [17] Tweedie M.C.K., (1957). Statistical properties of inverse Gaussian distributions I, II, *Annals of Mathematical Statistics*, 28: 362-377.
- [18] Robinson,D. (1974). Two stage replacement strategies and their application to manpower planning, *Management Science*, vol.21, pp.199-208.

PERFORMANCE ANALYSIS OF $M^{[X]}/G^B/1$ FEEDBACK RETRIAL QUEUE WITH VARIABLE SERVER MODEL

N. MICHEAL MATHAVAVISAKAN¹ AND K. INDHIRA^{1,*}

¹ Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore, 632014, Tamil Nadu, India
michealmathava.visakan2020@vitstudent.ac.in, kindhira@vit.ac.in

Abstract

In this article, a working vacation policy-based on bulk arrival feedback retrial queueing system with variable server capacity has been analyzed. The server can serve a minimum of one customer and a maximum of B customers in a batch in accordance with the variable server capacity bulk service rule. As soon as the orbit becomes empty at the time of service completion, the server goes for a working vacation. The server works at a lower speed during a working vacation period. In addition, the steady state probability generating function for system size and orbit size is generated by incorporating the supplementary variables technique (SVT). Further, the conditional decomposition law is shown for this retrial queueing system. Moreover, system performance metrics, and significant special instances are discussed. Finally, the effects of various parameters on the system performance are analyzed numerically.

Keywords: Retrial queue, variable server capacity, working vacation, supplementary variable technique.

1. INTRODUCTION

The study of vacation queues (VQs) and retrial queues (RQs) in queueing theory has been going on for a while. When a consumer arrives and discovers the server is busy, they are directed to depart the service area and join a retry line called "orbit." In a RQ system, this is referred to as a RQ with repeated tries. The orbiting consumers may attempt their service request again when some time has passed. Additionally, the consumers in the orbit are allowed to request the same service repeatedly without affecting the other consumers. Modified models can be explored in RQs from Artalejo and Gomez Corral [1] and in VQs from Ke et al. [10]. The application of these queues in computer and communication systems is distinct.

In a VQ system, the server serves consumers at a slower rate speed during the working vacation (WV) period but fully discontinues service during the regular vacation period. Major uses for this queueing system (QS) include delivering network service, online service, file transfer service, and mail service, among others. Gautam Choudhury [7] probed a bulk arrival queue with a vacation period under single vacation strategy. Shan Gao [20] discussed a batch arrival queue with delayed single WV. A concise explanation of WV queueing systems was given in recent years by Chandrasekaran et al.[4]. Rajadurai [17] developed a unique form of the RQ model, which contained WV and breaks. An exponentially distributed multiple WV and a bulk arrival RQ with feedback were both studied by Pazhani Bala Murugan and Vijaykrishnaraj [16]. Madhu Jain and Anshul Kumar [12] analysed the $M^{[X]}/G/1$ model with WV, balking, unreliable server.

One of the most important aspects of communication systems is the feedback phenomena. If the service provided to a consumer is unsatisfactory, it may be tried again until it is. Maragathasundari and Balamurugan [14] studied the $M^{[X]}/G/1$ feedback queue with two stages of

repair times, general delay times. Madhu Jain and Anshul Kumar [13] discovered the bulk arrival general service RQ subject to balking, feedback and vacation interruption under multiple WV policy.

In real-world circumstances like elevators, freight loading and unloading, big wheels, chemical industrial processes, communication networks, tourism, etc., bulk QS are frequently used. Bailey [2] invented batch service queueing techniques. Batch service queueing system have been researched by Sasikala and Indhira [18]. Jaiswal [9] is the source of the original research on the variable server capacity bulk service rule. Banerjee et al. [3] have thought about queueing models with a variable server capacity and bulk service rule. Recently, Sasikala et al. [19] discovered the bulk RQ system with Bernoulli vacation schedule and variable server capacity. In the WV queue for bulk arrival feedback, no work is being done. Therefore, we concentrated on batch arrival using a batch service feedback RQ system with variable server capacity while working on vacation.

The purpose of this research is to ascertain the queue length and orbit length distributions, which will be used to ascertain the system's other behaviour metrics. The structure of our article is as follows: We offer a detailed description of the queueing model in section 2 once the prerequisites have been met. In section 3, it has been clearly determined how the system behaves in steady-state (SS) conditions and what the probability generating function (PGF) of the queue size is at a random epoch. There are various important system behaviour indicators in section 4. Stochastic decomposition and some important specific occurrences are mentioned in section 5. There are both numerical and pictorial findings in section 6. Finally, the paper's key ideas are summarized in section 7.

2. DESCRIPTION OF THE MODEL AND ITS IMPLEMENTATION IN REAL WORLD

Under WVs policy, we provide a $M^{[X]}/G^B/1$ feedback RQ. The precise justification of our model is as follows:

The arrival process: According to the Poisson process, consumers are arriving for service at the rate α . where \mathcal{F} is the batch size random variable with probability mass function $P\{\mathcal{F} = n\} = f_n$, $n = 1, 2, 3, \dots$ probability generating function (PGF) $\mathcal{F}(\zeta) = \sum_{n=0}^{\infty} \zeta^n f_n$ and mean batch size $\mathcal{E}(\mathcal{I})$.

The retrial process: If arriving consumers not getting service immediately due to some reasons, they gather together, and this is known as an "orbit". After a certain random amount of time, customers will retry for service. Inter retrial times have a random dist., $\mathcal{H}(\omega)$ with corresponding "Laplace-Stieltjes Transform" (LST) $\mathcal{H}^*(\delta)$

The regular service process: Based on the variable server capacity bulk service regulation, the server will transmit the consumers. According to the variable server capacity batch service rule, the server will serve either fixed size, like a " \mathcal{B} ", or all of the consumers from the orbit, depending on which is lower. If there are more than or equal to " \mathcal{B} " consumers in the orbit after the group of consumers has been transmitted, the server will proceed to transmit " \mathcal{B} " packets in a batch. If less than " \mathcal{B} " packets remain in the orbit after transmission, the server will take whole consumers to send in a batch. After the service has begun, late entrants are not permitted to participate in the ongoing service, even if the batch size is smaller than " \mathcal{B} ". The service period represents a general dist. and it is marked by the arbitrary variable \mathcal{D} with dist. function $\mathcal{D}(\omega)$ having LST $\mathcal{D}^*(\delta)$.

Feedback rule: Unsatisfied consumers have the option to re-enter the orbit as feedback consumers once their normal service is complete in order to maybe receive another service with prob., β ($0 \leq \beta \leq 1$) will exit the system with prob., $\bar{\beta} = (1 - \beta)$.

The working vacation policy: When the orbit is free, the server periodically takes a WV. The vacation period takes an exponential dist. with variable ω . If a consumer enters during a vacation time, the server keeps running at a reduced rate. During the WV time, tasks are carried out at a slower pace. If any consumers are in the orbit at a slower service completion moment during the vacation period, the server will end the vacation and return to the normal busy time, interrupting the vacation. If not, the vacation, keeps going. When the vacation gets over, the server restores

normal operations if there are still customers in the orbit. During the WV period, the service period is assessed by a random variable \mathcal{D}_v with dist. function $\mathcal{D}_v(\omega)$ and LST $\mathcal{D}_v^*(\delta)$.

The system's stochastic processes are considered to be independent of one another.

2.1. Practical application of the model

As an example of the proposed paradigm in action, consider that telephone consultations are a significant component of medical care delivery systems, which illustrates how the suggested paradigm is used in real-life scenarios. We take into consideration a telephone consultation system with a primary server (the chief physician) and a physician assistant (the working breakdown server). Patients attempt to schedule appointments for treatment over the phone, but there is a restriction on the no.of appointments (variable server capacity) that may be made each day for treatment. The physician assistant provides service only when the primary physician is unavailable, and it is noted that the assistant's service delivery is frequently slower than that of the primary physician. Furthermore, when each patient's regular service is completed (feedback), the dissatisfied patient may re-enter the orbit.

In order to schedule the appointments, a phone operator will be available, who usually manages the patients and doctors. If the phone line is busy when a patient calls, he must wait and try again later (retrial); if not, he will be given an appointment right away to see the head physician or the physician assistant for treatment. However, the phone operator will call (or search for) the FCFS customers in orbit as soon as the service is finished. It is predicted that the search time will be evenly divided, which is in line with the general retry time policy.

3. OVERVIEW OF STEADY STATE PROBABILITIES

This division first develops the steady-state (SS) equations for the RQ system by considering the elapsed retrial period, the elapsed service time and the elapsed lower-speed service times as supplementary variable (SV). The PGF of the no. of consumers in the orbit and system, as well as the orbit length generating functions for numerous server states, are computed.

3.1. Probabilities and Notations

It is assumed in SS that $\mathcal{H}(0) = 0$, $\mathcal{H}(\infty) = 1$, $\mathcal{D}(0) = 0$, $\mathcal{D}(\infty) = 1$ and $\mathcal{D}_v(0) = 0$, $\mathcal{D}_v(\infty) = 1$ are cont., at $\omega = 0$. So that the func. $\chi(\omega)$, $\eta(\omega)$, $\eta_v(\omega)$, are the hazard rates (HR) for retrial, service and slower pace service respectively.

Further, the subsequent notations and probabilities were defined:

$\chi(\omega)$	-	HR for retrial (i.e.,) $\chi(\omega)d\omega = \frac{d\mathcal{H}(\omega)}{1-\mathcal{H}(\omega)}$
$\eta(\omega)$	-	HR for service (i.e.,) $\eta(\omega)d\omega = \frac{d\mathcal{D}(\omega)}{1-\mathcal{D}(\omega)}$
$\eta_v(\omega)$	-	HR for slower pace service (i.e.,) $\eta_v(\omega)d\omega = \frac{d\mathcal{D}_v(\omega)}{1-\mathcal{D}_v(\omega)}$
$\mathcal{Y}(\check{\tau})$	-	no.of consumers in the orbit
$\mathcal{H}^0(\check{\tau})$	-	elapsed retrial time
$\mathcal{D}^0(\check{\tau})$	-	elapsed service time
$\mathcal{D}_v^0(\check{\tau})$	-	elapsed WV times
$\Upsilon_n(\omega, \check{\tau})$	-	Prob. that at time $\check{\tau}$ there are precisely n consumers in the orbit with the consumer going through a retrial having served their whole service period is ω .
$\Omega_n(\omega, \check{\tau})$	-	Prob. that at time $\check{\tau}$ there are precisely n consumers in the orbit with the consumer going through normal service having served their whole service period is ω .

$\Psi_{v,n}(\omega, \check{\tau})$ - Prob. that at time $\check{\tau}$ there are precisely n consumers in the orbit with the consumer going through slower pace service having served their whole service period is ω .

Apart from it, let $\mathcal{H}^0(\check{\tau}), \mathcal{D}^0(\check{\tau})$ and $\mathcal{D}_v^0(\check{\tau})$ be the elapsed retrial period, the elapsed period of normal service and the elapsed slower-rate service period respectively at time $\check{\tau}$. Additionally, generate the random variable,

$$\Theta(\check{\tau}) = \begin{cases} 0, & \text{if the server is available and in WV time} \\ 1, & \text{if the server is available and in normal service time} \\ 2, & \text{if the server is unavailable and in normal service at time } \check{\tau} \\ 3, & \text{if the server is unavailable and in lower speed rate at time } \check{\tau} \end{cases}$$

Here, we highlight the usage of bivariate Markov process to describe the system's state at time $\{\Theta(\check{\tau}), \mathcal{Y}(\check{\tau}); \check{\tau} \geq 0\}$, where $\Theta(\check{\tau})$ signifies the server state (0, 1, 2, 3) depending on whether the server is free or busy on both normal service and WV periods. $\mathcal{Y}(\check{\tau})$ denotes the no. of consumers in the orbit. If $\Theta(\check{\tau}) = 1$ and $\mathcal{Y}(\check{\tau}) > 0$, then $\mathcal{H}^0(\check{\tau})$ is equivalent to the elapsed retrial time. If $\Theta(\check{\tau}) = 2$ and $\mathcal{Y}(\check{\tau}) \geq 0$, then $\mathcal{D}^0(\check{\tau})$ is equivalent to the elapsed time of the consumer served in normal busy period. If $\Theta(\check{\tau}) = 3$ and $\mathcal{Y}(\check{\tau}) \geq 0$, then $\mathcal{D}_v^0(\check{\tau})$ is equivalent to the elapsed time of the consumer being served in lower rate service period.

3.2. Ergodicity analysis of the model

We examine the embedded Markov chain's ergodicity during the departure and vacation epochs. Let $\{\check{\tau}_n; n = 1, 2, \dots\}$ be the series of epochs where either a service period completion or a shorter service period happens. $G_n = \{\Theta(\check{\tau}_n+), \mathcal{Y}(\check{\tau}_n+)\}$ sequence of random vectors. The Markov chain formed by embedded in the RQ system. It follows from Appendix A that is the embedded Markov chain $\{\mathcal{G}_m; m \in M\}$ is ergodic iff $\Lambda < \mathcal{B}$ for our system will be stable.

For the method $\{\mathcal{Y}(\check{\tau}), \check{\tau} \geq 0\}$, we specify the probabilities $\mathcal{Q}_0(\check{\tau}) = \mathcal{P}\{\Lambda(\check{\tau}) = 0, \mathcal{Y}(\check{\tau}) = 0\}$ and the prob. densities are

$$Y_n(\omega, \check{\tau})d\omega = \mathcal{P}\{\Lambda(\check{\tau}) = 1, \mathcal{Y}(\check{\tau}) = n, \omega \leq \mathcal{H}^0(\check{\tau}) < \omega + d\omega\},$$

for $\check{\tau} \geq 0, \omega \geq 0$ and $n \geq 1$.

$$\Omega_n(\omega, \check{\tau})d\omega = \mathcal{P}\{\Lambda(\check{\tau}) = 2, \mathcal{Y}(\check{\tau}) = n, \omega \leq \mathcal{D}^0(\check{\tau}) < \omega + d\omega\},$$

for $\check{\tau} \geq 0, \omega \geq 0$ and $n \geq 0$.

$$\Psi_{v,n}(\omega, \check{\tau})d\omega = \mathcal{P}\{\Lambda(\check{\tau}) = 3, \mathcal{Y}(\check{\tau}) = n, \omega \leq \mathcal{D}_v^0(\check{\tau}) < \omega + d\omega\},$$

for $\check{\tau} \geq 0, \omega \geq 0$ and $n \geq 0$.

We presume that the stability requirement is satisfied in the sequel, so we may assign $\mathcal{Q}_0 = \lim_{\check{\tau} \rightarrow \infty} \mathcal{Q}_0(\check{\tau})$ and limiting densities are

$$Y_n(\omega) = \lim_{\check{\tau} \rightarrow \infty} Y_n(\omega, \check{\tau}); \Omega_n(\omega) = \lim_{\check{\tau} \rightarrow \infty} \Omega_n(\omega, \check{\tau});$$

$$\Psi_{v,n}(\omega) = \lim_{\check{\tau} \rightarrow \infty} \Psi_{v,n}(\omega, \check{\tau});$$

Using the supplementary variable method, we build the following system of equations.

$$\begin{aligned} \alpha \mathcal{Q}_0 &= \bar{\beta} \int_0^\infty \Omega_0(\omega) \eta(\omega) d\omega + \bar{\beta} \int_0^\infty \Psi_{v,0}(\omega) \eta_v(\omega) d\omega \\ &+ \alpha \int_0^\infty \Omega_n(\omega) d\omega, n \geq 0 \end{aligned} \tag{1}$$

$$\frac{d}{d\omega} Y_n(\omega) + (\alpha + \chi(\omega)) Y_n(\omega) = 0, n \geq 1 \tag{2}$$

$$\frac{d}{d\omega} \Omega_0(\omega) + (\alpha + \eta(\omega)) \Omega_0(\omega) = 0, n = 0 \tag{3}$$

$$\frac{d}{d\omega} \Omega_n(\omega) + (\alpha + \eta(\omega)) \Omega_n(\omega) = \alpha \sum_{k=1}^n \Omega_{n-k} f_k(\omega), n \geq 1 \tag{4}$$

$$\frac{d}{d\omega} \Psi_{v,0}(\omega) + (\alpha + \omega + \eta_v(\omega)) \Psi_{v,0}(\omega) = 0, n = 0 \tag{5}$$

$$\frac{d}{d\omega} \Psi_{v,n}(\omega) + (\alpha + \omega + \eta_v(\omega)) \Psi_{v,n}(\omega) = \alpha \sum_{k=1}^n \Psi_{v,n-k} f_k(\omega), n \geq 0 \tag{6}$$

At $\omega = 0$ the SS boundary criteria are as follows:

$$Y_n(0) = \beta \int_0^\infty \Omega_n(\omega) \eta(\omega) d\omega + \bar{\beta} \int_0^\infty \Omega_{n-1}(\omega) \eta(\omega) d\omega \tag{7}$$

$$+ \beta \int_0^\infty \Psi_{v,n}(\omega) \eta_v(\omega) d\omega + \bar{\beta} \int_0^\infty \Psi_{v,n-1}(\omega) \eta_v(\omega) d\omega, n \geq 1$$

$$\Omega_n(0) = \int_0^\infty Y_{n+B}(\omega) \chi(\omega) d\omega + \alpha \int_0^\infty \sum_{k=1}^\infty f_k Y_{n-k+B}(\omega) d\omega \tag{8}$$

$$+ \omega \int_0^\infty \Psi_{v,n}(\omega) d\omega, n \geq 1$$

$$\Omega_0(0) = \int_0^\infty \sum_{n=1}^B Y_n(\omega) \chi(\omega) d\omega + \alpha \sum_{k=1}^B f_k Y_0 + \omega \int_0^\infty \Psi_{v,0}(\omega) d\omega, n = 0 \tag{9}$$

$$\Psi_{w,n}(0) = \begin{cases} \alpha Q_0, & n = 0 \\ 0, & n \geq 1 \end{cases} \tag{10}$$

The normalizing criteria is

$$Q_0 + \sum_{n=1}^\infty \int_0^\infty Y_n(\omega) d\omega + \sum_{n=0}^\infty \left(\int_0^\infty \Omega_n(\omega) d\omega + \int_0^\infty \Psi_{v,n}(\omega) d\omega \right) = 1 \tag{11}$$

3.3. The steady state solution

The GFs for $|\check{\zeta}| < 1$ in order to solve the aforementioned equations, are expressed in the form.

$$Y(\omega, \check{\zeta}) = \sum_{n=1}^\infty Y_n(\omega) \check{\zeta}^n; Y(0, \check{\zeta}) = \sum_{n=1}^\infty Y_n(0) \check{\zeta}^n;$$

$$\Omega(\omega, \check{\zeta}) = \sum_{n=0}^\infty \Omega_n(\omega) \check{\zeta}^n; \Omega(0, \check{\zeta}) = \sum_{n=0}^\infty \Omega_n(0) \check{\zeta}^n;$$

$$\Psi_v(\omega, \check{\zeta}) = \sum_{n=0}^\infty \Psi_{v,n}(\omega) \check{\zeta}^n; \Psi_v(0, \check{\zeta}) = \sum_{n=0}^\infty \Psi_{v,n}(0) \check{\zeta}^n;$$

Now multiply the SS equation and SS boundary criteria from (2) to (10) by $\check{\zeta}^n$ and summing over n , ($n = 0, 1, 2, \dots$)

$$\frac{\partial}{\partial \omega} Y(\omega, \check{\zeta}) + (\alpha + \chi(\omega)) Y(\omega, \check{\zeta}) = 0 \tag{12}$$

$$\frac{\partial}{\partial \omega} \Omega(\omega, \check{\zeta}) + (\alpha(1 - \mathcal{F}(\check{\zeta})) + \eta(\omega)) \Omega(\omega, \check{\zeta}) = 0 \tag{13}$$

$$\frac{\partial}{\partial \omega} \Psi_v(\omega, \check{\zeta}) + (\alpha(1 - \mathcal{F}(\check{\zeta})) + \omega + \eta_v(\omega)) \Psi_v(\omega, \check{\zeta}) = 0 \tag{14}$$

$$Y(0, \check{\zeta}) = (\beta + \check{\beta}\check{\zeta}) \int_0^\infty \Omega(\omega, \check{\zeta})\eta(\omega)d\omega + (\beta + \check{\beta}\check{\zeta}) \int_0^\infty \Psi_v(\omega, \check{\zeta})\eta_v(\omega)d\omega - \alpha Q_0 \quad (15)$$

$$\Omega(0, \check{\zeta}) = \frac{1}{\check{\zeta}^B} \int_0^\infty Y(\omega, \check{\zeta})\chi(\omega)d\omega + \frac{\alpha\mathcal{F}(z)}{\check{\zeta}^B} \int_0^\infty Y(\omega, \check{\zeta})d\omega + \omega \int_0^\infty \Psi_v(\omega, \check{\zeta})d\omega \quad (16)$$

$$\Psi_v(0, \check{\zeta}) = \alpha Q_0 \quad (17)$$

Solving the partial differential eqns. (12) to (14), we obtain

$$Y(\omega, \check{\zeta}) = Y(0, \check{\zeta})[1 - \mathcal{H}(\omega)]e^{-\alpha\omega} \quad (18)$$

$$\Omega(\omega, \check{\zeta}) = \Omega(0, \check{\zeta})[1 - \mathcal{D}(\omega)]e^{-\mathcal{S}(\check{\zeta})\omega} \quad (19)$$

$$\Psi_v(\omega, \check{\zeta}) = \Psi_v(0, \check{\zeta})[1 - \mathcal{A}_\omega(\omega)]e^{-\mathcal{S}_v(\check{\zeta})\omega} \quad (20)$$

where $\mathcal{S}(\check{\zeta}) = \alpha(1 - \mathcal{F}(\check{\zeta}))$, and $\mathcal{S}_v(\check{\zeta}) = \omega + \alpha(1 - \mathcal{F}(\check{\zeta}))$

Inserting the eqns. (17) to (20) in (8) after some computation, we eventually arrive to,

$$\Omega(0, \check{\zeta}) = \frac{Y(0, \check{\zeta})}{\check{\zeta}^B} \{ \mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta})[1 - \mathcal{H}^*(\alpha)] \} + \alpha Q_0 \mathcal{V}(\check{\zeta}) \quad (21)$$

where $\mathcal{V}(\check{\zeta}) = \frac{\omega}{\omega + \alpha(1 - \mathcal{F}(\check{\zeta}))} (1 - \mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})))$,

$$Y(0, \check{\zeta}) = (\beta + \check{\beta}\check{\zeta})\Omega(0, \check{\zeta})\mathcal{D}^*(\mathcal{S}(\check{\zeta})) + (\beta + \check{\beta}\check{\zeta})\Psi_v(0, \check{\zeta})\mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})) - \alpha Q_0 \quad (22)$$

Combining (10) and (21) in (22), we get

$$\begin{aligned} \Omega(0, \check{\zeta}) \{ \check{\zeta}^B - (\beta + \check{\beta}\check{\zeta})[\mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta})(1 - \mathcal{H}^*(\alpha))]\mathcal{D}^*(\mathcal{S}(\check{\zeta})) \} \\ = \alpha Q_0 \{ \check{\zeta}^B \mathcal{V}(\check{\zeta}) + [(\beta + \check{\beta}\check{\zeta})\mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})) - 1][\mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta})(1 - \mathcal{H}^*(\alpha))]\} \end{aligned} \quad (23)$$

In the following theorem, we are willing to exploring the marginal orbit size distributions caused by the server's system state.

Theorem 1. Under the stability requirement, $\Lambda < B$ provides the stationary dist., of the no. of customers in the orbit when the server is available, busy, reduced rate service, and the prob., that the server is available given by,

$$Y(\check{\zeta}) = \frac{Ne(\check{\zeta})}{De(\check{\zeta})} \quad (24)$$

$$Ne(\check{\zeta}) = \check{\zeta}^B Q_0 (1 - \mathcal{H}^*(\alpha)) \{ (\beta + \check{\beta}\check{\zeta})[\mathcal{D}^*(\mathcal{S}(\check{\zeta}))\mathcal{V}(\check{\zeta}) + \mathcal{D}_v^*(\mathcal{S}_v(z))] - 1 \}$$

$$De(\check{\zeta}) = \check{\zeta}^B - (\beta + \check{\beta}\check{\zeta}) \{ \mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta})[1 - \mathcal{H}^*(\alpha)] \} \mathcal{D}^*(\mathcal{S}(\check{\zeta}))$$

$$\Omega(\check{\zeta}) = \frac{\alpha Q_0 (1 - \mathcal{D}^*(\mathcal{S}(\check{\zeta})))}{\mathcal{S}(\check{\zeta}) De(\check{\zeta})} \{ \check{\zeta}^B \mathcal{V}(\check{\zeta}) + [(\beta + \check{\beta}\check{\zeta})\mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})) - 1][\mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta})[1 - \mathcal{H}^*(\alpha)]] \} \quad (25)$$

$$\Psi_v(\check{\zeta}) = \frac{\alpha Q_0}{\omega} \mathcal{V}(\check{\zeta}) \quad (26)$$

where

$$Q_0 = \frac{B - \{ \check{\beta} - \alpha \mathcal{E}(\mathcal{I}) \mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) \}}{\alpha \mathcal{E}(\mathcal{I}) \mathcal{E}(\mathcal{D}) \{ \frac{2\alpha}{\omega} (1 - \mathcal{D}_v^*(\omega)) - 2\mathcal{D}_v^*(\omega) \mathcal{H}^*(\alpha) + \mathcal{D}_v^*(\omega) + \mathcal{H}^*(\alpha) + 1 \} - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))[1 + \alpha \mathcal{E}(\mathcal{D}_v)] - \alpha \mathcal{E}(\mathcal{D})[1 + B(1 - \mathcal{D}_v^*(\omega))] + B(1 + \frac{\alpha}{\omega}(1 - \mathcal{D}_v^*(\omega)))} \quad (27)$$

Proof. Taking the eqns. (18)-(20) and integrate with respect to ω and compute the PG $Y(\check{\zeta}) = \int_0^\infty Y(\omega, \check{\zeta})d\omega$, $\Omega(\check{\zeta}) = \int_0^\infty \Omega(\omega, \check{\zeta})d\omega$, $\Psi_w(\check{\zeta}) = \int_0^\infty \Psi_w(\omega, \check{\zeta})d\omega$. We calculate the prob. that the server is empty using the normalization condition (Q_0) by establishing functions as, when there is no consumer in the orbit $\check{\zeta} = 1$ in (24)-(26) and whenever the condition of L'Hospital is needed, we get $Q_0 + Y(1) + \Omega(1) + \Psi_w(1) = 1$. ■

Theorem 2. Utilizing the PGF function, the no. of consumers in the system and the orbit size dist. at a stationary point of period are calculated under the stability constraint $\Lambda < \mathcal{B}$,

$$\mathcal{K}_s(\check{\zeta}) = \frac{Ne_s(\check{\zeta})}{De_s(\check{\zeta})} \tag{28}$$

$$\begin{aligned} Ne_s(\check{\zeta}) &= Q_0 \{ \mathcal{S}(\check{\zeta}) \{ \check{\zeta}^{\mathcal{B}} - (\beta + \bar{\beta}\check{\zeta}) \{ \mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta}) [1 - \mathcal{H}^*(\alpha)] \} \mathcal{D}^*(\mathcal{S}(\check{\zeta})) \} \\ &\quad [1 + \frac{\alpha}{\omega} \check{\zeta} \mathcal{V}(\check{\zeta})] \} + \check{\zeta}^{\mathcal{B}} \mathcal{S}(\check{\zeta}) (1 - \mathcal{H}^*(\alpha)) \{ (\beta + \bar{\beta}\check{\zeta}) [\mathcal{D}^*(\mathcal{S}(\check{\zeta})) \mathcal{V}(\check{\zeta}) + \mathcal{D}_v^*(\mathcal{S}_v(z))] - 1 \} \\ &\quad + \check{\zeta} \alpha (1 - \mathcal{D}^*(\alpha(1 - \mathcal{F}(\check{\zeta}))) \{ \check{\zeta}^{\mathcal{B}} \mathcal{V}(\check{\zeta}) + [(\beta + \bar{\beta}\check{\zeta}) \mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})) - 1] \\ &\quad [\mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta}) [1 - \mathcal{H}^*(\alpha)]] \} \} \\ De_s(\check{\zeta}) &= \mathcal{S}(\check{\zeta}) \{ \check{\zeta}^{\mathcal{B}} - (\beta + \bar{\beta}\check{\zeta}) \{ \mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta}) [1 - \mathcal{H}^*(\alpha)] \} \mathcal{D}^*(\mathcal{S}(\check{\zeta})) \} \end{aligned}$$

$$\mathcal{K}_0(\check{\zeta}) = \frac{Ne_0(\check{\zeta})}{De_s(\check{\zeta})} \tag{29}$$

$$\begin{aligned} Ne_0(\check{\zeta}) &= Q_0 \{ \mathcal{S}(\check{\zeta}) \{ \check{\zeta}^{\mathcal{B}} - (\beta + \bar{\beta}\check{\zeta}) \{ \mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta}) [1 - \mathcal{H}^*(\alpha)] \} \mathcal{D}^*(\mathcal{S}(\check{\zeta})) \} \\ &\quad [1 + \frac{\alpha}{\omega} \mathcal{V}(\check{\zeta})] \} + \check{\zeta}^{\mathcal{B}} \mathcal{S}(\check{\zeta}) (1 - \mathcal{H}^*(\alpha)) \{ (\beta + \bar{\beta}\check{\zeta}) [\mathcal{D}^*(\mathcal{S}(\check{\zeta})) \mathcal{V}(\check{\zeta}) + \mathcal{D}_v^*(\mathcal{S}_v(z))] - 1 \} \\ &\quad + \alpha (1 - \mathcal{D}^*(\alpha(1 - \mathcal{F}(\check{\zeta}))) \{ \check{\zeta}^{\mathcal{B}} \mathcal{V}(\check{\zeta}) + [(\beta + \bar{\beta}\check{\zeta}) \mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})) - 1] \\ &\quad [\mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta}) [1 - \mathcal{H}^*(\alpha)]] \} \} \end{aligned}$$

where Q_0 is denoted by eqn. (27). **Proof.** The PGF of the no. of consumer in the system ($\mathcal{K}_s(\check{\zeta})$) and in the orbit ($\mathcal{K}_0(\check{\zeta})$) is calculated by using $\mathcal{K}_s(\check{\zeta}) = Q_0 + Y(\check{\zeta}) + \Omega(\check{\zeta}) + \Psi_w(\check{\zeta})$. The eqns. (28) and (29) may be derived directly when the eqns. (24)-(27) are substituted in the earlier results. ■

4. SYSTEM PERFORMANCE MEASURES

In this section, different system states are used to derive a number of pertinent system probabilities, system efficiency metrics, and the model's mean busy time and mean busy cycle.

4.1. Probabilities of system states

Utilizing eqns, (24)-(26) we obtain the findings shown below, giving $\check{\zeta} \rightarrow 1$ and, if feasible, using L'Hospital's rule.

(i) Let Y be the SS prob. of the server is available during the retrial,

$$Y = Y(1) = Q_0 (1 - \mathcal{H}^*(\alpha)) \left\{ \frac{\bar{\beta} + \alpha \mathcal{E}(\mathcal{I}) [\mathcal{E}(\mathcal{D}) \mathcal{D}_v^*(\omega) + \frac{1}{\omega} (1 - \mathcal{D}_v^*(\omega)) - \mathcal{E}(\mathcal{D}_v)]}{\mathcal{B} - \{ \bar{\beta} - \alpha \mathcal{E}(\mathcal{I}) \mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I}) (1 - \mathcal{H}^*(\alpha)) \}} \right\} \tag{30}$$

(ii) Let Ω be the SS prob. that the server is full,

$$\Omega = \Omega(1) = \alpha \mathcal{E}(\mathcal{D}) Q_0 \left\{ \frac{\mathcal{E}(\mathcal{I}) (1 - \mathcal{D}_v^*(\omega)) [\mathcal{H}^*(\alpha) + \frac{\alpha}{\omega}] + (\bar{\beta} - \mathcal{B}) \mathcal{D}_v^*(\omega) + \mathcal{B} - 1}{\mathcal{B} - \{ \bar{\beta} - \alpha \mathcal{E}(\mathcal{I}) \mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I}) (1 - \mathcal{H}^*(\alpha)) \}} \right\} \tag{31}$$

(iii) Let Ψ_w be the SS prob. that the server is on WV,

$$\Psi_v = \Psi_v(1) = \frac{\alpha Q_0}{\omega} [1 - \mathcal{D}_v^*(\omega)] \quad (32)$$

(iv) Let Y_f be the SS prob. that the server is failure,

$$Y_f = \alpha \times \Omega(1) = \alpha^2 \mathcal{E}(\mathcal{D}) Q_0 \left\{ \frac{\mathcal{E}(\mathcal{I})(1 - \mathcal{D}_v^*(\omega))[\mathcal{H}^*(\alpha) + \frac{\alpha}{\omega}] + (\bar{\beta} - \mathcal{B})\mathcal{D}_v^*(\omega) + \mathcal{B} - 1}{\mathcal{B} - \{\bar{\beta} - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))\}} \right\} \quad (33)$$

4.2. Mean size of a system and orbit

When the system is in a steady state,

(i) With respect to ζ , (29) and providing $\zeta = 1$ yields the mean no. of consumers in the orbit (\mathcal{L}_q)

$$\mathcal{L}_q = \mathcal{K}'_0(1) = \lim_{\zeta \rightarrow 1} \frac{d}{d\zeta} \mathcal{K}_0(\zeta) = Q_0 \left[\frac{\mathcal{N}'''_q(1)\mathcal{D}''_q(1) - \mathcal{D}'''_q(1)\mathcal{N}''_q(1)}{3(\mathcal{D}''_q(1))^2} \right] \quad (34)$$

$$\begin{aligned} \mathcal{N}''_q(1) &= -2\alpha \mathcal{E}(\mathcal{I}) \left\{ \left[1 + \frac{\alpha}{\omega} (1 - \mathcal{D}_v^*(\omega)) \right] [\mathcal{B} - \bar{\beta} + \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))] \right. \\ &\quad \left. + (1 - \mathcal{H}^*(\alpha)) \left\{ \bar{\beta} + \alpha \mathcal{E}(\mathcal{I}) [\mathcal{E}(\mathcal{D})\mathcal{D}_v^*(\omega) + \frac{1}{\omega} (1 - \mathcal{D}_v^*(\omega)) - \mathcal{E}(\mathcal{D}_v)] \right\} \right. \\ &\quad \left. - \alpha \mathcal{E}(\mathcal{D}) \left\{ \mathcal{E}(\mathcal{I})(1 - \mathcal{D}_v^*(\omega)) \left[\mathcal{H}^*(\alpha) + \frac{\alpha}{\omega} \right] + (\bar{\beta} - \mathcal{B})\mathcal{D}_v^*(\omega) + \mathcal{B} - 1 \right\} \right\} \\ \mathcal{D}''_q(1) &= -2\alpha \mathcal{E}(\mathcal{I}) \left\{ \mathcal{B} + \bar{\beta} - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) + \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) \right\} \\ \mathcal{N}'''_q(1) &= -6\alpha \mathcal{E}(\mathcal{I}) [\mathcal{B} - \bar{\beta} + \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))] \left\{ \frac{\alpha}{\omega} \mathcal{E}(\mathcal{I})(1 + \omega \mathcal{E}(\mathcal{D}) \right. \\ &\quad \left. - \mathcal{D}_v^*(\omega)) \right\} + \mathcal{D}'''_q(1) \left[1 + \frac{\alpha}{\omega} (1 - \mathcal{D}_v^*(\omega)) \right] - 3\alpha \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) \\ &\quad \left\{ \bar{\beta} + \alpha \mathcal{E}(\mathcal{I})(1 + \mathcal{B}) \left[\frac{1}{\omega} (1 - \mathcal{D}_v^*(\omega)) + \mathcal{E}(\mathcal{D})\mathcal{D}_v^*(\omega) - \mathcal{E}(\mathcal{D}_v) \right] \right. \\ &\quad \left. + 2\bar{\beta} \left\{ \left[\frac{\alpha}{\omega} \mathcal{E}(\mathcal{I})(1 + \omega \mathcal{E}(\mathcal{D}_v) - \mathcal{D}_v^*(\omega)) \right] - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D})(1 - \mathcal{D}_v^*(\omega)) \right. \right. \\ &\quad \left. \left. - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}_v) \right\} - 2\alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) \left[\frac{\alpha}{\omega} \mathcal{E}(\mathcal{I})(1 + \omega \mathcal{E}(\mathcal{D}) - \mathcal{D}_v^*(\omega)) \right] + (1 - \mathcal{D}_v^*(\omega)) \right. \\ &\quad \left. \left[\alpha^2 \mathcal{E}(\mathcal{I})\mathcal{E}^2(\mathcal{D}) - \alpha \mathcal{E}(\mathcal{I}(\mathcal{I} - 1))\mathcal{E}(\mathcal{D}) \right] - \alpha \mathcal{E}(\mathcal{I}(\mathcal{I} - 1))\mathcal{E}(\mathcal{D}_v) - \alpha^2 \mathcal{E}(\mathcal{I})\mathcal{E}^2(\mathcal{D}_v) \right. \\ &\quad \left. + \mathcal{V}''(1) + 3\alpha \left\{ \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) \left\{ \mathcal{B}(\mathcal{B} - 1)(1 - \mathcal{D}_v^*(\omega)) + (\mathcal{B} + 1) \left[\frac{\alpha}{\omega} \mathcal{E}(\mathcal{I}) \right. \right. \right. \right. \\ &\quad \left. \left. \left. (1 + \omega \mathcal{E}(\mathcal{D}) - \mathcal{D}_v^*(\omega)) \right] + 2\alpha \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) [\bar{\beta} - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}_v)] \right. \right. \\ &\quad \left. \left. - 2\bar{\beta} \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}_v) + \alpha^2 \mathcal{E}(\mathcal{I})\mathcal{E}^2(\mathcal{D}_v) - \alpha \mathcal{E}(\mathcal{I}(\mathcal{I} - 1))\mathcal{E}(\mathcal{D}_v) + \mathcal{V}''(1) \right\} \right. \\ &\quad \left. \left. + \alpha \left[\mathcal{E}(\mathcal{I}(\mathcal{I} - 1))\mathcal{E}(\mathcal{D}) - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}^2(\mathcal{D}) \right] \left\{ \bar{\beta} + \alpha \mathcal{E}(\mathcal{I}) [\mathcal{E}(\mathcal{D})\mathcal{D}_v^*(\omega) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{\omega} (1 - \mathcal{D}_v^*(\omega)) - \mathcal{E}(\mathcal{D}_v) \right] \right\} \right\} \right\} \\ \mathcal{D}'''_q(1) &= -3\alpha \mathcal{E}(\mathcal{I}) \left\{ \mathcal{B}(\mathcal{B} - 1) - \mathcal{E}(\mathcal{I}(\mathcal{I} - 1))(1 - \mathcal{H}^*(\alpha)) + 2\{\alpha \bar{\beta} \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) \right. \\ &\quad \left. + \alpha \mathcal{E}(\mathcal{I}(\mathcal{I} - 1))\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) [\alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \bar{\beta}] \right\} \end{aligned}$$

where $\mathcal{V}''(1) = \frac{\alpha}{\omega} \mathcal{E}(\mathcal{I}(\mathcal{I} - 1)) [1 + \omega \mathcal{E}(\mathcal{D}_v) - \mathcal{D}_v^*(\omega)] + \frac{\mathcal{E}(\mathcal{I})}{\omega^3} \{ \omega^2 \mathcal{E}^2(\mathcal{D}_v) - 2\alpha \omega \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}_v) + \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}_v) \} + \alpha \mathcal{E}(\mathcal{I})(1 - \mathcal{D}_v^*(\omega))$

(ii) With regard to ζ , (28) and providing $\zeta = 1$ yields the mean no. of consumers in the system (\mathcal{L}_s)

$$\mathcal{L}_s = \mathcal{K}'_s(1) = \lim_{\zeta \rightarrow 1} \frac{d}{d\zeta} \mathcal{K}_s(\zeta) = Q_0 \left[\frac{\mathcal{N}'''_s(1)\mathcal{D}''_q(1) - \mathcal{D}'''_q(1)\mathcal{N}''_s(1)}{3(\mathcal{D}''_q(1))^2} \right] \quad (35)$$

$$\begin{aligned} \mathcal{N}_s'''(1) = & \mathcal{N}_q'''(1) + 6\alpha\mathcal{E}(\mathcal{I})\{\mathcal{E}(\mathcal{D})\{\mathcal{E}(\mathcal{I})(1 - \mathcal{D}_v^*(\omega))[\mathcal{H}^*(\alpha) + \frac{\alpha}{\omega}] + (\bar{\beta} - \mathcal{B})\mathcal{D}_v^*(\omega) \\ & + \mathcal{B} - 1\} - \frac{\alpha}{\omega}[1 - \mathcal{D}_v^*(\omega)]\{\mathcal{B} - \bar{\beta} + \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) - (1 - \mathcal{H}^*(\alpha))\}\} \end{aligned}$$

(iii) The mean period of the consumers in the system (\mathcal{W}_s) and the mean period of the consumers in the queue (\mathcal{W}_q) are estimated utilizing Little's method. $\mathcal{W}_s = \frac{\mathcal{L}_s}{\alpha\mathcal{E}(\mathcal{I})}$ and $\mathcal{W}_q = \frac{\mathcal{L}_q}{\alpha\mathcal{E}(\mathcal{I})}$, respectively.

4.3. Mean busy period and the busy cycle

Under SS circumstances, let the projected lengths of the busy period and busy cycle be $\mathcal{A}(\mathcal{T}_y)$ and $\mathcal{A}(\mathcal{T}_z)$, respectively. The conclusions are directly obtained from the analysis of an alternate renewal process [6], which leads to

$$\mathcal{Q}_0 = \frac{\mathcal{A}(\mathcal{T}_0)}{\mathcal{A}(\mathcal{T}_y) + \mathcal{A}(\mathcal{T}_0)}; \mathcal{A}(\mathcal{T}_y) = \frac{1}{\alpha} \left(\frac{1}{\mathcal{Q}_0} - 1 \right); \mathcal{A}(\mathcal{T}_z) = \frac{1}{\alpha\mathcal{Q}_0} = \mathcal{A}(\mathcal{T}_0) + \mathcal{A}(\mathcal{T}_y). \quad (36)$$

where \mathcal{T}_0 amount of time the system was in its empty condition. As the duration between the arrivals of two consumers differs exponentially. We have the equation $\mathcal{A}(\mathcal{T}_0) = (1/\alpha)$. with variable α . We may recover (27) by applying (36) the previously discovered results,

$$\mathcal{A}(\mathcal{T}_y) = \frac{1}{\alpha} \times \left\{ \frac{\alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D})\left\{\frac{2\alpha}{\omega}(1 - \mathcal{D}_v^*(\omega)) - 2\mathcal{D}_v^*(\omega)\mathcal{H}^*(\alpha) + \mathcal{D}_v^*(\omega) + \mathcal{H}^*(\alpha) + 1\right\} - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))[1 + \alpha\mathcal{E}(\mathcal{D}_v)] - \alpha\mathcal{E}(\mathcal{D})[1 + \mathcal{B}(1 - \mathcal{D}_v^*(\omega))] + \mathcal{B}(1 + \frac{\alpha}{\omega}(1 - \mathcal{D}_v^*(\omega)))}{\mathcal{B} - \{\bar{\beta} - \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))\}} - 1 \right\} \quad (37)$$

$$\mathcal{A}(\mathcal{T}_z) = \frac{1}{\alpha} \times \left\{ \frac{\alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D})\left\{\frac{2\alpha}{\omega}(1 - \mathcal{D}_v^*(\omega)) - 2\mathcal{D}_v^*(\omega)\mathcal{H}^*(\alpha) + \mathcal{D}_v^*(\omega) + \mathcal{H}^*(\alpha) + 1\right\} - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))[1 + \alpha\mathcal{E}(\mathcal{D}_v)] - \alpha\mathcal{E}(\mathcal{D})[1 + \mathcal{B}(1 - \mathcal{D}_v^*(\omega))] + \mathcal{B}(1 + \frac{\alpha}{\omega}(1 - \mathcal{D}_v^*(\omega)))}{\mathcal{B} - \{\bar{\beta} - \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))\}} \right\} \quad (38)$$

5. STOCHASTIC DECOMPOSITION AND SPECIAL CASES

Here, the stochastic decomposition aspect of the system size distribution is examined. In $M/G/1$ queueing models with server vacations, stochastic decomposition has been extensively explored by Fuhrman and Cooper [5]. The no. of consumers in the system at SS at a random point in period is distributed as the sum of two independent RVs, one of which is the no. of consumers in the corresponding standard QS at a random point in time without vacations, and the other of which may have different probabilistic interpretations depending on the scheduling of the vacations. Furthermore, it has been discovered by Krishnakumar and Arivudainambi[11] that stochastic decomposition is valid for a no. of $M/G/1$ RQ models.

Theorem 3. The PGF of no.of consumers in the system ($\mathcal{K}_s(\check{\zeta})$) can be executed as convolution of two independent RVs like, i.e., $\mathcal{K}_s(\check{\zeta}) = \mathcal{M}_a(\check{\zeta}) \cdot \mathcal{M}_b(\check{\zeta})$,

(i) The PGF of no.of consumers $\psi(\check{\zeta})$ in the $M^{[X]}/G^B/1$ feedback RQ with variable server capacity under WV policy.

(ii) The PGF of no.of consumers in the orbit given that the server is idle $\mathcal{M}_b(\check{\zeta})$. **Proof.** The PGF

of the system length can be decomposed as derived:

The stochastic decomposition law formulation is $K_s(\check{\zeta}) = \mathcal{M}_a(\check{\zeta}) \cdot \mathcal{M}_b(\check{\zeta})$,

(i) The system size distribution of the $M^{[X]}/G^B/1$ feedback RQ with variable server capacity under WV is $\psi(\check{\zeta})$ and its distribution can be assigned by $\mathcal{H}^*(\alpha) \rightarrow 1$ in the eqn., (28).

$$\psi(\check{\zeta}) = \mathcal{Q}_0 \left\{ \frac{(1 - \mathcal{F}(\check{\zeta}))\{\check{\zeta}^B - (\beta + \bar{\beta}\check{\zeta})\mathcal{D}^*(\mathcal{S}(\check{\zeta}))\}[1 + \frac{\alpha}{\omega}\check{\zeta}\mathcal{V}(\check{\zeta}) + \check{\zeta}(1 - \mathcal{D}^*(\alpha(1 - \mathcal{F}(\check{\zeta})))\{\check{\zeta}^B\mathcal{V}(\check{\zeta}) + [(\beta + \bar{\beta}\check{\zeta})\mathcal{D}_v^*(\mathcal{S}_v(\check{\zeta})) - 1]\}}{(1 - \mathcal{F}(\check{\zeta}))\{\check{\zeta}^B - (\beta + \bar{\beta}\check{\zeta})\mathcal{D}^*(\mathcal{S}(\check{\zeta}))\}} \right\}$$

(ii) The conditional distribution of the no.of consumers in the system at random point in period given the server is empty $\mathcal{M}_b(\check{\zeta})$.

$$\mathcal{M}_b(\check{\zeta}) = \frac{\mathcal{Q}_0 + \mathcal{Q}(\check{\zeta}) + \Psi_v(\check{\zeta})}{\mathcal{Q}_0 + \mathcal{Q}(1) + \Psi_v(1)}$$

$$\mathcal{M}_b(\check{\zeta}) = \left\{ \frac{\{\mathcal{B} - \{\bar{\beta} - \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))\}}{De(\check{\zeta})\{\mathcal{B} - \{\bar{\beta} - \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))\} + (1 - \mathcal{H}^*(\alpha))\{\bar{\beta} + \alpha\mathcal{E}(\mathcal{I})[\mathcal{E}(\mathcal{D})\mathcal{D}_v^*(\omega) + \frac{1}{\omega}(1 - \mathcal{D}_v^*(\omega)) - \mathcal{E}(\mathcal{D}_v)]\} + \frac{\alpha}{\omega}(1 - \mathcal{D}_v^*(\omega))\}} \right\} \\ \times \{[\check{\zeta}^B - (\beta + \bar{\beta}\check{\zeta})\{\mathcal{H}^*(\alpha) + \mathcal{F}(\check{\zeta})[1 - \mathcal{H}^*(\alpha)]\}\mathcal{D}^*(\mathcal{S}(\check{\zeta})) + \check{\zeta}^B\mathcal{Q}_0(1 - \mathcal{H}^*(\alpha))\{\beta + \bar{\beta}\check{\zeta}\}[\mathcal{D}^*(\mathcal{S}(\check{\zeta}))\mathcal{V}(\check{\zeta}) + \mathcal{D}_v^*(\mathcal{S}_v(z))] - 1\} + \frac{\alpha}{\omega}\mathcal{V}(\check{\zeta})\}$$

From the aforementioned stochastic decomposition law, we see that $\mathcal{K}_s(\check{\zeta}) = \mathcal{M}_a(\check{\zeta}) \cdot \mathcal{M}_b(\check{\zeta})$, which is consistent with the decomposition results of Geo et al. [6], are also applicable for this particular vacation system. ■

5.1. Special cases

In this section, we examine a few real-world examples of our strategy that are consistent with recent literature.

Case (i):

Let $\mathcal{Pr}[\mathcal{F} = 1] = 1$, $\mathcal{B} = 1$, $\omega, \bar{\beta} = 0$ and $\mathcal{H}^*(\alpha) \rightarrow 1$. Our model can be simplified to a $M/G/1$ queue. The results agree with Takagi [21].

$$\mathcal{K}_s(\check{\zeta}) = \mathcal{Q}_0 \left\{ \frac{\mathcal{N}e_s(\check{\zeta})}{\mathcal{D}e_s(\check{\zeta})} \right\} \quad (39)$$

$$\mathcal{N}e_s(\check{\zeta}) = (1 - \check{\zeta})\{\check{\zeta} - \mathcal{D}^*(\alpha(1 - \check{\zeta}))\} + \check{\zeta}(1 - \mathcal{D}^*(\alpha(1 - \check{\zeta})))\{\mathcal{D}_v^*(\alpha(1 - \check{\zeta}))\}$$

$$\mathcal{D}e_s(\check{\zeta}) = (1 - \check{\zeta})\{\check{\zeta} - \mathcal{D}^*(\alpha(1 - \check{\zeta}))\}$$

$$\text{where, } \mathcal{Q}_0 = \frac{1 + \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D})}{\alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) - \alpha\mathcal{E}(\mathcal{D})}$$

Case (ii):

Let $\mathcal{Pr}[\mathcal{F} = 1] = 1$, $\mathcal{B} = 1$, and $\omega, \bar{\beta} = 0$. Our model simplified to an $M/G/1$ RQ. Here are the results agree with Gao and Wang [6].

$$\mathcal{K}_s(\check{z}) = \mathcal{Q}_0 \left\{ \frac{\mathcal{N}e_s(\check{z})}{\mathcal{D}e_s(\check{z})} \right\} \quad (40)$$

$$\begin{aligned} \mathcal{N}e_s(\check{z}) &= (1 - \check{z})\{\check{z} - [\mathcal{H}^*(\alpha) + \check{z}(1 - \mathcal{H}^*(\alpha))]\mathcal{D}^*(\alpha(1 - \check{z}))\} + \check{z}\alpha(1 - \check{z})[1 - \mathcal{H}^*(\alpha)] \\ &\quad (\mathcal{D}_v^*(\alpha(1 - \check{z}) - 1)) + \alpha\check{z}[1 - \mathcal{D}^*(\alpha(1 - \check{z}))]\{\mathcal{D}_v^*(\alpha(1 - \check{z}) - 1) \\ &\quad [\mathcal{H}^*(\alpha) + \check{z}(1 - \mathcal{H}^*(\alpha))]\} \\ \mathcal{D}e_s(\check{z}) &= \alpha(1 - \check{z})\{\check{z} - [\mathcal{H}^*(\alpha) + \check{z}(1 - \mathcal{H}^*(\alpha))]\mathcal{D}^*(\alpha(1 - \check{z}))\} \\ \text{where, } \mathcal{Q}_0 &= \frac{1 + \alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))}{\alpha\mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D})\{1 - \mathcal{H}^*(\alpha)\} - \alpha\mathcal{E}(\mathcal{D}) - \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha))[1 + \alpha\mathcal{E}(\mathcal{D}_v)] + \mathcal{B}} \end{aligned}$$

Case (iii):

Let $\mathcal{P}r[\mathcal{F} = 1] = 1$, $\mathcal{B} = 1$, and $\bar{\beta} = 0$. our model simplified to an $M/G/1$ queue with WVs. Here are the results agree with Zhang and Hou [23].

$$\mathcal{K}_s(\check{z}) = \mathcal{Q}_0 \left\{ \frac{\mathcal{N}e_s(\check{z})}{\mathcal{D}e_s(\check{z})} \right\} \tag{41}$$

$$\begin{aligned} \mathcal{N}e_s(\check{z}) &= (1 - \check{z})\{\check{z} - [\mathcal{H}^*(\alpha) + \check{z}(1 - \mathcal{H}^*(\alpha))]\mathcal{D}^*(\alpha(1 - \check{z}))\} + \check{z}\alpha(1 - \check{z})[1 - \mathcal{H}^*(\alpha)] \\ &\quad (\mathcal{D}_v^*(\alpha(1 - \check{z}) - 1)) + \alpha\check{z}[1 - \mathcal{D}^*(\alpha(1 - \check{z}))]\{\mathcal{D}_v^*(\alpha(1 - \check{z}) - 1) \\ &\quad [\mathcal{H}^*(\alpha) + \check{z}(1 - \mathcal{H}^*(\alpha))]\} \\ \mathcal{D}e_s(\check{z}) &= \alpha(1 - \check{z})\{\check{z} - [\mathcal{H}^*(\alpha) + \check{z}(1 - \mathcal{H}^*(\alpha))]\mathcal{D}^*(\alpha(1 - \check{z}))\} \end{aligned}$$

6. NUMERICAL RESULTS

The various effects on system performance measurements are demonstrated using MATLAB in this section. We examine exponentially distributed retrial times, service times, and slower service times. The numerical measurements that satisfy the stability condition are chosen at random.

Table 2 clearly displays that arrival rate (α) escalates, \mathcal{L}_q , \mathcal{L}_s , Ψ_v are increases. Table 3 displays that feedback rate β escalates, \mathcal{L}_q , \mathcal{L}_s , are increases and \mathcal{Q}_0 decreases. Table 4 displays that lower service rate η_v escalates, \mathcal{L}_q , \mathcal{L}_s , Ψ_v and \mathcal{Q}_0 decreases.

With the impact of the parameters B , α , β , ω , $\chi(\omega)$, $\eta(\omega)$, $\eta_v(\omega)$, Fig. 1 illustrate the

Table 2: \mathcal{Q}_0 and \mathcal{L}_q for different arrival rate (α) for the values of $\mathcal{B} = 30$, $\beta = 0.5$, $\omega = 2$, $\chi(\omega) = 6$, $\eta(\omega) = 0.6$, $\eta_v(\omega) = 0.7$

Arrival rate (α)	\mathcal{Q}_0	\mathcal{L}_q	\mathcal{L}_s	Ψ_v	\mathcal{W}_q
1	0.8495	0.0120	0.0005	0.1274	0.0840
2	0.8606	0.0131	0.0015	0.2582	0.0459
3	0.8704	0.0142	0.0026	0.3917	0.3341
4	0.8789	0.0156	0.0036	0.5273	0.0272
5	0.8859	0.0168	0.0047	0.6644	0.0235
6	0.8914	0.0180	0.0058	0.8023	0.0210
7	0.8954	0.0192	0.0068	0.9402	0.0192

two-dimensional plot that depict the system performance measures. In Fig. 1(a), displays the escalation of the arrival rate (α), (\mathcal{L}_q) and (\mathcal{W}_q) increases. In Fig. 1(b), we found that (\mathcal{L}_s) increases while diminishing the feedback rate β and (Ψ_v).

The three-dimensional graph representing the system performance metrics is shown in Fig. 2. In Fig. 2(a), the surface displays the elevation the (b), we found that (\mathcal{W}_q) diminishes while increasing the feedback rate (β), (\mathcal{L}_s). In Fig. 2(c), we found that (\mathcal{Q}_0) and (Ψ_v) diminishes while increasing the lower service rate η_v .

The numerical findings above may be used to determine the impact of attributes on the system's assessment criteria, and we can be sure that the results are representative of actual conditions.

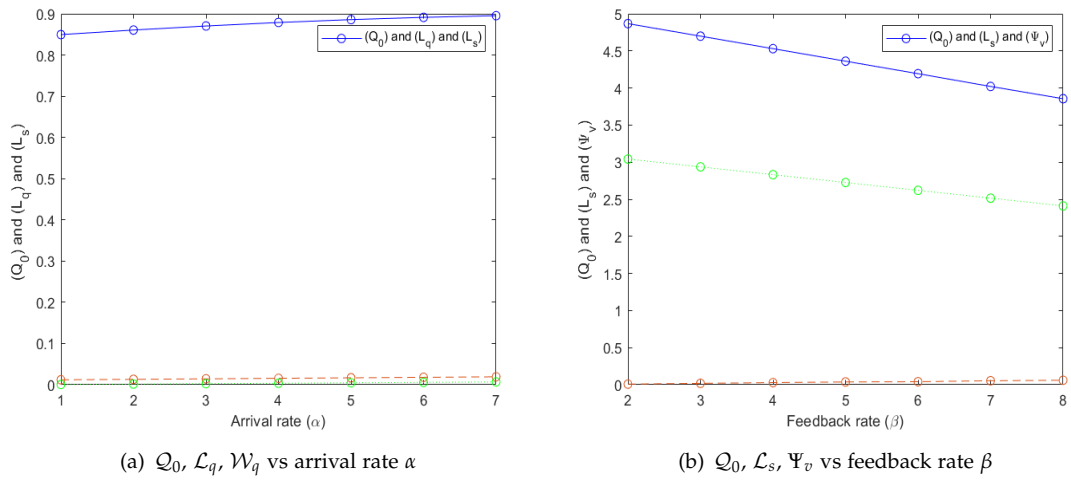


Figure 1: 2D visualization of α and β

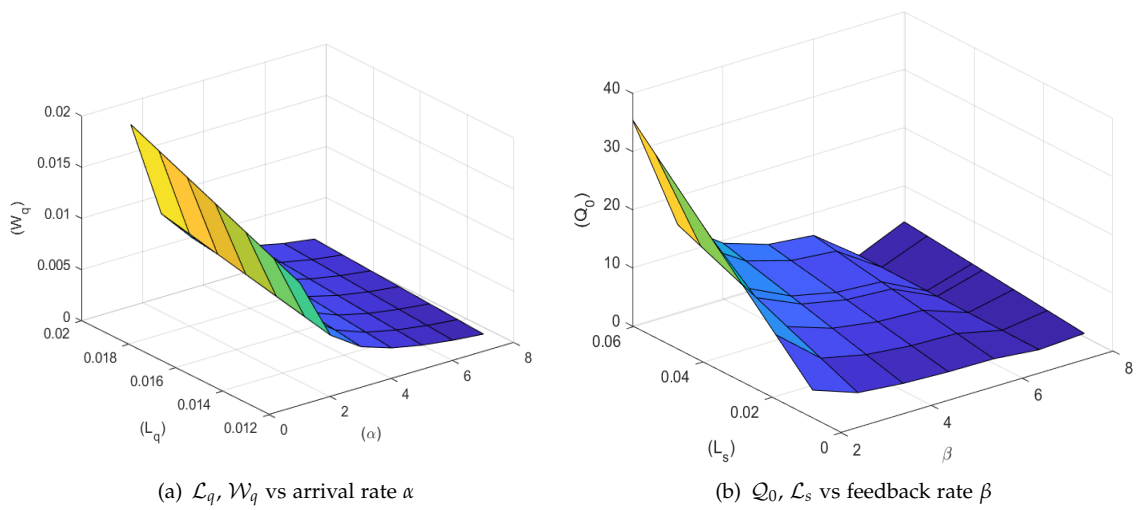


Figure 2: 3D visualization of $\alpha, \beta,$ and η_v

Table 3: Q_0 and \mathcal{L}_q for different arrival rate (α) for the values of $\mathcal{B} = 30, \alpha = 5, \omega = 4, \chi(\omega) = 6, \eta(\omega) = 0.6, \eta_v(\omega) = 0.5$

Feedback rate (β)	Q_0	\mathcal{L}_q	\mathcal{L}_s	Ψ_v	\mathcal{W}_q
2	4.8675	0.0205	0.0082	3.0421	0.0286
3	4.6988	0.0322	0.0189	2.9367	0.0450
4	4.5301	0.0431	0.0287	2.8313	0.0603
5	4.3614	0.0531	0.0377	2.7259	0.0743
6	4.1928	0.0623	0.0407	2.6205	0.0872
7	4.0209	0.0706	0.0536	2.5150	0.0989
8	3.8554	0.0781	0.0603	2.4096	0.0109

Table 4: Q_0 and \mathcal{L}_q for different lower service rate (η_v) for the values of $\mathcal{B} = 30, \alpha = 1, \omega = 4, \chi(\omega) = 4, \eta(\omega) = 0.6, \beta = 0.7$

Lower service rate (η_v)	Q_0	\mathcal{L}_q	\mathcal{L}_s	Ψ_v	\mathcal{W}_q
0.1	1.0937	0.0430	0.0030	0.2461	0.0998
0.2	1.0457	0.0138	0.0026	0.2091	0.0968
0.3	1.0018	0.0134	0.0021	0.1753	0.0941
0.4	0.9614	0.0131	0.0018	0.1442	0.0916
0.5	0.9242	0.0128	0.0015	0.1155	0.0893
0.6	0.8897	0.0124	0.0012	0.0889	0.0871
0.7	0.8577	0.0122	0.0009	0.0643	0.0851

7. CONCLUSION

We examined the $M^{[X]}/G^B/1$ feedback retrial queueing system with variable server capacity under working vacation in this article. If all essential and appropriate conditions are met, the system can be stabilized. When it is ideal, normally busy, and on lower rate service, the PGF of the no. of system consumers and its orbit are calculated using the PGF approach and the supplementary variable technique. Eventually, a wide range of numerical findings are presented to examine the impact of system parameters. The results of this study may be used in the design of different computer communication systems, packet switching networks, manufacturing lines, and postal systems by network and software engineers. Lastly, a study of a bulk service queueing system with priority consumers under working vacation could enhance this work. Additionally, it may be worthwhile to investigate in the future of transient solution for bulk service under a working vacation.

Acknowledgments: Not applicable

Funding information: Not applicable

Conflicts of interest: The authors declare no conflict of interest.

Data availability: Not applicable

Supplementary material: Not applicable

APPENDIX A

Theorem 4. The embedded Markov chain $\{\mathcal{G}_m; m \in M\}$ is ergodic iff $\Lambda < \mathcal{B}$ for our system will be stable, where $\Lambda = \beta - \alpha \mathcal{E}(\mathcal{I})E(\mathcal{D}) + E(\mathcal{I})(1 - \mathcal{H}^*(\alpha))$.

Proof. Foster's [15] criteria, which claim that the chain $\{\mathcal{G}_m; m \in M\}$ is an irreducible and

aperiodic chain, may be used to easily confirm the required condition of ergodicity. Assuming a non-negative measure $e(r)$, $r \in M$ and $\delta > 0$, the Markov chain is ergodic, and mean drift $v_r = \mathcal{E}[e(u_{m+1}) - e(u_m)]/v_m = r$ with a limited exception r 's, $r \in M$ and $v_r \leq -\delta \forall r \in M$. In this case, we're focusing on the function $e(r) = r$. Next, we obtain

$$v_r = \begin{cases} \bar{\beta} - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) - \mathcal{B}, & \text{if } r=0 \\ \bar{\beta} - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) - \mathcal{B}, & \text{if } r=1,2,\dots \end{cases}$$

In this case, $\bar{\beta} - \alpha \mathcal{E}(\mathcal{I})\mathcal{E}(\mathcal{D}) + \mathcal{E}(\mathcal{I})(1 - \mathcal{H}^*(\alpha)) < \mathcal{B}$ is undoubtedly a prerequisite for ergodicity.

As said by Humblett et al. [8], if the Markov chain $\{\mathcal{G}_m; m \in M\}$ matches Kaplan's status, specifically $v_r < \infty \forall r \geq 0$ and $\exists r_0 \in M$ such that $v_r \geq 0$ for $r \geq r_0$, the necessary condition is satisfied. $\mathcal{V} = (v_{qr})$ is the the unit-step transition matrix of $\{\mathcal{G}_m; m \in M\}$ for $r < q - j$ and $q > 0$. The Markov chain's non-ergodicity is suggested by $\Lambda \geq B$. ■

REFERENCES

- [1] Artalejo, J.R. and Gómez-Corral, A. (2008). Limiting Distribution of the System State, In Retrial Queueing Systems: A Computational Approach, Berlin, Heidelberg.
- [2] Bailey, N.T. (1954). On queueing processes with bulk service, *Journal of the Royal Statistical Society*, 16:80–87.
- [3] Banerjee, A. Gupta, U.C. and Sikdar, K. (2013). Analysis of finite-buffer bulk-arrival bulk service queue with variable service capacity and batch-size-dependent service, *International Journal of Mathematics in Operational Research*, 5:358–386.
- [4] Chandrasekaran, V. Indhira, K. Saravananarajan, M. and Rajadurai, R. (2016). A survey on working vacation queueing models, *International Journal of Pure and Applied Mathematics*, 106:33–41.
- [5] Fuhrmann, S.W. and Cooper, R.B. (1985). Stochastic decompositions in the $M/G/1$ queue with generalized vacation, *Operation Research*, 33:1117–1129.
- [6] Gao, S. Wang, J. and Li, W. (2014). An $M/G/1$ retrial queue with general retrial times, working vacations and vacation interruption, *Asia Pacific Journal of Operational Research*, 31:6–31.
- [7] Choudhury, G. (2002). A batch arrival queue with a vacation time under single vacation policy, *Computers and Operations Research*, 29:1941–1955.
- [8] Humblett, Sennott, and Tweedi, (1983). Average drifts and the non-Ergodicity of Markov chains, *Operation Research*, 31:783–789.
- [9] Jaiswal, N.K. (1961). A bulk service queueing problem with variable capacity, *Journal of the Royal Statistical Society*, 106:143–148.
- [10] Ke, J.C. Wu, C.H. and Zhang, Z.G. (2010). Recent developments in vacation queueing models: a short survey, *International Journal of Operation Research*, 7:3–8.
- [11] Krishnakumar, B. and Arivudainambi, D. (2002). The $M/G/1$ retrial queue with Bernoulli schedules and general retrial times, *Computers and Mathematics with Applications*, 43:15–30.
- [12] Madhu Jain, and Anshul Kumar, (2022). Unreliable Server $M^{[X]}/G/1$ Queue with Working Vacation and Multi-phase Repair with Delay in Verification, *International journal of applied engineering research*, 8:1–17.
- [13] Madhu Jain, and Anshul Kumar, (2022). Unreliable Server $M^{[X]}/G/1$ Retrial Feedback Queue with Balking, Working Vacation and Vacation Interruption, *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*.
- [14] Maragathasundari, S and Balamurugan, B. (2015). Analysis of an $M^{[X]}/G/1$ feedback queue with two stages of repair times, general delay time, *International journal of applied engineering research*, 10:25165–25174.
- [15] Pakes, A.G. (1969). Some conditions for Ergodicity and recurrence of Markov chains, *Operation Research*, 17:1058–1061.

- [16] Pazhani Bala Murugan, S. and Vijaykrishnaraj, R. (2019). A bulk arrival retrial queue with feedback and exponentially distributed multiple working vacation, *International Journal of Computing Sciences Research*, 10:81–91.
- [17] Rajadurai, P. (2018). Sensitivity analysis of an $M/G/1$ retrial queueing system with disaster under working vacations and working breakdowns, *RAIRO Operation Research*, 52:35–54.
- [18] Sasikala, S. and Indhira, K. (2016). Bulk service queueing models-a survey, *International Journal of Pure Applied Mathematics*, 106:43–56.
- [19] Sasikala, S. Indhira, K. and Chandrasekaran, V. M. (2017). A study on $M^X/G^B/1$ retrial queueing system with Bernoulli vacation schedule and variable server capacity, *International Journal of Knowledge Management in Tourism and Hospitality*, 1:263–277.
- [20] Shan Gao, (2011). An $M^{[X]}/G/1$ Queue System with Delayed Single Working Vacation, *Energy Procedia*, 13:1493–1498.
- [21] Takagi, H. (1991). Vacation and priority systems. Part I. Queueing analysis: A foundation of performance evaluation, North-Holland, Amsterdam, New York.
- [22] Yang, D.Y. and Wu, C.H. (2015). Cost-minimization analysis of a working vacation queue with N -policy and server breakdowns, *Computers and Industrial Engineering*, 82:151–158.
- [23] Zhang, M. and Hou, Z. (2012). $M/G/1$ queue with single working vacation, *Journal of Applied Mathematics and Computing*, 39:221–234.

STOCHASTIC MODELING AND PERFORMABILITY ANALYSIS OF REPAIRABLE SYSTEM OF A PLYWOOD INDUSTRY

Mr. Amit Kumar Singh ¹, Dr. P. C. Tewari ²

•

- (1) Research Scholar, Department of Mechanical Engineering, National Institute of Technology, Kurukshetra- 136119, India
E-mail: amitsingh54321@gmail.com
- (2) Professor, Department of Mechanical Engineering, National Institute of Technology, Kurukshetra- 136119, India
E-mail: pctewari1@gmail.com

Abstract

The current paper analyzes the performance behavior concerning the performability of the Veneer layup system in a plywood industry. A Markovian Approach is utilized to develop a process model for the system and enhance to evaluate system performability i.e. the function of system availability. The study investigates the impact of varying failure and repair rates on the availability of system, variation in the availability is also determined by varying available repair facilities, using a licensed software package. Particle Swarm Optimization (PSO) method has been employed to optimize the results. Additionally, a Decision Support System (DSS) has been proposed for making strategic decisions regarding financial investments and maintenance order priorities. The findings of the paper will aid the practitioners in deciding the maintenance order priorities among various subsystems.

Keywords: Performability, Markov Chain, Decision Support System, Particle Swarm Availability Optimization

I. Introduction

The manufacturing process of plywood involves several intricate stages, including veneer cutting, placing up and gluing operations, pressing, and finishing processes. As market competition intensifies, manufacturers must continually enhance the performance of production processes. The utilization of human labor offers flexibility, yet the need for varying sizes of the final product often disrupts the layup stage, a critical phase. This condition has adversely affected factors such as availability, production costs, quality, and in some cases, operator safety. However, modern business communities within these sectors have turned this challenge into a learning opportunity. Industrialists are fervently engaged in operating process plants and industries continuously, aiming to minimize the breakdowns in this competitive era. This endeavor is essential for maintaining maximum productivity and ensuring the highest profits to ensure the survival of the industry concerned.

The performance of a system is enhanced through proper design and maintenance throughout its service life. Through a case study analysis, the paper demonstrates how Markov techniques can provide insights into system performance to identify bottlenecks and suggest strategies for

improving production efficiency and product quality. The findings underscore the potential of Markov models as valuable tools for decision-making and process optimization in the plywood industry. Tewari and Khan [11] discussed by using quasi-independence Markov chain and entropy methods, demonstrate a predictable sequence of sedimentary structures, reflecting typical fluvial channel processes. Malik and Tewari [6] dealt with the performance modeling and maintenance order priorities for the Feed Water System in a thermal power plant based on coal and also analyzed the system process by using Chapman-Kolmogorov equations and Markov approaches. Abedi, Yoon, and Kwon [1] discussed a cyclic time-dependent Markov process and reinforcement learning for a battery energy storage control system. Wu and Hoa [12] optimized feature mappings and Markovian models using the Koopman operator's top singular components and introduces score functions for model optimization. Khan and Tewari [3] introduced the Kolmogorov criterion for analyzing transition matrices of reversible Markov processes.

Malik [5] developed a performability model for the Coal Ash Handling System (CAHS) in a thermal power plant operating at subcritical conditions. This model is constructed by aggregating state probabilities using a normalizing condition. Parkash [8] designed Performance Modeling and proposed a DSS to prioritize repairs tasks for an assembly line system. Kumar [4] proposed a Decision Support Priorities (DSP) framework, highlighting the criticality of different useful units. Singh and Tewari [9], Sheikh and Tewari [10] discussed the applications of Reliability, Availability, Maintainability and Safety (RAMS) concepts in various process industries for enhancement of performability. Stochastic processes deal with randomness in systems, like how things change over time in unpredictable ways. Performability analysis facilitates the performance behavior of the system concerned. This helps us make smarter choices about how to design and run systems to make them more reliable and effective.

The Markovian approach analyzes systems based on present states, regardless of past events. It's useful for systems with discrete states and helps prioritize maintenance in assembly lines by modeling subsystem performance. Hale et. al. [2] explored the use of quantitative or conceptual methods to create the Markov chain model of particular industrial unit. Marozzi and Mostarda [7] discussed stochastic processes for Byzantine Fault Tolerant performance evaluation.

Particle Swarm Optimization (PSO) is a computational optimization method inspired by the collective behavior of birds or fish in search of their food. In PSO, a group of potential solutions, called particles, navigates the search space to locate the optimal solution. Through iterative adjustments, PSO effectively converges towards the optimal solution. A DSS is a high-tech tool that helps decision-makers to analyze the data, generate the reports, and evaluate the alternatives to make the right decisions quickly and effectively.

II. System Description

In the plywood industry, there are typically nine primary steps involved in the production process. These include (a) log collection, (b) debarking, (c) steaming blocks, (d) peeling blocks and veneer cutting machine, (e) drying veneers, (f) gluing and stacking the veneers on top of each other, (g) pressing the veneers in hot and cold presses, (h) trimming the plywood, and (i) super finishing and grade stamping, as illustrated in Figure 1. In plywood industry veneer production system is accounts for approximately 37 to 42% of the total plant output. The system is being studied as plywood industry and base material used as poplar and eucalyptus wood, situated within the Ganga basin of Northern India, this area encompasses several subsystems. These subsystems include:

- **Debarking Machine:** Debarking machine removes bark from logs before to turned into plywood sheets. Logs go in, bark comes off using rotary cutters or water jets, and clean logs come out for further processing. This step ensures high-quality veneer or

chips free from bark-related defects and contaminants, extending the life of subsequent machinery.

- **Veneer Cutting Machine:** Veneer cutting machines are vital in woodworking and plywood manufacturing, turning logs into thin, even veneer sheets. Then use rotating blades or drums to slice layers from logs, ensuring consistent thickness and quality. These machines vary in design and features, customized for different wood types and production needs, with some equipped with automation for improved efficiency.
- **Veneer Drier:** A veneer dryer is a specialized subsystem in plywood production, extracting moisture from freshly cut veneer sheets to achieve the desired moisture content for further processing and storage. Through controlled heat and airflow, it prevents defects like warping or cracking. Techniques like hot air circulation or infrared radiation may be used depending on the dryer's design. Efficient moisture removal is vital for producing high-quality plywood sheets with uniform thickness and strength.
- **Gluing and Pasting:** Gluing and pasting are key processes in plywood production. Veneer surfaces are prepared and coated with adhesive before being stacked and pressed to create strong bonds. Curing ensures structural integrity, followed by trimming and finishing for precise dimensions and surface quality. This results in resilient plywood panels used in construction and furniture.

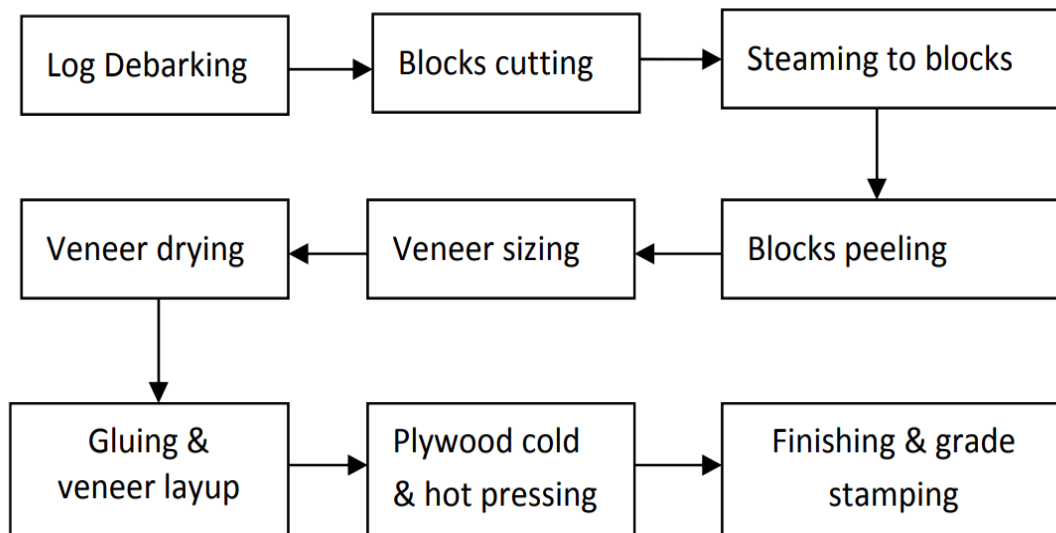


Figure 1: Flow Diagram of Plywood Manufacturing Process

The plywood - making process involves preparing veneer sheets from logs, sorting and grading them, sanding or cleaning for adhesion, applying adhesive, stacking with perpendicular grain orientations, pressing to activate the adhesive, curing, and finally trimming and finishing for precise dimensions and quality.

III. Assumptions and Notations

Markov chains rely on several assumptions to effectively model systems. Firstly, to assume stationary, meaning that transition probabilities between states remain consistent over time. This assumption is vital for the stability of the model, allowing us to make reliable predictions about future states based on current probabilities. In other term the probability of transitioning to the next state depends solely on the current state and is unaffected by the history of previous states.



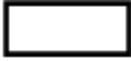
This simplifies the model and makes computations more manageable. The most important assumption is a finite state space, implying that the set of possible states the system can occupy is finite and well-defined. To construct a transition matrix that encapsulates all possible state transitions. Furthermore, assume homogeneity, meaning that transition probabilities are consistent across different time periods. This assumption is essential for making long-term predictions about the system's behavior without being influenced by short-term fluctuations.

Notifications play a vital role in keeping all involved informed about the system's dynamics and relevant updates. Hence to encompass alerts concerning state transitions, providing clear insights into the current state and the probabilities associated with transitioning to subsequent states. Regular updates on state probabilities, derived from observed transitions, enable continuous tracking of the system's behavior over time. The performance modeling of the system relies on certain assumptions and notations, which are as follows:

a) Assumptions:

- Failure and Repair rates are constant over time and statistically independent.
- The system has the potential to operate at a reduced capacity.
- The standby systems exhibit similar characteristics to the active system.
- Service encompasses both repair and replacement of components.
- Simultaneous failures do not take place.
- A subsystem that undergoes repair is considered to be in a condition equivalent to new for a specified period.
- Adequate repair facilities are available to commence repairs promptly, without any delay.

b) Notations:

-  : Denotes the system concerned is working at its full capacity state.
-  : Denotes the system concerned is working at reduced capacity state.
-  : Denotes the system concerned is working at failed state.
- A, Bi, C, D: Indicate that the subsystems are in a fully functioning condition.
- a, b, c, d: Denotes that subsystems A, Bi, C, and D are in a state of failure.
- P0(t): Probability of the system operating at full capacity at time t.
- P1(t) – P5(t): Probabilities associated with the system operating in a state of reduced capacity.
- P6(t) – P29(t): Probabilities of the system in failed state.
- $q_{i,i=1-4}$: Average failure rates for subsystems A, Bi, C, and D, respectively.
- $\mu_i, i=1-4$: Average repair rates for subsystems A, Bi, C, and D, respectively.
- d/dt : Characterizes derivative with respect to time (t).

The diagram illustrating the transitions between states of Veneer Layup System is given in Figure 1. In which state 0 denotes the working of system with full capacity, states 1,2,3,4 and 5 are working of systems with reduced capacity and states 6 – 29 have failed.

IV. Performance Modeling of System

The performance modeling for Veneer Layup System is carried out by Markov Birth-Death Process using a probabilistic approach and a differential equation related to the transition diagram. In performance modeling using Markov analysis, systems are depicted as transitioning between various states based on predefined probabilities. Initially, states representing different configurations or conditions of the system are identified.

These transitions are quantified through transition probabilities, reflecting the likelihood of moving from one state to another within a defined timeframe. Constructing a transition matrix encapsulates these probabilities, facilitating analysis of system behavior. Through this model, metrics like steady-state probabilities or mean time to absorption can be calculated, offering insights into system performance. These equations are determined to describe the steady-state performability of the system.

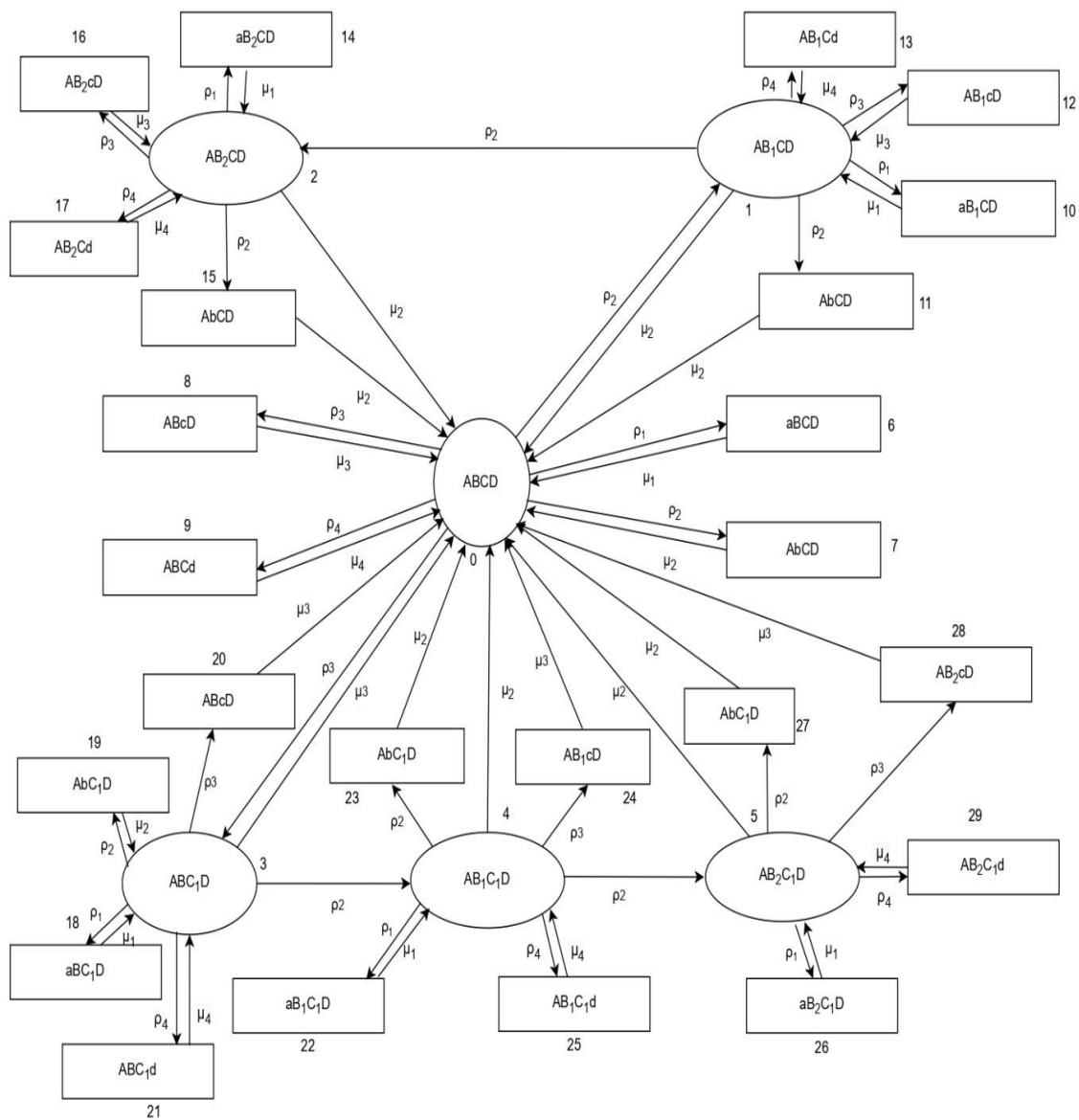


Figure 2: Performance Model of Veneer Cutting System of Plywood Industry

V. Performance Analysis

A system's first-order differential equations associated to the transition diagram (Figure 2) at time $(t+\Delta t)$, can be expressed as follows using the mnemonic rule:

$$P_0(t+\Delta t) - P_0(t) = [-Q_1\Delta t - 2Q_2\Delta t - 2Q_3\Delta t - Q_4\Delta t]P_0(t) + \mu_2P_1(t)\Delta t + \mu_2P_2(t)\Delta t + \mu_3P_3(t)\Delta t + \mu_2P_4(t)\Delta t + \mu_2P_5(t)\Delta t + \mu_1P_6(t)\Delta t + \mu_2P_7(t)\Delta t + \mu_3P_8(t)\Delta t + \mu_4P_9(t)\Delta t + \mu_2P_{11}(t)\Delta t + \mu_2P_{15}(t)\Delta t + \mu_2P_{20}(t)\Delta t + \mu_3P_{24}(t)\Delta t + \mu_2P_{27}(t)\Delta t + \mu_3P_{28}(t)\Delta t \quad (1)$$

After dividing both sides by Δt , the outcome is:

$$[P_0(t+\Delta t) - P_0(t)]/\Delta t = [-Q_1-2Q_2-2Q_3-Q_4]P_0(t) + \mu_2P_1(t) + \mu_2P_2(t) + \mu_3P_3(t) + \mu_2P_4(t) + \mu_2P_5(t) + \mu_1P_6(t) + \mu_2P_7(t) + \mu_3P_8(t) + \mu_4P_9(t) + \mu_2P_{11}(t) + \mu_2P_{15}(t) + \mu_2P_{20}(t) + \mu_3P_{24}(t) + \mu_2P_{27}(t) + \mu_3P_{28}(t) \quad (2)$$

After assuming that $\Delta t \rightarrow 0$ is the limit, this can be found as:

$$P'_0(t) = -X_0P_0(t) + \mu_2P_1(t) + \mu_2P_2(t) + \mu_3P_3(t) + \mu_2P_4(t) + \mu_2P_5(t) + \mu_1P_6(t) + \mu_2P_7(t) + \mu_3P_8(t) + \mu_4P_9(t) + \mu_2P_{11}(t) + \mu_2P_{15}(t) + \mu_2P_{20}(t) + \mu_3P_{24}(t) + \mu_2P_{27}(t) + \mu_3P_{28}(t)$$

or

$$P'_0(t) + X_0P_0(t) = \mu_2P_1(t) + \mu_2P_2(t) + \mu_3P_3(t) + \mu_2P_4(t) + \mu_2P_5(t) + \mu_1P_6(t) + \mu_2P_7(t) + \mu_3P_8(t) + \mu_4P_9(t) + \mu_2P_{11}(t) + \mu_2P_{15}(t) + \mu_2P_{20}(t) + \mu_3P_{24}(t) + \mu_2P_{27}(t) + \mu_3P_{28}(t) \quad (3)$$

Similarly

$$P'_1(t) + X_1P_1(t) = Q_2P_0(t) + \mu_1P_{10}(t) + \mu_3P_{12}(t) + \mu_4P_{13}(t) \quad (4)$$

$$P'_2(t) + X_2P_2(t) = Q_2P_1(t) + \mu_1P_{14}(t) + \mu_3P_{16}(t) + \mu_4P_{17}(t) \quad (5)$$

$$P'_3(t) + X_3P_3(t) = Q_3P_0(t) + \mu_1P_{18}(t) + \mu_2P_{19}(t) + \mu_4P_{21}(t) \quad (6)$$

$$P'_4(t) + X_4P_4(t) = Q_2P_3(t) + \mu_1P_{22}(t) + \mu_4P_{25}(t) \quad (7)$$

$$P'_5(t) + X_5P_5(t) = Q_2P_4(t) + \mu_1P_{26}(t) + \mu_4P_{29}(t) \quad (8)$$

Where

$$X_0 = Q_1 + 2Q_2 + 2Q_3 + Q_4$$

$$X_1 = Q_1 + 2Q_2 + Q_3 + Q_4 + \mu_2$$

$$X_2 = Q_1 + Q_2 + Q_3 + Q_4 + \mu_2$$

$$X_3 = Q_1 + 2Q_2 + Q_3 + Q_4 + \mu_3$$

$$X_4 = Q_1 + 2Q_2 + Q_3 + Q_4 + \mu_2$$

$$X_5 = Q_1 + Q_2 + Q_3 + Q_4 + \mu_2$$

$$P'_6(t) + \mu_1P_6(t) = Q_1P_0(t) \quad (9)$$

$$P'_7(t) + \mu_2P_7(t) = Q_2P_0(t) \quad (10)$$

$$P'_8(t) + \mu_3P_8(t) = Q_3P_0(t) \quad (11)$$

$$P'_9(t) + \mu_4P_9(t) = Q_4P_0(t) \quad (12)$$

$$P'_{10}(t) + \mu_1P_{10}(t) = Q_1P_1(t) \quad (13)$$

$$P'_{11}(t) + \mu_2P_{11}(t) = Q_2P_1(t) \quad (14)$$

$$P'_{12}(t) + \mu_3P_{12}(t) = Q_3P_1(t) \quad (15)$$

$$P'_{13}(t) + \mu_4P_{13}(t) = Q_4P_1(t) \quad (16)$$

$$P'_{14}(t) + \mu_1P_{14}(t) = Q_1P_2(t) \quad (17)$$

$$P'_{15}(t) + \mu_2P_{15}(t) = Q_2P_2(t) \quad (18)$$

$$P'_{16}(t) + \mu_3P_{16}(t) = Q_3P_2(t) \quad (19)$$

$$P'_{17}(t) + \mu_4P_{17}(t) = Q_4P_2(t) \quad (20)$$

$$P'_{18}(t) + \mu_1P_{18}(t) = Q_1P_3(t) \quad (21)$$

$$P'_{19}(t) + \mu_2P_{19}(t) = Q_2P_3(t) \quad (22)$$

$$P'_{20}(t) + \mu_3P_{20}(t) = Q_3P_3(t) \quad (23)$$

$$P'_{21}(t) + \mu_4P_{21}(t) = Q_4P_3(t) \quad (24)$$

$$P'_{22}(t) + \mu_1P_{22}(t) = Q_1P_4(t) \quad (25)$$

$$P'_{23}(t) + \mu_2P_{23}(t) = Q_2P_4(t) \quad (26)$$

$$P'_{24}(t) + \mu_3P_{24}(t) = Q_3P_4(t) \quad (27)$$

$$P'_{25}(t) + \mu_4P_{25}(t) = Q_4P_4(t) \quad (28)$$

$$P'_{26}(t) + \mu_1P_{26}(t) = Q_1P_5(t) \quad (29)$$

$$P'_{27}(t) + \mu_2P_{27}(t) = Q_2P_5(t) \quad (30)$$

$$P'_{28}(t) + \mu_3 P_{28}(t) = Q_3 P_5(t) \quad (31)$$

$$P'_{29}(t) + \mu_4 P_{29}(t) = Q_4 P_5(t) \quad (32)$$

It is a complex system, and every system must be accessible for a long time to achieve the maximum output. The steady-state behavior of the plywood plant can be investigated by finding $t \rightarrow \infty, - \rightarrow 0$.

With this, equations from (01) to (32) reduced to

$$X_0 P_0(t) = \mu_2 P_1(t) + \mu_2 P_2(t) + \mu_3 P_3(t) + \mu_2 P_4(t) + \mu_2 P_5(t) + \mu_1 P_6(t) + \mu_2 P_7(t) + \mu_3 P_8(t) + \mu_4 P_9(t) + \mu_2 P_{11}(t) + \mu_2 P_{15}(t) + \mu_2 P_{20}(t) + \mu_3 P_{24}(t) + \mu_2 P_{27}(t) + \mu_3 P_{28}(t) \quad (33)$$

Similarly

$$X_1 P_1(t) = Q_2 P_0(t) + \mu_1 P_{10}(t) + \mu_3 P_{12}(t) + \mu_4 P_{13}(t) \quad (34)$$

$$X_2 P_2(t) = Q_2 P_1(t) + \mu_1 P_{14}(t) + \mu_3 P_{16}(t) + \mu_4 P_{17}(t) \quad (35)$$

$$X_3 P_3(t) = Q_3 P_0(t) + \mu_1 P_{18}(t) + \mu_2 P_{19}(t) + \mu_4 P_{21}(t) \quad (36)$$

$$X_4 P_4(t) = Q_2 P_3(t) + \mu_1 P_{22}(t) + \mu_4 P_{25}(t) \quad (37)$$

$$X_5 P_5(t) = Q_2 P_4(t) + \mu_1 P_{26}(t) + \mu_4 P_{29}(t) \quad (38)$$

$$\mu_i P_j(t) = Q_i P_0(t), \text{ where, } i=1,2,3,4; j= 6,7,8,9 \quad (39)$$

$$\mu_i P_j(t) = Q_i P_1(t), \text{ where, } i=1,2,3,4; j= 10,11,12,13 \quad (40)$$

$$\mu_i P_j(t) = Q_i P_2(t), \text{ where, } i=1,2,3,4; j= 14,15,16,17 \quad (41)$$

$$\mu_i P_j(t) = Q_i P_3(t), \text{ where, } i=1,2,3,4; j= 18,19,20,21 \quad (42)$$

$$\mu_i P_j(t) = Q_i P_4(t), \text{ where, } i=1,2,3,4; j= 22,23,24,25 \quad (43)$$

$$\mu_i P_j(t) = Q_i P_5(t), \text{ where, } i=1,2,3,4; j= 26,27,28,29 \quad (44)$$

By solving these equations as:

Taking K as a constant, which is the ratio of failure rate to repair rate,

$$K = -$$

$$K_1 = -, \quad K_2 = -, \quad K_3 = -, \quad K_4 = -, \quad K_5 = \text{—————}, \quad K_6 = \text{—————},$$

$$K_7 = \text{—————}, \quad K_8 = \text{—————}, \quad K_9 = \text{—————}$$

$$P_1 = K_8 P_0 \quad (45)$$

$$P_2 = K_8 K_9 P_0 \quad (46)$$

$$P_3 = K_5 P_0 \quad (47)$$

$$P_4 = K_5 K_6 P_0 \quad (48)$$

$$P_5 = K_5 K_6 K_7 P_0 \quad (49)$$

$$P_6 = K_1 P_0 \quad (50)$$

$$P_7 = K_2 P_0 \quad (51)$$

$$P_8 = K_3 P_0 \quad (52)$$

$$P_9 = K_4 P_0 \quad (53)$$

$$P_{10} = K_1 K_8 P_0 \quad (54)$$

$$P_{11} = K_2 K_8 P_0 \quad (55)$$

$$P_{12} = K_3 K_8 P_0 \quad (56)$$

$$P_{13} = K_4 K_8 P_0 \quad (57)$$

$$P_{14} = K_1 K_8 K_9 P_0 \quad (58)$$

$$P_{15} = K_2 K_8 K_9 P_0 \quad (59)$$

$$P_{16} = K_3 K_8 K_9 P_0 \quad (60)$$

$$P_{17} = K_4 K_8 K_9 P_0 \quad (61)$$

$$P_{18} = K_1 K_5 P_0 \quad (62)$$

$$P_{19} = K_2 K_5 P_0 \quad (63)$$

$$P_{20} = K_3 K_5 P_0 \quad (64)$$

$$P_{21} = K_4 K_5 P_0 \quad (65)$$

$$P_{22} = K_1 K_5 K_6 P_0 \quad (66)$$

- $P_{23} = K_2K_5K_6P_0$ (67)
- $P_{24} = K_3K_5K_6P_0$ (68)
- $P_{25} = K_4K_5K_6P_0$ (69)
- $P_{26} = K_1K_5K_6K_7P_0$ (70)
- $P_{27} = K_2K_5K_6K_7P_0$ (71)
- $P_{28} = K_3K_5K_6K_7P_0$ (72)
- $P_{29} = K_4K_5K_6K_7P_0$ (73)

In accordance with the normalization principle, the sum of collective probabilities of all events should be equal to one that is:

$$\Sigma P_i = 1 \tag{74}$$

$$P_0 + P_1 + P_2 + \dots + P_{29} = 1 \tag{75}$$

$$P_0 [1 + (K_8 + K_8K_9 + K_5 + K_5K_6 + K_5K_6K_7 + K_1 + K_2 + K_3 + K_4 + K_1K_8 + K_2K_8 + K_3K_8 + K_4K_8 + K_1K_8K_9 + K_2K_8K_9 + K_3K_8K_9 + K_4K_8K_9 + K_1K_5 + K_2K_5 + K_3K_5 + K_4K_5 + K_1K_5K_6 + K_2K_5K_6 + K_3K_5K_6 + K_4K_5K_6 + K_1K_5K_6K_7 + K_2K_5K_6K_7 + K_3K_5K_6K_7 + K_4K_5K_6K_7)] = 1$$

or

$$P_0 = 1 / [1 + (K_8 + K_8K_9 + K_5 + K_5K_6 + K_5K_6K_7 + K_1 + K_2 + K_3 + K_4 + K_1K_8 + K_2K_8 + K_3K_8 + K_4K_8 + K_1K_8K_9 + K_2K_8K_9 + K_3K_8K_9 + K_4K_8K_9 + K_1K_5 + K_2K_5 + K_3K_5 + K_4K_5 + K_1K_5K_6 + K_2K_5K_6 + K_3K_5K_6 + K_4K_5K_6 + K_1K_5K_6K_7 + K_2K_5K_6K_7 + K_3K_5K_6K_7 + K_4K_5K_6K_7)] \tag{76}$$

Now, the system availability $A(\infty)$ can be found by using:

$$A(\infty) = P_0 + P_1 + P_2 + P_3 + P_4 + P_5 = [1 + K_8 + K_8K_9 + K_5 + K_5K_6 + K_5K_6K_7]P_0 \tag{77}$$

Using equation 77, the long-term availability for a range of permissible combinations of veneer manufacturing systems' failure and repair rates in a steady state can be ascertained. Table 1 provides an overview of how failure and repair rates affect the availability of the system. Availability impacts system performance by ensuring that each part of the system is ready to work when desired.

Table 1: Failure and Repair Rates of Veneer System

Sub-System's Name	Mean Failure Rate (Q_i)	Mean Repair Rate (μ_i)
Debarking Machine (A)	0.013 (Q_1)	0.15 (μ_1)
Veneer Cutting Machine (B)	0.004 (Q_2)	0.2 (μ_2)
Veneer Drier (C)	0.0024 (Q_3)	0.126 (μ_3)
Gluing and Pasting (D)	0.005 (Q_4)	0.19 (μ_4)

Table 2 and Figure 3 describe the effect of different failure and repair rates of a debarking machine on system performance, in terms of availability. It is observed that as the failure rate increases from 0.003 to 0.043, the system's performability declines from 0.8886 to 0.5194, marking a decrease of 41.5%. Likewise, with the repair rate increasing from 0.05 to 0.45, the system's performability improves from 0.8886 to 0.9328, reflecting a 4.7% increase.

Table 2: Effect of the Failure and Repair Rates of Debarking Machine subsystem on system Performability(%)

Failure Rates (Q_1)	Repair Rates of Debarking Machine (μ_1)					Constant Parameters
	0.05	0.15	0.25	0.35	0.45	
0.003	0.8886	0.9214	0.9282	0.9312	0.9328	$Q_2 = 0.004,$ $\mu_2 = 0.2,$ $Q_3 = 0.0024,$ $\mu_3 = 0.126,$ $Q_4 = 0.005,$ $\mu_4 = 0.19$
0.013	0.7545	0.8680	0.8950	0.9070	0.9139	
0.023	0.6556	0.8205	0.8640	0.8841	0.8957	
0.033	0.5796	0.7780	0.8352	0.8623	0.8782	
0.043	0.5194	0.7396	0.8082	0.8416	0.8614	

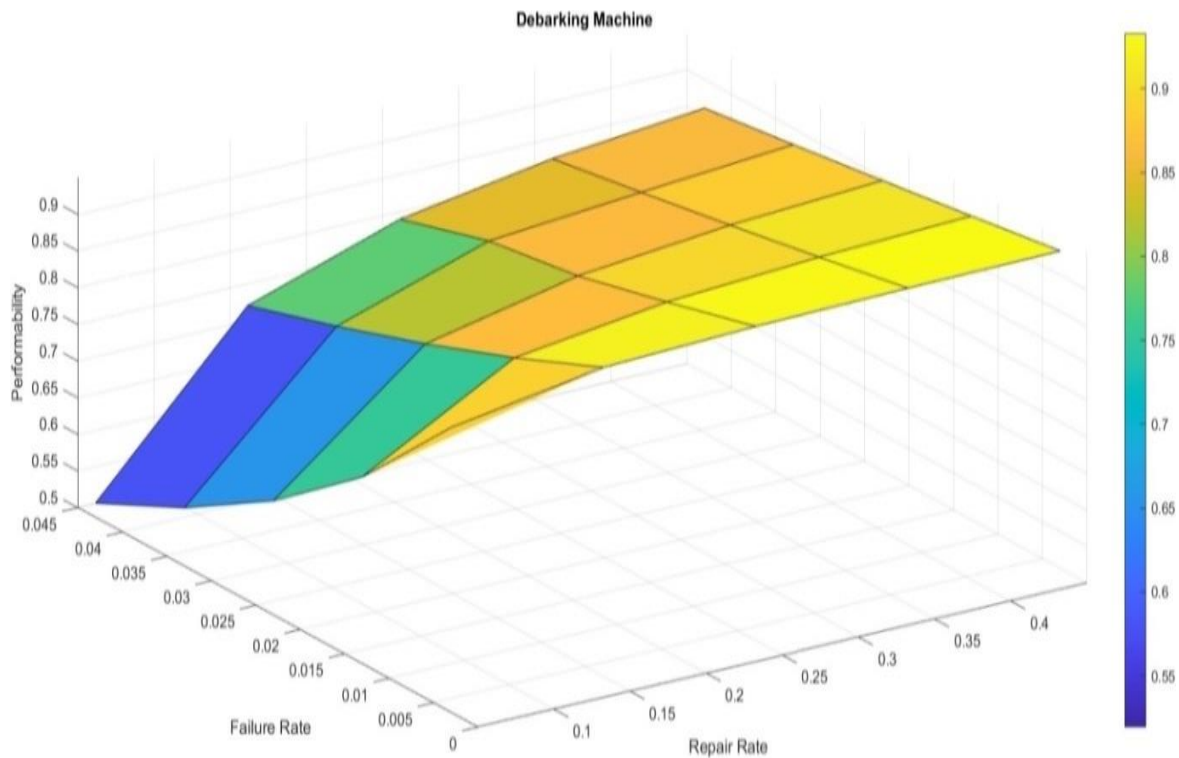


Figure 3: Performability Variation with respect to Failure and Repair Rates of Debarking Machine Subsystem

Similarly, in Table 3 and Figure 4 for the Veneer Cutting Machine subsystem, the performance of the subsystem in terms of availability varies between 3.35% and 1.175% for the respective failure (Q_2) and repair rates (μ_2) when all other factors stay the same.

Table 3: Effect of the Failure and Repair Rates of Veneer Cutting Machine Subsystem on System Performability (%)

Failure Rates (Q_2)	Repair Rates of Veneer Cutting Machine (μ_2)					Constant Parameters
	0.1	0.15	0.2	0.25	0.3	
0.002	0.8680	0.8731	0.8756	0.8772	0.8782	$Q_1 = 0.013,$ $\mu_1 = 0.15,$ $Q_3 = 0.0024,$ $\mu_3 = 0.126,$ $Q_4 = 0.005,$ $\mu_4 = 0.19$
0.003	0.8606	0.8680	0.8718	0.8741	0.8756	
0.004	0.8532	0.8630	0.8680	0.8711	0.8731	
0.005	0.8460	0.8581	0.8643	0.8680	0.8706	
0.006	0.8389	0.8532	0.8606	0.8650	0.8680	

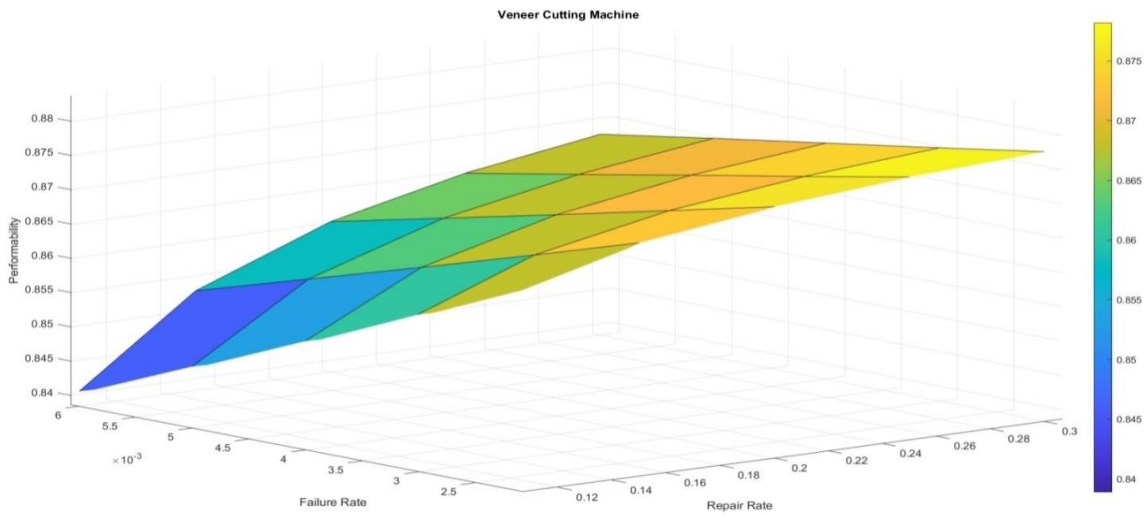


Figure 4: Performability Variation with respect to Failure and Repair Rates of Veneer Cutting Machine Subsystem

Table 4 and Figure 5 illustrate the impact of repair and failure rates of the Veneer Drier subsystem on its performability. It is observed that the known values of failure rate (ρ_3) and repair rate (μ_3) of Veneer drier, as the failure rate increases from 0.0012 to 0.0036, and the performability of the system decreases quickly from 0.8739 to 0.8569, i.e., 1.945%. Also, as the repair rate (ρ_3) increases from 0.106 to 0.146, the performability of the system increases considerably from 0.8739 to 0.8763, i.e., 0.28%.

Table 4: Effect of the Failure and Repair Rates for Veneer Drier Subsystem on System Performability(%)

Failure Rates (ρ_3)	Repair Rates of Veneer Drier (μ_3)					Constant Parameters
	0.106	0.116	0.126	0.136	0.146	
0.0012	0.8739	0.8746	0.8753	0.8758	0.8763	$Q_1 = 0.013,$ $\mu_1 = 0.15,$ $Q_2 = 0.004,$ $\mu_2 = 0.2,$ $Q_4 = 0.005,$ $\mu_4 = 0.19$
0.0018	0.8696	0.8707	0.8716	0.8724	0.8731	
0.0024	0.8653	0.8668	0.8680	0.8691	0.8700	
0.0030	0.8611	0.8629	0.8645	0.8658	0.8669	
0.0036	0.8569	0.8591	0.8609	0.8625	0.8638	

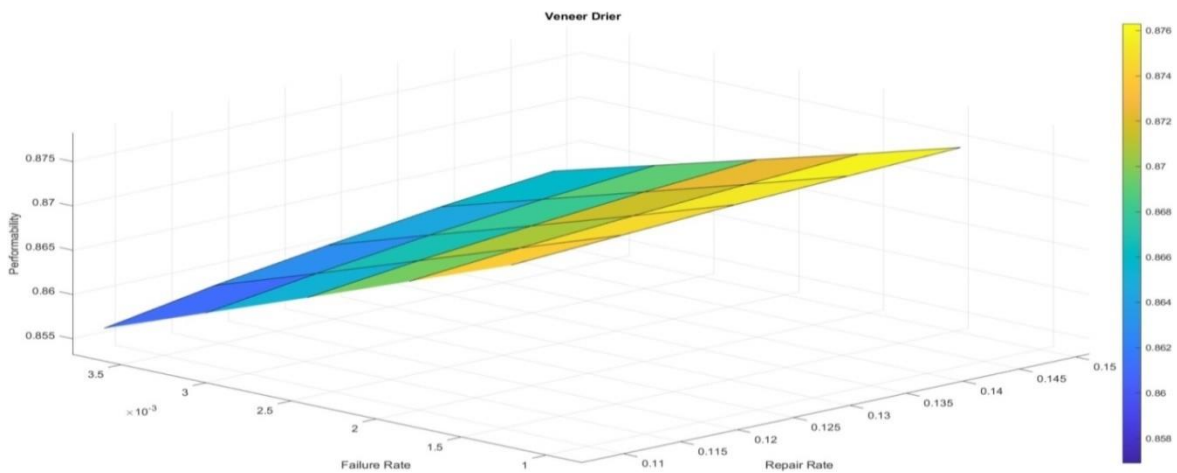


Figure 5: Performability Variation with respect to Failure and Repair Rates of Veneer Driers Subsystem

In the Table 5 and Figure 6 describe various combinations of repair and failure rates for gluing and pasting subsystem that influence their performability. It's clearly shown that for distinct values of failure rate (Q_4) and repair rate (μ_4), when the failure rate increases from 0.001 to 0.009, then performability decreases from 0.8729 to 0.7659 i.e. 12.26%. In the same way as the repair rate increases from 0.05 to 0.33, gluing and pasting performability increases drastically from 0.8729 to 0.8859 i.e. (1.49%).

Table 5: Effect of the Failure and Repair Rates of Gluing and Pasting Subsystem on System Performability(%)

Failure Rates	Repair Rates of Gluing and Pasting					Constant Parameters
	0.05	0.12	0.19	0.26	0.33	
0.001	0.8729	0.8818	0.8842	0.8853	0.8859	$Q_1 = 0.013,$ $\mu_1 = 0.15,$ $Q_2 = 0.004,$ $\mu_2 = 0.2,$ $Q_3 = 0.0024,$ $\mu_3 = 0.126$
0.003	0.8434	0.8690	0.8760	0.8793	0.8812	
0.005	0.8159	0.8566	0.8680	0.8734	0.8765	
0.007	0.7901	0.8446	0.8602	0.8676	0.8719	
0.009	0.7659	0.8328	0.8525	0.8618	0.8673	

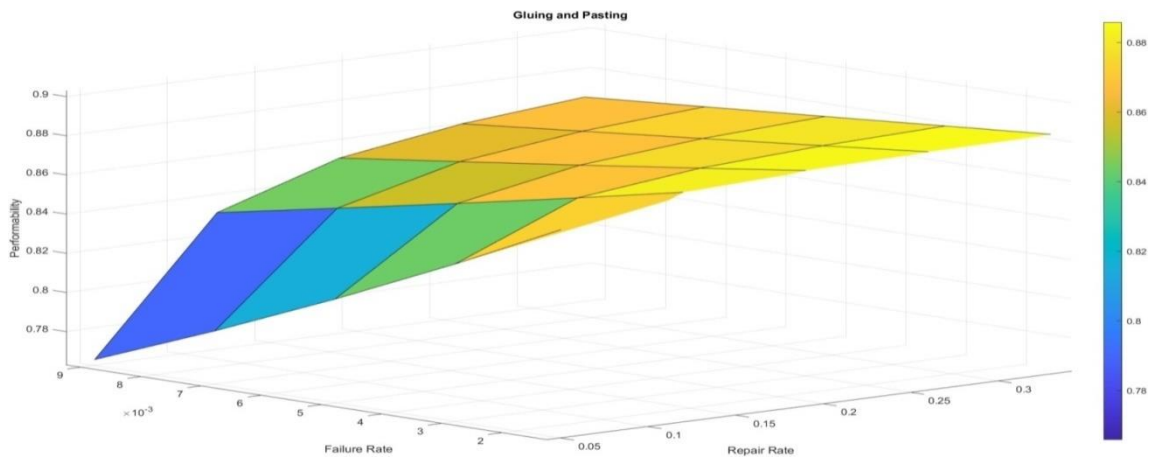


Figure 6: Performability Variation with respect to Failure and Repair Rates of Gluing and Pasting Subsystem

VI. Results

This paper explores the application of Markov techniques to evaluate the performance of production system in the plywood industry by providing DSS regarding maintainability order. According to Table 6, the study indicates that the Veneer Drier subsystem contributes the least to the system's performance, while the Debarking Machine subsystem is the most critical subsystem of the assembly line system.

Table 6: Effect of Subsystems Failure and Repair Rates Deviation on System Performance

Name of Sub-System	Variation in Failure Rates Q_i , (Repair Rates μ_i)	Impact of Variation on System Performance (%)	Proposed Maintenance Priority
Debarking Machine (A)	0.003-0.043, (0.05-0.45)	0.9328-0.5194 (44.318%)	I
Gluing and Pasting (D)	0.00-0.009, (0.05-0.33)	0.8859-0.7659 (13.546%)	II
Veneer Cutting Machine (B)	0.002-0.006, (0.1-0.3)	0.8782-0.8389 (4.475%)	III
Veneer Drier (C)	0.0012-0.0036, (0.106-0.146)	0.8763-0.8569 (2.214%)	IV

The Markov approach is used to analyze the performance in terms of availability. If there is a need to increase the performability of such systems, it should be recommended to enhance the system performance using optimization techniques like Ant Colony Algorithm, PSO and Teacher Learning Based Optimization etc.

References

- [1] Abedi, S., Yoon, S.W. and Kwon, S. (2022). International journal of electrical power and energy systems battery energy storage control using a reinforcement learning approach with cyclic time-dependent Markov process. *International Journal of Electrical Power and Energy Systems*, 134:107368.
- [2] Hale, A.R. et al. (1998). Evaluating safety in the management of maintenance activities in the chemical process industry. *Safety Science*, 28:21–44.
- [3] Khan, Z.A. and Tewari, R.C. (2021). Markov reversibility, quasi-symmetry and marginal homogeneity in cyclothymiacs geological successions. *ISSRG International Journal of Geo-informatics and Geological Science*, 8:9–25.
- [4] Kumar, N. (2024). Stochastic approach in performance modeling of a repairable industrial system, *Engineering Modeling*, 37:15–28.
- [5] Malik, S. (2023). Performability and maintenance decisions for coal ash handling system of a subcritical thermal power plant. *International Journal of System Assurance Engineering and Management*, 14:45–54.
- [6] Malik, S. and Tewari, P.C. (2018). Performance modeling and maintenance priorities decision for the water flow system of a coal-based thermal power plant. *International Journal of Quality and Reliability Management*, 35:996–1010.
- [7] Marcozzi, M. and Mostarda, L. (2024). Analytical model for performability evaluation of practical byzantine fault-tolerant systems. *Expert Systems with Applications*, 238:121838.
- [8] Parkash, S. (2022). Performance modeling and DSS for assembly line. *Reliability: Theory & Applications*, 17:403–412.
- [9] Sheikh, M. and Tewari, P.C. (2024). A critical literature review and future perspective of RAM approaches for complex. *Reliability: Theory & Applications*, 19:431–439.
- [10] Singh, A.K. and Tewari, P.C. (2023). An overview of reliability, availability, maintainability, and safety strategies for complex systems in various process industries. *International Journal of Performability Engineering*, 19:788–796.
- [11] Tewari, R.C. and Khan, Z.A. (2017). Structures and sequences in early permian fluvial barakar rocks of peninsular India Gondwana basins using binomial and Markov chain analysis. *Arabian Journal of Geosciences*, 10:13.
- [12] Wu, H. and Noé, F. (2020). Variational approach for learning Markov processes from time series data. *Journal of Nonlinear Science*, 30:23–66.

E-BAYESIAN ESTIMATION FOR BATHTUB-SHAPED LIFETIME DISTRIBUTION BASED ON UPPER RECORD VALUES

SANA¹ AND M. FAIZAN²

¹ Department of Science, Aligarh College of Education, Aligarh, India.

² Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, India.

¹ Corrospoding Author: sanarais233@gmail.com, ² mdfaizan02@gmail.com

Abstract

In this research paper, we presents the expected Bayesian (E-Bayesian) estimation of bathtub-shaped lifetime (BSL) distribution for scale parameter based on upper record values (URV) using a conjugate prior distribution. Also, we are considered different prior distributions for the E-Bayesian estimators. Some properties of the E-Bayesian estimators are discussed. A simulation study is given to compare the performance of the E-Bayesian estimators with Bayesian estimator. we notice that the E-Bayesian estimators are perform better than the Bayesian estimators. Moreover, the performance of the Bayesian estimators and E-Bayesian estimators for Prior II are better than Prior I. Also, we observe that if we increase the sample size n then the estimators are showing lesser mean square error (MSE).

Keywords: Bathtub-shaped lifetime distribution, Bayes estimation, E-Bayes estimation, Upper Record values.

1. INTRODUCTION

Suppose that X_1, X_2, \dots is a sequence of independent and identically distributed (*iid*) random variables with a cumulative distribution function (*cdf*) $F(x)$ and a probability density function (*pdf*) $f(x)$. Let $Y_m = \max(\min) \{X_1, X_2, \dots, X_m\}$ for $m \geq 1$. We say X_j is an URV or LRV (lower) of this sequence $\{X_m, m \geq 1\}$ if $Y_j > (<) Y_{j-1}, j \geq 2$. By definition, X_1 is URV as well as a LRV. One can transform the upper records to lower records by replacing the original sequence $\{X_j; j \geq 1\}$ by $\{-X_j; j \geq 1\}$ or $Pr(X_i > 0) = 1$ if for all i by $\{\frac{1}{X_i}, i \geq 1\}$, the LRV of this sequence will correspond to the URV of the original sequence, (Ahsanullah, 2004) [1].

The BSL distribution was introduced by Chen (2000) [2]. The probability density function (*pdf*) of the BSL distribution with parameters α and β ($Chen(\alpha, \beta)$) is

$$f(x; \alpha, \beta) = \alpha \beta x^{\beta-1} \exp[\alpha(1 - e^{x^\beta}) + x^\beta], \quad x > 0, \quad \alpha, \beta > 0 \quad (1)$$

with corresponding cumulative distribution function (*cdf*) as

$$F(x; \alpha, \beta) = 1 - \exp[\alpha(1 - e^{x^\beta})], \quad x > 0, \quad \alpha, \beta > 0. \quad (2)$$

The E-Bayesian method was introduced by Han (1997) [3]. In recent years, there has been a growing interest in the study of E-Bayesian estimation. E-Bayesian estimate and its properties were considered by Han (2007) [4], Han (2011) [6] and Han (2011) [7] for the case of one and two hyperparameters. E-Bayesian estimate for the parameters of the geometric distribution based on URV and their relations were obtained by Okasha and Wang (2016) [14]. Han (2009) [5] discussed

the properties of E-Bayes estimate with three different prior distributions of hyperparameters. Jaheen and Okasha (2011) [8] studied the E-Bayesian estimation for the Burr type XII distribution based on the type-II censoring. Okasha (2012) [11] discussed the E-Bayesian method for computing estimates for the unknown parameters of the Weibull distribution based on type-II-censored samples. Okasha (2014) [12] studied that E-Bayesian methods for estimating the parameters of Lomax distribution under the balanced squared error loss function based on type-II censored data. Okasha (2019) [13] used the Burr XII model for type-II censored data. He observed that the E-Bayes estimates performs better than the Bayes estimates. For more details, see, Kizilaslan (2017) [9], Kizilaslan (2019) [10] and Piriaei et al. (2020) [15].

The main object of this article is to discuss the E-Bayesian estimation of the BSL distribution based on URV. Bayesian estimators of the BSL distribution are given in Section 2. E-Bayesian estimation based on three different distributions to the hyperparameters is derived in Section 3. Properties of the E-Bayesian estimators based on squared error loss (SEL) function are discussed in Section 4. Simulation study is given in Section 5. Finally, the conclusion of this article is discussed in Section 6.

2. BAYESIAN ESTIMATION

We consider m URV $X_{U(1)} = x_1, X_{U(2)} = x_2, \dots, X_{U(m)} = x_m$, from $Chen(a, b)$, with pdf (1). In this case, Ahsanullah (2004) [1] gives the likelihood function as

$$l(\alpha, \beta | \underline{x}) = f(x_{U(m)}; \alpha, \beta) \prod_{i=1}^{m-1} \frac{f(x_{U(i)}; \alpha, \beta)}{1 - F(x_{U(i)}; \alpha, \beta)}. \quad (3)$$

Using (1), (2) and (3), we get

$$l(\alpha, \beta | \underline{x}) = \alpha^m \psi(\beta, \underline{x}) e^{-\alpha L}, \quad (4)$$

where

$$\underline{x} = (x_1, x_2, \dots, x_m), \quad \psi(\beta, \underline{x}) = \beta^m \prod_{i=1}^m x_i^{\beta-1} e^{x_i^\beta}$$

and

$$L = e^{x_m^\beta} - 1.$$

When β is known in the two-parameter BSL distribution, the maximum likelihood estimator (MLE) of the scale parameter α , can be written as

$$\hat{\alpha}^{MLE} = \frac{m}{L}. \quad (5)$$

The gamma conjugate prior density of the parameter α can be expressed as

$$h(\alpha) = \frac{\gamma^\lambda}{\Gamma(\lambda)} \alpha^{\lambda-1} e^{-\alpha\gamma}, \quad \alpha > 0. \quad (6)$$

Using (4) and (6), we get posterior density of α , i.e.

$$q(\alpha | \underline{x}) = A^* \alpha^{m+\beta-1} e^{-\alpha(L+\gamma)}, \quad \alpha > 0, \beta > 0 \quad (7)$$

where

$$A^* = \frac{(L + \gamma)^{m+\lambda}}{\Gamma(m + \lambda)}.$$

The Bayes estimate of α based on the SEL function can be expressed as

$$\hat{\alpha}^{BE}(\lambda, \gamma) = \frac{(m + \lambda)}{(L + \gamma)}. \quad (8)$$

3. E-BAYESIAN ESTIMATION

According to Han (1997) [3], the prior parameters λ and γ should be selected to guarantee that the prior $h(\alpha|\lambda, \gamma)$ in (6) is a decreasing function of α . The derivative of $h(\alpha|\lambda, \gamma)$ with respect to α is

$$\frac{dh(\alpha)}{\alpha} = \frac{\gamma^\lambda}{\Gamma\lambda} \alpha^{\lambda-2} e^{-\alpha\gamma} [(\lambda - 1) - \alpha\gamma]. \quad (9)$$

Thus, for $0 < \lambda < 1$, $\gamma > 0$ and $\alpha > 0$, the prior $h(\alpha|\lambda, \gamma)$ is a decreasing function of α . Assuming that the prior parameters λ and γ are independent random variables and their density functions are $\pi_1(\lambda)$ and $\pi_2(\gamma)$ respectively. Then the joint density function of λ and γ is

$$\pi(\lambda, \gamma) = \pi_1(\lambda)\pi_2(\gamma). \quad (10)$$

The E-Bayesian estimate of α can be written as

$$\hat{\alpha}^{EBE} = E(\alpha|\underline{x}) = \int \int_D \hat{\alpha}^{BE}(\lambda, \gamma) \pi(\lambda, \gamma) d\lambda d\gamma, \quad (11)$$

where D is the domain of λ and γ for which the prior density is decreasing in α . $\hat{\alpha}^{BE}(\lambda, \gamma)$ is the Bayes estimate of α as given in (8).

3.1. E-Bayesian Estimation of α under SEL function

In order to obtain E-Bayesian estimation of α , the following distributions of the hyperparameters λ and γ are used

$$\begin{cases} \pi_1(\lambda, \gamma) = \frac{1}{kB(s,g)} \lambda^{s-1} (1-\lambda)^{d-1}, & 0 < \lambda < 1, \quad 0 < \gamma < k, \\ \pi_2(\lambda, \gamma) = \frac{2}{k^2B(s,g)} (k-\gamma) \lambda^{s-1} (1-\lambda)^{d-1}, & 0 < \lambda < 1, \quad 0 < \gamma < k, \\ \pi_3(\lambda, \gamma) = \frac{2\gamma}{k^2B(s,g)} \lambda^{s-1} (1-\lambda)^{d-1}, & 0 < \lambda < 1, \quad 0 < \gamma < k. \end{cases} \quad (12)$$

The E-Bayesian estimates of the parameter α based on SEL function is obtained by using (8), (11) and (12) as

$$\begin{aligned} \hat{\alpha}^{EBE1} &= \int \int_D \hat{\alpha}^{BE}(\lambda, \gamma) \pi_1(\lambda, \gamma) d\lambda d\gamma = \frac{1}{kB(s,g)} \int_0^1 \int_0^c \left(\frac{m+\lambda}{L+\gamma} \right) \lambda^{s-1} (1-\lambda)^{d-1} d\gamma d\lambda, \\ &= \frac{1}{k} \left(m + \frac{s}{s+g} \right) \ln \left(\frac{L+k}{L} \right), \end{aligned} \quad (13)$$

$$\hat{\alpha}^{EBE2} = \frac{2}{k} \left(m + \frac{s}{s+g} \right) \left[\frac{L+k}{k} \ln \left(\frac{L+k}{L} \right) - 1 \right] \quad (14)$$

and

$$\hat{\alpha}^{EBE3} = \frac{2}{k} \left(m + \frac{s}{s+g} \right) \left[1 - \frac{L}{k} \ln \left(\frac{L+k}{L} \right) \right]. \quad (15)$$

4. PROPERTIES OF E-BAYESIAN ESTIMATION BASED ON SEL FUNCTION

In this section, we presents the relations among $\hat{\alpha}^{EBE1}$, $\hat{\alpha}^{EBE2}$ and $\hat{\alpha}^{EBE3}$.

Lemma. Let $0 < k < L$, $s > 0$, $g > 0$ and $\hat{\alpha}^{EBEi}$ ($i = 1, 2, 3$) be given by (13), (14) and (15). Then the following inequalities are:

(i) $\hat{\alpha}^{EBE2} < \hat{\alpha}^{EBE1} < \hat{\alpha}^{EBE3}$.

(ii) $\lim_{L \rightarrow \infty} \hat{\alpha}^{EBE1} = \lim_{L \rightarrow \infty} \hat{\alpha}^{EBE2} = \lim_{L \rightarrow \infty} \hat{\alpha}^{EBE3}$.

Proof. (i) From (13), (14) and (15), we have

$$\hat{\alpha}^{EBE2} - \hat{\alpha}^{EBE1} = \hat{\alpha}^{EBE1} - \hat{\alpha}^{EBE3} = \frac{1}{k} \left(m + \frac{s}{s+g} \right) \left[\frac{k+2L}{k} \ln \left(\frac{L+k}{L} \right) - 2 \right]. \quad (4.1)$$

For $-1 < t < 1$, we have: $\ln(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \dots = \sum_{p=1}^{\infty} (-1)^{p-1} \frac{t^p}{p}$.
Let $t = \frac{k}{L}$, when $0 < k < L$ and $0 < \frac{k}{L} < 1$, we get:

$$\begin{aligned} & \left[\frac{k+2L}{k} \ln \left(\frac{L+k}{L} \right) - 2 \right] \\ &= \frac{k+2L}{k} \left[\left(\frac{k}{L} \right) - \frac{1}{2} \left(\frac{k}{L} \right)^2 + \frac{1}{3} \left(\frac{k}{L} \right)^3 - \frac{1}{4} \left(\frac{k}{L} \right)^4 + \frac{1}{5} \left(\frac{k}{L} \right)^5 - \dots \right] - 2 \\ &= \left[\left(\frac{k}{L} \right) - \frac{1}{2} \left(\frac{k}{L} \right)^2 + \frac{1}{3} \left(\frac{k}{L} \right)^3 - \frac{1}{4} \left(\frac{k}{L} \right)^4 + \frac{1}{5} \left(\frac{k}{L} \right)^5 - \dots \right] - 2 \\ & \quad + \left(2 - \left(\frac{k}{L} \right) + \frac{2}{3} \left(\frac{k}{L} \right)^2 - \frac{2}{4} \left(\frac{k}{L} \right)^3 + \frac{2}{5} \left(\frac{k}{L} \right)^4 - \dots \right) \\ &= \left(\frac{k^2}{6L^2} - \frac{k^3}{6L^3} \right) + \left(\frac{3k^4}{20L^4} - \frac{2k^5}{15L^5} \right) + \dots \\ &= \frac{k^2}{6L^2} \left(1 - \frac{k}{L} \right) + \frac{k^4}{60L^4} \left(9 - \frac{8k}{L} \right) + \dots > 0. \end{aligned} \tag{4.2}$$

According to (4.1) and (4.2), we have

$$\hat{\alpha}^{EBE2} - \hat{\alpha}^{EBE1} = \hat{\alpha}^{EBE1} - \hat{\alpha}^{EBE3} > 0,$$

that is

$$\hat{\alpha}^{EBE2} < \hat{\alpha}^{EBE1} < \hat{\alpha}^{EBE3}.$$

(ii) From (4.1) and (4.2), we get

$$\begin{aligned} \lim_{L \rightarrow \infty} (\hat{\alpha}^{EBE2} - \hat{\alpha}^{EBE1}) &= \lim_{L \rightarrow \infty} (\hat{\alpha}^{EBE1} - \hat{\alpha}^{EBE3}) \\ &= \frac{1}{c} \left(m + \frac{s}{s+g} \right) \lim_{L \rightarrow \infty} \left\{ \frac{c^2}{6L^2} \left(1 - \frac{c}{L} \right) + \frac{c^4}{60L^4} \left(9 - \frac{8c}{L} \right) + \dots \right\} \\ &= 0. \end{aligned}$$

That is, $\lim_{L \rightarrow \infty} \hat{\alpha}^{EBE1} = \lim_{L \rightarrow \infty} \hat{\alpha}^{EBE2} = \lim_{L \rightarrow \infty} \hat{\alpha}^{EBE3}$.

Thus, the proof is complete.

5. SIMULATION STUDY

This section presents, a simulation study for a comparison of Bayes and E-Bayes methods of estimation .

The steps of the simulation are:

1. Sample sizes $n = 20, 30, 40, 50$, no. of records $m = 8$ and for the each case the parameters $(\alpha, \beta) = (1, 2)$.
2. $(s, g) = (0.5, 0.5)$ and $k = 12$.
3. Estimates are calculated by two types of priors:
 - for prior I, hyperparameter values, $(\lambda, \gamma) = (0.2, 2.5)$.
 - for prior II, hyperparameter values, $(\lambda, \gamma) = (0.5, 3)$.
4. The URV are generated from $Chen(\alpha, \beta)$, by using $X = [\log\{1 - \frac{\log(1-U)}{\alpha}\}]^{\frac{1}{\beta}}$, where U is uniform $(0, 1)$.
5. The estimates $\hat{\alpha}^{BE}$ and $\hat{\alpha}^{EBEi}$, $i = 1, 2, 3$ under the SEL function are computed from (8) and (14)-(16).
6. Repeat the above steps for 10,000 times. The average of all 10,000 estimated values are, respectively, calculated and summarized.
7. The computational results are displayed in Table 1. All computations were performed using R Software.

Table 1: Bayesian and E-Bayesian estimates of α (first row), average bias (second row) and mean square error (third row)

n	Par	Prior I				Prior II			
		BE	EBE1	EBE2	EBE3	BE	EBE1	EBE2	EBE3
20	α	1.3940	1.0751	1.3433	0.8068	1.3331	1.0789	1.3491	0.8087
	AB	0.3940	0.0751	0.3433	-0.1932	0.3331	0.0789	0.3491	-0.1913
	MSE	0.2229	0.0389	0.1901	0.0467	0.1646	0.0404	0.1961	0.0461
30	α	1.3086	1.0151	1.2545	0.7757	1.2530	1.0148	1.2543	0.7753
	AB	0.3086	0.0151	0.2545	-0.2243	0.2530	0.0148	0.2543	-0.2247
	MSE	0.1478	0.0255	0.1178	0.0581	0.1075	0.0268	0.1204	0.0586
40	α	1.2505	0.9748	1.1958	0.7537	1.1995	0.9726	1.1927	0.7526
	AB	0.2505	-0.0252	0.1958	-0.2463	0.19951	-0.0274	0.1927	-0.2474
	MSE	0.1079	0.0222	0.0826	0.0677	0.0755	0.0219	0.0806	0.0682
50	α	1.2088	0.9459	1.1542	0.7376	1.1631	0.9445	1.1521	0.7367
	AB	0.2088	-0.0541	0.1542	-0.2625	0.1631	-0.0556	0.1521	-0.2632
	MSE	0.0839	0.0222	0.0626	0.0754	0.0591	0.0222	0.0617	0.0758

6. CONCLUSION

In this paper, Bayes and E-Bayes methods are considered for estimating the unknown scale parameter of the BSL distribution based on URV. Under SEL function and three distributions of the hyperparameters, the E-Bayesian estimators are introduced. Properties of E-Bayesian Estimation based on SEL function are derived. A comparison of Bayes and E-Bayes estimates of scale parameter is performed through a simulation study. From Table [1], we observed that the performance of the E-Bayesian estimators are better than the Bayesian estimators. Moreover, Bayesian estimators and E-Bayesian estimators for Prior II are better than Prior I. Also, from Table [1], we observe that estimators are showing lesser MSEs, as increase the sample size n.

REFERENCES

- [1] M. Ahsanullah. *Record values–Theory and Applications*. University Press of America, 2004.
- [2] Z. Chen. A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics and Probability Letters*, 49(2):155–161, 2000.
- [3] M. Han. The structure of hierarchical prior distribution and its applications. *Chinese Operations Research and Management Science*, 6(3):31–40, 1997.
- [4] M. Han. E-bayesian estimation of failure probability and its application. *Mathematical and Computer Modelling*, 45(9-10):1272–1279, 2007.
- [5] M. Han. E-bayesian estimation and hierarchical bayesian estimation of failure rate. *Applied Mathematical Modelling*, 33(4):1915–1922, 2009.
- [6] M. Han. E-bayesian estimation and hierarchical bayesian estimation of failure probability. *Communications in Statistics-Theory and Methods*, 40(18):3303–3314, 2011.
- [7] M. Han. Estimation of failure probability and its applications in lifetime data analysis. *Journal of Quality and Reliability Engineering*, 2011, 2011.
- [8] Z. F. Jaheen and H. M. Okasha. E-bayesian estimation for the burr type XII model based on type-II censoring. *Applied Mathematical Modelling*, 35(10):4730–4737, 2011.
- [9] F. Kızılaslan. The E-bayesian and hierarchical bayesian estimations for the proportional reversed hazard rate model based on record values. *Journal of Statistical Computation and Simulation*, 87(11):2253–2273, 2017.
- [10] F. Kızılaslan. E-bayesian estimation for the proportional hazard rate model based on record values. *Communications in Statistics-Simulation and Computation*, 48(2):350–371, 2019.

- [11] H. Okasha. E-bayesian estimation of system reliability with Weibull distribution of components based on type-II censoring. *Journal of Advanced Research in Scientific Computing*, 4(4):34–45, 2012.
- [12] H. M. Okasha. E-bayesian estimation for the Lomax distribution based on type-II censored data. *Journal of the Egyptian Mathematical Society*, 22(3):489–495, 2014.
- [13] H. M. Okasha. E-bayesian estimation for the exponential model based on record statistics. *Journal of Statistical Theory and Applications*, 18(3):236–243, 2019.
- [14] H. M. Okasha and J. Wang. E-Bayesian estimation for the geometric model based on record statistics. *Applied Mathematical Modelling*, 40(1):658–670, 2016.
- [15] H. Piriaei, G. Yari, and R. Farnoosh. E-bayesian estimations for the cumulative hazard rate and mean residual life based on exponential distribution and record data. *Journal of Statistical Computation and Simulation*, 90(2):271–290, 2020.

A CLASS OF LOGARITHMIC-CUM-EXPONENTIAL ESTIMATORS FOR POPULATION MEAN WITH RISK ANALYSIS USING DOUBLE SAMPLING

DIWAKAR SHUKLA¹, ASTHA JAIN²

^{1,2} Department of Mathematics and Statistics,
Dr. Harisingh Gour Vishwavidyalaya, Sagar, M.P., 470003
¹diwakarshukla@rediffmail.com, ²asthain2597@gmail.com

Abstract

In order to improve upon the efficiency of an estimate in double sampling for estimating population mean of character under study using an auxiliary variable, a part of survey resources are used to collect the information on auxiliary variable. Some authors have suggested exponential-type estimators and some others advocated for log-type estimators. But combination of such is required for specific situation. This paper presents a class of logarithmic-cum-exponential ratio estimators in double sampling setup. The expressions for the mean squared error and bias of the proposed class of estimators are derived for two different cases (sub-sample and independent sample). Sometimes the persons involved in the sample survey have to undergo for risk on life. For example, data collection in naxalites area, working in intense forest, interview during spread of epidemic or data collection in politically disturbed region. Such risk may affect the accuracy, efficiency of estimation. A linear Risk function is used for the proposed class of estimators. Two cases of double sampling are compared in terms of relative efficiency in view to risk aspect. It is found that the proposed class of estimators has a lower mean squared error than the simple mean estimator, usual ratio, usual exponential, usual log estimators in the double sampling setup. In addition, these theoretical results are supported by a numerical example. Risk function based simulated study is performed for the support of findings of the content. Optimal sample sizes under risk are derived and compared under two cases.

Keywords: Exponential estimator, Logarithmic estimator, Mean squared error, Bias, Risk function, Risk Analysis, Survey sampling, Double sampling, Simple random sampling without replacement(SRSWOR).

1. INTRODUCTION

In double sampling, some part of the resources available for the survey are used to collect data for auxiliary variable. It is because the population mean of auxiliary variable is assumed unknown. Such are collected through sample at the preliminary level and then used to estimate population mean (or population total).

In recent study on the estimators in the double sampling Sahoo et al.[9] discussed the approach of estimating the population mean using regression-type estimator. It boosted the analytical approach of estimation for dealing with double sampling scheme. Bahal and Tuteja[2] developed exponential-type ratio and product estimator for the SRSWOR setup which later extended by the many authors in verity of other sampling schemes. Shashi Bhushan et al.[5] suggested double sampling ratio type estimator using two auxiliary variables. Authors discussed asymptotic properties of the estimators with bias and mean squared error. Shabbir and Sat Gupta[14] suggested exponential ratio-type estimator for estimating the population mean in the setup of

stratified sampling. Such proposal is found to perform better than the usual mean, usual ratio, usual exponential ratio, traditional regression estimators.

Zahoor et. al.[17] suggested regression estimator in double sampling using multi-auxiliary information in the presence of non-response and measurement error in the second phase sample. Such an extension of Azeem[8] who suggested ratio and ratio-cum-exponential estimators in double sampling for population mean incorporating the possibility of non-response and measurement error. The Wu and Luan[6] marked that major advantages of double sampling are the gain in high precision without much substantial increase in cost. Sanaullah et al.[10] suggested generalized exponential-type estimators for the stratified double sampling setup. Sanaullah et al.[12] developed the generalized exponential type estimators for estimating the population variance in double sampling with the help of two auxiliary variables. Zaman and Kadilar[18] proposed exponential ratio-type estimation procedures in the stratified two phase sampling setup. Shukla and Alim[1] proposed parameter estimation approach based an double sampling showing on application in big-data environment. Bhushan and Gupta[3] discussed some log-type estimators using attribute. In another useful contribution Bhushan and Kumar[4] proposed log-type estimators for population mean under the setup of ranked set sampling.

1.1. Risk in data collection

While the conduct of sample survey, using the personal interview method, some areas may be politically disturbed, some may dangerous due to being forest area, some may risky because of naxalites movement and few may under the risk of intense epidemic spread (like Covid-19). Such exposure of risk may possible on the life of field workers involved in data collection. Consider an example where area of a district exposed under risk are identified as A, B, C, D and each having different zones z_1, z_2, z_3, z_4, z_5 with percentage of risk varying over zones.

Table 1: Risk distribution as per area and zones

Area of District	Zones with risk (r_i)					Overhead Risk(r')
	z_1	z_2	z_3	z_4	z_5	
A	25%	10%	20%	30%	7%	8%
B	15%	13%	28%	12%	22%	10%
C	35%	14%	5%	25%	10%	11%
D	16%	11%	18%	19%	23%	13%

Risk per units (r_i) belongs to zones and overhead risk r' belong to the geographical areas of a district.

Deriving motivational idea and scientific approach from above contributions, this paper consider the development of new class of estimators under the risk of life of surveyor during data collection using double sampling.

1.2. Symbols used for population

Let a population of finite size N , D be the variable of main interest and A is an auxiliary variable correlated to D . The pair $(D_i, A_i), i = 1, 2, 3, \dots, N$ represents population values such that

$$\bar{D} = \frac{1}{N} \sum_{i=1}^N D_i, \quad \bar{A} = \frac{1}{N} \sum_{i=1}^N A_i \tag{1.1}$$

$$S_d^2 = \frac{1}{N-1} \sum_{i=1}^N (D_i - \bar{D})^2, \quad S_a^2 = \frac{1}{N-1} \sum_{i=1}^N (A_i - \bar{A})^2 \quad (1.2)$$

$$S_{da} = \frac{1}{N-1} \sum_{i=1}^N (D_i - \bar{D})(A_i - \bar{A}), \quad C_{da} = \frac{S_{DA}}{(\bar{D}\bar{A})} \quad (1.3)$$

$$C_d = \frac{S_d}{\bar{D}}, \quad C_a = \frac{S_a}{\bar{A}}, \quad \rho = \frac{S_{da}}{S_d S_a}, \quad M = \rho \frac{C_d}{C_a} \quad (1.4)$$

where C_d and C_a denote coefficient of variations, ρ correlation coefficient.

1.3. Notations in SRSWOR Setup:

Assumed that information about variable of main interest D is not available, so a simple random sampling is used, using sample of size $n (n < N)$, to predict about that. Further, in usual practice such assumes population mean of auxiliary variable \bar{A} available. All possible samples are $\binom{N}{n}$.

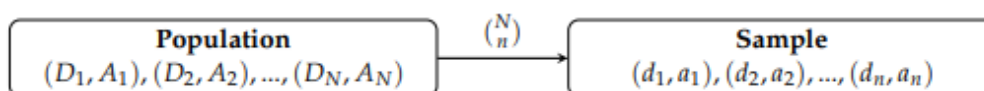


Figure 1: Population and Sample

Let values of random sample by SRSWOR are $(d_i, a_i), i = 1, 2, 3, \dots, n$ then one can define sample statistics as:

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i, \quad \bar{a} = \frac{1}{n} \sum_{i=1}^n a_i \quad (1.5)$$

$$s_d^2 = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2, \quad s_a^2 = \frac{1}{n-1} \sum_{i=1}^n (a_i - \bar{a})^2 \quad (1.6)$$

$$s_{da} = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})(a_i - \bar{a}), \quad \hat{M} = \frac{s_{da}}{s_a^2} \quad (1.7)$$

1.4. Some usual estimators in SRSWOR

- (a) Usual Ratio Estimator: $\hat{D}_R = \frac{\bar{d}}{\bar{a}} \bar{A}$
- (b) Usual Product Estimator: $\hat{D}_P = \frac{\bar{d}\bar{a}}{\bar{A}}$
- (c) Usual Regression Estimator: $\hat{D}_{Re} = \bar{d} + \hat{M}(\bar{A} - \bar{a})$
- (d) Usual Log Estimator: $\hat{D}_L = \bar{d} \left[1 + \log \left(\frac{\bar{A}}{\bar{a}} \right) \right]$
- (e) Usual Exponential Estimator: $\hat{D}_{Ex} = \bar{d} \left[\exp \left(\frac{\bar{A} - \bar{a}}{\bar{A} + \bar{a}} \right) \right]$

Some useful symbols are:

$$V_{qs} = \frac{E\{(\bar{d} - \bar{D})^q (\bar{a} - \bar{A})^s\}}{\bar{D}^q \bar{A}^s}, \quad V'_{qs} = \frac{E\{(\bar{d} - \bar{D})^q (\bar{a}' - \bar{A})^s\}}{\bar{D}^q \bar{A}^s}; q, s = 0, 1, 2$$

$$\begin{aligned}
 V_{20} &= \left(\frac{1}{n} - \frac{1}{N}\right) C_d^2, & V_{02} &= \left(\frac{1}{n} - \frac{1}{N}\right) C_a^2, & V'_{02} &= \left(\frac{1}{n'} - \frac{1}{N}\right) C_a^2 \\
 V_{11} &= \left(\frac{1}{n} - \frac{1}{N}\right) \rho C_d C_a, & V'_{11} &= \left(\frac{1}{n'} - \frac{1}{N}\right) \rho C_d C_a
 \end{aligned}$$

Symbols have their usual meaning as adopted by the survey practitioners in the concerned literature. The Bias $Bias(\cdot)$ and Mean Squared Error $MSE(\cdot)$ of above existing estimators under SRSWOR are expressed as under:

$$Bias(\hat{D}_R) = \bar{D} [V_{02} - V_{11}], \quad MSE(\hat{D}_R) = \bar{D}^2 [V_{20} - 2V_{11} + V_{02}] \quad (1.8)$$

$$Bias(\hat{D}_P) = \bar{D} [V_{02} + V_{11}], \quad MSE(\hat{D}_P) = \bar{D}^2 [V_{20} + 2V_{11} + V_{02}] \quad (1.9)$$

$$Bias(\hat{D}_{Re}) = \bar{D} [V_{02} - \hat{\beta} V_{11}], \quad MSE(\hat{D}_{Re}) = \bar{D}^2 [V_{20} - 2\hat{\beta} V_{11} + \hat{\beta}^2 V_{02}] \quad (1.10)$$

$$Bias(\hat{D}_L) = \bar{D} [V_{02} - V_{11}], \quad MSE(\hat{D}_L) = \bar{D}^2 [V_{20} - 2V_{11} + V_{02}] \quad (1.11)$$

$$Bias(\hat{D}_{Ex}) = \bar{D} \left[\frac{3}{8} V_{02} - \frac{1}{2} V_{11} \right], \quad MSE(\hat{D}_{Ex}) = \bar{D}^2 \left[V_{20} + \frac{1}{4} V_{02} - V_{11} \right] \quad (1.12)$$

2. DOUBLE SAMPLING APPROACH

When the information about population mean of variable is not available then during sample survey with the extra risk and efforts, the sample could be obtained using two different strategies.

Assume n' be the size of first sample with values $(a'_1, a'_2, \dots, a'_{n'})$ and $\bar{a}' = \frac{1}{n'} \sum_{i=1}^{n'} a'_i$

- **Case I:** When the second-phase sample of size n is a sub-sample of the first-phase sample of size n'

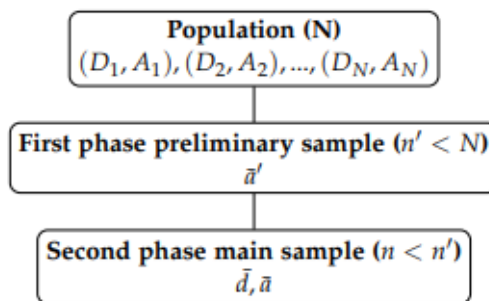


Figure 2: Sampling strategy under case I

- **Case II:** When the second-phase sample of size n is drawn independently of the first-phase sample of size n' .

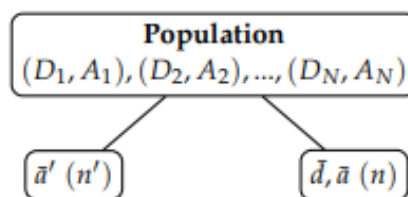


Figure 3: Sampling strategy under case II

2.1. Some existing estimators in double sampling

In Double sampling setup, the existing estimators with their respective bias $Bias(\cdot)_I, Bias(\cdot)_{II}$ and mean squared error $MSE(\cdot)_I$ & $MSE(\cdot)_{II}$ under case I and case II are as below.

(a) *Simple Random sample mean estimator:*

$$\hat{D} = \frac{1}{n} \sum_{i=1}^n d_i \tag{2.1}$$

$$V(\hat{D}) = \bar{D}^2 V_{20} \tag{2.2}$$

where $V(\cdot)$ denotes variance of estimators.

(b) *Usual Ratio Estimator:*

$$\hat{D}_{Rd} = \bar{d} \left(\frac{\bar{a}'}{\bar{a}} \right) \tag{2.3}$$

$$Bias(\hat{D}_{Rd})_I = \bar{D}[(V_{02} - V'_{02}) - (V_{11} - V'_{11})] \tag{2.4}$$

$$Bias(\hat{D}_{Rd})_{II} = \bar{D}[(V_{02} + V'_{02}) - V_{11}] \tag{2.5}$$

$$MSE(\hat{D}_{Rd})_I = \bar{D}^2[V_{20} + (V_{02} - V'_{02}) - 2(V_{11} - V'_{11})] \tag{2.6}$$

$$MSE(\hat{D}_{Rd})_{II} = \bar{D}^2[V_{20} + (V_{02} + V'_{02}) - 2V_{11}] \tag{2.7}$$

(c) *Usual Exponential Ratio Estimator:*

$$\hat{D}_{Exd} = \bar{d} \exp \left(\frac{\bar{a}' - \bar{a}}{\bar{a}' + \bar{a}} \right) \tag{2.8}$$

$$Bias(\hat{D}_{Exd})_I = \bar{D} \left[\frac{3}{8}(V_{02} - V'_{02}) - \frac{1}{2}(V_{11} - V'_{11}) \right] \tag{2.9}$$

$$Bias(\hat{D}_{Exd})_{II} = \bar{D} \left[\frac{1}{8}(3V_{02} - V'_{02}) - \frac{1}{2}V_{11} \right] \tag{2.10}$$

$$MSE(\hat{D}_{Exd})_I = \bar{D}^2 \left[V_{20} + \frac{1}{4}(V_{02} - V'_{02}) - (V_{11} - V'_{11}) \right] \tag{2.11}$$

$$MSE(\hat{D}_{Exd})_{II} = \bar{D}^2 \left[V_{20} + \frac{1}{4}(V_{02} + V'_{02}) - V_{11} \right] \tag{2.12}$$

(d) *Usual Log Ratio Estimator:*

$$\hat{D}_{Lod} = \bar{d} \left[1 + \log \left(\frac{\bar{a}'}{\bar{a}} \right) \right] \tag{2.13}$$

$$Bias(\hat{D}_{Lod})_I = \bar{D} [2(V_{02} - V'_{02}) - (V_{11} - V'_{11})] \tag{2.14}$$

$$Bias(\hat{D}_{Lod})_{II} = \bar{D} [2V_{02} + V'_{02} - V_{11}] \tag{2.15}$$

$$MSE(\hat{D}_{Lod})_I = \bar{D}^2 [V_{20} + (V_{02} - V'_{02}) - 2(V_{11} - V'_{11})] \tag{2.16}$$

$$MSE(\hat{D}_{Lod})_{II} = \bar{D}^2 [V_{20} + (V_{02} + V'_{02}) - 2V_{11}] \tag{2.17}$$

(e) *Usual Regression Estimators:*

$$\hat{D}_{Red} = \bar{d} + \hat{M}(\bar{a}' - \bar{a}) \tag{2.18}$$

$$Bias(\hat{D}_{Red})_I = \bar{D} [(V_{02} - V'_{02}) - (V_{11} - V'_{11})] \tag{2.19}$$

$$Bias(\hat{D}_{Red})_{II} = \bar{D} [V_{02} + V'_{02} - V_{11}] \tag{2.20}$$

$$MSE(\hat{D}_{Red})_I = \bar{Y}^2 [V_{20} + \hat{M}^2(V_{02} - V'_{02}) - 2\hat{M}(V_{11} - V'_{11})] \tag{2.21}$$

$$MSE(\hat{D}_{Red})_{II} = \bar{Y}^2 [V_{20} + \hat{M}^2(V_{02} + V'_{02}) - 2\hat{M}V_{11}] \tag{2.22}$$

where \hat{M} is the regression coefficient.

2.2. Motivation

Estimators suggested in simple random sampling, double sampling, stratified sampling may usual type or exponential type or log-type. Sometime the data may follow the pattern different that of exponential or log-type. It may be a mixture of log and exponential type (Fig4c). This motivates to look for a new combined class of log-cum-exponential type estimators. This paper considers the same in the setup of double sampling. Several authors have suggested estimators

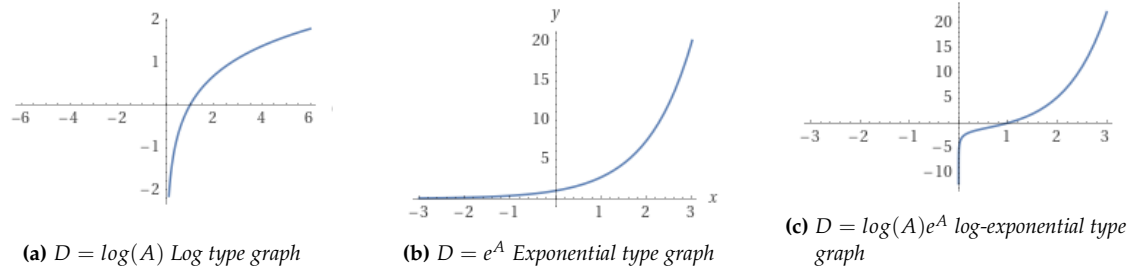


Figure 4: Graphical pattern of relationship

for relationship between D and A variables as shown in Fig(4a) and Fig (4b). But for relationship of type as in Fig (4c) yet needs to be explored. This paper is focused on proposing estimation methodologies with respect to mutual relation shown in fig 4c under the double sampling setup.

3. PROPOSED CLASS OF LOGARITHMIC-EXPONENTIAL TYPE ESTIMATORS

A family of estimators under the double sampling is proposed, to estimate the unknown population mean of the study variable D assuming the presence of auxiliary information A:

$$\hat{D}_{LEd} = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}'}{\bar{a}} \right)^\alpha \right) \left(1 + \log \left(\frac{\bar{a}'}{\bar{a}} \right)^\beta \right) \right\} \right] \tag{3.1}$$

assuming expo-log type relationship between D and A(fig4c), where α, β are constants may positive or negative real numbers.

Theorem 1. The bias of the proposed class of estimator for the sub-sample(Case I) and independent sample(Case II) respectively are:

$$Bias(\hat{D}_{LEd})_I = \alpha \bar{D} ((V_{11} - V'_{11}) - \beta (V_{02} - V'_{02})) \tag{3.2}$$

$$Bias(\hat{D}_{LEd})_{II} = \alpha \bar{D} (V_{11} - \beta (V_{02} + V'_{02})) \tag{3.3}$$

where $Bias(\cdot)_I, Bias(\cdot)_{II}$ are for case I and case II strategies respectively.

Proof. For large sample approximation, define some quantities $\epsilon_0, \epsilon_1, \epsilon_2$ with $|\epsilon_0| < 1, |\epsilon_1| < 1, |\epsilon_2| < 1$ such that

$$\bar{d} = \bar{D}(1 + \epsilon_0), \quad \bar{a} = \bar{A}(1 + \epsilon_1), \quad \bar{a}' = \bar{A}(1 + \epsilon_2)$$

where $\bar{a}' = \frac{1}{n'} (\sum_{i=1}^{n'} a'_i)$ and $(a'_1, a'_2, \dots, a'_n)$ is first phase sample of size n' .

$$E(\epsilon_0) = E(\epsilon_1) = E(\epsilon_2) = 0$$

Moreover,

$$E(\epsilon_0^2) = \left(\frac{1}{n} - \frac{1}{N} \right) C_d^2, \quad E(\epsilon_1^2) = \left(\frac{1}{n} - \frac{1}{N} \right) C_a^2, \quad E(\epsilon_2^2) = \left(\frac{1}{n'} - \frac{1}{N} \right) C_a^2$$

$$E(\epsilon_0\epsilon_1) = \left(\frac{1}{n} - \frac{1}{N}\right) \rho C_d C_a, \quad E(\epsilon_0\epsilon_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) \rho C_d C_a, \quad E(\epsilon_1\epsilon_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) C_a^2$$

General expression for bias for \hat{D}_{LEd} is

$$Bias(\hat{D}_{LEd}) = [E(\hat{D}_{LEd}) - \bar{D}]$$

Under large sampling approximation, upto first order ,

$$\hat{D}_{LEd} = \bar{D}(1 + \epsilon_0) \left[\exp \left\{ (1 - (1 + \epsilon_2)^\alpha (1 + \epsilon_1)^{-\alpha})(1 + \beta \log(1 + \epsilon_2)(1 + \epsilon_1)^{-1}) \right\} \right]$$

Since $|\epsilon_0| < 1, |\epsilon_1| < 1$ and $|\epsilon_2| < 1$, using Taylor series expansion upto the first order approximation, ignoring terms of higher order $(\epsilon_0^i, \epsilon_1^j, \epsilon_2^k)$ for $i > 2, j > 2, k > 2, (i + j + k) > 2$,

$$\hat{D}_{LEd} = \bar{D} \left[1 + \epsilon_0 + \alpha(\epsilon_1 - \epsilon_2) + \alpha\epsilon_0(\epsilon_1 - \epsilon_2) + \beta\alpha(\epsilon_1^2 + \epsilon_2^2 - 2\epsilon_1\epsilon_2) \right]$$

Using expectation $E(\epsilon_0)=E(\epsilon_1)=E(\epsilon_2)=0$, which leads to bias of proposed class of estimator,

$$Bias(\hat{D}_{LEd})_I = \alpha\bar{D}[(V_{11} - V'_{11}) - \beta(V_{02} - V'_{02})] \tag{3.4}$$

$$Bias(\hat{D}_{LEd})_{II} = \alpha\bar{D}[V_{11} - \beta(V_{02} + V'_{02})] \tag{3.5}$$

Since $E(\epsilon_0\epsilon_2)=V'_{11}=0$ for case II because of sample n' being independent to n . ■

Theorem 2. The mean squared error of the proposed class of estimator for the sub-sample(Case I) and independent sample(Case II) respectively are

$$MSE(\hat{D}_{LEd})_I = \bar{D}^2 \left[V_{20} + 2\alpha(V_{11} - V'_{11}) + \alpha^2(V_{02} - V'_{02}) \right] \tag{3.6}$$

$$MSE(\hat{D}_{LEd})_{II} = \bar{D}^2 \left[V_{20} + 2\alpha V_{11} + \alpha^2(V_{02} + V'_{02}) \right] \tag{3.7}$$

Proof. The proposed class in double sampling is,

$$\hat{D}_{LEd} = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}'}{\bar{a}} \right)^\alpha \right) \left(1 + \log \left(\frac{\bar{a}'}{\bar{a}} \right)^\beta \right) \right\} \right]$$

and above in terms of large sample approximation is,

$$\hat{D}_{LEd} = \bar{d} \left[\exp \left\{ (1 - (1 + \epsilon_2)^\alpha (1 + \epsilon_1)^{-\alpha})(1 + \beta \log(1 + \epsilon_2)(1 + \epsilon_1)^{-1}) \right\} \right]$$

Using $|\epsilon_0| < 1, |\epsilon_1| < 1$ and $|\epsilon_2| < 1$ and Taylor series expansion upto the first order of approximation, one can get

$$\hat{D}_{LEd} = \bar{D} [1 + \epsilon_0 + \alpha(\epsilon_1 - \epsilon_2)]$$

by ignoring terms of higher order $(\epsilon_0^i, \epsilon_1^j, \epsilon_2^k)$ for $i > 1, j > 1, k > 1, (i + j + k) > 1, i, j, k=0, 1, 2, \dots$
 Subtracting \bar{D} and squaring both sides one can get,

$$(\hat{D}_{LEd} - \bar{D})^2 = \bar{D}^2 \left[\epsilon_0^2 + 2\alpha\epsilon_0(\epsilon_1 - \epsilon_2) + \alpha^2(\epsilon_1 - \epsilon_2)^2 \right]$$

By taking expectation both sides,

$$E(\hat{D}_{LEd} - \bar{D})^2 = \bar{D}^2 E \left[\epsilon_0^2 + 2\alpha\epsilon_0(\epsilon_1 - \epsilon_2) + \alpha^2(\epsilon_1 - \epsilon_2)^2 \right]$$

So the mean squared error is for Case I and Case II are:

$$MSE(\hat{D}_{LEd})_I = \bar{D}^2 \left[V_{20} + 2\alpha(V_{11} - V'_{11}) + \alpha^2(V_{02} - V'_{02}) \right] \tag{3.8}$$

and

$$MSE(\hat{D}_{LEd})_{II} = \bar{D}^2 [V_{20} + 2\alpha V_{11} + \alpha^2(V_{02} + V'_{02})] \tag{3.9}$$

Since $E(\epsilon_0\epsilon_2) = V'_{11} = 0$ for case II. ■

Remark 1: Gain in precision under case I and case II

$$[MSE(\hat{D}_{LEd})_I - MSE(\hat{D}_{LEd})_{II}] = -2\bar{D}^2(\alpha^2 V'_{02} + \alpha V'_{11}) \tag{3.10}$$

The gain in precision depends on the sign of V'_{11} . In general, case I is better, but if $(\alpha V'_{02} < V_{11})$ then case II of double sampling is better than case I. It provides range when $0 < \alpha < \left(\frac{V_{11}}{V'_{02}}\right)$ then case II is more efficient than case I.

Remark 2: Some particular estimators in the proposed class are in table6:

Table 2: Estimators as member of proposed class.

Estimators	α	β
$\hat{D}_1 = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}}{\bar{a}'} \right) \right) \left(1 + \log \left(\frac{\bar{a}}{\bar{a}'} \right) \right) \right\} \right]$	-1	-1
$\hat{D}_2 = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}}{\bar{a}'} \right) \right) \right\} \right]$	-1	0
$\hat{D}_3 = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}}{\bar{a}'} \right) \right) \left(1 + \log \left(\frac{\bar{a}'}{\bar{a}} \right) \right) \right\} \right]$	-1	1
$\hat{D}_4 = \bar{d}$	0	-1
$\hat{D}_5 = \bar{d}$	0	0
$\hat{D}_6 = \bar{d}$	0	1
$\hat{D}_7 = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}'}{\bar{a}} \right) \right) \left(1 + \log \left(\frac{\bar{a}}{\bar{a}'} \right) \right) \right\} \right]$	1	-1
$\hat{D}_8 = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}'}{\bar{a}} \right) \right) \right\} \right]$	1	0
$\hat{D}_9 = \bar{d} \left[\exp \left\{ \left(1 - \left(\frac{\bar{a}'}{\bar{a}} \right) \right) \left(1 + \log \left(\frac{\bar{a}'}{\bar{a}} \right) \right) \right\} \right]$	1	1

Table 3: Mean Squared Error of Estimators under case I as members of proposed class

Mean Squared Error	α	β
$MSE(\hat{D}_1)_I = \bar{D}^2 [V_{20} - 2(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	-1	-1
$MSE(\hat{D}_2)_I = \bar{D}^2 [V_{20} - 2(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	-1	0
$MSE(\hat{D}_3)_I = \bar{D}^2 [V_{20} - 2(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	-1	1
$V(\hat{D}_4) = \bar{D}^2 V_{20}$	0	-1
$V(\hat{D}_5) = \bar{D}^2 V_{20}$	0	0
$V(\hat{D}_6) = \bar{D}^2 V_{20}$	0	1
$MSE(\hat{D}_7)_I = \bar{D}^2 [V_{20} + 2(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	1	-1
$MSE(\hat{D}_8)_I = \bar{D}^2 [V_{20} + 2(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	1	0
$MSE(\hat{D}_9)_I = \bar{D}^2 [V_{20} + 2(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	1	1

3.1. Optimal sub-class of estimators

Differentiating $MSE(\cdot)$ with respect to α , one can obtain optimum value of α as

Case I

$$\hat{\alpha} = \frac{(V_{11} - V'_{11})}{(V_{02} - V'_{02})} = \left(-\rho \frac{C_d}{C_a} \right) = (-M) \tag{3.11}$$

Table 4: Mean Squared Error of Estimators under case II as members of proposed class

Mean Squared Error	α	β
$MSE(\hat{D}_1)_{II} = \bar{D}^2 [V_{20} - 2V_{11} + (V_{02} + V'_{02})]$	-1	-1
$MSE(\hat{D}_2)_{II} = \bar{D}^2 [V_{20} - 2V_{11} + (V_{02} + V'_{02})]$	-1	0
$MSE(\hat{D}_3)_{II} = \bar{D}^2 [V_{20} - 2V_{11} + (V_{02} + V'_{02})]$	-1	1
$V(\hat{D}_4) = \bar{D}^2 V_{20}$	0	-1
$V(\hat{D}_5) = \bar{D}^2 V_{20}$	0	0
$V(\hat{D}_6) = \bar{D}^2 V_{20}$	0	1
$MSE(\hat{D}_7)_{II} = \bar{D}^2 [V_{20} + 2V_{11} + (V_{02} + V'_{02})]$	1	-1
$MSE(\hat{D}_8)_{II} = \bar{D}^2 [V_{20} + 2V_{11} + (V_{02} + V'_{02})]$	1	0
$MSE(\hat{D}_9)_{II} = \bar{D}^2 [V_{20} + 2V_{11} + (V_{02} + V'_{02})]$	1	1

Table 5: Bias of Estimators under case I as members of proposed class

Bias	α	β
$Bias(\hat{D}_1)_I = -\bar{D} [(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	-1	-1
$Bias(\hat{D}_2)_I = -\bar{D}(V_{11} - V'_{11})$	-1	0
$Bias(\hat{D}_3)_I = -\bar{D} [(V_{11} - V'_{11}) - (V_{02} - V'_{02})]$	-1	1
$Bias(\hat{D}_4) = 0$	0	-1
$Bias(\hat{D}_5) = 0$	0	0
$Bias(\hat{D}_6) = 0$	0	1
$Bias(\hat{D}_7)_I = \bar{D} [(V_{11} - V'_{11}) - (V_{02} - V'_{02})]$	1	-1
$Bias(\hat{D}_8)_I = \bar{D}(V_{11} - V'_{11})$	1	0
$Bias(\hat{D}_9)_I = \bar{D} [(V_{11} - V'_{11}) + (V_{02} - V'_{02})]$	1	1

Table 6: Bias of Estimators under case II as members of proposed class

Bias	α	β
$Bias(\hat{D}_1)_{II} = -\bar{D} [V_{11} + (V_{02} + V'_{02})]$	-1	-1
$Bias(\hat{D}_2)_{II} = -\bar{D} [V_{11}]$	-1	0
$Bias(\hat{D}_3)_{II} = -\bar{D} [V_{11} - (V_{02} + V'_{02})]$	-1	1
$Bias(\hat{D}_4) = 0$	0	-1
$Bias(\hat{D}_5) = 0$	0	0
$Bias(\hat{D}_6) = 0$	0	1
$Bias(\hat{D}_7)_{II} = \bar{D} [V_{11} - (V_{02} + V'_{02})]$	1	-1
$Bias(\hat{D}_8)_{II} = \bar{D} V_{11}$	1	0
$Bias(\hat{D}_9)_{II} = \bar{D} [V_{11} - (V_{02} + V'_{02})]$	1	1

Case II

$$\hat{\alpha} = \left[\frac{V_{11}}{V_{02} + V'_{02}} \right] = - \left[\frac{1}{(1 + \delta)} \left(\rho \frac{C_d}{C_a} \right) \right] = - \left[\frac{M}{(1 + \delta)} \right] \tag{3.12}$$

where, $\delta = \frac{\left(\frac{1}{n'} - \frac{1}{N} \right)}{\left(\frac{1}{n} - \frac{1}{N} \right)}$

The mean squared error under the optimum value of $\alpha = \hat{\alpha}$ [as per (3.8), (3.9)] are

Case I

$$[MSE(\hat{D}_{LEd})_I]_{opt} = \bar{D}^2 C_d^2 \left\{ \left(\frac{1}{n} - \frac{1}{N} \right) - \left(\frac{1}{n} - \frac{1}{n'} \right) \rho^2 \right\} \tag{3.13}$$

Case II

$$[MSE(\hat{D}_{LEd})_{II}]_{opt} = \bar{D}^2 C_d^2 \left\{ \left(\frac{1}{n} - \frac{1}{N} \right) - \left(\frac{1}{n} - \frac{1}{n'} \right) \left(\frac{\rho^2}{1 + \delta} \right) \right\} \tag{3.14}$$

4. COMPARISON WITH EXISTING ESTIMATORS

The existing estimators will be less efficient to the proposed estimators for case I and case II respectively under the following conditions:

(1) *Simple random sample mean estimator* (\bar{d}):

$$\text{Case I: } \alpha \leq \frac{-2(V_{11} - V'_{11})}{(V_{02} - V'_{02})}, \quad \text{Case II: } \alpha \leq \frac{-2V_{11}}{(V_{02} + V'_{02})}$$

(2) *Usual Ratio Estimator* (\hat{D}_{Rd})[eq(2.3)]

$$\text{Case I: } \alpha \leq \left[1 - 2 \left(\rho \frac{C_d}{C_a} \right) \right], \quad \text{Case II: } \alpha \leq \left[1 - \frac{2}{(1 + \delta)} \left(\rho \frac{C_d}{C_a} \right) \right]$$

(3) *Usual Exponential Ratio estimator* (\hat{D}_{Ed})[eq(2.8)]

$$\text{Case I: } \alpha \leq \frac{1}{2} \left[1 - 4\rho \frac{C_d}{C_a} \right], \quad \text{Case II: } \alpha \leq \frac{1}{2} \left[1 - \frac{4}{(1 + \delta)} \left(\rho \frac{C_d}{C_a} \right) \right]$$

(4) *Usual Log Ratio Estimator* (\hat{D}_{Ld})[eq(2.13)]

$$\text{Case I: } \alpha \leq \left[1 - 2 \left(\rho \frac{C_d}{C_a} \right) \right], \quad \text{Case II: } \alpha \leq \left[1 - \frac{2}{(1 + \delta)} \left(\rho \frac{C_d}{C_a} \right) \right]$$

(5) *Usual Regression Estimator* (\hat{D}_{Red})[eq(2.18)]

$$\text{Case I: } \alpha \leq -2 \left(\rho \frac{C_d}{C_a} \right), \quad \text{Case II: } \alpha \leq \left(1 - \frac{2}{(1 + \delta)} \right) \left(\rho \frac{C_d}{C_a} \right)$$

5. RISK FUNCTION AND THE PROPOSED ESTIMATOR

The risk in data collection for dangerous area while implementing a sampling procedure is defined as

- (a) Total Risk
- (b) Per unit respondent contact risk (infection, injury, life risk)
- (c) General risk (area dependent risk)

Risk is associated to various ground conditions like risk in hilly area during data collection, risk of reaching to the household, risk of non-response, risk of dangerous situations, risk of attack on the life of surveyor, risk of epidemic etc.

Let us use symbols for risk as:

r' : Overhead risk

r_0 : Total risk

r_1 : Risk per unit for information collection on variable D and A using second sample n.

r_2 : Risk per unit for first sample for collecting information on auxiliary variable A.

Linear risk function for collecting information is:

$$r_0 = r' + r_1n + r_2n'$$

It is matter of interest to determine the n and n' for a given risk r_0 at the situation when MSE of \hat{D}_{LEd} is minimum. To minimize risk function under risk constraint ϕ and optimum MSE, one can get,

Case I

$$\phi = [MSE(\hat{D}_{LEd})_I]_{opt} + \lambda(r' + r_1n + r_2n' - r_0)$$

where λ is a Lagrange's multiplier. Differentiating with respect to n and n' , equating it to zero, the optimum values of n and n' are

$$n_{opt} = \frac{(r_0 - r')\sqrt{r_1R}}{r_1M_1}, \quad n'_{opt} = \frac{(r_0 - r')\sqrt{-r_2(R - C_d^2)}}{r_2M_1} \tag{5.1}$$

where

$$M_1 = [\sqrt{r_1R} + \sqrt{-r_2(R - C_d^2)}], \quad R = [C_d^2 + 2\alpha C_{da} + \alpha^2 C_a^2]$$

Case II

$$\phi = [MSE(\hat{D}_{LEd})_{II}]_{opt} + \lambda(r' + r_1n + r_2n' - r_0)$$

where λ is a Lagrange's multiplier. Now differentiating with respect to n and n' , equating it to zero, the optimum values of n and n' under case II are

$$n_{opt} = \frac{(r_0 - r')\sqrt{r_1R}}{r_1M_2}, \quad n'_{opt} = \frac{(r_0 - r')\alpha C_a \sqrt{r_2}}{r_2M_2} \tag{5.2}$$

where

$$M_2 = [\sqrt{r_1R} + \sqrt{r_2(\alpha^2 C_a^2)}], \quad R = [C_d^2 + 2\alpha C_{da} + \alpha^2 C_a^2]$$

The ratio of optimal selection of n and n' under fixed risk c_0 is

Case I

$$\left(\frac{n_{opt}}{n'_{opt}}\right) = \frac{r_2(\sqrt{r_1R})}{r_1(\sqrt{-r_2(R - C_d^2)})}$$

Case II

$$\left(\frac{n_{opt}}{n'_{opt}}\right) = \frac{r_2(\sqrt{r_1R})}{r_1\alpha C_a \sqrt{r_2}}$$

6. EMPIRICAL RISK BASED STUDY

Consider a positively correlated population with two variables D and A(Data source -6th Minor Irrigation Census - Village Schedule - Assam)[19] with N=100.

The values of variable D and A are shown in Table 7, where A represents geographical area and D represents the net shown area in hectares.

Table 7: Population Undertaken.

D_i	152	98	75	68	60	295	72	125	16	260
A_i	165	111	80	79	78	319	86	189	26	380
D_i	62	95	210	95	175	180	100	37	87	96
A_i	74	123	220	123	185	197	120	48	105	109
D_i	80	148	85	98	38	95	200	84	18	38
A_i	110	158	121	108	40	110	350	95	28	46
D_i	53	69	30	55	29	75	78	48	81	75
A_i	71	81	45	63	45	89	110	59	95	92
D_i	103	97	82	25	76	70	57	182	55	85
A_i	113	105	96	35	94	81	70	192	65	122
D_i	70	24	190	53	190	158	80	93	176	81
A_i	75	34	200	67	232	169	100	103	186	89

Moreover, population parameters are in the Table 8.

Table 8: Population Parameters

$\bar{D} = 135$	$S_d^2 = 82327$	$C_d^2 = 4.534$	$S_{ad} = 96274.91$
$\bar{A} = 161$	$S_a^2 = 113076.5$	$C_a^2 = 4.356$	$C_{ad} = 4.43$

Table 9: PREs of different estimators with respect to proposed estimator in double sampling

Estimators	PRE	
	Case I	Case II
Simple Random sampling (\hat{D})	22.13%	56.083%
Ratio Estimator (\hat{D}_{Rd})	0.009%	41.493%
Exponential ratio estimator (\hat{D}_{Exd})	6.856%	2.011%
Log ratio Estimator (\hat{D}_{Lod})	0.009%	41.493%

where PRE is Percentage Relative Efficiency defined as:

$$(PRE)_{I,II} = \frac{MSE(T)_{I,II} - (MSE(\hat{D}_{LEd})_{I,II})_{opt}}{MSE(T)} \times 100 \tag{6.1}$$

and T represents estimators like usual ratio, usual expo-ratio, usual log-ratio estimators. It is observed that in case I, at the α_{opt} , the proposed is 22.13% efficient over sample mean estimator, 6.85% better over exponential ratio estimator and same to the usual ratio usual log ratio estimator. Moreover, in case II, at value α_{opt} , the proposed is 56% efficient to sample mean estimator, 41.4% efficient over ratio estimator, 2% efficient over to exponential estimator and 41.4% over log-ratio estimator.

In Figure 5, while general variation of α values, the case I bears lower MSE than case II. But while reaching to α_{opt} , both cases achieve the same MSE level equivalent to that of Regression estimator in double sampling.

Figure 6, reveals the variation of total risk r_0 over the optimum sample sizes (n_{opt} & n'_{opt}). It is observed that increasing fixed risk r_0 leads to larger n'_{opt} (first sample) in comparison to second sample optimum n_{opt} . Low level risk indicates for equal (but small) n and n' to be used by the survey practitioners.

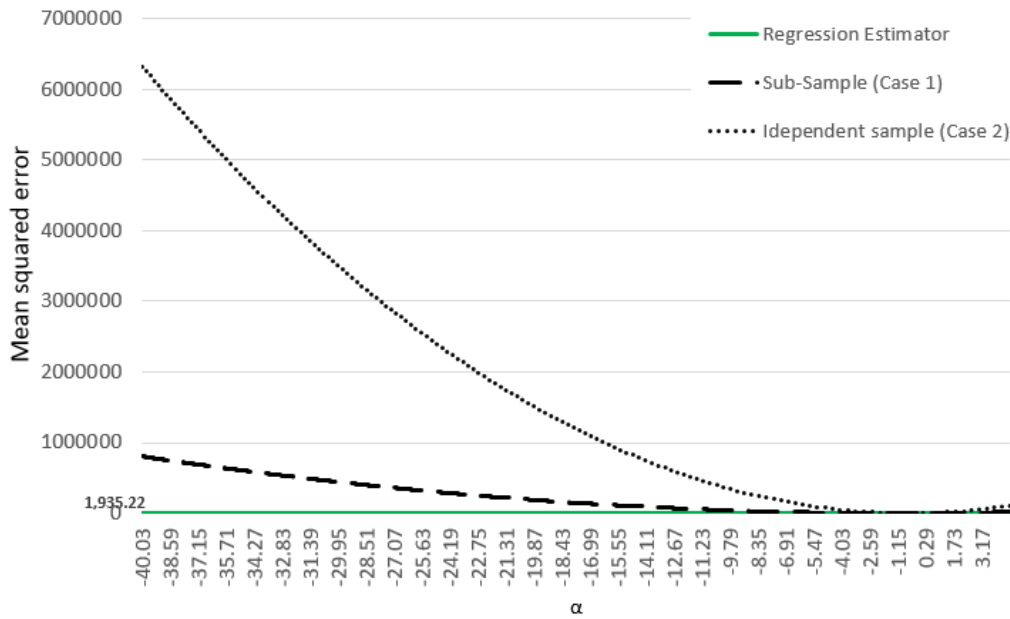


Figure 5: Comparison between MSE's of the proposed class under case I and case II over variation of alpha

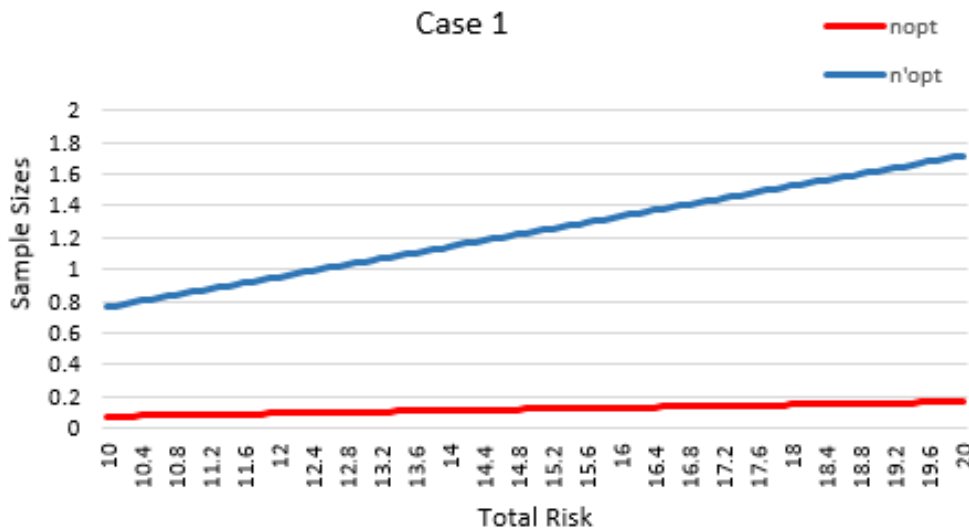


Figure 6: n_{opt} and n'_{opt} for case I over change to total risk r_0

Figure 7, depicts similar pattern among n'_{opt} and n_{opt} while considering variation of total risk r_0 . But interesting is that with the increment in total risk r_0 , the case II needs smaller optimum first phase (preliminary) sample than case I.

The Figure 8, reveals some interesting features of two cases I and II as when ratio $\left(\frac{n_{opt}}{n'_{opt}}\right)$ than case I. This feature confirms that if r_2 increases over fixed r_1 then n_{opt} increases over fixed n'_{opt} . But such increment is high in case II rather than case I.

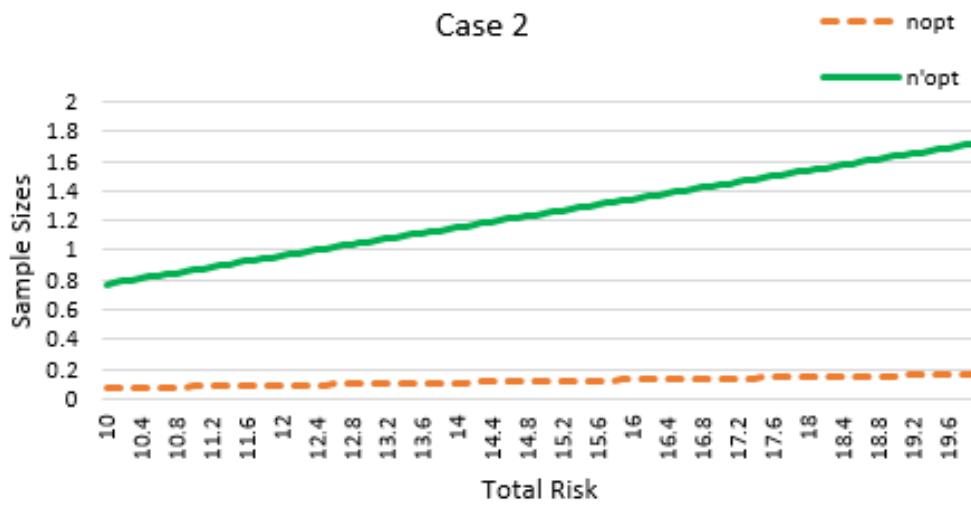


Figure 7: Variation of n_{opt} and n'_{opt} for case II over change to total risk r_0

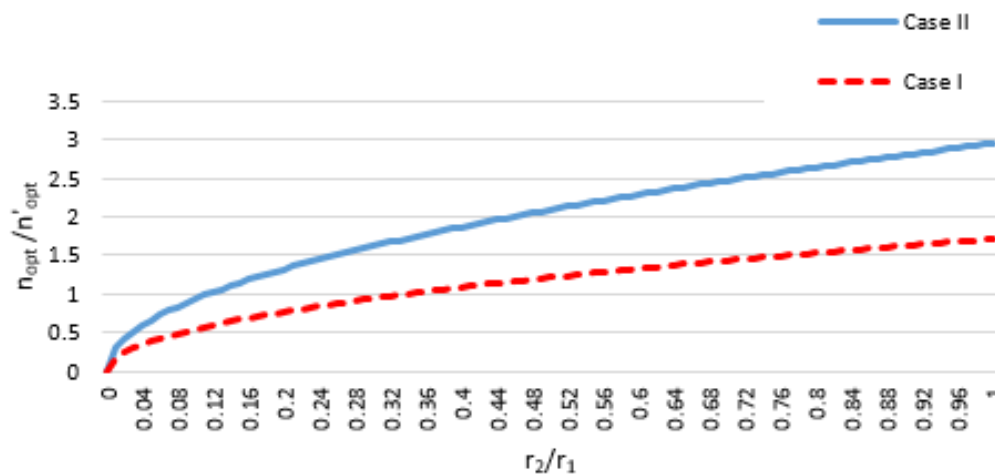


Figure 8: $\left(\frac{n'_{opt}}{n_{opt}}\right)$ with respect to ratio of $\left(\frac{r_2}{r_1}\right)$

7. CONCLUSION

On recapitulation, this paper presents a new class of estimators for estimating the unknown population mean in double sampling in the presence of auxiliary information. Some authors in literature have proposed exponential-type and some others proposed log-type estimators. The suggested estimation procedure is a combo-type class of estimators incorporating both expo and log-type structure. Its properties are discussed and compared in the set up of double sampling, under case I and case II sampling strategies. The proposed is found conditional efficient over usual expo-type and usual log-type estimators (Table 9). Moreover, a linear risk function is used in the paper with three risks parameters r_0, r_1, r_2 and expressions for optimal sample sizes n_{opt} and n'_{opt} are derived. Risk based simulation study reveals that increasing the fixed risk r_0 leads to larger n_{opt} (first sample) in comparison to equal (but small) n and n' to be used by the survey practitioner over incrementing r_0 . Case II needs smaller preliminary sample size in comparison

to case I. While considering variation of optimum ratio of sample sizes (n_{opt}/n'_{opt}) with respect to the risk ratio (r_2/r_1) variation, the case I graph of such ratio constantly lower than the case II, graph indicating lesser need of comparative optimum sample ratio in double sampling using the suggested expo-log estimator at $\alpha = \alpha_{opt}$ choice.

REFERENCES

- [1] Alim A. and Shukla D. (2021). Double sampling based parameter estimation in Big Data and application in Control Charts, *Reliability(RT&M)*, 16(62):72-86.
- [2] Bahl, S. and Tuteja, R.K. (1991). Ratio and product type estimators, *Journal of Information and Optimization Sciences*, 12(1):159-164.
- [3] Bhushan, S. and Gupta, R (2019). Some log-type classes of estimator using auxiliary variable attribute, *Advances in Computational Science and Technology*, 12(2):99-108.
- [4] Bhushan, S. and Kumar, A. (2020). Log-type estimators of population mean in rank set sampling, *Predictive analysis using statistics and Big Data: Concepts and Modeling*, 28:47-74.
- [5] Bhushan, S., Pandey, A. and Shubra, K. (2008). A class of estimators in double sampling using two auxiliary variables, *Journal of Reliability and Statistical studies*, 1(1):67-73.
- [6] Changbao Wu and Ying L. (2003). Optimal calibration estimators under two-phase sampling, *Quality Engineering*, 19(2):119-131.
- [7] Kumari, C. and Thakur, R. K. (2021). An efficient log-type class of estimators using auxiliary information under double sampling, *Journal of Statistics Application and Probability*, 10(1):197-202.
- [8] Muhammad A. and Muhammad H. (2014). On estimation of population mean in the presence of measurement error and non-response, *Pak. J. Statist.*, 31(5):657-670.
- [9] Sahoo J., Sahoo L.N. and Mohanty S. (1993). A regression approach for estimation in two-phase sampling using auxiliary variables, *Current Science*, 65:73-75.
- [10] Sanaullah, A., Ali, H. A., Amits, Muhammad Noor ul and Hanif, M. (2014). Generalised exponential chain ratio estimators under stratified two-phase random sampling, *Applied Mathematics and Computation*, 226: 541-547.
- [11] Sanaullah, A., Amin, Muhammad Noor-ul and Hanif, M. (2015). Generalized exponential-type ratio-cum-ratio and product-cum-product estimators for population mean in presence of non-response under stratified two-phase random sampling, *Pakistan Journal of Statistics*, 31(1):71-94.
- [12] Sanaullah, A., Hanif, M. and Asghar, A. (2016). Generalized exponential estimators for population variance under two-phase sampling, *International Journal of Applied Computation and Mathematics*, 2:75-84.
- [13] Shabbir, J., Ahmed S., Sanaullah, A. and Onyange, R. (2021). Measuring performance of ratio-exponential-log type general class of estimators using two auxiliary variables, *Mathematical Problems in Engineering*, 2021(3):1-12.
- [14] Shabbir, J. and Sat G. (2011). On estimating finite population mean in simple and stratified random sampling, *Communication in Statistics(Theory and Methods)*, 40:199-212.
- [15] Shabbir, J., Sat G. and Masood, S. (2022). An improved class of estimators for finite population mean in simple random sampling, *Communication in Statistics, Theory and Method*, 51(11):3508-3520.
- [16] Shukla, D. (2002). FT estimator under two phase sampling, *Metron (International Journal of Statistics)*, 60(1-2):97-106.
- [17] Zahoor, A., Iqra M. and Muhammad H. (2014). Regression estimator in two phase sampling using Multi-auxiliary information in the presence of non-response and measurement error at second phase, *Journal of Applied Probability and Statistics*, 9(2):41-5.
- [18] Zaman, T. and Kadilar, C. (2021). Exponential ratio and product type estimators of mean in stratified two-phase sampling, *AIMS Mathematics*, 6(5):4265–4279.
- [19] <https://data.gov.in/> Data source -6th Minor Irrigation Census - Village Schedule - Assam

STOCHASTIC ANALYSIS OF A GAS TURBINE SYSTEM WITH PRIORITY AND RANDOM INSPECTION BY SINGLE SERVER UNDER DIFFERENT HUMID CONDITIONS

Pinki, Vijeta Kumari and Dalip Singh

•

Department of Mathematics, Maharshi Dayanand University Rohtak, India
pinkichahar95@gmail.com, vijetadahya1993@gmail.com and dsmdur@gmail.com

Abstract

In this study, we investigated the impact of two different humid levels on the reliability measures of a stochastic model for a gas turbine system composed of a gas turbine and a steam turbine. To enhance the system's overall performance, we prioritize gas turbine repair over steam turbine repair in addition to a combined inspection and preventative maintenance approach. To find some reliability measures, such as the mean time to system failure, availability, etc., semi-Markov process and regenerating point technique are utilized. These measures are analysed graphically based on the data obtained from a gas turbine power plant in Delhi, India.

Keywords: Gas Turbine, Steam Turbine; Reliability; Cost-Benefit; Maintenance;

1. Introduction

The growing global demand for electrical energy, driven by factors such as rapid industrialization, urbanization and the increasing number of electronic gadgets has put significant pressure on our existing power generation systems. Researchers from all around the world have conducted substantial research into the complex dynamics of energy supply, demand, and security in both developed and emerging economies [1-3]. The nation of India, which has to cope with its own unique issues and potential is the focus of in-depth research on electricity demand [4]. The study made by Zhang et al. goes beyond typical clustering algorithms to investigate novel approaches to analyse electricity usage trends [5]. Optimizing demand response initiatives within smart grids is studied by Derakhshan et al. using TBLO and SFL algorithms [6]. Furthermore, [7] points out the impact of these demand patterns on both power system costs and supply sufficiency. These research articles provide a comprehensive understanding of the complicated relationship between power demand and supply, as well as the issues created by shifting consumption patterns.

To address these issues, researchers and policymakers need to design and improve power generation systems that can satisfy the rising energy demands profitably and ensuring sustainability. Though, there are various ways to generate electricity, such as using water, sunlight, thermal energy from sources like coal, and harnessing nuclear reactions, leading to different types of power plants. In today's competitive energy markets, a new approach called the "Risk-Based Approach" is gaining attention for managing Virtual Power Plants. This approach is all about figuring out smart and efficient ways to schedule the activities of these

virtual power plants [8]. Operating power generating systems in demanding environments indeed presents several challenges [9]. The performance of a system inevitably deteriorates when operated over lengthy periods of time under adverse conditions. When this degradation exceeds a certain threshold, it may lead components or subsystems to fail, compromising the overall safety of the system. As a result, one of the key objectives of engineering systems is to provide timely maintenance. Preventive maintenance and corrective maintenance are two essential approaches for maintaining and managing equipment and systems in various industries [10]. The primary objective of any maintenance strategy is to uphold the system functionality to the greatest extent possible while striking a balance between downtime and maintenance expenses, thereby avoiding catastrophic breakdowns. Zaho et al. studied the preventive maintenance scheduling on gas turbine power plant through a sequential approach [11].

The key objective of this research paper is to develop a comprehensive operational stochastic model for a combined cycle power plant. Gas turbines, a crucial part of combined cycle power plant, are responsible for this decision since they have amazing qualities including great efficiency, adaptability, and quick start-up times. The significance of our research lies in addressing a previously unexplored aspect of the literature. Although there is a lot of information on the reliability of combined cycle power plants, none of the existing studies have considered the impact of humidity with priority and random inspection within a stochastic model using the semi-Markov approach. Recognizing the importance of humidity in power plant management, we have addressed a research gap by incorporating this critical aspect into our analysis.

Statistical methods play a fundamental role in the development of reliability/stochastic models, offering valuable insights that can inform maintenance and repair planning for technical systems. These reliability criteria are used to measure the system's potential for maintenance and repair. Table 1 provides a comprehensive summary of the foundational statistical techniques employed during the creation of reliability/stochastic models for various gas turbine and combined cycle power plants within the domains of the energy sector in recent years. Many researchers in Table 1 studied the reliability models for gas turbine systems under different conditions using different methods but none of the existing studies have considered the impact of humidity with priority and random inspection within a stochastic model using the semi-Markov approach. Reliability models assist in figuring out how reliable a system is, how frequently it may fail, and how quickly it may recover from those failures. Creating such models for gas turbine power plants allows engineers and researchers to foresee future faults, develop effective maintenance procedures, and maximize the overall performance and operating efficiency of these systems.

This study aims to conduct a thorough investigation into the effects of two distinct humidity levels (i.e., humidity less than or equal to 50% and humidity greater than 50%) on the reliability measures of a stochastic model for a gas turbine system. The system comprises a gas turbine and a steam turbine and has been developed under specific assumptions that have not been addressed in the existing literature. Through a comprehensive investigation into the impact of humidity variations on the reliability measures of gas turbine systems, we aim to deepen our understanding and provide invaluable insights in this field. Thus it will provide a comprehensive analysis of our research methodology, the experimental setup, data collection, and, ultimately, the results and implications of our findings. By doing so, we aim to promote the efficient and reliable utilization of gas turbine systems, thus furthering the cause of sustainable energy generation and contributing to the broader goals of the industry.

Table 1: Summary of literature review in recent years

Methods	Model Structure	Characteristic of Study	References
Statistical Methods	Mathematical modeling	Impact of ambient conditions on CCPP in Syria	[12]
	AI-coherent modeling	Data-driven forecasting for a CCPP using BFGS algorithm	[13]
	Thermodynamic modelling	Effect of temperature and relative humidity on gas turbine	[14]
	Mathematical modeling	Effects of the intake air humidity on the gas turbine	[15]
Monte Carlo and MLE method	Stochastic modeling	CCPP study under temperature fluctuations in Tehran	[16]
PJM method	Four-state reliability model	Combined heat and power plants	[17]
Semi-Markov Process and RGT	Reliability modeling	CCPP with schedule inspection	[18]
	Reliability modeling	CCPP with random inspection	[19]
	Three-unit CCPP reliability modeling	Effect of ambient temperature with FCFS repair pattern on CCPP	[20]
	Reliability modeling	Effect of humidity on CCPP in Delhi	[21]

2. Modeling of System

In this paper, we discuss the impact of two different humid levels (i.e., humidity less than or equal to 50% and humidity greater than 50%) on the reliability measures of a stochastic model for a gas turbine system composed of a gas turbine and a steam turbine. To enhance the system's overall performance, we prioritize gas turbine repair over steam turbine repair in addition to a combined inspection and preventative maintenance approach as shown in Figure 1. To find some reliability measures, such as the mean time to system failure, availability, etc., we employ the semi-Markov process and regenerating point technique, which are well-suited for this type of analysis. At initial stage, both units, the gas turbine and the steam turbine are up and completely operational, operating together in a combined cycle. Steam turbine failure keeps the system in upstate mode with partially working and termed as single cycle. However, if the gas turbine fails, the system transitions to a downstate mode.

The following reasonable assumptions are used to create the model:

- The failure time distribution is presumed to be exponential, whereas the repair/maintenance time distribution is arbitrary.
- After each maintenance/repair activity, the unit is stated to be as satisfactory as new.
- The system's repair sequence adheres to a first come, first serve basis, except in cases of complete system failure, where priority is given to gas turbine repair over steam turbine repair.
- System failure is asserted when both units fail.

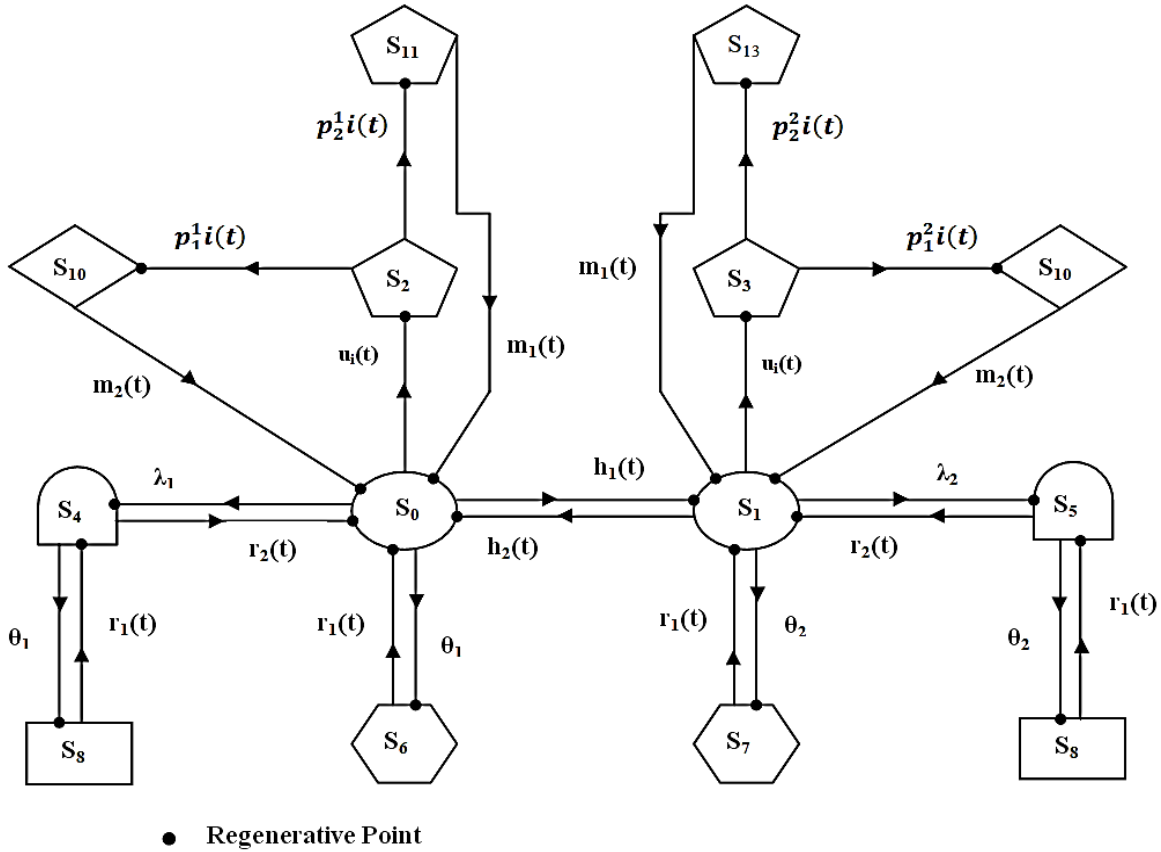


Figure 1: State Transition Diagram of the System

Description of the states in Figure 1:

S_0/S_1 : Both units are operational when humidity is $\leq 50\%$ / $> 50\%$.

S_2/S_3 : System is down due to inspection when humidity is $\leq 50\%$ / $> 50\%$.

S_4/S_5 : System is operational with the gas turbine running and the steam turbine under repair when humidity is $\leq 50\%$ / $> 50\%$.

S_6/S_7 : System is down with the gas turbine under repair and the steam turbine also down when humidity is $\leq 50\%$ / $> 50\%$.

S_8/S_9 : System has failed, with the gas turbine under repair and the steam turbine awaiting repair when humidity is $\leq 50\%$ / $> 50\%$.

S_{10}/S_{11} : System is operational with the gas turbine running and the steam turbine under maintenance when humidity is $\leq 50\%$ / $> 50\%$.

S_{12}/S_{13} : System is down with the gas turbine under repair and the steam turbine also down when humidity is $\leq 50\%$ / $> 50\%$.

2.1 Notations

θ_1/θ_2 : rate of gas turbine failure when humidity is $\leq/\geq 50\%$.

λ_1/λ_2 : rate of steam turbine failure when humidity is $\leq/\geq 50\%$.

$BH_1^1(t)/BH_1^2(t)$: server is busy at a particular time t when the humidity is $\leq/\geq 50\%$.

$DH_1^1(t)/DH_1^2(t)$: system is in a down state at specific time t when the humidity is $\leq/\geq 50\%$.

$h_1(t)/H_1(t)$: pdf/cdf of time changing humidity from $\leq 50\%$ to $> 50\%$.

- $h_2(t)/H_2(t)$: pdf/cdf of time changing humidity from $> 50\%$ to $\leq 50\%$.
- $i(t)/I(t)$: pdf/cdf of the examination to identify the type of maintenance required.
- $IH_1^1(t)/IH_1^2(t)$: system is under inspection at a particular time t when the humidity is $\leq/\geq 50\%$.
- $m_1(t)/M_1(t)$: pdf/cdf of gas turbine maintenance.
- $m_2(t)/M_2(t)$: pdf/cdf of steam turbine maintenance.
- p_1^1/p_1^2 : probability that an inspection will indicate the need for gas turbine maintenance when humidity is $\leq/\geq 50\%$.
- p_2^1/p_2^2 : probability that an inspection will indicate the need for steam turbine maintenance when humidity is $\leq/\geq 50\%$.
- $q_{ij}(t)/Q_{ij}(t)$: pdf/cdf of the first-passage time without visiting any other regenerative state from regenerative state i to a regenerative state j or a failed state j in $(0, t)$.
- $q_{ij}^{(k)}/Q_{ij}^{(k)}(t)$: pdf/cdf of first-passage time from regenerative state i to a regenerative state j , visiting state k one time in $(0, t]$
- $r_1(t)/R_1(t)$: time for gas turbine repair in pdf/cdf respectively.
- $r_2(t)/R_2(t)$: time for steam turbine repair in pdf/cdf respectively.
- $u_i(t)/U_i(t)$: time required to inspect the system in pdf/cdf.
- $VH_1^1(t)/VH_1^2(t)$: server's expected number of visits when humidity is $\leq/\geq 50\%$.
- \otimes/\oplus : Laplace convolution/ Laplace Stieltjes convolution

3. State Transition Probabilities and Mean Sojourn Time

The expression $dQ_{ij}(t)$ for all essential combinations of i and j is generated based on state transition diagram and the transition probabilities p_{ij} are computed by applying Laplace transform and utilizing $p_{ij} = \lim_{s \rightarrow 0} q_{ij}^*(s)$.

Table 2: State Transition Probabilities

$dQ_{01} = e^{-(\theta_1+\lambda_1)t}F_1$	$dQ_{02} = e^{-(\theta_1+\lambda_1)t}F_2$	$dQ_{04} = \lambda_1 e^{-(\theta_1+\lambda_1)t}F_4$
$dQ_{06} = \theta_1 e^{-(\theta_1+\lambda_1)t}F_4$	$dQ_{10} = e^{-(\theta_2+\lambda_2)t}F_0$	$dQ_{13} = e^{-(\theta_2+\lambda_2)t}F_3,$
$dQ_{15} = \lambda_2 e^{-(\theta_2+\lambda_2)t}F_5$	$dQ_{17} = \theta_2 e^{-(\theta_2+\lambda_2)t}F_5$	$dQ_{2,10} = p_1^1(t)$
$dQ_{2,11} = p_2^1(t)$	$dQ_{3,12} = p_1^2(t)$	$dQ_{3,13} = p_2^2(t)$
$dQ_{40} = e^{-\theta_1(t)}r_2(t)$	$dQ_{48} = \theta_1 e^{-\theta_1(t)}\overline{R_2(t)}$	$dQ_{51} = e^{-\theta_2(t)}r_2(t)$
$dQ_{59} = \theta_2 e^{-\theta_2(t)}\overline{R_2(t)}$	$dQ_{60} = r_1(t)$	$dQ_{71} = r_1(t)$
$dQ_{84} = r_1(t)$	$dQ_{95} = r_1(t)$	$dQ_{10,0} = m_2(t)$
$dQ_{11,0} = m_1(t)$	$dQ_{12,1} = m_2(t)$	$dQ_{13,1} = m_1(t)$

where, $F_0 = h_2(t)\overline{U_i(t)}$, $F_1 = h_1(t)\overline{U_i(t)}$, $F_2 = u_i(t)\overline{H_1(t)}$, $F_3 = u_i(t)\overline{H_2(t)}$
 $F_4 = \overline{H_1(t)U_i(t)}$, $F_5 = \overline{H_2(t)U_i(t)}$

Mean Sojourn Time (μ_i) is the time the system expects to spend in state i . The expressions for μ_i are produced by using $\mu_i = \int_0^\infty P[T_i > t] dt$ where T_i denotes the system's stay time in state i .

Table 3: Mean Sojourn Time		
$\mu_0 = F_4^*(\theta_1 + \lambda_1)$	$\mu_1 = F_5^*(\theta_2 + \lambda_2)$	$\mu_2 = \int_0^\infty \overline{I(t)} dt = \mu_3$
$\mu_4 = \frac{1}{\theta_1} [1 - r_2^*(\theta_1)]$	$\mu_5 = \frac{1}{\theta_2} [1 - r_2^*(\theta_2)]$	$\mu_6 = \int_0^\infty \overline{R_1(t)} dt = \mu_7$
$\mu_{10} = \int_0^\infty \overline{M_2(t)} dt = \mu_{12}$	$\mu_{11} = \int_0^\infty \overline{M_1(t)} dt = \mu_{13}$	

4. Reliability Measures

4.1 Mean Time to System Failure (MTSF)

Assuming that $\phi_i(t)$ represents the cumulative distribution function of the initial-passage time from a failed state to a regenerative state i . The recursive relations listed below are employed to compute the system's mean time to failure.

$$\phi_0(t) = Q_{01}(t) \otimes \phi_1(t) + Q_{02}(t) \otimes \phi_2(t) + Q_{04}(t) \otimes \phi_4(t) + Q_{06}(t) \otimes \phi_6(t) \quad (1)$$

$$\phi_1(t) = Q_{10}(t) \otimes \phi_0(t) + Q_{13}(t) \otimes \phi_3(t) + Q_{15}(t) \otimes \phi_5(t) + Q_{17}(t) \otimes \phi_7(t) \quad (2)$$

$$\phi_2(t) = Q_{2,10}(t) \otimes \phi_{10}(t) + Q_{2,11}(t) \otimes \phi_{11}(t) \quad (3)$$

$$\phi_3(t) = Q_{3,12}(t) \otimes \phi_{12}(t) + Q_{3,13}(t) \otimes \phi_{13}(t) \quad (4)$$

$$\phi_4(t) = Q_{40}(t) \otimes \phi_0(t) + Q_{48}(t) \quad (5)$$

$$\phi_5(t) = Q_{51}(t) \otimes \phi_1(t) + Q_{59}(t) \quad (6)$$

$$\phi_6(t) = Q_{60}(t) \otimes \phi_0(t) \quad (7)$$

$$\phi_7(t) = Q_{71}(t) \otimes \phi_1(t) \quad (8)$$

$$\phi_{10}(t) = Q_{10,0}(t) \otimes \phi_0(t) \quad (9)$$

$$\phi_{11}(t) = Q_{11,0}(t) \otimes \phi_0(t) \quad (10)$$

$$\phi_{12}(t) = Q_{12,1}(t) \otimes \phi_1(t) \quad (11)$$

$$\phi_{13}(t) = Q_{13,1}(t) \otimes \phi_1(t) \quad (12)$$

Using Laplace Stieltjes Transform on both sides of aforementioned relations and Cramer's Rule to solve them, we get

$$MTSF = \lim_{s \rightarrow 0} \frac{1 - \phi_0^*(s)}{s} = \frac{N}{D} \quad (13)$$

$$\text{where, } N = (p_{10} + p_{15}p_{59})(\mu_0 + p_{02}\mu_2 + p_{04}\mu_4) + p_{01}p_{13}\mu_3 + p_{01}\mu_1 + p_{15}\mu_5(p_{01} + p_{04}p_{48})$$

$$D = p_{15}p_{59}(p_{01} + p_{04}p_{48}) + p_{04}p_{10}p_{48}$$

4.2 Steady State Availability

$AH_i^1(t)/AH_i^{1s}(t)$ and $AH_i^2(t)/AH_i^{2s}(t)$ indicates how likely it is that the system will be in a combined cycle or single cycle at any given time t , assuming that it was in a regenerative condition at time $t=0$ when the humidity is \leq and $> 50\%$. We obtain the equations for availability in both combined and single cycles by studying empirical argumentation and solving the resulting equations by using the Laplace Transform, we get

$$V_1 = \begin{bmatrix} 1 & -p_{01} & -p_{02} & 0 & -p_{04} & 0 & -p_{06} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -p_{10} & 1 & 0 & -p_{13} & 0 & -p_{15} & 0 & -p_{17} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -p_{2,10} & -p_{2,11} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -p_{3,12} & -p_{3,13} \\ -p_{40} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -p_{48} & 0 & 0 & 0 & 0 & 0 \\ 0 & -p_{51} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -p_{59} & 0 & 0 & 0 & 0 \\ -p_{60} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -p_{71} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -p_{84} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -p_{95} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -p_{10,0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -p_{11,0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -p_{12,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -p_{13,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Or $V_1 = |C_1 \ C_2 \ C_3 \ C_4 \ C_5 \ C_6 \ C_7 \ C_8 \ C_9 \ C_{10} \ C_{11} \ C_{12} \ C_{13} \ C_{14}|$
 where C_i ($1 \leq i \leq 14$), represents the i^{th} column of the V_1 .

$$U_1 = |C_1^{n1} \ C_2 \ C_3 \ C_4 \ C_5 \ C_6 \ C_7 \ C_8 \ C_9 \ C_{10} \ C_{11} \ C_{12} \ C_{13} \ C_{14}|$$

$$C_1^{n1} = [\mu_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

$$AH_0^1(t) = \frac{U_1}{V}, \text{ where } V = V_1'$$

$$\text{Similarly, } AH_0^2(t) = \frac{U_2}{V}, \quad AH_0^{1s}(t) = \frac{U_3}{V}, \quad AH_0^{2s}(t) = \frac{U_4}{V} \tag{14}$$

where, $U_j = |C_1^{nj} \ C_2 \ C_3 \ C_4 \ C_5 \ C_6 \ C_7 \ C_8 \ C_9 \ C_{10} \ C_{11} \ C_{12} \ C_{13} \ C_{14}|$
 and ($2 \leq j \leq 4$)

$$C_1^{n2} = [0 \ \mu_1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

$$C_1^{n3} = [0 \ 0 \ 0 \ 0 \ \mu_4 \ 0 \ 0 \ 0 \ 0 \ 0 \ \mu_{10} \ 0 \ 0 \ 0]$$

$$C_1^{n4} = [0 \ 0 \ 0 \ 0 \ 0 \ \mu_5 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \mu_{10} \ 0]$$

4.3 Other Performance Measures

$$DH_0^1(t) = \frac{U_5}{V}, \quad DH_0^2(t) = \frac{U_6}{V}, \quad IH_0^1(t) = \frac{U_7}{V}, \quad IH_0^2(t) = \frac{U_8}{V}, \quad BH_0^1(t) = \frac{U_9}{V}, \quad BH_0^2(t) = \frac{U_{10}}{V}$$

$$VH_0^1(t) = \frac{U_{11}}{V}, \quad VH_0^2(t) = \frac{U_{12}}{V} \tag{15}$$

where, ($5 \leq j \leq 12$)

$$U_j = |C_1^{nj} \ C_2 \ C_3 \ C_4 \ C_5 \ C_6 \ C_7 \ C_8 \ C_9 \ C_{10} \ C_{11} \ C_{12} \ C_{13} \ C_{14}|$$

$$C_1^{n5} = [0 \ 0 \ \mu_2 \ 0 \ 0 \ 0 \ \mu_6 \ 0 \ 0 \ 0 \ 0 \ \mu_{11} \ 0 \ 0]$$

$$C_1^{n6} = [0 \ 0 \ 0 \ \mu_3 \ 0 \ 0 \ 0 \ \mu_6 \ 0 \ 0 \ 0 \ 0 \ \mu_{11}]$$

$$C_1^{n7} = [0 \ 0 \ \mu_2 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

$$C_1^{n8} = [0 \ 0 \ 0 \ \mu_2 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

$$C_1^{n9} = [0 \ 0 \ 0 \ 0 \ \mu_4 \ 0 \ \mu_6 \ 0 \ 0 \ 0 \ \mu_{10} \ \mu_{11} \ 0 \ 0]$$

$$C_1^{n10} = [0 \ 0 \ 0 \ 0 \ 0 \ \mu_5 \ 0 \ \mu_6 \ 0 \ 0 \ 0 \ 0 \ \mu_{10} \ \mu_{11}]$$

$$C_1^{n11} = [p_{02} + p_{04} + p_{06} \ 0 \ 0 \ \mu_2 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

$$C_1^{n12} = [0 \ p_{13} + p_{15} + p_{17} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

4.4 Profit of the System

$$P = C_1 * AH_0^1 + C_2 * AH_0^2 + C_3 * AH_0^{1s} + C_4 * AH_0^{2s} - C_5 * BH_0^1 - C_6 * BH_0^2 - C_7 * VH_0^1 - C_8 * VH_0^2 - PE$$

C_1/C_2 : Revenue earned per unit when system works in combined cycle for humidity $\leq/ > 50\%$

C_3/C_4 : Revenue earned per unit when system works in single cycle for humidity $\leq/ > 50\%$

C_5/C_6 : Expense per unit time when server is busy for humidity $\leq/ > 50\%$

C_7/C_8 : Cost per visit by server when humidity is $\leq/ > 50\%$

PE : Additional expenses of Plant

5. Results and Discussion

For numerical calculations, we study the specific situation, in which all temporal distributions are assumed to be exponential which are best fitted over real time data, as established by Singh [22]. We examined one-year real-time data from a gas turbine power plant in Delhi, India, restricted the temperature range to up to 25°C, and the methodology used to obtain the values of all the parameters is provided in appendix, which are used to assess the graphical behaviour of reliability measures. The following distributions have been assumed for various times.

$$i(t) = \gamma e^{-\gamma(t)}, r_1(t) = \alpha_1 e^{-\alpha_1(t)}, r_2(t) = \alpha_2 e^{-\alpha_2(t)}, m_1(t) = \alpha e^{-\alpha(t)}, m_2(t) = \beta e^{-\beta(t)}$$

$$h_1(t) = \beta_1 e^{-\beta_1(t)}, h_2(t) = \beta_2 e^{-\beta_2(t)}, u_i(t) = \theta e^{-\theta(t)}$$

5.1 MTSF V/s Failure rate λ_1 for different values of θ_2

Figure 2 illustrate the behaviour of mean time to system failure v/s failure rate of steam turbine λ_1 for different values of θ_2 . MTSF decreases with increase in any one of the failure rate $\lambda_1, \theta_1, \lambda_2$ and θ_2 .

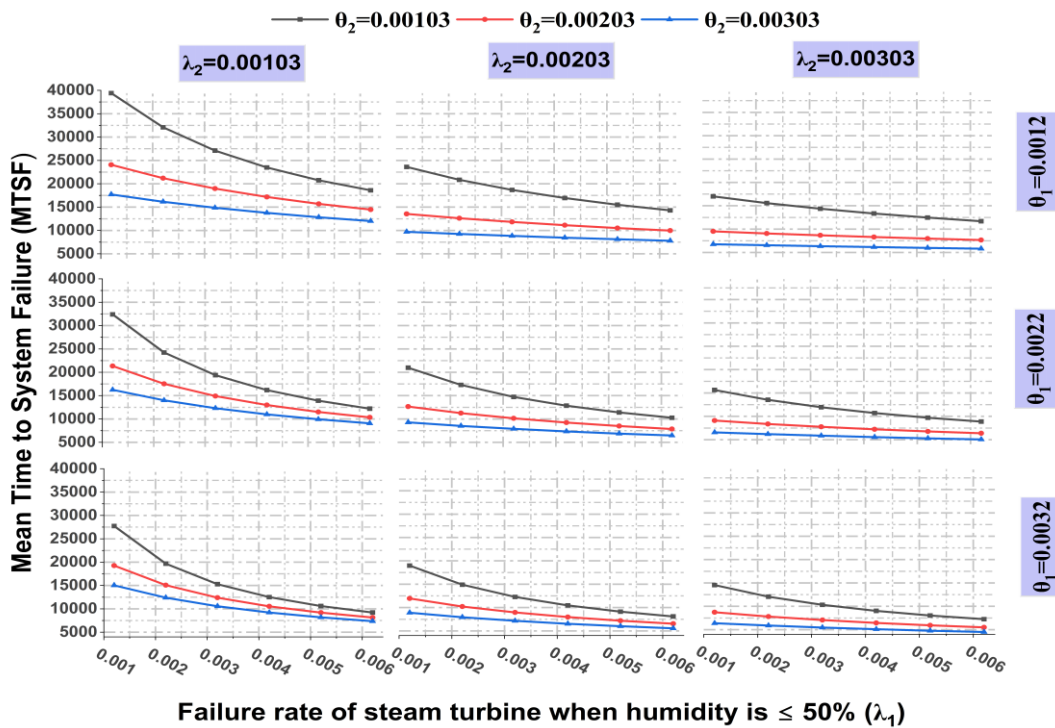


Figure 2: MTSF Vs Failure rates $\theta_1, \theta_2, \lambda_1, \lambda_2$

5.2 Availability in Steady State

Figure 3 demonstrates the availability in combined cycle when humidity is $\leq 50\%$ and when humidity is $> 50\%$

- Both availabilities (when humidity is $\leq / > 50\%$) of combined cycle decreases as we increase any one of the failure rates.
- Availability in combined cycle when humidity is $> 50\%$ is higher than availability in combined cycle when humidity is $\leq 50\%$.

Figure 4 demonstrates the availability in single cycle when humidity is $\leq 50\%$ and when humidity is $> 50\%$

- Availability when humidity is $\leq 50\% / > 50\%$ of single cycle increases with increase in failure rate λ_1 / λ_2 respectively.

- Availability in single cycle (when humidity is > 50%) decreases smoothly with increase in failure rate λ_1 .

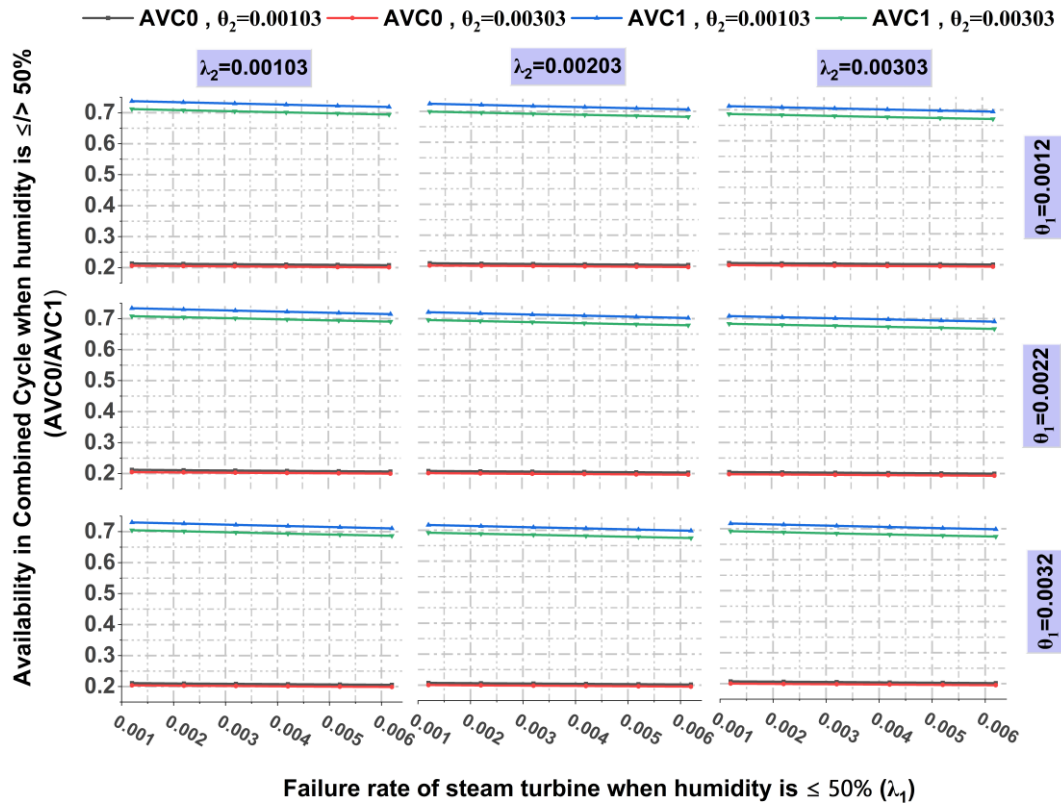


Figure 3: Availability in Combined Cycle Vs Failure rates $\theta_1, \theta_2, \lambda_1$ and λ_2

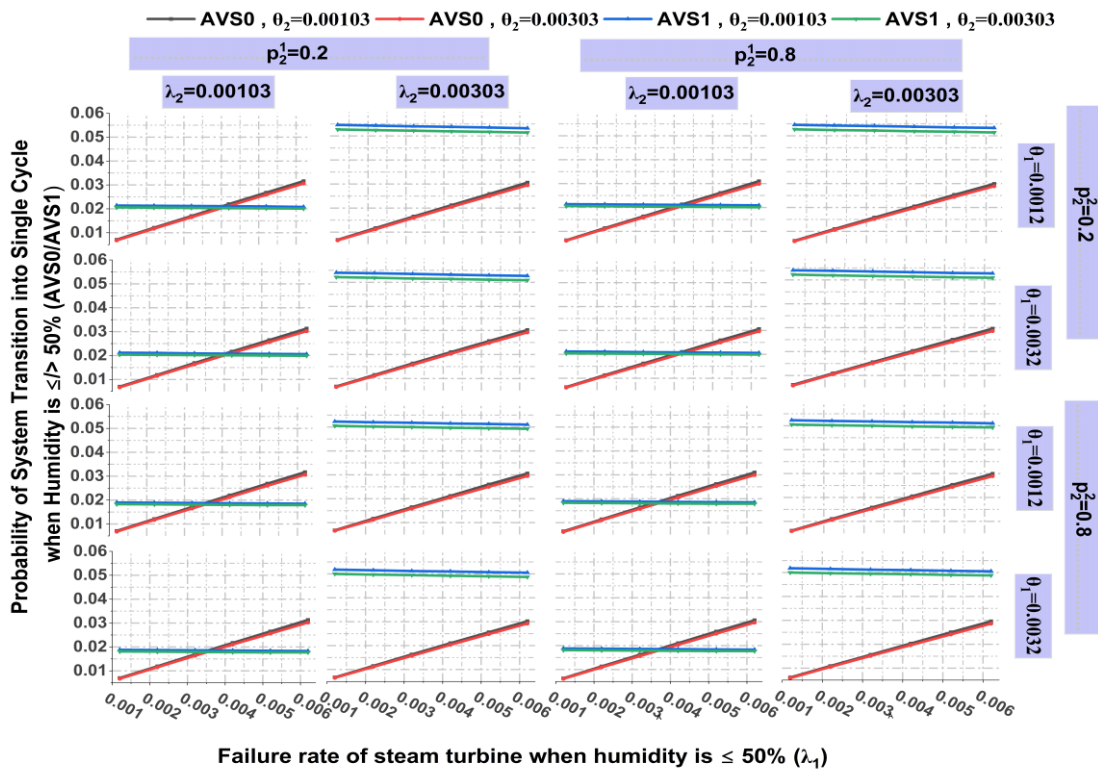


Figure 4: Availability in Single Cycle Vs Failure rates $\theta_1, \theta_2, \lambda_1$ and λ_2

5.3 Profit V/s Plant Expenses (PE) for different values of Price of Electricity (P)

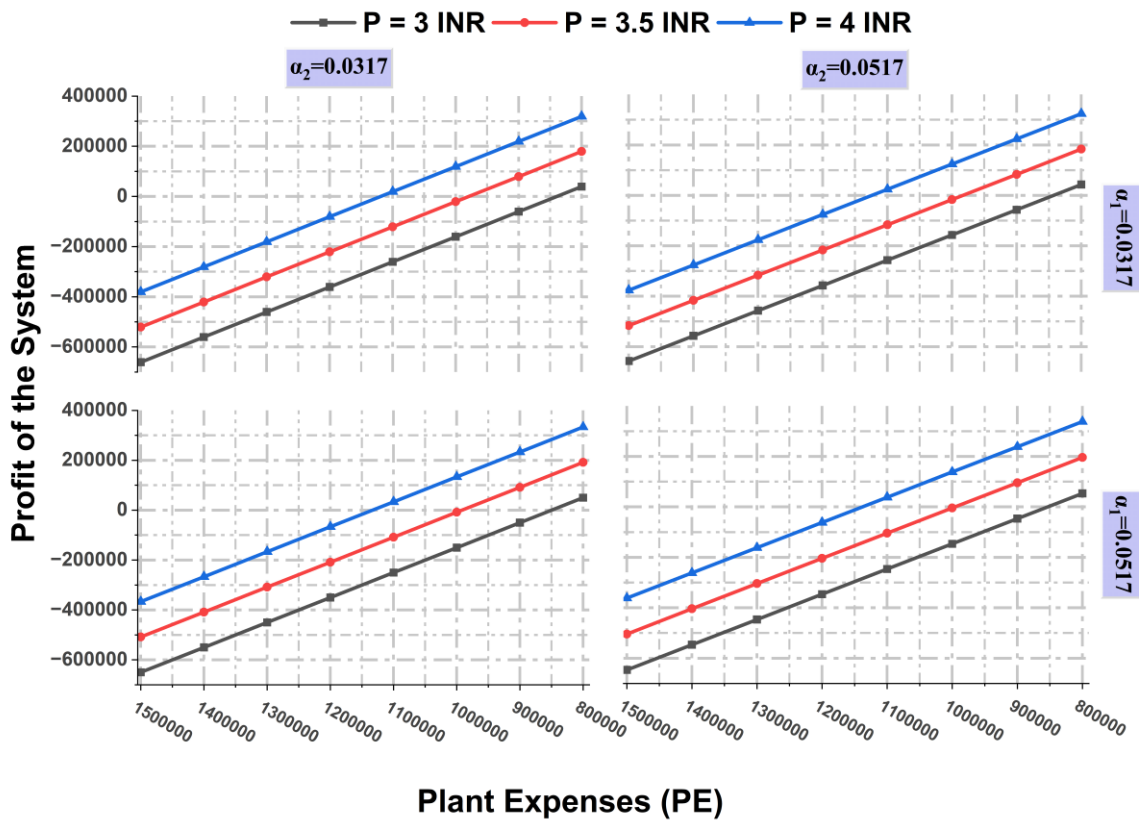


Figure 5: Profit Vs Plant Expenses for different P Values

Figure 5 illustrates the behavior of Profit of plant with respect to Plant Expenses.

- Profit increases with decrease in plant expenses.
- Profit increases with increase in values of P.

Table 4: Cut-off Values of PE for different values of α_1 and α_2

Price Per Unit	$\alpha_2 = 0.0317$		$\alpha_2 = 0.0517$	
	$\alpha_1 = 0.0317$	$\alpha_1 = 0.0517$	$\alpha_1 = 0.0317$	$\alpha_1 = 0.0517$
P=3 INR	839173.20	842787.19	849834.18	853495.72
P=3.5 INR	979100.50	983291.64	991540.49	995786.31
P=4 INR	1119027.81	1123796.09	1133246.81	1138076.91

Table 4 shows threshold points of plant expenses at particular price of electricity to achieve profit.

6. Conclusion

For two different humidity conditions (i.e., humidity less than or equal to 50% and humidity greater than 50%), a stochastic model of a gas turbine system composed of one gas turbine and one steam turbine is developed by prioritizing repair of gas turbine over steam turbine and applying random inspection and maintenance policy of a system using single service facility. Various reliability measures like system’s mean time to failure, availability for steady state, etc. have been obtained and the graphical analysis of the effects of failure rates of steam turbines when humidity is $\leq/ >$ 50%. Finding shows that mean time to system failure declines as failure rate increases. Trends in

availability for both cycles and varied humidity levels, i.e. when humidity is \leq / $>$ 50%, have been depicted with respect to steam turbine failure rate, and many interesting results about availability have been found. Profit for the plant is shown, which declines as the price of electricity decreases. Furthermore, a thorough study of gas turbine systems may be beneficial to people involved in the industry of electricity generation.

References

- [1] Asif, M. and Muneer, T. (2007). Energy supply, its demand and security issues for developed and emerging economies. *Renewable and Sustainable Energy Reviews*, 11:1388-1413.
- [2] Filippov, S.P., Dil'man M.D. and Ionov, M.S. (2017). Demand of the power industry of Russia for gas turbines: the current state and prospects. *Thermal Engineering*, 64:829–840.
- [3] Taylor, L.D. (1975). The Demand for Electricity: A Survey. *The Bell Journal of Economics*, 6:74-110.
- [4] Gaur, K., Rathour, H.K., Agarwal, P.K., Baba K.V.S. and Soonee, S.K. Analysing the Electricity Demand Pattern. National Power Systems Conference (NPSC), Bhubaneswar, India, 2016.
- [5] Zhang, X., Ramirez-Mendiola, J.L., Li M. and Guo, L. (2022). Electricity consumption pattern analysis beyond traditional clustering methods: A novel self-adapting semi-supervised clustering method and application case study. *Applied Energy*, 308:1-32.
- [6] Derakhshan, G., Shayanfar, H.A. and Kazemi, A. (2016). The optimization of demand response programs in smart grids. *Energy Policy*, 94:295–306.
- [7] Kan, X., Reichenberg L. and Hedenus, F. (2021). The impacts of the electricity demand pattern on electricity system cost and the electricity supply mix: A comprehensive modeling analysis for Europe. *Energy*, 235:1-14.
- [8] Mahdavi, S.S., Saebi, J. and Ghasemi, A. (2023). Risk-based approach for self-scheduling of virtual power plants in competitive power markets. *Journal of Operation and Automation in Power Engineering*, 11:94–104.
- [9] Fernandez, D.A.P., Foliaco, B., Padilla, R.V., Bula, A. and Quiroga, A.G. (2021). High ambient temperature effects on the performance of a gas turbine-based cogeneration system with supplementary fire in a tropical climate. *Case Studies in Thermal Engineering*, 26:101206-101219.
- [10] Emami-Mehrgani, B., Neumann, W.P., Nadeau, S. and Bazrafshan, M. (2016). Considering human error in optimizing production and corrective and preventive maintenance policies for manufacturing systems. *Applied Mathematical Modelling*, 40:2056-2074.
- [11] Zhao, Y., Volovoi, V., Waters, M. and Mavris, D. (2006). A Sequential Approach for Gas Turbine Power Plant Preventive Maintenance Scheduling. *Journal of Engineering for Gas Turbines and Power*, 128:796-805.
- [12] Sammour, A.A., Komarov, O.V., Qasim, M.A., Almalghouj, S., Dakkak, A.M.A. and Du, Y. (2023). Ambient conditions impact on combined cycle gas turbine power plant performance. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 45:557-574.
- [13] Danish, M.S.S., Nazari, Z. and Senjyu, T. (2023). AI-coherent data-driven forecasting model for a combined cycle power plant. *Energy Conversion and Management*, 286:1-20.
- [14] Shukla, A.K. and Singh, O. (2014). Effect of compressor inlet temperature & relative humidity on gas turbine cycle performance. *International Journal of Scientific & Engineering Research*, 5:664-670.
- [15] Hanachi, H., Liu, J., Banerjee, A. and Chen, Y. Effects of the intake air humidity on gas turbine performance monitoring. *ASME Turbo Exposition*, Quebec, Canada, 2015.
- [16] Yeganeh, A.H.J., Behbahaninia, A. and Ghadamabadi, P. (2022). Monte Carlo Simulation of

- a Combined-Cycle Power Plant Considering Ambient Temperature Fluctuations. *Journal of Power and Energy Engineering*, 10:1-21.
- [17] Ghaedi, A., Gorginpour, H. and Noroozi, E. (2021). Operation Studies of the Power Systems Containing Combined Heat and Power Plants. *Journal of Operation and Automation in Power Engineering*, 9:160–171.
- [18] Singh, D. and Taneja, G. (2013). Reliability analysis of a power generating system through gas and steam turbines with scheduled inspection. *Aryabhatta Journal of Mathematics and Informatics*, 5:373-380.
- [19] Singh, D. and Taneja, G. (2014). Reliability and economic analysis of a power generating system comprising one gas and one steam turbine with random inspection. *Journal of Mathematics and Statistics*, 10:436-442.
- [20] Rajesh, Taneja, G. and Prasad, J. (2018). Reliability and availability analysis for a three-unit gas turbine power generating system with seasonal effect and FCFS repair pattern. *International Journal of Applied Engineering Research*, 13:10948-10964.
- [21] Pinki and Singh, D. (2023). Reliability modeling of two-unit gas turbine system considering the effect of humidity. *Reliability: Theory and Applications*, 18:607-617.
- [22] Singh, D. Reliability and economic analysis of some models on gas turbine power plants, Ph.D. Thesis, Maharshi Dayanand University Rohtak, 2013.

PYTHON IMPLEMENTATION OF FUZZY LOGIC FOR ZERO-INFLATED POISSON SINGLE SAMPLING PLANS

Kavithanjali S* and Sheik Abdullah A¹

Department of Statistics, Salem Sowdeswari College, Salem-636 010, Tamil Nadu, India

* kavithanjalis2018@gmail.com, ¹ sheik.stat@gmail.com

Correspondence Email: kavithanjalis2018@gmail.com

Abstract

Acceptance sampling is used in Statistical Quality Control (SQC) to conduct lot quality evaluations through sample inspections which involve probability theory and fuzzy sets. It aims to optimize quality, costs, and productivity, frequently applying linguistic variables when accurate parameter values are not good enough which is handled using fuzzy set theory. This research analyses single sampling plans (SSP) in the presence of fuzzy number non-conformities, modelling them with the Zero-inflated Poisson (ZIP) distribution structure. This study presents a unique method to single sampling plans (SSP) inside the Zero-inflated Poisson (ZIP) distribution framework that makes use of fuzzy logic approaches. In addition, we show how to apply this method using a Python programme, providing practical suggestions for real-world quality control complications.

Keywords: Acceptance sampling plan, Single sampling plan (SSP), ZIP distribution, OC functions, Fuzzy Parameter.

I. Introduction

Acceptance sampling, which is critical in industrial sectors, maintains product quality while balancing time and cost restrictions. It categorizes features and variables, including single, double, and sequential sample plans, which are critical for raw material and product inspections. Designing single sampling plans (SSP) requires balancing producer and consumer interests using criteria such as lot size, sample size, and acceptance rates.

Stephens [16] and Schilling &Neubauer [15] provide details on SSP determination, while Duncan [4] and Schilling &Neubauer [15] expound on approaches based on the Poisson distribution. Technological developments seek towards zero defects, yet random fluctuations require models such as the zero-inflated Poisson (ZIP) distribution, which is a hybrid of the zero-inflated and Poisson distributions. ZIP finds applications across disciplines, from agriculture to manufacturing, detailed in Bohning et.al., [1], Lambert [10], Naya et al. [13], and Ridout et al. [14]. Under the assumption of a zero-inflated Poisson distribution, Loganathan and Shalini [11] created single sample plans based on characteristics. Xie et al. [18] address the construction of control charts using the ZIP distribution. In McLachlan and Peel [12], several theoretical elements of ZIP distributions are discussed. The ZIP (ω, λ) distribution's probability mass function (p.m.f.) may be found in Lambert [10] and McLachlan and Peel [12]. Kavithanjali and Sheik Abdullah [8] review a various sampling plans.

Lotfi A. Zadeh [9] invented fuzzy set theory. Many authors, including Tamaki, Kanagawa and Ohta [17], Grzegorzewski [6], Hrniewicz [7], Chakraborty [3], Buckley [2], EzzatallahBaloui et al. [5], have developed fuzzy statistical theory and statistical applications-based challenges in

recent years. In the ensuing sections, will take a look at the methods used, show the final results of the study, and explain the relevance of our findings for quality control professionals. We are optimistic that our research will contribute considerably to the developing spectrum of statistical approaches designed to address the problems posed by challenging distributions in industrial scenarios. This work investigates finding SSPs based on characteristics within ZIP distribution conditions. Section 2 introduces the approach and terminology. Section 3 describes how to design SSPs fuzzy OC functions and a Python script. Section 4 displays how to pick sample plans, FOC bands, and conclusions, which highlight the study's findings regarding SSP determination.

II. Methodology

2.1 Basic Definitions:

2.1.1 Fuzzy Number: If and only if (i) \tilde{N} is normal (ii) \tilde{N} is fuzzy convex (iii) μ_N is upper semi continuous (iv) $\text{Supp}(\tilde{N})$ is bounded, the fuzzy subset \tilde{N} of the real line R with the membership function $\mu_N : R \rightarrow [0,1]$ is a fuzzy number.

2.1.2 Triangular Fuzzy: A triangular fuzzy number is a fuzzy number \tilde{N} with a membership function given by three numbers $a < c < d$, with the interval $[a, b]$ as the base and $x=c$ as the vertex.

2.1.3 Fuzzy α Cut: The α -cut of a fuzzy integer \tilde{N} is defined as $N[\alpha] = \{X \in R; \mu_N(x) \geq \alpha\}$ in a non-fuzzy set. Consequently, we have $\tilde{N}[\alpha] = [N^L(\alpha), N^U(\alpha)]$

Where $N^L[\alpha] = \inf\{X \in R; \mu_N(x) \geq \alpha\}$ (Infimum of lower limit α -cut)

$N^U[\alpha] = \sup\{X \in R; \mu_N(x) \geq \alpha\}$ (supremum of lower limit α -cut)

2.1.4 ZIP Distribution: The ZIP (φ, λ) distribution's probability mass function (p.m.f.) found in Lambert [9] and McLachlan and Peel [11]. $P(X = d / \varphi, \lambda) = \varphi f(d) + (1 - \varphi)P(X = d / \lambda)$,

where $f(d) = \begin{cases} 1 & \text{if } d = 0 \\ 0 & \text{if } d \neq 0 \end{cases}$ and $P(X = d / \lambda) = \begin{cases} \frac{e^{-\lambda} \lambda^d}{d!} & \text{if } d = 0, 1, 2, \dots, \text{ and } \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$

Given the ZIP (φ, λ) distribution, the probability mass function of the distribution can be written as

$$\tilde{P}(X = d / \varphi, \lambda) = \tilde{P}(d) = \begin{cases} \varphi + (1 - \varphi)e^{-\lambda} & \text{When } d = 0 \\ (1 - \varphi) \frac{e^{-\lambda} \lambda^d}{d!} & \text{When } d = 1, 2, \dots, 0 < \varphi < 1, \lambda > 0 \end{cases}$$

To obtain the fuzzy ZIP probability mass function, replace λ with the fuzzy number $\tilde{\lambda} > 0$. Let $\tilde{P}(d)$ be the approximate probability that D equals d. Next, we get this fuzzy number's α -cut as

$$\tilde{P}(X = d / \varphi, \lambda) = \tilde{P}(d)[\alpha] = \begin{cases} \varphi + (1 - \varphi)e^{-\lambda} & \text{When } d = 0 \\ (1 - \varphi) \frac{e^{-\lambda} \lambda^d}{d!} & \text{When } d = 1, 2, \dots, 0 < \varphi < 1, \lambda > 0 \end{cases} \quad |\lambda \in \lambda(\alpha)$$

For every $\alpha \in [0,1]$ So that $\exists \tilde{p}[\alpha]$. The fuzzy parameter $\tilde{p}(d)[\alpha]$ has been supplanted by $\tilde{P}[a, b][\alpha]$.

$$\tilde{P}[a, b][\alpha] = \begin{cases} \varphi + (1 - \varphi)e^{-\lambda} & \text{When } d = 0 \\ (1 - \varphi) \frac{e^{-\lambda} \lambda^d}{d!} & \text{When } d = 1, 2, \dots, 0 < \varphi < 1, \lambda > 0 \end{cases} \quad |\lambda \in \tilde{p}(\alpha)$$

The ZIP (φ, λ) has a mean of $(1 - \varphi)$ and a variance of $(1 - \varphi)(1 + \lambda \varphi)$.

III. OC function of SSP in ZIP distribution conditions

A Single Sampling Plan (SSP) with characteristics is defined by three parameters N, n, and c. A random sample of size n is taken from a large number of N units, and the number of

nonconforming units, $X = d$, is counted. If $d < c$, the lot is accepted, otherwise it is rejected. Evaluating the performance of a sample plan entails examining its Operating Characteristic (OC) function, which indicates its ability to discern acceptable from non-acceptable lots based on certain criteria. $\tilde{P}_a(p) = [X \leq c]$

Using the Zero-Inflated Poisson model, the probability mass function of the number of defects in the lot is given by

$$\tilde{P}(X = d / \varphi, \lambda) = \tilde{P}(d) = \begin{cases} \varphi + (1 - \varphi)e^{-\lambda} & \text{When } d = 0 \\ (1 - \varphi) \frac{e^{-\lambda} \lambda^d}{d!} & \text{When } d = 1, 2, \dots, 0 < \varphi < 1, \lambda > 0 \end{cases}$$

Given a sample size of n , the probability of finding no deficiencies will be

$$\tilde{P}(X = 0) = \tilde{P}_a(p) = \varphi + (1 - \varphi)e^{-n\tilde{p}} \tag{1}$$

This is the single sample plan's OC function when $c=1$. Then equation becomes

$$\tilde{P}_a(p) = \varphi + (1 - \varphi)e^{-n\tilde{p}} (1 + n\tilde{p}) \tag{2}$$

Which is the single sampling plan's OC function for $c=1$

3.1 Python Programming

Python programming was used in this study on statistical quality control to create the Fuzzy OC Band table's upper and lower bounds. The Fuzzy Operating Characteristic (OC) and Fuzzy Probability of Acceptance curves were also drawn using Python. Python's extensive numerical calculation capabilities and flexible modules made it easy to use these statistical approaches inside the study framework.

Illustration 1: According to the company's experience, 0.5 percent of packages are empty. A department store has 60 items of this product on hand, and many customers select and browse asking if they can buy that item. If our search shows that this sample has only one mismatch, the customer gets away buy every item in the store, otherwise, The fuzzy number where the customer can choose not to buy that product can be taken as $\tilde{P} = (0, 0.005, 0.01)$. Consequently, the probability of purchase is similar to that to be described.

$$n=60, c=1, \tilde{P}=(0,0.005,0.01), \tilde{\lambda} = np, \varphi = 0.0001$$

$$\tilde{\lambda} = [0.3, 0.6], \tilde{\lambda}[\alpha] = [0.3\alpha, 0.6 - 0.3\alpha]$$

$$\tilde{P}_a(p) = \varphi + (1 - \varphi)e^{-n\tilde{p}} + (1 - \varphi)e^{-n\tilde{p}} (n\tilde{p}) \quad |n\tilde{p} \in \tilde{\lambda}(\alpha)$$

Therefore, the $\varphi + (1 - \varphi)e^{-n\tilde{p}} (1 + n\tilde{p})$ decreasing, then

$$\tilde{P}_a(p) = \varphi + (1 - \varphi)e^{-(0.6-0.3\alpha)} (1 + (0.6 - 0.3\alpha)), \varphi + (1 - \varphi)e^{-(0.3\alpha)} (1 + 0.3\alpha)$$

Under $\alpha = 0, 0.005, 0.01$ discover $\tilde{P}_a[p] = [0.8781, 1], [0.87860, 0.9999], [0.8790, 0.9999]$ In

Figure1 it shows is expected that 88 to 100 lots out of every 100 lots in this process will be accepted.

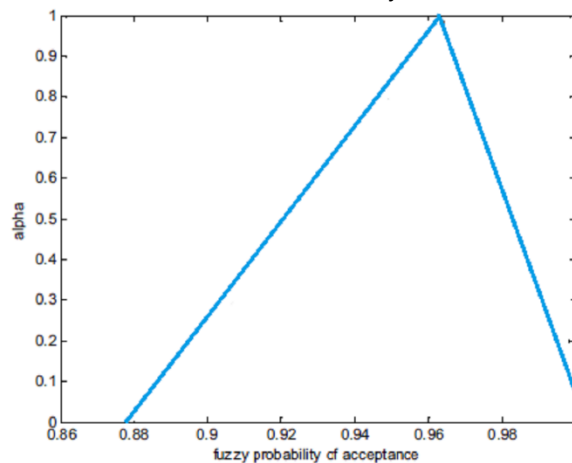


Figure 1: Fuzzy Probability of acceptance with $\tilde{P}=(0,0.005,0.01)$

IV. FOC band

The SSP operational characteristic curve is built using fuzzy parameters. The operational characteristic curve represents both the proportion of defective p and the probability of acceptance $(Pa(p))$. The sampling plan's operational characteristic curve can be utilized to identify excellent and challenging lots. When a consumer rejects a product that meets the established conditions (i.e., the product quality is good), this risk is referred to as producer's risk; when a consumer accepts a product that does not meet the conditions (i.e., the product quality is bad), this risk is referred to as consumer's risk. An upper and lower band fuzzy parameter may be used to calculate the proportion defective, if the values of the upper and lower band are equal, this is referred to as superior state.

$$\text{In firm related to example we had } n=60, \varphi = 0.0001, \tilde{P}=(0,0.005,0.01) \\
c = 1, a_2 = 0.005, a_3 = 0.01, \alpha = 0, \tilde{\lambda}[0] = [nk, nk + 0.01n], 0 \leq k \leq 0.99$$

$$\text{From equation (2) } \tilde{P}_a(p) = \varphi + (1 - \varphi)e^{-(\tilde{\lambda}_2\alpha)} (1 + (\tilde{\lambda}_2\alpha)), \varphi + (1 - \varphi)e^{-(\tilde{\lambda}_1\alpha)} (1 + (\tilde{\lambda}_1\alpha)) \\
= \varphi + (1 - \varphi)e^{-(nk+0.01n)} (1 + nk + 0.01n), \varphi + (1 - \varphi)e^{-(nk)} (1 + nk)$$

Table 1: Fuzzy Probability of Acceptance $\varphi = 0.0001, c=1, n=60$

k	$\tilde{P}_a(p)$
0	0.878111, 1.000000
0.01	0.662661, 0.878111
0.02	0.462891, 0.662661
0.03	0.308510, 0.462891
0.04	0.199228, 0.308510
0.05	0.125777, 0.199228
0.06	0.078069, 0.125777
0.07	0.047828, 0.078069
0.08	0.029003, 0.047828
0.09	0.017450, 0.029003
0.1	0.010438, 0.017450

Example 1. Table 1 and Figure 2 illustrate the OC curve. This graphic shows how process quality will drop from extremely good to moderate, while the OC curve will expand.

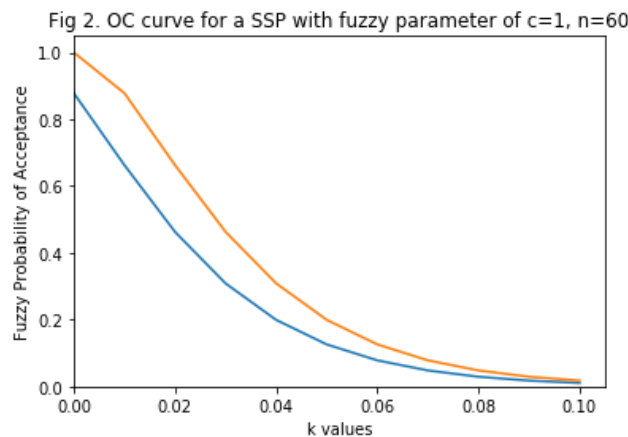


Figure 2: OC Curve for SSP with Fuzzy Parameter of $c=1, n=60$

Illustration 2: Had $\varphi = 0.0001, c=0$ and $a_2=0.005, a_3=0.01, \tilde{\lambda}[0] = [20k + 0.2], 0 \leq k \leq 0.99,$

leading to OC curve in terms of fuzzy ZIP distribution. From equation (1),

$$\begin{aligned} \tilde{P}_a(p) &= \varphi + (1 - \varphi)e^{-n\tilde{p}} \quad | \lambda \epsilon \lambda(0) = \varphi + (1 - \varphi)e^{-(0.01n+nk)}, \varphi + (1 - \varphi)e^{-(nk)} \\ &= \varphi + (1 - \varphi)e^{-(0.2+20k)}, \varphi + (1 - \varphi)e^{-(20k)} \end{aligned}$$

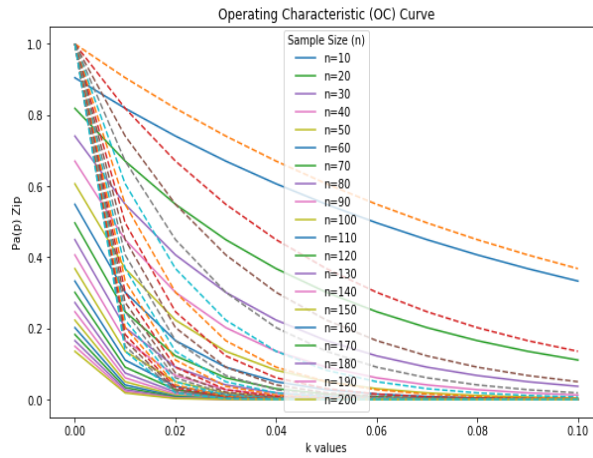


Figure 3: OC Curve for a SSP with fuzzy parameter of c=0

Table 2: Fuzzy Probability of Acceptance $\varphi = 0.0001, c=0$ with different sample size(n)

n=10		n=20		n=30		n=40	
K	Pa(p)	k	Pa(p)	k	Pa(p)	k	Pa(p)
0	0.9048, 1.0000	0	0.8187, 1.0000	0	0.7408, 1.0000	0	0.6704, 1.0000
0.01	0.8187, 0.9048	0.01	0.6704, 0.8187	0.01	0.5489, 0.7408	0.01	0.4494, 0.6704
0.02	0.7408, 0.8187	0.02	0.5489, 0.6704	0.02	0.4066, 0.5489	0.02	0.3013, 0.4494
0.03	0.6704, 0.7408	0.03	0.4494, 0.5489	0.03	0.3013, 0.4066	0.03	0.2020, 0.3013
0.04	0.6066, 0.6704	0.04	0.3679, 0.4494	0.04	0.2232, 0.3013	0.04	0.1354, 0.2020
0.05	0.5489, 0.6066	0.05	0.3013, 0.3679	0.05	0.1654, 0.2232	0.05	0.0908, 0.1354
0.06	0.4966, 0.5489	0.06	0.2467, 0.3013	0.06	0.1225, 0.1654	0.06	0.0609, 0.0908
0.07	0.4494, 0.4966	0.07	0.2020, 0.2467	0.07	0.0908, 0.1225	0.07	0.0409, 0.0609
0.08	0.4066, 0.4494	0.08	0.1654, 0.2020	0.08	0.0673, 0.0908	0.08	0.0274, 0.0409
0.09	0.3679, 0.4066	0.09	0.1354, 0.1654	0.09	0.0499, 0.0673	0.09	0.0184, 0.0274
0.1	0.3329, 0.3679	0.1	0.1109, 0.1354	0.1	0.0370, 0.0499	0.1	0.0124, 0.0184
n=60		n=80		n=100		n=120	
K	Pa(p)	k	Pa(p)	k	Pa(p)	k	Pa(p)
0	0.5489, 1.0000	0	0.4494, 1.0000	0	0.3679, 1.0000	0	0.3013, 1.0000
0.01	0.3013, 0.5489	0.01	0.2020, 0.4494	0.01	0.1354, 0.3679	0.01	0.0908, 0.3013
0.02	0.1654, 0.3013	0.02	0.0908, 0.2020	0.02	0.0499, 0.1354	0.02	0.0274, 0.0908
0.03	0.0908, 0.1654	0.03	0.0409, 0.0908	0.03	0.0184, 0.0499	0.03	0.0083, 0.0274
0.04	0.0499, 0.0908	0.04	0.0184, 0.0409	0.04	0.0068, 0.0184	0.04	0.0026, 0.0083
0.05	0.0274, 0.0499	0.05	0.0083, 0.0184	0.05	0.0026, 0.0068	0.05	0.0008, 0.0026
0.06	0.0151, 0.0274	0.06	0.0038, 0.0083	0.06	0.0010, 0.0026	0.06	0.0003, 0.0008
0.07	0.0083, 0.0151	0.07	0.0018, 0.0038	0.07	0.0004, 0.0010	0.07	0.0002, 0.0003
0.08	0.0046, 0.0083	0.08	0.0008, 0.0018	0.08	0.0002, 0.0004	0.08	0.0001, 0.0002
0.09	0.0026, 0.0046	0.09	0.0004, 0.0008	0.09	0.0001, 0.0002	0.09	0.0001, 0.0001
0.1	0.0015, 0.0026	0.1	0.0003, 0.0004	0.1	0.0001, 0.0001	0.1	0.0001, 0.0001
n=140		n=160		n=180		n=200	
K	Pa(p)	k	Pa(p)	k	Pa(p)	k	Pa(p)

0	0.2467,1.0000	0	0.2020,1.0000	0	0.1654,1.0000	0	0.1354,1.0000
0.01	0.0609,0.2467	0.01	0.0409,0.2020	0.01	0.0274,0.1654	0.01	0.0184,0.1354
0.02	0.0151,0.0609	0.02	0.0083,0.0409	0.02	0.0046,0.0274	0.02	0.0026,0.0184
0.03	0.0038,0.0151	0.03	0.0018,0.0083	0.03	0.0008,0.0046	0.03	0.0004,0.0026
0.04	0.0010,0.0038	0.04	0.0004,0.0018	0.04	0.0002,0.0008	0.04	0.0001,0.0004
0.05	0.0003,0.0010	0.05	0.0002,0.0004	0.05	0.0001,0.0002	0.05	0.0001,0.0001
0.06	0.0002,0.0003	0.06	0.0001,0.0002	0.06	0.0001,0.0001	0.06	0.0001,0.0001
0.07	0.0001,0.0002	0.07	0.0001,0.0001	0.07	0.0001,0.0001	0.07	0.0001,0.0001
0.08	0.0001,0.0001	0.08	0.0001,0.0001	0.08	0.0001,0.0001	0.08	0.0001,0.0001
0.09	0.0001,0.0001	0.09	0.0001,0.0001	0.09	0.0001,0.0001	0.09	0.0001,0.0001
0.1	0.0001,0.0001	0.1	0.0001,0.0001	0.1	0.0001,0.0001	0.1	0.0001,0.0001

Table 2 and Figure 3 show that separate curves in the plot indicate sample sizes ranging from 10 to 200. Each curve represents a unique sample size and indicates how the probability of accepting the null assumption varies with effect size (k) for different sample sizes.

- The probability of rejecting the null hypothesis rises with an increase in effect size, k.
- smaller sample sizes (lower (n)) typically have lesser (k) discriminative powers to identify differences, which raises the likelihood of adopting the null hypothesis.
- With larger sample sizes, the ability to detect differences improves and the null theory is less likely to be accepted.
 - The OC curve, in essence, represents the relationship between sample size, effect magnitude, and the chance of not rejecting the hypothesis. As a result of this, statisticians can evaluate their assumptions and make choices with greater certainty.
 - Finally, the fuzzy ZIP distribution can be used to approximate the OC curve. In this regard, a plan of this such can be created using the OC fuzzy ZIP distribution. The OC curve shows that zero convergences with the acceptance number (c), which causes a rapid fall in the fuzzy probability of accepting the proportion of faulty goods with small fuzzy numbers. This is why there is the increase in n.

V. Conclusions

In this research introduces a new way to design SSP by combining fuzzy logic with the ZIP distribution. This improves how we control quality. In this method manages risks for both producers and consumers well. Importantly, In this plans work smoothly alongside traditional ones when damage is rare, making them adaptable. The suggested OC curves have clear restrictions, no acceptance values, and are simple to interpret, making them extremely helpful. This new approach improves the way we choose samples in quality control, particularly when dealing with complex distribution patterns. We seek to improve our goods and make users better through using fuzzy logic and ZIP distribution, particularly in competitive marketplaces.

References:

- [1] Bohning D, Dietz E, Schlattmann P.(1999). The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of royal statistical society, Series A*, 162(2),195-209.
- [2] Buckley J. J. Fuzzy probability: New approach and application, physica-velage, Heidelberg, Germany, (2003).
- [3] Chakraborty T.K.(1994) A class of single sampling plan based on fuzzy Optimization, *Opsearch*, 63(1), 35-43.

-
- [4] Duncan, A. J. Quality Control and Industrial Statistics. Homewood: Richard D. Irwin, Inc. 1986.
- [5] EzzatallahBalouiJamkhaneh et. Al.,(2009). Acceptance Single Sampling Plan with fuzzy parameter with the using of Poisson distribution, International journal of mathematical, computational, physical, *electrical and Computer Engineering*, 3(1), 1017-1021.
- [6] Grzegorzewski P.(1008) A soft design of acceptance sampling by attributes, Proceedings of the VIth International Workshop on Intelligent statistical quality control. *Wurzburg*, 14(16), 29-38.
- [7] Hrniewicz. O.(1994). Statistical acceptance sampling with uncertain information from a sample and fuzzy quality criteria, Working Paper of SRI PAS, Warsaw,(in Polish), 16(1), 197-206.
- [8] Kavithanjali, S Sheik Abdullah, A. (2024). A Review study on Various Sampling Plans, *GVN Publication*, 1(2), 15-28.
- [9] L.A Zadeh. (1965). Fuzzy Sets, *Information and Control*, 38(4), 338-353.
- [10] Lambert D.(1992). Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics*, 34(1), 1-14.
- [11] Loganathan A, Shalini K.(2014). Determination of Single Sampling Plans by Attributes Under the Conditions of Zero-Inflated Poisson Distributions. *Communication in Statistics-Simulation and Computation*, 43(3), 538-548.
- [12] McLachlan G, Peel D. Finite Mixture Models. New York: John Wiley & Sons, 2000.
- [13] Naya H, Urioste JI, Chang YM, Motta MR, Kremer R, Gianola D.(2008). A comparison between Poisson and zero inflated Poisson regression models with an application to number of black spot in corriedale sheep. *Genetics Selection Evolution*. 40(1), 379-394.
- [14] Ridout M, Demtrio CGB, Hinde J.(1998). Models for Count Data with Many Zeros. Cape Town, *International Biometric Conference*, 34(1), 531-547.
- [15] Schilling, E. G., Neubauer, D. V. Acceptance Sampling in Quality Control. Boca Raton: CRC Press. 2009.
- [16] Stephens, K. S. The Handbook of Applied Acceptance Sampling Plans, Procedures and Principles. Wisconsin: ASQ Quality Press. 2001.
- [17] Tamaki f., Kanagawa a. And Ohta h. (1991) A Fuzzy Design of Sampling Inspection plans by Attributes, Japanese journal of fuzzy theory and systems, 3(4), 767-776.
- [18] Xie M, He B, Goh TN.(2001). Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis*, 38(2),191-201.

ANALYSIS OF NON-MARKOVIAN BATCH ARRIVAL RETRIAL QUEUE WITH PRIORITY SERVICES, IMMEDIATE FEEDBACK, PUSH OUT, DIFFERENTIATED BREAKDOWNS, DELAYED REPAIR, RANDOMIZED VACATION

¹G. AYYAPPAN, ²S. NITHYA

•

Department of Mathematics
Puducherry Technological University
Puducherry-605014, India
ayyappan@ptuniv.edu.in, nithyamouttouraman@gmail.com

Abstract

Priority and ordinary customers arrive according to Poisson processes, and their service time based on the general distribution. The server constantly offers a single service for both priority and ordinary customers. We compute the Laplace transforms of the time-dependent probabilities of system states using the probability generating function and supplementary variable technique. Numerical results are obtained which are also examined to facilitate the sensitivity analysis of system descriptions.

Keywords: Batch Arrivals; Priority Queues; Immediate Feedback; Push out; Differentiated Breakdowns; Delayed Repair; Randomized Vacation.

AMS Subject Classification (2010): 60K25, 68M30, 90B22.

1. INTRODUCTION

We see several queueing situations every day where customers must wait for service and there is a delay in providing it. Retrial queues in queueing theory have been the topic of a lot of exciting research over the last two decades. The concept of retrial queues has attracted the attention of numerous scholars and received important contributions from them. An $M/M/1$ retrial queueing system with Poisson arrival flows, impatient customers, breakdown, collisions was studied by Danilyuk et al. [9]. Nazarov et al. [16] investigated how, depending on whether the server is busy or idle, it is dependent on random failures and repairs in a retry queueing system with a finite number of sources and customer collision. For the aggregation of the customers and their group service, D'Arienzo et al. [10] created a single-server retrial queue with a MAP flow, PH service times, and a finite capacity. Ahuja et al. [1] explored the retrial queueing system with an optional service and finite population subject to balking. Pavai Madheswari et al. [17] analysed an $M/G/1$ retrial queueing system with two service phases, the second of which is optional, and a server working on a Bernoulli vacation schedule. Innovative applications for performance study of various systems in telecommunications, data split networks, traffic management on high-speed networks, and production engineering make use of these queueing models.

The literature on retrial queueing has extensively researched retrial queues with various customer categories. An important component of priority discipline is preemptive and non-preemptive priorities. D'Apice et al. [11] considered a priority queueing model with many types of requests and restricted processor sharing. Ayyappan and Thilgavathy [3] determined priority queueing system with breakdown, repair, discouragement, single vacation, standby server,

negative arrival and impatient customers. Li et al. [15] investigate equilibrium queueing strategies in an unobservable non-preemptive priority queue with homogeneous customers. Ammar and Rajadurai [2] introduced preemptive priority retrial queueing system with disaster under working breakdown services.

Most of the time, while discussing queueing, it is generally thought that the server is always accessible. However, server failure made the significant impact in queueing system. Therefore, in order to establish retrial queueing models, it is essential to carry out investigation on the retrial queue with breakdowns. Choudhury and Kalita [8] described a non-Markovian queueing system with breakdown, delayed repair and two general heterogeneous service, optional service. Krishna Kumar et al. [14] examined a Markovian retrial queue where the server is subject to breakdowns and repairs. Gao et al. [12] studied an $M/G/1$ retrial queue with two types of breakdowns. Ayyappan and Gowthami [5] researched the single server classical queueing system $MAP/PH/1$ with breakdown, repair, Bernoulli vacation and setup time. Begum and Choudhry [7] investigated a $M/(G1, G2)/1$ queue with service interruption consisting of a definite repairability.

Customers may be serviced more than once for particular reasons in several queueing situations. Customers have to re-join the queue and wait in queue after the service is completed. Optional re-service is a concept that can be considered as immediate feedback in this regard. The customer receives their service in the first step, and if there is any issue with it or they need it again, they will receive it immediately without having to wait in queue. Re-service has several practical applications in places like bank desks, functioning ATMs, large supermarkets, and medical facilities etc., The idea of immediate feedback (re-service) has been addressed by some authors, including Azhagappan and Deepa [6], Ayyappan and Deepa [4] and Jose and Deepthi [13]. According to the previously mentioned literature, a customer who wants to receive more service must visit the server once more at that moment.

The interesting parameter in this chapter is the randomized vacation policy. It is described as follows: After the vacation completion, if there is at least one unit present in the system, then the server immediately commence the service. Otherwise, the server will decide either remains idle or go for another vacation, if no units present in the system. The concept of variant vacation policy was proposed by Takagi [18], which is a generalization of the single and multiple vacation for the $M/G/1$ queueing system. Ke et al. [19] studied an $M^X/G/1$ queueing system with a randomized vacation policy and at most J vacations. Geo and Yao [20] developed this vacation policy for an $M^X/G/1$ queueing system, in which the server takes randomized vacation policy and at most J working vacations.

There are many papers dealing with unit's abandoned behaviour. Recently, Gao et al. [21] studied an $M/G/1$ retrial queue with abandoned customers and multi-optional vacations. Krishnamoorthy et al. [22] presented an $M_1, M_2/PH/1$ retrial queueing system with pre-emptive priority service, orbital search and abandoned units in which the retrial is failed, then the failed units abandoned the system with certain probabilities. In this model, the arriving ordinary unit may remove the ordinary unit, who is getting service from the system. Here, the interrupting ordinary unit is referred as the abandoned unit.

We consider non-Markovian batch arrival retrial queue with priority services, discouragement, re-service, differentiate breakdown, restoration and delayed vacation. Priority customers and ordinary customers arrive according to Poisson processes, and their service time based on the general distribution. The server consistently provides a single service for both priority customers and ordinary customers. In the event of the server being unavailable, ordinary customers may choose to balk the system. Server failure may happen at any time during normal engaged period. The two types of system failure are hard and soft failures. Hard failure can be characterised as an equipment breakdown which demands the availability of a skilled repair person, which is an extensive process. Soft failure is described as breakdown based on by circumstances as instead of

mechanical components, and it can be generally resolved by restarting the system. Customer may re-enter the system as a feedback customer for receiving normal service due to inadequate quality of service after every priority service is completed. The server goes on vacation after priority services are completed; the time it takes the server to go on vacation is known as delay time.

The article's remaining content is formatted as follows: In Section 2, the mathematical model is presented. and the distribution of queue sizes is analysed in Section 3. Section 4 contains the exact expression for the governing equation. Section 5 of this article discusses steady state analysis. Section 6 lists stability condition. Section 7 provides an illustration of how system performance measures have an impact. Section 8 exhibits particular cases. In Sections 9 and 10, conclusions are drawn after deriving numerical and graphical results.

1.1. Integrating the Model into Real-life Situations

In the online food delivery network, independent contractors are responsible for delivering food orders to customers in their designated service areas. These contractors are self-employed individuals who have chosen to work with the food delivery platform, offering their transportation services to ensure that customers receive their meals promptly and efficiently. The nature of the food delivery business can sometimes lead to fluctuations in the availability of delivery orders within a specific area. This could be due to various factors, such as changes in customer demand, local events, or even the time of day. When independent contractors find that there are no orders available in their assigned service area, they may face a few challenges and considerations. To maintain their income and professional engagement during periods of low order availability, these contractors might choose to take a break from their work or explore other temporary job opportunities. By doing so, they can keep themselves occupied and ensure a steady income while waiting for orders to become available in their area again. This approach allows them to manage their workload and personal commitments more effectively.

Once the independent contractors have completed their temporary work or vacation, they can return to the online food delivery network and resume their services when orders are available in their area. This flexibility is crucial for contractors, as it enables them to balance their professional and personal lives while staying connected to the platform and being ready to serve customers when needed. In conclusion, the online food delivery networks independent contractors operate within separate service areas, and when delivery orders are scarce in their region, they may opt for vacations or other work to maintain their income and professional engagement. Upon the return of available orders, these contractors can resume their food delivery services, ensuring a balance between their work and personal lives. This flexibility helps them adapt to the dynamic nature of the food delivery business and maintain a sustainable work arrangement.

2. MATHEMATICAL DESCRIPTION

- **Arrival Process :**

Two distinct customer arrive in batches through separate Poisson compound processes. $\lambda_p, \lambda_o > 0$ are used to indicate, for PC and OC, the respective arrival rates. Assume the initial order probability for both priority and ordinary customers $\lambda_p c_i dt$ ($i = 1, 2, 3, \dots$) and $\lambda_o c_j dt$ ($j = 1, 2, 3, \dots$) respectively. The system has i and j batch size customers enters within a short period of time $(t, t + dt)$. Here, $0 \leq c_i, c_j \leq 1, \sum_{i/j=1}^{\infty} c_{i/j} = 1$.

- **Retrial Service Process :**

Customers who are on retrial are treated the same as ordinary customers. Customers on a retrial are regular consumers. These customers will eventually return to orbit and seek their services again if the server is engaged or unavailable. Retrial service time is characterised by

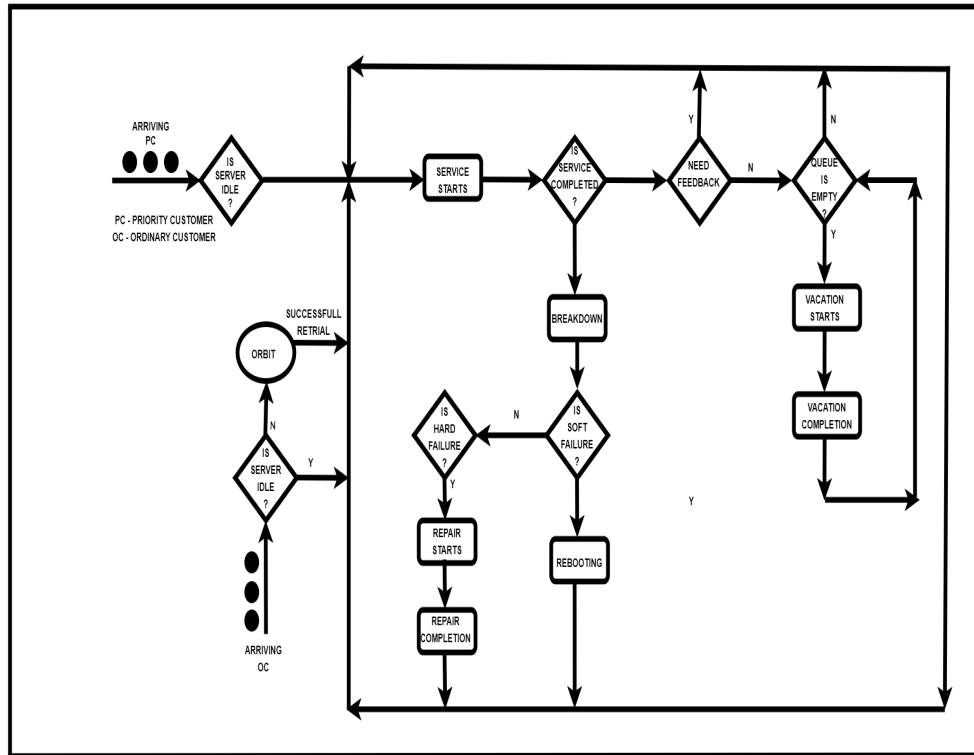


Figure 1: Schematic representation

a rate of $\beta(u)$, which is defined by the probability density function $m(u)$ and the probability distribution function $M(u)$. This rate follows a general distribution.

• **Regular Service Process :**

Customers with different queues are served in batches, including priority and ordinary customers. The server offers a single service at rates of $\mu_i(u)$, $i = 1, 2$ and service rates are distributed generally, characterized by probability distribution function $B_i(s)$ and probability density function $b_i(s)$, where $i = 1, 2$. If the priority queue is empty, ordinary customers can begin receiving service.

• **Immediate Feedback:**

Only customers with a priority are offered the feedback service. Once every customer in the priority queue have received their services, if customers are dissatisfied with the service they received. Customers could either abandon the system with probability $(1 - r)$ or they could get a re-service with probability 'r' without joining in queue.

• **Push Out:**

When the server is attending to an ordinary customers, a newly arriving ordinary customers has the potential to disrupt the ongoing service. It can either immediately take over the service area with probability 'q' or joins the orbit with probability ' \bar{q} ' ($= 1 - q$).

• **Differentiate Breakdown:**

The service channel is susceptible to failure at any moment, even when the server is operating at its usual engaged pace during any phase of service. Consequently, the server will be inaccessible for a brief duration. Both hard and soft failure rates follow exponential distributions, with rates denoted by α_1 & α_2 respectively.

• **Delayed repair and Repair :**

The server is subject to hard failure. The breakdown server does not send for repair instantly. There exists a delay period before the repair process initiates. Following this delay, the repair process commences to restore functionality. The probability distribution function $D(s)$ and $R^{(2)}(s)$, along with the probability density function $d(s)$ and $r^{(2)}(s)$ are employed

to characterize the delay time and the duration of hard failure repairs, respectively. Let $\zeta(u)$ and $\eta_2(u)$ be the completion rate for delay repair time and hard failure repair time. Soft failure repair time distributed exponentially with rate η_1 .

• **Randomized Vacation:**

When the server determines that the system is empty upon completing a service, it opts for a vacation. Upon concluding the vacation, if the server observes an empty system again it decides whether to embark on another vacation with prob. 'p' or to remain idle with a probability of '(1-p)'. The duration of the vacation adheres to a general distribution characterized by a rate of $\gamma(u)$. The time spent on vacation follows a probability distribution function $V(s)$, accompanied by its corresponding probability density function $v(s)$.

3. ANALYSIS OF QUEUE SIZE DISTRIBUTION

The formation of governing equations is the main focus of this section. This model has been solved using the probability generating function and supplementary variable technique with respect to the non-Markovian queueing system.

Assuming that $M(0) = 0, M(\infty) = 1, B_i(0) = 0, B_i(\infty) = 1, V(0) = 0, V(\infty) = 1, D(0) = 0, D(\infty) = 1,$ and $R^{(2)}(0) = 0, R^{(2)}(\infty) = 1$ are continuous at $u = 0$ for $i = 1, 2$. So that the function $\beta(u), \mu_1(u), \mu_2(u), \gamma(u), \zeta(u)$ and $\eta_2(u)$ are the conditional completion rates for retrial, priority and ordinary customers service rate, vacation and delay time to hard failure repair, hard failure repair respectively.

Also, $\beta(u) = \frac{dM(u)}{1-M(u)}, \mu_i(u) = \frac{dB_i(u)}{1-B_i(u)}, \gamma(u) = \frac{dV(u)}{1-V(u)}, \zeta(u) = \frac{dD(u)}{1-D(u)}$ and $\eta_2(u) = \frac{dR^{(2)}(u)}{1-R^{(2)}(u)}$; $i = 1, 2$. are the hazard rate functions of $M(\cdot), B_i(\cdot)$ for $i = 1, 2, V(\cdot), D(\cdot)$ and $R^{(2)}(\cdot)$ respectively.

Markov process for the given model is $\{N_p(t), N_o(t), Y(t), M^0(t), B_1^0(t), B_2^0(t), (V)^0(t), (D)^0(t), (R^{(2)})^0(t)\}$, where $N_p(t)$ and $N_o(t)$ denote the number of customers in the priority queue and ordinary queue respectively. $M^0(t), B_1^0(t), B_2^0(t), (V)^0(t), (D)^0(t), (R^{(2)})^0(t)$ are the elapsed retrial, service, vacation, delay time to hard failure repair and hard failure repair time of the server at time 't'.

$Y(t)$ represents the server state. Here $Y(t) = (0, 1, 2, 3, 4, 5, 6)$, denotes: 0, the server is idle; 1, retrial state; 2, engaged with PC; 3, engaged with OC; 4, on vacation ; 5, delayed to hard failure repair, and 6, hard failure repair.

Let's represent the probability as $I_{0,n_2}(t)$, indicating the probability that at time $t, I_{0,n_2}(t)$ equals the event that $N_p(t) = 0, N_o(t) = n_2,$ and $Y(t) = 0$, where $t > 0$. We consider probability densities for this scenario.

$$\begin{aligned}
 I_{0,n_2}(u, t)du &= \Pr\{N_p(t) = 0, N_o(t) = n_2, Y(t) = 1; u \leq I^0(t) \leq u + du\}, n_2 \geq 1 \\
 P_{n_1,n_2}^{(1)}(u, t)du &= \Pr\{N_p(t) = n_1, N_o(t) = n_2, Y(t) = 2; u \leq B_1^0(t) \leq u + du\}, \\
 P_{n_1,n_2}^{(2)}(u, t)du &= \Pr\{N_p(t) = n_1, N_o(t) = n_2, Y(t) = 3; u \leq B_2^0(t) \leq u + du\}, \\
 V_{n_1,n_2}(u, t)du &= \Pr\{N_p(t) = n_1, N_o(t) = n_2, Y(t) = 4; u \leq V^0(t) \leq u + du\}, \\
 D_{n_1,n_2}(u, t)du &= \Pr\{N_p(t) = n_1, N_o(t) = n_2, Y(t) = 5; u \leq D^0(t) \leq u + du\}, \\
 R_{n_1,n_2}^{(2)}(u, t)du &= \Pr\{N_p(t) = n_1, N_o(t) = n_2, Y(t) = 6; u \leq R^0(t) \leq u + du\}, \\
 &\text{for } u \geq 0, t \geq 0, n_1 \geq 0 \text{ and } n_2 \geq 0.
 \end{aligned}$$

4. EQUATION GOVERNING THE SYSTEM

$$\frac{d}{dt} I_{0,0}(t) = -(\lambda_p + \lambda_o) I_{0,0}(t) + (1 - p) \int_0^\infty V_{0,0}(u, t) \gamma(u) du, \tag{1}$$

$$\frac{\partial}{\partial t} I_{0,n}(u, t) + \frac{\partial}{\partial u} I_{0,n}(u, t) = -(\lambda_p + \lambda_o + \beta(u)) I_{0,n}(u, t), \tag{2}$$

$$\begin{aligned} \frac{\partial}{\partial t} P_{n_1, n_2}^{(1)}(u, t) + \frac{\partial}{\partial u} P_{n_1, n_2}^{(1)}(u, t) &= -(\lambda_p + \lambda_o + \alpha_1 + \alpha_2 + \mu_1(u)) P_{n_1, n_2}^{(1)}(u, t) \\ &+ \lambda_p (1 - \delta_{0n_1}) \sum_{i=1}^{n_1} c_i P_{n_1-i, n_2}^{(1)}(u, t) + \lambda_o (1 - \delta_{0n_2}) \sum_{j=1}^{n_2} c_j P_{n_1, n_2-j}^{(1)}(u, t), \end{aligned} \tag{3}$$

$$\begin{aligned} \frac{\partial}{\partial t} P_{n_1, n_2}^{(2)}(u, t) + \frac{\partial}{\partial u} P_{n_1, n_2}^{(2)}(u, t) &= -(\lambda_p + b\lambda_o + \alpha_1 + \alpha_2 + \mu_1(u)) P_{n_1, n_2}^{(2)}(u, t) \\ &+ \lambda_p (1 - \delta_{0n_1}) \sum_{i=1}^{n_1} c_i P_{n_1-i, n_2}^{(2)}(u, t) + \lambda_o (1 - \delta_{0n_2}) \sum_{j=1}^{n_2} c_j P_{n_1, n_2-j}^{(2)}(u, t), \end{aligned} \tag{4}$$

$$\begin{aligned} \frac{d}{dt} R_{n_1, n_2}^{(1)}(u, t) + \frac{d}{du} R_{n_1, n_2}^{(1)}(u, t) &= -(\lambda_p + b\lambda_o + \eta_1) R_{n_1, n_2}^{(1)}(t) + \lambda_p (1 - \delta_{0n_1}) \sum_{i=1}^{n_1} c_i R_{n_1-i, n_2}^{(1)}(t) \\ &+ \alpha_1 \int_0^\infty (P_{n_1, n_2}^{(1)}(u, t) + P_{n_1, n_2}^{(2)}(u, t)) du + \lambda_p (1 - \delta_{0n_2}) \sum_{j=1}^{n_2} c_j R_{n_1, n_2-i}^{(1)}(t), \end{aligned} \tag{5}$$

$$\begin{aligned} \frac{\partial}{\partial t} D_{n_1, n_2}(u, t) + \frac{\partial}{\partial u} D_{n_1, n_2}(u, t) &= -(\lambda_p + b\lambda_o + \zeta(u)) D_{n_1, n_2}(u, t) \\ &+ \lambda_p (1 - \delta_{0n_1}) \sum_{i=1}^{n_1} c_i D_{n_1-i, n_2}(u, t) + \lambda_o (1 - \delta_{0n_2}) \sum_{j=1}^{n_2} c_j D_{n_1, n_2-i}(u, t), \end{aligned} \tag{6}$$

$$\begin{aligned} \frac{\partial}{\partial t} R_{n_1, n_2}^{(2)}(u, t) + \frac{\partial}{\partial u} R_{n_1, n_2}^{(2)}(u, t) &= -(\lambda_p + b\lambda_o + \eta_2(u)) R_{n_1, n_2}^{(2)}(u, t) \\ &+ \lambda_p (1 - \delta_{0n_1}) \sum_{i=1}^{n_1} c_i R_{n_1-i, n_2}^{(2)}(u, t) + \lambda_o (1 - \delta_{0n_2}) \sum_{j=1}^{n_2} c_j R_{n_1, n_2-i}^{(2)}(u, t), \end{aligned} \tag{7}$$

$$\begin{aligned} \frac{\partial}{\partial t} V_{n_1, n_2}(u, t) + \frac{\partial}{\partial u} V_{n_1, n_2}(u, t) &= -(\lambda_p + b\lambda_o + \gamma(u)) V_{n_1, n_2}(u, t) \\ &+ \lambda_p (1 - \delta_{0n_1}) \sum_{i=1}^{n_1} c_i V_{n_1-i, n_2}(u, t) + \lambda_o (1 - \delta_{0n_2}) \sum_{j=1}^{n_2} c_j V_{n_1, n_2-j}(u, t). \end{aligned} \tag{8}$$

The preceding set of equations must be solved under the following boundary conditions at $u = 0$,

$$\begin{aligned} P_{n_1, n_2}^{(1)}(0, t) &= (1 - r) \int_0^\infty P_{n_1+1, n_2}^{(1)}(u, t) \mu_1(u) du + r \int_0^\infty P_{n_1, n_2}^{(1)}(u, t) \mu_1(u) du \\ &+ \int_0^\infty P_{n_1+1, n_2}^{(2)}(u, t) \mu_2(u) du + R_{n_1+1, n_2}^{(1)}(t) \eta_1 + \lambda_p c_{n_1+1} I_{0, n_2}(t) \\ &+ \int_0^\infty V_{n_1+1, n_2}(u, t) \gamma(u) du + \int_0^\infty R_{n_1+1, n_2}^{(2)}(u, t) \eta(u) du, \end{aligned} \tag{9}$$

$$\begin{aligned} P_{0, n_2}^{(2)}(0, t) &= \lambda_o c_{n_2+1} I_{0,0}(t) + \int_0^\infty I_{0, n_2+1}(u, t) \beta(u) du + \sum_{i=1}^{n_2} \lambda_o C_i(u, t) \\ &+ \int_0^\infty I_{0, n_2+1-i}(u, t) du + \lambda_o q \int_0^\infty P_{0, n_2}^{(2)}(u, t) \mu_2(u) du, \end{aligned} \tag{10}$$

$$R_{n_1, n_2}^{(2)}(0, t) = \int_0^\infty D_{n_1, n_2}(u, t) du, \tag{11}$$

$$D_{n_1, n_2}(0, t) = \alpha_2 \int_0^\infty P_{n_1-1, n_2}^{(1)}(u, t) du + \alpha_2 \int_0^\infty P_{n_1, n_2}^{(2)}(u, t) du, \tag{12}$$

$$V_{0,0}(0, t) = (1-r) \int_0^\infty P_{0,0}^{(1)}(u, t) \mu_1(u) du + \int_0^\infty P_{0,0}^{(2)} \mu_2(u)(u, t) du \\ + \int_0^\infty R_{0,0}^{(2)}(u, t) \eta_2(u) du + R_{0, n_2}^{(1)}(t) \eta_1 + p \int_0^\infty V_{0,0}(u, t) \gamma(u) du, \tag{13}$$

$$V_{n_1, n_2}(0, t) = 0, \tag{14}$$

$$I_{0, n_2}(0, t) = (1-r) \int_0^\infty P_{0, n_2}^{(1)}(u, t) \mu_1(u) du + \int_0^\infty P_{0, n_2}^{(2)} \mu_2(u)(u, t) du \\ + \int_0^\infty R_{0, n_2}^{(2)}(u, t) \eta_2(u) du + R_{0, n_2}^{(1)}(t) \eta_1 \int_0^\infty V_{0, n_2}(u, t) \gamma(u) du. \tag{15}$$

$$P_{n_1, n_2}^{(1)}(0) = P_{n_1, n_2}^{(2)}(0) = D_{n_1, n_2}(0) = V_{n_1, n_2}(0) = R_{n_1, n_2}^{(1)}(0) = R_{n_1, n_2}^{(2)}(0) = 0, \\ I_{0,0} = 1, I_{0, n_2}(0) = 0, \text{ for } n_2 \geq 1 \text{ are the initial conditions.} \tag{16}$$

The PGF defined as,

$$I(u, t, z_0) = \sum_{n_2=1}^\infty z_p^{n_2} I_{0, n_2}(u, t); \quad A(u, t, z_p, z_0) = \sum_{n_1=0}^\infty \sum_{n_2=0}^\infty z_0^{n_1} z_p^{n_2} A_{n_1, n_2}(u, t); \\ A(u, t, z_p) = \sum_{n_1=0}^\infty z_p^{n_1} A_{n_1}(u, t); \quad A(u, t, z_p) = \sum_{n_2=0}^\infty z_0^{n_2} A_{n_2}(u, t); \tag{17}$$

here $A = P^{(1)}, P^{(2)}, D, V, R^{(1)}, R^{(2)}$.

We derive the following equations by applying Laplace transforms to equations (1) to (15) along with (16) and (17).

$$\bar{I}_0(u, s, z_0) = \bar{I}_0(0, s, z_0) e^{-(s+\lambda_p+\lambda_o)u - \int_0^u \beta(t) dt}, \tag{18}$$

$$\bar{P}^{(1)}(u, s, z_p, z_0) = \bar{P}^{(1)}(0, s, z_p, z_0) e^{-\phi_1(s, z)u - \int_0^u \mu_1(t) dt}, \tag{19}$$

$$\bar{P}^{(2)}(u, s, z_p, z_0) = \bar{P}^{(2)}(0, s, z_p, z_0) e^{-\phi_2(s, z)u - \int_0^u \mu_2(t) dt}, \tag{20}$$

$$\bar{V}(u, s, z_p, z_0) = \bar{V}(0, s, z_p, z_0) e^{-\phi_2(s, z)u - \int_0^u \gamma(t) dt}, \tag{21}$$

$$\bar{D}(u, s, z_p, z_0) = \bar{D}(0, s, z_p, z_0) e^{-\phi_2(s, z)u - \int_0^u \xi(t) dt}, \tag{22}$$

$$\bar{R}^{(2)}(u, s, z_p, z_0) = \bar{R}^{(2)}(0, s, z_p, z_0) e^{-\phi_2(s, z)u - \int_0^u \eta_2(t) dt}. \tag{23}$$

where,

$$\psi_1(s, z) = s + \lambda_p + \lambda_o(1 - C(z_0)) + \alpha_1 + \alpha_2,$$

$$\psi_2(s, z) = s + \lambda_p + \lambda_o(1 - C(z_0)),$$

$$\psi_3(s, z) = s + \lambda_p + \lambda_o(1 - C(z_0)) + \eta_1,$$

$$\phi_1(s, z) = s + \lambda_p(1 - C(z_p)) + \lambda_o(1 - C(z_0)) + \alpha_1 + \alpha_2,$$

$$\phi_2(s, z) = s + \lambda_p(1 - C(z_p)) + \lambda_o(1 - C(z_0)),$$

$$\phi_3(s, z) = s + \lambda_p(1 - C(z_p)) + \lambda_o(1 - C(z_0)) + \eta_1,$$

$$\sigma_1(s, z) = s + \lambda_p(1 - C(g(z_0))) + \lambda_o(1 - C(z_0)) + \alpha_1 + \alpha_2,$$

$$\sigma_2(s, z) = s + \lambda_p(1 - C(g(z_0))) + \lambda_o(1 - C(z_0)),$$

$$\sigma_3(s, z) = s + \lambda_p(1 - C(g(z_0))) + \lambda_o(1 - C(z_0)) + \eta_1.$$

$$\bar{I}_0(0, s, z_0) = \frac{\left\{ \begin{aligned} & [(1 - \bar{V}(\sigma_2(s, z)))\bar{V}_{0,0} + (s + \lambda_p + \lambda_o)\bar{I}_{0,0} - 1] \\ & \bar{P}^{(2)}(0, s, z_0)[\bar{B}_2(\sigma_1(s, z)) + [\alpha_2\bar{R}^{(2)}(\sigma_2(s, z))\bar{D}(\sigma_2(s, z))] \\ & + \frac{\alpha_1\eta_1}{\sigma_3(s, z)} \left[\frac{1 - \bar{B}_2(\sigma_1(s, z))}{\sigma_1(s, z)} \right] \end{aligned} \right\}}{\left\{ [(1 - C(g(z_0)))\lambda_p \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right]] \right\}},$$

$$\bar{P}^{(2)}(0, s, z_0) = \frac{\left\{ \begin{aligned} & \lambda_o C(z_0)\bar{I}_{0,0}(s)1 - \lambda_p C(g(z_0)) \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right] \\ & - [\bar{I}_{0,0}(s)(s + \lambda_p + \lambda_o) - 1 + [(1 - \bar{V}(\sigma_2(s, z)))\bar{V}_{0,0}]] \\ & [\bar{M}(s + \lambda_p + \lambda_o) + C(z_0)\lambda_o \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right]] \end{aligned} \right\}}{\left\{ \begin{aligned} & (z_0 - \lambda_p q z_0) \left[\frac{1 - \bar{B}_2(\psi_1(s, z))}{\psi_1(s, z)} \right] [(1 - C(g(z_0)))\lambda_p \\ & \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right]] - [\bar{M}(s + \lambda_p + \lambda_o) + C(z_0)\lambda_o \\ & \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right]] [\bar{B}_2(\sigma_1(s, z)) [\alpha_2\bar{R}^{(2)}(\sigma_2(s, z))\bar{D}(\sigma_2(s, z))] \\ & + \frac{\alpha_1\eta_1}{\sigma_3(s, z)} \left[\frac{1 - \bar{B}_2(\sigma_1(s, z))}{\sigma_1(s, z)} \right]] \end{aligned} \right\}}, \tag{24}$$

$$\bar{P}^{(1)}(0, s, z_p, z_0) = \frac{\left\{ \begin{aligned} & \lambda_p [C(z_p) - C(g(z_0))] \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right] \bar{I}_0(0, s, z_0) (\bar{V}(\phi_2(s, z))) \\ & - \bar{V}(\sigma_2(s, z))\bar{V}_{0,0} + \bar{P}^{(2)}(0, s, z_0) + [\bar{B}_2(\phi_1(s, z)) - \bar{B}_2(\sigma_1(s, z))] \\ & + [\alpha_2\bar{R}^{(2)}(\phi_2(s, z))\bar{D}(\phi_2(s, z)) + \frac{\alpha_1\eta_1}{\phi_3(s, z)} \left[\frac{1 - \bar{B}_2(\sigma_1(s, z))}{\sigma_1(s, z)} \right]] \\ & - [\alpha_2\bar{R}^{(2)}(\sigma_2(s, z))\bar{D}(\sigma_2(s, z)) + \frac{\alpha_1\eta_1}{\sigma_3(s, z)} \left[\frac{1 - \bar{B}_2(\sigma_1(s, z))}{\sigma_1(s, z)} \right]] \end{aligned} \right\}}{\left\{ \begin{aligned} & [z_p - ((rz_p + (1 - r))\bar{B}_1(\phi_1(s, z)) + (z_p\alpha_2\bar{R}^{(2)}(\phi_2(s, z))) \\ & \left[\bar{D}(\phi_2(s, z)) + \frac{\alpha_1\eta_1}{\phi_3(s, z)} \left[\frac{1 - \bar{B}_1(\phi_1(s, z))}{\phi_1(s, z)} \right]] \right] \end{aligned} \right\}}. \tag{25}$$

Theorem.1 When the system is operating normally, experiencing a breakdown, going on a randomized vacation, delay time to repair, or being repair, the probability generating function of the number of customers in the relevant queue will be provided using Laplace transforms.

$$\bar{I}_0(s, z_0) = \bar{I}_0(0, s, z_0) \left[\frac{1 - \bar{M}(s + \lambda_p + \lambda_o)}{s + \lambda_p + \lambda_o} \right], \tag{26}$$

$$\bar{P}^{(1)}(s, z_p, z_0) = \bar{P}^{(1)}(0, s, z_p, z_0) \left[\frac{1 - \bar{B}_1(\phi_1(s, z))}{\phi_1(s, z)} \right], \tag{27}$$

$$\bar{P}^{(2)}(s, z_p, z_0) = \bar{P}^{(2)}(0, s, z_0) \left[\frac{1 - \bar{B}_2(\phi_1(s, z))}{\phi_1(s, z)} \right], \tag{28}$$

$$\bar{V}(s, z_p, z_0) = \bar{V}(0, s, z_0) \left[\frac{1 - \bar{V}(\phi_2(s, z))}{\phi_2(s, z)} \right], \quad (29)$$

$$\bar{D}(s, z_p, z_0) = \bar{D}(0, s, z_0) \left[\frac{1 - \bar{D}(\phi_2(s, z))}{\phi_2(s, z)} \right], \quad (30)$$

$$\bar{R}^{(2)}(s, z_p, z_0) = \bar{R}^{(2)}(0, s, z_p, z_0) \left[\frac{1 - \bar{R}^{(2)}(\phi_2(s, z))}{\phi_2(s, z)} \right]. \quad (31)$$

Proof: The following result is reached by applying the renewal theory's solution and solving the previous equations (26) to (31) with respect to u through integration.

$$\int_0^\infty [1 - H(u)] e^{-su} du = \frac{1 - \bar{h}(s)}{s}. \quad (32)$$

The LST of the $H(u)$ distribution function of a random variable is expressed by $\bar{h}(s)$. The following states' precise probability generating function results are as follows: $\bar{I}_0(s, z_0)$, $\bar{P}^{(1)}(s, z_p, z_0)$, $\bar{P}^{(2)}(s, z_p, z_0)$, $\bar{D}(s, z_p, z_0)$, $\bar{V}(s, z_p, z_0)$ and $\bar{R}^{(2)}(s, z_p, z_0)$ are obtained by using equation (26) to (31).

5. STEADY STATE ANALYSIS

Steady state analysis refers to the examination of a system's behavior once it has reached a stable condition where its key parameters remain relatively constant over time. In this state, the system's inputs and outputs balance out, resulting in a consistent and unchanging pattern of behavior. Steady state analysis is often used in various fields such as engineering, economics and physics to understand long-term behavior and performance characteristics of systems.

According to Tauberian property,

$$\lim_{s \rightarrow 0} s \bar{f}(s) = \lim_{t \rightarrow \infty} f(t).$$

The queue size's PGF is as follows, in spite of the system's current state:

$$W_q(z_p, z_0) = \frac{Nr(z_p, z_0)}{Dr(z_p, z_0)}, \quad (33)$$

$$Nr(z_p, z_0) = N_1(z)D_2(z)D_3(z)\phi_1(z)\phi_2(z)\phi_3(z) \left[\frac{1 - \bar{M}(\lambda_1 + \lambda_2)}{\lambda_1 + \lambda_2} \right] + N_2(z)D_1(z)D_3(z)$$

$$f_1(z)(1 - \bar{B}_2\phi_1(z)) + \left[\frac{1 - \bar{M}(s + \lambda_1 + \lambda_2)}{s + \lambda_1 + \lambda_2} \right] N_3(z)D_1D_2(1 - \bar{B}_2\phi_1(z))f_1(z),$$

$$Dr(z_p, z_0) = D_1(z)D_2(z)D_3(z)D_4(z)\phi_1(z)\phi_2(z)\phi_3(z),$$

where,

$$N_1(z) = \bar{V}_{00}[\bar{V}\sigma_2(z) - 1] - (\lambda_p + \lambda_o)\bar{I}_{00} + [\bar{B}_2\sigma_1(z) + (\alpha_2\bar{R}^{(2)}\sigma_2(z)\bar{D}\sigma_2(z) + \frac{\alpha_1}{\sigma_3(z)})\left[\frac{1 - \bar{B}_2\sigma_1(z)}{\sigma_1(z)}\right]],$$

$$D_1(z) = 1 - C(g(z_o))\lambda_p\left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o}\right],$$

$$N_2(z) = \lambda_o C(z_o)\bar{I}_{00}\left[1 - C(g(z_o))\lambda_p\left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o}\right]\right] + [1 - (\lambda_p + \lambda_o)\bar{I}_{00}$$

$$\bar{V}_{00}[\bar{V}\sigma_2(z) - 1]]\left[\bar{M}(\lambda_p + \lambda_o) - \lambda_o C(z_o)\left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o}\right]\right],$$

$$D_2(z) = [z_o - \lambda_o q z_o\left[\frac{1 - \bar{B}_2\psi_1(z)}{\psi_1(z)}\right]][1 - C(g(z_o))\lambda_p\left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o}\right]][[\bar{B}_2\sigma_1(z)$$

$$+ (\alpha_2\bar{R}^{(2)}\sigma_2(z)\bar{D}\sigma_2(z) + \frac{\alpha_1}{\sigma_3(z)})\left[\frac{1 - \bar{B}_2\sigma_1(z)}{\sigma_1(z)}\right]]$$

$$\left[\bar{M}(\lambda_p + \lambda_o) - \lambda_o C(z_o)\left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o}\right]\right],$$

$$N_3(z) = \lambda_p[C(z_p) - C(g(z_o))]\left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o}\right]\bar{I}_0(0, z_o) + \bar{V}_{00}[\bar{V}\phi_2(z) - \bar{V}\sigma_2(z)]$$

$$\left[\bar{B}_2\phi_1(z) - \bar{B}_2\sigma_1(z)(\alpha_2\bar{R}^{(2)}\phi_2(z)\bar{D}\phi_2(z) + \frac{\alpha_1}{\phi_3(z)})\left[\frac{1 - \bar{B}_2\phi_1(z)}{\phi_1(z)}\right]\right]$$

$$- (\alpha_2\bar{R}^{(2)}\sigma_2(z)\bar{D}\sigma_2(z) + \frac{\alpha_1}{\sigma_3(z)})\left[\frac{1 - \bar{B}_2\sigma_1(z)}{\sigma_1(z)}\right]\bar{P}^{(2)}(0, z_o),$$

$$D_3(z) = z_p - [((1 - r) + rz_p)\bar{B}_1\phi_1(z) + (\alpha_2\bar{R}^{(2)}\phi_2(z)\bar{D}\phi_2(z) + \frac{\alpha_1}{\phi_3(z)})\left[\frac{1 - \bar{B}_1\phi_1(z)}{\phi_1(z)}\right]]$$

$$\psi_1(z) = \lambda_p + \lambda_o(1 - C(z_o)) + \alpha_1 + \alpha_2,$$

$$\psi_2(z) = \lambda_p + \lambda_o(1 - C(z_o)),$$

$$\psi_3(z) = \lambda_p + \lambda_o(1 - C(z_o)) + \eta_1,$$

$$\phi_1(z) = \lambda_p(1 - C(z_p)) + \lambda_o(1 - C(z_o)) + \alpha_1 + \alpha_2,$$

$$\phi_2(z) = \lambda_p(1 - C(z_p)) + \lambda_o(1 - C(z_o)),$$

$$\phi_3(z) = \lambda_p(1 - C(z_p)) + \lambda_o(1 - C(z_o)) + \eta_1,$$

$$\sigma_1(z) = \lambda_p(1 - C(g(z_o))) + \lambda_o(1 - C(z_o)) + \alpha_1 + \alpha_2,$$

$$\sigma_2(z) = \lambda_p(1 - C(g(z_o))) + \lambda_o(1 - C(z_o)),$$

$$\sigma_3(z) = \lambda_p(1 - C(g(z_o))) + \lambda_o(1 - C(z_o)) + \eta_1.$$

6. STABILITY CONDITION

The stability requirement is a criterion that establishes whether a QS can manage incoming traffic without increasing indefinitely over time. A stable QS maintains consistent queue length and performance measurements over time, even with changing arrival rates.

We apply the normalising condition to determine $I_{0,0}$.

$$I_{0,0} + I_0I + P^{(1)}(1,1) + P_{II}(1,1) + V(1,1) + R_I(1,1) + D(1,1) + R^{(2)}(1,1) = 1. \quad (34)$$

$$I_{0,0} = \frac{\left\{ \begin{aligned} & [D_1(1,1)D_2(1,1)D_3'(1,1)(\alpha_1 + \alpha_2)\eta_1(\lambda_p + \lambda_o)(-E(X))] \\ & - \left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o} \right] N_1(1,1)D_3'(1,1)D_2(1,1) \\ & (\alpha_1 + \alpha_2)(\lambda_p + \lambda_o)(-E(X))\eta_1 \\ & (\lambda_p + \lambda_o)(-E(X)) + N_2(1,1)D_3'(1,1)D_1(1,1)f_1'(1,1) \\ & (1 - B_2(\alpha_1 + \alpha_2)) + N_3'(1,1)D_1(1,1)D_2(1,1) \\ & f_1'(1,1)D_3(1,1)(1 - B_1(\alpha_1 + \alpha_2)) \end{aligned} \right\}}{\left\{ D_1(1,1)D_2(1,1)D_3'(1,1)(\alpha_1 + \alpha_2)\eta_1(\lambda_p + \lambda_o)(-E(X)) \right\}}, \tag{35}$$

and the utilization factor is given by

$$\rho = \frac{\left\{ \begin{aligned} & \left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o} \right] N_1(1,1)D_3'(1,1)D_2(1,1) \\ & (\alpha_1 + \alpha_2)(\lambda_p + \lambda_o)(-E(X))\eta_1 \\ & (\lambda_p + \lambda_o)(-E(X)) + N_2(1,1)D_3'(1,1)D_1(1,1)f_1'(1,1) \\ & (1 - B_2(\alpha_1 + \alpha_2)) + N_3'(1,1)D_1(1,1)D_2(1,1) \\ & f_1'(1,1)D_3(1,1)(1 - B_1(\alpha_1 + \alpha_2)) \end{aligned} \right\}}{\left\{ D_1(1,1)D_2(1,1)D_3'(1,1)(\alpha_1 + \alpha_2)\eta_1(\lambda_p + \lambda_o)(-E(X)) \right\}}. \tag{36}$$

The steady state stability requirement for the model is $\rho < 1$.

where,

$$N_3'(1) = \lambda_p \left[\frac{1 - \bar{M}(\lambda_p + \lambda_o)}{\lambda_p + \lambda_o} \right] \bar{I}_0(0,1)[1 - E(X_1)]E(X),$$

$$D_3'(1) = 1 - [r\bar{B}_2(\lambda_1 + \lambda_2) + (\alpha_2(E(R^{(2)}) + E(D)) - \frac{\alpha_1}{\eta_1} - 1) \left[\frac{1 - \bar{B}_1(\alpha_1 + \alpha_2)}{(\alpha_1 + \alpha_2)} \right]],$$

$$f_1'(1) = [\eta_1 + \alpha_1 - \alpha_2\eta_1(E(R^{(2)}) + E(D))][-(\lambda_p + \lambda_o)E(X)],$$

7. PERFORMANCE ASSESSMENTS

Performance measures in QS are metrics used to evaluate and quantify various aspects of system behavior, efficiency and effectiveness. These measures help assess how well a QS is performing and provide insights into its operational characteristics.

The following is the expected queue size for PC and orbit size for OC

$$L_{q1} = \frac{d}{dz_p} W_q(z_p, 1)|_{z_p=1}, \tag{37}$$

$$L_{q2} = \frac{d}{dz_o} W_q(1, z_o)|_{z_o=1}. \tag{38}$$

Where,

$$L_{q1} = \frac{Dr_1''(1)Nr_1'''(1) - Dr_1'''(1)Nr_1''(1)}{3(Dr_1''(1))^2}, \tag{39}$$

$$L_{q2} = \frac{Dr_2''(1)Nr_2'''(1) - Dr_2'''(1)Nr_2''(1)}{3(Dr_2''(1))^2}, \tag{40}$$

The following is the expected waiting time for priority queue:

$$W_{q1} = \frac{Lq_1}{\lambda_p} \tag{41}$$

The following is the expected waiting time for orbit:

$$W_{q2} = \frac{Lq_2}{\lambda_o} \tag{42}$$

8. PARTICULAR CASES

Case 1: In the absence of priority queue, ordinary customers arrive individually without retrials, breakdowns, or push-out mechanisms. In this scenario, the model can be simplified of queue type $M/G/1$ with a general randomized vacation policy.

$$P_{II}(z) = \frac{I_{0,0}[1 - \bar{B}_2(\lambda_o(1 - z_o))]\{\bar{V}(\lambda_o(1 - z_o)) - (1 - z_o)(1 - p)\bar{V}(\lambda_o) - 1\}}{(1 - z_o)(1 - p)\bar{V}(\lambda_o)\{z - \bar{B}_2(\lambda_o(1 - z_o))\}},$$

$$V(z_o) = \frac{I_{0,0}[1 - \bar{V}(\lambda_o(1 - z_o))]}{(1 - p)(1 - z_o)\bar{V}(\lambda_o)}.$$

The aforementioned outcome bears resemblance to the findings of Chen et al. [23] albeit without the inclusion of a second optional service.

Case 2: In the absence of priority queue, ordinary customers arrive individually without any breakdowns or vacations. In such a scenario, this model can be simplified a RQ of type $M/G/1$ with abandoned customers.

$$I_0(z) = \frac{I_{0,0}z(1 - z_o)[1 - \bar{B}_2(\lambda_p(1 - z_o) + \lambda_oqz_o)][1 - \bar{M}(\lambda_o)]}{D(z_o)},$$

$$P_{II}(z) = \frac{I_{0,0}\bar{M}(\lambda_o)(1 - z_o)[1 - \bar{B}_2(\lambda_o(1 - z_o) + \lambda_oqz_o)]}{D(z_o)},$$

where,

$$D(z) = \bar{B}_2(\lambda_o(1 - z_o) + \lambda_oqz_o)\{(1 - z_o + qz_o)(z_o + (1 - z_o)\bar{M}(\lambda_o)) - z_o^2q\} - z_o(1 - z_o).$$

These findings align with the results reported by Krishna Kumar et al. [24].

9. NUMERICAL RESULTS

The numerical and graphical analyses of this model are covered in this section. We assumed that the distribution of service period, failure, repair and vacation period are all exponential.

Table 1: The impact of the priority arrival rate (λ_p)

λ_p	I_0	ρ	Lq_1	Wq_1	Lq_2	Wq_2
0.5	0.9096	0.0904	3.3151	4.2339	1.3073	0.6535
0.6	0.8464	0.1536	3.6657	4.4002	1.4388	0.7194
0.7	0.7722	0.2278	4.0216	4.5624	1.5698	0.7849
0.8	0.6847	0.3153	4.3764	4.7233	1.6996	0.8498
0.9	0.5809	0.4191	4.7243	4.8869	1.8275	0.9138
1.0	0.4573	0.5427	5.0592	5.0592	1.9526	0.9763
1.1	0.3090	0.6910	5.3756	5.2492	2.0737	1.0368
1.2	0.1299	0.8701	5.6679	5.4705	2.1890	1.0945

Table 1 exhibit that when an arrival rate (λ_p) for PQ escalates, then (L_{q_1}/L_{q_2}) and (W_{q_1}/W_{q_2}) also rises at $\lambda_o = 2, \alpha_1 = 2, \alpha_2 = 3, \mu = 4, \eta_1 = 5, \eta_2 = 8, \gamma = 20, p = 0.3, \xi = 15, \beta = 8, q = 0.4, r = 0.1$ and $\lambda_p = 0.5$ to 1.3.

Table 2: Impact of hard failure repair rate (η_2)

η_2	I_0	ρ	L_{q_1}	W_{q_1}	L_{q_2}	W_{q_2}
5.0	0.0435	0.9565	3.2910	3.2910	1.6708	0.8354
5.5	0.1807	0.8193	3.0328	3.0328	1.6612	0.8306
6.0	0.2742	0.7258	2.8290	2.8290	1.6507	0.8253
6.5	0.3421	0.6579	2.6642	2.6642	1.6392	0.8196
7.0	0.3937	0.6063	2.5283	2.5283	1.6267	0.8133
7.5	0.4344	0.5656	2.4145	2.4145	1.6128	0.8064
8.0	0.4673	0.5327	2.3177	2.3177	1.5975	0.7987
8.5	0.4944	0.5056	2.2346	2.2346	1.5804	0.7902
9.0	0.5173	0.4827	2.1623	2.1623	1.5612	0.7806

Table 2 exhibit that when an hard failure repair rate (η_2) escalates, then (L_{q_1}/L_{q_2}) and (W_{q_1}/W_{q_2}) also decreases at $\lambda_p = 1, \lambda_o = 2, \alpha_1 = 1, \alpha_2 = 5, \mu = 4, \eta_1 = 4, \gamma = 15, p = 0.4, \xi = 10, \beta = 12, q = 0.3, r = 0.1$ and $\eta_2 = 5.0$ to 10.0.

Table 3: Impact of ordinary arrival rate (λ_o)

λ_o	I_0	ρ	L_{q_1}	W_{q_1}	L_{q_2}	W_{q_2}
1.0	0.8423	0.1577	1.6669	3.3339	0.1541	0.1541
1.1	0.8402	0.1598	1.7448	3.4896	0.1954	0.1777
1.2	0.8371	0.1629	1.8228	3.6455	0.2440	0.2034
1.3	0.8329	0.1671	1.9008	3.8016	0.3005	0.2311
1.4	0.8275	0.1725	1.9790	3.9579	0.3655	0.2610
1.5	0.8204	0.1796	2.0572	4.1144	0.4396	0.2931
1.6	0.8116	0.1884	2.1355	4.2710	0.5237	0.3273
1.7	0.8004	0.1996	2.2139	4.4277	0.6182	0.3636

Table 3 exhibit that when an ordinary arrival rate (λ_o) for PQ escalates, then the (L_{q_1}/L_{q_2}) and the (W_{q_1}/W_{q_2}) also rises at $\lambda_p = 0.5, \alpha_1 = 0.8, \alpha_2 = 4, \mu = 3, \eta_1 = 3, \gamma = 17, p = 0.3, \xi = 10, \beta = 12, q = 0.2, r = 0.2, \eta_2 = 7$ and $\lambda_o = 1.0$ to 2.0.

We obviously follow the exponential distribution for service time, breakdown, repair and vacation time in graphical representations. Figures 2 - 4 illustrate the 2D graphs. (L_{q_1}, L_{q_2}) increases when the priority arrival rate (λ_p) increases, as demonstrates in Figure 2. Figure 3 demonstrates the behaviour of the queue sizes (L_{q_1}, L_{q_2}), which depends on the hard failure repair rate (η_2). The length of the queue grows as the soft failure rate improves. Figure 4 depicts the behaviour of the queue sizes (L_{q_1}, L_{q_2}), which is affected by the average customer arrival rate (λ_o).

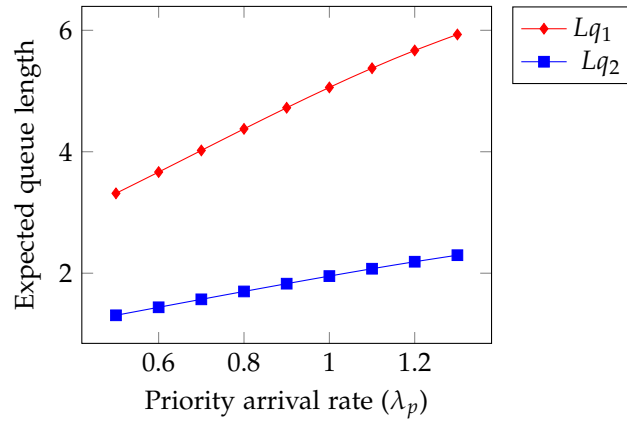


Figure 2: Lq Vs λ_p

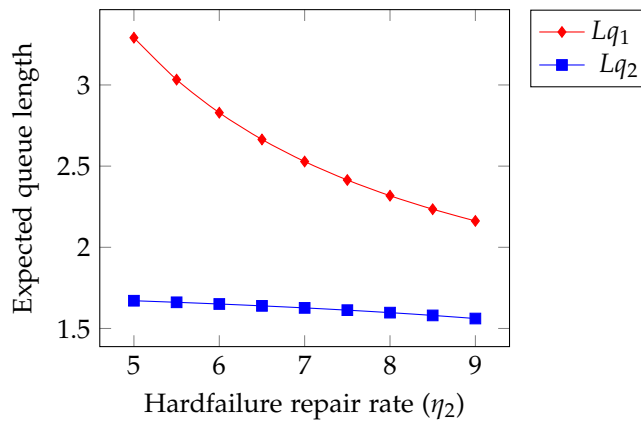


Figure 3: Lq Vs η_2

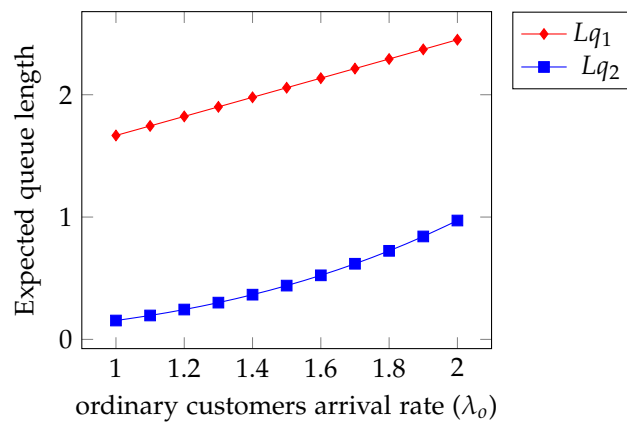


Figure 4: Lq Vs λ_o

10. CONCLUSION

In this research, we examined a single server retrial queueing system with non-preemptive priority service, immediate feedback, push out, differentiated breakdowns, delayed repair, randomized vacation. The analytical findings that are supported by numerical examples can be used to a wide

range of real-world situations to produce results. The supplementary variable technique is used to determine the PGFs for the number of users in the system when it is free, busy, and under repair. The system's and orbit's average queue lengths contain many expressions. The mean busy period and other significant system performance measures are obtained. The conditional decomposition law is finally demonstrated to be effective for this retrial queueing system. In real-world queueing scenarios, our queueing system is more flexible when dealing with real-time systems used by many industries.

REFERENCES

- [1] Ahuja, A., Anamika Jain and Madhu Jain. (2018). Finite population multi-server retrial queueing system with an optional service and balking, *International Journal of Computers and Applications*, 25(3):1-8.
- [2] Ammar, S.I. and Rajadurai, P (2019). Performance Analysis of Preemptive Priority Retrial Queueing System with Disaster under Working Breakdown Services, *Symmetry*, 11(3): 419.
- [3] Ayyappan, G and Thilagavathy, K. (2021). Analysis of $MAP(1), MAP(2)/PH/1$ Non-preemptive Priority Queueing Model Under Classical Retrial Policy with Breakdown, Repair, Discouragement, Single Vacation, Standby Server, Negative Arrival and Impatient Customers, *International Journal of Applied and Computational Mathematics*, 7:174-184
- [4] Ayyappan, G. and Deepa, T. (2021). Analysis of an Admission Control $M^{[X]}/G(a, b)/1$ Queue with Multiple Vacation, Restricted Re-service, Closedown and Setup Times, *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 91:35-47.
- [5] Ayyappan, G. and Gowthami, R. (2021). A $MAP/PH/1$ queue with Setup time, Bernoulli vacation, Reneging, Balking, Bernoulli feedback, Breakdown and repair, *Reliability: Theory & Applications*, 16(2):191- 221.
- [6] Azhagappan, A. and Deepa, T. (2023). Variant Impatient Behavior Of A Markovian Queue With Balking Reserved Idle Time And Working Vacation, *RAIRO Operations Research*, 54:783-793.
- [7] Begum, A. and Choudhury, G. (2022). Analysis of an $M/(G1, G2)/1$ Queue with Bernoulli Vacation and Server Breakdown, *International Journal of Applied and Computational Mathematics*, 9(6):1-22.
- [8] Choudhury, G. and Kalita, C.R. (2018). An $M/G/1$ queue with two types of general heterogeneous service and optional repeated service subject to server's breakdown and delayed repair, *Quality Technology & Quantitative Management*, 15(5):622-654.
- [9] Danilyuk, E., Plekhanov, A., Moiseeva, S.P and Janos Sztrik. (2022). Asymptotic Diffusion Analysis of Retrial Queueing System $M/M/1$ with Impatient Customers, Collisions and Unreliable Servers, *MDPI Journals*, 11(12):699.
- [10] D'Arienzo, M.P., Dudin, A.N., Dudin, S.A and Manzo, A. (2020). Analysis of a retrial queue with group service of impatient customers, *Journal of Ambient Intelligence and Humanized Computing*, 11:2591-2599.
- [11] D'Apice, C., Dudin, A., Dudin, S. and Manzo, R. (2022). Priority queueing system with many types of requests and restricted processor sharing, *Journal of Ambient Intelligence and Humanized Computing*, 14:12651-12662.
- [12] Gao, S., Zhang, J., and Wang, X. (2020). Analysis of a Retrial Queue With Two-Type Breakdowns and Delayed Repairs, *IEEE Access*, 8:172428-172442.
- [13] Jose, J.K. and Deepthi, V. (2021). $M/PH/1$ queueing model with re-servicing, *Communications in Statistics - Simulation and Computation*, 52(7):2865-2885.
- [14] Krishna Kumar, B., Rukmani, R., Thanikachalam, A. and Kanakasabapathi, V. (2018). Performance analysis of retrial queue with server subject to two types of breakdowns and repairs, *Operational Research*, 18:521-559.
- [15] Li, Q., Guo, P. and Yulan Wang (2020). Equilibrium analysis of unobservable $M/M/n$ priority queues with balking and homogeneous customers, *Operations Research Letters*, 48(5):674-681.

- [16] Nazarov, A., Janos Sztrik, Anna Kvach and Adam Toth. (2022). Asymptotic Analysis of Finite-Source $M/GI/1$ Retrial Queueing Systems with Collisions and Server Subject to Breakdowns and Repairs, *Methodology and Computing in Applied Probability*, 24:1503-1518.
- [17] Pavai Madheswari, S., Krishna Kumar, B. and Suganthi, P. (2019). Analysis Of $M/G/1$ Retrial Queues With Second Optional Service And Customer Balking Under Two Types Of Bernoulli Vacation Schedule, *RAIRO Operations Research*, 53:415-443.
- [18] Takagi, H. Queueing Analysis: *A Foundation of Performance Evaluation*, Vol. 1: Vacation and Priority Systems. Elsevier Science Ltd, North Holland, Amsterdam, New York, (1991).
- [19] Ke, J. C., Wu C. H. and Zhang, Z. G. (2010). Recent developments in vacation queueing models : A short survey. *International Journal of Operations Research*, 7(4):3-8.
- [20] Gao, S. and Yao, Y. (2014). An $MX/G/1$ queue with randomized working vacations and at most j vacations. *International Journal of Computer Mathematics*, 91(3): 368-383.
- [21] Gao, S., Liu B. and Wang, J. (2018). On a retrial queue with abandoned customers and multi-optional vacations. *International Journal of Computer Mathematics: Computer Systems Theory*, 3(3):177-195.
- [22] Krishnamoorthy, A., Joshua V. C. and Mathew, Ambily P. (2017). A retrial queueing system with abandonment and search for priority customers, In Vishnevskiy, V. M., Samouylov K. E. and Kozyrev, D. V., editors, *Distributed Computer and Communication Networks*. Springer International Publishing. Cham. pages 98-107.
- [23] Chen, Y., Lin X. W. Wei C. M. and Fan, Z. (2018). An $M/G/1$ queue with second optional service and general randomized vacation policy, In Cao, Bing-Yuan, editor, *Fuzzy Information and Engineering and Decision*. Springer International Publishing. Cham. pages 297-307.
- [24] Krishnakumar, B., Vijayakumar A. and Arivudainambi, D. (2002). An $M/G/1$ retrial queueing system with two-phase service and preemptive resume. *Annals of Operations Research*, 113(1):61-79.

PERSONALIZED FEATURES-BASED STRESS DETECTION WITH HYPERPARAMETER TUNING USING GENETIC ALGORITHM

Jigna Jadav¹, Uttam Chauhan²

1 Research Scholar, Computer Engineering, Gujarat Technological University, Ahmedabad, India.
jigna.j.jadav@gmail.com

2 Assistant Professor, Computer Engineering department, Vishwakarma Government Engineering
College Chandkheda, Ahmedabad, India. ug_chauhan@gtu.edu.in

Abstract

In recent years, there have been considerable improvements in how we keep track of mental health, especially with devices you can wear, which give us a better chance of spotting and dealing with problems like stress before they become serious. This research paper presents an innovative approach. Experimental validation uses a comprehensive dataset of 15 subjects working as multinational company employees. Heart Rate Variability(HRV) was obtained from wearable sensors using Apple Watch during working hours. We have calculated time, frequency and non-linear domains as well and added personalized features like a person's age, height, weight, etc. Recurrent Neural Network(RNN)and Long Short-Term Memory (LSTM)models are applied and get an accuracy of 87% and 90%, respectively. To enhance stress detection accuracy by optimizing hyperparameters using a genetic algorithm (GA) explicitly targeting the configuration of LSTM models. Key hyperparameters, including the number of units in the LSTM layer and the number of training epochs, are optimized to maximize stress detection accuracy. Model Through 5 generations of evolution, the GA identifies optimal hyperparameter settings of 45 units in the LSTM layer 49 epochs, significantly improving stress detection accuracy compared to baseline configurations. It gives 92 % accuracy with optimized hyperparameters. Analyzing recorded data, we observe that the time per training step decreases gradually, indicating efficient convergence during optimization. Simultaneously, stress detection accuracy steadily improves over epochs, showcasing the model's effectiveness in learning patterns from physiological data. So, This study provides insights into the practical application of genetic algorithms for hyperparameter optimization in healthcare contexts, contributing to advancements in personalized monitoring and intervention strategies for mental well-being.

Keywords: Genetic Algorithm, HRV, Hyperparameters, LSTM, Personalised Features, Stress Detection

I. Introduction

Stress is the feeling of being unable to handle a lot of mental or emotional pressure. Some things that cause stress are worrying about your career, meeting tight deadlines, feeling pressure from your peers, making bad choices, job demands, work pressure, health worries, etc. Acute stress is the most common type of stress. All the challenges usually cause it and demand everyone to deal with it constantly. Some clear signs are excessive sweating, headaches, trouble focusing, changes in hunger, a weakened immune system, difficulty sleeping, etc. Stress can have effects that last for a long time. Long-term worry can lead to health problems like high blood pressure, heart disease, and memory loss. Many trackers like the Empatica E4, the Apple Watch, the FitBit, and others are on the market. The suggested way to find stress for working employees is to create a wristband that predicts stress based on constant, real-time data from physiological sensors. The role of stress management systems in identifying the levels of tension that disrupt our socioeconomic lifestyle is crucial. According to the World Health Organization (WHO), one in every four

citizens is afflicted with stress, which is a mental health issue [1]. Human stress gives rise to psychological and socio-financial complications, impaired task performance, and unclear thinking. Relationship difficulties, melancholy, and, in extreme cases, suicide. This requires counselling to assist overwhelmed individuals in managing their stress. Although stress avoidance is unattainable, taking preventative measures can help to surmount it [2]. At this time, distressing states of depression (such as tension) can only be identified by medical and physiological specialists. One of the conventional approaches to stress detection involves using questionnaires [3]. This method depends entirely on the responses provided by the participants; individuals may exhibit trepidation when asked whether they are experiencing tension or feeling it every day. Automated stress detection reduces the likelihood of health complications and enhances societal welfare. This creates the conditions for developing a scientific instrument that automates the detection of stress levels in individuals via physiological signals.

Stress is also known as the flight-or-flight response, as it evolves as a survival mechanism, enabling people to react speedily to life-threatening or challenging situations. When met with a threat or challenge, an individual's body activates resources for self-protection. These resources either help face the situation or provide an expedited escape route. This flight-or-flight response is the reaction of the body's sympathetic nervous system that reacts to a stressor by producing larger quantities of chemicals like cortisol, adrenaline, and noradrenaline[4]. When you feel scared or threatened, your body enters an emergency mode. Your heart beats faster, your muscles tense up, your breathing gets faster, and your senses become sharper. This helps you react quickly in dangerous situations. It gives you more strength, energy, and focus to decide whether to fight or run away faster. It's like your body's way of helping you stay safe when you're in trouble.

The structure of this document is as follows. Section II provides an overview of the field's current state of the art. Section III provides an account of the research conducted in this study. Section IV presents the findings from the analysis of the Apple Watch dataset. The conclusions can be found in section V.

II. Related Work

Smartwatches have emerged as tools for stress detection and management [5] by measuring heart rate, heart rate variability (HRV), sleep quality, physical activity, and other stress-related variables. Real-time stress monitoring may provide immediate biofeedback to individuals and allow for early self-intervention [6]. How pervasive are smartwatches with stress detection among undergraduates? How effective are they as an unobtrusive way of providing biofeedback to students about their stress levels to lower their anxiety and increase their stress-management skills to be more effective in their Learning.

This exploratory pilot study examined the effect of smartwatches equipped with HRV sensors for stress detection on undergraduate students' anxiety. The study's research question was the following: Are smartwatches with stress detection sensors effective for supporting students in reducing their anxiety? A quasi-experimental pre-test post-test control group design was used. Thirteen students of an experimental group, who were self-selected, used the same commercially available smartwatch over 3-4 weeks and had access to their measured stress on a 24/7 basis. Nineteen students of a control group did not have access to smartwatches and used other means of stress management over the same period. Students' anxiety before and after the experiment was measured using a standardized instrument (GAD-7), which is based on participants' recollections of the frequency of experiencing specific anxiety symptoms over the last two weeks. GAD-7 scores range from 0 (no anxiety) to 21 (severe anxiety). The experimental group (M=6.00/21, SD=6.58) and the control group (M=8.18/21, SD=5.67) had "mild" anxiety levels before the study, based on GAD-7. An independent samples t-test compared students' anxiety before the experiment and established that the two groups were equivalent ($t_{27}=-0.97$, $p=0.341$). A slight, non-significant decrease in anxiety was observed in the experimental group from the pre-test (M= 6.00, SD= 6.57) to the post-test (M=5.67 SD= 3.26). On the contrary, the control group's anxiety increased from the pre-test (M= 7.84, SD= 5.43) to the post-test (M= 10.11, SD= 5.24), indicating a "moderate" level of anxiety post-intervention. An independent samples t-test ($t_{27}=-2.5$, $p=0.019$) for a comparison of students' post-test anxiety scores showed that the experimental

group had significantly lower anxiety ($M=5.67$, $SD=3.26$) compared to the control group ($M=10.06$, $SD=5.41$)[8].

Authors[9] propose a novel approach for predicting stress severity by measuring sleep phasic heart rate variability (HRV) using a smart device. This device can potentially be applied for stress self-screening in large populations. Using a Holter electrocardiogram (ECG) and a Huawei smart device, we conducted 24-h dual recordings of 159 medical workers working regular shifts. Based on photoplethysmography (PPG) and accelerometer signals acquired by the Huawei smart device, we sorted episodes of cyclic alternating pattern (CAP; unstable sleep), non-cyclic alternating pattern (NCAP; stable sleep), wakefulness, and rapid eye movement (REM) sleep based on cardiopulmonary coupling (CPC) algorithms. We further calculated the HRV indices during NCAP, CAP and REM sleep episodes using the Holter ECG and smart-device PPG signals. This exploratory pilot[10] study showed that smartwatches with stress detection sensors are somewhat effective in helping students reduce their anxiety. Without any structured intervention for stress management, students' anxiety may increase over time. The sample size and duration of the study were too small to allow for the generalizability of findings. More research is needed on how smartwatches that detect stress can be used either in conjunction with stress management interventions, such as mobile apps for supporting resilience, to maximize their effectiveness, or as additional ways of measuring stress, complimenting self-reported data in interventions that target stress management, to optimize Learning. The author developed a machine learning model to predict stress severity based only on the smart device data obtained from the participants and a clinical evaluation of emotion and stress conditions. Sleep phasic HRV indices predict individual stress severity with better CAP or REM sleep performance than in NCAP. Using the smart device data only, the optimal machine learning-based stress prediction model exhibited an accuracy of 80.3 %, sensitivity of 87.2 %, and 63.9 % for specificity. Sleep phasic heart rate variability can be accurately evaluated using a smart device and subsequently used for stress prediction.

Driving in urban areas can be challenging, and one can encounter acute stress. Collecting data on real roads without interfering with the driver is preferred to detect driver stress. A smartphone-based data collection protocol was developed to support a naturalistic driving study. Sixty-one participants drove on predetermined actual road routes, and driving information and physiological, psychological, and facial data were collected. The algorithm identified potentially stressful events based on the collected data. Participants classified these events as low, medium, or highly stressful by watching recorded videos after the experiment. These events were then used to train prediction models. The best model achieved an accuracy of 92.5% in classifying low/medium/highly stressful events. The contribution of physiological, psychological, and facial expression indices and individual profile information was evaluated. The method can be applied to visualize the geographical distribution of stressors, monitor driver behaviour, and help drivers regulate their driving habits [11].

III. Proposed Approach

We have prepared a data set using an optical heart sensor in the Apple Watch SE, which measures your heart rate and heart rhythm. Utilize the Breath application to calculate your stress with maximum precision. The Apple Watch has numerous capabilities that can be used to track stress levels. For instance, it features a heart rate monitor that can detect variations in the wearer's heart rate and heart rate variability, which can signal stress levels. A breathing app for the Watch also leads users through breathing exercises to lower stress. The gadget also monitors sleep patterns, physical activity levels, and other health indicators that may assist in pinpointing stress origins and offer insights into general well-being. It's crucial to remember that these features shouldn't be used to diagnose or treat any medical conditions and aren't intended to replace expert medical advice.

The study involved 15 employees from a multinational company working as developers, who were observed over ten working days. Raw data was collected using Apple Watch SE devices worn on the participants' non-dominant wrists. Participants performed a specific gesture (double tapping) with their non-dominant hand to ensure accurate data collection, generating a characteristic pattern in the

acceleration signal for data synchronization. The Apple Watch SE offers various health monitoring features suitable for the study to analyze Heart Rate under three distinct mental health conditions:

- Stress condition: During company meetings or instances of sudden extra work.
- Daily work condition: Routine work activities during regular working hours.
- No stress condition: Periods of relaxation or when no work-related tasks were being performed.

By examining Heart Rates across these conditions, the study aimed to understand how stress impacts physiological responses during work, contributing to a deeper understanding of employee well-being in multinational workplace settings.

It mainly focuses on how stress impacts employees' physiological responses during work hours, contributing to understanding employee well-being in workplace environments.

The raw data collected for Heart Rate Bit per minute, initially in XML format, was converted into CSV format for analysis. Gathering data in real-life contexts remains uncommon due to challenges such as limited context and reliance on self-reported information. Real-world data collection possesses both advantages and challenges. While it maintains ethical constraints and context awareness, it lacks a clear ground truth and introduces noisy data. Investigated HRV in real-world scenarios and highlighted its small relationship with stress compared to controlled lab settings[12]. The heart rate is extracted using The Apple Watch and the Health application on your paired iPhone. It can accurately measure your heart rate.

3.1. Feature selection

Core functionality involves iterating through XML files from an Apple Health export, identifying and extracting heart rate data within a specified date range. Each 'Record' element is examined to isolate valid heart rate measurements by cross-referencing their timestamps. From these values, RR intervals—the temporal gaps between consecutive heartbeats—are calculated. These intervals serve as the foundation for HRV analysis, which includes addressing missing values and performing both frequency and time domain analyses using an HRV analysis library. The process uses the `xml.etree.ElementTree` module for XML parsing, and incorporates the Malik rule[13] via the `hrvanalysis.remove_ectopic_beats` function to clean the RR interval data of ectopic beats.

Time domain features quantify RR interval variability, revealing insights into heart rate fluctuations over specific time spans. The Mean NN Interval (Mean NNI) portrays the average duration between successive normal heartbeats [14]. The Standard Deviation of NN Intervals (SDNN) characterizes overall RR interval variability, indicative of autonomic modulation. The Root Mean Square of Successive Differences (RMSSD) reflects short-term variability with parasympathetic sensitivity [15]. The Percentage of NN50 Intervals (pNN50) gauges parasympathetic influence by identifying RR intervals differing by over 50 ms. Frequency domain analysis dissects HRV into frequency bands. Low Frequency (LF) power signifies both sympathetic and parasympathetic activity, whereas High Frequency (HF) power primarily denotes parasympathetic modulation [16]. The LF/HF ratio quantifies sympathetic-parasympathetic balance [17].

The non-linear analysis captures intricate patterns. Sample Entropy (SampEn) gauges HRV complexity based on pattern repetition. Poincaré plots visually explore RR interval relationships, providing insights into autonomic dynamics [18].

Additional PhysioBank, PhysioToolkit, and PhysioNet furnish resources for physiological signal access and analysis. Advanced HRV analysis methods are exhaustively covered, offering insights into diverse techniques [19].

So, Time-frequency, Frequency Domain and Non-linear domain features are calculated to understand physiological responses to stress. These features can provide additional insights into the dynamics and patterns of the stress response, enabling more accurate and personalized stress detection algorithms and systems. We divided all experimental data into 60-second windows and independently calculated each

window's accuracy. We selected eight features because they are the most relevant Features. Twenty-two features have been extracted from 149,000 records collected from 15 subjects. These features encompass various aspects of heart rate variability (HRV).

IV. Experimental Results and Discussion

We Have used the architecture of the LSTM model [20] using the Keras API provided by TensorFlow[21]. The model consists of an LSTM layer with 128 units, a fully connected (Dense) layer with three output units (matching the number of classes) and a softmax activation function. The model is compiled with the Adam optimizer and categorical cross-entropy loss function, which is suitable for multiclass classification. The experimental result is presented in Figure 1, where we compare our proposed LSTM model with existing approaches in stress detection studies. In our experimental setup, we achieved an overall accuracy of 88% by applying the LSTM model, considering the time-series property inherent in the data. Notably, neither of the previously referenced authors leveraged the time-series property to obtain their results.

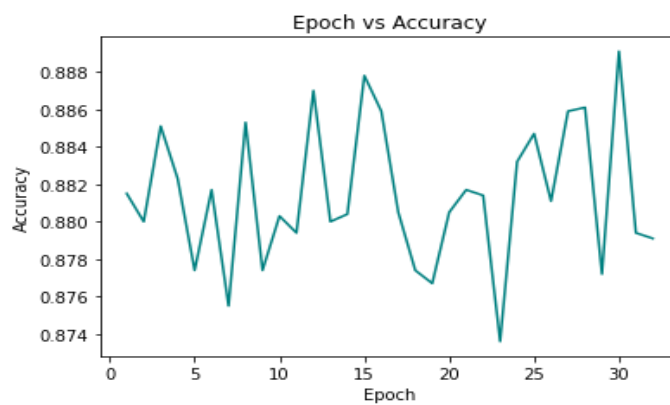


Figure 1: Epochwise Accuracy Plot

4.1. Personalized Features

As different users have relatively different responses to stress conditions, examining the individuals' heart rate variability ranges, the dataset and machine learning model should be designed carefully. We have added personalized features as defined in Figure 2, like a person's gender, age, country, height, body mass, resting heart rate, VO2MaxL/min·kg measures the maximum amount of oxygen a person can use during intense exercise, relative to body weight. It's expressed in millilitres per minute per kilogram and



Figure 2: Personalized Features

It is a key indicator of cardiovascular fitness and aerobic endurance. Walking Heart Rate Average refers to the average heart rate observed during walking activities. It helps assess cardiovascular health at a moderate intensity, with lower averages suggesting better fitness and heart efficiency. So, eight personalized features and other 22-time domain, frequency domain and non-linear features are added.

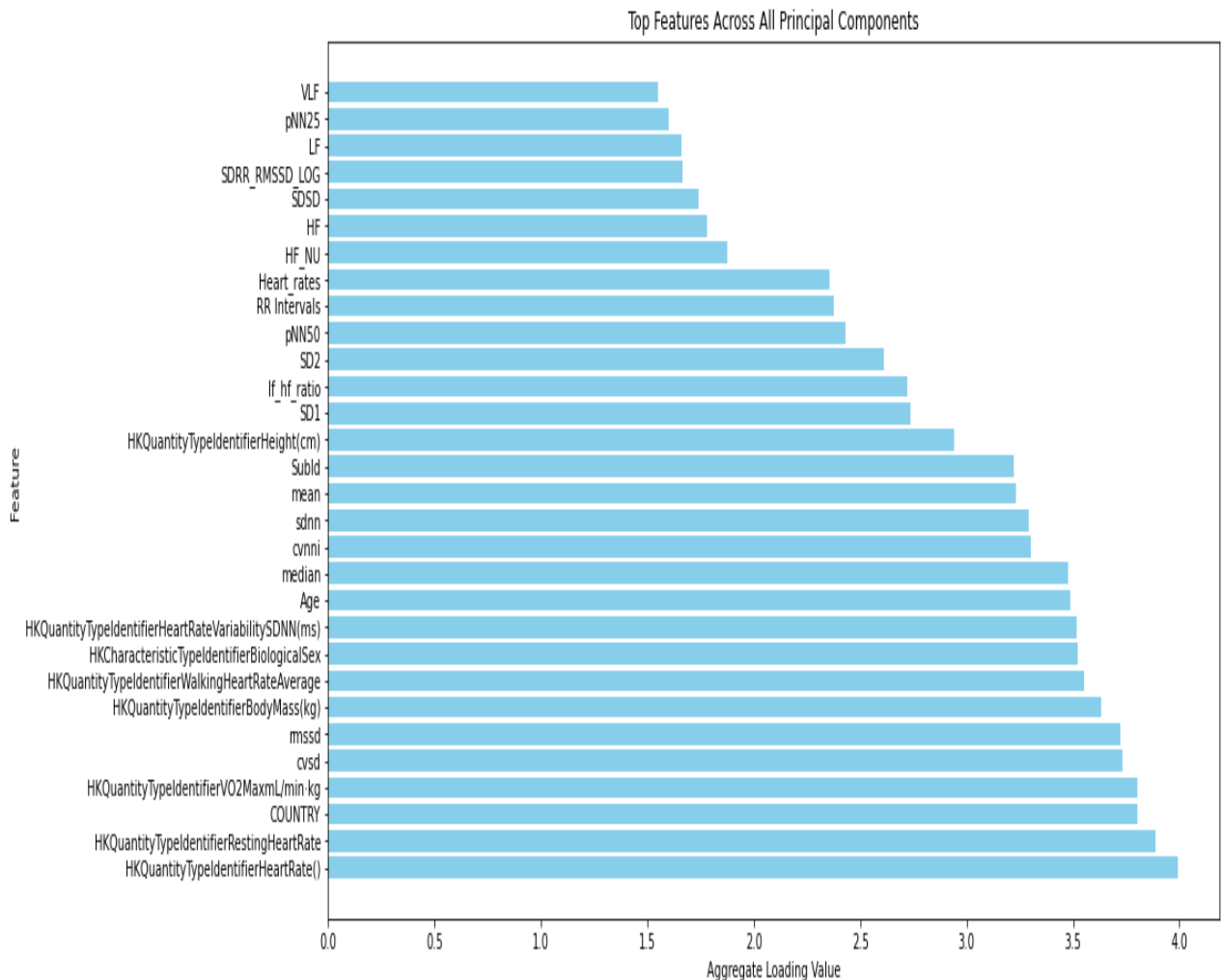


Figure 3: Features According to PCA

Figure 3 indicates the most influential features by examining the Aggregate values associated with each principal component. These ratios indicate the proportion of total variance in each component's dataset. A heatmap of the correlation matrix is created to visualize the relationships and dependencies between these features.

This heatmap illustrates in figure 4 the strength and direction of linear relationships between each pair of features. Correlation values range from -1 to 1, where 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no linear correlation. By examining the heatmap, we can identify which features are strongly correlated, either positively or negatively, which can provide insights into potential redundancies or synergies among the features. Personalized features are tailored to individual characteristics or behaviours, while time domain features typically involve statistical measures like mean, variance, and standard deviation calculated over time-series data. Frequency domain features

involve transformations like Fourier transforms to analyze the frequency components of the data. Non-linear features capture more complex, non-linear relationships in the data that are not apparent through linear analysis alone.

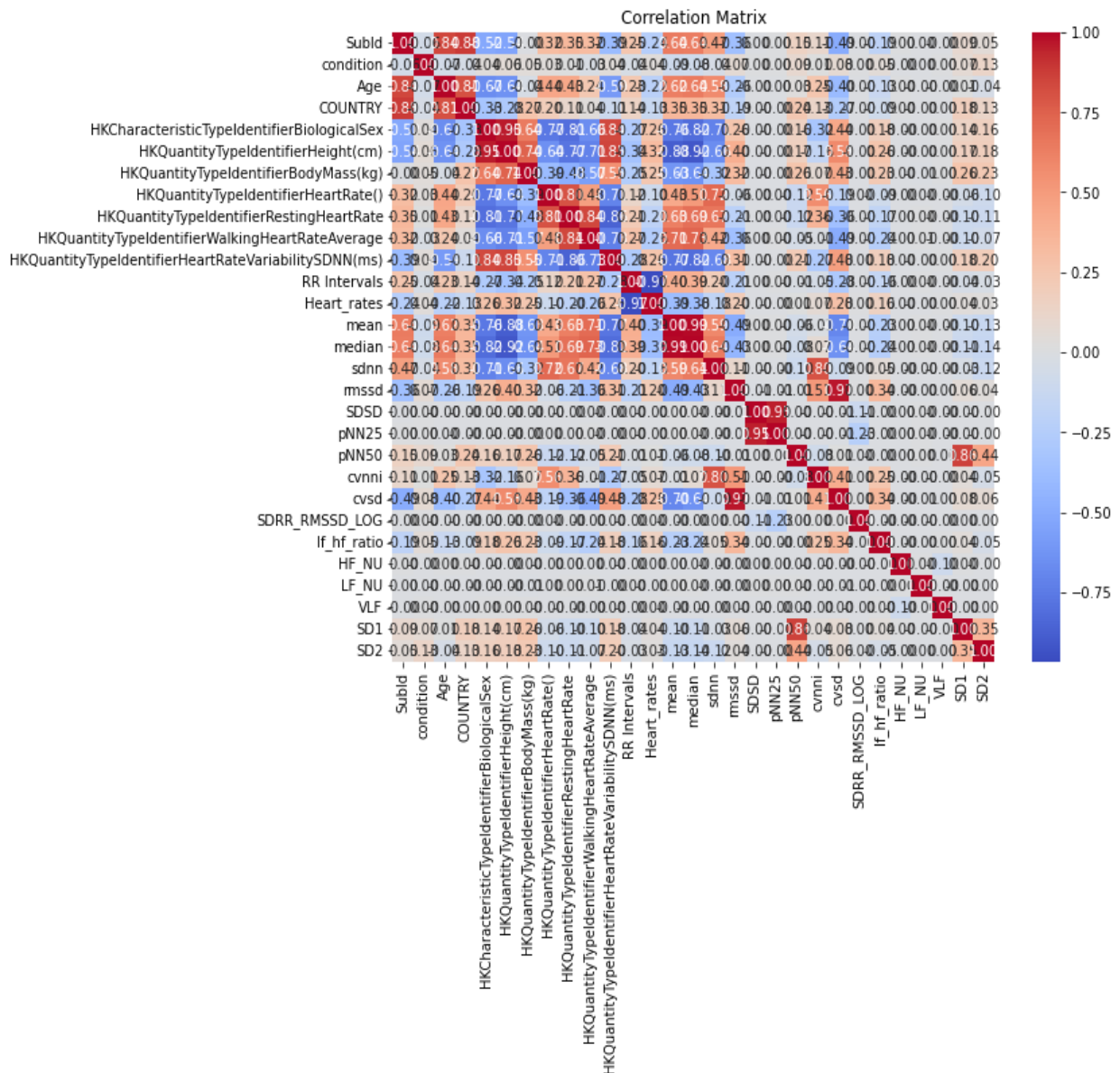


Figure 4: Heatmap of Correlation of Features

The indices of the most impactful features are identified by sorting these ratios in descending order. The actual feature names are then extracted from the original dataset columns.

The Long Short-Term Memory (LSTM) network, a type of recurrent neural network (RNN), was introduced by [22] in 1997 to address real-world time-series problems. LSTM networks have been shown to effectively learn long-term dependencies and overcome issues such as vanishing and exploding gradients [23]. The analysis involving LSTM (Long Short-Term Memory) and RNN (Recurrent Neural Network) models on a dataset of heart rate variability (HRV) data from an Apple Watch effectively preprocesses, trains, and evaluates both models. After preprocessing steps that include label encoding,

one-hot encoding, and standardizing the data, the models are reshaped for LSTM and RNN inputs. Two distinct neural network architectures are employed: an LSTM model and an RNN model, each designed with a recurrent layer comprising 128 units followed by a dense output layer with softmax activation to handle multiclass classification. Both models are compiled using the Adam optimizer and categorical cross-entropy loss, emphasizing their suitability for sequence data processing. The models are trained using a StratifiedShuffleSplit approach, ensuring balanced class distribution in both training and testing datasets. This setup allows for robust training over 32 epochs with a batch size of 32. Upon evaluation, the LSTM model demonstrates a superior performance with an accuracy of 90%, compared to the RNN model, which achieves 87% accuracy. Figure 5 presents the training and validation accuracy of LSTM and RNN models over a series of epochs. The plot is designed to illustrate how well each model learns from the training data and generalizes it to the validation data.

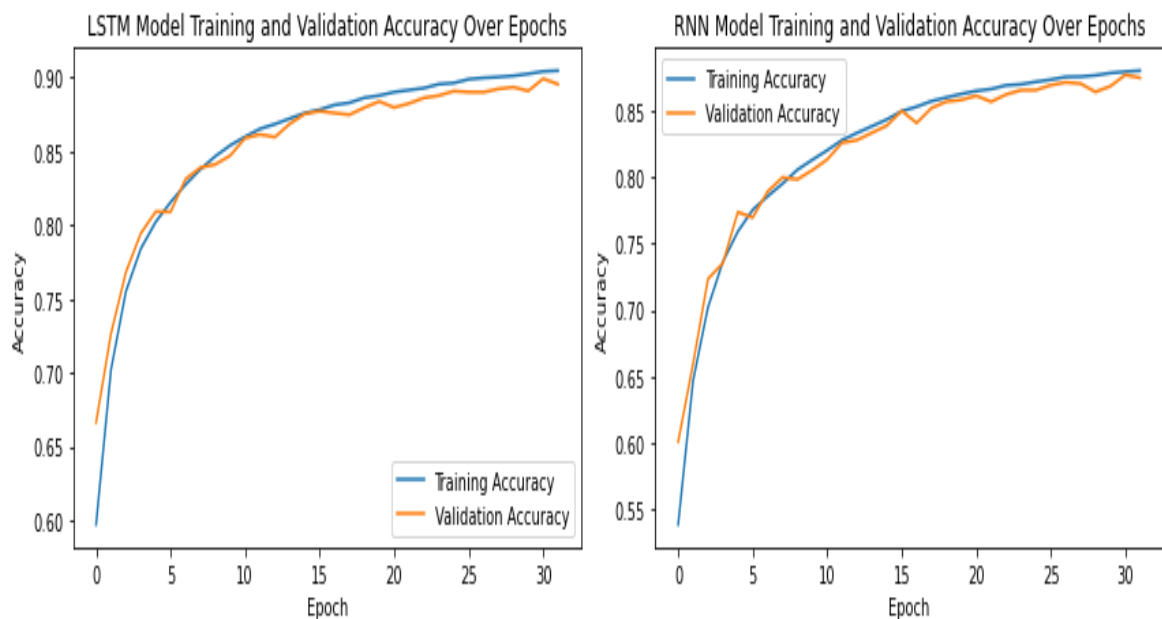


Figure 5: LSTM & RNN Training and Validation Accuracy

Visual analysis through various plots—accuracy and loss graphs over epochs and confusion matrices—provides deeper insights into each model's performance. The accuracy plots confirm the LSTM's slightly better capability than RNN in capturing and leveraging long-term dependencies, which is crucial for the temporal dynamics inherent in HRV data from wearable devices, as shown in figure 6. These detailed evaluations and visual representations are critical for understanding the model. Dynamics, guiding further improvements in model architecture or training strategies for enhanced performance in medical data analysis tasks.

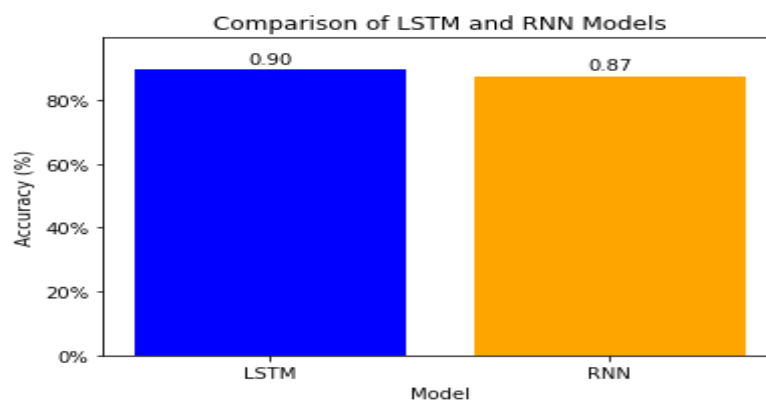


Figure 6: Comparison of LSTM & RNN

4.2. Genetic Algorithm

The genetic algorithm (GA) operates on the principle of evolution, where individuals with superior traits have a higher chance of survival and pass on their characteristics to the next generation. Each individual in the population represents a set of hyperparameters, with its genetic makeup determining the specific values of these parameters. The algorithm seeks to optimize these hyperparameters for a given problem through selection, crossover, and mutation. GA continually refines the population through these mechanisms, gradually converging towards optimal hyperparameter values. By iteratively selecting, recombining, and mutating individuals, the algorithm explores the hyperparameter space to discover configurations that maximize performance. Recent studies have shown the effectiveness of GA for optimizing LSTM hyperparameters in various applications, such as COVID-19 dataset classification [24]. In GA-based hyperparameter optimization, each hyperparameter is analogous to a gene within an individual's chromosome. The population encompasses a range of potential parameter values, and the fitness function evaluates how well a set of parameters performs. By selecting individuals with high fitness values, the algorithm ensures that favourable traits are carried over to subsequent generations for sepsis prediction [25]. Researchers [26] introduce a straightforward genetic algorithm method for hyperparameter tuning in a standard language model. This approach efficiently optimizes the parameters without relying on exhaustive search techniques. GA continually refines the population through these mechanisms, gradually converging towards optimal hyperparameter values. By iteratively selecting, recombining, and mutating individuals, the algorithm explores the hyperparameter space to discover configurations that maximize performance and also provides a robust mechanism for handling the complex hyperparameter space.

GA-based hyperparameter optimization, each hyperparameter is analogous to a gene within an individual's chromosome. The population encompasses a range of potential parameter values, and the fitness function evaluates how well a set of parameters performs. By selecting individuals with high fitness values, the algorithm ensures that favourable traits are carried over to subsequent generations. Crossover involves combining genetic material from two individuals to create offspring with a mix of their traits, while mutation introduces random changes to individual genes, promoting diversity in the population. GA continually refines the population through these mechanisms, gradually converging towards optimal hyperparameter values. By iteratively selecting, recombining, and mutating individuals, the algorithm explores the hyperparameter space to discover configurations that maximize performance.

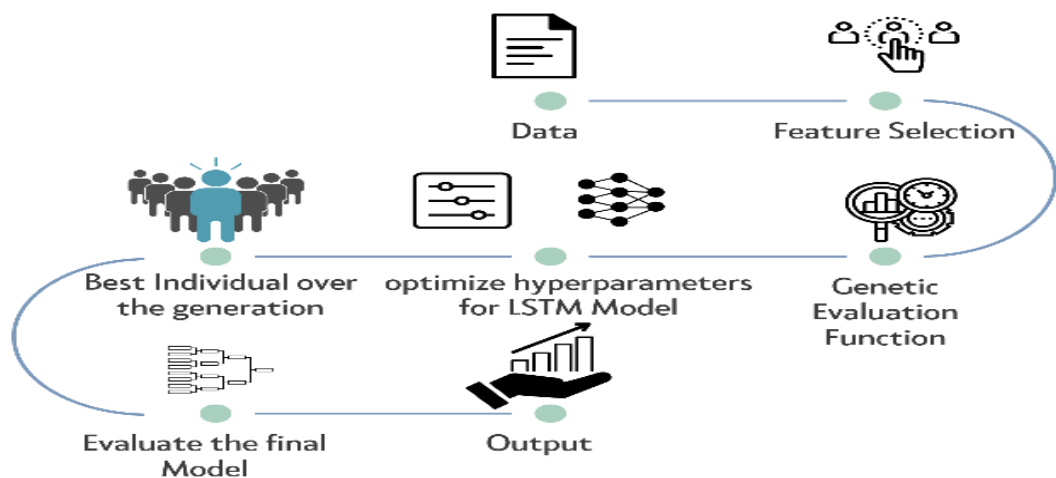


Figure 7: Flow of Hyperparameter tuning With GA

Figure 7 shows a flowchart of a process involving a genetic algorithm to optimize hyperparameters in a sequence model that is LSTM. The data that is relevant to the model should be collected or aggregated. After that, feature selection, which means feature selection, should be done from the data that will be input into the LSTM model learning. A genetic algorithm is applied to learn the optimal hyperparameters. From the genetic algorithm, an evaluating function judges the fitness of each approach in its ability to solve a problem; each individual, usually referred to with a hyperparameter set, is referred to as an individual. The result of the evaluating functions is used to adjust the hyperparameters using GA. The GA optimizes the hyperparameters to maximize the fitness function, which is done iteratively and without reinvention by simulating the process of natural selection. The Individual over the Generation: It says the best individual is used to highlight the best hyperparameter set found over all the generations. With the results of the best hyperparameters, the LSTM model is finalized and evaluated, and this has been done on a test dataset to check accuracy. So, this flowchart describes the whole process, from preparing the data to hyperparameter tuning, and then outputs are produced. The interesting aspect is the close relationship between the output and the final model to be evaluated.

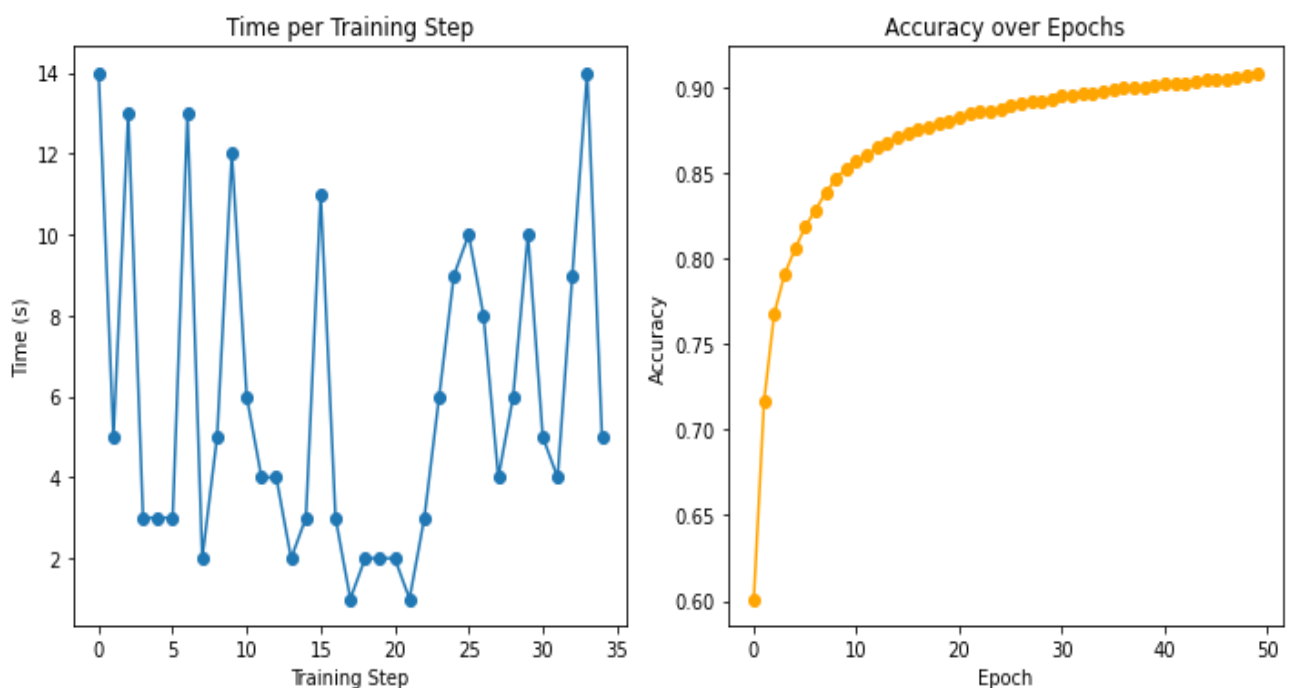


Figure 8. Hyperparameter Tuning setting

Figure 8 shows hyperparameter tuning in that the left plot shows each training step's time (in seconds). The times appear to vary significantly, ranging from as low as 1 second to as high as 14 seconds per step. There is no clear trend in the data, indicating that the variation in step times may be due to different computational demands at each step or other varying factors during the training process. The right plot shows the accuracy of the model over several epochs, plotted in orange. The accuracy starts at around 60% and steadily increases, reaching above 90% in later epochs. The plot shows a typical learning curve where the model initially improves rapidly before the gains in accuracy begin to diminish, suggesting that the model is approaching its performance limit. So, after setting the hyperparameter at epoch 49, we get 92 % accuracy.

V. Conclusion and Future Work

By employing HRV data from Apple Watches worn by employees during work, the study bridges the gap between controlled laboratory conditions and the variability in everyday environments. This approach enhances the model's applicability in real-life scenarios, providing insights into its practical deployment. The study's focus on feature selection and incorporating personalized features such as age, gender, and physical metrics further tailors stress detection to individual profiles, potentially increasing the model's sensitivity and accuracy in diverse populations. We presented an innovative approach to enhancing stress detection accuracy using wearable devices, focusing on optimizing LSTM models with GA. By employing a comprehensive dataset from 15 multinational company employees, we gathered HRV data through Apple Watches during working hours and calculated time, frequency, and non-linear domain features, supplemented by personalized characteristics such as age, height, and weight. The application of RNN and LSTM models yielded accuracy rates of 87% and 90%, respectively. Through GA optimization, we targeted key hyperparameters, including the number of units in the LSTM layer and the number of training epochs. The GA identified optimal settings of 45 units in the LSTM layer and 49 epochs, achieving a significant improvement in stress detection accuracy, reaching 92%. Our analysis indicated that the training time per step decreased, suggesting efficient convergence, while stress detection accuracy improved steadily over epochs, demonstrating the model's effectiveness in learning from physiological data. These findings underscore the potential of genetic algorithms in optimizing hyperparameters for LSTM models, contributing to advancements in personalized mental health monitoring and intervention strategies. As future work remedies can be suggested, and a wider population can be used for real-life experiments. Incorporating additional physiological and environmental data sources, such as skin conductance, body temperature, and ambient noise levels, could improve the robustness and accuracy of stress detection.

References

- [1] Communications, N. World health report. 2001. URL: http://www.who.int/whr/2001/media_centre/press_release/en/
- [2] Bakker, J., Holenderski, L., Kocielnik, R., Pechenizkiy, M., Sidorova, N.. Stess@ work: From measuring stress to its understanding, prediction and handling with personalized coaching. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium. ACM; 2012, p. 673–678.
- [3] Deng, Y., Wu, Z., Chu, C.H., Zhang, Q., Hsu, D.F.. Sensor feature selection and combination for stress identification using combinatorial fusion. *International Journal of Advanced Robotic Systems* 2013;10(8):306.
- [4] S. A. Singh, P. K. Gupta, M. Rajeshwari, and T. Janumala, "Detection of stress using biosensors," *Mater. Today*, vol. 5, no. 10, pp 21003_21010, 2018.
- [5] Jerath R, Syam M, Ahmed S. The Future of Stress Management: Integration of Smartwatches and HRV Technology. *Sensors*. 2023; 23(17):7314. <https://doi.org/10.3390/s23177314> 2nd ed., vol. 3, J. Peters, Ed. New York, NY, USA: McGraw-Hill, 1964, pp. 15–64.
- [6] Chalmers T, Hickey BA, Newton P, Lin CT, Sibbritt D, McLachlan CS, Clifton-Bligh R, Morley JW, Lal S. Associations between Sleep Quality and Heart Rate Variability: Implications for a Biological Model of Stress Detection Using Wearable Technology. *Int J Environ Res Public Health*. 2022 May 9;19(9):5770. doi: 10.3390/ijerph19095770. PMID: 35565165; PMCID: PMC9103972.
- [7] I. Nicolaidou (2024) CAN SMARTWATCHES WITH STRESS DETECTION LOWER STUDENTS' ANXIETY? AN EXPLORATORY PILOT STUDY USING WEARABLES, INTED2024 Proceedings, pp. 3080-3083.
- [8] Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006 May 22;166(10):1092-7. doi: 10.1001/archinte.166.10.1092. PMID: 16717171.

- [9] Fan J, Mei J, Yang Y, Lu J, Wang Q, Yang X, Chen G, Wang R, Han Y, Sheng R, Wang W, Ding F. Sleep-phasic heart rate variability predicts stress severity: Building a machine learning-based stress prediction model. *Stress Health*. 2024 Feb 27:e3386. doi: 10.1002/smi.3386. Epub ahead of print. PMID: 38411360.
- [10] Zhu, Lili & Spachos, P. & Ng, Pc & Yu, Yuanhao & Wang, Yang & Plataniotis, Konstantinos & Hatzinakos, Dimitrios. (2023). Stress Detection Through Wrist-Based Electrodermal Activity Monitoring and Machine Learning. *IEEE Journal of Biomedical and Health Informatics*. PP. 1-11. 10.1109/JBHI.2023.3239305.
- [11] Healey, Jennifer A, & Picard, R. W. (2008). Stress Recognition in Automobile Drivers [Data set]. *physionet.org*. <https://doi.org/10.13026/C2SG6B>.
- [12] I. Noura, A. Ben Abdallah, and M. H. Bedoui, "A Robust R Peak Detection Algorithm Using Wavelet Transform for Heart Rate Variability Studies," *Int. J. Electr. Eng. Informatics*, vol. 5, no. 3, pp. 270–284, Sep. 2013, doi: 10.15676/ijeii.2013.5.3.3.
- [13] Malik, M., et al. (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93(5), 1043-1065.
- [14] Thayer, J. F., & Brosschot, J. F. (2005). Psychosomatics and psychopathology: looking up and down from the brain. *Psychoneuroendocrinology*, 30(10), 1050-1058.
- [15] Mietus, J. E., et al. (2002). The pNNx files: re-examining a widely used heart rate variability measure. *Heart*, 88(4), 378-380.
- [16] Pomeranz, B., et al. (1985). Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology*, 248(1), H151-H153.
- [17] Tarvainen, M. P., et al. (2010). Advanced methods for heart rate variability analysis. In *The Handbook of Behavioral Medicine* (pp. 161-181). Springer .
- [18] Brennan, M., et al. (2001). Do existing measures of Poincaré plot geometry reflect non-linear features of heart rate variability? *IEEE Transactions on Biomedical Engineering*, 48(11), 1342-1347.
- [19] Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215-e220.
- [20] Fandango A. *Mastering TensorFlow 1. x: Advanced machine learning and deep learning concepts using TensorFlow 1. x and Keras*. Packt Publishing Ltd; 2018 Jan 22.
- [21] Gulli A, Kapoor A, Pal S. *Deep learning with TensorFlow 2 and Keras: regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API*. Packt Publishing Ltd; 2019 Dec 27.
- [22] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [23] Andersen, R. S., Peimankar, A., & Puthusserypady, S. (2019). A deep learning approach for real-time detection of atrial fibrillation. *Expert Systems with Applications*, 115, 465–473. <https://doi.org/10.1016/j.eswa.2018.08.011>
- [24] M. Parvizimosaed, M. Esnaashari, A. Damia and M. T. Paband, "Hyper-parameter Optimization of LSTM Network Using Genetic Algorithm and Q-Learning Algorithm for Classification of COVID-19 Dataset," 2023 9th International Conference on Web Research (ICWR), Tehran, Iran, Islamic Republic of, 2023, pp. 167-172, doi: 10.1109/ICWR57742.2023.10139161.
- [25] Nejedly P, Plesinger F, Viscor I, Halamek J, Jurak P. Prediction of sepsis using LSTM neural network with hyperparameter optimization with a genetic algorithm. In 2019 Computing in Cardiology (CinC) 2019 Sep 8 (pp. Page-1). IEEE.
- [26] Gorgolis N, Hatzilygeroudis I, Istenes Z, Gyenne LG. Hyperparameter optimization of LSTM network models through genetic algorithm. In 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA) 2019 Jul 15 (pp. 1-4). IEEE.

ANALYSIS OF SINGLE SERVER FEEDBACK RETRIAL QUEUE WITH BERNOULLI WORKING VACATION AND STARTING FAILURE

KEERTHIGA S¹ AND INDHIRA K*

^{1,*}Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore - 632 014, Tamil Nadu, India.
keerthiga.2020@vitstudent.ac.in, kindhira@vit.ac.in.

Abstract

The suggested queueing model describes a single-server feedback retrial queueing system with starting failure, Bernoulli working vacation and vacation interruptions. The server departs on a working vacation as soon as orbit is empty. During the working vacation period, the server provides a slower level of service. The supplementary variable method was utilized to determine the steady-state probability-generating functions for the system and its orbit. If there are consumers in the system at the end of each vacation, the server becomes idle and ready to serve new customers. The average busy time and the average busy cycle are presented as important system performance indicators. Additionally, the adaptive neuro-fuzzy interface system has compared the numerical results with the neuro-fuzzy results. Finally, particle swarm optimization (PSO) were utilized to obtain the best (optimal) cost for the system in this study. We have examined the convergence of these optimization strategies.

Keywords: Retrial queues, Feedback, Supplementary variable technique, Starting Failure and Working Vacation,ANFIS.

1. INTRODUCTION

In a queueing system (QS), queues involving continuous tries occur when a consumer comes and identify the server is occupied. The client is instructed to leave the service region and join a virtual area referred to as the 'orbit'. Subsequently, the customer within the orbit can make a service request after a period of time. In a vacation periods, the server halts its service entirely, becoming unavailable to the primary clients for a short duration, which is termed a "vacation." However, during the working vacation (WV) period, the server provides services to consumers, albeit at a reduced service rate. Also, the server's vacation may be ignored if customers arrive during the vacation period, and the server may resume operation in its regularly scheduled manner. It is known as the vacation interruption(VI) strategy. Major uses for this QS include delivering network services, online services, file transfer services, mail services and so on. A more realistic RQ with feedback happens in many real-world scenarios; for instance, in multiple-access telecommunications systems, where data returned as failures is forwarded again, it may be treated as a retrial queue with feedback.

1.1. Survey of Literature

In an $M/G/1$ retrial queue (RQ) with general retrial times, consumers who find the server busy join the orbit according to the first-come,first-served (FCFS) principle as studied by Gomez-Corral [1]. Such an instance occurs in certain communication protocols, in production lines at

stores, etc. The RQ has been extensively studied by Falin and Templeton [2], Artalejo and Corral [3], Artalejo [4], etc. Many authors have investigated a single server retrial queue (SSRQ) with WVs and VIs, including Zhang and Hou [16], Gao and Liu [6], Gao et al. [5], Zhang and Liu [22], and Rajadurai et al. ([7], [8], [9]). Mokaddis et al. [10] explored the $M/G/1$ retrial queue with Bernoulli feedback, Starting Failure (SF) and a single vacation (SV). Clients in orbit connect to the server via FCFS discipline, and an arbitrary distribution is assumed for the retry time. The server goes on vacation when there are no clients on the system. If the server comes back from vacation and there are no consumers, it waits for the first client to arrive on the system from the outside.

Krishna Kumar et al. [11] researched a RQ with feedback and a server exposed to SF, as well as a general stochastic decomposition rule for $M/G/1$ vacation models. Rajadurai [12] investigated a single server preemptive priority RQ based on Bernoulli working vacation (BWV) and VIs. Rajadurai et al. [13] explored a SSRQ system with BWV and VI. Performance indicators and analytical illustrations are provided. Jain and Kumar [26] analyzed bulk arrival general service RQ subject to balking, feedback and vacation interruption under multiple WV policy. Pazhani Bala Murugan and Keerthana [27] investigated an $M/G/1$ feedback RQ with WV and a waiting server. Keerthiga and Indhira [28] examined SSRQ with two phases of service for retrial customers. Agarwal et al. [29] discussed detection of optimal WV service rate for retrial priority G-queue with immediate Bernoulli feedback. Rachita Sethi et al. [17] researched a threshold-based repair facility for machining systems with a WV approach. WV was established to allow repairmen to offer service at a reduced rate as opposed to entirely discontinuing operations. The idea of F-policy is used to govern its arrival in the system. The implementation of a threshold N-policy to start the repair reduces the system's cost. Performance measurements are computed utilizing the 4th-order Runge-Kutta approach, and the numerical findings obtained are compared to the adaptive neuro fuzzy inference system (ANFIS).

Charu Bhargava and Madhu Jain [18] studied the "Modelling and Analysis of a Markovian multi server queue with an (e,d) SV procedure, server failures and repairs". Some stationary performance indicators are established after service completion using the matrix geometric technique. Additionally, the direct search method is utilised to estimate the best no. of idle, vacationing, and total no. of servers at the most affordable price. Also, the acquired numerical outcomes have been compared using a soft computing technique (SCT) based on an ANFIS. Radhika Agarwal et al. [19] analyzed the performance metrics that are used in improving service standards using the SVT and compared the analytical outcomes to the neuro fuzzy outcomes via the ANFIS (SCT). In addition, single and bi-objective minimization issues are explored with minimum attained via "PSO and a multi-objective GA" respectively.

In this research, we have extended the work of Rajadurai et al. [13] by including the ideas of feedback and SF. By using PSO, we have also performed a cost analysis of the model under consideration. Because the suggested solution improves repeatedly and the system gives us the best option that is feasible, this approach has gained a lot of reputation in recent years. When it comes to queueing analysis, this technique may be used to get productive outcomes, whether the goal is to save overall costs or maximize performance metrics. This framework aids in our analysis of various real-world queueing scenarios, allowing us to enhance the customer experience. To that extent, this article contributes. In areas with heavy traffic and congestion, this kind of project is highly pertinent and beneficial.

To the best of author's knowledge, there has been no previous research that has examined in this work. Therefore, to fill up this gap, in this article, we consider the feedback RQ with WV and VI subjected to server breakdown and repair. SVT has been used, and for some of the variables, a 3D graphical representation has also been provided. To attain optimal operating conditions, minimize expected costs, and maximize economic performance PSO a well-known meta-heuristic technique are applied. The aforementioned framework may be used in a wide variety of situations, including but not limited to: telephone switching, telecommunications, computer networks, online ticket booking centers, aviation traffic control, quality control procedures, and inspection testing of items. The purpose of this investigation is to estimate the queue length and orbit size dist., which will be implemented to calculate the system's performance metrics.

The following is an overview of our article: Section 2 provides a detailed discussion of the queueing paradigm. Section 3 specifically determines the system's steady state (SS) behavior and the queue length's PGF at a random epoch. Section 4 includes various substantial indicators of system behavior. Section 5 discussed particular cases. Sections 6 and 7 provide numerical outcomes and cost optimization. Finally, Section 8 provides a conclusion and overview of the study.

2. MODEL DESCRIPTION

A comprehensive explanation of this framework is given below:

The arrival process: New customers join the system from the outside, according to a Poisson process (PP), at a rate of μ .

The retrial process: According to FCFS discipline, when a customer visits while the server is occupied or unavailable, the consumer departs the service area and joins a group of blocked clients known as "orbit." The server appears to be accessible to the clients at the front of the orbit queue. Every customer's successive inter-retrial duration's are determined by an arbitrary probability distribution function (PDF) $B(x)$, with an associated density function (df) $b(x)$ and the "Laplace-Stieltjes transform" (LST) $\beta^*(\theta)$.

The service process: When a new or repeated consumers enters at the server while it's idle, the server promptly begins its regular service for the incoming customers. The service time follows a general dist., and PDF $C(x)$, a df $c(x)$, a LST $\alpha^*(\theta)$ & the 1st and 2nd moments are C_1 and C_2 .

The Bernoulli working vacation process: The server goes on a WV whenever the orbit is empty, and the duration of this vacation follows an exponential dist., with a specified parameter γ . If a consumer visits during vacation time, the server will continue to operate at a slower service rate. The WV time is a slower-paced operating period. In the event that any clients in the orbit reach the instant of service completion during the vacation time, the server will end the vacation and return to its normally busy state, which is known as VI. On the other hand, if there are no clients in the system at the completion of the vacation, the server rejoins the system and waits to serve a new client with prob. r_1 (SWV) or makes for another WV with prob. $r_2 = 1 - r_1$ (MWV). When a vacation is over and there are still consumers in its orbit, the server resumes usual operation. During the WV period, the service time is determined by a general random variable H_v with a dist., function $H_v(t)$, LST $H_v^*(\varphi)$ & the 1st and 2nd moments are h_1 and h_2 , respectively.

Feedback Procedure: After getting their normal services, dissatisfied customers have two options: they may either exit the system with probability $\bar{\omega} = (1 - \omega)$ or they can return to the orbit as unsatisfied clients and get a service again with probability ω .

Starting Failure: The customer will almost definitely start receiving service right away if the server is successfully activated. If the server is unable to start the service, the consumer exits the service area, enters the orbit, and repeats the request for the service after some time. The server is instantly repaired if a failure occurs. SF happens with prob. $\bar{\lambda}$ and successful service begins with prob. λ .

Repair Process: If the server fails to start, the repair process begins immediately. During the repair process, the server refuses to serve external or repeat consumers. Repair times have a distribution function $I(x)$ and a corresponding density function $i(x)$, and the first two moments are I_1 and I_2 , respectively.

The system's stochastic processes are considered to be independent of each other.

2.1. Practical justification of the recommended paradigm

The suggested scenario has beneficial applications in a telecommunications. For example, we investigate a communication system designed for making reservations at restaurants. Let us consider a scenario in which a restaurant uses a phone system to accept reservations and provide a range of other services. client can use this system to reserve a table for themselves. The supervisor who answers all calls is in charge of this phone system. The consumer is able to pick up the phone (leave the system) or inquire regarding event reservations, purchase tickets for an upcoming musical performance, etc. after reserving a table. A caller must reaffirm their reservations if there is a possibility of a misinterpretation stemming from an unclear network or other related difficulties (feedback). When the manager is occupied overseeing other areas of the restaurant, he is unable to answer calls (vacation mode). In these circumstances, the junior manager often serves, albeit somewhat more slowly (WV).In this stage, the supervisor returns right away (i.e., a vacation interruption happens) if there are any calls in the system after the phone call is over (at service completion). However, the supervisor continues to take care of other restaurant-related matters if no calls come in after completing his secondary work (vacation mode). It is likely that when a consumer calls, the line is busy and that the client will call back after some period of time (retrial). It is possible that during a phone conversation, a bad signal, inadequate network coverage, or a virus attack (SF) might occur, causing the client to lose service. Once the communication system's signal is repaired, it functions flawlessly.

3. SCRUTINY OF THE STEADY STATE PROBABILITIES

In steady state (SS), we presume that $B(0) = 0, B(\infty) = 1, C(0) = 0, C(\infty) = 1,$ and $\mathcal{H}_v(0) = 0, \mathcal{H}_v(\infty) = 1, \mathcal{I}(0) = 0, \mathcal{I}(\infty) = 1,$ are continuous at $\tilde{\varphi} = 0$. So that the function $\eta(\tilde{\varphi}), \zeta(\tilde{\varphi}), \kappa(\tilde{\varphi}),$ and $v(\tilde{\varphi}),$ are the hazard rates of the conditions (retrial, normal service, vacation and repair) are

$$\begin{aligned} \eta(\tilde{\varphi})d\tilde{\varphi} &= \frac{dB(\tilde{\varphi})}{1 - B(\tilde{\varphi})} \\ \zeta(\tilde{\varphi})d\tilde{\varphi} &= \frac{dC(\tilde{\varphi})}{1 - C(\tilde{\varphi})} \\ \kappa(\tilde{\varphi})d\tilde{\varphi} &= \frac{d\mathcal{H}_v(\tilde{\varphi})}{1 - \mathcal{H}_v(\tilde{\varphi})} \\ v(\tilde{\varphi})d\tilde{\varphi} &= \frac{d\mathcal{I}(\tilde{\varphi})}{1 - \mathcal{I}(\tilde{\varphi})} \end{aligned}$$

$$L(\xi) = \begin{cases} 0, & \text{if the server is free} \\ 1, & \text{if the server is active period} \\ 2, & \text{if the server is operative mode on WV period} \\ 3, & \text{if the server is on repair} \end{cases}$$

Thus, the state of the system $B^0(\xi), C^0(\xi), \mathcal{H}_v^0(\xi),$ and $\mathcal{I}^0(\xi)$ are required to construct a bivariate Markov process $\{N(\xi); \xi \geq 0\},$ where $L(\xi)$ belongs to the server stage $(0, 1, 2, 3)$ based on if the server is idle, typical operative period, slow service and repair time.

3.1. Ergodicity Condition

Let $\{\xi_\sigma; \sigma = 1, 2, \dots\}$ represent a series of epochs in which either a service time is reduced or completed. $U_\sigma = \{L(\xi_\sigma+), X(\xi_\sigma+)\}$ is a random vector sequence. The embedded Markov chain generated by the RQ system. Its state space is $S = \{0, 1, 2, 3\} \times \mathbb{N}.$

3.2. Theorem

The embedded Markov chain $\{U_\sigma; \sigma \in N\}$ is ergodic iff $\rho < \bar{B}(\mu)$ for our system to be stable, where $\rho = \lambda\mu C_1 + \bar{\lambda}(1 + \mu I_1) + \lambda\omega$.

3.3. System of governing equations

For the procedure $\{N(\xi), \xi \geq 0\}$, we specify the prob., $\phi_0(\xi) = P\{L(\xi) = 0, X(\xi) = 0\}$ and $\chi_0(\xi) = P\{L(\xi) = 1, X(\xi) = 0\}$ the probability densities,

$$\chi_\sigma(\tilde{\varphi}, \xi)d\tilde{\varphi} = P\{L(\xi) = 1, X(\xi) = \sigma, \tilde{\varphi} \leq \mathcal{B}^0(\xi) < \tilde{\varphi} + d\tilde{\varphi}\},$$

for $\xi \geq 0, \tilde{\varphi} \geq 0$ and $\sigma \geq 1$.

$$\Psi_\sigma(\tilde{\varphi}, \xi)d\tilde{\varphi} = P\{L(\xi) = 2, X(\xi) = \sigma, \tilde{\varphi} \leq \mathcal{C}^0(\xi) < \tilde{\varphi} + d\tilde{\varphi}\},$$

for $\xi \geq 0, \tilde{\varphi} \geq 0, \sigma \geq 0$.

$$\Lambda_{v,\sigma}(\tilde{\varphi}, \xi)d\tilde{\varphi} = P\{L(\xi) = 3, X(\xi) = \sigma, \tilde{\varphi} \leq \mathcal{H}_v^0(\xi) < \tilde{\varphi} + d\tilde{\varphi}\},$$

for $\xi \geq 0, \tilde{\varphi} \geq 0$ and $\sigma \geq 0$.

$$\Pi_\sigma(\tilde{\varphi}, \xi)d\tilde{\varphi} = P\{L(\xi) = 4, X(\xi) = \sigma, \tilde{\varphi} \leq \mathcal{I}^0(\xi) < \tilde{\varphi} + d\tilde{\varphi}\},$$

for $\xi \geq 0, \tilde{\varphi} \geq 0, \sigma \geq 1$.

In subsequent parts, the following probabilities are applied:

1. The prob., of the server being idle and on WV at time ξ is denoted by $\phi_0(\xi)$.
2. The prob., of the server being idle and on typical active period at time ξ is denoted by $\chi_0(\xi)$.
3. If there are accurately σ clients in the orbit at time ξ and the elapsed retrial time of the test clients undergoing retrial is between $\tilde{\varphi}$ and $\tilde{\varphi} + d\tilde{\varphi}$, then the prob., that this is the case is $\chi_\sigma(\tilde{\varphi}, \xi)$.
4. When there are σ consumers in the orbit, the prob., of the test customer's elapsed regular service time ranging between $\tilde{\varphi}$ and $\tilde{\varphi} + d\tilde{\varphi}$ is $\Psi_\sigma(\tilde{\varphi}, \xi)$.
5. $\Lambda_{v,\sigma}(\tilde{\varphi}, \xi)d\tilde{\varphi}$ and $\Pi_\sigma(\tilde{\varphi}, \xi)d\tilde{\varphi}$ is the prob., that there are precisely σ patrons in the orbit, with the elapsed (reduced service time and repair time) of the test patron being between $\tilde{\varphi}$ and $\tilde{\varphi} + d\tilde{\varphi}$ at time ξ .

Suppose that the sequel fulfills the stability condition, thus we can provide $\chi_0 = \lim_{\xi \rightarrow \infty} \chi_0(\xi)$ and limiting densities are

$$\chi_\sigma(\tilde{\varphi}) = \lim_{\xi \rightarrow \infty} \chi_\sigma(\tilde{\varphi}, \xi) \text{ for } \tilde{\varphi} \geq 0 \text{ and } \sigma \geq 1.$$

$$\Psi_\sigma(\tilde{\varphi}) = \lim_{\xi \rightarrow \infty} \Psi_\sigma(\tilde{\varphi}, \xi) \text{ for } \tilde{\varphi} \geq 0 \text{ and } \sigma \geq 0.$$

$$\Lambda_{v,\sigma}(\tilde{\varphi}) = \lim_{\xi \rightarrow \infty} \Lambda_{v,\sigma}(\tilde{\varphi}, \xi) \text{ for } \tilde{\varphi} \geq 0 \text{ and } \sigma \geq 0.$$

$$\Pi_\sigma(\tilde{\varphi}) = \lim_{\xi \rightarrow \infty} \Pi_\sigma(\tilde{\varphi}, \xi) \text{ for } \tilde{\varphi} \geq 0 \text{ and } \sigma \geq 1.$$

Applying the SVT, we create the following system of equations.

$$\mu\chi_0 = \gamma r_1 \phi_0 \tag{1}$$

$$(\mu + \gamma)\phi_0 = \gamma r_2 \phi_0 \int_0^\infty \Lambda_{v,0}(\tilde{\varphi})\kappa(\tilde{\varphi})d\tilde{\varphi} + \int_0^\infty \Psi_\sigma(\tilde{\varphi})\zeta(\tilde{\varphi})d\tilde{\varphi} \tag{2}$$

$$\frac{d}{d\tilde{\varphi}}\chi_\sigma(\tilde{\varphi}) + (\mu + \eta(\tilde{\varphi}))\chi_\sigma(\tilde{\varphi}) = 0, \sigma \geq 1 \tag{3}$$

$$\frac{d}{d\tilde{\varphi}}\Psi_0(\tilde{\varphi}) + (\mu + \zeta(\tilde{\varphi}))\Psi_0(\tilde{\varphi}) = 0, \sigma = 0. \tag{4}$$

$$\frac{d}{d\tilde{\varphi}}\Psi_\sigma(\tilde{\varphi}) + (\mu + \zeta(\tilde{\varphi}))\Psi_\sigma(\tilde{\varphi}) = \mu\Psi_{\sigma-1}(\tilde{\varphi}), \sigma \geq 1 \tag{5}$$

$$\frac{d}{d\tilde{\varphi}}\Lambda_{0,v}(\tilde{\varphi}) + (\mu + \gamma + \kappa(\tilde{\varphi}))\Lambda_{0,v}(\tilde{\varphi}) = 0, \sigma = 0. \tag{6}$$

$$\frac{d}{d\tilde{\varphi}}\Lambda_{\sigma,v}(\tilde{\varphi}) + (\mu + \gamma + \kappa(\tilde{\varphi}))\Lambda_{\sigma,v}(\tilde{\varphi}) = \mu\Lambda_{\sigma-1}(\tilde{\varphi}), \sigma \geq 1. \tag{7}$$

$$\frac{d}{d\tilde{\varphi}}\Pi_0(\tilde{\varphi}) + (\mu + v(\tilde{\varphi}))\Pi_0(\tilde{\varphi}) = 0, \sigma = 0. \tag{8}$$

$$\frac{d}{d\tilde{\varphi}}\Pi_\sigma(\tilde{\varphi}) + (\mu + v(\tilde{\varphi}))\Pi_\sigma(\tilde{\varphi}) = \mu\Pi_{\sigma-1}(\tilde{\varphi}), \sigma \geq 1. \tag{9}$$

At $\tilde{\varphi} = 0$ the steady state boundary conditions are as follows:

$$\begin{aligned} \chi_\sigma(0) &= \bar{\omega} \int_0^\infty \Psi_\sigma(\tilde{\varphi})\zeta(\tilde{\varphi})d\tilde{\varphi} + \omega \int_0^\infty \Psi_{\sigma-1}(\tilde{\varphi})\zeta(\tilde{\varphi})d\tilde{\varphi} + \bar{\omega} \int_0^\infty \Lambda_{v,\sigma}(\tilde{\varphi})\kappa(\tilde{\varphi})d\tilde{\varphi} \\ &+ \omega \int_0^\infty \Lambda_{v,\sigma-1}(\tilde{\varphi})\kappa(\tilde{\varphi})d\tilde{\varphi} + \int_0^\infty \Pi_\sigma(\tilde{\varphi})v(\tilde{\varphi})d\tilde{\varphi} \end{aligned} \tag{10}$$

$$\Psi_0(0) = \lambda \int_0^\infty \chi_1(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \lambda\bar{\mu}\chi_0 + \gamma \int_0^\infty \Lambda_{0,v}(\tilde{\varphi})d\tilde{\varphi}, \sigma = 0 \tag{11}$$

$$\Psi_\sigma(0) = \lambda \int_0^\infty \chi_{\sigma+1}(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \lambda\mu \int_0^\infty \Psi_\sigma(\tilde{\varphi})d\tilde{\varphi} + \gamma \int_0^\infty \Lambda_{\sigma,v}(\tilde{\varphi})d\tilde{\varphi}, \sigma \geq 1 \tag{12}$$

$$\Lambda_{v,\sigma}(0) = \begin{cases} \mu\phi_0, & \sigma = 0 \\ 0, & \sigma \geq 1 \end{cases} \tag{13}$$

$$\Pi_1(0) = \bar{\lambda} \int_0^\infty \chi_1(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \bar{\lambda}\mu\chi_0 \tag{14}$$

$$\Pi_\sigma(0) = \bar{\lambda} \int_0^\infty \chi_\sigma(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \bar{\lambda}\mu \int_0^\infty \chi_{\sigma-1}(\tilde{\varphi})d\tilde{\varphi}, \sigma \geq 2 \tag{15}$$

The normalizing condition is

$$\begin{aligned} \chi_0 + \phi_0 + \sum_{\sigma=1}^\infty \int_0^\infty \chi_\sigma(\tilde{\varphi})d\tilde{\varphi} + \sum_{\sigma=0}^\infty \int_0^\infty \Psi_\sigma(\tilde{\varphi})d\tilde{\varphi} + \sum_{\sigma=0}^\infty \int_0^\infty \Lambda_{\sigma,v}(\tilde{\varphi})d\tilde{\varphi} \\ + \sum_{\sigma=1}^\infty \int_0^\infty \Pi_\sigma(\tilde{\varphi})d\tilde{\varphi} = 1 \end{aligned} \tag{16}$$

3.4. The steady state solution

The PGF is used to compute the steady state solution for the RQ model. To solve the aforementioned equations, the generating functions for $|\vartheta| < 1$ are described as below:

$$\begin{aligned}\chi(\tilde{\varphi}, \vartheta) &= \sum_{\sigma=1}^{\infty} \chi_{\sigma}(\tilde{\varphi})\vartheta^{\sigma}; \chi(0, \vartheta) = \sum_{\sigma=1}^{\infty} \chi_{\sigma}(0)\vartheta^{\sigma}; \\ \Psi(\tilde{\varphi}, \vartheta) &= \sum_{\sigma=0}^{\infty} \Psi_{\sigma}(\tilde{\varphi})\vartheta^{\sigma}; \Psi(0, \vartheta) = \sum_{n=0}^{\infty} \Psi_0(0)\vartheta^{\sigma}; i = 1, 2 \\ \Lambda_v(\tilde{\varphi}, \vartheta) &= \sum_{\sigma=0}^{\infty} \Lambda_{v,\sigma}(\tilde{\varphi})\vartheta^{\sigma}; \Lambda_v(0, \vartheta) = \sum_{\sigma=0}^{\infty} \Lambda_{v,\sigma}(0)\vartheta^{\sigma}; \\ \Pi(\tilde{\varphi}, \vartheta) &= \sum_{\sigma=1}^{\infty} \Pi_{\sigma}(\tilde{\varphi})\vartheta^{\sigma}; \Pi(0, \vartheta) = \sum_{\sigma=1}^{\infty} \Pi_{\sigma}(0)\vartheta^{\sigma}\end{aligned}$$

Next multiply the SS eqn. and SS boundary conditions from (3) to (15) by ϑ^{σ} and adding over σ , ($\sigma = 0, 1, 2, \dots$)

$$\frac{\partial}{\partial \tilde{\varphi}} \chi(\tilde{\varphi}, \vartheta) + [\mu + \eta(\tilde{\varphi})]\chi(\tilde{\varphi}, \vartheta) = 0 \tag{17}$$

$$\frac{\partial}{\partial \tilde{\varphi}} \Psi(\tilde{\varphi}, \vartheta) + [\mu(1 - \vartheta) + \zeta(\tilde{\varphi})]\Psi(\tilde{\varphi}, \vartheta) = 0 \tag{18}$$

$$\frac{\partial}{\partial \tilde{\varphi}} \Lambda_v(\tilde{\varphi}, \vartheta) + [\gamma + \mu(1 - \vartheta) + \kappa(\tilde{\varphi})]\Lambda_v(\tilde{\varphi}, \vartheta) = 0 \tag{19}$$

$$\frac{\partial}{\partial \tilde{\varphi}} \Pi(\tilde{\varphi}, \vartheta) + [\mu(1 - \vartheta) + v(\tilde{\varphi})]\Pi(\tilde{\varphi}, \vartheta) = 0 \tag{20}$$

Solving the partial differential eqns. (17) to (20), we obtain

$$\chi(\tilde{\varphi}, \vartheta) = \chi(0, \vartheta)[1 - \mathcal{B}(\tilde{\varphi})]e^{-\mu\tilde{\varphi}} \tag{21}$$

$$\Psi(\tilde{\varphi}, \vartheta) = \Psi(0, \vartheta)[1 - \mathcal{C}(\tilde{\varphi})]e^{-\mathcal{F}(\vartheta)\tilde{\varphi}} \tag{22}$$

$$\Lambda_v(\tilde{\varphi}, \vartheta) = \Lambda_v(0, \vartheta)[1 - \mathcal{H}_v(\tilde{\varphi})]e^{-\mathcal{F}_v(\vartheta)\tilde{\varphi}} \tag{23}$$

$$\Pi(\tilde{\varphi}, \vartheta) = \Pi(0, \vartheta)[1 - \mathcal{I}(\tilde{\varphi})]e^{-\mathcal{F}(\vartheta)\tilde{\varphi}} \tag{24}$$

where $\mathcal{F}(\vartheta) = \mu(1 - \vartheta)$, $\mathcal{F}_v(\vartheta) = \gamma + \mu(1 - \vartheta)$

Multiplying equation (10) and (12,13,15) by appropriate powers of ϑ , adding over n with few mathematical manipulations, we obtain

$$\begin{aligned}\chi(0, \vartheta) &= (\bar{\omega} + \omega\vartheta) \int_0^{\infty} \Psi(\tilde{\varphi}, \vartheta)\zeta(\tilde{\varphi})d\tilde{\varphi} + (\bar{\omega} + \omega\vartheta) \int_0^{\infty} \Lambda_v(\tilde{\varphi}, \vartheta)\kappa(\tilde{\varphi})d\tilde{\varphi} \\ &\quad + \int_0^{\infty} \Pi(\tilde{\varphi}, \vartheta)v(\tilde{\varphi})d\tilde{\varphi} - (\mu + \gamma r_1)\phi_0\end{aligned} \tag{25}$$

$$\Psi(0, \vartheta) = \frac{\lambda}{\vartheta} \int_0^{\infty} \chi(\tilde{\varphi}, \vartheta)\eta(\tilde{\varphi})d\tilde{\varphi} + \lambda\mu \int_0^{\infty} \chi(\tilde{\varphi}, \vartheta)d\tilde{\varphi} + \gamma \int_0^{\infty} \Lambda_v(\tilde{\varphi}, \vartheta)d\tilde{\varphi} + \lambda\mu\chi_0 \tag{26}$$

$$\Lambda_v(0, \vartheta) = \mu\phi_0 \tag{27}$$

$$\Pi(0, \vartheta) = \bar{\lambda}\vartheta\mu \int_0^{\infty} \chi(\tilde{\varphi}, \vartheta)d\tilde{\varphi} + \bar{\lambda} \int_0^{\infty} \chi(\tilde{\varphi}, \vartheta)\eta(\tilde{\varphi})d\tilde{\varphi} + \vartheta\mu\bar{\lambda}\chi_0 \tag{28}$$

Using eqn (21,23 and 27) in eqn (26)

$$\Psi(0, \vartheta) = \lambda\chi(0, \vartheta) \left[\frac{\vartheta + (1 - \vartheta)\bar{\mathcal{B}}(\mu)}{\vartheta} \right] + \mu\phi_0\mathcal{V}(\vartheta) + \lambda\gamma r_1\phi_0 \tag{29}$$

Similarly using equation (21) in (28)

$$\Pi(0, \vartheta) = \vartheta \gamma r_1 \bar{\lambda} \chi_0 + \bar{\lambda} \chi(0, \vartheta) [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)] \quad (30)$$

Substituting equations (22),(23) and (24) in (25), we obtain

$$\chi(0, \vartheta) = (\bar{\omega} + \omega \vartheta) \Psi(0, \vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + (\bar{\omega} + \omega \vartheta) \Lambda_v(0, \vartheta) \bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \Pi(0, \vartheta) \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0 \quad (31)$$

Using equations (27),(29) and (30) in equation (31)

$$\chi(0, \vartheta) = \vartheta \left\{ \frac{(\bar{\omega} + \omega \vartheta) \mu \phi_0 [\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega \vartheta) \lambda \gamma r_1 \phi_0 \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \gamma r_1 \phi_0 \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0}{\vartheta - [(\bar{\omega} + \omega \vartheta) \lambda \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \bar{\mathcal{I}}(\mathcal{F}(\vartheta))] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)]} \right\} \quad (32)$$

substituting equation (32) in (29) and (30),we obtain

$$\Psi(0, \vartheta) = \left\{ \frac{\lambda [(\bar{\omega} + \omega \vartheta) \mu \phi_0 [\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega \vartheta) \lambda \gamma r_1 \phi_0 \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \gamma r_1 \phi_0 \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)] + \mu \phi_0 \mathcal{V}(\vartheta) + \lambda \gamma r_1 \phi_0}{\vartheta - [(\bar{\omega} + \omega \vartheta) \lambda \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \bar{\mathcal{I}}(\mathcal{F}(\vartheta))] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)]} \right\} \quad (33)$$

$$\Pi(0, \vartheta) = \left\{ \frac{\bar{\lambda} [(\bar{\omega} + \omega \vartheta) \mu \phi_0 [\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega \vartheta) \lambda \gamma r_1 \phi_0 \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \gamma r_1 \phi_0 \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0] + \vartheta \bar{\lambda} \gamma r_1 \phi_0}{\vartheta - [(\bar{\omega} + \omega \vartheta) \lambda \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \bar{\mathcal{I}}(\mathcal{F}(\vartheta))] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)]} \right\} \quad (34)$$

Substituting equations (27) and (32) to (34) in (21) to (24)

$$\chi(\tilde{\varphi}, \vartheta) = \vartheta \left\{ \frac{(\bar{\omega} + \omega \vartheta) \mu \phi_0 [\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega \vartheta) \lambda \gamma r_1 \phi_0 \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \gamma r_1 \phi_0 \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0}{\vartheta - [(\bar{\omega} + \omega \vartheta) \lambda \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \bar{\mathcal{I}}(\mathcal{F}(\vartheta))] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)]} \right\} \quad (35)$$

$$\times [1 - \mathcal{B}(\tilde{\varphi})] e^{-\mu \tilde{\varphi}}$$

$$\Psi(\tilde{\varphi}, \vartheta) = \left\{ \frac{\lambda [(\bar{\omega} + \omega \vartheta) \mu \phi_0 [\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega \vartheta) \lambda \gamma r_1 \phi_0 \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \gamma r_1 \phi_0 \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)] + \mu \phi_0 \mathcal{V}(\vartheta) + \lambda \gamma r_1 \phi_0}{\vartheta - [(\bar{\omega} + \omega \vartheta) \lambda \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \bar{\mathcal{I}}(\mathcal{F}(\vartheta))] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)]} \right\} \quad (36)$$

$$\times [1 - \mathcal{C}(\tilde{\varphi})] e^{-\mathcal{F}(\vartheta) \tilde{\varphi}}$$

$$\Lambda_v(\tilde{\varphi}, \vartheta) = \mu \phi_0 [1 - \mathcal{H}_v(\tilde{\varphi})] e^{-\mathcal{F}_v(\vartheta) \tilde{\varphi}} \quad (37)$$

$$\Pi(\tilde{\varphi}, \vartheta) = \left\{ \frac{\bar{\lambda} [(\bar{\omega} + \omega \vartheta) \mu \phi_0 [\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta) \bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega \vartheta) \lambda \gamma r_1 \phi_0 \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \gamma r_1 \phi_0 \bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu \phi_0 - \gamma r_1 \phi_0] + \vartheta \bar{\lambda} \gamma r_1 \phi_0}{\vartheta - [(\bar{\omega} + \omega \vartheta) \lambda \bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta \bar{\lambda} \bar{\mathcal{I}}(\mathcal{F}(\vartheta))] [\vartheta + (1 - \vartheta) \bar{\mathcal{B}}(\mu)]} \right\} \quad (38)$$

$$\times [1 - \mathcal{I}(\tilde{\varphi})] e^{-\mathcal{F}(\vartheta) \tilde{\varphi}}$$

3.5. Theorem

The stationary dist. of the no. of clients in the orbit while the server is free,normal operative service, slow service and the prob. that the server is idle is described by $\rho < \bar{B}(\mu)$ under the stability condition

$$\chi(\vartheta) = \vartheta \left\{ \frac{(\bar{\omega} + \omega\vartheta)\phi_0[\bar{H}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{C}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega\vartheta)\frac{\lambda}{\mu}\gamma r_1\phi_0\bar{C}(\mathcal{F}(\vartheta)) + \vartheta\frac{\bar{\lambda}}{\mu}\gamma r_1\phi_0\bar{I}(\mathcal{F}(\vartheta)) - \phi_0 - \frac{\gamma r_1}{\mu}\phi_0}{\vartheta - [(\bar{\omega} + \omega\vartheta)\lambda\bar{C}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{I}(\mathcal{F}(\vartheta))][\vartheta + (1 - \vartheta)\bar{B}(\mu)]} \right\} \quad (39)$$

$$\times [1 - \bar{B}(\mu)]$$

$$\Psi(\vartheta) = \left\{ \frac{\lambda[(\bar{\omega} + \omega\vartheta)\mu\phi_0[\bar{H}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{C}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega\vartheta)\lambda\gamma r_1\phi_0\bar{C}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{I}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0][\vartheta + (1 - \vartheta)\bar{B}(\mu)] + \mu\phi_0\mathcal{V}(\vartheta) + \lambda\gamma r_1\phi_0}{\vartheta - [(\bar{\omega} + \omega\vartheta)\lambda\bar{C}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{I}(\mathcal{F}(\vartheta))][\vartheta + (1 - \vartheta)\bar{B}(\mu)]} \right\} \quad (40)$$

$$\times \frac{[1 - \bar{C}(\mathcal{F}(\vartheta))]}{\mu(1 - \vartheta)}$$

$$\Lambda_v(\vartheta) = \frac{\mu\phi_0\mathcal{V}(\vartheta)}{\gamma} \quad (41)$$

$$\Pi(\vartheta) = \left\{ \frac{\bar{\lambda}[(\bar{\omega} + \omega\vartheta)\mu\phi_0[\bar{H}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{C}(\mathcal{F}(\vartheta))] + (\bar{\omega} + \omega\vartheta)\lambda\gamma r_1\phi_0\bar{C}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{I}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0] + \vartheta\bar{\lambda}\gamma r_1\phi_0}{\vartheta - [(\bar{\omega} + \omega\vartheta)\lambda\bar{C}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{I}(\mathcal{F}(\vartheta))][\vartheta + (1 - \vartheta)\bar{B}(\mu)]} \right\} \quad (42)$$

$$\times \frac{[1 - \bar{I}(\mathcal{F}(\vartheta))]}{\mu(1 - \vartheta)}$$

Proof. Taking the equations. (35) – (38) and integrating them with regard to $\tilde{\varphi}$ and obtain the partial PGF's $\chi(\vartheta) = \int_0^\infty \chi(\tilde{\varphi}, \vartheta)d\tilde{\varphi}$, $\Psi(\vartheta) = \int_0^\infty \Psi(\tilde{\varphi}, \vartheta)d\tilde{\varphi}$, $\Lambda_v(\vartheta) = \int_0^\infty \Lambda_v(\tilde{\varphi}, \vartheta)d\tilde{\varphi}$, $\Pi(\vartheta) = \int_0^\infty \Pi(\tilde{\varphi}, \vartheta)d\tilde{\varphi}$.

We can find the prob. that the server is free by using the normalisation condition (χ_0) and (ϕ_0) by establishing functions as, when there is no consumer in the orbit $\vartheta = 1$ in (3.39) – (3.42) and using the "L'Hospital rule" if it is required, we examine $\chi_0 + \phi_0 + \chi(1) + \Psi(1) + \Lambda_v(1) + \Pi(1) = 1$. ■

3.6. Theorem

The stability constraint $\rho < \bar{B}(\mu)$ used to determine the PGF of the no. of clients in the system and the orbit size dist. at a stationary point in time is given by

$$H_s(\vartheta) = \frac{Ne_s(\vartheta)}{De_s(\vartheta)} \quad (43)$$

$$H_0(\vartheta) = \frac{Ne_0(\vartheta)}{De_s(\vartheta)} \quad (44)$$

Proof. The "PGF of the no.of consumer in the system ($H_s(\vartheta)$) and in the orbit ($H_0(\vartheta)$)" is calculated by applying $H_s(\vartheta) = \chi_0 + \phi_0 + \chi(\vartheta) + \vartheta\{\Psi(\vartheta) + \Lambda_v(\vartheta)\} + \Pi(\vartheta)$. and $H_0(\vartheta) = \chi_0 + \phi_0 + \chi(\vartheta) + \{\Psi(\vartheta) + \Lambda_v(\vartheta)\} + \Pi(\vartheta)$. Insert the eqns. (39) – (42) in the earlier results,then the eqns. (43) and (44) may be computed immediately. ■

4. MEASURES OF SYSTEM PERFORMANCE

This part calculates many appropriate system prob., system efficiency metrics, and signifies the mean busy period and cycle that occur while the system is in various phases.

4.1. System state probabilities

By putting $\vartheta \rightarrow 1$ in equations. (39) – (42) and applying “L Hospital’s rule” wherever possible. we obtain the following findings.

(i)Pr(The server being available for the duration of the retrial)

$$\chi(1) = \phi_0[1 - \bar{B}(\mu)] \left\{ \frac{[[\frac{\mu}{\gamma}[1 - \bar{H}_v(\gamma)] - \mu C_1[1 - \bar{H}_v(\gamma)]] - \frac{\lambda}{\mu} \gamma r_1[\omega + \mu C_1] + \frac{\bar{\lambda}}{\mu} \gamma r_1[1 - \mu \mathcal{I}_1]]}{\bar{B}(\mu) + \lambda \omega - \lambda \mu C_1 - \bar{\lambda}(1 + \mu \mathcal{I}_1)} \right\} \quad (45)$$

(ii)Pr(The server is operative on usual service period)

$$\Psi(1) = \left\{ \frac{\phi_0 \lambda C_1 [[\frac{\mu}{\gamma}[1 - \bar{H}_v(\gamma)] - \mu C_1[1 - \bar{H}_v(\gamma)]] - \lambda \gamma r_1[\omega + \mu C_1 + \bar{\lambda} \gamma r_1[1 - \mu \mathcal{I}_1] + \mu[\mu \mathcal{H}_1 + \frac{\mu}{\gamma}[1 - \bar{H}_v(\gamma)]]]}{\bar{B}(\mu) + \lambda \omega - \lambda \mu C_1 - \bar{\lambda}(1 + \mu \mathcal{I}_1)} \right\} \quad (46)$$

(iii)Pr(The server is on WV)

$$\Lambda_v(1) = \frac{\phi_0 \mu [1 - \bar{H}_v(\gamma)]}{\gamma} \quad (47)$$

(iv)Pr(The server is under repair time during usual active period)

$$\Pi = \Pi(1) = \left\{ \frac{\phi_0 \bar{\lambda} \mathcal{I}_1 [[\frac{\mu}{\gamma}[1 - \bar{H}_v(\gamma)] - \mu C_1[1 - \bar{H}_v(\gamma)]] - \lambda \gamma r_1[\omega + \mu C_1 + \bar{\lambda} \gamma r_1[1 - \mu \mathcal{I}_1] + \bar{\lambda} \gamma r_1]]}{\bar{B}(\mu) + \lambda \omega - \lambda \mu C_1 - \bar{\lambda}(1 + \mu \mathcal{I}_1)} \right\} \quad (48)$$

4.2. Average system size and its orbit

In a steady state, the system,

(i) Differentiating the equation (44) and the predicted no. of clients in the orbit (L_q) is established with regard to ϑ and $\vartheta = 1$.

$$L_q = H'_o(1) = \lim_{\vartheta \rightarrow 1} \frac{d}{d\vartheta} H_o(\vartheta) = \phi_0 \left[\frac{Ne_q'''(1)De_q''(1) - De_q'''(1)Ne_q''(1)}{3(De_q''(1))^2} \right] \quad (49)$$

(ii) The predicted no. of clients in the system (L_s) is determined by differentiating the eqn. (43) with regard to ϑ and giving $\vartheta = 1$ yields.

$$L_s = H'_s(1) = \lim_{\vartheta \rightarrow 1} \frac{d}{d\vartheta} H_s(\vartheta) = \phi_0 \left[\frac{Ne_s'''(1)De_q''(1) - De_q'''(1)Ne_q''(1)}{3(De_q''(1))^2} \right] \quad (50)$$

(iii) The mean waiting time of consumers in the system and queue [W_s and W_q] are computed utilizing “Little’s method” $W_s = \frac{L_s}{\mu}$ and $W_q = \frac{L_q}{\mu}$ respectively.

4.3. Mean busy period and the busy cycle

Let $A(T_b)$ and $A(T_c)$ be the predicted sizes of the busy period and cycle, respectively under steady state conditions. The outcomes are derived directly from the justification of a different renewal procedure [5], which concludes in

$$\phi_0 = \frac{A(T_0)}{A(T_b) + A(T_0)}; A(T_b) = \frac{1}{\mu} \left(\frac{1}{\phi_0} - 1 \right); A(T_c) = \frac{1}{\mu\phi_0} = A(T_0) + A(T_b). \quad (51)$$

where T_0 is the period of time spent in the system's null state. Because there is an exponential difference in time between the arrivals of two customers and $A(T_0) = (1/\mu)$ with the parameter μ .

5. PARTICULAR CASES

We examine a few real-world examples of our technique that are consistent with the existing research in this area.

Case (i): No feedback, No VI and No SF

If $\omega = 1, \lambda = 1$, and $\gamma = 0$. The model may be lowered to a $M/G/1$ RQ with WV and the findings match those of Arivudainambi et.al.[14]

Case (ii): No retrial, No feedback and No starting failure.

Let $r_2 = 0, \omega = 1, \lambda = 1$ and $\tilde{B}(\mu) \rightarrow 1$. our framework has been simplified to an "M/G/1 queue with WVs and VI". Our results agree with Zhang and Hou [15].

6. NUMERICAL ANALYSIS

This section will demonstrate the different settings for system performance measures by using MATLAB. We investigate exponentially distributed retrial, service, slower pace service, vacation and repair periods. Numerical measurements are selected at random in order to fulfil the stability criteria. Tables 1 to 3 provides assessed outcomes of the idle prob., χ_0 , ϕ_0 the "mean queue size (L_q), mean waiting time in the queue (W_q)" in our QM.

Table 1 shows that the retrial rate (η) escalates, χ_0 escalates, but L_q, W_q decreases for the value of $\omega = 0.19, \mu = 0.9, \lambda = 0.8, \gamma = 3, \tilde{H}_v(\gamma) = 0.9, r_1 = 0.5$.

Table 2 demonstrates that the vacation rate (γ) mounts, ϕ_0 increases, L_q, W_q subsides for the value of $\omega = 0.19, \mu = 1.5, \lambda = 0.19, \tilde{H}_v(\gamma) = 0.9, r_1 = 0.9$.

Table 3 clearly displays that feedback rate (ω) mounts, χ_0, L_q, W_q diminishes for the value of $\mu = 0.9, \lambda = 0.10, \tilde{H}_v(\gamma) = 0.9, r_1 = 0.19, \gamma = 0.9$.

Table 1: The impact of Retrial rate (η) on χ_0, L_q, W_q

Retrial rate (η)	χ_0	L_q	W_q
2.0	2.0718	0.0883	0.0982
2.5	2.2132	0.0885	0.0984
3.0	2.3172	0.0821	0.0912
3.5	2.3969	0.0728	0.0809
4.0	2.4599	0.0621	0.0690
4.5	2.5110	0.0507	0.0564
5.0	2.5532	0.0391	0.0434

The Figure 1 (a) indicates that retrial rate (η) escalates, (L_q) and (W_q) increases. The Figure 1 (b) displays that vacation rate (γ) escalates, (L_q) and (W_q) decreases. The Figure 1 (c)

Table 2: The impact of Vacation rate (γ) on ϕ_0, L_q, W_q

Vacation rate (γ)	ϕ_0	L_q	W_q
0.31	0.3724	0.6878	1.3756
0.32	0.3756	0.4931	0.9863
0.33	0.3785	0.3460	0.6921
0.34	0.3811	0.2890	0.4780
0.35	0.3835	0.1658	0.3316
0.36	0.3857	0.1211	0.2422
0.37	0.3878	0.1005	0.2010

Table 3: The impact of Feedback rate (ω) on χ_0, L_q, W_q

Feedback rate (ω)	χ_0	L_q	W_q
0.10	0.8306	0.2211	0.2457
0.20	0.8269	0.2008	0.2231
0.30	0.8233	0.1812	0.2013
0.40	0.8197	0.1623	0.1804
0.50	0.8161	0.1443	0.1603
0.60	0.8125	0.1270	0.1411
0.70	0.8090	0.1105	0.1227

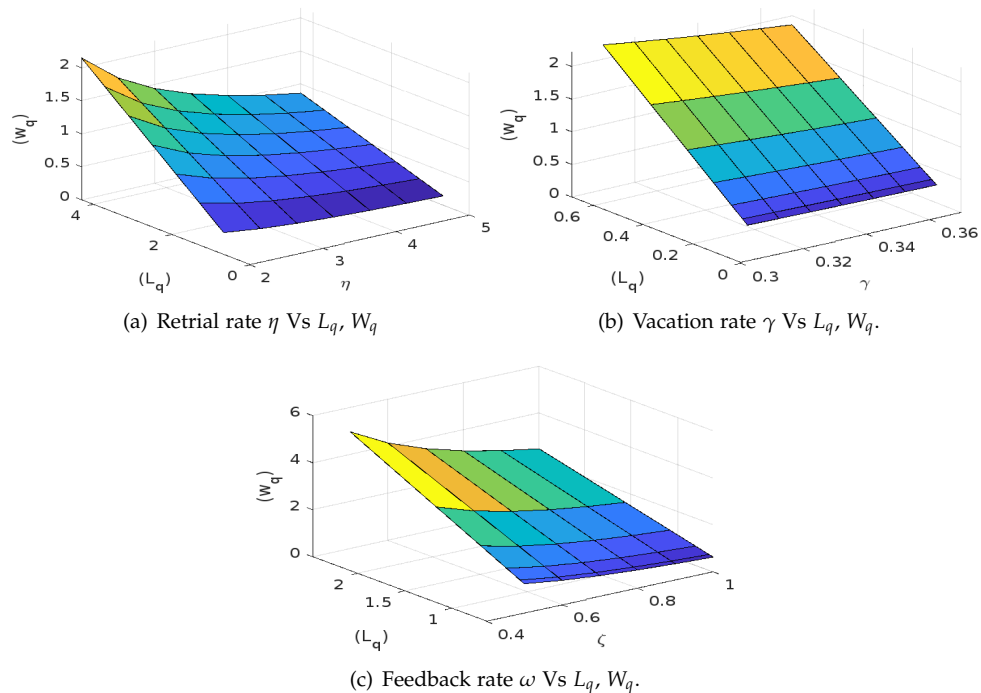


Figure 1: Effects of a few parameters on 3D representation.

demonstrates that feedback rate (ω) increases, (L_q) and (W_q) diminishes.

We may use the numerical findings above to determine the influence of features on the

system's assessment criteria with certainty, the outcomes correspond to actual circumstances.

6.1. ANFIS Computing

Using the fuzzy toolbox of MATLAB software, an ANFIS network can be executed to compare outcomes from analyses. ANFIS, a soft computing approach, is an effective tool for identifying important results that are useful in busy everyday environments. This approach aids in the identification of approximate solutions for measurements whose definite outcomes would otherwise be difficult to determine.

In our framework, a neuro-fuzzy technique is used to compute the expected no. of consumers in the queue (W_q) by changing the retrial rate (η), vacation rate (γ) and feedback rate (ω), as shown in the numerical results in Figure 2 (a – c). We consider the parameters (η, γ and ω) as linguistic variables (LV) that are performed for four epochs each. The analytic (ANFIS) outcomes are exhibited by the solid (dashed) lines. In the context of fuzzy systems, these factors are regarded as LV and are used as input variables in ANFIS networks.

The Gaussian function provides the membership functions for each of these input variables. The following are the linguistic values for each parameter: low, average, high, and excessive. The diagrams demonstrate agreement between the analytical findings for the paradigm and the neuro fuzzy results achieved through the ANFIS approach.

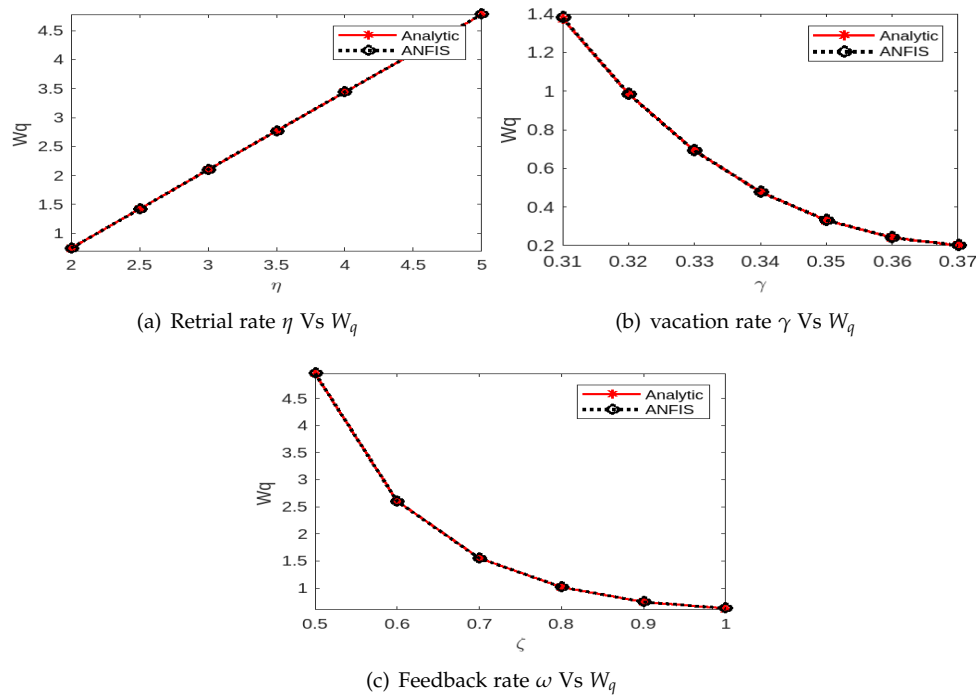


Figure 2: Effects of a few parameters on 2D representation.(ANFIS)

7. COST OPTIMIZATION

Our research aims to maintain system accessibility while optimizing system costs. Consequently, we establish the predicted cost function for system performance metrics and then accomplish a numerical analysis of the machining system under study. In order to calculate the best average cost per unit of time (TC), the parameters must be determined. This section discusses the best cost construct for the suggested approach using the standard cost notation form, and it provides the estimated total cost per unit of time as follows:

$$ETC = S_h L_s + S_b \Psi + S_v \Lambda + S_r \Pi + S_1 \zeta + S_2 \kappa \tag{52}$$

where,

- S_h = Holding cost per unit consumer.
- S_b = Cost per unit time while the server provides service during a usual busy period.
- S_v = Cost per unit time in the system when the server is on vacation.
- S_r = Cost per unit time for providing repair to the failed server.
- S_1 = Cost per unit time consumer served by the mean service rate ζ .
- S_2 = Cost per unit time consumer served by the mean vacation rate κ .

Equation (52) has an estimated total cost function that is multivariate and nonlinear. Therefore, developing an analytical solution for optimal parameter values say ζ^* and κ^* is problematic. In order to determine the most suitable numerical value for the decision parameters, the well-recognised meta heuristic technique: The optimization approach used is called Particle Swarm Optimization (PSO).

Choosing at random the present values for the cost element and the parameters are : $\mu = 0.05$, $\bar{B}(\mu) = 4.5$, $\bar{\lambda} = 5.9$, $\bar{H}_v(\gamma) = 0.75$, ($\zeta^* = 0.3670$, $\kappa^* = 0.3900$). Our goal is to identify the best values that will allow us to minimise the cost function. The five sets of cost factors that we have chosen are listed below.

Table 4: Cost Sets values for different cost aspects

Cost sets	S_h	S_b	S_v	S_1	S_2	S_r
Set 1	\$15	\$75	\$20	\$19	\$12	\$10
Set 2	\$20	\$85	\$25	\$17	\$19	\$23
Set 3	\$25	\$95	\$30	\$15	\$15	\$29

Applying the PSO algorithm using MATLAB software to the previously specified cost factors. In this research, we have 100 candidates, 500 iterations throughout, and a range of parameters between 0.006 and 0.65 for the lower and upper bounds.

Tables 5 demonstrates that the effects of μ, ω, γ on TEC^* using PSO.

Table 5: The PSO approach is executed by changing μ, ω and γ to determine the minimal cost for different cost sets.

Parameters	(TEC^*)			
	Cost set 1	Cost set 2	Cost set 3	
μ	0.20	\$75.1471	\$59.7276	\$66.7365
	0.25	\$115.1554	\$87.6434	\$92.1116
	0.30	\$151.7660	\$114.6657	\$119.3708
ω	1.00	\$73.5533	\$58.6272	\$65.7761
	1.15	\$74.5054	\$59.2872	\$66.3535
	1.20	\$74.8255	\$59.5074	\$66.5452
γ	1.25	\$78.6575	\$61.4278	\$67.9418
	1.30	\$76.0953	\$60.0313	\$66.8629
	1.35	\$73.5533	\$58.6272	\$65.7761

By changing a few of the variables, we were able to determine the entire system's cost, and we found that for $\mu = 1.00, \gamma = 1.35$ and ($\zeta^* = 0.3670, \kappa^* = 0.3900$) the lowest cost was \$58.6272.

7.1. Particle Swarm Optimization

A precise evaluation of the QM is highly essential to offer adequate service and decrease congestion with the increasing expansion of computer networking and communications. If the greatest number of consumers can access an affordable system, then this is feasible. As a result, solutions that incorporate cost optimisation are very beneficial and advised. Jain et al. [25] and Jain and Meena [24] presented a cost investigation of a queueing model that includes an unreliable server and vacation periods. Utilizing the *particle swarm optimization* (PSO) approach, we have attempted to deal with the cost constraints in networking systems. This approach may sort through a very vast number of possible solutions to identify the most appropriate one. PSO has an additional benefit over a variety of optimisation strategies in that it does not require the objective function to be possible to differentiate.

Using this procedure, we first initiate a given population consisting of several particles or candidates. These candidates are then forced to travel inside the search space while adhering to the specified parameters and the goal function over their location and velocity. Every particle's fitness value is computed and to assess the values of the global best (gbest) and personal best (pbest) in more detail. The new gbest value is the particle whose pbest value is greater than gbest. This technique keeps on going until the predetermined number of iterations is reached. The algorithmic rule for PSO was first proposed by Kennedy and Eberhart [20]. The price optimization of a discrete-time RQ with SF utilizing this method has been examined by Upadhyaya [21]. Zhang et al.[22] examined set up cost and numerical answers for a single server recurrent model with state-dependent service using the PSO algorithmic approach. We have cited Malik et al.[23] investigation as it pertains to the operation of the PSO and GA algorithm.

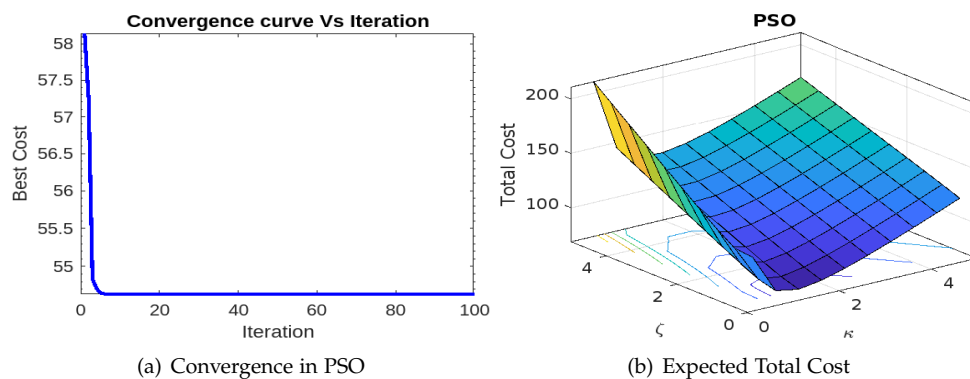


Figure 3: (a)2D and 3D visualization of PSO optimization

7.2. Convergence in PSO

Convergence holds significant importance within meta-heuristic optimization algorithms, representing the gradual improvement of potential solutions towards an optimal or near-optimal solution. The convergence pattern of an algorithm reflects its efficacy in exploring the solution space adeptly and moving closer to the global optimum. The findings from these figures suggest that employing the concept of a working vacation enhances the system's stability and reliability. This is due to the consistent availability of the server during this period. When a machine fails, the server responds quickly to the problem, but it provides a slower rate of service than when the machine is operating normally. Within PSO, particles converge toward the most optimal solution they are aware of, coming together as their movements become restricted and the finest solution steadies. Fig 3(a) Demonstrates that PSO achieves convergence towards the optimal cost. Fig 3(b)Displays the convexity and optimality of the cost function concerning the cost sets utilized in the optimization analysis.

8. CONCLUSION

In this research, a single server feedback retrial queueing system with starting failures, Bernoulli working vacations, and vacation interruptions was investigated. The number of customers in the system and its orbit are used to find the PGFs. This is done by using the "supplementary variable method". The average orbital queue length and the system average queue length have precise expressions. Numerical examples are used to verify the analytical conclusions. The mean busy period as well as other significant system performance indicators are determined. Also, numerical outcomes are compared to ANFIS. We have demonstrated how to optimize the functioning of a real-world service system using the PSO meta-heuristic algorithm. This suggested paradigm may be used in communication networks, supermarkets, management and production industries, etc. Basically, it is nearly impossible to construct a paradigm in which the server never defects or deactivates in any of these enormous sectors. As a result, this analysis is pertinent and in favour of scenarios in which a server can remain idle to maximise the consumption of resources. This model's construction helps to prevent the regular overcrowding issues that networking and communication systems suffer. The suggested model may be expanded in the future to incorporate other factors, such as modified vacation policy, randomized policy, consumer impatience, priorities, and setup times.

DECLARATIONS

- **Acknowledgments** : Not applicable.
- **Competing interest** :The authors declare that they have no competing interests.
- **Funding** :The authors declare that there is no funding source for this research.
- **Author's contributions** : All the authors made substantial contributions to the conception or design of the work.

REFERENCES

- [1] Gomez-Corral,A.(1999).Stochastic analysis of a single server retrial queue with general retrial times.*Naval Res. Logist.*, Vol.46 (5):561-581.
- [2] Falin,G.I. and Templeton, J.G.C.(1997).Retrial queues.*London:Chapman and Hall,CRC Press*,75.
- [3] Artalejo, J., and Corral,A.G.(2008).Retrial Queueing Systems, *Springer, Berlin, Germany*.
- [4] Artalejo,J.R.(2010).Accessible bibliography on retrial queues,progress in 2000-2009.*Mathematical and computer modelling*. 51:1071-81 .
- [5] Gao, Wang, Li.(2014). An $M/G/1$ retrial queue with general retrial times, working vacations and vacation interruption.*Asia-Paci J. Oper. Res.*,31 (2):1440006. <https://doi.org/10.1142/S0217595914400065>.
- [6] Gao.S., and Liu,Z.(2013). An $M/G/1$ queue with single working vacation and vacation interruption under Bernoulli schedule, *Applied Mathematical Modelling*.37(3):1564-1579.
- [7] Rajadurai,P., Saravananarajan,M.C. and Chandrasekaran,V.M.(2016). Analysis of an $M/G/1$ feedback retrial queue with unreliable server, non-persistent customers, single working vacation and vacation interruption.*Int. J. Services and Operations Management*. 24(2):235-266.
- [8] Rajadurai,P., Chandrasekaran,V.M. and Saravananarajan,M.C.(2016). Analysis of an $M^{[X]}/G/1$ unreliable retrial G-queue with orbital search and feedback under Bernoulli vacation schedule.*OPSEARCH*.53(1):197-223.
- [9] Rajadurai,P.,Saravananarajan,M.C. and Chandrasekaran,V.M.(2018). Analysis of an unreliable retrial G-queue with working vacations and vacation interruption under Bernoulli schedule.*Ain Shams Engineering Journal*.9(4):567-580. <http://DOI.10.1016/j.asej.2016.03.008>.
- [10] Mokaddis,G.S.,Metwally,S.A. and Zaki,B.M.(2007).A Feedback Retrial Queueing System with Starting Failures and Single Vacation.*Tamkang Journal of Science and Engineering*.10(3): 183-192.

- [11] Krishna Kumar,B.,Pavai Madheswari, S. and Vijayakumar,A.(2002).The M/G/1 re-trial queue with feedback and starting failures.*Applied Mathematical Modelling*26:1057-1075.[https://doi.org/10.1016/S0307-904X\(02\)00061-6](https://doi.org/10.1016/S0307-904X(02)00061-6).
- [12] P.Rajadurai.(2019).A study on M/G/1 preemptive priority retrial queue with Bernoulli working vacations and vacation interruption.*Int. J. Process Management and Benchmarking*.9 (2):193-215. <https://doi.org/10.1504/IJPMB.2019.099331>.
- [13] Rajadurai,P.,Saravananarajan,M. C. and Chandrasekaran, V. M.(2016).single server retrial queue with Bernoulli working vacation and vacation interruption.*International Journal of Applied Engineering Research*.11 (1).
- [14] Arivudainambi,D.,Godhandaraman,P. and Rajadurai,P.(2013).Performance analysis of a single server retrial queue with working vacation, *OPSEARCH*.51(3):434-462. <https://doi.org/10.1007/s12597-013-0154-1>.
- [15] Zhang,M., and Hou,Z.(2010).Performance analysis of M/G/1 queue with work-ing vacations and vacation interruption.*J. Comput. Appl. Math.*234(10):2977-2985.<https://doi.org/10.1016/j.cam.2010.04.010>.
- [16] Zhang, M., and Hou,Z.(2012). M/G/1 queue with single working vacation.*Journal of Appl. Math.Comput.*.39:221-234.
- [17] Rachita Sethi,Amrita Bhagat and Deepika Garg.(2019). ANFIS based Machine Repair Model with Control Policies and Working Vacation. *International Journal of Mathematical, Engineering and Management Sciences*.4(6):1522.<https://dx.doi.org/10.33889/IJMEMS.2019.4.6-120>.
- [18] Charu Bhargava and Madhu Jain.(2014). Unreliable multiserver queueing system with modified vacation policy, *OPSEARCH*, 2(51):159-182. DOI 10.1007/s12597-013-0138-1.
- [19] Radhika Agarwal,Divya Agarwal,Shweta Upadhyaya and Izhar Ahmad.(2023).Optimization of a stochastic model having erratic server with immediate or delayed repair,*Annals of Operations Research*.331(2):605-628. <https://doi.org/10.1007/s10479-022-04804-2>.
- [20] Kennedy, J.,and Eberhart, R.(1995). Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks* (Perth, Australia): pp. 1942-1948.
- [21] Upadhyaya, S.(2020). Cost optimization of a discrete-time retrial queue with Bernoulli feedback and starting failure. *Int. J. Ind. Syst. Eng.*36:165-196.
- [22] Zhou, M.,Liu, L.,Chai,X. and Wang,Z.(2018).Equilibrium strategies in a constant retrial queue with setup time and the N-policy. *Commun. Stat. Theory Methods*.49(7):1695-1711.
- [23] Malik, G.,Upadhyaya,S.and Sharma,S.(2021). Cost inspection of a Geo/G/1 retrial model using particle swarm optimization and Genetic algorithm. *Ain Shams Eng. J.* 12(2):2241-2254.
- [24] Jain, M., and Meena, R.K.(2020).Availability analysis and cost optimization of M/G/1 fault-tolerant machining system with imperfect fault coverage. *Arabian Journal for Science and Engineering*. 45(3):2281-2295.
- [25] Jain, M.,Kumar,P.and Meena,R.K.(2020).Fuzzy metrics and cost optimization of a fault tolerant system with vacationing and unreliable server. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-020-01951-x>.
- [26] Jain, M.and Kumar,A.(2023). Unreliable Server $M^{[X]}/G/1$ Retrial Feedback Queue with Balking, Working Vacation and Vacation Interruption. *Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.* 93: 57-73. <https://doi.org/10.1007/s40010-022-00777-w>.
- [27] Pazhani Bala Murugan,S., and Keerthana,R.(2023). An M/G/1 Feedback retrial queue with working vacation and a waiting server. *Journal of computational analysis and applications*.31(1).
- [28] Keerthiga, S.,and Indhira, K.(2023).Two phase of service in M/G/1 queueing system with retrial customers. *Journal of Analysis* .<https://doi.org/10.1007/s41478-023-00635-x>.
- [29] Agarwal, D., Agarwal,R. and Upadhyaya,S.(2024).Detection of optimal working vacation service rate for retrial priority G-queue with immediate Bernoulli feedback. *Results in Control and Optimization*.14:100397.<https://doi.org/10.1016/j.rico.2024.100397>

IMPROVING THE SPECTRAL EFFICIENCY IN DOWNLINK MULTIPLE USER MULTIPLE INPUT MULTIPLE OUTPUT TRANSMISSION FOR FIFTH GENERATION AND BEYOND WIRELESS COMMUNICATIONS

Abdulmujeeb Akajewole Masud, Donatus Uchechukwu Onyishi

•
Department of Electrical and Electronics Engineering
Federal University of Petroleum
Effurun, Nigeria

abdulmujeebakajewole@gmail.com, onyishi.donatus@fupre.edu.ng

Abstract

This research presents a solution in Multiple Input Multiple Output (MIMO) wireless systems to meet the growing demand for high data rates in cellular networks. Although MIMO systems offer greater capacity, the higher frequencies used have caused interference problems, especially for mobile User Equipment (UE). This research aims to reduce interference problems in the downlink of Multi-user MIMO (MU-MIMO) systems, with a specific focus on improving Quality of Service (QoS) metrics, such as outage probability and Signal-to-Interference plus Noise Ratio (SINR). Existing solutions to these challenges are complex due to the dynamic nature of the factors involved in modelling real-world scenarios. As such, an Improved Downlink MU-MIMO (ID-MU-MIMO) algorithm is developed as a solution to these problems. The ID-MU-MIMO method employs both single antenna users and multiple transmitter antennas. The performance of the suggested algorithm is compared to the IEEE 802.11ax standard specification and a previous research work for validation and evaluation. Performance measures considered to aid validation included outage probability, spectrum efficiency, and communication connection reliability. On this premise, the outcomes showed that the proposed ID-MU-MIMO scheme outperforms both the IEEE 802.11ax standard and current MD-MU-MIMO systems. In particular, compared to IEEE 802.11ax, the ID-MU-MIMO technique achieved a 7.71% reduction in interference. When compared to the performance of the random and uniform MD-MU-MIMO algorithms, the proposed ID-MU-MIMO scheme showed a reduction in interference in percentages of 8.90% and 2.28%, respectively. The ID-MU-MIMO scheme outperformed the random and uniform MD-MU-MIMO algorithms in terms of Signal-to-Interference Noise Ratio (SINR), outperforming them by 4.27% and 2.75%, respectively, and resource block use, outperforming them by 20.05% and 3.89%, respectively.

Keywords: ID-MU-MIMO, Interference, MU-MIMO, Resource Utilization, SINR

I. Introduction

The growing demand for wireless data, driven by the pervasive adoption of smart devices and mobile internet applications, is mirroring the exponential growth predicted by Martin Cooper's Law [1]. This law forecasts a doubling of wireless data usage every 2.5 years, a trend expected to accelerate in the coming years, with data rate requirements projected to surge by a staggering 5,000-fold by 2030 [2]. To address this growing demand, Fifth Generation (5G) technology emerged as a game-changer, surpassing the capabilities of its predecessor, Fourth Generation (4G), to meet the escalating needs [3]. One of its primary objective, as outlined in [4], is to achieve near-ubiquitous coverage while minimizing the likelihood of outages. This necessitates the development of innovative approaches that enhance spectral efficiency without compromising energy and bandwidth requirements [5]. Prominent standardization bodies, like the 3GPP and its affiliated groups, actively strive to establish standards that meet these objectives. These standards are designed to provide adaptable, easily accessible, and user-centric wireless data services, fulfilling the aforementioned commitments while adapting to evolving user requirements [6]. Consequently, technologies such as NOMA, massive MIMO, hybrid precoding, mm-wave, OFDM, beamforming, and D2D are utilized to enhance performance after data transmission [4], [7]. Since its inception in the era of 3G wireless networks, MIMO technology has continuously revolutionized the performance of wireless transceivers, enabling unprecedented data rates and spectral efficiency [8]. In cellular networks, base stations are categorized based on their ability to serve multiple User Equipment (UE) simultaneously, giving rise to SU-MIMO and MU-MIMO systems [9]. In SU-MIMO, the base station (BS) communicates with a single UE, transmitting one or more data streams using its antenna array. In contrast, MU-MIMO systems harness the power of BS antenna arrays to simultaneously broadcast multiple data streams to different UEs in distinct beams, sharing the same frequency resource without interference. mMIMO, a cutting-edge advancement of MIMO technology, leverages MU-MIMO principles to significantly enhance network capacity and user experience [10].

Recent research has delved into the downlink performance of MU-MIMO systems in 5G networks [11]. On this premise, this study focuses on improving downlink performance by analyzing MU-MIMO data transmission [12]. The categorization of user UE reception behavior based on inter-beam interference and the number of interference beams received is employed to achieve this improvement. Additionally, the network architecture is considered, and different scenarios within a specific time frame similar to the work of [13] are taken into account. These scenarios are influenced by users' mobility or stationariness relative to their sending and receiving devices. Although, several researcher have employed several systems to improve the downlink MU-MIMO systems, most of them are computationally complex. This complexity arises from the necessity of incorporating real-world factors such as interference, outages, consumption of energy, security, link reliability, and capacity limitations. Some of these factors stem from the shared utilization of the same frequency band by diverse technologies, including Bluetooth, Zigbee, and other WLAN systems [14]. To address these complexities, many studies have opted to trade-off computational complexity for in-depth analysis [10], [12], [15], or simplified factors like UE mobility to reduce complexity and focus on downlink Mu-MIMO [7]. To effectively serve the maximum number of UEs while considering their demands, locations, and mobility conditions, it is crucial to evaluate the available bandwidth, potential interference, and outage probability. This research, therefore, focuses on improving a downlink MU-MIMO technique for 5G networks that accounts for outage odds, enhances spectral efficiency, and strengthens link reliability. The rest of the paper discusses the followings: section 2 discusses literature that are related to the study. Section 3 delineates the methodology adopted in carrying out this research. Section 4 discusses the results obtained utilizing the ID-MU-MIMO algorithm. Section 5 concludes this paper.

II. Related Literature

This section reviews a few studies that have been shown to be helpful in enhancing Mu-MIMO's spectrum efficiency. Additionally, a summary of the methods employed by the assessed works to increase spectral efficiency, with a particular emphasis on Mu-MIMO outage mitigation, is discussed. With the knowledge gained from these works, this research was able to model an effective mechanism for MU-MIMO wireless communication system to limit outages in 5G and beyond wireless communications.

The energy efficiency of mmWave large MIMO systems was demonstrated through authors [4] evaluation of achievable sum-rate. Compared to traditional digital precoding techniques that assume a dedicated RF chain for each antenna and employ streamlined beamforming algorithms like ZF and MRC, mmWave large MIMO systems exhibited significant sum-rate and power reductions. The extended SOMP algorithm was proposed as a practical solution for hybrid precoding optimization, and its performance was shown to be near-optimal compared to standard MIMO. Massive MIMO technology has been found to significantly enhance system throughput when combined with basic signal processing techniques. The BER performance of SM and SSK schemes was analyzed, revealing that SM offers higher data rates but requires complex decoders and is more susceptible to errors, while SSK offers lower data rates but requires simpler decoders and exhibits lower BER. However, the impact of inter-beam interference was not considered in this study, which could potentially limit the effectiveness of strategies aimed at improving spectral efficiency.

The work of [15] explored the simultaneous alignment of Hermitian matrices representing desired signals and co-channel interference. Their approach simplified the analysis of critical performance measures like probability of outage, ergodic capacity, and spectral efficiency by leveraging joint unitary eigenvectors and their corresponding eigenvalues. These authors integrated digital baseband beamforming vectors into transmitter-side channel weight matrices, facilitating the representation of SINR in a standard quadratic form. Moreover, they devised a real scalar objective function to quantify the correlation loss linked with joint-diagonalization. Acknowledging the hardware limitations of mmWave systems, they employed baseband beamforming to optimize the objective function. The results delineated from the simulation showed that the proposed beamforming algorithm outperformed various nonlinear optimization methods in terms of time complexity and correlation assessment, surpassing, notably demonstrating efficiency in the "active-set" approach. While the targeted signal and co-channel interference were considered, inter-beam interference remained unaddressed. Moreover, the computational complexity was relatively substantial, potentially affecting the communication link's reliability.

The increasing number of transmit antennas and planned users on the same frequency response led to a significant rise in the computing complexity of outage probability and coverage probability calculations, as observed by [12]. To assess the coverage performance of downlink cellular MU-MIMO networks, the authors proposed combining exponential functions to approximate the complementary cumulative distribution function of received signal gain. Prony's technique was employed to simplify the calculation of expectations related to intra-cell and inter-cell interference. Simulation results based on this method indicated that downlink coverage probability remained unaffected by an increase in the number of base stations. While a key objective was to reduce the computational complexity of coverage probability determination, the method did not consider the SIR values of users connected to the same cell. These shortcomings are likely to hinder the method's performance in real-world scenarios.

Digital federated learning and over-the-air compression techniques were employed in [16] to enhance signal processing optimization. The authors formulated problems tailored to user

demands while minimizing Mean Square Error (MSE) of parameter vectors. These problems were addressed using block coordinated descent-based iterative methods. Moreover, the authors optimized the precoding matrices of the over-the-air compression system. The proposed methodology demonstrated superior performance compared to the standalone digital federated learning scheme, particularly with increasing user participation. However, a drawback of this approach was its heightened computational complexity, which could potentially jeopardize the communication link's reliability.

To enhance spectral efficiency and communication link reliability in MU-MIMO for 5G and future wireless systems [7], a mechanism was developed. Despite user demands, the system was designed to handle MU-MIMO in downlink with multiple antenna users, delivering optimal QoS to all the devices that are connected. Three user allocation scenarios were considered: random, uniform, and default allocation for antennas serving multiple users. The uplink and downlink were influenced by the random demand on these scenarios. Throughput utilization, network capacity, data rate for each device and the network, distances between users and antennas and from other devices, path loss, channel gain, and interference were additional variables estimated to model these scenarios. The uniform allocation algorithm outperformed default and random allocation algorithms in most trials, surpassing all three scenarios. In other words, RB utilization and SINR were more effective with uniform allocation. The authors did not consider user mobility, which could negatively impact communication link reliability in real-world scenarios.

In their work [11], the authors introduced an innovative, low-complexity approach combining multi-beamforming and Maximal Likelihood-Multi-User Detection (ML-MUD). They utilized a Radix Factorization-based FFT (RF-FFT) and integrated sub-detector systems to significantly reduce complexity without compromising error rate performance. The proposed architecture, known as RF-FFT - Multi-Beamforming (RF-FFT-MBF), holds immense promise in mitigating hardware complexity, energy consumption, and fulfil the throughput needs of 5G devices. To showcase the efficiency of the scaled ML-sub detector system on the downlink, simulation results of this detector were compared against conventional ML detectors. The outcomes delineated that the proposed detector outperformed the conventional ML detectors in terms of hardware and energy efficiency, boasting minimal performance.

Extensive research on downlink MU-MIMO systems has predominantly focused on ad hoc network topologies, with limited attention given to cellular networks and UE mobility. While some studies have explored cellular networks, they either disregard UE mobility or employ computationally demanding approaches. Despite significant efforts to address interference, throughput, SINR, outage probability, spectral efficiency, and link communication reliability, challenges like interference and high computational complexity persist. This research presents an improved MU-MIMO Downlink mechanism for 5G and beyond wireless cellular networks, leveraging on the work of [12]. The proposed mechanism is expected to enhance spectral efficiency, link communication reliability, as well as reduce the probability of outage in comparison to existing systems. Section 3 outlines the methods and system model designed to achieve these goals.

III. Methodology

The development process for the Improved Downlink MU-MIMO Downlink (ID-MU-MIMO) scheme is described in this section. The methods used for developing the system model for the ID-MU-MIMO algorithm are explained in subsections I to V.

I. ID-MU-MIMO Algorithm

Building upon the foundations established in [7] and [12], this work adopts a hybrid approach that merges techniques from both works to address the limitations of [7], namely its overlook for UE mobility, while simultaneously reducing computational complexity. The proposed ID-MU-MIMO algorithm is the culmination of these considerations. To achieve these objectives, this work expands the network scenarios for UEs beyond those presented in [7] by introducing two additional network case scenarios, which are represented in the steps outlined below.

II. Defining the Mobile and Stationary Events of UEs at the Start of Each Time Frame

In a real-world setting, it is assumed that UE distribution throughout a coverage region is random. Inter-beam interference and the quantity of received interference in a MU-MIMO broadcast are both impacted by the mobility status of the UEs within the coverage area. According to published findings, measuring the state of a channel continuously would probably be computationally challenging because the channel's state, the power delivered, etc., are all dynamic. As such, the data transfer to UEs is taken into account within a time frame to reduce this complexity. It is assumed, based on the literature, that the channel's condition and transmitted power will remain constant over this time. In this work, certain premises are adopted. It is simpler to look into co-channel and inter-beam interference at this specific time. A closed form outage probability assessment of an instantaneous MU-MIMO transmission is also permitted by this consideration. In light of the aforementioned, this work considers the placement of UEs in a coverage area to be randomly allocated according to Spatial Poisson Point Process (SPPP) with intensity (κ). The set of BS are represented as $S(\kappa) = \{B_i\}$, where B_i is the position of the BS, at an instant in time, i .

The received signal of the receiving UE, k , in the MU-MIMO data transmission phase is mathematically stated as follows [14, 17]:

$$y_k = \sqrt{\frac{P_t}{k}} D_k^{-\beta} \mathbf{h}_k^H \mathbf{w}_k s_k + \sum_{j \in K, j \neq k} \sqrt{\frac{P_t}{k}} D_k^{-\beta} \mathbf{h}_k^H \mathbf{w}_j s_j + \sum_{\tau \in \Phi / T_k} \sqrt{P_t |Z_\tau|^{-\beta}} \mathbf{g}_{\tau k} s_\tau + \eta_k \quad (1)$$

Where y_k represents the received signal by UE, k , P_t denotes the power transmitted by k , T_k denotes the BS, $\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ represents the channel gain vector between the k and T_k , $\mathbf{g}_{\tau, k} \in \mathbb{C}^{1 \times 1}$ represents the channel gain vector between the k and τ , $\mathbf{s}_k \in \mathbb{C}^{K \times 1}$ signifies the transmitted symbol from the BS, n_k denotes the AWGN with zero mean and variance of σ^2 , $\sqrt{\frac{P_t}{k}} D_k^{-\beta} \mathbf{h}_k^H \mathbf{w}_k s_k$ denotes the desired signal, $\sum_{j \in K, j \neq k} \sqrt{\frac{P_t}{k}} D_k^{-\beta} \mathbf{h}_k^H \mathbf{w}_j s_j$ denotes the inter-beam interference signal, $\sum_{\tau \in \Phi / T_k} \sqrt{P_t |Z_\tau|^{-\beta}} \mathbf{g}_{\tau k} s_\tau$ depicts the co-channel interference signal.

The presence or absence of inter-beam interference for a receiving UE, k , is subject to the variations, that is, the change or stationariness the channels of the wireless experiences with respect to the location of the BS during the time period. In the case of stationary UEs, the orthogonality of the beams ensures the elimination of inter-beam interference. However, for moving UEs, their own beams interfere with other beams, and the beams of stationary UEs interfere with the beams of moving UEs. To evaluate the performance of MU-MIMO transmission, we characterize the UE behaviour based on the occurrence and intensity of inter-beam interference. By defining a set of events that represent different UE behaviours, we simplify calculations and gain insight into the impact of inter-beam interference on MU-MIMO performance. The beam of a motionless UE interferes with the beams for M moving UEs when the event \mathcal{H}_2 with $0 < M < K$ occurs, but the beam of a moving UE interferes with the beams for $K-1$ UEs. In light of this, the

proposed measurement of the instantaneous received SINR for stationary and moving UEs is denoted as event \mathcal{H}_2^s and \mathcal{H}_2^m , respectively [12].

Case One

In this case, which is referred to as \mathcal{H}_1 both the UEs in the MIMO system are stationary during the considered time frame.

Case Two

Event \mathcal{H}_2 refers to a circumstance in which the UEs are mobile within the time period. For this event $M(1 < M \leq K)$. Since a mobile user was present at every time point in the study period, $M \geq 0$ in this instance. The probability that the instantaneous received SINR falls below the required SINR threshold is called the outage probability of a MU-MIMO transmission. The predicted probability of outage is given as in (2) as delineated in the work of [12].

$$P_k^{\text{exp}}(\xi_{\text{thres}}) = \sum_{i=1}^3 \wp(H_i) \wp(\xi_{k,\xi_i}^{\text{ins}} \leq \xi_{\text{thres}}) = \sum_{i=1}^2 \wp(H_i) \wp(\xi_{k,\xi_i}^{\text{ins}} \leq \xi_{\text{thres}}) + \left[\wp(H_3) \times \left[\wp(H_3^s) \wp(\xi_{K,H_3^s}^{\text{ins}} \leq \xi_{\text{thres}}) + \wp(H_3^m) \wp(\xi_{K,H_3^m}^{\text{ins}} \leq \xi_{\text{thres}}) \right] \right] \quad (2)$$

where $\wp(\mathcal{H}_i)$ denotes the likelihood of the event \mathcal{H}_i occurring and $\xi_{k,\mathcal{H}_i}^{\text{ins}}$ signifies the instantaneous SINR received by k at the time of the event \mathcal{H}_i . However, an event when both the UE and the BS station are mobile is not taken into account in this work's total of the outage probability weighted by the likelihood of defined occurrences. As a result, only a total of \mathcal{H}_2 events, as identified in cases one and two, can be accommodated by the probability of weighted events \mathcal{H}_i , which is decreased ($i-1$) as a result. The requirements as they apply to scenario one take into account the likelihood that the UE and the gNB are stationary, and as a result $M=0$. The second event, on the other hand, takes into account a scenario in which the UE is mobile and the gNB is stationary. In this case, $M \geq 0$ since a UE is always mobile at any point in the time fame considered.

When the two aforementioned events are taken into account and applied to the system chosen for this study, the events reduce from \mathcal{H}_3 to \mathcal{H}_2 , and the influence on the outage probability sum multiplied by the probabilities of the defined events in (2) results in (3).

$$P_k^{\text{exp}}(\xi_{\text{thres}}) = \sum_{i=1}^2 \wp(H_i) \wp(\xi_{k,\xi_i}^{\text{ins}} \leq \xi_{\text{thres}}) = \wp(H_i) \wp(\xi_{K,\xi_{\text{thres}}}^{\text{ins}}) \times \left[\wp(H_2) \wp(\xi_{K,H_2^s}^{\text{ins}} \leq \xi_{\text{thres}}) + \wp(H_2^m) \wp(\xi_{K,H_2^m}^{\text{ins}} \leq \xi_{\text{thres}}) \right] \quad (3)$$

The implications of the lower weighted sum affect both the performance factors that are examined and the probability of events.

III. System Model

To model the occurrences \mathcal{H}_1 and \mathcal{H}_2^M , it is crucial for this study to first replicate the MD-MU-MIMO scheme model before modifying it to take the mobility of the UEs into account. When taking into account the UEs' mobility, care must be given to elements that could raise the likelihood that a UE would suffer a loss of signal quality or be unable to establish a trustworthy connection because of elements like interference, fading, and channel circumstances. The term "outage probability" is used to describe this likelihood. When simulating the MD-MU-MIMO scheme, it is crucial to keep in mind that Mu-MIMO gives a great capacity boost by utilizing multiple antennas and supporting single-antenna users without the need to increase network

bandwidth. Literature studied earlier reveal that the integration of multiple antennas at the communication terminals significantly enhances the communication system's spectral efficiency and ensures a more reliable communication link. A MIMO system typically uses several broadcast and receive antennas. Each antenna, while intended to receive specific components of the signal, also captures indirect components intended for other antennas on the same channel. To maximize system performance, data intended for transmission is divided into separate, independent data streams, a number that is typically equal to or less than the number of available antennas. This design results in a linear increase in system capacity as the number of flows grows, aligning with the principles described in the Shannon-Hartley theorem for MIMO, denoted as (4) [7].

$$C \leq MB \log_2 \left(1 + \frac{S}{N} \right) \quad (4)$$

where C is the channel capacity in bps, B is the bandwidth in Hz, S is the average received signal power in Watts, and N is the average power of the noise in Watts. The maximum feasible data rate of a system is obtained by multiplying the number of pulse levels (M) by the system bandwidth and the base 2 logarithm of the SNR + 1 [7].

$$MaxDR = B \log_2 \left(1 + \frac{S}{N} \right) \quad (5)$$

In contrast, the Maximum Data Rate (MaxDR) rises roughly linearly when the SNR is low. Although, simply aiming for a high SNR is not very effective; a better strategy involves dividing the SNR throughout the streams, which multiplies the maximum data rate that can be transmitted. One method to achieve this distribution is through spread spectrum techniques like Code Division Multiple Access (CDMA), which spread the signal over a broader bandwidth. MU-MIMO technology enhances system capacity and communication reliability. MU-MIMO is particularly advantageous for the uplink since UE employs just a single antenna for transmission, reducing complexity. This work aims to leverage multiple antennas at both the user and base station (BS) ends of the communication link to enhance the achievable spectral efficiency and overall reliability of the communication system and link. Spectral efficiency (SE) refers to the total amount of data transmissions that can occur over a given bandwidth within a cell of a cellular network. The cell throughput, expressed in bit/s/Hz, is obtained by multiplying the SE by the bandwidth. Since the bandwidth is fixed, increasing cell throughput is the preferable solution. For the system to simultaneously support the maximum number of UEs within a cell while maintaining the same bandwidth, it is crucial to achieve higher spectral efficiency. The spectral efficiency of 5G New Radio (NR) (bits/sec/Hz) can be approximated as follows [7]:

$$SE = \frac{\text{Throughput (bps)}}{\text{Channel Bandwidth (Hz)}} \quad (6)$$

It is essential to follow the 3GPP standard suggestion in order to get accurate results from the computed spectral efficiency in a network. This involves doing an extensive throughput calculation that takes into account a variety of factors, including the bandwidth, number of MIMO layers, modulation type, and frequency range. Only one aggregated component carrier is required in order to guarantee reliable results. Following that, the channel bandwidth must be divided by the throughput. A variable number of users are served by 25 antennas that make up the ID-MU-MIMO architecture. The 5G NR Frequency Radio 2 (FR2) type's user count varies depending on the deployment environment. As a result of the proposed topology's consideration of both micro cells and metro cells, micro cells can accommodate a maximum of 256 users, whilst metro cells can accommodate more than 250 people. The objective of this ID-MU-MIMO technique is to determine

the optimal number of UEs that should be scheduled and served per cell within the network to maximize spectral efficiency. However, multi-cell systems create difficulties that require a variety of criteria and parameters for validation of results. The simulation takes into account several different variables, including pilot allocation, hardware, hardware configuration, and block lengths. In the simulated network design, the UL and DL transmission modes have a substantial impact on UE performance.

In an MU-MIMO system, UE equipped with M_a single antennas is connected to the BS using N antennas during the downlink. Different processing strategies are taken into consideration, including Zero-Forcing (ZF), Maximum Ratio (MR) combining/transmission, and full-Pilot ZF (P-ZF) strategy. In a fully distributed coordinated beamforming system, the P-ZF actively suppresses inter-cell interference. The usage of the same matrix for scheduled users who are in the waiting list is made possible by the assumption that each scheduled user's precoding matrix is almost similar. Additionally, every UE engages in single-user identification, and its bit rate is determined by its bandwidth and Signal-to-Noise Ratio (SNR), where f_l is essentially a function that characterizes the performance of the link [7].

$$R_b = B \times F_l(SNR) \quad (7)$$

This work addresses the complexities of multi-cell systems and considers various features and processing algorithms in the simulated network topology, with the ultimate goal of optimizing network performance. The SINR(x) function allows the definition of SINR based on the user's location. Spectral efficiency at a UE location is determined by the relationship between the bit rate and the user bandwidth at that location. Each BS serving a cell allocates subcarriers of the total bandwidth to individual users based on the methodology described above. This assignment ensures that different users within the same cell, served by the same BS, receive different subsets of subcarriers. Consequently, since each BS transmits at a constant power, users within the cell experience interference from their respective base stations.

Modifying the above equations to consider events (\mathcal{H}_2^M) that take into account the mobility of the UEs

$$C \leq M_a B \log_2 \left(1 + \xi_{k, \mathcal{H}_2^m}^{ins} \right) \quad (8)$$

where $\xi_{k, \mathcal{H}_2^m}^{ins}$ is defined as Eq. 8

$$\xi_{k, \mathcal{H}_2^m}^{ins} = \frac{P_t / K D_k^{-\beta} |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sigma^2 + \phi \sum_{j \in U} (P_t / K) D_k^{-\beta} |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sum_{\tau \in \Phi / T_k} P_t |Z_\tau|^{-\beta} |g_{\phi, k}|^2} \quad (9)$$

where $P_t / K D_k^{-\beta}$ represents desired signal power, $|\mathbf{h}_k^H \mathbf{w}_k|^2$ denotes the squared magnitude of the channel coefficient between the transmitter and the intended receiver, multiplied by the squared magnitude of the weight vector for the intended user, σ^2 represents noise power, $\phi \sum_{j \in U} (P_t / K) D_k^{-\beta} |\mathbf{h}_k^H \mathbf{w}_j|^2$ represents the interference caused by other users in the system, and $\sum_{\tau \in \Phi / T_k} P_t |Z_\tau|^{-\beta} |g_{\phi, k}|^2$ represents the interference caused by other cells in the system, particularly focusing on the antennas in those cells that interfere with the intended receiver.

As such, the MaxDR computed is defined as in Eq. 10

$$MaxDR = B \log_2 \left(1 + \xi_{k, \mathcal{H}_2^m}^{ins} \right) \quad (10)$$

Furthermore, the definition of spectral efficiency (SE) is as follows:

$$SE = \frac{iTh_{tot}^{exp}}{\text{Channel Bandwidth (Hz)}} \quad (11)$$

Since the proposed expected throughput (iTh_{tot}^{exp}) proposed is denoted in Eq. 12 as:

$$iThr_{tot}^{exp} = \sum_{k=1}^K \left(\wp(\mathcal{H}_1)Thr_{k,\mathcal{H}_1}^{exp} + \wp(\mathcal{H}_2) \left(\wp(\mathcal{H}_2^s)Thr_{k,\mathcal{H}_2^s}^{exp} + \wp(\mathcal{H}_2^m)Thr_{k,\mathcal{H}_2^m}^{exp} \right) \right) \quad (12)$$

As such, substituting (12) into (11) yields:

$$SE = \frac{\sum_{k=1}^K \left(\wp(\mathcal{H}_1)Thr_{k,\mathcal{H}_1}^{exp} + \wp(\mathcal{H}_2) \left(\wp(\mathcal{H}_2^s)Thr_{k,\mathcal{H}_2^s}^{exp} + \wp(\mathcal{H}_2^m)Thr_{k,\mathcal{H}_2^m}^{exp} \right) \right)}{\text{Channel Bandwidth (Hz)}} \quad (13)$$

While the resource block usage is given in Eq. 14 as:

$$R_b = B \times F_l \left(\xi_{k,\mathcal{H}_2^m}^{ins} \right) \quad (14)$$

Equation (8) to (14) form some of the fundamental algorithms used to model the ID-MU-MIMO algorithm.

Equation (15) provides the estimated probability of outage for the instantaneous MU-MIMO transmission. The average of the outage probabilities for the cases of the UEs' events is known as the expected outage probability.

$$P_k^{exp}(\xi_{thres}) = \sum_{i=1}^2 \wp(H_i) \wp(\xi_{k,\xi_i}^{ins} \leq \xi_{thres}) = \wp(H_i) \wp(\xi_{K,\xi_{thres}}^{ins}) \times \left[\wp(H_2) \wp(\xi_{K,H_2^s}^{ins} \leq \xi_{thres}) + \wp(H_2^m) \wp(\xi_{K,H_2^m}^{ins} \leq \xi_{thres}) \right] \quad (15)$$

Where $P_k^{exp}(\xi_{thres})$ is the outage probability (P_{outage}) based on a given threshold (ξ_{thres}), $\wp(\mathcal{H}_i)$ represents the probability of event \mathcal{H}_i given as $P(M)$ and $\xi_{k,\mathcal{H}_i^s}^{ins}$ depicts the instantaneous received SINR of the receiving node (k) when the event \mathcal{H}_i occurs given as $P((SINR < \xi_{thres}) | M)$.

The ramifications of the reduced weighted sum impact the probability of events and the performance factors measured. Outage probability represents the probability that a wireless communication link will fail due to various factors, including interference and signal quality. Equations (16) and (17) evaluate the PDF on event \mathcal{H}_i using equations (16) and (17).

$$\wp(\mathcal{H}_1) = \left(P_{state} e^{-\theta T_{frame}} \right)^{K+1} \quad (16)$$

Where $\wp(\mathcal{H}_1)$ represents the outage for \mathcal{H}_1 , P_{state} is a parameter denoting the probability that a specific state (e.g., a certain channel state or interference level) occurs during the communication. $e^{-\theta T_{frame}}$ denotes the exponential function where T_{frame} represents the duration of time for which the communication is observed. And $K + 1$ is the exponent that models the outage probability for $K+1$ receiving users.

$$\wp(\mathcal{H}_2) = \left(P_{state} e^{-\theta T_{frame}} \right)^{K-M+1} \left(1 - P_{state} e^{-\theta T_{frame}} \right)^M \quad (17)$$

Equation (17) models the outage probability for a more complex scenario where there are $K+1$ users in the MU-MIMO system, and it depends on the probability of a specific state (P_{state}), the time duration (T_{frame}), the number of users ($K+1$), and the Mobile UEs. These equations are crucial for assessing the system's performance and understanding the probability of outage in complex wireless scenarios.

These equations are the equations for the PDFs on event \mathcal{H}_i . The PDFs are the probabilities that the UEs are in good or bad channel condition, given that the UEs are in event \mathcal{H}_i . The outage probability of instantaneous MU-MIMO transmission using the appropriate equation, depending on the value of ϕ . If $\phi = 0$, then the outage probability is obtained using equations (18).

$$P_k^{ins}(\xi_{thres}) = 1 - \sum_{m=0}^{N_i-K} \frac{(-\omega)^m}{m!} \frac{(d)^m}{d\omega^m} \mathcal{L}_{I_{k,CCI}(\omega)} \quad (18)$$

Where $\mathcal{L}_{\rho I_{k,IBI}}(\omega)$ signifies the Laplace transform of the interference seen by user K due to Inter Block Interference (IBI), scaled by ρ . Equation (16) combines both IBI and ICI to compute the outage probability for ξ_{thres} . It quantified the probability that ξ_{thres} experiences by the k -th user when it falls below the defined ξ_{thres} due to interference.

To enhance the pragmatism of the simulation studies, additional complex channel models that account for realistic propagation conditions, including Non-Line-of-Sight (NLOS) components and variable mobility patterns were considered by further modifying the Laplace transforms $\mathcal{L}_{\rho I_{k,IBI}}(\omega)$ and $\mathcal{L}_{I_{k,CCI}}(\omega)$. As such, the parameter ρ_{NLOS} was introduced to capture the effects of NLOS components and $f_{mobility}(t)$ was introduced to represent the mobility patterns of users over time (t). Hence, the redefined Laplace are:

$$\mathcal{L}_{\rho LOS I_{k,IBI}}(\omega) = \mathcal{L}\{\rho LOS \cdot I_{k,IBI}(t)\} \quad (19)$$

$$\mathcal{L}_{k,CCI}(\omega, t) = \mathcal{L}\{I_{k,CCI}(t) \cdot f_{mobility}(t)\} \quad (20)$$

Hence, equation (18) is redefined as:

$$P_k^{ins}(\xi_{thres}) = 1 - \sum_{m=0}^{N_i-K} \frac{(-\omega)^m}{m!} \frac{(d)^m}{d\omega^m} \mathcal{L}\{\rho NLOS \cdot I_{k,IBI}(t)\} \cdot \mathcal{L}\{I_{k,CCI}(t) \cdot f_{mobility}(t)\} \quad (21)$$

IV. Algorithm of the ID-MU-MIMO Scheme

The ID-MU-MIMO algorithm is defined in Algorithm 1. The network design in the algorithm takes into account the mobility states of the user equipment. The MU-MIMO system's overall performance in a realistic environmental context is significantly influenced by the movement of the user equipment. Consequently, the suggested network model regards UEs as being distributed at random over the BS's coverage area. Both of the ID-MU-MIMO algorithm's cases are delineated in Algorithm 1. The first event denoted as \mathcal{H}_1 indicates that the UEs are immobile at the start of a set time frame and do not move during the duration not less than T_{Frame} . The second scenario, \mathcal{H}_2

considers the movement of the UE during T_{Frame} . The input parameters required to simulate and model the network topology of the wireless cellular network encompass the variables associated with mobility events. In the case of UEs linked to the gNB, the work assesses the mobility state of these events prior to calculating the resource block of each user. Concerning the two states, the resources are computed for one in which the effects of interference are removed, that is, where $\phi = 0$ and another in which the interference is considered, that is, $\phi = 1$. Whether inter-beam interference is taken into account or not, the goal of this modification is to reduce processing time. An outage in the network would be less likely using the suggested ID-MU-MIMO mechanism.

Algorithm 1 ID-MU-MIMO algorithm

Input $\mathcal{H}_i, T_{Frame}, i, \mathcal{H}_2^s$ and \mathcal{H}_2^m

Input Modulation type, bandwidth, Number of MIMO layers, Bandwidth, frequency range

Output Outage probabilities, spectral efficiency

- 1: **For** spatial distribution of users at time instant, U_i in the simulated network topology are uniformly or randomly distributed
 - 2: **For** each U_i
 - 3 **If** i is connected
 - 4 **For** each U_i trying to connect
 - 5 Utilizing Eq. (6) and Eq. (12), determine the maximum data rate the UE is capable of achieving
 - 6 $i++$
 - 7 **End for**
 - 8 **Else**
 - 9 $i++$
 - 10 Evaluate the expected outage probability using equation Eq. (3)
 - 11 Compute PDF on events \mathcal{H}_i
 - 12 **If** $\phi = 0$
 - 13 Compute outage probability
 - 14 Determine the RB's for each user using Eq. (6)
 - 15 **Else**
 - 16 **If** $\phi = 1$ and $0 < M < K$
 - 17 Compute outage probability
 - 18 Compute each user's RB using Eq. (12)
 - 19 **End if**
 - 20 **End if**
 - 21 **End if**
 - 22 **End for**
 - 23 **For** each cell topology
 - 24 Calculate BSs power allocated to the connected users
 - 25 Contrast performance with existing approach
 - 26 $K++$
 - 27 **End for**
 - 28 **End for**
-

V. Simulation Parameters

The simulation parameters adopted for the ID-MU-MIMO mechanism are presented in Table 1. These values are used to evaluate the performance of the conventional IEEE 802.11ax approach, the mechanism for downlink MU-MIMO, and the modified ID-MU-MIMO mechanism adopted in this research.

Table 1: *Extracted Parameters* [7, 12]

Parameters	Values
Bandwidth, B	160MHz
Capacity speed, v_c	3.39 Gb/s
Carrier bandwidth, B_c	400MHz
Downlink rates	20Gbps
Link rate	867Mbps
Path loss exponent	4
Mobile receiving nodes	1, $K/2$, K
Receiver noise power	-32dBm
Resource block frames, DL	275
Resource block frames, UL	275
Number of discrete rates	4
Average time interval	1s
Scenario	4 antennas AP
SINR threshold	-30dB
Sub Carrier Spacing	120KHz
Users for each scenarios	200, 250, 300
Distance between Tx and Rx, R_x	100m

IV. Results and Discussion

A thorough overview of the study's results is provided in this section by the presenting findings and their impacts.

I. Interference vs Users

Figure 1 represents the effect of interference as it affects the performance of the UEs in the MU-MIMO system. The events considered to weigh the impact are $H_1 = 1 - (M/K)$ and H_2^M at $M (0 < M < K)$ where H_1 represents the event where the UEs are stationary and H_2^M represents the events where some or all of the receiving UEs (K) are mobile. The impact of this event as applied in the developed scheme, the IEEE 802.11ax scheme, the random MD-MU-MIMO scheme and the uniform MD-MU-MIMO scheme are represented in Figure 1. In all schemes, both the BS and UEs deployed are considered to be stationary. For the receiving UEs, inter-beam interference is present or absent depending on whether the wireless channel changes with the UE's locations over the predetermined time period. Orthogonality of beams and the elimination of inter-beam interference are both possible because all nodes are kept fixed throughout the time period. Since the ID-MU-MIMO algorithm takes the mobility of the UEs into consideration, it examines the event H_2^M where some nodes can be fixed and others mobile. Here, $0 < M < K$ is appropriate. Therefore, $K-1$ UEs in the system could cause inter-beam interference with mobile UE (M). For both IEEE 802.11ax and the MD-MU-MIMO system, user deployment of 200, 250, and 300 are taken into account consistently. For the ID-MU-MIMO algorithm, the UEs are randomly deployed and this is also considered for one of the deployment scenarios for the MD-MU-MIMO scheme.

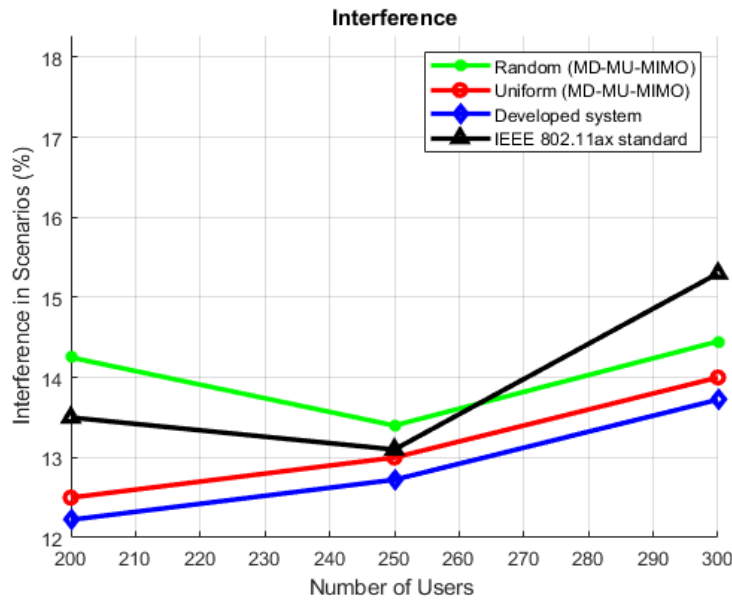


Figure 1: Interference vs Users

The performance of the designed ID-MU-MIMO scheme is shown in Figure 1, and it takes into account the aforementioned deployment situations as well as the UEs' mobility in order to model the system to actual environmental settings. The IEEE802.11ax algorithm and the existing MD-MU-MIMO algorithm did not take the UE's mobility into account. Figure 1 shows that, for the existing random MD-MU-MIMO algorithm, interference values decreased from 14.4% to 13.5% between the use cases distributions of users from 200 to 250, and for the IEEE 802.11ax values, interference values decreased from 13.5% to 13.1%. The source of this behaviour is a change in the measured distances between users placed in the network due to the random spatial deployment of users. As a result, there will inevitably be fluctuating channel conditions that could either raise or decrease interference for network users, as evidenced by the work's random reduction of interference. However, it is typically noted that when the number of users deployed in the system increases, interference of the system within the provided time frame generally increases for the system. In comparison to the IEEE 802.11ax protocol, the percentage gains in interference for the random and uniform distributions of users, respectively, were found to be 4.98% and 5.55% from the results obtained per use case scenario taken into account for the users. Additionally, comparing the ID-MU-MIMO method to the IEEE 802.11ax scheme, interference was also improved by 7.71%. The percentage reductions obtained for the designed and existing schemes both demonstrated that they outperform the current IEEE 802.11ax system. When compared to the results of the random and uniform MD-MU-MIMO algorithms, the created ID-MU-MIMO scheme demonstrated interference reduction in percentages of 8.90% and 2.28%, respectively.

II. SINR vs Users

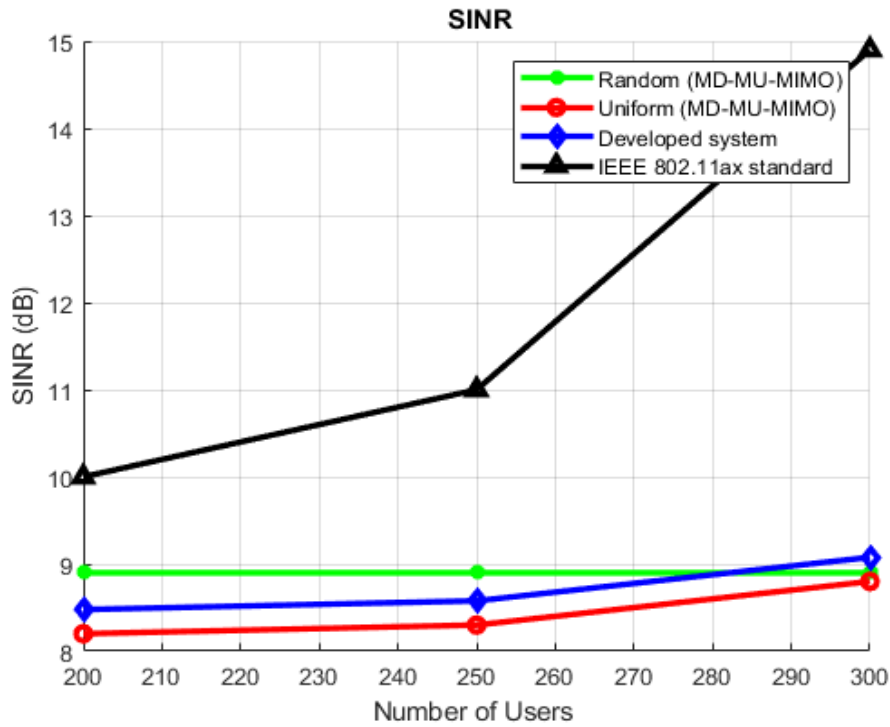


Figure 2: SINR vs Users

In maintaining the conditions for H_1 and H_2^m , Figure 2, illustrates the change in SINR, which revealed a considerable increase in SINR percentages from 11.0% to 14.9% in the IEEE 802.11ax values between the use case distributions of users from 200 to 250. This is probably caused by the low SINR. As a result, the maximum data rate rises approximately linearly. However, focusing just on high SINR percentages is ineffective because there is a greater likelihood of signal overlap and interference when more users broadcast at the same time while using the same time and frequency resources. Additionally, it was found that the performance of the random user deployment over a rising user count remained steady in this iteration of the SINR performance for the MD-MU-MIMO scheme. The interference as seen in this iteration of the simulation is probably low relative to the noise, which is the likely explanation for this stability as the focus is on the receiving UEs rather than the transmitting BS. As a result, the noise power and channel gain, which are unaffected by the user count, dictate the SINR. In comparison to the IEEE 802.11ax method, percentage gains in SINR of 23.45% and 27.43% for the random and uniform distribution of users, respectively, were seen from the findings produced per use case scenario taken into account for the users. Additionally, as compared to the IEEE 802.11ax system, the ID-MU-MIMO technique demonstrated a SINR improvement of 22.55%. The percentage reductions obtained for the designed and existing schemes both demonstrated that they outperform the current IEEE 802.11ax system. When the performance of the new ID-MU-MIMO scheme was compared to that of the random and uniform MD-MU-MIMO algorithms, the SINR was reduced by percentages of 4.27% and 2.75%, respectively.

III. Resource Block Usage vs Users

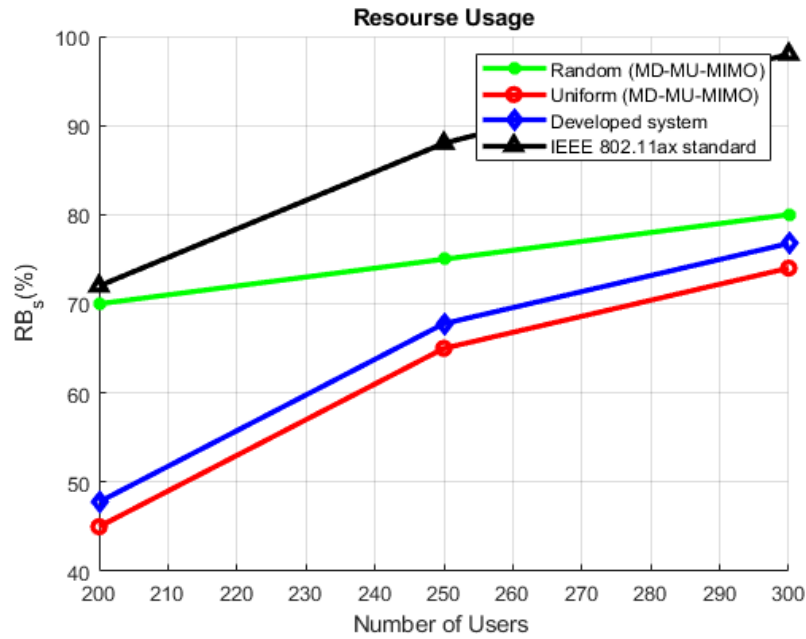


Figure 3: Resource Block Usage vs Users

In Figure 3, the results obtained per use case scenario considered for users, percentage improvements of 11.65% and 28.63% were observed in the use of resource blocks for random and uniform distribution of users respectively, compared to the IEEE 802.11ax scheme. Furthermore, the ID-MU-MIMO scheme showed a 25.87% improvement in interference compared to the IEEE 802.11ax scheme. The percentage reduction obtained from the existing and developed scheme showed that both improve over the existing IEEE 802.11ax scheme. The developed ID-MU-MIMO scheme showed a reduction in resource usage of 20.05% and 3.89% compared to the performance of the random and uniform MD-MU-MIMO algorithm, respectively.

V. Conclusion

Using a closed-form outage probability of a discrete time frame of MU-MIMO transmission, this work presented an effective mechanism for downlink MU-MIMO for 5G networks. This work examined the interference, SINR, and resource block usage over a specified time period, the outage probability of the receiving UEs in a specified area, and spectral efficiency in a modified MU-MIMO system to confirm the model's performance. The outcomes showed that the proposed ID-MU-MIMO scheme outperforms both the IEEE 802.11ax standard and current MD-MU-MIMO systems. In particular, compared to IEEE 802.11ax, the ID-MU-MIMO technique achieved a 7.71% reduction in interference. When compared to the performance of the random and uniform MD-MU-MIMO algorithms, the proposed ID-MU-MIMO scheme showed a reduction in interference in percentages of 8.90% and 2.28%, respectively. The ID-MU-MIMO scheme outperformed the random and uniform MD-MU-MIMO algorithms in terms of SINR, outperforming them by 4.27% and 2.75%, respectively, and resource block use, outperforming them by 20.05% and 3.89%, respectively. Future work is recommended to examine how a use case scenario of increasing the number of receiving UEs (K) can impact the probability of outage over a defined time frame and over a defined area, as well as determining the spectral efficiency under the conditions considered for a higher number of receiving UEs. Additionally, Machine Learning (ML) approaches can be adopted to modify the ID-MU-MIMO algorithm to dynamically adjust to varying network conditions and UE densities in real time to maximize performance under varying conditions.

References

- [1] Struzak, R. Tjelta, T. and Borrego, J. P. (2015). On Radio-frequency Spectrum Management. *URSI Radio Science Bulletin*, 354: 11-35.
- [2] Dala, P. S. V. Dester, P. S. Soares, P. F. M. Gomes Silva, D. and De Figueiredo, F. A. P. Tejerina, G. R. L. ... and Cardieri, P. (2023). Emerging MIMO Technologies for 6G Networks. *Sensors*, 23, 4: 1-20.
- [3] Wong, V. W., Schober, R. and Ng, D. W. K. (Eds.). (2017). *Key technologies for 5G wireless systems*. Cambridge university press.
- [4] Nalband, A. H., Sarvagya, M. and Ahmed, M. R. (2020). Optimal hybrid precoding for millimeter wave massive MIMO systems. *Procedia Computer Science*, 171: 810-819.
- [5] Singh, J. and Kedia, D. (2020). Performance improvement in large-scale MU-MIMO system with multiple antennas on user side in a single-cell downlink system. *Arabian Journal for Science and Engineering*, 45(8): 6769-6789.
- [6] Akpakwu, G. A., Silva, B. J., Hancke, G. P. and Abu-Mahfouz, A. M. (2017). A survey on 5G networks for the Internet of Things: Communication technologies and challenges. *IEEE access*, 6, 3619-3647.
- [7] Barri, E., Bouras, C., Kokkinos, V. and Koukouvela, A. (2021). A Mechanism for Improving the Spectral Efficiency in mu-MIMO for 5G and Beyond Networks. In *Proceedings of the 19th ACM International Symposium on Mobility Management and Wireless Access*, 11-16.
- [8] Albreem, M. A., Juntti, M. and Shahabuddin, S. (2019). Massive MIMO detection techniques: A survey. *IEEE Communications Surveys & Tutorials*, 21(4): 3109-3132.
- [9] Nguyen, M. (2018). Massive MIMO: a survey of benefits and challenges. *ICSES Trans. Comput. Hardw. Electr. Eng*, 4: 1-4.
- [10] Arsal, A., Civanlar, M. R. and Uysal, M. (2021). Coverage analysis of downlink MU-MIMO cellular networks. *IEEE Communications Letters*, 25(9): 2859-2863.
- [11] Kinol, A. M. J., Nisha, A. S. A., Marshiana, D. and Krishnamoorthy, N. R. (2022). Hybrid Multi Beamforming and Multi-User Detection Technique for MU MIMO System. *Wireless Personal Communications*, 124(4): 3375-3385.
- [12] Zhang, W. and Jiang, S. (2021). Effect of node mobility on MU-MIMO transmissions in mobile Ad Hoc networks. *Wireless Communications and Mobile Computing*, 2021: 1-9.
- [13] Spencer, Q. H., Peel, C. B., Swindlehurst, A. L. and Haardt, M. (2004). An introduction to the multi-user MIMO downlink. *IEEE communications Magazine*, 42(10): 60-67.
- [14] Lee, T., Kim, H. S., Park, S. and Bahk, S. (2014, June). Mitigation of sounding pilot contamination in massive MIMO systems. In *2014 IEEE International Conference on Communications (ICC)*, 1191-1196.
- [15] Hassan, A. K., Moinuddin, M., Al-Saggaf, U. M., Aldayel, O., Davidson, T. N. and Al-Naffouri, T. Y. (2020). Performance analysis and joint statistical beamformer design for multi-user MIMO systems. *IEEE Communications Letters*, 24(10): 2152-2156.
- [16] Huh, M., Yu, D. and Park, S. H. (2021). Signal processing optimization for federated learning over multi-user MIMO uplink channel. In *2021 International Conference on Information Networking (ICOIN)*, 495-498.
- [17] Masud, A. A., Uchekukwu, O. D., Adikpe, A. O. and Ibikunle, F. (2023). Mitigating Interference and Improving the SINR in a Discrete Time Frame of a Downlink MU-MIMO Transmission in 5G and Beyond Wireless Networks. In *2023 International Conference on Science, Engineering and Business for Sustainable Development Goals (SEB-SDG)*, 1: 1-6.

ENHANCING PROCESS CAPABILITY ANALYSIS FOR LOGNORMAL DATA UTILIZING BOX COX TRANSFORMATION AND GOODNESS OF FIT TESTS

J. Krishnan¹ and R. Vijayaraghavan²

(1). Department of Mathematics, Sri Krishna Adithya College of Arts and Science
Coimbatore – 641042, Tamil Nadu, INDIA

(2). Department of Statistics, Bharathiar University, Coimbatore 641 046,
Tamil Nadu, INDIA

¹krrishme92@gmail.com, ²vijaystatbu@gmail.com

Abstract

Process capability analysis is a valuable tool in quality assurance, but deviations from normal distribution necessitate adjustments to basic process capability indices. Process control literature offers solutions for non-normality, with data transformation being a common approach. The Box-Cox transformation (BCT) is often used to normalize non-normal data, relying on maximum likelihood estimation (MLE) to determine the transformation parameter, lambda. Alternative methods exist for estimating the single transformation parameter lambda, employing goodness-of-fit tests instead of the MLE method. This study explores two expressions within the Box-Cox transformation (BCT), encompassing both optimal and rounded values of lambda. The primary goal is to identify an effective method for transforming non-normal data into a distribution closer to normality through goodness-of-fit tests, aiming to obtain accurate estimates for process capability analysis in alignment with six sigma standards. Furthermore, this study focuses on the influence of utilizing both optimal and rounded values of lambda when transforming non-normal data to normal, and how these lambda values impact the estimates of process capability analysis. The findings reveal that methods such as Shapiro-Wilk's (SW) and Artificial Covariate (AC) outperform the MLE method. Moreover, employing the optimal lambda value during data transformation leads to improved estimates of process capability. Data simulation and analysis were conducted using Minitab software and the R programming language.

Keywords: Goodness of fit tests, Box-Cox Transformation, Asymmetric, MLE, Lognormal distribution, Six sigma.

I. Introduction

Process capability indices (PCIs) are essential tools in quality control, commonly utilized across manufacturing industries to ensure processes meet required standards. Process capability analysis (PCA) evaluates how effectively a manufacturing process adheres to specified targets. However, traditional PCIs assume a normal distribution, which may inaccurately assess non-normal processes. Kane (1986) suggests that transforming data to preserve a somewhat normal distribution improves the accuracy of process capability analysis [5]. Empirical studies have shown that transformed data yields superior results compared to original data [4]. Based on many literature surveys transformation methods, especially for non-normal distributions like Lognormal and Weibull, consistently outperform Non-Transformation (NT) methods. NT methods are inadequate for assessing process capability when distributions deviate significantly from normal [15]. Hence, transformation methods are preferred, as they provide more reliable assessments,

even for distributions distanced from normality.

In process capability analysis (PCA), the variability of a process is measured using the standard deviation. This variability can be divided into short-term and long-term variations. Short-term variability is determined by the estimated standard deviation obtained from random sample observations, which is then used in calculating process capability indices. On the other hand, long-term variability is assessed for computing process performance indices. Consequently, capability indices are computed using short-term variation, while performance indices utilize all data points, considering long-term variation. The commonly used capability indices are denoted as C_p and C_{pk} , while the respective performance indices are represented as P_p and P_{pk} . Various methods for handling non-normality in calculating process capability indices are discussed in [13]. Among these, the most widely applied indices in the manufacturing industry are the process capability index C_p and process capability ratio C_{pk} , as shown in Table 1 below, along with their respective performance indices. Here, \bar{x} denotes the sample mean, USL refers to the upper specification limit, and LSL indicates the lower specification limit.

Table 1: Process Capability and Process Performance Indices

Process capability indices	Process performance indices
$C_p = \frac{USL - LSL}{6\sigma_W}$	$P_p = \frac{USL - LSL}{6\sigma_{overall}}$
$C_{pk} = \min(C_{PU}, C_{PL})$	$P_{pk} = \min(C_{PU}, C_{PL})$
$C_{PU} = \frac{USL - \bar{x}}{3\sigma_W}, C_{PL} = \frac{\bar{x} - LSL}{3\sigma_W}$	$P_{PU} = \frac{USL - \bar{x}}{3\sigma_{overall}}, P_{PL} = \frac{\bar{x} - LSL}{3\sigma_{overall}}$

In [2], researchers employed the method of maximum likelihood estimation (MLE) to determine the optimal parameter λ in the Box-Cox transformation. Other approaches to the MLE methods, which rely on goodness of fit tests (specifically normality tests), were developed in [1], [3], [9], [10] and [14]. Through an examination of the impact of transforming non-normal data into normal data using different goodness of fit tests, [3] illustrated that the MLE method for estimating the λ parameter in BCT could be biased and inefficient. Furthermore, as indicated in [18] employing various goodness of fit tests instead of the MLE method for estimating the BCT parameter λ leads to improved estimates of process capability and process performance for non-normal data. The effectiveness of different goodness of fit tests was also assessed in [3] using various error measures, estimates of process capability and process performance indices, and defective parts per million (PPM) products. The results of different goodness of fits tests are recorded and presented to help the practitioner to choose the method which will produce the improvised results in various asymmetric situations, *viz.*, low, moderate and high. Thus, the objectives of this paper is to examine the effectiveness of the different goodness of fit tests involving transformation of non-normal data into normal data using BCT and to recommend a superior test that will produce higher values of process capability with minimum of error and PPM values particularly, for lognormal distribution. Additionally this paper focuses on the impact of optimal and rounded value of transforming parameter λ in BCT. It also verifies whether the proposed method produce the results within the standard of six sigma level.

II. Methodology

Converting non-normal data into a normal distribution is a common practice when observed data fail to meet normality assumptions. Several methods are employed for this purpose in practical applications, including Johnson's system of transformation (JST), Box-Cox transformation (BCT), and Rosenblatt transformation (RT). While both JST and BCT approaches are effective, BCT is generally preferred over JST, particularly in situations where computer-assisted analysis is

available, as it tends to outperform other methods [12]. Additionally, BCT is noted for its superior accuracy and precision compared to the JST method. BCT offers a range of power transformations designed to optimally normalize specific variables. According to [2], the BCT method transforms non-normal data into normal data for positive response variable x , as expressed below:

$$x^\lambda = \begin{cases} \frac{x^\lambda - 1}{\lambda}, & \text{for } \lambda \neq 0 \\ \log x, & \text{for } \lambda = 0 \end{cases} \quad (1)$$

It may be noted that since an analysis of variance is unchanged by a linear transformation, the expressions given (1) is equivalent to

$$x^\lambda = \begin{cases} x^\lambda, & \text{for } \lambda \neq 0 \\ \log x, & \text{for } \lambda = 0 \end{cases} \quad (2)$$

The form (1) is slightly preferable for theoretical analysis because it is continuous at $\lambda = 0$, refer [2]. The major effort in Box-Cox transformation is connected to the transformation X to X^λ , with the parameter λ possibly a vector describing a specific transformation. A single transforming parameter λ is the main source of dependence for this family of transformations, and its value is determined using maximum likelihood estimation [2].

The results of the earlier studies presented in the literature, particularly in [1], [6], [9], [10], [13], [14], [15] and [17] would be useful to understand the significance of tests of goodness of fit while transforming non-normal data into normal data. The estimation of λ is done through various goodness of tests for normality, that are available in the literature, which includes tests, such as Shapiro - Wilk (SW), Anderson Darling (AD), Cramer Von Mises (CVM), Pearson Chi-square (PC), Shapiro - Francia (SF), Lillefors (Kolmogorov - Simirnov) (LT / KS), Jarque - Bera (JB), and artificial covariate method (AC). Additionally, Minitab Transformation (M_T) is included in the evaluation, which employs a rounded value of λ compared to the optimal value of λ used in goodness-of-fit tests. Since, the choice of the value for lambda (λ) in a Box-Cox transformation might have a significant impact on the result of process capability or process performance analysis. [9] Shows that the test based on SW statistic is a powerful test of normality for a variety of non-normal distributions, the SW statistic is reliable for small samples and in regression applications, the statistic would yield higher R2. It is asserted in [6] that the test based on SW statistic is the most powerful test for non-normal distributions. According to [18], the current MLE technique could be effectively substituted by using goodness of fits tests in Box-Cox transformation to get data close to normal as possible and achieving desired results in estimating process capability analysis.

III. Lognormal Distribution

The log-normal distribution is a probability distribution of a random variable whose logarithm is normally distributed. When the logarithm of a log-normal distributed variable is taken, it results in a normal distribution. However, when looking at the original data itself, it doesn't follow a normal distribution. It typically exhibits skewness and can have a long tail on one side. It's often used to model phenomena where the logarithm of the variable is normally distributed, such as stock prices, incomes, and certain biological measurements.

$$f(x | \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}} \quad (3)$$

Where,

$x > 0$ is the value of the random variable.

μ is the mean of the natural logarithm of the variable.

σ is the standard deviation of the natural logarithm of the variable.

The mean and variance of the log-normal distribution is given by

$$E(x) = e^{\mu + \frac{\sigma^2}{2}} \tag{4}$$

$$V(x) = (e^{\sigma^2} - 1) \cdot e^{2\mu + \sigma^2} \tag{5}$$

The lognormal distribution was examined at various asymmetric levels, characterized by different mean and standard deviation pairs: (0, 0.25), (0, 0.50), and (0, 1). These parameter sets were grouped to evaluate the impact of low, moderate and high asymmetry in transforming non-normal data to normal data and conducting process capability analysis. Figure 1 illustrates the shape of the density function for each parameter set.

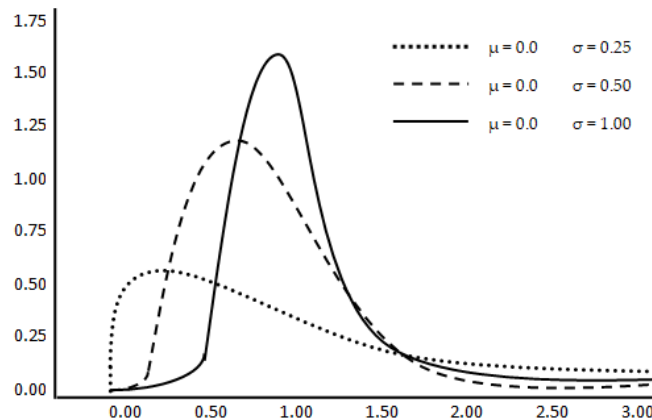


Figure 1: The asymmetric behavior of lognormal distribution used for simulation study

IV. Numerical Illustrations

The log normal distribution is applicable to a wide range of non-normal processes because it is capable of generating a variety of distinct curves based on its parameters. A log-normally distributed random variable only accepts positive real values. It is an easy-to-use model that can be applied to measurements in the exact sciences, engineering, medicine, economics, and other fields (such as energies, concentrations, lengths, prices of financial instruments, and other metrics). For simulation set-up, the data set of the size is taken as 100 and generated using different asymmetric levels of lognormal distribution. The lower and upper specification limits were taken as 0.01 and 10. The defined specification limit in this study of process capability analysis might be appropriate in certain situations where the process parameter being analyzed is bounded by very low values and 0.01 represents a meaningful lower limit for the process output. Here are some scenarios where these specification limits could be reasonable such as chemical concentrations, precision engineering, analytical instruments, environmental monitoring, biomedical applications and so on.

The study evaluates the effectiveness of the method using a combination of box plots, descriptive statistics, measures of errors (Bias, Percentage Bias, Median Absolute Error (MdAE), Root Mean Square Error (RMSE)), and radar charts. Due to space limitations, only error measures and radar plots are included. Bias, MdAE, and RMSE serve as error metrics for transforming non-normal data into normal data using various goodness-of-fit tests in Box Cox transformation. After transformation, the data are utilized to estimate process capability and performance index, aiding in the selection of the most effective approach among different goodness-of-fit tests.

A process is considered to be under six sigma controls if both process capability and performance indices such as Cp and Cpk and Pp and Ppk are greater than or equal to 2 and 1.5, respectively. In the automotive industry, a Cpk value of 1.33 is used as a benchmark for assessing process capability. According to Pearn W.L. and Chen K.S. (2002), a process is considered

inadequate if its process capability index (PCI) is less than 1.00, capable if PCI is between 1.00 and 1.33, satisfactory if PCI is between 1.33 and 1.50, excellent if PCI is between 1.50 and 2.00, and super if PCI is equal to or greater than 2.00 [7]. As outlined by Sibalija TV and Majstorovic VD (2010), the primary objective for quality and industry practitioners is to achieve 6σ limits, with a corresponding defect rate of 3.4 PPM associated with the process [11]. One may refer to [8] and [16] for the details on the concepts of six-sigma tools and process capability analysis for non-normal data, respectively. Table 2 displays the process fallout in defective parts per million products alongside the proportion of good items and PPM values for different sigma levels.

Table 2: Process fallout in defective parts per million with respect to different sigma levels

Sigma Level	Percentage	PPM Values
6	99.9997%	3.4
5	99.98%	233
4	99.4%	6,210
3	93.3%	66,807
2	69.1%	308,537
1	30.9%	691,462

I. Low Asymmetric Distribution

The simulation study focuses on utilizing a low asymmetric lognormal distribution with a skewness of 0.36 and 0.63, where the mean and standard deviation are 0 and 0.25, respectively. To assess the effectiveness of various methods in transforming non-normal data into a normal distribution, two sets of data are analyzed. One with a Skewness (Sk) of 0.36 and another with a skewness of 0.63. For the dataset with $\ln(0, 0.25)(1)$, methods like M_T, PT, LT, and JB transforms data more like a normal distribution with fewer errors. Likewise, for the dataset with $\ln(0, 0.25)(2)$, methods such as SF, JB, SW, AC, and MLE are transforms data getting closer to a normal with fewer errors. For further details, refer to Table 3 and Figure 2. Subsequently, the transformed data from different goodness-of-fit tests are used to estimate process capability/performance. This analysis helps identify the most effective method for handling non-normal, low asymmetric distributions.

Table 3: Various measures of error values for low asymmetric data after the data transformation

Goodness of fit tests	Low Asymmetry (Sk=0.36)			Low Asymmetry (Sk=0.63)		
	Lognormal distribution ($\mu=0, \sigma=0.25$)			Lognormal distribution ($\mu=0, \sigma=0.25$)		
	Bias	MdAE	RMSE	Bias	MdAE	RMSE
SW	1.0156	1.0068	1.0158	1.0231	1.0104	1.0236
AD	1.0156	1.0068	1.0158	1.0373	1.0168	1.0388
CVM	1.0180	1.0078	1.0182	1.0423	1.0191	1.0442
PT	0.9916	0.9962	0.9916	1.0484	1.0218	1.0510
SF	1.0159	1.0069	1.0161	1.0223	1.0100	1.0228
LT	1.0147	1.0065	1.0149	1.0355	1.0160	1.0368
JB	1.0152	1.0067	1.0153	1.0228	1.0102	1.0233
AC	1.0161	1.0071	1.0163	1.0245	1.0110	1.0251
MLE	1.0161	1.0070	1.0163	1.0246	1.0111	1.0252
M_T	-0.0008	0.0735	0.1049	1.0262	1.0118	1.0269

All the methods of data transformation used in this study results within the standard of six sigma but the only few methods produces better estimates of Pp/Cp and Ppk/Cpk, such methods are CVM, AC, MLE, SF, SW and AD for data set $\ln(0, 0.25)(1)$ and SF, JB, SW, AC and MLE for data set $\ln(0, 0.25)(2)$. Though only the tests SF, SW, AC, and MLE were considered appropriate

procedures to deal non-normal low asymmetric distributions in order to obtain desirable results with less errors, better estimates, and PPM values within the six sigma limits. For more information see the table 2, 4 and 5.

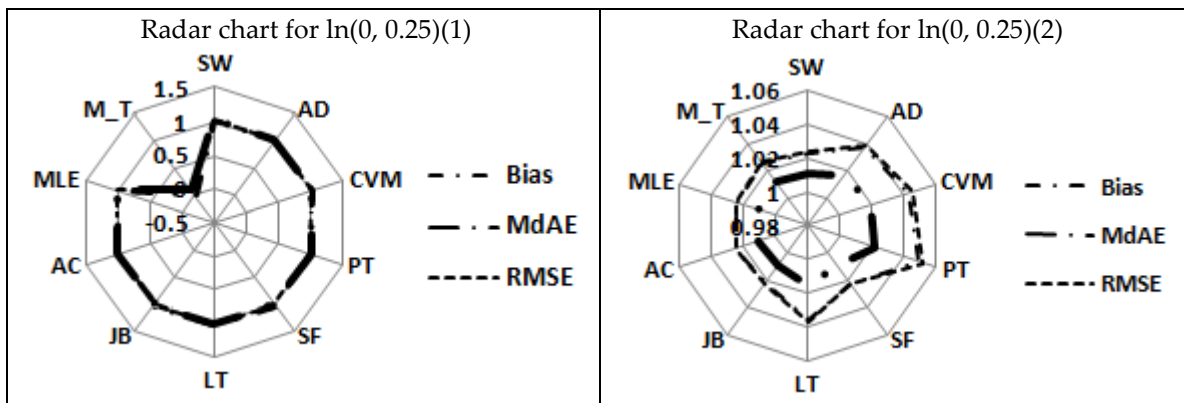


Figure 2: Radar chart for various measures of errors after the normalization of low asymmetric distribution

Table 4: Estimates of process capability and process performance indices for ln(0, 0.25)(1) data after normalization via goodness of fit tests

Goodness of fit tests	λ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
ln(0, 0.25)(1)	-	0.01	10	7.62	1.49	3.87	7.893	1.544	1.81
SW	0.31	-2.45	3.36	4.38	3.65	0.00	4.54	3.79	0.00
AD	0.31	-2.45	3.36	4.38	3.65	0.00	4.54	3.79	0.00
CVM	0.21	-2.95	2.96	4.44	4.39	0.00	4.6	4.55	0.00
PT	1.38	-0.72	16.66	13.2	1.09	5.64	13.7	1.12	373
SF	0.3	-2.50	3.32	4.39	3.73	0.00	4.55	3.86	0.00
LT	0.35	-2.29	3.54	4.40	3.42	0.00	4.57	3.54	0.00
JB	0.33	-2.37	3.45	4.39	3.54	0.00	4.55	3.67	0.00
AC	0.29	-2.54	3.27	4.38	3.79	0.00	4.54	3.92	0.00
MLE	0.29	-2.54	3.28	4.39	3.79	0.00	4.55	3.92	0.00
M_T	0.5	0.100	3.16	4.65	2.69	0.00	4.82	2.79	0.00

Table 5: Estimates of process capability and process performance indices for ln(0, 0.25)(2) data after normalization via goodness of fit tests

Goodness of fit tests	λ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
ln(0, 0.25)(2)	-	0.01	10	7.041	1.436	8.25	7.020	1.432	8.73
SW	0.12	-3.54	2.65	4.55	3.89	0.00	4.51	3.86	0.00
AD	-0.43	-14.5	1.46	11.68	2.15	0.00	11.55	2.12	0.00
CVM	-0.62	-26.4	1.23	20.03	1.80	0.00	19.78	1.78	0.00
PT	-0.85	-57.7	1.01	41.93	1.47	5.27	41.37	1.45	6.90
SF	0.15	-3.33	2.75	4.47	4.03	0.00	4.43	4	0.00
LT	-0.36	-11.8	1.57	9.80	2.31	0.00	9.69	2.28	0.00
JB	0.13	-3.47	2.68	4.52	3.92	0.00	4.48	3.9	0.00
AC	0.065	-3.98	2.48	4.75	3.64	0.00	4.71	3.61	0.00
MLE	0.06	-4.02	2.47	4.78	3.63	6.10	4.73	3.6	0.00
M_T	0.0	-4.61	2.31	5.086	3.387	0.00	5.039	3.356	0.00

II. Moderate Asymmetric Distribution

The lognormal distribution, characterized by parameters $\mu = 0$ and $\sigma = 0.50$, offers a means to generate moderately asymmetric data with respective skewness values of 0.96 and 1.32. Through a simulation study, it is confirmed that for data set $\ln(0, 0.5)(1)$ the LT, AD, SW, SF, CVM and JB and for data set $\ln(0, 0.5)(2)$, the M_T, LT, AC, MLE, SF and SW methods of goodness of fit tests effectively transform the non-normal data into normal distributions with minimal errors. Consequently, the transformed datasets are subjected to further examination to evaluate their efficiency in estimating process capability/performance for moderately asymmetric distributions. See the table 6 and figure 3.

Table 6: Various measures of error values for moderate asymmetric data after the data transformation

Goodness of fit tests	Moderate Asymmetry (Sk=0.96)			Moderate Asymmetry (Sk=1.32)		
	Lognormal distribution ($\mu=0, \sigma=0.5$)			Lognormal distribution ($\mu=0, \sigma=0.5$)		
	Bias	MdAE	RMSE	Bias	MdAE	RMSE
SW	1.1050	1.0392	1.1141	1.1634	1.0599	1.1856
AD	1.1036	1.0386	1.1124	1.1703	1.0627	1.1938
CVM	1.1065	1.0397	1.1158	1.1648	1.0605	1.1873
PT	1.1079	1.0403	1.1174	1.1648	1.0605	1.1873
SF	1.1050	1.0392	1.1141	1.1621	1.0594	1.1840
LT	1.1007	1.0375	1.1091	1.1538	1.0561	1.1741
JB	1.1065	1.0397	1.1158	1.1675	1.0616	1.1905
AC	1.1086	1.0406	1.1182	1.1608	1.0589	1.1825
MLE	1.1079	1.0403	1.1174	1.1607	1.0588	1.1823
M_T	0.1071	0.1667	0.3422	1.1483	1.0539	1.1675

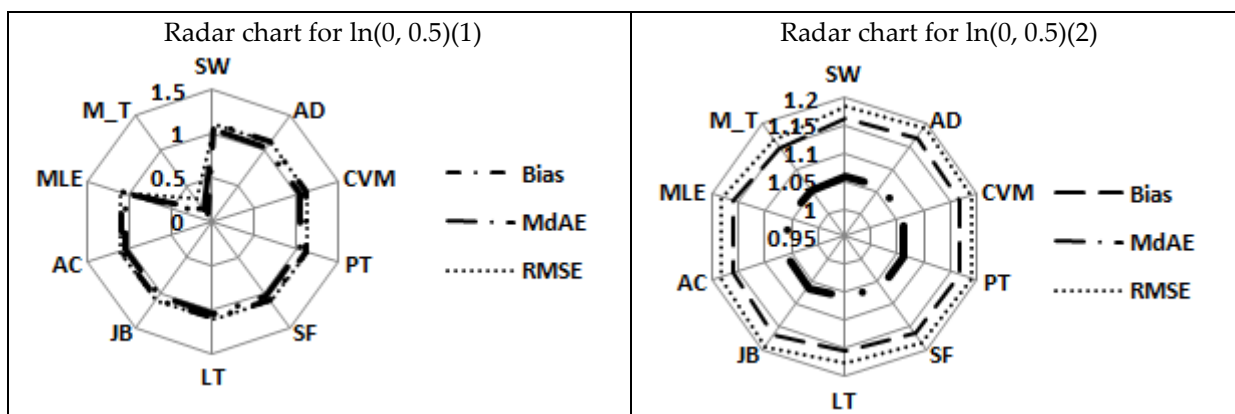


Figure 3: Radar chart for various measures of errors after the normalization of moderate asymmetric distribution

In the simulation study, transformed data yields improved estimates of Pp/Cp and Ppk/Cpk using methods such as AC, PT, MLE, JB, CVM, SW, and SF for dataset $\ln(0, 0.5)(1)$, and AC, MLE, SF, and SW for dataset $\ln(0, 0.5)(2)$. Furthermore, the PPM values indicate that for dataset $\ln(0, 0.5)(1)$, the results fall within the standard six sigma limits, while for dataset $\ln(0, 0.5)(2)$, PPM values range between 5σ and 6σ limits (with recorded values ranging from 17 to 128, against benchmark values of 233 for 5σ and 3.4 for 6σ , as detailed in Table 2). This close alignment with the six sigma standard is promising. Upon considering various measures of errors, process capability/performance indices, and associated PPM values, it becomes evident that the AC, MLE, SW, and SF approaches outshine other methods, as depicted in Tables 2, 7, and 8.

Table 7: Estimates of process capability and process performance indices for $\ln(0, 0.5)(1)$ data after normalization via goodness of fit tests

Goodness of fit tests	λ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
$\ln(0, 0.5)(1)$	-	0.01	10	2.85	0.64	22623	2.81	0.64	28271
SW	0.29	-2.54	3.28	1.80	1.59	0.88	1.80	1.59	0.95
AD	0.3	-2.50	3.32	1.80	1.57	1.24	1.80	1.56	1.34
CVM	0.28	-2.59	3.23	1.80	1.62	0.56	1.80	1.62	0.60
PT	0.27	-2.64	3.19	1.80	1.65	0.36	1.80	1.65	0.38
SF	0.29	-2.54	3.28	1.80	1.59	0.88	1.80	1.59	0.95
LT	0.32	-2.41	3.40	1.80	1.52	2.71	1.79	1.51	2.93
JB	0.28	-2.59	3.23	1.80	1.62	0.56	1.80	1.62	0.60
AC	0.27	-2.66	3.17	1.80	1.66	0.300	1.80	1.66	0.32
MLE	0.27	-2.64	3.19	1.80	1.65	0.36	1.80	1.65	0.38
M_T	0.5	0.10	3.16	1.88	1.15	285	1.87	1.14	310

Table 8: Estimates of process capability and process performance indices for $\ln(0, 0.5)(2)$ data after normalization via goodness of fit tests

Goodness of fit tests	λ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
$\ln(0, 0.5)(2)$	-	0.01	10	2.59	0.59	38718	2.60	0.59	37847
SW	-0.11	-6.00	2.03	2.51	1.28	62.67	2.54	1.30	49.46
AD	-0.16	-6.81	1.93	2.72	1.22	127.9	2.77	1.24	102.4
CVM	-0.12	-6.15	2.01	2.55	1.27	72.39	2.59	1.29	57.30
PT	-0.12	-6.15	2.01	2.55	1.27	72.39	2.59	1.29	57.30
SF	-0.1	-5.85	2.06	2.47	1.30	50.39	2.51	1.32	39.42
LT	-0.04	-5.06	2.20	2.27	1.38	17.58	2.30	1.40	13.52
JB	-0.14	-6.47	1.97	2.63	1.24	93.68	2.67	1.26	76.73
AC	-0.091	-5.72	2.08	2.44	1.31	43.33	2.47	1.33	33.92
MLE	-0.09	-5.71	2.08	2.43	1.31	43.33	2.47	1.33	33.96
M_T	0.0	-4.61	2.30	2.16	1.44	8.00	2.19	1.46	6.00

III. High Asymmetric Distribution

The lognormal distribution, with parameters $\mu = 0$ and $\sigma = 1$, can lead to highly skewed distributions. Through numerical examples, it's clear that methods like SF, AC, SW, MLE, and JB effectively transform non-normal data into normal distributions with minimal errors for dataset $\ln(0, 1)(1)$, and methods M_T, SF, AC, MLE, and SW do the same for dataset $\ln(0, 1)(2)$. Refer to Table 9 and Figure 4 for more details. In terms of producing accurate estimates of Pp/Cp and Ppk/Cpk, methods like SF, AC, JB, SW, and MLE perform well for dataset $\ln(0, 1)(1)$, and SW, SF, AC, and MLE perform well for dataset $\ln(0, 1)(2)$. Considering various error measures, estimates of process capability/performance indices, and corresponding PPM values, it's evident that methods like AC, MLE, SW, and SF perform better than others. The estimates and their PPM values in this study fall within the standard 4σ and 5σ limits. For dataset $\ln(0, 1)(1)$, values range from a minimum of 424 to a maximum of 495, while for dataset $\ln(0, 1)(2)$, values range from a minimum of 5664 to a maximum of 6230. The standard PPM value for 4σ limits is 6210, and for 5σ limits is

233. These values closely approach the standard of 5σ limits for highly asymmetric distributions. Refer to Tables 2, 10, and 11 for further details.

Table 9: Various measures of error values for high symmetric data after the data transformation

Goodness of fit tests	High Asymmetry (Sk=1.81)			High Asymmetry (Sk=2.45)		
	Lognormal distribution ($\mu=0, \sigma=1$)			Lognormal distribution ($\mu=0, \sigma=1$)		
	Bias	MdAE	RMSE	Bias	MdAE	RMSE
SW	1.4662	1.1958	1.6223	1.6526	1.3108	1.9611
AD	1.5072	1.2156	1.6798	1.6903	1.3319	2.0109
CVM	1.5335	1.2301	1.7174	1.7012	1.3379	2.0253
PT	1.5715	1.2509	1.7729	1.7514	1.3655	2.0912
SF	1.4612	1.1935	1.6153	1.6472	1.3072	1.9540
LT	1.5124	1.2185	1.6872	1.6848	1.3288	2.0038
JB	1.4662	1.1958	1.6223	1.6686	1.3198	1.9824
AC	1.4657	1.1956	1.6216	1.6506	1.3096	1.9585
MLE	1.4662	1.1958	1.6223	1.6526	1.3108	1.9611
M_T	1.5335	1.2301	1.7174	1.6102	1.2810	1.9043

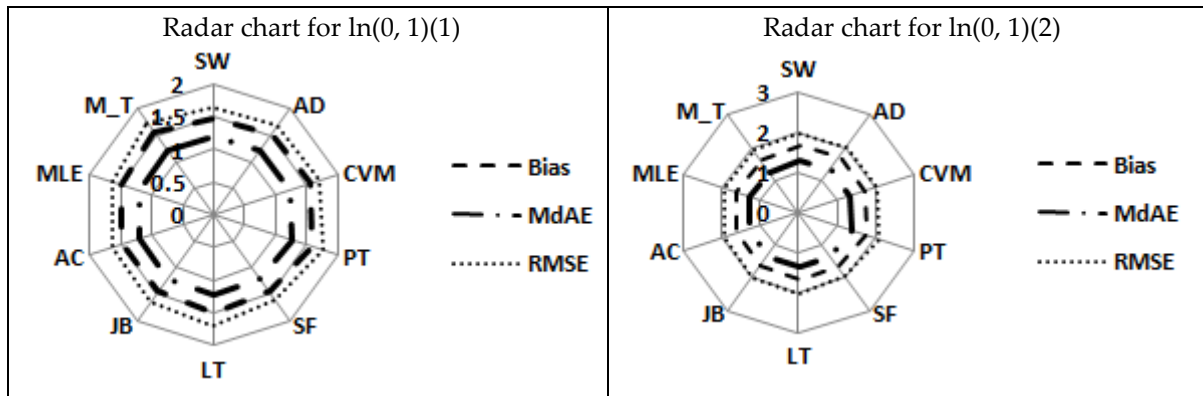


Figure 4: Radar chart for various measures of errors after the normalization of high asymmetric distribution

Table 10: Estimates of process capability and process performance indices for $\ln(0, 1)(1)$ data after normalization via goodness of fit tests

Goodness of fit tests	λ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
				$\ln(0, 1)(1)$	-	0.01	10	1.143	0.324
SW	0.13	-3.47	2.68	1.25	1.10	492.37	1.28	1.13	355.35
AD	0.05	-4.11	2.44	1.31	1.00	1325.51	1.35	1.03	1016.34
CVM	0	-	-	-	-	-	-	-	-
PT	-0.07	-5.43	2.13	1.47	0.86	4947.72	1.50	0.88	4127.10
SF	0.14	-3.39	2.72	1.24	1.12	424.31	1.27	1.15	302.77
LT	0.04	-4.21	2.41	1.33	0.99	1548.60	1.36	1.01	1206.37
JB	0.13	-3.47	2.68	1.25	1.10	488.93	1.28	1.13	352.46
AC	0.13	-3.46	2.69	1.25	1.10	475.10	1.28	1.13	342.26
MLE	0.13	-3.47	2.68	1.25	1.10	494.90	1.28	1.13	357.47
Minitab	0.0	-4.61	2.30	1.096	0.766	10808	1.12	0.780	9660

Table 11: Estimates of process capability and process performance indices for $\ln(0, 1)(2)$ data after normalization via goodness of fit tests

Goodness of fit tests	λ Value	LSL	USL	PCI (Within Capability)			PPI (Overall Capability)		
				Cp	Cpk	PPM	Pp	Ppk	PPM
				$\ln(0, 1)(2)$	-	0.01	10	0.970	0.284
SW	-0.08	-5.57	2.10	1.47	0.83	6211.92	1.44	0.82	6896.12
AD	-0.15	-6.64	1.95	1.62	0.78	9864.76	1.60	0.77	10845.7
CVM	-0.17	-6.99	1.91	1.66	0.76	11394.4	1.64	0.75	12448.9
PT	-0.26	-8.89	1.73	1.94	0.68	19977.2	1.90	0.67	21670.8
SF	-0.07	-5.43	2.13	1.45	0.84	5663.84	1.43	0.83	6298.88
LT	-0.14	-6.47	1.97	1.59	0.78	9466.45	1.57	0.77	10412.8
JB	-0.11	-6.00	2.03	1.53	0.81	7697.25	1.51	0.79	8552.19
AC	-0.08	-5.52	2.11	1.46	0.84	6036.33	1.44	0.82	6704.05
MLE	-0.08	-5.57	2.10	1.47	0.83	6230.13	1.44	0.82	6918.05
Minitab	0.0	-4.61	2.30	1.189	0.839	5916	1.13	0.80	8528

V. Result and Discussion

This study investigates two main areas, focusing on data transformation and the estimation of process capability analysis. The effectiveness of different goodness-of-fit tests is evaluated based on various error measures, estimates of process capability/performance, and PPM values that closely adhere to the six sigma standard. It also explores the impact of optimal and rounded values of lambda when transforming non-normal data into normal data for estimating process capability analysis. To achieve desired outcomes, it is essential that the transformed data closely resemble a normal distribution with minimal errors. Additionally, consistency in producing standard estimates and lower PPM values from the extended transformed data serves as evidence that the methodology employed in this study yields the desired results.

In each of the three distinct asymmetric scenarios examined a range of goodness-of-fit tests, notably SW, SF, AC, and MLE, exhibit proficiency in converting non-normal datasets into normal distributions with minimal error values, the estimated values of Pp/Cp and Ppk/Cpk meet or exceed benchmark standards. The corresponding PPM values fall within or near the 6σ limits, only when low and moderate asymmetric distributions. For highly asymmetric distributions, the transformed dataset demonstrates reduced errors, yet the estimates of Pp/Cp and Ppk/Cpk deviate from standard results, and the corresponding PPM values do not align with the 6σ benchmark. It is noteworthy that across all asymmetric scenarios, error values are minimized for the JB, M_T, LT, and PT methods of goodness-of-fit tests. However, the associated estimates and PPM values do not correspond with the desired outcomes.

The primary objective of this paper is to obtain improved estimates of process capability or process performance indices using Box-Cox Transformation (BCT) through goodness-of-fit tests. When applying BCT to convert non-normal data into a normal distribution, selecting the transformation parameter λ becomes crucial. BCT provides an optimal and rounded value of lambda for data transformation. In this study, goodness-of-fit tests utilize the optimal value of λ , whereas M_T employs the rounded value of λ . Based on the numerical illustrations, it is observed that the PPM values estimated in this study using the optimal value of λ are higher than those estimated using other methods. Notably, for moderately asymmetric distributions, employing the M_T method results in higher PPM values compared to methods utilizing the optimal value of lambda for $\ln(0, 0.5)(1)$. However, for $\ln(0, 0.5)(2)$, PPM values of minimum compared to goodness of fit tests but results in lesser values of Pp/Cp and Ppk/Cpk compared to benchmark standards.

Similarly, for highly asymmetric distributions, the M_T method produces higher PPM values compared to methods utilizing the optimal value of λ for $\ln(0, 1)(1)$. Nonetheless, for $\ln(0, 1)(2)$ PPM values are minimum compared to goodness of fit tests but results in lesser values of Pp/Cp and Ppk/Cpk compared to benchmark standards.

Table 12: Efficiency comparison over different goodness of fit tests in data transformation and estimation of process capability and process performance indices for lognormal distribution

Goodness of fit tests	Efficiency in data transformation						Efficiency in estimation of PCI					
	Low		Moderate		High		Low		Moderate		High	
	Asymmetric		Asymmetric		Asymmetric		Asymmetric		Asymmetric		Asymmetric	
Skewness	0.36	0.63	0.96	1.32	1.81	2.45	0.36	0.63	0.96	1.32	1.81	2.45
SW	✓	✓	✓	✓	✓	✓	✓*	✓*	✓*	✓*	✓*	✓*
AD	✓		✓				✓					
CVM			✓				✓\$		✓			
PT	✓								✓\$			
SF		✓	✓	✓	✓	✓	✓	✓*	✓*	✓*	✓*	✓*
LT	✓		✓	✓						✓		
JB	✓	✓	✓		✓	✓		✓	✓		✓*	✓*
AC		✓		✓	✓	✓	✓\$	✓*	✓\$	✓*	✓*	✓*
MLE		✓		✓	✓	✓	✓\$	✓*	✓\$	✓*	✓*	✓*
M_T	✓		✓	✓		✓	✓	✓\$				
DME												

DME – Direct Minitab Estimation | ✓ - Less Error and/or Better Estimate | ✓* - Better Estimate with less error and lesser PPM values | ✓\$ - Better Estimates with less PPM and higher error values.

This clearly indicates that using rounded values of λ is somewhat less efficient in transforming non-normal data into normal data, resulting in corresponding estimates of process capability/performance that do not meet the benchmark standards compared to the results obtained from methods utilizing the optimal value of λ . This discrepancy arises because the use of rounded value of λ does not accurately reflect the transforming pattern needed to achieve a close approximation to a normal distribution, as opposed to the optimal value of λ . Therefore, it is evident that opting for the optimal value of λ to attain improved estimates in process capability analysis would be the superior choice. One may refer table 4, 5, 7, 8, 10 and 11. A table of data is formed for the better understanding of the efficiency of different normality tests under various asymmetric behaviors of lognormal distribution based on the numerical example, result and discussion. See table 12 for more information.

V. Conclusion

The core objective of this research work is to analyze the impact of the transformation parameter lambda on enhancing the process capability assessment for lognormal distribution. A test that satisfies all the requirements including data closely adhering to a normal distribution with less error, enhanced estimates of process capability/performance and reduced PPM values, to achieve the desired result is looked at utilizing low, moderate, and high asymmetric log normal distribution. Accordingly, based on the findings, result and discussion, SW, SF, AC, and MLE methodologies of goodness of fit tests have more intense power to estimate process capability/performance indices with smaller PPM values and also have higher accuracy in data

transformation. The SW test performs better than the other approaches in every way to produce enhanced estimates of process capability analysis. However, other test methods, like the SF, AC, and MLE methods, position themselves subsequent places for dealing with non-normal quality characteristics, particularly lognormal distribution and delivering remarkably good results.

M_T (Minitab Transformation) utilizes the rounded value of λ instead of the optimal value to ascertain the transforming parameter. Optimal λ is typically required for superior results as it accurately reflects the transforming pattern, unlike a rounded value. This approach ensures that all values are brought as close to normal as possible. Based on the numerical illustrations, this study produces large amount of error values while using rounded value of λ , except in low asymmetry situations, resulting in the transformed data not close enough to normal distribution and less efficient estimates when compared to methods that are utilizing an ideal value of λ during data transformation and estimation. Furthermore, it is concluded and recommended that when dealing with non-normal data specifically lognormal distribution, utilizing an optimum value of λ is typically required for better results and Shapiro Wilk's (SW) test is one such method among the different goodness of fit tests to transform non normal data into normal and estimating process capability/ performance in order to get enhanced results.

References

- [1] Asar O, Ilk O and Dag O (2013). Estimating Box Cox power transformation parameter via goodness of fit tests. *Communication in statistics – simulation and computation*, 46(1), pp 91 – 105.
- [2] Box.G.E.P., Cox. D.R., (1964). An analysis of Transformations. *Journal of the royal statistical society. Series B(Methodological)*, 26(2), 211-252.
- [3] Dag O, Asar O and Ilk O (2013). A Methodology to Implement Box-Cox Transformation When No Covariate is Available. *Communication in statistics – simulation and computation*, 43(7), pp 1740 – 1759.
- [4] Gunter, B.H. (1989). The use and abuse of Cpk ,1-4, *Quality Progress*, 22(1), 72-73(3), 108-109(5), 79-80(7).
- [5] Kane V E (1986), Process capability indices. *Journal quality technology*, 18(1), pp 41 – 52.
- [6] Oztuna D, Elhan A and Tuccar E (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions, *Turkish Journal of Medical Sciences*, 36(3), pp 171-176.
- [7] Pearn, W. L., & Chen, K. S. (2002). One-sided capability indices CPU and CPL: decision making with sample information. *International Journal of Quality & Reliability Management*, 19(3), 221–245.
- [8] Pyzdek T, *The Six Sigma Handbook*, ISBN 0-07-141015-5, McGraw-Hill Companies, Inc, New York 2003.
- [9] Rahman M (1999). Estimating the box cox transformation via Shapiro wilk W statistic. *Communication in statistics – simulation and computation*, 28(1), pp 223 – 241.
- [10] Rahman M and Pearson LM (2008). Anderson-darling statistic in estimating the box-cox transformation parameter. *Journal of applied probability & statistics*, 3(1), pp 23 – 35.
- [11] Sennaroglu B and Senvar O (2015). Performance comparison of Box-Cox transformation and weighted variance methods with weibull distribution. *Journal of Aeronautics and Space Technologies*, 8(2), pp 49-55.
- [12] Sibalija TV and Majstorovic VD (2010). Process performance analysis for non normal data distribution. *International Journal “Total Quality Management & Excellence”*, 38(3), pp 1-4.
- [13] Tang LC and Than SU EE, (1999). Computing process capability indices for non normal data: A review and comparative study. *Quality and Reliability Engineering International*, 15, pp 339 – 353.
- [14] Thadewald T and Buning H (2007), Jarque-Bera Test and its Competitors for Testing

Normality - A Power Comparison, *Journal of Applied Statistics*, 34(1), pp 87-105.

[15] Yang Y and Zhu H (2018). A study on process capability analysis based on Box-Cox transformation, *IEEE*, pp 240 – 243.

[16] Yap B W and Sim C H (2011), Comparisons of various types of normality tests. *Journal statistical computation and simulation*, 18(12), pp 2141 – 2155.

[17] Yoap T (2006). Process capability analysis for non normal data with Minitab, *Six Sigma: Advances Tools for Black Belts and Master Black Belts*, pp 131 – 149.

[18] Krishnan J and Vijayaraghavan R (2024), Process capability analysis for non normal data Based on box-cox transformation through Tests of goodness of fit, *Reliability: Theory & Applications*. March 1(77), 297-309.

[19] Swamy, D. R., Nagesh, P., and Wooluru, Y. (2016). Process Capability Indices for Non-normal Distribution – A Review, *Proceedings of the International Conference on Operations Research and Management*, January 21 – 22, Mysuru, India.

INFERENCE ON THE INVERSE POWER BURR-HATKE DISTRIBUTION UNDER TYPE II CENSORING

¹Pavitra Kumari, ²Vinay Kumar

•

¹ Department of Statistics, Central University of Haryana, India

² Department of Mathematics and Statistics, Chaudhary Charan Singh Haryana Agricultural
University, India

[¹pavitrastats@gmail.com](mailto:pavitrastats@gmail.com)

[²vinay.stat@gmail.com](mailto:vinay.stat@gmail.com)

Abstract

There are many real-life situations, where data require probability distribution function which have decreasing or upside-down bathtub (UBT) shaped failure rate function. The inverse power burr hatke distribution consists both decreasing and UBT shaped failure rate functions. Here, we address the different estimation methods of the parameter and reliability characteristics of the inverse Pareto distribution from both classical and Bayesian approaches. We consider classical estimation procedures to estimate the unknown parameter of inverse power burr-hatke distribution, such as maximum likelihood. Also, we consider Bayesian estimation using squared error loss function based joint priors. The Monte Carlo simulations are performed to compare the performances of the obtained estimators in mean square error sense. Finally, the flexibility of the proposed distribution is illustrated empirically using one real-life datasets. The analyzed data shows that the introduced distribution provides a superior fit than some important competing distributions such as the Weibull, inverse Pareto and Burr-Hatke distributions.

Keywords: Burr-Hatke Distribution, Inverse Power Burr- Hatke Distribution, Type II censoring, Bayesian estimation, Lindley's Approximation technique.

I. Introduction

Statistical distributions can be used to model many real-life scenarios, such as reliability, actuarial science, survival analysis and lifetime data. Different lifetime distributions have been introduced in the statistical literature to provide greater flexibility in modelling data in these applied sciences. One of the important features of generalized distributions is their capability for providing superior fit for various life-time data encountered in the applied fields. Hence, the statisticians have been interested in constructing new families of distributions to model such data. Recently, several new distributions and regression models to provide inferences on these distributions have been developed for modeling health and biomedical data, among other fields. Some distributions and classes of distributions developed include exponentiated Burr XII Poisson distribution by da Silva et al. [1], Weibull Burr XII (WBXII) distribution by Afify et al. [2], odd log logistic Topp–Leone G family of distributions by Alizadeh et al. [3], Burr-Hatke exponential (BHE) distribution by Abouelmagd [4] and Yadav et al. [5], odd generalized gamma-G family of distributions by Nasir et al. [6], Chen-G family of distributions by Anzagra et al. [7], inverse-power Burr-Hatke distribution by Afify et al. [8], harmonic mixture Weibull-G family of distributions by Zamanah et al. [9], harmonic mixture G family of distributions by Kharazmi et al. [10] and Alshenawy R. [11] studied Progressive Type-II Censoring Schemes of Extended Odd Weibull Exponential Distribution with Applications in

Medicine and Engineering. Ahmed et. al. [12] studied Bayesian and Classical Inference under Type-II Censored Samples of the Extended Inverse Gompertz Distribution with Engineering Applications. Hassan [13] studied Statistical Inference of Chen distribution Based on Two Progressive Type-II Censoring Schemes. Burr Hatke model provides only a decreasing hazard rate (HR) shape; hence, its use will be limited to modelling the data that exhibits only increasing failure rate. IPBH model can accommodate right-skewed shape, symmetrical shape, reversed J shape and left-skewed shape densities. Its hazard rate (HR) can be an increasing shape, a unimodal shape, or a decreasing shape. IPBH distribution provides more accuracy and flexibility in fitting engineering and medicine data. The IPBH distribution was constructed using the inverse-power (IP) transformation. The aim of this article is to develop the classical and Bayesian estimation procedures for the parameters of the IPBH. The rest of the article is organized as follows: IPBH is discussed in Section 2. Also, mathematical formulation is given for type II censoring with failure and censoring time distributions in this section. Section 3 deals with the maximum likelihood estimation and asymptotic confidence intervals of the parameters. Section 4 describes asymptotic confidence interval. Sections 5 describe the formulation of Bayes estimation procedure using Markov chain Monte Carlo (MCMC) methods under SELF loss function using gamma informative priors. Section 6 deals with a Monte Carlo simulation study to explore the properties of various estimates developed in this article. Real life dataset is analyzed for illustration purposes in Section 7. Finally, conclusive remarks are given in section 8. Also, it is essential to mention that the statistical software R 3.5.2, [R Core Team (2018)] is used for computation purposes throughout the article.

II. The Model

If a random variable X follows IPBH with parameter (λ, θ) the cdf is given by:

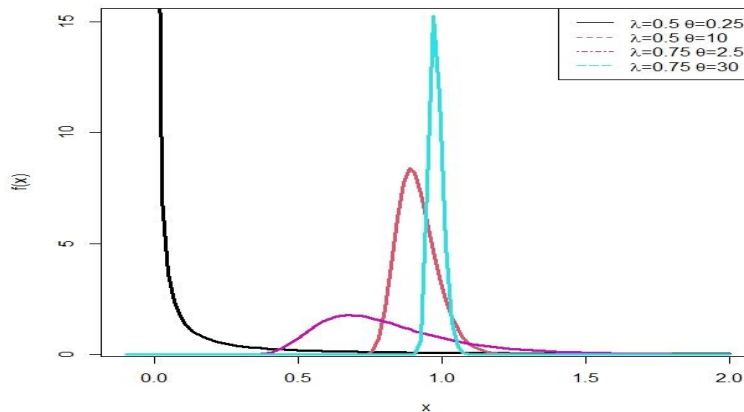
$$F(x; \lambda, \theta) = \frac{\exp(-\lambda x^{-\theta})}{x^{-\theta} + 1}, \quad \lambda, \theta > 0 \quad (2.1)$$

Therefore, the corresponding probability density function is given by

$$f(x; \theta, \lambda) = \frac{\theta \exp(-\lambda x^{-\theta}) [\lambda + (1 + \lambda)] x^{-\theta}}{x(x^\theta + 1)^2}, \quad \lambda, \theta > 0 \quad (2.2)$$

Where θ and λ are shape parameters, respectively.

Figure 1. Possible density shapes of the IPBH distribution for several values of λ and θ .



The survival function (SF) and HR function of the IPBH distribution take the following forms, respectively:

$$S(x; \lambda, \theta) = 1 - \frac{x^\theta \exp(-\lambda x^\theta)}{x^{-\theta} + 1} \quad (2.3)$$

$$h(x; \lambda, \theta) = \frac{\theta[\lambda + (\lambda + 1)x^{-\theta}]}{x(x^\theta + 1)[(x^\theta + 1) \exp(-\lambda x^{-\theta}) - x^\theta]} \quad (2.4)$$

Figure 2. Possible failure rate shape of the IPBH distribution for values of $\lambda = 0.75$ and $\theta = 3$

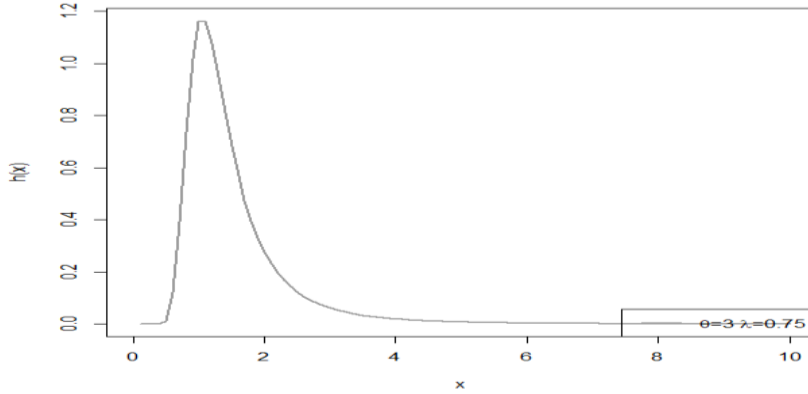


Figure 3. Possible failure rate shape of the IPBH distribution for values of $\lambda = 0.5$ and $\theta = 0.25$

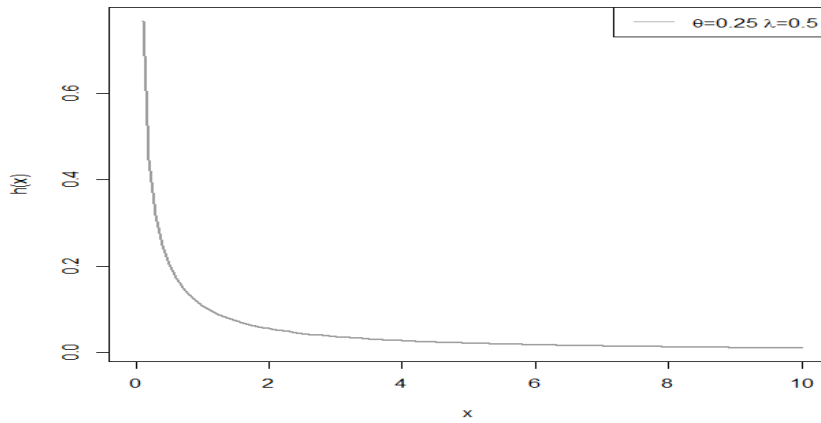
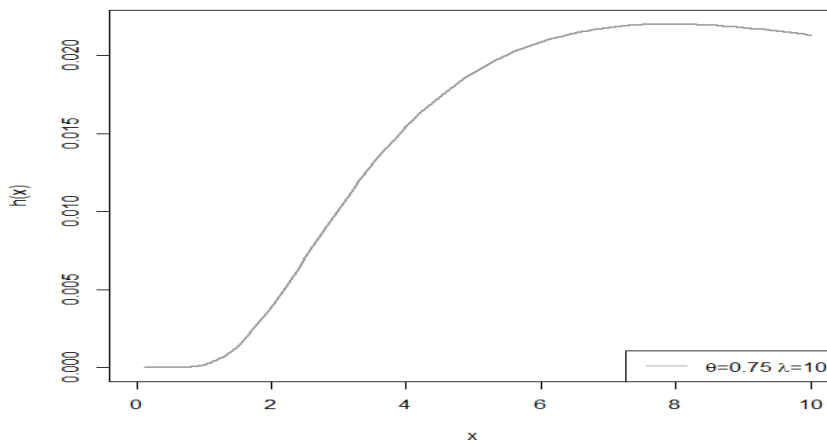


Figure 4. Possible failure rate shape of the IPBH distribution for values of $\lambda = 10$ and $\theta = 0.75$



III. Maximum Likelihood Estimation

In the literature, Several censoring schemes have been discussed. Even though, Type-I and Type-II censoring schemes are most popular censoring. Consider a life test where n independent units taken from a IPBH distribution are placed under observation and failure time of each unit is recorded. Suppose that the test is terminated when r th, ($1 \leq r \leq n$), r is prefixed unit fails. These observed failure times, say (x_1, x_2, \dots, x_r) is a Type-II censored sample of size r . In this censoring scheme $n-r$ units remain unobserved and survive beyond the time of termination. In Type-II censoring the time of termination is a random variable and the likelihood function based on (x_1, x_2, \dots, x_r) is given by Cohen [14].

$$L(\lambda, \theta | \underline{x}) = \frac{n}{(n-r)} \prod_{i=0}^r f(x_i) [1 - F(x_{(r)})]^{n-r} \quad (2.5)$$

Assume that n independent observed values taken of IPBH distribution as presented in (2) are put on a test. Using the Type-II censoring, we obtained the ordered r failures. If the ordered r failures are then (x_1, x_2, \dots, x_r) the likelihood function of (λ, θ) under Type-II censored data drawn of an IPBD distribution, is obtained as follows:

$$L(\lambda, \theta | \underline{x}) = \frac{n}{(n-r)} \prod_{i=0}^r f(x_i) [1 - F(x_{(r)})]^{n-r}$$

$$L(\lambda, \theta | \underline{x}) = r \log(\theta) - \lambda \sum_{i=1}^r x_i^{-\theta} + \sum_{i=1}^r \log[\lambda + (\lambda + 1)x_i^\theta] - 2 \sum_{i=1}^r \log(x_i^\theta + 1) - \sum_{i=1}^r \log(x_i) - \eta$$

MLEs of λ and θ is a solution of equation (2.5) accomplished by addressing the first partial derivatives of the total log-likelihood to be zero. So, we consider the equation as follows,

$$\frac{d \log L}{d \lambda} = \sum_{i=1}^r \frac{x_i^\theta + 1}{\lambda + (\lambda + 1)x_i^\theta} + \sum_{i=1}^r x_i^\theta + \frac{(n-r)}{(x_{(r)}^\theta + 1)e^{\lambda x_{(r)}^{-\theta}} - x_{(r)}^\theta}$$

$$\frac{d \log L}{d \theta} = -\lambda \sum_{i=1}^r x_i^{-\theta} \log(x_i) + \sum_{i=1}^r \frac{(\lambda + 1)x_i^\theta \log(x_i)}{\lambda + (\lambda + 1)x_i^\theta} - 2 \sum_{i=1}^r \frac{x_i^\theta \log(x_i)}{x_i^\theta + 1} + \eta_1(x)$$

The closed form solutions to the nonlinear Equations are difficult to reach and a numerical method must be applied to solve these simultaneous equation for obtaining the MLE of λ and θ .

IV. Asymptotic Confidence Intervals

The maximum likelihood estimators of the unknown parameters are not in closed form, it is not easy to drive the exact distributions of the MLEs. Thus, we use the asymptotic distribution of MLEs for the constructions of asymptotic confidence intervals of the parameters based on observed Fisher information matrix. Let $\hat{\alpha} = (\hat{\lambda}, \hat{\theta})$, be the MLE of $\alpha = (\lambda, \theta)$. The observed Fisher information matrix is given by:

$$I(\alpha) = \begin{bmatrix} \frac{\partial \ln L(\theta, \lambda)}{\partial \lambda^2} & \frac{\partial \ln L(\theta, \lambda)}{\partial \lambda \partial \theta} \\ \frac{\partial \ln L(\theta, \lambda)}{\partial \theta \partial \lambda} & \frac{\partial \ln L(\theta, \lambda)}{\partial \theta^2} \end{bmatrix}$$

$$\frac{\partial \ln L(\theta, \lambda)}{\partial \lambda^2} = \sum_{i=1}^r \frac{(x_i^\theta + 1)^2}{((\lambda + (\lambda + 1)x_i^\theta))^2} + \frac{(n-r)(x_{(r)}^\theta + 1)x_{(r)}^\theta e^{\lambda x_{(r)}^{-\theta}}}{(x_{(r)}^\theta - (x_{(r)}^\theta + 1)e^{\lambda x_{(r)}^{-\theta}})^2}$$

$$\frac{\partial \ln L(\theta, \lambda)}{\partial \theta \partial \lambda} = \sum_{i=1}^r \left(\frac{x_i^{-\theta} \log(x_i)}{\lambda + (\lambda + 1)x_i^\theta} - \frac{(\lambda + 1)x_i^\theta (x_i^\theta + 1) \log(x_i)}{(\lambda + (\lambda + 1)x_i^\theta)^2} \right) - \sum_{i=1}^r x_i^{-\theta} (-\log(x_i))$$

Thus, the observed variance-covariance matrix becomes $I^{-1}(\hat{\alpha})$. The asymptotic distribution of MLE $\hat{\alpha}$ is a bivariate normal distribution as $\hat{\alpha} \sim N(0, I^{-1}(\hat{\alpha}))$. Consequently, two sided equal tailed $100(1-\eta)\%$

asymptotic confidence intervals for the parameters λ and θ are given by $\left[\hat{\lambda} + Z_{\frac{\eta}{2}} \sqrt{\text{var}(\hat{\lambda})} \right]$ and $\left[\hat{\theta} + Z_{\frac{\eta}{2}} \sqrt{\text{var}(\hat{\theta})} \right]$ respectively. Here, $\text{Var}(\hat{\lambda})$ and $\text{Var}(\hat{\theta})$ are diagonal elements of the observed variance-covariance matrix $I^{-1}(\hat{\alpha})$ and $Z_{\frac{\eta}{2}}$ is the upper $\left(Z_{\frac{\eta}{2}} \right)^{th}$ percentile of the standard normal distribution.

V. The Bayesian Estimation

In this section, we discuss the Bayes estimators of the unknown parameters of the model in (2) under square error loss function (SELF). In order to select the best decision in decision theory, an appropriate loss function must be specified. SELF is generally used for this purpose. The use of the SELF is well justified when over estimation and under estimation of equal magnitude has the same consequences. When the true loss is not symmetric with respect to over estimation and under estimation, asymmetric loss functions are used to represent the consequences of different errors. If all parameters of the model are unknown, a joint conjugate prior for the parameters does not exist. In such conditions there are numerous ways to choose the priors. Hence, we choose to consider the piecewise independent priors. The proposed priors for the parameters λ and θ may be taken as:

$$\begin{aligned} g_1(\lambda) &= \lambda^{a_1-1} e^{-\lambda b_1}, & a_1, b_1 > 0 \\ g_2(\theta) &= \lambda^{a_1-1} e^{-\lambda b_1}, & a_2, b_2 > 0 \end{aligned}$$

Thus, the joint prior distribution of λ and θ can be written as:

$$g(\lambda, \theta) = \lambda^{a_1-1} \theta^{a_2-1} e^{-(\lambda b_1 + \theta b_2)} \tag{4.1}$$

Now we derive the Bayes estimators for the unknown parameters λ and θ under squared error loss function. If μ is the parameter to be estimated by an estimator $\hat{\mu}$ then the squared error loss function is defined as $L_s(\mu, \hat{\mu}) = (\mu - \hat{\mu})^2$. The joint posterior distribution of λ and θ after simplification is:

$$\Pi(\lambda, \theta | \underline{x}) = \frac{\frac{n}{(n-r)} \lambda^{a_1-1} \theta^{a_2-1} e^{-(\lambda b_1 + \theta b_2)} \prod_{i=0}^r f(x_i) (1 - F(x))^{n-r}}{\int_0^\infty \int_0^\infty \frac{n}{(n-r)} \lambda^{a_1-1} \theta^{a_2-1} e^{-(\lambda b_1 + \theta b_2)} \prod_{i=0}^r f(x_i) (1 - F(x))^{n-r} \partial \lambda \partial \theta} \tag{4.2}$$

Therefore, the Bayes estimator of any function of λ and θ , say $\alpha(\hat{\lambda}, \hat{\theta})$ under squared error loss function is.

I. Subsection One

Lindley's Approximation

It is difficult to compute Eq. (4.2) analytically. Lindley's [15] approximation is used to compute the ratio of integrals of the form Eq. (4.3). Based on Lindley's approximation, the approximate Bayes estimator of λ under the squared error loss function is:

$$\hat{\lambda}_{\text{lindley}} = \hat{\lambda} + \frac{1}{2} [\mu_1 (2\rho_1 \sigma_{11} + 2\rho_2 \sigma_{21} + \sigma_{11}^2 L_{111} + 2\sigma_{12} \sigma_{21} L_{111} + \sigma_{11} \sigma_{22} L_{211} + \sigma_{12} \sigma_{22} L_{222})] \tag{4.4}$$

$$\hat{\theta}_{\text{lindley}} = \hat{\theta} + \frac{1}{2} [\mu_2 (2\rho_2 \sigma_{22} + 2\rho_1 \sigma_{21} + \sigma_{22}^2 L_{222} + 2\sigma_{12} \sigma_{11} L_{111} + 3\sigma_{12} \sigma_{22} L_{122})] \tag{4.5}$$

Here $L(\lambda, \theta)$ is the log-likelihood and $\rho(\lambda, \theta)$ is the log of prior distribution $\pi(\lambda, \theta)$, $\hat{\lambda}$ and $\hat{\theta}$ are the MLEs of λ and θ respectively.

VI. Simulation Study

This section deals with a Monte Carlo simulation study. Here, we compare various estimators developed in the previous sections with the help of Monte Carlo simulation study. Six different sample sizes $n = 50, 60, 70, 80$ and 90 are considered in the simulation study. Following combination of the true values of the parameters $(\lambda, \theta) = (0.5, 1)$ and $(\lambda, \theta) = (1.5, 1)$ are taken. In each case the ML

and Bayes estimates of the unknown parameters are computed. The whole process is simulated 1000 times. Tables 1–2 report the simulation results including Average Estimate (AE), MSE of the IPBH parameters.

Table 1: Bayes estimate of the parameter λ and θ when $\theta = 1$ and $\lambda = 0.5$

n	r	Prior1		Prior2		Prior1		Prior2	
		$\hat{\lambda}$				$\hat{\theta}$			
		AE	MSE	AE	MSE	AE	MSE	AE	MSE
50	46	0.5332	0.018	0.5714	0.0152	1.0862	0.0082	1.0778	0.2459
50	48	0.5321	0.0137	0.5318	0.0142	1.0571	0.0715	1.0754	0.0821
60	56	0.5263	0.0124	0.5268	0.0122	1.0655	0.00615	1.0553	0.00567
60	58	0.5195	0.0102	0.5257	0.0099	1.0525	0.0516	1.0529	0.0588
70	66	0.5173	0.0089	0.5224	0.0091	1.0491	0.0506	1.0551	0.0511
70	68	0.5171	0.0084	0.5223	0.0092	1.0468	0.0485	1.0492	0.0492
80	76	0.5168	0.0071	0.5152	0.007	1.0423	0.0447	1.0468	0.0429
80	78	0.5156	0.0061	0.5187	0.0074	1.0387	0.0394	1.0271	0.0366
90	86	0.5078	0.0054	0.5162	0.0063	1.0311	0.0343	1.0327	0.0364
90	88	0.5115	0.0049	0.511	0.0053	1.0296	0.0316	1.0329	0.0349

Table 2: Bayes estimate of the parameter λ and θ when $\theta = 1$ and $\lambda = 1.5$

n	r	Prior1		Prior2		Prior1		Prior2	
		$\hat{\lambda}$				$\hat{\theta}$			
		AE	MSE	AE	MSE	AE	MSE	AE	MSE
50	46	1.7311	0.464	1.7088	0.394	1.0504	0.0528	1.0498	0.0492
50	48	1.6536	0.2447	1.6543	0.2456	1.0369	0.0407	1.0439	0.0423
60	56	1.6038	0.1654	1.6382	0.1643	1.0295	0.0313	1.0453	0.0338
60	58	1.5855	0.1197	1.5290	0.1250	1.0244	0.0286	1.0306	0.0298
70	66	1.5841	0.1195	1.5771	0.1181	1.0258	0.0248	1.0221	0.0258
70	68	1.5723	0.0956	0.1181	0.0922	1.0243	0.0244	1.0202	0.0231
80	76	1.5636	0.0958	1.5771	0.1127	1.0143	0.0199	1.0285	0.0237
80	78	1.573	0.0856	1.5639	0.0827	1.0228	0.0198	1.0191	0.002
90	86	1.5614	0.0821	1.5587	0.0792	1.0186	0.0182	1.0162	0.0198
90	88	1.5534	0.0712	1.5489	0.0710	1.0199	0.0171	1.0276	0.0175

VII. Real-Life Applications

In this section, we illustrate estimation procedures discussed in the previous sections with the help of one real datasets. Here, we consider a real dataset namely the strengths of glass fibres The Data I, respectively are given below:

Data set:

This dataset consists of 63 observations which are generated to simulate the strengths of glass fibres [18]. The 63 observations of the dataset are as follows: "1.014, 1.081, 1.082, 1.185, 1.223, 1.248, 1.267, 1.271, 1.272, 1.275, 1.276, 1.278, 1.286, 1.288, 1.292, 1.304, 1.306, 1.355, 1.361, 1.364, 1.379, 1.409, 1.426, 1.459, 1.460, 1.476, 1.481, 1.484, 1.501, 1.506, 1.524, 1.526, 1.535, 1.541, 1.568, 1.579, 1.581, 1.591, 1.593, 1.602, 1.666, 1.670, 1.684, 1.691, 1.704, 1.731, 1.735, 1.747, 1.748, 1.757, 1.800, 1.806, 1.867, 1.876, 1.878,

1.910, 1.916, 1.972, 2.012, 2.456, 2.592, 3.197, and 4.121”.

We calculate MLEs of the unknown parameters together with some useful measure of goodness-of-fit tests for one dataset, namely, the negative log likelihood function $-\ln L$, the Akaike information criterion denoted by $AIC = 2k - 2\ln L$, proposed by Akaike [16] and Bayesian information criterion denoted by $BIC = k\ln(n) - 2\ln L$, proposed by Schwarz [17], where k is the number of parameters in the model, n is the number of observations in the given datasets, L is the maximized value of the likelihood function for the estimated model and Kolmogorov-Smirnov (K-S) statistic with its p -value. The best distribution corresponds to the lowest $-\ln L$, AIC , BIC and K-S statistic and the highest p values. The K-S statistic with its p -value is obtained using ks test function in statistical software R. The results of the MLEs and measures of goodness-of-fit tests are reported in Tables 3 and 4, respectively. These results show that IPBH distribution is the best choice for the considered datasets. However, for Data I, according to K-S test IPBH is better than the BH.

Table 3: Data Summary for the Data Set

Min	1 st Qu.	Median	mean	3 rd Qu.	Max
1.014	1.305	1.526	1.616	1.741	4.121

Table 4: Goodness of Fit criterions on the data set

Distribution	Estimates	$-\log L$	AIC	BIC	K-S (stat)	P-value
IPBD	$\hat{\theta} = 5.7408$ $\hat{\lambda} = 6.03415$	15.403	26.8066	39.0929	0.08507	0.7197
BR	$\hat{\theta} = 0.2325$ 0	113.364	222.7286	235.0148	0.77052	< 0.001
Weibull	$\hat{\theta} = 3.0521$ $\hat{\lambda} = 1.7873$	43.254	96.7345	101.12	0.2051	0.009
Exponential	$\hat{\lambda} = 0.6189$	93.222	187.432	190.523	0.4721	< 0.001

VIII. Conclusion

This article deals with the classical and Bayesian estimation procedures for parameters of inverse power Burr-Hatke distribution using second type censoring. The maximum likelihood estimators and corresponding asymptotic confidence intervals based on observed Fisher information matrix of the unknown parameters were derived. The Bayes estimates of the parameters under square error loss function were approximated using Lindley’s approximation. The performance of these estimators was examined by extensive Monte Carlo simulation study, which indicated that the MLEs can be obtained easily and quickly with satisfactory estimates. For more efficient estimators, Bayes estimation method with available prior information or convenient non-informative priors in the absence of prior information is recommended.

References

- [1] Da Silva, R. V., Gomes-Silva, F., Ramos, M. W. A., & Cordeiro, G. M. (2015). The exponentiated Burr XII Poisson distribution with application to lifetime data. *International Journal of Statistics and Probability*, 4(4): 112-131.
- [2] Afify, A. Z., Cordeiro, G. M., Ortega, E. M., Yousof, H. M. and Butt, N. S. (2018). The four parameter Burr XII distribution: properties, regression model, and applications. *Communications in*

Statistics-Theory and Methods, 47(11): 2605–2624.

[3] Alizadeh, M., Lak, F., Rasekhi, M., Ramires, T. G., Yousof, H. M., and Altun, E. (2018). The odd log-logistic Topp–Leone G family of distributions: heteroscedastic regression models and applications. *Computational Statistics*, 33(3): 1217-1244.

[4] Abouelmagd, T. H. M. (2018). The Logarithmic Burr-Hatke Exponential Distribution for Modeling Reliability and Medical Data. *International Journal of Statistics and Probability*, 7(5): 73-85.

[5] Yadav, A. S., Altun, E. and Yousof, H. M. (2021). Burr–Hatke Exponential Distribution: A Decreasing Failure Rate Model, Statistical Inference and Applications. *Annals of Data Science*, 8(2): 241–260.

[6] Nasir, M. A., Tahir, M. H., Chesneau, C., Jamal, F., and Shah, M. A. A. (2020). The odds generalized gamma-G family of distributions: Properties, regressions and applications. *Statistica*, 80(1): 3-38.

[7] Anzagra, L., Sarpong, S., and Nasiru, S. (2020). Chen-G class of distributions. *Cogent Mathematics and Statistics*, 7(2):1721.

[8] Afify, A. Z., Aljohani, H. M., Alghamdi, A. S., Gemeay, A. M. and Sarg, A. M. (2021). A new two-parameter burr-hatke distribution: properties and bayesian and non-bayesian inference with applications. *Journal of Mathematics*. 2021: 16.

[9] Zamanah, E., Nasiru, S., and Luguterah, A. (2022). Harmonic Mixture Weibull-G Family of Distributions: Properties, Regression and Applications to Medical Data. *Computational and Mathematical Methods*, 2022: 24.

[10] Khaazmi, O., Nik, A. S., Hamedani, G. G., and Altun, E. (2022). Harmonic Mixture-G Family of Distributions: Survival Regression, Simulation by Likelihood, Bootstrap and Bayesian Discussion with MCMC Algorithm. *Austrian Journal of Statistics*, 51(2): 1-27.

[11] Altun, E., Alizadeh, M., & Yousof, H. M. (2022). The Odd Log-Logistic Weibull-G Family of Distributions with Regression and Financial Risk Models. *Journal of the Operations Research Society of China*, 10(1): 133-158.

[12] Ahmed, E., Hassan M. A. and Ahmed, Z. A. (2021). Bayesian and Classical Inference under Type-II Censored Samples of the Extended Inverse Gompertz Distribution with Engineering Applications. *Entropy*, 23: 1578.

[13] Hassan, M. A. (2021). Statistical inference of chen distribution based on two progressive type-ii censoring schemes. *Computers, Materials & Continua*, 66: 2797–2814.

[14] Cohen, A. C. (1965). Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and Censored Samples. *Technometrics*, 7: 579–588.

[15] Lindley, D. V. (1980). Approximate Bayesian methods (with discussions), *Trabajos de Estadística*. 31: 232–245.

[16] Akaike, H. A. (1974). New look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716-723.

[17] Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6: 461-464.

ANALYSIS OF TWO NON-IDENTICAL UNIT SYSTEM HAVING SAFE AND UNSAFE FAILURES WITH REBOOTING AND PARAMETRIC ESTIMATION IN CLASSICAL AND BAYESIAN PARADIGMS

POONAM SHARMA AND PAWAN KUMAR

•
Department of Statistics, University of Jammu, J&K
sharmapoonam1038@gmail.com, pkk_skumar@yahoo.co.in

Abstract

The present paper aims at the study of a two non-identical system model having safe and unsafe failures and rebooting. The focus centers on the analysis w.r.t important reliability measures and estimation of parameters in Classical and Bayesian paradigms. At first one of the units is operational whereas other one is confined to standby mode. Any unit may suffer safe or unsafe failure. A safe failure is immediately taken up for remedial action by a repairman available with the system all the time, while the case of unsafe failure cannot be dealt directly but first rebooting is performed to convert the unsafe failure to safe failure mode so as to start repair normally. A switching device is used to make the repaired and standby units operational. The lifetime of both the units and switching device are taken to be exponentially distributed random variables whereas the distribution of repair times are assumed to be general. Regenerative point technique is employed to derive associated measures of effectiveness. To make the study more elaborative and visually attractive, some of the derived characteristics have been studied graphically too. A simulation study has also been undertaken to exhibit the behaviour of obtained characteristics in Classical and Bayesian setup. Valuable inferences about MLE and Bayes estimates have been drawn from the tables and graphs for varying values of failure and repair parameters.

Keywords: Reliability, Availability, Mean Time to System Failure, Regenerative Point Technique, Rebooting, Coverage Probability, Bayesian Estimation, Maximum Likelihood Estimation.

1. INTRODUCTION

Reliability is a fundamental concept that underpins the dependability and consistency of systems, processes, products, or services. It is the assurance that something will perform its purposeful function or deliver expected outcomes consistently and without failure over a specified period or under specific conditions. In a world where technological advancements and complex interdependencies are ever-increasing, reliability has become a critical factor in determining the success, safety, and satisfaction of individuals, businesses, and societies at large. We observe that machine failure, which results in significant losses, frequently follows unit failure. The incorporation of standby units is one strategy for enhancing reliability. Also, there are cases where the root cause of a unit failure is not immediately identified, leading to inadequate coverage that must be fixed by rebooting. Depending on the complexity, the length of the reboot time varies from system to system. Recent times have seen extensive and rigorous research on reliability, availability, standby systems, inadequate coverage, reboot, etc. Sharma & Kumar[1] examined the concept of two similar units with one switching device and imperfect coverage. In case of unsafe failure, repair cannot begin immediately but first rebooting is done which transforms the unsafe failure

to safe failure and then repair is carried out. Trivedi[2] gave the concept of reboot in his work "Probability & Statistics With Reliability, Queuing and Computer Science Applications" Gupta et al.[3] carried a study about two dissimilar unit parallel system accompanied by correlated lifetimes. The system stops functioning when both of the units fail. Wang & Chen [4] provided a comparative analysis by computing the availability of three systems with General Repair times and Reboot delay. Pham [5] performed analysis of reliability of a system with high voltage having dependent failures and insufficient coverage. A high voltage (HV) system that consists of a power supply and two transmitters is considered. Also a model of the HV system and a detailed development of the reliability function are presented. Ke et al. [6] examined a resolvable system with insufficient coverage and reboot. As a unit fails, it can be immediately detected, and replaced with a coverage probability c . Kumar P & Jain M [7] proposed the machine having multi-components with service interruption, imperfect coverage, and reboot. Kadyan & Malik [8] performed a stochastic study on non-identical units with cold standby units operating at the same time. The idea of Classical and Bayesian estimation in a two non-identical unit parallel system is given by Saxena et.al[9], where the Bayesian estimates are calculated by taking different priors. Also, a comparative study is done to determine the performance of Maximum likelihood estimation and Bayesian estimation methods. Kishan & Jain [10] put forth the idea of study of system model both in classical and Bayesian perspectives and some important measures of reliability characteristics of a two nonidentical unit standby system model with repair, inspection and post repair are obtained using regenerative point technique.

Keeping above ideas in mind, this paper deals with the performance measures and estimation of parameters of a two non-identical units system with switching device and rebooting having safe and unsafe failures. Switch is used to turn on the unit from standby to operational mode and initially is assumed to be in good condition. Unsafe failures occur when the cause of any of the breakdowns is unknown and can be resolved by rebooting. Reboot delay times and failure times for both units and switch are assumed to be exponentially distributed, whereas the repair time distributions are taken to be general in nature. Other measures, such as mean time to system failure, reliability, availability, and expected number of repairs, have been calculated using the regenerating point technique. Furthermore, a simulation study is carried out to examine the given system model in both the Classical and Bayesian setups. Finally, numerous noteworthy conclusions are drawn from the tables and graphs.

2. SYSTEM DESCRIPTION AND ASSUMPTIONS

- The system is composed of two non-identical units, A and B, coupled by a switch, S.
- Initially, one of the units is functioning, while the other remains in standby mode. A switch assists to turn on the repaired and standby components. During the early stage, switch is supposed to be in operable condition.
- There may be both safe and unsafe failures among the units but only a regular switch failure. If any of the unit fails safely, it can be identified with coverage probability c , and repaired instantly if the repairman is present.
- In the event of an unsafe failure, repair can't begin instantly; instead, a reboot is first performed to convert the unsafe failure to a safe failure, followed by a usual normal repair. Reboot delay periods are taken as exponentially distributed random variables with varying parameters.
- The system has a dedicated repair facility and is constantly accessible to repair and reboot failed items. Switch repair has priority over failed items in the system.
- The failure times of the units and switch follow an exponential distribution, whereas the repair time distributions are general.
- A repaired item functions as new.

3. NOTATIONS AND SYMBOLS

- α_1 : Failure rate of Unit A
- α_2 : Failure rate of Unit B
- α_3 : Failure rate of Switch
- c: Coverage probability
- γ_1 : Rebooting delay rate for unsafe failure of Unit A
- γ_2 : Rebooting delay rate for unsafe failure of Unit B.
- $F_1(\cdot)$: Repair rate of unit A
- $F_2(\cdot)$: Repair rate of Unit B
- $F_3(\cdot)$: Repair rate of Switch

3.1. SYMBOLS FOR THE STATES OF THE SYSTEM

- A_0/B_0 :Units in operative mode
 - A_r/B_r :Units under repair
 - A_s/B_s :Units in standby mode
 - S_g/S_r :Switch under good/repair condition
 - A_{wr}/B_{wr} :Units waiting for repair
 - A_{usf}/B_{usf} :Units having unsafe failure
- Using the symbols provided above, the achievable states of the system are:
- $S_0 = [A_0, B_s, S_g]$
 - $S_1 = [A_r, B_0, S_g]$
 - $S_2 = [A_{usf}, B_g, S_g]$
 - $S_3 = [A_0, B_s, S_r]$
 - $S_4 = [A_{usf}, B_g, S_{wr}]$
 - $S_5 = [A_{wr}, B_0, S_r]$
 - $S_6 = [A_{wr}, B_{usf}, S_{wr}]$
 - $S_7 = [A_{wr}, B_{wr}, S_r]$
 - $S_8 = [A_{wr}, B_{usf}, S_g]$
 - $S_9 = [A_r, B_{wr}, S_g]$
 - $S_{10} = [A_0, B_r, S_g]$
 - $S_{11} = [A_{usf}, B_{wr}, S_g]$
 - $S_{12} = [A_g, B_{wr}, S_r]$
 - $S_{13} = [A_{wr}, B_g, S_r]$

The transition diagram of the model is shown in Figure 1.

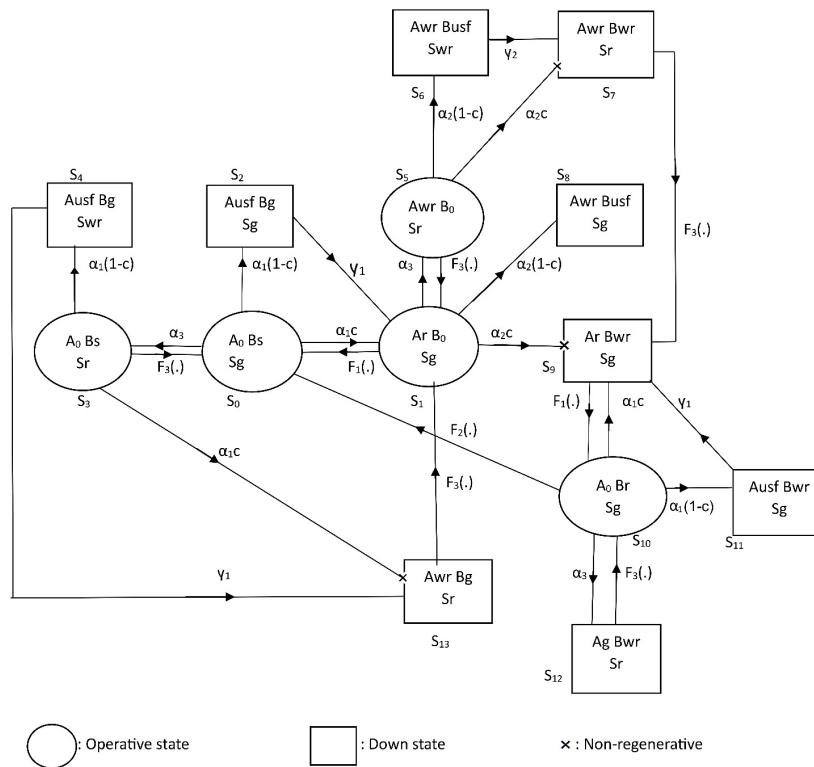


Figure 1: Transition Diagram

4. TRANSITION PROBABILITIES AND SOJOURN TIMES

The long-run or of the steady state probabilities are obtained as under,

$$p_{ij} = \lim_{t \rightarrow \infty} Q_{ij}(t) = \int q_{ij}(t) dt \quad p_{ij}^{(k)} = \lim_{t \rightarrow \infty} Q_{ij}^{(k)}(t) \quad \text{and} \quad p_{ij}^{(k,l)} = \lim_{t \rightarrow \infty} Q_{ij}^{k,l}(t).$$

In particular we have

$$p_{01}(t) = \int \alpha_1 c e^{-\alpha_1 c t} e^{-\alpha_1(1-c)t} e^{-\alpha_3 t} dt = \frac{\alpha_1 c}{(\alpha_1 + \alpha_3)}$$

Similarly,

$$\begin{aligned} p_{02} &= \frac{\alpha_1(1-c)}{\alpha_1 + \alpha_3} & p_{03} &= \frac{\alpha_3}{\alpha_1 + \alpha_3} \\ p_{15} &= \frac{\alpha_3}{\alpha_2 + \alpha_3} [1 - \tilde{F}_1(\alpha_2 + \alpha_3)] & p_{18} &= \frac{\alpha_2(1-c)}{\alpha_2 + \alpha_3} [1 - \tilde{F}_1(\alpha_2 + \alpha_3)] \\ p_{10} &= \tilde{F}_1(\alpha_2 + \alpha_3) & p_{1,10}^{(9)} &= \frac{\alpha_2 c}{\alpha_2 + \alpha_3} [1 - \tilde{F}_1(\alpha_2 + \alpha_3)] \\ p_{30} &= \tilde{F}_3(\alpha_1) & p_{34} &= (1-c)(1 - \tilde{F}_3(\alpha_1)) \\ p_{31}^{(13)} &= c(1 - \tilde{F}_3(\alpha_1)) & p_{51} &= \tilde{F}_3(\alpha_2) \\ p_{56} &= (1-c)[1 - \tilde{F}_3(\alpha_2)] & p_{59}^{(7)} &= c[1 - \tilde{F}_3(\alpha_2)] \\ p_{10,11} &= \frac{\alpha_1(1-c)}{\alpha_1 + \alpha_3} [1 - \tilde{F}_2(\alpha_1 + \alpha_2)] & p_{10,9} &= \frac{\alpha_1 c}{\alpha_1 + \alpha_3} [1 - \tilde{F}_2(\alpha_1 + \alpha_3)] \\ p_{10,12} &= \frac{\alpha_3}{\alpha_1 + \alpha_3} [1 - \tilde{F}_2(\alpha_1 + \alpha_3)] \end{aligned}$$

Thus, the following relationships can be established

$$\begin{aligned} p_{01} + p_{02} + p_{03} &= 1 & p_{30} + p_{34} + p_{31}^{(13)} &= 1 \\ p_{15} + p_{18} + p_{10} + p_{1,10}^{(9)} &= 1 & p_{51} + p_{56} + p_{59}^{(7)} &= 1 \\ p_{10,0} + p_{10,9} + p_{10,11} + p_{10,12} &= 1 \end{aligned}$$

$$p_{67} = p_{79} = p_{89} = p_{13,1} = p_{4,13} = p_{11,9} = p_{21} = p_{12,10} = p_{9,10} = 1$$

4.1. Mean Sojourn times

In reliability, Mean Sojourn time ψ_i , is the expected length of time a system spends in a certain state before moving to another. There is never any transition from S_i to any other state, as long as the system is in state S_i . We utilize this knowledge to determine ψ_i for state S_i . Given T_i as the sojourn time in state S_i , the mean sojourn time ψ_i is as follows.

$$\psi_i = E[T_i] = \int P(T_i > t) dt$$

Hence, using the above formula following values for mean sojourn time are obtained:

$$\begin{aligned} \psi_0 &= \frac{1}{(\alpha_1 + \alpha_3)} & \psi_1 &= \frac{1}{(\alpha_2 + \alpha_3)} [1 - \tilde{F}_1(\alpha_2 + \alpha_3)] \\ \psi_3 &= \frac{1}{\alpha_1} [1 - \tilde{F}_3(\alpha_1)] & \psi_5 &= \frac{1}{\alpha_2} [1 - \tilde{F}_3(\alpha_1)] \\ \psi_2 &= \psi_4 = \psi_{11} = \frac{1}{\gamma_1} & \psi_6 &= \psi_8 = \frac{1}{\gamma_2} \\ \psi_9 &= \int \tilde{F}_1(t) dt & \psi_{10} &= \frac{1}{(\alpha_1 + \alpha_3)} [1 - \tilde{F}_2(\alpha_1 + \alpha_3)] \end{aligned}$$

$$\psi_7 = \psi_{12} = \psi_{13} = \int \tilde{F}_3(t) dt$$

5. ANALYSIS OF RELIABILITY AND MTSF

Let random variable T_i represents the life time of system when it initiate from state $S_i \in E_i$, the system's reliability is determined by:

$$R_i(t) = P[T_i > t]$$

To calculate $R_i(t)$, we treat failed states as absorbing states. The recursive relations between $R_i(t)$ can be established using probabilistic arguments by referring to the state transition diagram. Using the Laplace transform and determining the set of equations for $R_0^*(s)$, we get

$$R_0^*(s) = \frac{N_1(s)}{D_1(s)} \tag{1}$$

¹ where,

$$N_1(s) = Z_0^* + Z_3^*q_{03}^* + Z_5^*q_{01}^*q_{15}^* - Z_0^*q_{15}^*q_{51}^* - Z_3^*q_{03}^*q_{15}^*q_{51}^*$$

$$D_1(s) = 1 - q_{01}^*q_{10}^* - q_{15}^*q_{51}^* - q_{03}^*q_{30}^* + q_{03}^*q_{15}^*q_{30}^*q_{51}^*$$

We obtain the system's reliability by taking the inverse Laplace transform of (1). To obtain MTSF, we use the given formula

$$E(T_0) = \int R_0(t)dt = \lim_{s \rightarrow 0} R_0^*(s) = \frac{N_1(0)}{D_1(0)} \tag{2}$$

where,

$$N_1(0) = \psi_0 + \psi_1p_{01} + \psi_5p_{01}p_{15} - \psi_0p_{15}p_{51} - \psi_3p_{03}p_{15}p_{51}$$

and

$$D_1(0) = 1 - p_{01}p_{10} - p_{15}p_{51} - p_{03}p_{30} + p_{03}p_{15}p_{30}p_{51}$$

Since we've $q_{ij}^*(0) = p_{ij}$ and $\lim_{s \rightarrow 0} Z_i^*(s) = \int Z_i(t)dt = \psi_i$

6. AVAILABILITY ANALYSIS

The probability that a system is able to perform its intended task at time 't' if it initiates from $S_i \in E_i$ is known as Availability. Point wise availability refers to a system's availability at a specified time. It is a measure of system performance that reflects whether a system is potentially operational and able to provide the expected service at a given time. Using stochastic reasoning, recurrence relations between different point-wise availabilities are established. Using the Laplace transformations and solving the equations for $A_0^*(s)$, we obtain

$$A_0^*(s) = \frac{N_2(s)}{D_2(s)}$$

where,

$$\begin{aligned} N_2(s) = & Z_0^*q_{15}^*q_{51}^*b_1 + Z_3^*(q_{02}^* - q_{03}^*b_1) - (Z_0^* + Z_1^*Y_1 + Z_5^*Y_3)(q_{10,9}^*q_{9,10}^*) + Z_{10}^*q_{1,10}^{(9)*}Y_1 - (Z_0^* + Z_1^*Y_1) \\ & (q_{10,12}^*q_{12,10}^*) - Z_3^*q_{03}^* + Z_5^*Y_1 + (Z_1^* + Z_5^*q_{15}^*)q_{03}^*q_{31}^{(13)*}q_{11,9}^*q_{10,11}^*q_{9,10}^* + Z_{10}^*q_{9,10}^*b_3(Z_1^* + Z_5^*q_{15}^*) \\ & (Z_0^* + Z_1^*Y_2)q_{11,9}^*q_{10,11}^*q_{9,10}^* + Z_{10}^*q_{18}^*q_{89}^*q_{9,10}^*Y_1 + q_{03}^*q_{34}^*q_{131}^*q_{4,13}^*b_2 + q_{15}^*q_{9,10}^*Z_{10}^*b_4 + (Z_1^* + Z_3^*) \\ & q_{03}^* + Z_5^*q_{01}^*q_{15}^* \end{aligned}$$

Here,

¹Limits of integration whenever they are 0 to ∞ are not mentioned.

$$Y_1 = q_{01}^* + q_{02}^* q_{21}^* + q_{03}^* q_{31}^{(13)*} + q_{34}^* q_{13,1}^* q_{4,13}^*$$

$$Y_2 = q_{01}^* + q_{02}^* q_{21}^* q_{03}^* q_{34}^* q_{13,1}^* q_{4,13}^*$$

$$Y_3 = q_{02}^* q_{15}^* q_{21}^* + q_{03}^* q_{15}^* q_{31}^{(13)*} q_{03}^* q_{15}^* q_{34}^* q_{13,1}^* q_{4,13}^*$$

$$b_1 = 1 - q_{10,9}^* q_{9,10}^* - q_{10,12}^* q_{12,10}^* - q_{11,9}^* q_{9,10}^* q_{10,11}^*$$

$$b_2 = Z_1^* q_{15}^* Z_5^* + q_{15}^* (q_{59}^{(7)*} q_{9,10}^* Z_{10}^* + q_{56}^* q_{67}^* q_{9,10}^* Z_{10}^*)$$

$$b_3 = q_{18}^* q_{89}^* + q_{15}^* (q_{59}^{(7)*} + q_{56}^* q_{67}^*)$$

$$b_4 = (q_{59}^{(7)*} (q_{01}^* + q_{02}^* q_{21}^*) + q_{01}^* q_{56}^* q_{67}^*)$$

and,

$$D_2(s) = q_{10,0}^* - q_{10}^* q_{10,0}^* + q_{01}^* q_{15}^* q_{10,0}^* - (q_{56}^* + q_{59}^{(7)*}) [-q_{15}^* q_{10,0}^* (q_{01}^* + q_{02}^* + q_{03}^* (q_{31}^{(13)*} + q_{34}^*)) - q_{02}^* q_{10,0}^*] \\ - q_{03}^* q_{10,0}^* (q_{30}^* - q_{30}^* q_{13}^* q_{51}^* + q_{10}^* q_{31}^{(13)*} + q_{10}^* q_{34}^*) - q_{15}^* q_{51}^* q_{10,0}^* - q_{10,0}^* (q_{34}^* + q_{31}^{(13)*}) (q_{03}^* (q_{1,10}^{(9)*} + q_{18}^*)) \quad (3)$$

The steady state availability is given as under

$$A_0 = \lim_{t \rightarrow \infty} A_0(t) = \lim_{s \rightarrow 0} s A_0^*(s) = \frac{N_2(0)}{D_2(0)}$$

Furthermore, its a well known fact that $q_{ij}(t)$ is the pdf of the time of transition from state S_i to S_j and $q_{ij}^*(s)$ is the probability of a transition from state S_i to state S_j during the interval $(t, t + dt)$, thus

$$q_{ij}^*(s)|_{s=0} = q_{ij}^*(0) = p_{ij}$$

We also know that

$$\lim_{s \rightarrow 0} Z_i^*(s) = \int Z_i(t) dt = \psi_i$$

Therefore,

$$N_2(0) = \psi_0 p_{15} p_{51} b_1 + \psi_3 (p_{02} - p_{03} b_1) - (\psi_0 + \psi_1 Y_1 + \psi_5 Y_3) (p_{10,9} p_{9,10}) + \psi_{10} p_{1,10}^{(9)} Y_1 - (\psi_0 + \psi_1 Y_1) \\ (p_{10,12} p_{12,10}) - \psi_3 p_{03} + \psi_5 Y_1 + (\psi_1 + \psi_5 p_{15}) p_{03} p_{31}^{(13)} p_{11,9} p_{10,11} p_{9,10} + \psi_{10} p_{9,10} b_3 (\psi_1 + \psi_5 p_{15}) \\ (\psi_0 + \psi_1 Y_2) p_{11,9} p_{10,11} p_{9,10} + \psi_{10} p_{18} p_{89} p_{9,10} Y_1 + p_{03} p_{34} p_{13,1} p_{4,13} b_2 + p_{15} p_{9,10} \psi_{10} b_4 + (\psi_1 + \psi_3) \\ p_{03} + \psi_5 p_{01} p_{15}$$

Here,

$$Y_1 = p_{01} + p_{02} p_{21} + p_{03} (p_{31}^{(13)} + p_{34} p_{13,1} p_{4,13})$$

$$Y_2 = p_{01} + p_{02} p_{21} p_{03} p_{34} p_{13,1} p_{4,13}$$

$$Y_3 = p_{02} p_{15} p_{21} + p_{03} p_{15} p_{31}^{(13)} p_{03} p_{15} p_{34} p_{13,1} p_{4,13}$$

$$b_1 = 1 - p_{10,9} p_{9,10} - p_{10,12} p_{12,10} - p_{11,9} p_{9,10} p_{10,11}$$

$$b_2 = \psi_1 p_{15} \psi_5 + p_{15} (p_{59}^{(7)} p_{9,10} \psi_{10} + p_{56} p_{67} p_{9,10} \psi_{10})$$

$$b_3 = p_{18}p_{89} + p_{15} (p_{59}^{(7)} + p_{56}p_{67})$$

$$b_4 = (p_{59}^{(7)}(p_{01} + p_{02} p_{21}) + p_{01}p_{56}p_{67})$$

$$D_2(0) = p_{10,0} - p_{01}p_{10,0} + p_{01}p_{15}p_{10,0} - (p_{56} + p_{59}^{(7)})[-p_{15}p_{10,0}(p_{01} + p_{02} + p_{03}(p_{31}^{(13)} + p_{34})) - p_{02}p_{10,0}] \\ - p_{03}p_{10,0}(p_{30} - p_{30}p_{13}p_{51} + p_{10}p_{31}^{(13)} + p_{10}p_{34}) - p_{15}p_{51}p_{10,0} - p_{10,0}(p_{34} + p_{31}^{(13)})(p_{03}(p_{1,10}^{(9)} + p_{18}))$$

For a given system, the steady-state probability of its long-term operation is given by

$$A_0 = \lim_{t \rightarrow \infty} A_0(t) = \lim_{s \rightarrow 0} sA_0^*(s)$$

$$\lim_{s \rightarrow 0} \frac{sN_2(s)}{D_2(s)} = \lim_{s \rightarrow 0} N_2(s) \lim_{s \rightarrow 0} \frac{s}{D_2(s)}$$

Since as $s \rightarrow 0$, $D_2(s)$ becomes zero. Therefore, applying L'Hospital's rule, A_0 becomes

$$A_0 = \frac{N_2(0)}{D_2'(0)} \tag{4}$$

where,

$$D_2'(0) = p_{10,0}(\psi_0 + \psi_1) - p_{10,0}p_{03}[(\psi_3)(1 - p_{15}p_{51}) + p_{30}(\psi_1 + p_{15}\psi_5) - p_{34} + p_{34}p_{15}p_{51}(\psi_4 + \psi_{13}) \\ + p_{18}p_{30}\psi_8 + \psi_9(p_{10}(1 - p_{30}) - p_{30}p_{15}) - p_{15}p_{10,0}[(p_{51}\psi_0 - \psi_5 - \psi_6(p_{56}(1 - p_{03}p_{30}))) + \psi_9 \\ (1 - p_{51}) + p_{02}p_{51}\psi_2] - p_{15}p_{51}[(1 - p_{03}p_{30})(\psi_{10} + \psi_{11}p_{10,11}) - \psi_4] + p_{10,12}\psi_{12} + \psi_9(p_{10,9} + \\ p_{10,11} - p_{03}p_{30}[(1 - p_{10})(\psi_{10} + \psi_{11}p_{10,11}) + \psi_{12}p_{10,12}(1 - p_{15}p_{51}) + \psi_9(p_{10,9} + p_{10,11} - p_{10}(1 - \\ p_{10,12}) + p_{18}p_{10,0})] - p_{10}(\psi_{10} + \psi_{11}p_{10,11}) + \psi_{12}(p_{10,12})(1 - p_{34}) + \psi_9[(1 - p_{03})(p_{10,9} + p_{10,11}) \\ + p_{03}(1 - p_{10,12} + p_{34})] + p_{10,0}[(p_{02}\psi_2 + p_{18}(\psi_8 + \psi_9))] + p_{10,11}(\psi_{11} + \psi_9) + p_{10,12}\psi_{12} + \psi_{10} \tag{5}$$

Using $N_2(0)$ and $D_2'(0)$ in equation[4], the expression for A_0 can be determined. The system's expected uptime for $(0,t]$ is provided by

$$\mu_{up}(t) = \int_0^t A_0(u)du$$

So that,

$$\mu_{up}^*(s) = \frac{A_0^*(s)}{s}$$

7. BUSY PERIOD ANALYSIS

$B_i(t)$ is defined as the probability that, at time $t=0$, the system, which begins in the regenerative state $S_i \in E$, is undergoing repair as a result of a unit failure. To estimate these probabilities, we utilize simple probabilistic logics, on taking Laplace transformation and solving the consequent set of equations for $B_0^*(s)$, we have

$$B_0^*(s) = \frac{N_3(s)}{D_2(s)}$$

$$\begin{aligned}
 N_3(s) = & q_{03}^* q_{34}^* [\psi_4 - q_{15}^* q_{51}^* b_1 - q_{10,9}^* q_{9,10}^* b_5 - q_{10,12}^* q_{12,10}^* b_6 + q_{13,1}^* q_{4,13}^* b_7 + q_{1,10}^{(9)*} (\psi_{10} \\
 & + q_{10,9}^* \psi_9 + q_{10,11}^* \psi_{11} + q_{11,9}^* q_{10,11}^* \psi_9) + q_{15}^* q_{56}^* b_8 + q_{15}^* q_{59}^{(7)*} (\psi_9 + q_{9,10}^* \psi_{10} \\
 & + q_{9,10}^* q_{10,11}^* \psi_{11}) + \psi_1 (1 - q_{11,9}^* q_{9,10}^* q_{10,11}^*)] + \psi_4 (1 - q_{11,9}^* q_{9,10}^* q_{10,11}^*) + q_{1,10}^{(9)*} q_{01}^* b_9 \\
 & + q_{02}^* q_{21}^* b_9 + q_{03}^* q_{31}^{(13)*} b_9 + q_{31}^{(13)*} [q_{03}^* q_{18}^* (\psi_8 b_1)] + q_{15}^* q_{56}^* [q_{03}^* \psi_6 b_1 + q_{03}^* q_{67}^* q_{9,10}^* (\psi_{10} \\
 & + q_{10,11}^* \psi_{11})] + q_{15}^* q_{59}^{(7)*} (q_{03}^* q_{9,10}^* (\psi_{10} + q_{10,11}^* \psi_{11}) - q_{03}^* q_{10,12}^* q_{12,10}^* \psi_9) + q_{03}^* \psi_1 b_1 \\
 & + q_{18}^* [q_{02}^* q_{21}^* (\psi_8 b_1) + q_{89}^* q_{9,10}^* (\psi_{10} + q_{10,11}^* \psi_{11})] + q_{01}^* \psi_8 b_1 + q_{15}^* [q_{01}^* q_{56}^* (\psi_6 b_1 \\
 & + q_{67}^*) \psi_9 (1 - q_{10,12}^* q_{12,10}^*) + q_{67}^* q_{9,10}^* (\psi_{10} + q_{10,11}^* \psi_{11})] + q_{01}^* q_{59}^{(7)*} (\psi_9 (1 - q_{10,12}^* q_{12,10}^*) \\
 & + q_{9,10}^* \psi_{10}) + q_{02}^* q_{21}^* [\psi_6 (q_{51}^* - q_{56}^* b_1)] + q_{59}^{(7)*} \psi_9 (1 - q_{10,12}^* q_{12,10}^*) + q_{56}^* q_{67}^* b_{10} \\
 & + q_{02}^* q_{51}^* (-\psi_2 b_1 + q_{03}^* q_{59}^{(7)*} q_{31}^{(13)*} \psi_9) + q_{9,10}^* b_{11} - q_{02}^* (q_{10,9}^* + q_{11,9}^* q_{10,11}^*) (\psi_2 + \psi_1 q_{21}^*) \\
 & + q_{01}^* \psi_1 + q_{02}^* (\psi_2 + q_{21}^* \psi_1) (1 - q_{10,12}^* q_{12,10}^*)
 \end{aligned}$$

here,

$$b_1 = 1 - q_{10,9}^* q_{9,10}^* - q_{10,12}^* q_{12,10}^* - q_{11,9}^* q_{9,10}^* q_{10,11}^*$$

$$b_5 = \psi_4 - q_{13,1}^* q_{4,13}^* (\psi_1 + q_{18}^* \psi_8 + q_{15}^* q_{56}^* \psi_6)$$

$$b_6 = \psi_4 - q_{13,1}^* q_{4,13}^* (\psi_1 q_{18}^* \psi_8 + q_{18}^* q_{89}^* \psi_9 + q_{15}^* q_{59}^{(7)*} \psi_9 + q_{15}^* q_{56}^* q_{67}^* \psi_9)$$

$$b_7 = q_{18}^* (\psi_8 + q_{89}^* \psi_9 + q_{89}^* q_{9,10}^* \psi_{10} + q_{89}^* q_{9,10}^* q_{10,11}^* \psi_{11} - q_{11,9}^* q_{9,10}^* q_{10,11}^* \psi_8)$$

$$b_8 = \psi_6 + q_{67}^* \psi_9 + q_{67}^* q_{9,10}^* \psi_{10} - q_{10,12}^* q_{12,10}^* \psi_6 - q_{11,9}^* q_{9,10}^* q_{10,11}^* \psi_6 + q_{67}^* q_{9,10}^* q_{10,11}^* \psi_{11}$$

$$b_9 = \psi_{10} + \psi_9 q_{10,9}^* + q_{10,11}^* (\psi_{11} + q_{11,9}^* \psi_9)$$

$$b_{10} = \psi_9 + q_{9,10}^* \psi_{10} + q_{9,10}^* q_{10,11}^* \psi_{11} - q_{10,12}^* q_{12,10}^* \psi_9$$

$$b_{11} = q_{01}^* (\psi_1 (q_{10,9}^* + q_{11,9}^* q_{10,11}^*)) + q_{15}^* q_{59}^{(7)*} q_{10,11}^* \psi_{11}$$

and, $D_2(s)$ is same as given in equation [3].

The probability that the repairman will be busy in the long run is as follows:

$$B_0 = \lim_{t \rightarrow \infty} B_0(t) = \lim_{s \rightarrow 0} s B_0^*(s) = \frac{N_3(0)}{D_2(0)}$$

where,

$$\begin{aligned}
 N_3(0) = & p_{03} p_{34} [\psi_4 - p_{15} p_{51} b_1 - p_{10,9} p_{9,10} b_5 - p_{10,12} p_{12,10} b_6 + p_{13,1} p_{4,13} b_7 + p_{1,10}^{(9)} (\psi_{10} \\
 & + p_{10,9} \psi_9 + p_{10,11} \psi_{11} + p_{11,9} p_{10,11} \psi_9) + p_{15} p_{56} b_8 + p_{15} p_{59}^{(7)} (\psi_9 + p_{9,10} \psi_{10} \\
 & + p_{9,10} p_{10,11} \psi_{11}) + \psi_1 (1 - p_{11,9} p_{9,10} p_{10,11})] + \psi_4 (1 - p_{11,9} p_{9,10} p_{10,11}) + p_{1,10}^{(9)} p_{01} b_9 \\
 & + p_{02} p_{21} b_9 + p_{03} p_{31}^{(13)} b_9 + q_{31}^{(13)} [p_{03} p_{18} (\psi_8 b_1)] + p_{15} p_{56} [p_{03} \psi_6 b_1 + p_{03} p_{67} p_{9,10} (\psi_{10} \\
 & + p_{10,11} \psi_{11})] + p_{15} p_{59}^{(7)} (p_{03} p_{9,10} (\psi_{10} + p_{10,11} \psi_{11}) - p_{03} p_{10,12} p_{12,10} \psi_9) + p_{03} \psi_1 b_1 \\
 & + p_{18} [p_{02} p_{21} (\psi_8 b_1) + p_{89} p_{9,10} (\psi_{10} + p_{10,11} \psi_{11})] + p_{01} \psi_8 b_1 + p_{15} [p_{01} p_{56} (\psi_6 b_1 \\
 & + p_{67}) \psi_9 (1 - p_{10,12} p_{12,10}) + p_{67} p_{9,10} (\psi_{10} + p_{10,11} \psi_{11})] + p_{01} p_{59}^{(7)} (\psi_9 (1 - p_{10,12} p_{12,10}) \\
 & + p_{9,10} \psi_{10}) + p_{02} p_{21} [\psi_6 (p_{51} - p_{56} b_1)] + p_{59}^{(7)} \psi_9 (1 - p_{10,12} p_{12,10}) + p_{56} p_{67} b_{10} \\
 & + p_{02} p_{51} (-\psi_2 b_1 + p_{03} p_{59}^{(7)} p_{31}^{(13)} \psi_9) + p_{9,10} b_{11} - p_{02} (p_{10,9} + p_{11,9} p_{10,11}) (\psi_2 + \psi_1 p_{21}) \\
 & + p_{01} \psi_1 + p_{02} (\psi_2 + p_{21} \psi_1) (1 - p_{10,12} p_{12,10})
 \end{aligned}$$

here,

$$b_1 = 1 - p_{10,9}p_{9,10} - p_{10,12}p_{12,10} - p_{11,9}p_{9,10}p_{10,11}$$

$$b_5 = \psi_4 - p_{13,1}p_{4,13}(\psi_1 + p_{18}\psi_8 + p_{15}p_{56}\psi_6)$$

$$b_6 = \psi_4 - p_{13,1}p_{4,13}(\psi_1 p_{18}\psi_8 + p_{18}p_{89}\psi_9 + p_{15}p_{59}^{(7)}\psi_9 + p_{15}p_{56}p_{67}\psi_9)$$

$$b_7 = p_{18}(\psi_8 + p_{89}\psi_9 + p_{89}p_{9,10}\psi_{10} + p_{89}p_{9,10}p_{10,11}\psi_{11} - p_{11,9}p_{9,10}p_{10,11}\psi_8)$$

$$b_8 = \psi_6 + p_{67}\psi_9 + p_{67}p_{9,10}\psi_{10} - p_{10,12}p_{12,10}\psi_6 - p_{11,9}p_{9,10}p_{10,11}\psi_6 + p_{67}p_{9,10}p_{10,11}\psi_{11}$$

$$b_9 = \psi_{10} + \psi_9 p_{10,9} + p_{10,11}(\psi_{11} + p_{11,9}\psi_9)$$

$$b_{10} = \psi_9 + p_{9,10}\psi_{10} + p_{9,10}p_{10,11}\psi_{11} - p_{10,12}p_{12,10}\psi_9$$

$$b_{11} = p_{01}(\psi_1(p_{10,9} + p_{11,9}p_{10,11})) + p_{15}p_{59}^{(7)}p_{10,11}\psi_{11}$$

and $D_2'(0)$ is same as obtained in [5].

During $(0,t]$, the repairman's expected busy time is given by

$$\mu_b(t) = \int_0^t B_0(u)du$$

So that,

$$\mu_b^*(s) = \frac{B_0^*(s)}{s}$$

8. EXPECTED NUMBER OF REPAIRS

When the system begins from regenerative state S_i , $V_i(t)$ is described as the expected number of repairs over the time range $(0,t]$ of the failed units. Furthermore, given the definition of $V_i(t)$, the recurrence relations can be framed easily and, taking their Laplace- Stieltjes transformations and solving the consequent set of equations for $\tilde{V}_0(s)$, we get

$$\tilde{V}_0(s) = N_4(s)/D_3(s)$$

where,

$$\begin{aligned} N_4(s) = & \tilde{Q}_{02}\tilde{Q}_{21}[-\tilde{Q}_{10}b_2 - \tilde{Q}_{1,10}^{(9)}b_2\tilde{Q}_{18}\tilde{Q}_{89}\tilde{Q}_{9,10}b_2 - \tilde{Q}_{15}\tilde{Q}_{59}^{(7)}\tilde{Q}_{9,10}b_2 - \tilde{Q}_{15}\tilde{Q}_{56}\tilde{Q}_{67}\tilde{Q}_{9,10}b_2] \\ & + \tilde{Q}_{03}\tilde{Q}_{34}[\tilde{Q}_{13,1}\tilde{Q}_{4,13}b_1 - \tilde{Q}_{10}b_1 - \tilde{Q}_{1,10}^{(9)}b_2 + \tilde{Q}_{15}\tilde{Q}_{51}b_1 - \tilde{Q}_{18}\tilde{Q}_{89}\tilde{Q}_{9,10}b_2 - \tilde{Q}_{15}\tilde{Q}_{56} \\ & \tilde{Q}_{67}\tilde{Q}_{9,10}b_2 - \tilde{Q}_{15}\tilde{Q}_{59}^{(7)}\tilde{Q}_{9,10}b_2] + \tilde{Q}_{18}\tilde{Q}_{89}[-\tilde{Q}_{01}\tilde{Q}_{9,10}b_2 - \tilde{Q}_{03}\tilde{Q}_{9,10}\tilde{Q}_{31}^{(13)}b_2]\tilde{Q}_{15} \\ & [-\tilde{Q}_{01}\tilde{Q}_{59}^{(7)}\tilde{Q}_{9,10}b_2 - \tilde{Q}_{01}\tilde{Q}_{56}\tilde{Q}_{67}\tilde{Q}_{9,10}b_2] - \tilde{Q}_{01}\tilde{Q}_{10}b_1 \end{aligned}$$

here,

$$b_1 = 1 - \tilde{Q}_{10,9}\tilde{Q}_{9,10} - \tilde{Q}_{10,12}\tilde{Q}_{12,10} - \tilde{Q}_{11,9}\tilde{Q}_{9,10}\tilde{Q}_{10,11}$$

$$b_2 = 1 - \tilde{Q}_{10,0} - \tilde{Q}_{10,12}\tilde{Q}_{12,10}$$

and $D_3(s)$ is written by replacing q_{ij}^* and $q_{ij}^{(k)*}$ by \tilde{Q}_{ij} and $\tilde{Q}_{ij}^{(k)}$ respectively in equation[3]. The expected number of repairs per unit over time in the steady state is represented as

$$V_0 = \lim_{t \rightarrow \infty} V_0(t) = \lim_{s \rightarrow 0} s \tilde{V}_0(s) = \frac{N_4(0)}{D_2(0)}$$

where,

$$N_4(0) = p_{02}p_{21}[p_{10}b_2 - p_{1,10}^{(9)}b_2 - p_{18}p_{89}p_{9,10}b_2 - p_{15}p_{59}p_{9,10}^{(7)}b_2 - p_{15}p_{56}p_{67}p_{9,10}b_2] + p_{03}p_{34} \\ [p_{13,1}p_{4,13}b_1 - p_{10}b_1 - p_{1,10}^{(9)}b_2 + p_{15}p_{51}b_1 - p_{18}p_{89}p_{9,10}b_2 - p_{15}p_{56}p_{67}p_{9,10}b_2 - p_{15}p_{59}^{(7)} \\ p_{9,10}b_2] + p_{18}p_{89}[p_{01}p_{9,10}b_2 - p_{03}p_{9,10}p_{31}^{(13)}b_2]p_{15}[-p_{01}p_{59}p_{9,10}^{(7)}b_2 - p_{01}p_{56}p_{67}p_{9,10}b_2] - p_{01}p_{10}b_1)$$

here,

$$b_1 = 1 - p_{10,9}p_{9,10} - p_{10,12}p_{12,10} - p_{11,9}p_{9,10}p_{10,11}$$

$$b_2 = 1 - p_{10,0} - p_{10,12}p_{12,10}$$

9. PROFIT FUNCTION ANALYSIS

Having determined the reliability characteristics, the profit function P(t) can be calculated. Profit is defined as excess of revenue over the cost, hence the expected total profit made during(0,t] is expressed as :

$$P(t) = \text{Expected total revenue in}(0,t] - \text{Expected total expenditure in}(0,t]$$

$$= K_0\mu_{up}(t) - K_1\mu_b(t) - K_2V_0(t)$$

where,

K_0 = revenue per unit up time of the system.

K_1 = The cost per unit during which the repairman is engaged to fix the failed unit.

K_2 = Cost of repair of each unit.

The expected total gain per unit of time in steady state is provided by:

$$P = \lim_{t \rightarrow \infty} \frac{P(t)}{t} = \lim_{s \rightarrow 0} s^2 P^*(s)$$

Therefore, we have

$$P = K_0A_0 - K_1B_0 - K_2V_0 \tag{6}$$

10. ESTIMATION OF THE PARAMETERS, MTSF, AND PROFIT FUNCTION

10.1. Classical Estimation

10.1.1 ML Estimation

Let us take

$$\tilde{X}_1 = (x_{11}, x_{12}, \dots, x_{1n_1}), \quad \tilde{X}_2 = (x_{21}, x_{22}, \dots, x_{2n_2}), \quad \tilde{X}_3 = (x_{31}, x_{32}, \dots, x_{3n_3}), \\ \tilde{X}_4 = (x_{41}, x_{42}, \dots, x_{4n_4}), \quad \tilde{X}_5 = (x_{51}, x_{52}, \dots, x_{5n_5}), \quad \tilde{X}_6 = (x_{61}, x_{62}, \dots, x_{6n_6}), \\ \tilde{X}_7 = (x_{71}, x_{72}, \dots, x_{7n_7}) \text{ and } \tilde{X}_8 = (x_{81}, x_{82}, \dots, x_{8n_8})$$

Therefore, Likelihood function of combined sample is :

$$L = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8 | \alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2, \lambda_3, \gamma_1, \gamma_2)$$

The pdf of exponential distribution is $f(x, \lambda) = \lambda \exp(-\lambda x)$, $x > 0$, $\lambda > 0$

$$L = \alpha_1^{n_1} \alpha_2^{n_2} \alpha_3^{n_3} \lambda_1^{n_4} \lambda_2^{n_5} \lambda_3^{n_6} \gamma_1^{n_7} \gamma_2^{n_8} \exp - (\alpha_1 W_1 + \alpha_2 W_2 + \alpha_3 W_3 + \lambda_1 W_4 + \lambda_2 W_5 \\ + \lambda_3 W_6 + \gamma_1 W_7 + \gamma_2 W_8)$$

Here, $W_i = \sum_{j=1}^{n_i} x_{ij}$; $i = 1,2,3,4,5,6,7,8$
On solving, we get

$$\log L = n_1 \log \alpha_1 + n_2 \log \alpha_2 + n_3 \log \alpha_3 + n_4 \log \lambda_1 + n_5 \log \lambda_2 + n_6 \log \lambda_3 + n_7 \log \gamma_1 + n_8 \log \gamma_2 \quad (7)$$

$$-(\alpha_1 W_1 + \alpha_2 W_2 + \alpha_3 W_3 + \lambda_1 W_4 + \lambda_2 W_5 + \lambda_3 W_6 + \gamma_1 W_7 + \gamma_2 W_8)$$

The, MLE $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\gamma}_1, \hat{\gamma}_2)$ of the parameters $(\alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2, \lambda_3, \gamma_1, \gamma_2)$ are as under

$$\hat{\alpha}_1 = \frac{n_1}{W_1}, \quad \hat{\alpha}_2 = \frac{n_2}{W_2}$$

$$\hat{\alpha}_3 = \frac{n_3}{W_3}, \quad \hat{\lambda}_1 = \frac{n_4}{W_4}$$

$$\hat{\lambda}_2 = \frac{n_5}{W_5}, \quad \hat{\lambda}_3 = \frac{n_6}{W_6}$$

$$\hat{\gamma}_1 = \frac{n_7}{W_7}, \quad \hat{\gamma}_2 = \frac{n_8}{W_8}$$

The asymptotic distribution of $(\hat{\alpha}_1 - \alpha_1, \hat{\alpha}_2 - \alpha_2, \hat{\alpha}_3 - \alpha_3, \hat{\lambda}_1 - \lambda_1, \hat{\lambda}_2 - \lambda_2, \hat{\lambda}_3 - \lambda_3, \hat{\gamma}_1 - \gamma_1, \hat{\gamma}_2 - \gamma_2) \sim N_8(0, I^{-1})$, where I is the Fisher Information matrix with diagonal elements as

$$I_{11} = \frac{n_1}{\alpha_1^2}, \quad I_{22} = \frac{n_2}{\alpha_2^2}, \quad I_{33} = \frac{n_3}{\alpha_3^2}, \quad I_{44} = \frac{n_4}{\lambda_1^2}, \quad I_{55} = \frac{n_5}{\lambda_2^2}, \quad I_{66} = \frac{n_6}{\lambda_3^2}, \quad I_{77} = \frac{n_7}{\gamma_1^2}, \quad I_{88} = \frac{n_8}{\gamma_2^2}$$

and all non-diagonal elements are zero. Using MLE's invariance property, we can extract The MLE \hat{M} & \hat{P} of MTSF and Profit function. Also, asymptotic distribution of $(\hat{M} - M) \sim N(0, A' I^{-1} A)$ & that of $(\hat{P} - P) \sim N(0, B' I^{-1} B)$, where

$$A' = \left(\frac{\delta M}{\delta \alpha_1}, \frac{\delta M}{\delta \alpha_2}, \frac{\delta M}{\delta \alpha_3}, \frac{\delta M}{\delta \lambda_1}, \frac{\delta M}{\delta \lambda_2}, \frac{\delta M}{\delta \lambda_3}, \frac{\delta M}{\delta \gamma_1}, \frac{\delta M}{\delta \gamma_2} \right)$$

$$B' = \left(\frac{\delta P}{\delta \alpha_1}, \frac{\delta P}{\delta \alpha_2}, \frac{\delta P}{\delta \alpha_3}, \frac{\delta P}{\delta \lambda_1}, \frac{\delta P}{\delta \lambda_2}, \frac{\delta P}{\delta \lambda_3}, \frac{\delta P}{\delta \gamma_1}, \frac{\delta P}{\delta \gamma_2} \right)$$

10.2. Bayesian Estimation

Bayesian estimation is a statistical approach which is utilized to determine the impact of prior knowledge as well as the sample information on prior distributions of the parameters under study. The parameters involved in the model are taken to be random variables having independent Gamma prior distribution. Here, we estimate the unknown parameters taking into account the gamma prior distribution and the corresponding PDFs as

$$\alpha_1 \sim \text{Gamma}(a_1, b_1) \quad (\alpha_1, a_1, b_1) > 0, \quad (8)$$

$$\alpha_2 \sim \text{Gamma}(a_2, b_2) \quad (\alpha_2, a_2, b_2) > 0, \quad (9)$$

$$\alpha_3 \sim \text{Gamma}(a_3, b_3) \quad (\alpha_3, a_3, b_3) > 0, \quad (10)$$

$$\lambda_1 \sim \text{Gamma}(a_4, b_4) \quad (\lambda_1, a_4, b_4) > 0, \quad (11)$$

$$\lambda_2 \sim \text{Gamma}(a_5, b_5) \quad (\lambda_2, a_5, b_5) > 0, \quad (12)$$

$$\lambda_3 \sim \text{Gamma}(a_6, b_6) \quad (\lambda_3, a_6, b_6) > 0, \quad (13)$$

$$\gamma_1 \sim \text{Gamma}(a_7, b_7) \quad (\gamma_1, a_7, b_7) > 0, \quad (14)$$

$$\gamma_2 \sim \text{Gamma}(a_8, b_8) \quad (\gamma_2, a_8, b_8) > 0, \quad (15)$$

Here, a_i and b_i ($i = 1,2,3,4,5,6,7,8$) denotes the shape and scale parameters

Now using likelihood function and taking prior distributions, the posterior distributions of these parameters are calculated as given below:

$$\alpha_1 | X_1 \sim \text{Gamma}(n_1 + a_1, b_1 + W_1) \quad (16)$$

$$\alpha_2 | X_2 \sim \text{Gamma}(n_2 + a_2, b_2 + W_2) \quad (17)$$

$$\alpha_3 | X_3 \underset{\sim}{\sim} \text{Gamma}(n_3 + a_3, b_3 + W_3) \tag{18}$$

$$\lambda_1 | X_4 \underset{\sim}{\sim} \text{Gamma}(n_4 + a_4, b_4 + W_4) \tag{19}$$

$$\lambda_2 | X_5 \underset{\sim}{\sim} \text{Gamma}(n_5 + a_5, b_5 + W_5) \tag{20}$$

$$\lambda_3 | X_6 \underset{\sim}{\sim} \text{Gamma}(n_6 + a_6, b_6 + W_6) \tag{21}$$

$$\gamma_1 | X_7 \underset{\sim}{\sim} \text{Gamma}(n_7 + a_7, b_7 + W_7) \tag{22}$$

$$\gamma_2 | X_8 \underset{\sim}{\sim} \text{Gamma}(n_8 + a_8, b_8 + W_8) \tag{23}$$

To derive width of HPD intervals and Bayes estimates for parameters, we generate observations from the posterior distributions listed above. To obtain Bayesian estimation of MTSF and profit function, the above draws are put directly into the equations [2] & [6]. Using a squared error loss function, Bayesian estimates of parameters and reliability characteristics are derived from the sample means of the relevant drawings.

11. SIMULATION STUDY

To explore the behaviour of parameters, estimates and reliability aspects, a simulation study is carried out. The values of the Standard Error (SE)/Posterior Standard Error (PSE) and the width of confidence/HPD intervals are shown in table 1-6. Samples of sizes $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = n_7 = n_8 = 100$ were taken from the six investigated distributions while presuming various parameter values as shown in Tables 1-6. The number of iterations used is 10000. R software is used for the computations purpose.

Table 1: MTSF values for fixed $\lambda_1 = 0.05$ and varying α_1

α_1	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	13.438	10.101	0.0107	0.0078	10.024	0.00070	0.00051
0.2	5.188	5.075	0.0099	0.0074	5.021	0.00063	0.00046
0.3	3.689	3.420	0.0100	0.0073	3.353	0.00062	0.00046
0.4	2.986	2.565	0.0099	0.0074	2.520	0.00061	0.00044
0.5	2.565	2.091	0.0102	0.0074	2.020	0.00061	0.00044
0.6	2.281	1.756	0.0104	0.0078	1.686	0.00061	0.00045
0.7	2.075	1.518	0.0106	0.0078	1.448	0.00059	0.00044
0.8	1.918	1.343	0.0109	0.0079	1.269	0.00061	0.00045
0.9	1.795	1.204	0.0111	0.0082	1.131	0.00062	0.00045
1	1.696	1.123	0.0111	0.0083	1.020	0.00046	0.00044

Table 2: MTSF values for fixed $\lambda_1=0.45$ and varying α_1

α_1	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	16.785	10.257	0.0023	0.017	10.063	0.0016	0.0012
0.2	5.934	5.178	0.0016	0.011	5.043	0.0011	0.00082
0.3	4.163	3.477	0.0013	0.010	3.368	0.00094	0.00069
0.4	3.341	2.626	0.0012	0.0093	2.531	0.00085	0.00062
0.5	2.846	2.111	0.0012	0.0089	2.028	0.00078	0.00058
0.6	2.511	1.76	0.0011	0.0086	1.693	0.00076	0.00056
0.7	2.268	1.528	0.0011	0.0086	1.454	0.00074	0.00053
0.8	1.082	1.354	0.0011	0.0086	1.275	0.00072	0.00053
0.9	1.936	1.230	0.0011	0.0081	1.135	0.00070	0.00052
1	1.817	1.114	0.0011	0.0086	1.024	0.00071	0.00051

Table 3: MTSF values for fixed $\lambda_1=0.85$ and varying α_1

α_1	True MTSF	MLE.MTSF	SE	C.I	Bayes MTSF	PSE	HPD Interval
0.1	19.686	10.359	0.032	0.024	10.111	0.0026	0.0019
0.2	6.500	5.219	0.020	0.015	5.064	0.0016	0.0011
0.3	4.524	3.479	0.016	0.012	3.382	0.0012	0.00091
0.4	3.612	2.661	0.014	0.010	2.541	0.0010	0.00078
0.5	3.063	2.135	0.013	0.0098	2.037	0.00095	0.00071
0.6	2.691	1.804	0.012	0.0093	1.700	0.00091	0.00066
0.7	2.419	1.548	0.012	0.0091	1.460	0.00085	0.00062
0.8	2.211	1.377	0.011	0.0088	1.280	0.00081	0.00060
0.9	2.047	1.219	0.011	0.0088	1.140	0.00080	0.00059
1	1.913	1.104	0.011	0.0087	1.028	0.00077	0.00057

Table 4: Profit values for fixed $\lambda_1=0.05$ and varying α_1

α_1	True profit	MLE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	824.23	119.48	1.43	2.09	43.20	1.34	0.99
0.2	593.60	114.18	1.37	1.99	42.37	1.31	0.80
0.3	496.60	104.81	1.30	1.99	41.14	1.17	0.74
0.4	429.96	100.09	1.26	1.82	39.83	1.13	0.73
0.5	377.55	95.95	1.17	1.76	38.73	1.07	0.80
0.6	334.00	87.59	1.12	1.66	37.80	0.88	0.72
0.7	296.75	83.67	1.07	1.59	36.66	1.06	0.98
0.8	264.30	80.17	1.05	1.51	35.84	0.93	0.87
0.9	235.69	85.81	0.99	1.46	34.96	0.96	0.85
1	210.22	83.14	0.96	1.40	33.76	0.93	0.86

Table 5: Profit values for fixed $\lambda_1=0.45$ and varying α_1

α_1	True profit	MLE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	852.97	121.30	1.47	2.12	43.42	1.41	1.02
0.2	638.57	115.93	1.40	2.04	42.23	1.34	1.09
0.3	539.29	110.87	1.33	1.98	41.41	1.09	0.96
0.4	469.04	106.09	1.29	1.84	39.97	1.19	0.95
0.5	413.21	101.62	1.21	1.76	39.12	1.03	0.82
0.6	366.64	97.00	1.16	1.69	37.91	0.93	0.98
0.7	326.78	92.87	1.11	1.62	36.84	0.99	0.70
0.8	292.10	89.00	1.06	1.54	35.87	1.01	0.96
0.9	261.57	84.73	1.01	1.50	34.84	0.77	0.85
1	210.22	81.18	0.97	1.44	33.90	0.79	0.83

Table 6: Profit values for fixed $\lambda_1 = 0.85$ and varying α_1

α_1	True profit	MLE.Profit	SE	C.I	Bayes Profit	PSE	HPD Interval
0.1	871.88	123.18	1.49	2.20	43.45	1.32	0.99
0.2	671.99	117.34	1.45	2.10	42.33	1.26	0.75
0.3	572.73	112.39	1.35	2.02	41.31	1.05	0.90
0.4	500.62	107.48	1.29	1.93	40.07	1.18	0.90
0.5	442.57	103.13	1.24	1.82	39.12	1.009	0.82
0.6	393.82	98.24	1.19	1.74	38.12	1.008	0.77
0.7	351.93	94.19	1.13	1.66	36.93	1.06	0.87
0.8	215.41	89.78	1.07	1.60	35.84	1.05	0.74
0.9	283.24	85.91	1.03	1.53	35.26	0.87	0.70
1	254.65	82.00	0.96	1.45	34.02	0.85	0.84

12. GRAPHICAL STUDY

A graphical analysis of the system model provides a more insightful and vivid representation of system behaviour. So for more concrete study, we plot MTSF and Profit function wrt α_1 failure rate of unit A for different values of λ_1 repair rate of unit A as 0.05, 0.45 and 0.85. Here all other parameters are fixed $\alpha_2 = 0.9, \alpha_3 = 0.15, \lambda_2 = 0.35, \lambda_3 = 0.45, c = 0.7, \gamma_1 = 0.6$ and $\gamma_2 = 0.8$.

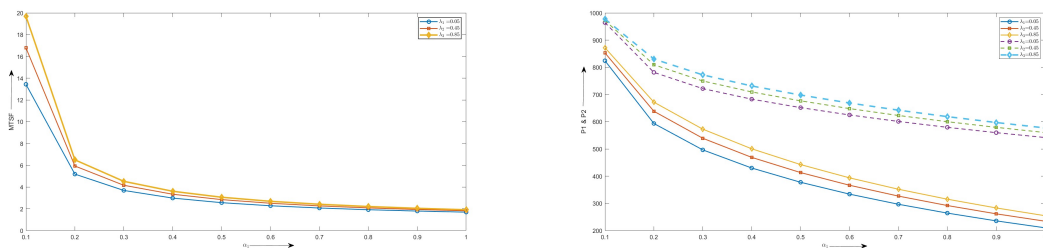


Figure 2: (a) Behaviour of MTSF wrt to α_1 for different values of λ_1 and (b) Behaviour of P_1 & P_2 wrt to α_1 for different values of λ_1

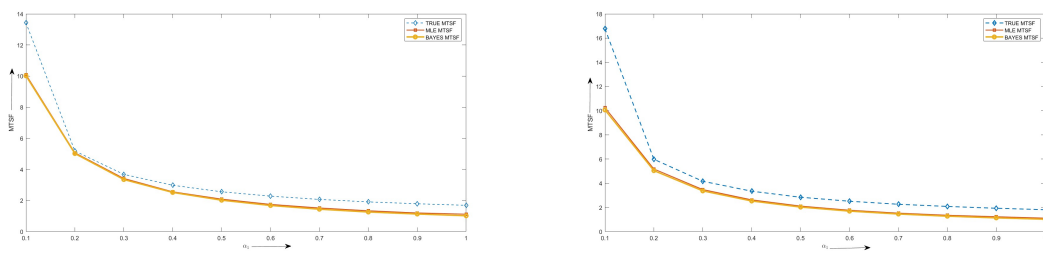


Figure 3: (a) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_1 for $\lambda_1=0.05$ and (b) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_1 for $\lambda_1=0.45$

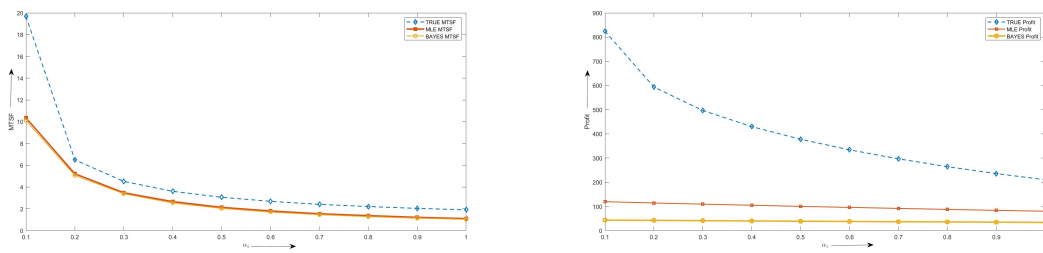


Figure 4: (a) Behaviour of True MTSF, MLE MTSF & Bayes MTSF wrt to α_1 for $\lambda_1=0.85$ and (b) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_1 for $\lambda_1=0.05$

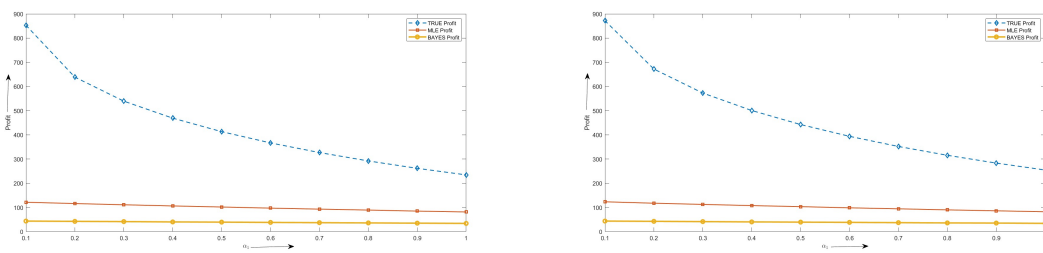


Figure 5: (a) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_1 for $\lambda_1=0.45$ and (b) Behaviour of True Profit, MLE Profit & Bayes Profit wrt to α_1 for $\lambda_1=0.85$

13. DISCUSSION AND CONCLUSION

1. Tables and figures exhibits that MTSF decreases as the failure rate α_1 increases, but increases as the repair rate λ_1 increases. The same trend is followed for the profit function.
2. Tables 1-6 indicate that for fixed and variable parameters, Bayes estimates of the MTSF and profit function perform better than MLEs in terms of SE as well as in terms of the width of the confidence intervals as they have lower PSE and the width of HPD intervals.
3. Based on the above discussions, we conclude that for estimating the MTSF and Profit function of the analyzed model, Bayes approach outperforms the Classical approach.

14. ACKNOWLEDGEMENT

A special thanks to Department of Science and Technology, Govt. of India for providing financial help in the form of INSPIRE fellowship.

REFERENCES

- [1] Sharma, Akshita and Kumar, Pawan (2019). Analysis of Reliability Measures of two Identical Unit System with One Switching Device and Imperfect Coverage *Reliability: Theory & Applications*, 14(1):44–52.
- [2] Trivedi, Kishor S(2002). Probability & Statistics with Reliability, Queuing and Computer Science Applications 2nd Edition John Wiley & Sons, New York.
- [3] Gupta, Rakesh and Kishan, Ram and Kumar, Pawan (1999). A two Non-Identical-Unit Parallel System with Correlated Lifetimes *International Journal of Systems Science*, 30(10):1123–1129.
- [4] Wang, Kuo-Hsiung and Chen, Yu-Ju(2009). Comparative Analysis of Availability Between three Systems with General Repair Times, Reboot Delay and Switching Failures *Applied Mathematics and Computation*, 215(1):384–394.
- [5] Pham, Hoang(1992). Reliability Analysis of a High Voltage System with Dependent Failures and Imperfect Coverage *Reliability Engineering & System Safety*, 37(1):25–28.
- [6] Ke, Jau-Chuan and Liu, Tzu-Hsin(2014). A Repairable System with Imperfect Coverage and Reboot *Applied Mathematics and Computation*, 246:148–158.
- [7] Kumar, Pankaj and Jain, Madhu(2020). Reliability Analysis of a Multi-Component Machining System with Service Interruption, Imperfect Coverage, and Reboot *Reliability Engineering & System Safety*, 202:106991.
- [8] Kadyan, MS and Malik, Suresh Chander and others(2020). Stochastic Analysis of a three Unit Non-Identical Repairable System with Simultaneous Working of Cold Standby Units *Journal of Reliability and Statistical Studies*, 385–400.
- [9] Saxena, Vashali and Gupta, Rakesh and Singh, Bhupendra(2023). Classical and Bayesian Stochastic Analysis of a two Unit Parallel System with Working and Rest Time of Repairman *Reliability: Theory & Applications*, 18(1(72))43–55.
- [10] Kishan, Ram and Jain, Divya(2012). A two Non-Identical Unit Standby System Model with Repair, Inspection and Post-Repair Under Classical and Bayesian Viewpoints *Journal of reliability and statistical studies*, 85–103.

RELIABILITY ESTIMATION OF STRESS-STRENGTH MODEL USING FUZZY DISTORTION FUNCTION UNDER UNCERTAINTY IN ENVIRONMENTAL FACTORS

K SRUTHI AND M KUMAR



National Institute of Technology, Calicut
mahesh@nitc.ac.in

Abstract

In the reliability estimation of stress-strength models, external factors such as temperature, humidity, etc. may influence the distribution of stress and strength random variables. In traditional reliability analysis, these external factors are accounted for by introducing a real-valued distortion function, which replaces the original distribution with a distorted one. However, it's important to note that the effect of these external factors is not always adequately represented by a single real-valued function. To address this issue, we propose the use of fuzzy numbers within the distortion function. In this paper, we introduce the concept of a "fuzzy distortion function" to incorporate the uncertainty stemming from external factors when estimating the reliability of stress-strength relationships. We present a methodology for estimating fuzzy reliability by employing this fuzzy distortion function. Through an illustrative example, we demonstrate how this approach to estimating fuzzy reliability offers a wider range of possibilities for system reliability and provides more comprehensive insights into the system's behaviour. Throughout our exploration, we have delved into the diverse properties inherent in fuzzy distortion functions. These properties highlight the versatility and adaptability of such functions in capturing uncertainty within data sets. Moreover, we have scrutinized several methods for constructing fuzzy distortion functions from pre-existing ones. By examining these methods, we gain valuable insights into how fuzzy distortion functions can be tailored to specific contexts and applications, thereby enhancing the accuracy and robustness of reliability analysis in complex systems. Additionally, in the conventional stress-strength model, reliability is determined without considering the uncertainty in the parameters of the distribution function. The drawback of existing methods in the literature is that they do not consider the uncertainty or fuzziness in the parameters of the distribution. Therefore, we estimate the system reliability in the presence of fuzzy parameters in the distribution function of corresponding random variables. The method we discuss in this paper provides a reliability estimate of the given system under realistic situations. A sensitivity analysis study is carried out to examine the behaviour of mean square errors (MSE) of estimated system reliability under various scenarios. It is observed that MSE can be significantly reduced by a suitable choice of parameters in the membership function of fuzzy parameters.

Keywords: Reliability, Fuzzy reliability, Distortion function, Fuzzy triangular number

1. INTRODUCTION

An et al. [1] used the universal generating function (UGF) approach to develop a discrete SSI model. And they handled strength and stress as discrete random variables. Considering a unilateral dependence of strength on stress found in some real-world circumstances. Huang et al. [2] provide a discrete SSI model with SDS based on a UGF approach. This model treats a

structure's SDS as a discrete random variable that, depending on the amplitude of the applied stress, has a different conditional probability mass function (pmf). In his study of reliability of stress-strength for a general coherent system, Eryilmaz [3] offered the exact formula as well as approximations and limitations. The estimation process for exponential stress-strength distributions was further illustrated by the author. In order to estimate the reliability of stress-strength model with multicomponent, Rao et al. [4] showed that the BLUE method of estimation shows the least MSE when compared to exact MLE, Method of Moments and TMMLE.

Kizilaslan [5] considered the system of multi-component with k statistically identical and independently distributed components of strength and each component subjected to a shared random stress. In order to determine the system's reliability in both known and unknown instances for the common second parameter λ , he used both classical and Bayesian methodologies. The system of multicomponent with k uniformly distributed and statistically independent strength components was explored by Kizilaslam [6]. Each element in the system was exposed to a common random stress. When the common scale parameter λ is known in some situations but not in others, the reliability of the system can be further assessed using frequentist and Bayesian approaches.

Dey et al. [7] compared the reliability of Bayes estimators and MLEs with respect to the mean squared errors and the average biases in their study of reliability of the stress-strength model of multicomponent for 238 two parameter Kumaraswamy distribution, then the distribution of strength and stress is the same. When both the stress and strength variations follow the same population, Rao et al. [8] investigated the reliability of stress-strength model with multicomponent for exponentiated Weibull distribution. Additionally, the predicted asymptotic confidence interval for the reliability of several components under stress. For the investigation of structural dependability using Copulas, Zhang et al. [9] presented a stress-strength time-varying correlation interference model.

By assuming that both the strength and stress variables follow a Chen distribution with a common shape parameter that may or may not be known, Tanmay et al. [10] obtained point and interval estimates of reliability of the multi-component stress-strength model of a s -out-of- j system using both Bayesian and classical approaches. The multi-component stress-strength reliability was evaluated by Amal et al. [11] based on the recorded data. When the stress and strength variables follow separate Weibull distributions with distinct scale parameters, the system's dependability is established. When samples are taken from distributions of stress and strength, and measurements are made in terms of upper record values, the reliability in MSS is evaluated using the maximum likelihood and Bayesian techniques of estimation. Amer et al. [12] assessed the reliability when the variables strength and stress are independent and follow the exponentiated Pareto distribution. The simple random sampling (SRS), median ranked set sampling (MRSS) and ranked set sampling (RSS) methods are used to calculate the maximum likelihood estimators in R. In four separate circumstances, the dependability estimate based on MRSS is taken into consideration. When the strength and stress variables are modeled by two separate but not identically distributed random variables from the generalized inverted exponential distributions, Amal et al. [13] assessed reliability of the stress-strength model. When evaluating the stress strength reliability estimator, MRSS is primarily used as opposed to RSS and SRS. A fresh addition to stress-strength models was made by Saber et al. [14]. The extended exponential distribution is used to apply the new model. The asymptotic distribution, the Bayesian estimation and the maximum likelihood estimator are derived. Shubham et al. [15] investigated both conventional and Bayesian techniques of reliability estimation of stress-strength model with multi-component and arrived at a maximum likelihood estimate of dependability. Additionally, the confidence intervals for asymptotic, boot-p, and boot-t data were built. Zhang et al. [16] investigated how well the multi-component stress-strength model, which includes one stress and two associated strength components from a parallel system, could predict dependability.

The rest of the paper is organized as follows. Section 2 gives basic definitions connected with fuzzy numbers and distortion functions. Section 3 introduces the concept of fuzzy distortion function. Section 3.1 describes with the estimation of the fuzzy reliability using fuzzy distortion function and a numerical illustration of this method is presented. Section 3.2 deals with some

interesting properties of distortion function. Section 3.3 deals with some basic methods for the construction of distortion function. Section 4 deals with the estimation of reliability of the stress strength model using weighted distributions. In Section 5, we present a sensitivity analysis study to check the behaviour of MSE of reliability estimate. Section 6 concludes with presenting the findings of this study.

2. SOME BASIC DEFINITIONS

In this section, we now provide some key definitions that are necessary to comprehend the findings in the sections that follow.

Definition 1. (see [17]) A fuzzy set D defined on a set S is a mapping from S to the unit interval $[0, 1]$, denoted by

$$D = \{(x, \mu_D(x)); \quad x \in S\}$$

or

$$D = \{x, \mu_D(x)\},$$

where $\mu_D(x)$ is the membership function of the set D .

Definition 2. (see [17]) A fuzzy set D defined on the real line \mathfrak{R} is convex if and only if $\forall x_1 \in S, \forall x_2 \in S$ and $\forall \lambda \in [0, 1]$, there holds

$$\mu_D(\lambda x_1 + (1 - \lambda)x_2) \geq \min(\mu_D(x_1), \mu_D(x_2)),$$

or equivalently, a fuzzy set is said to be convex if all of its cut sets are convex. If the \geq sign is replaced by $>$ sign, then we say that the fuzzy set is strictly convex.

Definition 3. (see [17]) For any $\alpha \in [0, 1]$, an α -cut set of D , denoted by D_α , is a classic set defined by

$$D_\alpha = \{x \in S, \mu_D(x) \geq \alpha\}.$$

Obviously, $D_{\alpha_1} \subseteq D_{\alpha_2}$ if $\alpha_1 \geq \alpha_2$.

Definition 4. (see [17]) If the membership function of fuzzy number D is determined by

$$\mu_D(x) = \begin{cases} 0 & ; x \leq a_1, \\ (x - a_1) / (a_2 - a_1) & ; a_1 \leq x \leq a_2, \\ (a_3 - x) / (a_3 - a_2) & ; a_2 \leq x \leq a_3, \\ 0 & ; x \geq a_3. \end{cases} \quad x, a_1, a_2, a_3, \in \mathfrak{R},$$

then D is referred to as a triangular fuzzy number, denoted $D = (a_1, a_2, a_3)$.

Suppose $D = (a_1, a_2, a_3)$. Then

$$D_\alpha = [a_1 + \alpha(a_2 - a_1), a_3 - \alpha(a_3 - a_2)].$$

Definition 5. (see [18]) A function $\nu(u)$ is called a distortion function if the following conditions hold:

- (i) $\nu(u)$ is a non-decreasing function on the interval $[0, 1]$,
- (ii) $\nu(0) = 0$ and $\nu(1) = 1$,
- (iii) except a finite number of points, $\varphi(u) = \frac{d}{du}\nu(u)$ exists on the interval $[0, 1]$.

Definition 6. (see [19]) Let V denote the strength random variable of the system and W denote the stress random variable. If V and W are independent with respective distribution functions G and F , then the traditional stress-strength reliability can be estimated as

$$R = P\{V > W\} = \iint_{v>w} dF(w)dG(v). \tag{1}$$

3. THE CONCEPT OF FUZZY DISTORTION FUNCTION

In traditional stress-strength model, we are not incorporating the uncertainty in the external factor. So it is necessary to incorporate the uncertainty in the external factor. Consider a traditional stress-strength model, where system reliability is estimated by using a stress-strength relation. The system's stress is represented by the random variable W , which has the cumulative distribution function F . With a cumulative distribution function of G , the random variable V represents the system's strength. Then, the reliability P of the stress-strength model of the system is given by $P = P(W < V)$. The drawback of this model is that the reliability estimate is unrealistic, since the normal working condition of the system is not considered. In other words, the uncertainty in the external factors is not considered. In general, reliability of the system is affected by environmental factors. To make it more clear, let us consider an example: think about creating a bridge in a city. Let W denote the weight stress of the bridge with a distribution of F and V represent the leg strength of a bridge with a distribution of G . With time, environmental elements including vibration, humidity, and high temperatures are exposed to the random strength and random stress. Now it is to be observed that the effect of external factor modelled by distortion function $\nu(\cdot)$ need not be a simple real valued function. In other words, the distortion force can not be considered as a constant force acting on the system at any given time. Hence, one may not obtain the realistic results. Hence, it is necessary to incorporate this vagueness in the external factors. It is not possible to represent the uncertainty in the external factor by a single real valued distortion function. Therefore, we let the distortion function can take fuzzy value also. In that concern we introduce the concept of fuzzy distortion function, which is more practicable. We are defining the fuzzy distortion function as follows,

Fuzzy distortion function

A function ν from $[0, 1]$ to set of fuzzy numbers in $[0, 1]$, is called fuzzy distortion function if the following conditions hold:

1. For fixed $\alpha \in [0, 1]$, the functions $A_\alpha(u)$ and $B_\alpha(u)$ are non decreasing where A_α and B_α are end points of the α -cut $\nu_\alpha(u) = [A_\alpha(u), B_\alpha(u)]$,
2. $\nu(0)$ is a fuzzy number with 0 having only membership value 1,
3. $\nu(1)$ is a fuzzy number with 1 having only membership value 1,
4. For fixed $\alpha \in [0, 1]$, the functions $A_\alpha(u)$ and $B_\alpha(u)$ are differentiable except a finite number of points on the interval $[0, 1]$.

Then we can have some interesting result for these fuzzy distortion functions.

Theorem The end points of α -cuts of fuzzy distortion functions are real distortion function. That is, for fixed $\alpha \in [0, 1]$, the above functions $A_\alpha(u)$ and $B_\alpha(u)$ are real distortion functions

Proof: For Fixed $\alpha \in [0, 1]$, the function $A_\alpha(u)$ is non decreasing. Since $\nu(0)$ is a fuzzy number with 0 having only membership value 1, it shows that $A_\alpha(0) = 0$. Similarly $\nu(1)$ is a fuzzy number with 1 having only membership value, shows that $A_\alpha(1) = 1$. Finally from the definition of fuzzy distortion function, it is clear that $A_\alpha(u)$ is differentiable With the exception of a few points on the range $[0, 1]$. Hence $A_\alpha(u)$ satisfy all the conditions for distortion function.

Similarly one can prove that $B_\alpha(u)$ is a distortion function. Since $\nu(0)$ is a fuzzy number with 0 having only membership value 1, it shows that $B_\alpha(0) = 0$. Similarly $\nu(1)$ is a fuzzy number with 1 having only membership value, shows that $B_\alpha(1) = 1$. Finally from the definition of fuzzy distortion function, it is clear that $B_\alpha(u)$ is differentiable with the exception of a few points on the range $[0, 1]$. Hence $B_\alpha(u)$ satisfy all the conditions for distortion function.

3.1. Estimation of fuzzy reliability using fuzzy distortion function

In traditional stress-strength model, we are not incorporating the uncertainty in the external factor. So it is necessary to incorporate the uncertainty in the external factor. It is not possible to represent the uncertainty in the external factor by a single real valued distortion function. In that

concern we are defining the fuzzy distortion function as follows, which is more reliable. Let the random variables V and W stand in for the system's strength with a cumulative distribution function of G and stress with a cumulative distribution function of F , respectively. Let $\nu(\cdot)$ be a fuzzy distortion function. Then fuzzy reliability of the system can be estimated as follows.

For fixed $\alpha \in [0, 1]$, we get two real valued function $A_\alpha(u)$ and $B_\alpha(u)$. Since both $A_\alpha(u)$ and $B_\alpha(u)$ are real distortion function.

Consider the function $A_\alpha(u)$, then we can estimate the system reliability using the distortion function $A_\alpha(u)$. Let $R_{\alpha,a}$ denote the reliability estimated using the function $A_\alpha(u)$ and is given by

$$R_{\alpha,a} = \int_0^1 A_\alpha \left(F \left(G^{-1}(u) \right) \right) dA_\alpha(u). \quad (2)$$

Let $R_{\alpha,b}$ denote the reliability estimated using the function $B_\alpha(u)$ and is given by

$$R_{\alpha,b} = \int_0^1 B_\alpha \left(F \left(G^{-1}(u) \right) \right) dB_\alpha(u). \quad (3)$$

We can estimate the α -cuts of fuzzy reliability of the system as follows

$$R_\alpha = [R_{\alpha,a}, R_{\alpha,b}]. \quad (4)$$

Similarly for each $\alpha \in [0, 1]$, we can estimate the α -cut of R .

3.1.1 Illustration

Let the strength of a bridge leg be represented by the random variable V with distribution G . The bridge's tension and weight are represented by the random variable W with distribution F . As time goes on, the random strength and stress are subjected to temperature-related external conditions. Suppose u^3, u^2 and V are the distortion function corresponding to the varying temperature. Let $\nu(u) = [u^3, u^2, u]$ be the distortion function. Then $A_\alpha(u)$ and $B_\alpha(u)$ can be estimated as follows

$$A_\alpha = u^3 + \alpha [u^2 - u^3] \quad (5)$$

and

$$B_\alpha = u - \alpha [u - u^2]. \quad (6)$$

The system's stress is represented by the random variable W , which has the cumulative distribution function $F(x) = 1 - e^{-\lambda_1 x}$. With a cumulative distribution function of $G(y) = 1 - e^{-\lambda_2 y}$, the random variable V represents the system's strength.

Then $A_{\alpha,a}$ and $B_{\alpha,a}$ can be estimated as follows

$$A_{\alpha,a} = \int_0^1 \left((1 - \alpha) \left(1 - (1 - u)^{\lambda_1/\lambda_2} \right)^3 + \alpha \left(1 - (1 - u)^{\lambda_1/\lambda_2} \right)^2 \right) \left(3(1 - \alpha)u^2 + 2\alpha u \right) du \quad (7)$$

and

$$B_{\alpha,b} = \int_0^1 \left((1 - \alpha) \left[1 - (1 - u)^{\lambda_1/\lambda_2} \right] + \alpha \left[1 - (1 - u)^{\lambda_1/\lambda_2} \right]^2 \right) \left((1 - \alpha) + 2\alpha u \right) du. \quad (8)$$

For each α , we can estimate the α -cut of R as

$$R_\alpha = [A_{\alpha,a}, B_{\alpha,b}]. \quad (9)$$

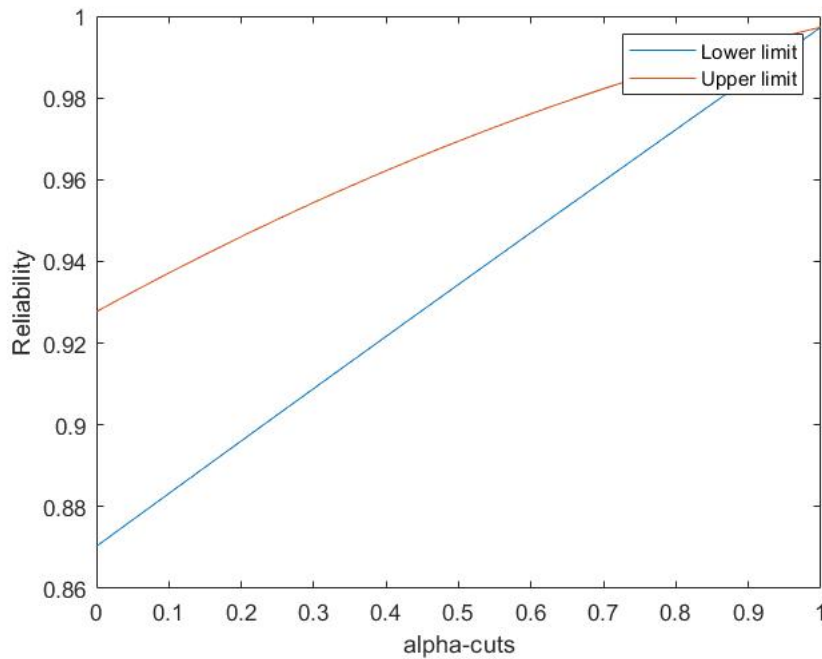


Figure 1: Fuzzy system reliability of for various values of α -cuts when $\lambda_1 = 0.33724$ and $\lambda_2 = 0.02628$.

3.2. Characterization and some properties of fuzzy distortion function

Result 1: The average of two distortion functions is again a distortion function.

Proof : Let $v_1(t)$ and $v_2(t)$ be two distortion function.

Define $f(t) = \frac{v_1(t)+v_2(t)}{2}$. Since both $v_1(t)$ and $v_2(t)$ are non-decreasing functions. Clearly $f(t)$ is a non-decreasing function.

Since $v_1(0) = 0$ and $v_2(0) = 0$, so we have $f(0) = 0$. Similarly $v_1(1) = 1$ and $v_2(1) = 1$, shows that $f(1) = 1$.

Since both $v_1(t)$ and $v_2(t)$ are differentiable with the exception of a few points, it holds for $f(t)$ also.

Result 2: The average of finite number of distortion functions is again a distortion function.

Proof: Let $v_1(t), v_2(t), \dots, v_n(t)$ be n distortion functions.

Define $f(t) = \frac{v_1(t)+v_2(t)+\dots+v_n(t)}{n}$. We have $v_1(t), v_2(t), \dots, v_n(t)$ are non-decreasing functions. Then $f(t)$ is also a non-decreasing function.

We have $v_1(0) = 0, v_2(0) = 0, \dots, v_n(0) = 0$, so we have $f(0) = 0$. Similarly $v_1(1) = 1, v_2(1) = 1, \dots, v_n(1) = 1$, imply $f(1) = 1$.

Since all the functions $v_1(t), v_2(t), \dots, v_n(t)$ are differentiable with the exception of a limited number of points, then the function $f(t)$ is also differentiable with the exception of a few points.

Result 3: The product of two distortion functions is again a distortion function.

Proof : Let $v_1(t)$ and $v_2(t)$ be two distortion function.

Define $f(t) = v_1(t).v_2(t)$. Since both $v_1(t)$ and $v_2(t)$ are non-decreasing functions. Clearly $f(t)$ is a non-decreasing function.

Since $v_1(0) = 0$ and $v_2(0) = 0$, so we have $f(0) = 0$. Similarly $v_1(1) = 1$ and $v_2(1) = 1$, imply $f(1) = 1$.

Since both $v_1(t)$ and $v_2(t)$ are differentiable with the exception of a few points, it holds for $f(t)$ also.

Result 4: The product of finite number distortion functions is again a distortion function.

Proof: Let $v_1(t), v_2(t), \dots, v_n(t)$ be n distortion functions.

Define $f(t) = v_1(t) \cdot v_2(t) \dots v_n(t)$. We have $v_1(t), v_2(t), \dots, v_n(t)$ are non-decreasing functions. Then $f(t)$ is also a non-decreasing function.

We have $v_1(0) = 0, v_2(0) = 0, \dots, v_n(0) = 0$, so we have $f(0) = 0$. Similarly $v_1(1) = 1, v_2(1) = 1, \dots, v_n(1) = 1$, imply $f(1) = 1$.

Since all the functions $v_1(t), v_2(t), \dots, v_n(t)$ are differentiable with the exception of a limited number of points, hence it holds for $f(t)$ also.

3.3. Methods for construction of fuzzy distortion function

Result 1 Let $v_1(t), v_2(t)$ and $v_3(t)$ are three distortion function with $v_1(t) \leq v_2(t) \leq v_3(t)$. Then $v(t)$ defined by $v(t) = [v_1(t), v_2(t), v_3(t)]$ is a fuzzy distortion function.

Proof: Define $v(t) = [v_1(t), v_2(t), v_3(t)]$. Then

$$A_\alpha(t) = v_1(t) + \alpha(v_2(t) - v_1(t)) \tag{10}$$

and

$$B_\alpha(t) = v_3(t) - \alpha(v_3(t) - v_2(t)). \tag{11}$$

First we have to prove that $A_\alpha(t)$ and $B_\alpha(t)$ are non-decreasing. For that we made a rearrange as $A_\alpha(t) = (1 - \alpha)v_1(t) + \alpha v_2(t)$ and $B_\alpha(t) = (1 - \alpha)v_2(t) + \alpha v_3(t)$.

Since $v_1(t), v_2(t)$ and $v_3(t)$ are non-decreasing functions and $1 - \alpha > 0$. Both $A_\alpha(t)$ and $B_\alpha(t)$ are non-decreasing.

We have $v_1(0) = 0, v_2(0) = 0$ and $v_3(0) = 0$.

Then

$$A_\alpha(0) = (1 - \alpha)v_1(0) + \alpha(v_2(0)) = (1 - \alpha)0 + \alpha 0 = 0, \tag{12}$$

and

$$B_\alpha(0) = (1 - \alpha)v_2(0) + \alpha(v_3(0)) = (1 - \alpha)0 + \alpha 0 = 0. \tag{13}$$

Similarly $v_1(1) = 1, v_2(1) = 1$ and $v_3(1) = 1$.

Then

$$A_\alpha(1) = (1 - \alpha)v_1(1) + \alpha(v_2(1)) = (1 - \alpha)1 + \alpha 1 = 1 \tag{14}$$

and

$$B_\alpha(1) = (1 - \alpha)v_2(1) + \alpha(v_3(1)) = (1 - \alpha)1 + \alpha 1 = 1. \tag{15}$$

Since the three functions $v_1(t), v_2(t)$ and $v_3(t)$ are differentiable with the exception of a limited number of points, $A_\alpha(t)$ and $B_\alpha(t)$ are also differentiable with with the exception of a few points. Hence the function $v(t)$ satisfy all the conditions for fuzzy distortion function.

Result 2 Let $v_1(t), v_2(t), v_3(t)$ and $v_4(t)$ are four distortion function with $v_1(t) \leq v_2(t) \leq v_3(t) \leq v_4(t)$. Then $v(t)$ defined by $v(t) = [v_1(t), v_2(t), v_3(t), v_4(t)]$ is a fuzzy distortion function

Proof: Define $v(t) = [v_1(t), v_2(t), v_3(t), v_4(t)]$. Then

$$A_\alpha(t) = v_1(t) + \alpha(v_2(t) - v_1(t)) \tag{16}$$

and

$$B_\alpha(t) = v_4(t) - \alpha(v_4(t) - v_3(t)). \tag{17}$$

First we will prove that $A_\alpha(t)$ and $B_\alpha(t)$ are non-decreasing. For that we made a rearrange as $A_\alpha(t) = (1 - \alpha)v_1(t) + \alpha v_2(t)$ and $B_\alpha(t) = (1 - \alpha)v_4(t) + \alpha v_3(t)$.

Since $v_1(t), v_2(t), v_3(t)$ and $v_4(t)$ are non-decreasing functions and $1 - \alpha > 0$. Both $A_\alpha(t)$ and

$B_\alpha(t)$ are non-decreasing.

We have $\nu_1(0) = 0, \nu_2(0) = 0, \nu_3(0) = 0$ and $\nu_4(0) = 0$,

Then

$$A_\alpha(0) = (1 - \alpha)\nu_1(0) + \alpha(\nu_2(0)) = (1 - \alpha)0 + \alpha 0 = 0, \tag{18}$$

and

$$B_\alpha(0) = (1 - \alpha)\nu_4(0) + \alpha\nu_3(0) = (1 - \alpha)0 + \alpha 0 = 0. \tag{19}$$

Similarly $\nu_1(1) = 1, \nu_2(1) = 1, \nu_3(1) = 1$ and $\nu_4(1) = 1$.

Then

$$A_\alpha(1) = (1 - \alpha)\nu_1(1) + \alpha(\nu_2(1)) = (1 - \alpha)1 + \alpha 1 = 1 \tag{20}$$

and

$$B_\alpha(1) = (1 - \alpha)\nu_4(1) + \alpha\nu_3(1) = (1 - \alpha)1 + \alpha 1 = 1. \tag{21}$$

Since the four functions $\nu_1(t), \nu_2(t), \nu_3(t)$ and $\nu_4(t)$ are differentiable except a finite number of points, $A_\alpha(t)$ and $B_\alpha(t)$ are also differentiable with the exception of a few points.

It is proved that the function $\nu(t)$ satisfy all the conditions for fuzzy distortion function.

4. ESTIMATION OF STRESS-STRENGTH RELIABILITY USING THE WEIGHTED PROBABILITY DENSITY FUNCTION

Consider a traditional stress-strength model, where system reliability is estimated by using a stress-strength relation. Let the random variable Y represent the strength of the system with cumulative distribution function G and the random variable X represent the stress of the system with cumulative distribution function F . Then, the stress-strength reliability P of the system is given by $P = P(X < Y)$. The drawback of this model is that the reliability estimate is unrealistic, since the normal working condition of the system is not considered. In other words, the fuzziness of parameters of the distribution is not taken into account, whereas, on the other hand, state of the system is highly dependent on the state of the parameters in the lifetime distribution. This is due to the fact that there exists an uncertainty in the parameters of distribution function. Hence, we incorporate this uncertainty factor by suitably modifying the density function and obtain the weighted probability density function and which can be used to get reliability estimate of the given system. Hence, we proceed as follows:

Let the random variable X have the probability density function $f(x, \theta)$. For the function $H(\theta)$, $\theta \in \Theta$, where Θ is the domain of definition of the parameter θ , the weighted probability density function of X is defined by

$$f(x) = \int_{\Theta} H^*(\theta) f(x, \theta) d\theta, \tag{22}$$

where $H^*(\theta)$ is called the pseudo-membership function and is defined by

$$H^*(\theta) = \frac{H(\theta)}{\int_{\Theta} H(\theta) d\theta}. \tag{23}$$

The definition and construction of above membership function is explained in detail by authors in [20].

Now let X be the stress random variable having the probability density function $f(x, \theta_1)$, $\theta_1 \in \Theta_1$, where Θ_1 is the domain of definition of the parameter θ_1 . Then the weighted probability density function (*wpdf*) of X (see, [20] for further information) is given by

$$f^*(x) = \int_{\theta_1 \in \Theta_1} H^*(\theta_1) \cdot f(x, \theta_1) d\theta_1 \tag{24}$$

Let Y be the strength random variable having the probability density function $g(x, \theta_2)$, $\theta_2 \in \Theta_2$, where Θ_2 is the domain of definition of the parameter θ_2 . Then the weighted probability density function of Y , similar to the definition of *wpdf* of X , is given by

$$g^*(x) = \int_{\theta_2 \in \Theta_2} H^*(\theta_2) \cdot g(x, \theta_2) d\theta_2 \tag{25}$$

Let X and Y have exponential distribution with parameters θ_1 and θ_2 respectively. Then the stress random variable X has probability density function $f(x) = \theta_1 e^{-\theta_1 x}$, $\theta_1 > 0$ and the strength random variable Y has probability density function $g(y) = \theta_2 e^{-\theta_2 y}$, $\theta_2 > 0$.

Let $H_1(\theta_1)$ and $H_2(\theta_2)$ represent the membership function for the parameters θ_1 and θ_2 respectively. Then we may write (See, [21])

$$H_1(\theta_1) = \begin{cases} \frac{1+\cos\left[\frac{a\pi(\theta_1-\frac{1}{b})}{2}\right]}{2} & \text{if } \frac{1}{b} - \frac{1}{a} \leq \theta_1 \leq \frac{1}{b} + \frac{1}{a} \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

$$H_2(\theta_2) = \begin{cases} \frac{1+\cos\left[\frac{c\pi(\theta_2-\frac{1}{d})}{2}\right]}{2} & \text{if } \frac{1}{d} - \frac{1}{c} \leq \theta_2 \leq \frac{1}{d} + \frac{1}{c} \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

It is easy to check from the definition of pseudo-membership function given in equation (23) that

$$H_1^*(\theta_1) = \frac{a}{2} \begin{cases} \frac{1+\cos\left[\frac{a\pi(\theta_1-\frac{1}{b})}{2}\right]}{2} & \text{if } \frac{1}{b} - \frac{1}{a} \leq \theta_1 \leq \frac{1}{b} + \frac{1}{a} \\ 0, & \text{otherwise} \end{cases} \quad (28)$$

$$H_2^*(\theta_2) = \frac{c}{2} \begin{cases} \frac{1+\cos\left[\frac{c\pi(\theta_2-\frac{1}{d})}{2}\right]}{2} & \text{if } \frac{1}{d} - \frac{1}{c} \leq \theta_2 \leq \frac{1}{d} + \frac{1}{c} \\ 0, & \text{otherwise} \end{cases} \quad (29)$$

Then reliability of the system under stress-strength model is given by

$$\begin{aligned} R^* &= P(X^* < Y^*) \\ &= \iint_{x < y} f^*(x)g^*(y)dx dy \\ &= \int_0^\infty \int_0^y f^*(x)g^*(y)dx dy \\ &= \int_0^\infty \int_0^y \left[\int_{\theta_1 \in \Theta_1} H_1^*(\theta_1) f(x, \theta_1) d\theta_1 \int_{\theta_2 \in \Theta_2} H_2^*(\theta_2) g(y, \theta_2) d\theta_2 \right] dx dy \\ &= \int_{\theta_1 \in \Theta_1} \int_{\theta_2 \in \Theta_2} H_1^*(\theta_1) H_2^*(\theta_2) \left[\int_0^\infty \int_0^y f(x, \theta_1) g(y, \theta_2) dx dy \right] d\theta_2 d\theta_1 \\ &= \int_{\theta_1 \in \Theta_1} \int_{\theta_2 \in \Theta_2} \frac{ac}{4} \left[1 + \cos\left(a\pi\left(\theta_1 - \frac{1}{b}\right)\right) \right] \left[1 + \cos\left(c\pi\left(\theta_2 - \frac{1}{d}\right)\right) \right] \frac{\theta_1}{\theta_1 + \theta_2} d\theta_2 d\theta_1, \end{aligned} \quad (30)$$

where the parameters θ_1 and θ_2 are non-negative. Since the closed form expression for indefinite integrals do not exist in (30), we resort to evaluate the integrals using numerical integration. The integrals are computed using two-dimensional quadrature method with the help of the operator quad2d in MATLAB for numerical integration. The following section gives the details of numerical results.

4.1. Numerical Results

In this section, we present some numerical results to illustrate the reliability of the system via equation (30) for various choices of the parameters, namely, a , b , c , and d in equation (28) and (29), which are the part of the integral in (30). The Table 1 illustrates the results obtained in Section 4.

Table 1 Reliability estimation using the weighted probability density function

<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	Reliability
9.5	9.3	450.0	444.0	0.9758
9.6	9.0	452.0	451.0	0.9774
12.0	10.0	456.0	450.0	0.9759
11.0	9.0	456.0	454.0	0.9785
10.0	8.0	457.0	454.0	0.9810
11.0	7.0	456.0	455.0	0.9840
19.0	9.0	456.0	455.0	0.9816
10.0	8.0	456.0	454.0	0.9810

It is observed from the computational experience that for the exponential lifetime distribution of the system, the fuzziness of the parameters modelled using membership function has significant advantage in describing the vagueness in the system parameters. Note that the above results are obtained from reliability equation given in (30), where the closed form expression of reliability does not exist, and a numerical integration is carried out. Hence, the choice of membership function works well under the situation that the exact expression for reliability is not possible to compute. This is evident from the examples computed in the above Table 1, where the minimum reliability obtained is about 97%. The reliability of the system without using the weighted probability density function with $\theta_1 = 0.00227$ and $\theta_2 = 0.00447$ is estimated as 0.3368. But from Table 1, it is observed that when we use weighted probability density function with parameters $a = 9.5, b = 9.3, c = 450.0$ and $d = 444.0$, the reliability of the system is estimated as 0.9758. It is observed that, when we use the weighted probability density function, the reliability of the system is increased by 60%.

5. SENSITIVITY ANALYSIS

In previous sections, several reliability estimates developed by incorporating the fuzziness in the data involve number of parameters which come from either membership functions or from distortion functions applied for estimating the stress-strength reliability of the system under consideration. Therefore it is necessary to study the worthiness of these reliabilities estimated in terms of their mean square errors (MSE). Hence this section discuss MSE of the reliability estimate developed in Section 4.

We study via numerical computation, how MSE is sensitive to changes in the parameters a, b, c and d of the membership function of parameters of distribution functions. For illustration, we let $\theta_1 = 0.2128$ and $\theta_2 = 0.0045$. First, we vary one parameter keeping the remaining parameters fixed. There will be four different cases. Figure 2 shows the variation of MSE when one of the parameters is varied while keeping other parameters fixed. In Figure 2, for the fixed combination $b = 1, c = 30, d = 10$ (see, Graph (i)), and varying the parameter a , it can be observed that there is a sudden decrease in MSE. Similar observation is true for set of $a = 10, b = 9, c = 450$ and for varying d (see, graph (iv)). Note that from Graph (ii) and (iii) we see that the MSE is showing increasing trend.

Next, study the changes in MSE by varying two parameters while keeping any two of the four remaining parameters held fixed. Figure 3 illustrate one such case for the fixed combination of $b = 10, c = 60$ while varying the remaining parameters. It is observed that there is a gradual decrease in MSE.

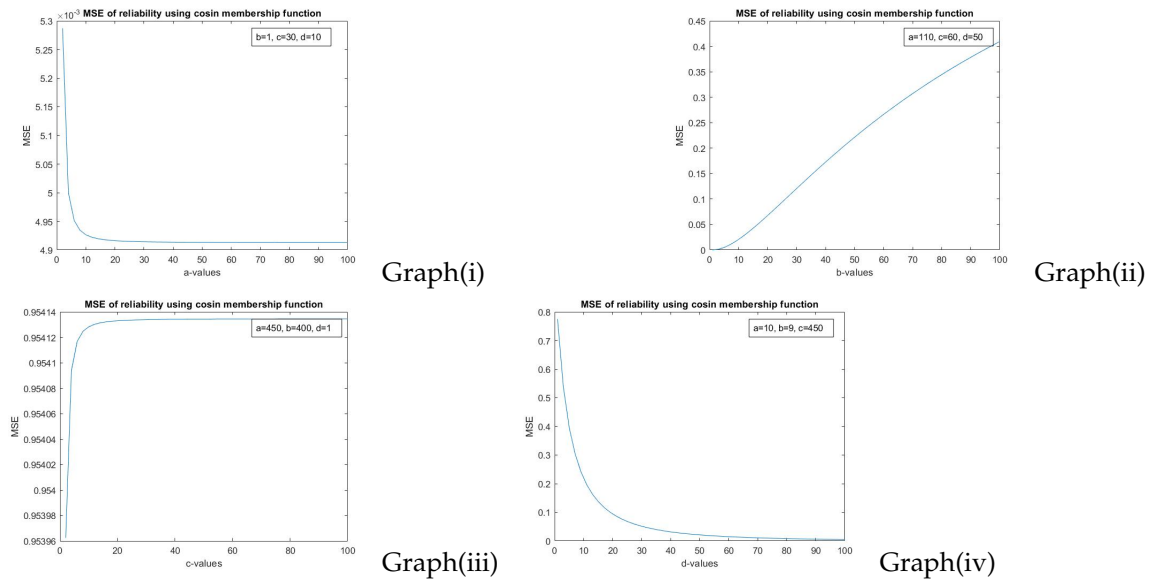


Figure 2: MSE of reliability estimator under membership function, when one parameter is varied and remaining parameters are held fixed.

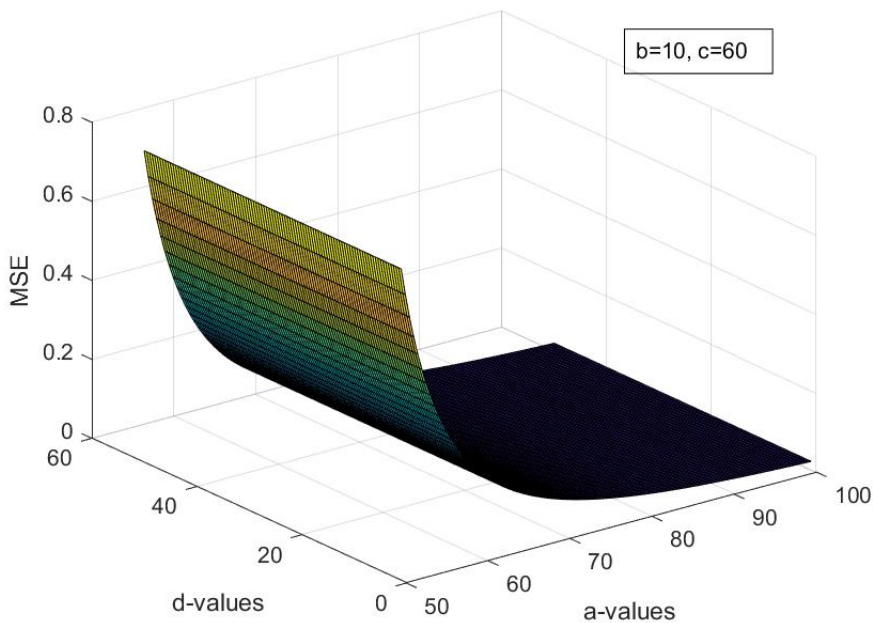


Figure 3 MSE of reliability using membership function when b and c are fixed.

Finally, from this sensitivity analysis study, it is noted that the accuracy of reliability estimate depends upon the choice of the parameters in the membership function, which can again depends upon availability of the type of data. Further, observe that the reliability estimates obtained are based upon numerical integration, since the closed form of expression does not exist. However, we strongly believe that the choice of membership in modelling vagueness play very important role in obtaining fuzzy reliability estimate.

6. CONCLUSIONS

In the conventional stress-strength model, reliability is determined without considering the variability in environmental factors. In our research, we introduced the notion of a fuzzy distortion function to account for the uncertainty in these environmental factors. Our approach involved assessing the reliability of the stress-strength model by employing the fuzzy distortion function. We illustrated this method with an example and also examined various characteristics of the distortion function. Additionally, we derived techniques for constructing a fuzzy distortion function based on existing actual distortion functions. The drawbacks of existing methods in the literature are that it does not consider the uncertainty or fuzziness in data and nature to estimate the system reliability under realistic situations. But in this work system reliability is estimated using the weighted probability density function by incorporating the fuzziness in data, which is practical and realistic. Finally, a sensitivity analysis study of MSE of reliability estimate obtained using cosine membership function is presented. It is observed that MSE as a function of parameters of membership function of a fuzzy parameter, can be minimized significantly by careful choice of parameters of membership functions.

REFERENCES

- [1] An, Z. W., Huang, H. Z. and Liu, Y. (2008). A discrete stress strength interference model based on universal generating function. *Reliability Engineering and System Safety*, 93:1485–1490.
- [2] Huang, H. Z. and An, Z. W. (2008). A discrete stress-strength interference model with stress dependent strength. *IEEE Transactions on Reliability*, 58:118–122.
- [3] Eryilmaz, S. (2010). On system reliability in stress strength setup. *Statistics and Probability Letters*, 80:834–839.
- [4] Rao, G. S. (2012). Estimation of system reliability for log logistic distribution. *Journal of Atoms and Molecules*, 2:182.
- [5] Kizilaslan, F. (2018). Classical and bayesian estimation of reliability in a multicomponent stress strength model based on a general class of inverse exponentiated distributions. *Statistical Papers*, 59:1161–1192.
- [6] Kizilaslan, F. (2017). Classical and bayesian estimation of reliability in a multicomponent stress-strength model based on the proportional reversed hazard rate mode. *Mathematics and Computers in Simulation*, 136:36–62.
- [7] Dey, S., Mazucheli, J. and Anis, M. (2017). Estimation of reliability of multicomponent stress-strength for a kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, 46:1560–1572.
- [8] Srinivasa Rao, G., Aslam, M. and Arif, O. H. (2017). Estimation of reliability in multicomponent stress-strength based on two parameter exponentiated weibull distribution. *Communications in Statistics-Theory and Methods*, 46:7495–7502.
- [9] Zhang, J., Ma, X. and Zhao, Y. (2017). A stress-strength time-varying correlation interference model for structural reliability analysis using copulas. *IEEE Transactions on Reliability*, 66:351–365.
- [10] Kayal, T., Tripathi, Y. M., Dey, S. and Wu, S. J. (2020). On estimating the reliability in a multicomponent stress-strength model based on chen distribution. *Communications in Statistics- Theory and Methods*, 49:2429–2447.
- [11] Hassan, A. S., Nagy, H. F., Muhammed, H. Z. and Saad, M. S. (2020). Estimation of multicomponent stress-strength reliability following weibull distribution based on upper record values. *Journal of Taibah University for Science* 14:244–253.
- [12] Al-Omari, A. I., Almanjahie, I. M., Hassan, A. S. and Nagy, H. F. (2020). Estimation of the stress- strength reliability for exponentiated pareto distribution using median and ranked set sampling methods. *CMC-Computers, Materials and Continua* 64:835–857.

- [13] Hassan, A. S., Al-Omari, A. and Nagy, H. F. (2021). Stress-strength reliability for the generalized inverted exponential distribution using mrss. *Iranian Journal of Science and Technology, Transactions A: Science*, 45:641–659.
- [14] Saber, M. M., Mohie El-Din, M. M. and Yousof, H. M. (2021). Reliability estimation for the remained stress-strength model under the generalized exponential lifetime distribution. *Journal of Probability and Statistics*, 2021:1–10.
- [15] Saini, S., Tomer, S. and Garg, R. (2022). On the reliability estimation of multicomponent stress- strength model for burr xii distribution using progressively first-failure censored samples. *Journal of Statistical Computation and Simulation*, 92:667–704.
- [16] Zhang, L., Xu, A., An, L. and Li, M. (2022). Bayesian inference of system reliability for multicomponent stress-strength model under marshall-olkin weibull distribution. *Systems*, 10:196.
- [17] Zimmermann, H. J. Fuzzy set theory and its applications, Springer Science and Business Media, 2011.
- [18] Pakdaman, Z. and Ahmadi, J. (2018). Some results on the stress-strength reliability under the distortion functions. *International Journal of Reliability, Quality and Safety Engineering*, 25:1850028.
- [19] Eryilmaz, S. and Tutuncu, G. Y. (2015). Stress strength reliability in the presence of fuzziness. *Journal of computational and applied mathematics*, 282:262–267.
- [20] Jamkhaneh, E. B. and Gildeh, B. S. (2010). Sequential sampling plan by variable with fuzzy parameters. *J. Math. Comput. Sci*, 1:392-401.
- [21] Ross, T. J. Fuzzy logic with engineering applications, John Wiley and Sons, 2005.

ANALYTICAL AND COMPUTATIONAL ASPECTS OF A MULTI-SERVER QUEUE WITH IMPATIENCE UNDER DIFFERENTIATED WORKING VACATIONS POLICY

¹AIMEN DEHIMI, ²MOHAMED BOUALEM, ³AMINA ANGELIKA BOUCHENTOUF,
⁴SOFIANE ZIANI, ⁵LOUIZA BERDJOUJ

¹University of Bejaia, Faculty of Exact Sciences, Applied Mathematics Laboratory, 06000 Bejaia, Algeria

²University of Bejaia, Faculty of Technology, Research Unit LaMOS, 06000 Bejaia, Algeria

³Laboratory of Mathematics, Djillali Liabes University of Sidi Bel Abbes, 22000 Sidi Bel Abbes, Algeria

⁴University of Bejaia, Research Unit LaMOS, 06000 Bejaia, Algeria

⁵University of Bejaia, Faculty of Exact Sciences, Research Unit LaMOS, 06000 Bejaia, Algeria

¹aimen.dehimi@univ-bejaia.dz, ²mohammed.boualem@univ-bejaia.dz, ³bouchentouf_amina@yahoo.fr

⁴sofiane.ziani@univ-bejaia.dz, ⁵louiza.berdjoudj@univ-bejaia.dz

Abstract

A multi-server queueing system with synchronous differentiated working vacation policy, Bernoulli schedule vacation interruption, and customer impatience (balking and reneging) is studied. The system consists of c servers and a finite capacity N , where customers arrive according to a Poisson process and are served in the chronological order of their arrival. When the system becomes empty, servers wait for a random duration before entering a type-1 working vacation, during which service is provided at a reduced rate. If customers are present in the system at the moment of service achievement during this period, the vacation is interrupted. With a certain probability, servers return to the regular busy period; otherwise, they continue the working vacation. Upon completion of the working vacation, if the system is still empty, servers can take another working vacation of shorter duration, named type-2 working vacation; otherwise, they switch to the regular busy period. Customer impatience is considered during both the normal busy period and working vacations. A recursive analysis method is used to find the steady-state probabilities of the system. Then, some important performance measures are obtained. Furthermore, an optimal operational policy for the model is developed to minimize the total expected cost. The Grey Wolf Optimization (GWO) meta-heuristic approach is employed to determine the optimal service rates for both working vacations and normal busy periods. Finally, several numerical examples are provided to validate and support the theoretical findings.

Keywords: Multi-server queue, differentiated vacations, impatience, GWO algorithm, cost optimization.

1. INTRODUCTION

Queueing models have gained considerable attention due to their significance in shaping and evaluating telecommunication systems, computer systems, and production management [8, 22, 23].

The concept of a server vacation queue has garnered extensive research attention, primarily due to its unique characteristic of allowing the server to utilize idle time for various tasks, such as

maintenance, service industries, production and manufacturing systems, or just taking a break [7]. For example, the growing utilization of wireless cellular networks has led to a substantial surge in energy consumption. To address this issue and promote the development of energy-efficient wireless cellular networks, researchers have introduced the concept of hibernation or sleeping for base stations (BS) during periods of inactivity. This approach is akin to the concept of a server going on vacation, where the BS temporarily reduces its power consumption when there are no active users in the network. In the classical server vacation queue, a server temporarily ceases its service during a designated vacation period [9, 10, 21]. It is important to note that some systems are designed with the presence of an alternate server that operates at a different, often lower, service rate when the primary server takes a vacation. Such a system is commonly referred to as a working vacation queue. In most working vacation policies, the server typically returns to its regular service rate after the vacation period ends, but only if there are customers waiting in the system. The idea of a working vacation was initially introduced by [24], where they proposed that the server does not entirely cease its operations during a vacation but continues providing service to the queueing system at a reduced rate. This concept has paved the way for various working vacation policies, enhancing the flexibility and efficiency of queueing system designs. These models have been discussed by different authors [3, 4, 12, 16, 26].

In numerous practical scenarios involving congestion, there are occasions when urgent events take place during vacation. As a result, the servers must interrupt their vacation and resume work instead of utilizing the remaining vacation time. Otherwise, such a situation incurs a substantial cost in terms of waiting customers. The concept of vacation interruption was initially introduced by [14]. Subsequently, [15] conducted a study on a $GI/M/1$ queue utilizing a supplementary variable method. Further discussions on an $M/PH/1$ queue, considering working vacations and vacation interruption, can be found in [2]. For more in-depth studies, additional references include [11, 13, 17, 18], and references therein.

In the past two decades, there has been a significant focus on the subject of impatient customers within queueing theory. This research area has proven to be intriguing and challenging, particularly in the context of globalization, hospital emergency rooms handling critical patients, and other relevant domains. As a result, the topic of queueing models with server vacations and impatient customers has garnered significant attention in the literature [1, 5, 8, 20, 25].

The main aim of this work is to conduct an analytical and optimization analysis of a finite capacity queue with multi-server and impatient customers (balking and reneging behaviors), incorporating vacation interruption and differentiated working vacations. The suggested queueing model presents promising applications across diverse sectors, including call centers, telecommunications and manufacturing, where servers can experience periods of downtime. In this research, the steady-state probabilities of queue length when servers are in working vacations period (type-1 and type-2), and in normal busy period are investigated using the recursive analysis approach. Several important performance measures are derived from these probabilities. Optimization in queueing systems is crucial in practical applications. In this study, the optimization problem tackled is complex and challenging, as the objective function is nonlinear on the service rates. To address this issue, the GWO algorithm is employed to determine the optimal service rates for both working vacations and normal busy periods, aiming to minimize the expected total cost. The GWO algorithm is known for its high performance in both unconstrained and constrained problems [19]. It has shown competitive results compared to well-established heuristics in swarm intelligence. Notably, the application of the GWO algorithm in queueing theory is relatively scarce in the existing literature. The present work can be considered as an extension to the research in reported [6], where the steady-state distributions were investigated in the case of a single server. By applying the GWO algorithm and considering the multi-server case, this paper contributes to the understanding and analysis of the considered model. Finally, numerical examples are presented to evaluate the behavior and performance of the proposed queueing system. These numerical results provide insights and support our findings.

The structure of the paper is as follows: Section 2 provides a detailed description of the queueing model being studied. In Section 3, we derive the steady-state distributions of queue

sizes during different server periods, including working vacation and normal busy periods. Section 4 gives explicit formulas for different performance measures of the queueing model. Moving on to Section 5, we analyze the effect of various system parameters on the performance measures through graphics. In Section 6, we address an optimization problem related to service rates using the GWO algorithm and present numerical results. Finally, section 7 gives a general conclusion and perspectives.

2. OVERVIEW AND ANALYSIS OF THE PROPOSED FRAMEWORK

We investigate an $M/M/c/K$ queue with impatience, operating under the differentiated working vacations along with vacation interruption. The fundamental assumptions underpinning this queueing system are outlined as:

- Customers enter the system in line with a Poisson process characterized by a rate of α .
- The time during a normal busy period of each server follows an exponential distribution and is denoted by service rate μ_1 .
- Customers are served in accordance with FCFS (First-Come-First-Served) discipline and the capacity of the system is considered to be finite, say K .
- The time during working vacations of each server follows an exponential distribution and is denoted by service rate μ_2 ($\mu_2 < \mu_1$).
- The queueing system under consideration involves multiple servers, denoted by c , when the system has no customers the servers wait for a random duration of time before leaving collectively for type-1 working vacation. Subsequently, When the servers return from their working vacation and find the system non-empty, they change their service rate from μ_2 to μ_1 and a normal busy period starts. If the servers return to find an empty queue, they immediately leave for another working vacation.
- The waiting time for the servers follows an exponential distribution with rate Δ .
- Following the completion of the waiting time duration, they begin an initial type-1 working vacation exponentially distributed with parameter Φ_1 . Once they return from the initial type-1 working vacation, if there are no customers in the queue, they transition to type-2 vacation which follows an exponential distributions characterized by parameter Φ_2 . Otherwise, they return to the normal busy period and start serving customers in the queue.
- Upon a customer's arrival during the vacation period, within this phase. Upon completing a service, if there are customers in the queue, the servers follow the Bernoulli distribution. They may opt to interrupt the vacation and move to the normal busy period, a choice determined by the probability denoted as β' . Alternatively, the servers may choose to continue the vacation, a decision made with the complementary probability $\beta = 1 - \beta'$. It is crucial to note that the vacation service rate exclusively applies to the first arriving customer during the working vacation period.
- Upon customer arrival, a decision is made based on the following probabilities: The customer opts to either join the queue with a probability denoted as ψ_k or decide not to and balk, with the complementary probability expressed as $\psi'_k = 1 - \psi_k$. This decision-making process occurs when there are already k customers ahead in the queue, where $c \leq k \leq K$. It is important to note that the probabilities ψ_k satisfy the conditions $0 \leq \psi_{k+1} \leq \psi_k \leq 1$, $c \leq k \leq K - 1$, $\psi_0 = 1, \dots, \psi_{c-1} = 1$, and $\psi_K = 0$.
- During the normal busy period or either type-1 or type-2 working vacations, customers are governed by impatience timers: T_0, T_1 , or T_2 , respectively. These timers follow exponential distributions with parameters ζ_0, ζ_1 , and ζ_2 . In practical terms, if a customer's service doesn't commence before the timer expires, they will abandon the queue (renege), and their return is not anticipated.

The variables introduced are mutually independent.

2.1. Real-world implementation of the model

The considered queueing system finds practical application in technical software product support centers. Customers seeking assistance with technical issues contact the support center, arriving randomly over time according to a Poisson process (α). During regular operating hours, support agents attend to customers, with service times following an exponential distribution at rate μ_1 . when there is no call in the system, the support agents are allowed to remain in an inactive state for a random period (waiting time). After that, support agents enter type-1 working vacation, where service capacity decreases to μ_2 . Upon return from a working vacation, if there are non-calls, agents transition to a type-2 working vacation. At the time during both type-1 and type-2 working vacation modes, if some calls are present in the system, the support agents can continue operating with probability β or they will switch to the normal busy period with probability $\beta' = 1 - \beta$ and be processed immediately (working vacation interruption). Calls decide whether to join the queue with probability ψ_k and $1 - \psi_k$ denotes the probability that they decide to balk when there are $n \geq c$ incoming calls in front of them in the system. Additionally, during various operational phases, customers are subject to impatience timers (T_0, T_1, T_2), abandoning the queue if service doesn't commence before timer expiration (reneging).

3. EXAMINATION OF THE PROBABILITIES IN A STEADY-STATE

We consider the bi-variate process $(S(t), L(t))_{(t \geq 0)}$, where $L(t)$ is the number of customers in the system at time t , and $S(t)$ defines the state of the servers at time t and takes one of three values, such as $S(t) = 0$: when the servers are in normal busy period at time t , and $S(t) = 1$ (resp. $S(t) = 2$): when the servers are in type-1 (resp. in type-2) working vacation period at time t .

The joint probability $P_{j,k} = \lim_{t \rightarrow \infty} P\{S(t) = j, L(t) = k, (j,k) \in \Omega\}$, denote the steady-state probabilities of the system. Figure 1 shows the transition diagram of the considered model. Next, to avoid overloading mathematical expressions, the following notations are used:

$$\zeta_k = \begin{cases} 0, & k = 0, 1, \\ k\beta'\mu_2, & 2 \leq k \leq c-1, \\ c\beta'\mu_2, & k \geq c, \end{cases} \quad \varphi_{0,k} = \begin{cases} \mu_1, & k = 1, \\ k\mu_1 + (k-1)\zeta_0, & 2 \leq k \leq c, \\ c\mu_1 + (k-1)\zeta_0, & k \geq c+1, \end{cases}$$

$$\zeta_{j,k} = \begin{cases} \mu_2, & k = 1, \\ k\beta\mu_{2j} + (k-1)\zeta_j, & 2 \leq k \leq c-1, \\ c\beta\mu_{2j} + (k-1)\zeta_j, & k \geq c. \end{cases}$$

Using the principle of balance equations

$$(\alpha + \Delta)P_{0,0} = \mu_1 P_{0,1}, \quad k = 0, \tag{1}$$

$$(\alpha + k\mu_1 + (k-1)\zeta_0)P_{0,k} = \alpha P_{0,k-1} + ((k+1)\mu_1 + k\zeta_0)P_{0,k+1} + \Phi_1 P_{1,k} + (k+1)\beta'\mu_2 P_{1,k+1} + \Phi_2 P_{2,k}, \quad 1 \leq k \leq c-1, \tag{2}$$

$$(\alpha\psi_c + c\mu_1 + (c-1)\zeta_0)P_{0,c} = (c\mu_1 + c\zeta_0)P_{0,c+1} + \alpha P_{0,c-1} + \Phi_1 P_{1,c} + c\beta'\mu_2 P_{1,c+1} + \Phi_2 P_{2,c} + c\beta'\mu_2 P_{2,c+1}, \tag{3}$$

$$(\alpha\psi_k + c\mu_1 + (k-1)\zeta_0)P_{0,k} = \alpha\psi_{k-1} P_{0,k-1} + (c\mu_1 + k\zeta_0)P_{0,k+1} + \Phi_1 P_{1,k} + c\beta'\mu_2 P_{1,k+1} + \Phi_2 P_{2,k} + c\beta'\mu_2 P_{2,k+1}, \quad c+1 \leq k \leq K-1, \tag{4}$$

$$(c\mu_1 + (K-1)\zeta_0)P_{0,K} = \alpha\psi_{K-1} P_{0,K-1} + \Phi_1 P_{1,K} + \Phi_2 P_{2,K}, \tag{5}$$

$$(\alpha + \Phi_1)P_{1,0} = \Delta P_{0,0} + \mu_2 P_{1,1}, \quad k = 0, \tag{6}$$

$$(\alpha + k\mu_2 + (k-1)\zeta_1 + \Phi_1)P_{1,k} = \alpha P_{1,k-1} + ((k+1)\beta\mu_2 + k\zeta_1)P_{1,k+1}, \quad 1 \leq k \leq c-1, \tag{7}$$

$$(\alpha\psi_c + c\mu_2 + (c-1)\zeta_1 + \Phi_1)P_{1,c} = \alpha P_{1,c-1} + (c\beta\mu_2 + c\zeta_1)P_{1,c+1}, \tag{8}$$

$$(\alpha\psi_k + c\mu_2 + (k-1)\zeta_1 + \Phi_1)P_{1,k} = \alpha\psi_{k-1} P_{1,k-1} + (c\beta\mu_2 + k\zeta_1)P_{1,k+1}, \quad c+1 \leq k \leq K-1, \tag{9}$$

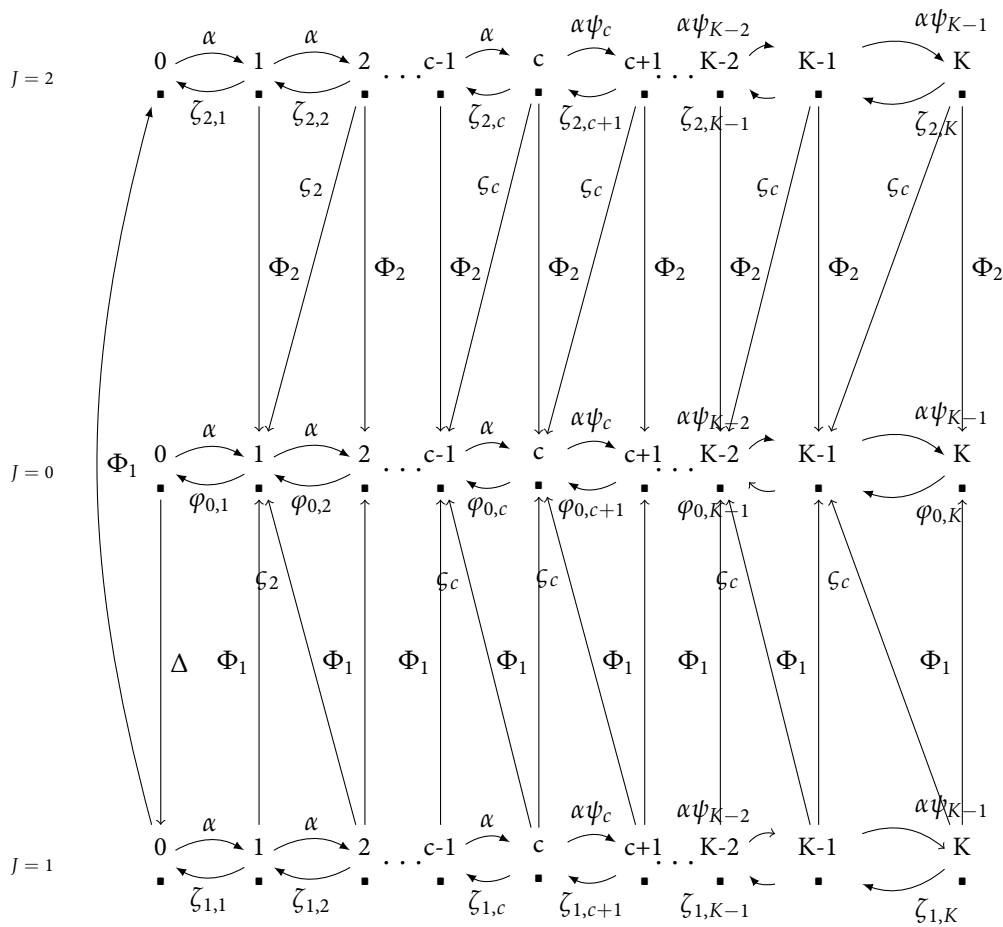


Figure 1: State transition rate diagram

$$(c\mu_2 + (K - 1)\zeta_1) + \Phi_1)P_{1,K} = \alpha\psi_{K-1}P_{1,K-1}, \tag{10}$$

$$\alpha P_{2,0} = \Phi_1 P_{1,0} + \mu_2 P_{2,1}, \quad k = 0, \tag{11}$$

$$(\alpha + k\mu_2 + (k - 1)\zeta_2 + \Phi_2)P_{2,k} = \lambda P_{2,k-1} + ((k + 1)\beta\mu_2 + k\zeta_2)P_{2,k+1}, \quad 1 \leq k \leq c - 1, \tag{12}$$

$$(\alpha\psi_c + c\mu_2 + (c - 1)\zeta_2) + \Phi_2)P_{2,c} = \alpha P_{2,c-1} + (c\beta\mu_2 + c\zeta_2)P_{2,c+1}, \tag{13}$$

$$(\alpha\psi_k + c\mu_2 + (k - 1)\zeta_2) + \Phi_2)P_{2,k} = \alpha\psi_{k-1}P_{2,k-1} + (c\beta\mu_2 + k\zeta_2)P_{2,k+1}, \quad c + 1 \leq k \leq K - 1, \tag{14}$$

$$(c\mu_2 + (K - 1)\zeta_2 + \Phi_2)P_{2,K} = \alpha\psi_{K-1}P_{2,K-1}, \tag{15}$$

The normalizing condition is

$$\sum_{k=0}^K (P_{0,k} + P_{1,k} + P_{2,k}) = 1. \tag{16}$$

Now, we present the solution of the equations above in the following theorem.

Theorem 1. The probabilities describing the system size in different operational periods, namely the type-2 working vacation period ($P_{2,k}$), type-1 working vacation period ($P_{1,k}$), and normal busy period ($P_{0,k}$), in the steady-state are respectively expressed as follows:

$$P_{2,k} = \theta_k P_{2,K} = \theta_k \left(\sum_{k=0}^K (\theta_k + \Theta_1 \delta_k + \Theta_2 \omega_k - \Gamma_k) \right)^{-1}, \quad k = 0, 1, 2, \dots, K. \tag{17}$$

$$P_{1,k} = \Theta_1 \delta_k P_{2,K}. \tag{18}$$

$$P_{0,k} = (\Theta_2 \omega_k - \Gamma_k) P_{2,K}, \tag{19}$$

where

$$\theta_k = \begin{cases} 1, & k = K, \\ \frac{c\mu_2 + (K-1)\xi_2 + \Phi_2}{\alpha\psi_{K-1}}, & k = K - 1, \\ \frac{\alpha\psi_{k+1} + c\mu_2 + \Phi_2 + k\xi_2}{\alpha\psi_k} \theta_{k+1} - \frac{(c\beta\mu_2 + (k+1)\xi_2)}{\alpha\psi_k} \theta_{k+2}, & c \leq k < K - 1, \\ \frac{\alpha\psi_{k+1} + (k+1)\mu_2 + \Phi_2 + k\xi_2}{\alpha} \theta_{k+1} - \frac{((k+1)\beta\mu_2 + (k+1)\xi_2)}{\alpha} \theta_{k+2}, & k = c - 1, \\ \frac{\alpha + (k+1)\mu_2 + \Phi_2 + k\xi_2}{\alpha} \theta_{k+1} - \frac{((k+2)\beta\mu_2 + (k+1)\xi_2)}{\alpha} \theta_{k+2}, & 0 \leq k \leq c - 2, \end{cases} \tag{20}$$

$$\delta_k = \begin{cases} 1, & k = K, \\ \frac{c\mu_2 + (K-1)\xi_1 + \Phi_1}{\alpha\psi_{K-1}}, & k = K - 1, \\ \frac{\alpha\psi_{k+1} + c\mu_2 + \Phi_1 + k\xi_1}{\alpha\psi_k} \delta_{k+1} - \frac{(c\beta\mu_2 + (k+1)\xi_1)}{\alpha\psi_k} \delta_{k+2}, & c \leq k < K - 1, \\ \frac{\alpha\psi_{k+1} + (k+1)\mu_2 + \Phi_1 + k\xi_1}{\alpha} \delta_{k+1} - \frac{((k+1)\beta\mu_2 + (k+1)\xi_1)}{\alpha} \delta_{k+2}, & k = c - 1, \\ \frac{\alpha + (k+1)\mu_2 + k\xi_1 + \Phi_1}{\alpha} \delta_{k+1} - \frac{((k+2)\beta\mu_2 + (k+1)\xi_1)}{\alpha} \delta_{k+2}, & 0 \leq k \leq c - 2, \end{cases} \tag{21}$$

$$\Theta_1 = \frac{\alpha\theta_0 - \mu_2\theta_1}{\Phi_1\delta_0}. \tag{22}$$

$$\omega_k = \begin{cases} 1, & k = K, \\ \frac{c\mu_1 + (K-1)\xi_0}{\alpha\psi_{K-1}}, & k = K - 1, \\ \frac{\alpha\psi_{k+1} + c\mu_1 + k\xi_0}{\alpha\psi_k} \omega_{k+1} - \frac{(c\mu_1 + (k+1)\xi_0)}{\alpha\psi_k} \omega_{k+2}, & c \leq k < K - 1, \\ \frac{\alpha\psi_{k+1} + (k+1)\mu_1 + k\xi_0}{\alpha} \omega_{k+1} - \frac{((k+1)\mu_1 + (k+1)\xi_0)}{\alpha} \omega_{k+2}, & k = c - 1, \\ \frac{\alpha + (k+1)\mu_1 + k\xi_0}{\alpha} \omega_{k+1} - \frac{((k+2)\mu_1 + (k+1)\xi_0)}{\alpha} \omega_{k+2}, & 0 \leq k \leq c - 2, \end{cases}$$

$$\Gamma_k = \begin{cases} 0, & k = K, \\ \frac{\Phi_1\Theta_1 + \Phi_2}{\alpha\psi_{K-1}}, & k = K - 1, \\ \frac{\Theta_1(\Phi_1\delta_{k+1} + c\beta'\mu_1\delta_{k+2}) + (\Phi_2\theta_{k+1} + c\beta'\mu_2\theta_{k+2})}{\alpha\psi_k}, & c \leq k < K - 1, \\ \frac{\Theta_1(\Phi_1\delta_{k+1} + (k+1)\beta'\mu_1\delta_{k+2}) + (\Phi_2\theta_{k+1} + (k+1)\beta'\mu_2\theta_{k+2})}{\alpha}, & k = c - 1, \\ \frac{\Theta_1(\Phi_1\delta_{k+1} + (k+2)\beta'\mu_2\delta_{k+2}) + (\Phi_2\theta_{k+1} + (k+2)\beta'\mu_2\theta_{k+2})}{\alpha}, & 0 \leq k \leq c - 2, \end{cases}$$

$$\Theta_2 = \frac{\Theta_1(\alpha + \Phi_1)\delta_0 - \Theta_1\mu_1\delta_1 + \Delta\Gamma_0}{\Delta\omega_0}, \tag{23}$$

and

$$P_{2,K} = \left(\sum_{k=0}^K (\theta_k + \Theta_1 \delta_k + \Theta_2 \omega_k - \Gamma_k) \right)^{-1}. \quad (24)$$

Proof. The stationary probabilities, denoted as $P_{2,k}$, $P_{1,k}$, and $P_{0,k}$, are determined using equations (1) – (15), expressed in terms of $P_{2,K}$. To calculate $P_{2,k}$, we use a recursive approach to solve equations (12) – (15). This leads us to derive expressions (17) and (20).

For $P_{1,k}$, we find it to be equal to $\Theta_1 \delta_k P_{1,K}$, with δ_k defined by (21). Utilizing equation (11), we obtain equations (18) and (22).

By solving equations (2) – (5), we can express $P_{0,k}$ in terms of both $P_{0,K}$ and $P_{2,K}$. Further, with the assistance of equation (6), we deduce $P_{0,k}$ as a function of $P_{2,K}$, as given in (19).

Finally, we ensure that these probabilities satisfy the normalization condition (see equation (16)), which leads us to equation (24). ■

4. METRICS OF SYSTEM PERFORMANCE

▷ The probabilities associated with different server states—normal busy period, type-1 working vacation, and type-2 working vacation—are defined as follows:

$$P_{bn} = P_{2,K} \sum_{k=0}^K (\Theta_2 \omega_k - \Gamma_k), \quad P_{wv1} = \Theta_1 P_{2,K} \sum_{k=0}^K \delta_k, \quad P_{wv2} = P_{2,K} \sum_{k=0}^K \theta_k.$$

▷ The probabilities of the servers being idle during the busy period (P_{id}) and actively working during the normal busy period (P_{wn}) are expressed as follows:

$$P_{id} = (\Theta_2 \omega_0 - \Gamma_0) P_{2,K}.$$

$$P_{wn} = 1 - \left[P_{2,K} \left((\Theta_2 \omega_0 - \Gamma_0) + \Theta_1 \sum_{k=0}^K \delta_k + \sum_{k=0}^K \theta_k \right) \right]. \quad (25)$$

▷ The expressions for the expected number of customers in the system (L_s) and in the queue (L_q) are defined as follows:

$$L_s = P_{2,K} \left[\sum_{k=0}^K (\Theta_2 k \omega_k - k \Gamma_k + \Theta_1 k \delta_k + k \theta_k) \right]. \quad (26)$$

$$L_q = P_{2,K} \left[\sum_{k=c}^K (\Theta_2 (k - c) \omega_k - (k - c) \Gamma_k + \Theta_1 (k - c) \delta_k + (k - c) \theta_k) \right]. \quad (27)$$

▷ The expression for E_{cs} (expected number of customers served per time unit) is given by:

$$E_{cs} = P_{2,K} \left[\mu_1 \Theta_2 \sum_{k=1}^{c-1} k \omega_k - \mu_1 \sum_{k=1}^{c-1} k \Gamma_k + c \mu_1 \Theta_2 \sum_{n=c}^K \omega_k - c \mu_1 \sum_{k=c}^K \Gamma_k \right]$$

$$+ P_{2,K} \left[\mu_2 \Theta_1 \sum_{k=1}^{c-1} k \delta_k + \mu_2 \sum_{k=1}^{c-1} k \theta_k + c \mu_2 \Theta_1 \sum_{k=c}^K \delta_k + c \mu_2 \sum_{k=c}^K \theta_k \right]. \quad (28)$$

▷ The expressions for the expected waiting time of customers in the system (W_s) and in the queue (W_q) are given by:

$$W_s = \frac{L_s}{\lambda'}, \quad \text{and} \quad W_q = \frac{L_q}{\lambda'}, \quad \text{where } \lambda' = \lambda - B_r. \quad (29)$$

▷ The expected reneging rate:

$$R_r = P_{2,K} \left[\sum_{k=1}^K (\zeta_0 \Theta_2 (k - 1) \omega_k - \zeta_0 (k - 1) \Gamma_k + \zeta_1 \Theta_1 (k - 1) \delta_k) + \zeta_2 \sum_{k=1}^K (k - 1) \theta_k \right]. \quad (30)$$

▷ The expected balking rate:

$$B_r = \alpha P_{2,K} \left[\sum_{k=c}^K (\Theta_2 \psi'_k \omega_k - \psi'_k \Gamma_k + \Theta_1 \psi'_k \delta_k + \psi'_k \theta_k) \right]. \quad (31)$$

5. NUMERICAL RESULTS

This section presents various numerical examples to illustrate the influence of different parameters, including $\alpha, \Phi_1, \zeta_0, \Phi_2, c, \zeta_1, K, \zeta_2$, on the performance metrics of the queueing model ($P_{wv1}, P_{wv2}, P_{wn}, P_{id}, B_r, R_r, E_{cs}, L_s, L_q, W_s, \lambda'$). To do this, we use the probability of non-balking defined as: $\psi_k = 1 - \frac{k}{K}$.

- Scenario 1: We fix $\alpha = 0.01 : .01 : 5, \beta' = 0.3, \Phi_1 = 1.15, \Phi_2 = 1.8, \zeta_0 = 0.7, \zeta_1 = 1.1, \zeta_2 = 1.5$. We consider the following cases:
 - Case 1: $\mu_1 = 2.5, \mu_2 = 1, \Delta = 0.3, c = [1; 2; 3; 4], K = 10$.
 - Case 2: $\mu_1 = 2.5, \mu_2 = 1, \Delta = 0.3, c = 3, K = [10; 15; 20; 25]$.
- Scenario 2: We fix $\Phi_1 = 0.01 : .01 : 2.5, \alpha = 2, \mu_1 = 2.5, \mu_2 = 1, \Delta = 0.3, \Phi_2 = 1.8, c = 3, K = 10$. We study the following cases :
 - Case 1: $\zeta_0 = [0.6; 0.9; 1.2; 1.5], \zeta_1 = 1.1, \zeta_2 = 1.5, \beta' = 0.3$.
 - Case 2: $\zeta_0 = 0.7, \zeta_1 = [0.8; 1.1; 1.4; 1.7], \zeta_2 = 1.5, \beta' = 0.3$.
 - Case 3: $\zeta_0 = 0.7, \zeta_1 = 1.1, \zeta_2 = [1.5; 1.8; 2.1; 2.4], \beta' = 0.3$.
- Scenario 3: We fix $\Phi_2 = 0.01 : .01 : 3, \alpha = 2, \mu_1 = 2.5, \mu_2 = 1, \Delta = 0.3, \Phi = 1.15, c = 3, K = 10$. We study the following cases :
 - Case 1: $\zeta_0 = [0.6; 0.9; 1.2; 1.5], \zeta_1 = 1.1, \zeta_2 = 1.5, \beta' = 0.3$.
 - Case 2: $\zeta_0 = 0.7, \zeta_1 = [0.8; 1.1; 1.4; 1.7], \zeta_2 = 1.5, \beta' = 0.3$.
 - Case 3: $\zeta_0 = 0.7, \zeta_1 = 1.1, \zeta_2 = [1.5; 1.8; 2.1; 2.4], \beta' = 0.3$.

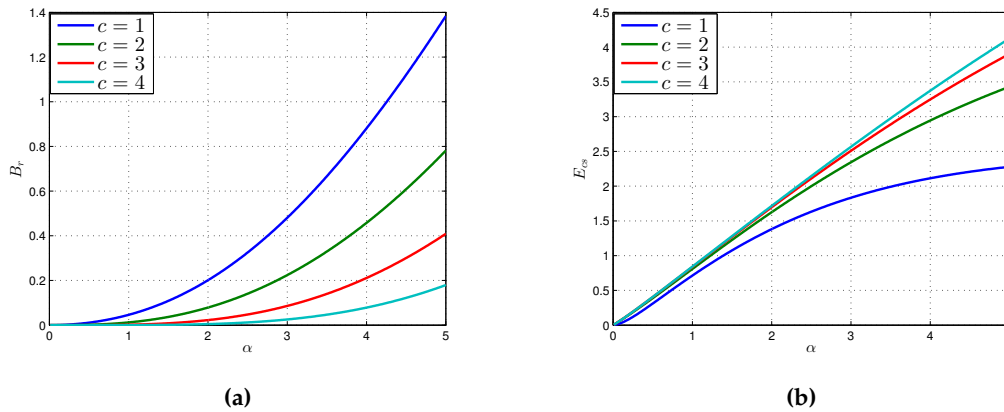


Figure 2: B_r and E_{cs} vs. α for different values of c

Discussion of Results

- ▷ Effect of α (arrival rate): Along with the increasing value of α , several factors are significantly affected. The system size increases, leading to an augmentation in the probability of working during the normal busy period P_{wn} . Additionally, the average balking B_r (see Figures 2a and 3a), mean number of served customer E_{cs} (see Figure 2b), mean number of customers

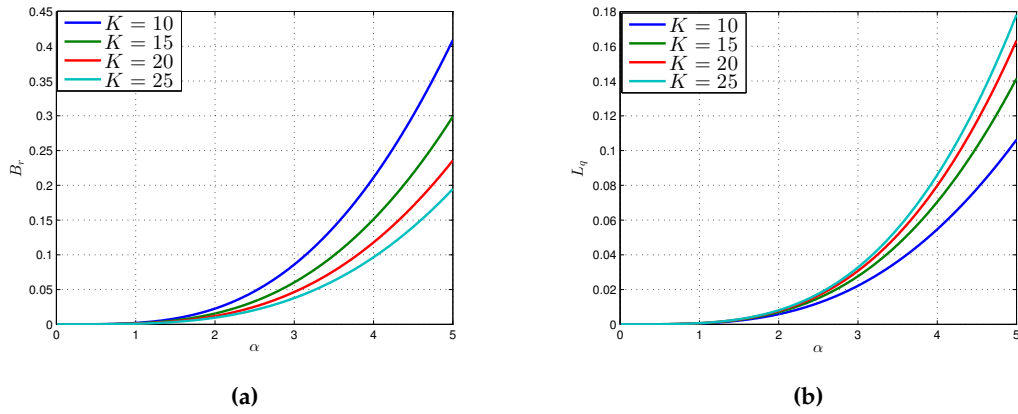


Figure 3: B_r and L_q vs. α for different values of K

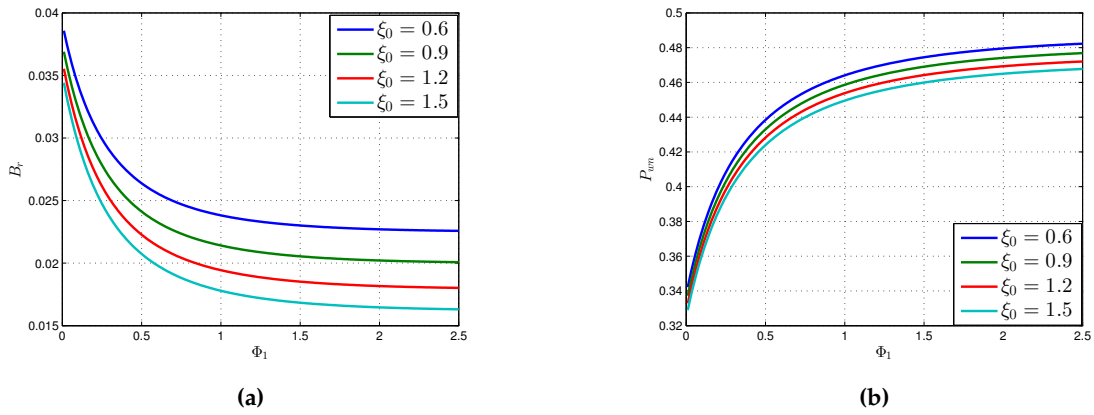


Figure 4: B_r and P_{wm} vs. Φ_1 for different values of ξ_0

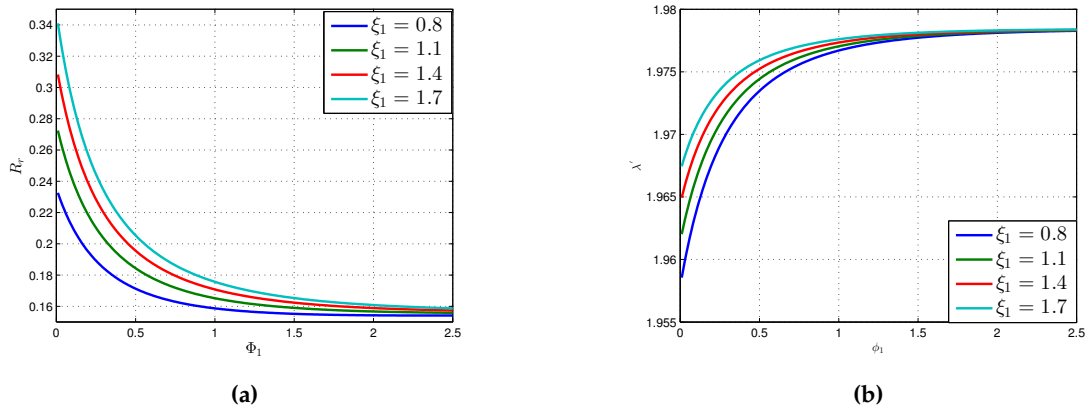


Figure 5: R_r and λ' vs. Φ_1 for different values of ξ_1

in the queue L_q (see Figure 3b) all increase. Conversely, the probabilities P_{wv1} , P_{wv2} and P_{id} decrease. As a result, the average waiting time of a customer in the system decreases. This can be attributed to the effective arrival rate λ' increasing faster than the mean number of customers in the system (L_s).

▷ Effect of c (number of servers): There is clear evidence that as the parameter c increases, the

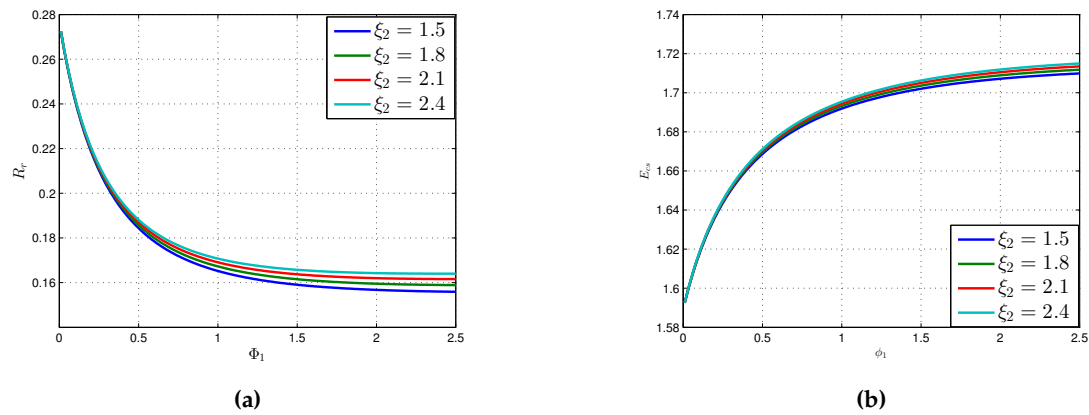


Figure 6: R_r and E_{cs} vs. Φ_1 for different values of ξ_2

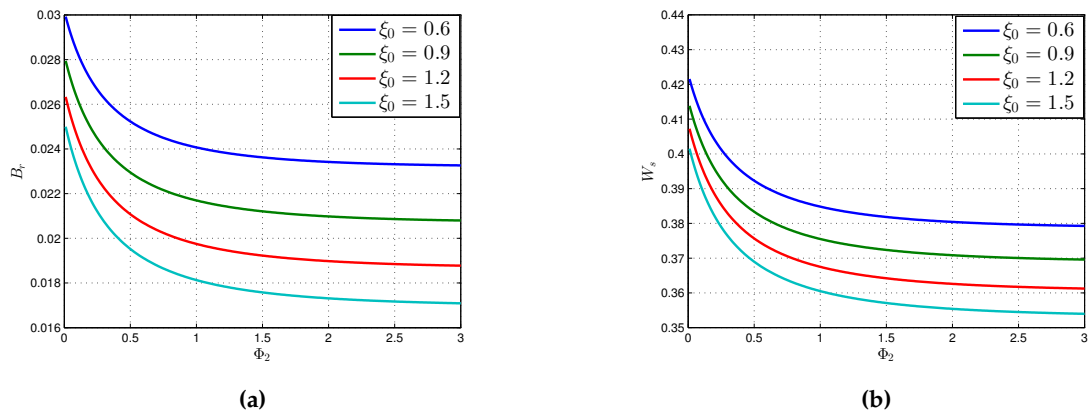


Figure 7: B_r and W_s vs. Φ_2 for different values of ξ_0

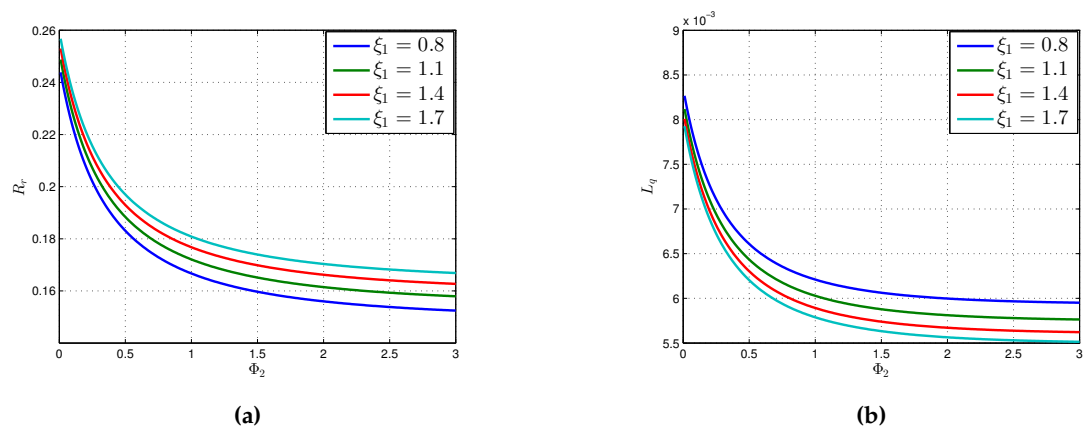


Figure 8: R_r and L_q vs. Φ_2 for different values of ξ_1

quantity L_q decreases. Moreover, a larger number of servers leads to a higher number of customers being served (see Figure 2b), thereby resulting in a reduced average balking rate (cf. Figure 2a).

- ▷ Effect of K (system capacity): The system's large capacity of the parameter K encourages more customers to join the queue, hoping to be served, which leads to a decrease in the average

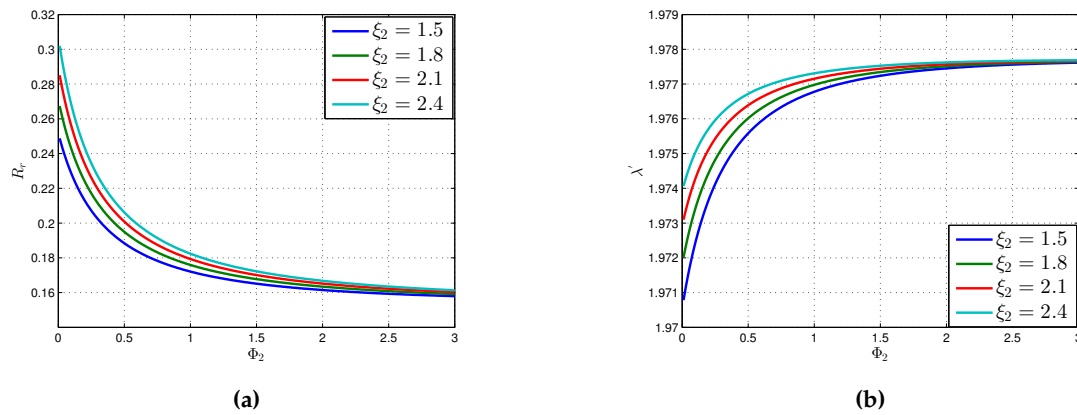


Figure 9: R_r and λ' vs. Φ_2 for different values of ξ_2

value B_r (see Figure 3a). Furthermore, as the systems capacity increases, the average number of customers in the queue also increases (cf. Figure 3b). Thus, there is a significant increase in the mean waiting time for customers.

- ▷ Effect of working vacation rates (Φ_i): By increasing the working vacations rates ϕ_i , ($i = 1, 2$), the system tends to transition quickly to the normal busy period (see Figure 4b) where customers are served much faster (see Figure 6b). This leads to a decrease in the mean waiting time of the customers (cf. Figure 7b). Then, the system becomes rapidly empty. Consequently, the average value B_r decreases (see Figures 4a and 7a), which implies a growth in effective arrivals (cf. Figures 5b, and 9b). Moreover, higher working vacation rates correspond to, lower average renegeing rate (see Figures 5a, 6a, 8a, 9a, and 9b), resulting in smaller mean number of customers in the queue L_q (cf. Figure 8b).
- ▷ Effect of parameters ξ_0 , ξ_1 , and ξ_2 (impatience rates) : Increasing impatience rates, whether during busy normal period or working vacations period, results in increased average value R_r (see Figures 5a, 6a, 8a, and 9a) as well as increased mean number E_{cs} (cf. Figure 6b). Additionally, higher impatience rates lead to, a decrease in the average value W_s (see Figure 7b). Consequently, due to this impatience, there is a decrease in the number of customers both L_s and L_q (cf. Figure 8b). This results in a reduced average balking rate and an increased effective arrival rate (see Figures 4a, 7a, 5b, and 9b).

6. COST OPTIMIZATION

6.1. Cost model

In this section, we propose a model for the costs incurred in our queueing model. In this context, we start by defining the total expected cost per unit of time of the system as:

$$Y(\mu_1, \mu_2) = C_{wn}P_{wn} + C_{id}P_{id} + C_{wv}(P_{wv1} + P_{wv2}) + C_qL_q + C_rR_r + C_bB_r + c\mu_1C_{\mu_1} + c\mu_2C_{\mu_2},$$

where,

- C_{wn} (resp. C_{id}) denotes the cost per unit time when the servers are working (resp. idle) during normal busy period,
- C_{wv} (resp. C_q) is the cost per unit time when the servers are on type-1 or type-2 working vacation period (resp. when a customer joins the queue and waits for service),
- C_r (resp. C_b) is the cost per unit time when a customer reneges (resp. balks),

- C_{μ_1} (resp. C_{μ_2}) denotes the cost per service per unit time during normal busy period (resp. during type-1 or type-2 working vacation period).

6.2. Grey Wolf Optimizer

The GWO algorithm is one of the recent advancements in swarm intelligence optimization (see [19]). It is inspired by grey wolves in nature, which search for the optimal way to hunt prey. The GWO algorithm uses the same mechanism found in nature, where it follows the hierarchy of the pack to organize the different roles in the pack of wolves. In addition, the GWO algorithm is promising for complex optimization problems. This meta-heuristic algorithm efficiently explores the search space and converges to the optimal solution by simulating the hunting behavior of grey wolves. Its simplicity, versatility and proven success make it an invaluable tool for researchers in a variety of fields. We use this novel technique to globally search (μ_1, μ_2) until the minimum value of $Y(\mu_1, \mu_2)$ is achieved.

6.3. Numerical Cost Optimum

The main goal is to identify optimal service rates μ_1 and μ_2 in order to minimize the expected cost function. Because optimization problems are complex and highly non-linear, they are challenging to solve analytically. However, we can utilize appropriate nonlinear optimization techniques to determine the optimal solutions in the cost model. In this case, we fix the parameters and employ the grey wolf optimization algorithm to search for the optimal values (μ_1^*, μ_2^*) for the service rates. The optimization problem can be written as:

$$\begin{aligned} & \min_{\mu_1, \mu_2} Y(\mu_1, \mu_2) \\ & \text{s.t.} \begin{cases} \mu_1 - \mu_2 > 0, \\ \mu_2 > 0, \\ (\mu_1, \mu_2) \in \mathbb{R}_+^2. \end{cases} \end{aligned}$$

The objective is to evaluate the cost function Y in accordance to parameters μ_1 and μ_2 to minimize the total expected cost incurred by the system using Grey Wolf Optimizer.

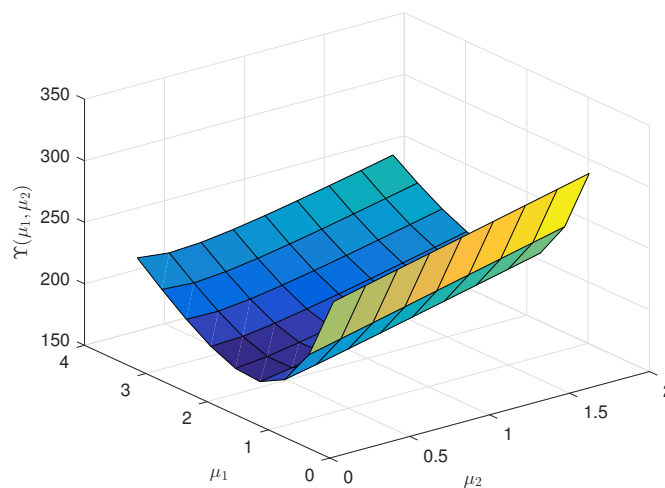


Figure 10: $Y(\mu_1, \mu_2)$ vs. μ_1 and μ_2

Figure 10 effectively visualizes the convexity of the objective function Y according to service rates μ_1 and μ_2 .

Then, in what follows, the optimal solutions are given by applying the GWO meta-heuristic for various system parameters. To do this, we fix the parameters as: $C_s = 45$, $C_{id} = 20$, $C_{wv} = 30$, $C_q = 40$, $C_r = 35$, $C_b = 25$, $C_{\mu_1} = 10$, $C_{\mu_2} = 5$.

Table 1: The optimal (μ_1^*, μ_2^*) and $Y^*(\mu_1^*, \mu_2^*)$ for various values of α and K , when $\alpha = 8 : 1 : 10$, $\Delta = 0.5$, $\beta' = 0.5$, $\Phi_1 = 0.4$, $\Phi_2 = 0.8$, $K = [20; 24; 28]$, $c = 3$, $\zeta_0 = 0.6$, $\zeta_1 = 0.9$, $\zeta_2 = 1.4$.

K	α	μ_1^*	μ_2^*	$Y^*(\mu_1^*, \mu_2^*)$
20	8	3.2914	0.4088	214.5547
	9	3.6627	0.4387	232.1802
	10	4.0280	0.4648	249.4850
24	8	3.3526	0.4278	215.1095
	9	3.7278	0.4581	232.7175
	10	4.0999	0.4857	249.9891
28	8	3.3939	0.4384	216.5381
	9	3.7759	0.4713	233.1391
	10	4.1548	0.5005	250.3924

Table 2: The optimal (μ_1^*, μ_2^*) and $Y^*(\mu_1^*, \mu_2^*)$ for various values of Δ when $\alpha = 9$, $\Delta = 0.2 : 0.2 : 0.8$, $\beta' = 0.4$, $\Phi_1 = 0.4$, $\Phi_2 = 0.8$, $K = 24$, $c = 3$, $\zeta_0 = 0.6$, $\zeta_1 = 0.9$, $\zeta_2 = 1.4$.

Δ	μ_1^*	μ_2^*	$Y^*(\mu_1^*, \mu_2^*)$
0.2	3.7858	0.2393	228.1624
0.4	3.7433	0.3986	231.5360
0.6	3.7129	0.5071	233.7100
0.8	3.6941	0.5948	235.3126

Table 3: The optimal (μ_1^*, μ_2^*) and $Y^*(\mu_1^*, \mu_2^*)$ for various values of β' , when $\alpha = 9$, $\Delta = 0.5$, $\beta' = 0.3 : 0.2 : 0.9$, $\Phi_1 = 0.4$, $\Phi_2 = 0.8$, $K = 24$, $c = 3$, $\zeta_0 = 0.6$, $\zeta_1 = 0.9$, $\zeta_2 = 1.4$.

β'	μ_1^*	μ_2^*	$Y^*(\mu_1^*, \mu_2^*)$
0.3	3.6668	0.5143	235.5446
0.5	3.7264	0.4575	232.7175
0.7	3.7621	0.4140	230.9342
0.9	3.7850	0.3798	229.6789

- From Table 1, it can be clearly seen that the optimum expected cost $Y^*(\mu_1^*, \mu_2^*)$ exhibits a significant increase as the values of the arrival rate α and finite capacity K increases.
- From Table 2, can be observed that as Δ , value increases, the minimum expected cost increases. This observation clearly indicates that increasing the waiting rate of servers is an expensive endeavor.
- From Table 3, when interruption probability (β') increases, the minimum expected cost decreases. So, a higher interruption probability positively affects the overall expected cost of the system.

7. CONCLUSION

This paper focused on the analysis of a finite-space multi-server queue where customers exhibit impatience under a synchronous differentiated working vacation policy. Specifically, the customers are assumed to be impatient during the normal busy period, as well as during type-1 and type-2 working vacations. The main goal of the analysis is to determine the steady-state probabilities of the system size under different server states, including the normal busy period and the working vacations (type-1 and type-2). This is achieved through the application of recursive analysis techniques. We derived important system performance measures that provide valuable insights into the behavior and efficiency of the considered multi-server queueing system.

A GWO algorithm is performed to determine the optimal service rates for both working vacations and normal busy periods aiming to minimize the expected total cost. The problem at hand is formulated as a nonlinear optimization problem, and several numerical examples are provided to illustrate the effectiveness of the proposed approach. The focus of the analysis is on conducting a cost optimization study, where the effect of different system parameters and cost elements is investigated. The numerical examples and cost optimization analysis presented in this study shed light on the significance of system parameters and cost elements in queueing systems. Overall, this study contributes to the understanding and optimization of queueing systems, highlighting the potential advantages of cost optimization techniques in various real-life and industrial settings.

The model discussed in the paper can be extended to handle more complex scenarios, such as an unreliable multi-server queue with heterogeneous customers, which introduces additional complexity to the problem. While this extension increases the dimension of the problem significantly. It is also possible to relax the exponential assumptions by considering phase-type distributions for service times.

REFERENCES

- [1] Afroun, F., Aïssani, D., Hamadouche, D., and Boualem, M. (2018). Q-matrix method for the analysis and performance evaluation of unreliable $M/M/1/N$ queueing model. *Mathematical Methods in the Applied Sciences*, 41:9152–9163.
- [2] Baba, Y. (2010). The $M/PH/1$ queue with working vacations and vacation interruption. *Journal of Systems Science and Systems Engineering*, 19:496–503.
- [3] Bouchentouf, A.A. and Guendouzi A., and Kandoucib, A. (2019). Performance and economic study of heterogeneous $M/M/2/N$ feedback queue with working vacation and impatient customers. *ProbStat Forum*, 12:15–35.
- [4] Bouchentouf, A.A., Boualem, M., Yahiaoui, L., and Ahmad, H. (2022). A multi-station unreliable machine model with working vacation policy and customers impatience. *Quality Technology & Quantitative Management*, 19:766–796.
- [5] Bouchentouf, A.A., Cherfaoui, M., and Boualem, M. (2019). Performance and economic analysis of a single server feedback queueing model with vacation and impatient customers. *Opsearch*, 56:300–323.
- [6] Bouchentouf, A. A., Guendouzi, A., and Majid, S. (2020). On impatience in Markovian $M/M/1/N/DWV$ queue with vacation interruption. *Croatian Operational Research Review*, 11:21–37.
- [7] Bouchentouf, A.A., Cherfaoui, M., and Boualem, M. (2021). Analysis and performance evaluation of Markovian feedback multi-server queueing model with vacation and impatience. *American Journal of Mathematical and Management Sciences*, 40:261–282.
- [8] Chettouf, A., Bouchentouf, A.A., and Boualem, M. (2024). A Markovian Queueing Model for Telecommunications Support Center with Breakdowns and Vacation Periods. *Oper. Res. Forum* 5, article number 22. <https://doi.org/10.1007/s43069-024-00295-y>
- [9] Doshi, B. T. (1986). Single server queues with vacation: A survey. *Queueing Systems*, Vol. 1, No. 1, 29–66.

- [10] Doshi, B. T. (1990). Single server queues with vacations. *Stochastic Analysis of the Computer and Communication Systems*, 217–264.
- [11] Jyothsna, K., Laxmi, P. V., and Kumar, V. P. (2022). Analysis of $GI/M/1/N$ and $GI/Geo/1/N$ queues with balking and vacation interruptions. *Journal of Mathematical Modeling*, 10:569-585.
- [12] Prakati, P. (2024). $M/M/C$ queue with multiple working vacations and single working vacation under encouraged arrival with impatient customers. *Reliability: Theory and Applications*, 19 :650-662.
- [13] Laxmi, P. V., and Jyothsna, K. (2015). Impatient customer queue with Bernoulli schedule vacation interruption. *Computers and Operations Research*, 56:1–7.
- [14] Li, J., and Tian, N. (2007). The $M/M/1$ queue with working vacations and vacation interruptions. *Journal of Systems Science and Systems Engineering*, 16:121–127.
- [15] Li, J. H., Tian, N. S., and Ma, Z. Y. (2008). Performance analysis of $GI/M/1$ queue with working vacations and vacation interruption. *Applied Mathematical Modelling*, 32:2715–2730.
- [16] Majid, S., and Manoharan, P. (2018). Impatient customers in an $M/M/c$ queue with single and multiple synchronous working vacations. *Pakistan Journal of Statistics and Operation Research*, 14:571–594.
- [17] Majid, S., Bouchentouf, A. A., and Guendouzi, A. (2021). Analysis and optimisation of a $M/M/1/WV$ queue with Bernoulli schedule vacation interruption and customers impatience. *Acta Universitatis Sapientiae, Mathematica*, 13:367–395.
- [18] Majid, S., Manoharan, P., and Ashok, A. (2018). An $M/M/1$ queue with working vacation and vacation interruption under Bernoulli schedule. *International Journal of Engineering and Technology*, 7:448–454.
- [19] Mirjalili, S., Mirjalili, S. M., and Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, 69:46–61.
- [20] Sasikala, S., and Abinaya, V. (2023). The $M/M/c/N$ inter dependent inter Arrival queueing model with controllable arrival rates, reverse balking and impatient Customers. *Arya Bhatta Journal of Mathematics and Informatics*, 15:1–10.
- [21] Shekhar, C., Varshney, S., and Kumar, A. (2021). Matrix-geometric solution of multi-server queueing systems with Bernoulli scheduled modified vacation and retention of reneged customers: A meta-heuristic approach. *Quality Technology & Quantitative Management*, 18:39–66.
- [22] Schwarz, M., Sauer, C., Daduna, H., Kulik, R., and Szekli, R. (2006). $M/M/1$ queueing systems with inventory. *Queueing Systems*, 54:55–78.
- [23] Stolletz, R. (2012). Performance analysis and optimization of inbound call centers. *Springer Science and Business Media*, 528:1–25. <https://doi.org/10.1007/978-3-642-55506-0>.
- [24] Servi, L. D., and Finn, S. G. (2002). $M/M/1$ queues with working vacations ($M/M/1/WV$). *Performance Evaluation*, 50:41–52.
- [25] Yue, D., and Yue, W. (2009). Analysis of an $M/M/c/N$ queueing system with balking, reneging, and synchronous vacations. *Advances in Queueing Theory and Network Applications*, 165–180.
- [26] Ziad, I., Laxmi, P. V., Bhavani, E. G., Bouchentouf, A. A., and Majid, S. (2023). A matrix geometric solution of a multi-server queue with waiting servers and customers impatience under variant working vacation and vacation interruption. *Yugoslav Journal of Operations Research*, 3:389–407. <https://doi.org/10.2298/YJOR220315001Z>.

COSINE MARSHAL-OLKIN-G FAMILY OF DISTRIBUTION: PROPERTIES AND APPLICATIONS

Akeem Ajibola Adepoju¹, Alhaji Modu Isa², Olalekan Akanji Bello³

¹Department of Statistics, Faculty of Computing and Mathematical Sciences, Aliko Dangote
University of Science and Technology, Wudil, 713281, Wudil, Kano. Nigeria.

²Department of Mathematics and Computer Science, Borno State University, Maiduguri, Nigeria.

³Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.
Email: akeebola@gmail.com ; alhajimoduisa@bosu.edu.ng ; olalekan4@gmail.com

Abstract

Trigonometric distributions have recently been emphasized due to its applicability and relevance for modeling different phenomena. This article contributes to the existing literature on trigonometric family by introducing and investigating new trigonometric family of distribution which is developed by compounding the cosine family of distribution with Marshall-olkin family of distribution to form a new Cosine Marshall-Olkin family of distribution (CMO). Graphical, numerical and analytical approach was explored to study the properties and applicability of the new CMO family of distribution. Special representations and important reliability properties and other statistical properties were defined. Simulation study was conducted in order to have an insight on the estimates of the three parameters model using maximum products of spacing (MPS). Emphases on the greater flexibility of the new CMO family of distribution beyond the cosine-G family and other top models of the Cosine related family was made through Weibull distribution. The results revealed the superiority of the Cosine Marshall-Olkin Weibull model (CMO-W) over others via two data sets.

Keywords: Cosine-G family, Marshall-Olkin-G family, Maximum Products of Spacing, Hazard function, Survival function.

I. Introduction

Recently, many authors have introduced various approaches to develop flexible continuous distributions from classical continuous distributions. The statisticians' attentions have been drawn to various applications of these continuous distributions in environment, physics, medicine, biology, finance, insurance, engineering and economy to mention few. The classical distributions are induced by adding parameter(s) to enhance the asymmetry, kurtosis, tails properties, central and dispersion parameters. This idea is considered as generalization of the classical distributions. These generalized distributions belong to particular families defined by transformation of the baseline cumulative distribution function (cdf). The values of the newly introduced parameter(s) can enhance the statistical capacities of the baseline distribution. for instance, families such as Weibull-G [1], Exp-G [2], Topp-Leone generated (TL-G) [3] Type I Half Logistic-G [4], new power

TL-G [5], Type II half Logistic-G [6], truncated inverted Kumaraswamy-G [7], a new alpha power transformed-G [8], a new extended alpha power transformed-G [9], type II power TL-G [10], Odd Beta prime-G [11].

A recent approach involves defining families of distributions by using the trigonometric transformation, be it parametric or not. Kumar et al. [12] and Souza [13] launched this trigonometric family exploring the use of the sine function, resulting to the sine-G family. The [14] and [15] extended the exponential and weibull distribution through sine-G family. The non trigonometric compounding families of distributions seen in the literature include but not limited to [16], [17], its extension is found in [18], [19]. The trigonometric compounded families include [20], [21], [22], [23], [24], [25] [26], [27], [28], [29], [30], [31], [32], [33].

The Marshall-olkin-G family of distribution was proposed by [34] and it was used to extended flexibility of Exponential and weibull distribution The Cosine –G family of distribution was proposed by [35]. Now, this article intends to compound the two families to form a new family of distribution called Cosine Marshall-olkin-G family of distribution.

The motivations behind CMO-G family are to develop models with improved shapes for the pdf and hazard function, improve symmetrical and asymmetrical distributions, construct heavy-tailed distributions, improve the flexibility of the baseline model through skewness, kurtosis, mean and variance, provide better fits than other Cosine family of distribution with the same baseline distribution and possibly with the same number of parameters and more complexity.

II. Methods

2.1 The Marshal-Olkin-G Family of Distribution

Definition 1: Suppose $X \sim MO(x; \theta, \xi)$ with corresponding cdf and pdf given by:

$$H_{MO}(x; \theta, \xi) = \frac{G(x; \xi)}{\theta + (1 - \theta)G(x; \xi)} \quad (1)$$

and

$$h_{MO}(x; \theta, \xi) = \frac{\theta g(x; \xi)}{(\theta + (1 - \theta)G(x; \xi))^2} \quad -\infty < x < \infty \text{ where } \theta > 0, \text{ and it is a shape parameter} \quad (2)$$

2.2 The Cosine-G Family of Probability Distribution

Definition 2: Suppose $X \sim COS(x; \Psi)$ with corresponding cdf and pdf given by:

$$F(x; \Psi) = 1 - \cos\left[\frac{\pi}{2}H(x)\right] \quad (3)$$

and

$$f(x; \Psi) = \frac{\pi}{2}h(x)\sin\left[\frac{\pi}{2}H(x)\right] \quad (4)$$

2.3 The proposed Cosine Marshal Olkin-G family of distribution

Definition 3: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf expressed below, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter is defined as the Cosine Marshal-Olkin-G Family

$$F_{CMO}(x; \theta) = 1 - \cos \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right] \quad (5)$$

It is important to note that for any baseline distribution, signified as $G(x)$, CMO cdf satisfy the following;

- a. $g(x) = \frac{dG(x)}{dx}$
- b. $\int_0^{\infty} g(x) dx = 1$
- c. The survival function $1 - G(x)$

Definition 4: Suppose $X \sim CMO(x; \theta, \xi)$ with pdf expressed below, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter is defined as the Cosine Marshal-Olkin-G Family

$$f_{CMO}(x; \theta) = \frac{\pi}{2} \frac{\theta g(x)}{(\theta + (1-\theta)G(x))^2} \sin \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right] \quad (6)$$

2.4 Special Representation

The pdf of the proposed Cosine Marshall-olkinG family can be expanded using the tailor series and binomial expansion; thus

$$f_{CMO}(x; \theta) = \frac{\pi}{2} \frac{\theta g(x)}{(\theta + (1-\theta)G(x))^2} \sin \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right]$$

$$\sin \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right] = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i+1)!} \frac{\pi^{2i+1}}{2^{2i+1}} G(x)^{2i} (\theta + (1-\theta)G(x))^{-2i}$$

Consider $(\theta + (1-\theta)G(x))^{-2i}$ and $(\theta + (1-\theta)G(x))^{-2}$

$$(\theta + (1-\theta)G(x))^{-2(i+1)} = \sum_{j=0}^{\infty} \theta^j (-1)^j \binom{2(i+1)-1+j}{j} (1-\theta)^j G(x)^{2ij}$$

$$f_{CMO}(x; \theta) = \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j}}{(2i+1)!} \frac{\pi^{2i+2}}{2^{2i+2}} \binom{2(i+1)-1+j}{j} \theta^{i+1} (1-\theta)^j g(x) G(x)^{2ij}$$

Hence the expansion of the pdf is expressed as

$$f_{CMO}(x; \theta) = \sum_{i,j=0}^{\infty} \Psi_{i,j} g(x) G(x)^{2ij} \tag{7}$$

Where

$$\Psi_{i,j} = \frac{(-1)^{i+j} \pi^{2i+2}}{(2i+1)! 2^{2i+2}} \binom{2(i+1)-1+j}{j} \theta^{j+1} (1-\theta)^j$$

The cdf can also be expanded as follows:

$$\begin{aligned} F_{CMO}(x; \theta) &= 1 - \cos \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right] = 1 - \sum_{k=0}^{\infty} \frac{(-1)^k \pi^{2k}}{(2k)! 2^{2k}} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right)^{2k} \\ &= 1 - \sum_{k=0}^{\infty} \frac{(-1)^k \pi^{2k}}{(2k)! 2^{2k}} G(x)^{2k} (\theta + (1-\theta)G(x))^{-2k} \\ &= \sum_{l=0}^{\infty} \theta^l (-1)^l \binom{2(k+1)-1+l}{l} (1-\theta)^l G(x)^l \end{aligned}$$

$$F_{CMO}(x; \theta) = 1 - \sum_{k,l=0}^{\infty} \frac{(-1)^{k+l} \pi^{2k}}{(2k)! 2^{2k}} \binom{2(k+1)-1+l}{l} \theta^l (1-\theta)^l G(x)^{2k+l}$$

Therefore,

$$F_{CMO}(x; \theta, \xi) = 1 - \sum_{k,l=0}^{\infty} \Phi_{k,l} G(x)^{2k+l} \tag{8}$$

Where

$$\Phi_{k,l} = \frac{(-1)^{k+l} \pi^{2k}}{(2k)! 2^{2k}} \binom{2(k+1)-1+l}{l} \theta^l (1-\theta)^l$$

Definition 5: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf and pdf well defined, where $\theta > 0$ and is a shape parameter and ξ is a baseline vector parameter. Then the survival function of X, signified by $SF_{CMO}(x; \theta, \xi) = 1 - F_{CMO}(x; \theta)$, the survival function for the CMO family of distribution, can be

represented by $SF_{CMO}(x; \theta, \xi) = \cos \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right]$

Definition 6: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf and pdf well defined, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter. Then the hazard rate function of X, signified by $HRF_{CMO}(x; \theta, \xi) = f_{CMO}(x; \theta, \xi) / SF_{CMO}(x; \theta)$, the hazard rate function for the CMO family of distribution, can be represented by

$$HRF_{CMO}(x; \theta, \xi) = \frac{\pi \theta g(x)}{2 (\theta + (1-\theta)G(x))^2} \tan \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right] \tag{10}$$

Definition 7: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf and pdf well defined, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter. Then the Qunatile function of X , signified by

$QF_{CMO}(x; \theta, \xi) = F_{CMO}^{-1}(x; \theta, \xi)$, the Qunatile function for the CMO family of distribution can be obtained as follows:

$$u = 1 - \cos \left[\frac{\pi}{2} \left(\frac{G(x)}{\theta + (1-\theta)G(x)} \right) \right], u \in (0,1),$$

$$\Phi(u) = G^{-1} \left\{ \frac{\theta \cos^{-1}(1-u)}{\frac{\pi}{2} - (1-\theta) \cos^{-1}(1-u)} \right\} \quad (11)$$

Definition 8: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf and pdf well defined, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter. Then the r^{th} Moments of X can be obtained as follow

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x) dx$$

$$\mu'_r = \sum_{i,j=0}^{\infty} \Psi_{i,j} \int_0^{\infty} x^r g(x) G(x)^{2ij} dx$$

$$\mu'_r = \sum_{i,j=0}^{\infty} \Psi_{i,j} \Phi \quad (12)$$

where

$$\Phi = \int_0^{\infty} x^r g(x) G(x)^{2ij} dx$$

Definition 9: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf and pdf well defined, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter. The r^{th} Moment generating function of X is obtained through

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Thus, the moment generating function of the Cosine Marshall-olkin-G family of distribution is given by:

$$M_x(t) = \sum_{i,j=0}^{\infty} \Psi_{i,j} \int_0^{\infty} e^{tx} g(x) G(x)^{2ij} dx$$

$$M_x(t) = \sum_{i,j=0}^{\infty} \Psi_{i,j} Y \quad (13)$$

where

$$Y = \int_0^{\infty} e^{tx} g(x) G(x)^{2ij} dx$$

Definition 10: Suppose $X \sim CMO(x; \theta, \xi)$ with cdf and pdf well defined, where $\theta > 0$ and θ is a shape parameter and ξ is a baseline vector parameter. The entropy is obtained as given below

$$I_{\theta}(x) = \frac{1}{1-\theta} \log \int_0^{\infty} f(x)^{\theta} dx$$

$$f(x)^{\theta} = \left(\sum_{i,j=0}^{\infty} \Psi_{i,j} g(x) G(x)^{2ij} \right)^{\theta}$$

$$f(x)^\theta = \left(\sum_{i,j=0}^{\infty} \Psi_{i,j} \right)^\theta \left(g(x)G(x)^{2ij} \right)^\theta$$

Let $\omega = g(x)G(x)^{2ij}$

Therefore,

$$f(x)^\theta = \left(\sum_{i,j=0}^{\infty} \Psi_{i,j} \right)^\theta \omega^\theta \tag{14}$$

$$I_\theta(x) = \frac{1}{1-\theta} \left[\left(\sum_{i,j=0}^{\infty} \Psi_{i,j} \right)^\theta \int_0^\infty \omega^\theta dx \right]$$

2.5 Cosine Marshall-olkinWeibull Distribution

Supposed the baseline distribution is Weibull distribution with cdf and pdf given by:

$$G(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha} \tag{15}$$

$$g(x) = \frac{\alpha}{\lambda} \left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-\left(\frac{x}{\lambda}\right)^\alpha} \tag{16}$$

Where α is a shape parameter and λ is a scale parameter, then the cumulative distribution, probability distribution, hazard and survival function of the Cosine Marshall-olkinWeibull

(CMO-W) distribution is given as: $F_{CMO}(x;\theta) = 1 - \cos \left[\frac{\pi}{2} \frac{1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha}}{\theta + (1-\theta) \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha} \right)} \right]$ (17)

and the associated pdf is given as:

$$f_{CMO}(x;\theta) = \frac{\frac{\pi}{2} \frac{\theta \alpha}{\lambda} \left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-\left(\frac{x}{\lambda}\right)^\alpha}}{\left[\theta + (1-\theta) \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha} \right) \right]^2} \sin \left\{ \frac{\pi}{2} \frac{1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha}}{\left[\theta + (1-\theta) \left(1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha} \right) \right]} \right\}$$

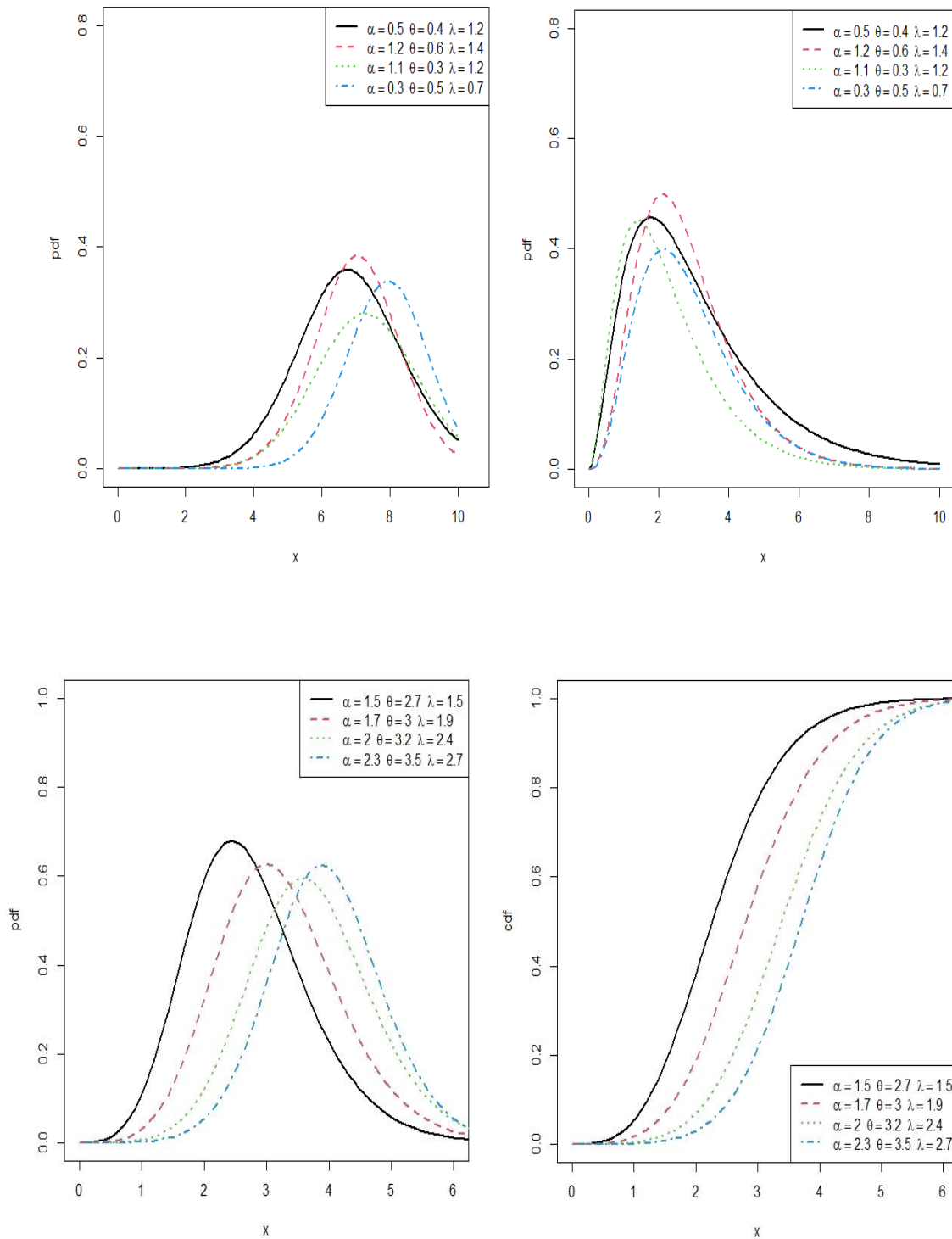


Figure 1: Plots of pdf and cdf of CMO-W distribution

The figure 1 above reveals left skewness, right skewness and approximately symmetric pdf shapes. The cdf shape converges to one, validating the CMO-W distribution.

The Hazard and reliability function of the CMO-W distribution is obtained as:

$$HRF_{CMO}(x; \theta, \xi) = \frac{\pi \theta \left(\frac{\alpha}{\lambda} \left(\frac{x}{\lambda} \right)^{\alpha-1} e^{-\left(\frac{x}{\lambda} \right)^\alpha} \right)}{2 \left(\theta + (1-\theta) \left(1 - e^{-\left(\frac{x}{\lambda} \right)^\alpha} \right) \right)^2} \tan \left[\frac{\pi}{2} \frac{1 - e^{-\left(\frac{x}{\lambda} \right)^\alpha}}{\theta + (1-\theta) \left(1 - e^{-\left(\frac{x}{\lambda} \right)^\alpha} \right)} \right] \quad (19)$$

and

$$SF_{CMO}(x; \theta, \xi) = \cos \left[\frac{\pi}{2} \frac{1 - e^{-\left(\frac{x}{\lambda} \right)^\alpha}}{\theta + (1-\theta) \left(1 - e^{-\left(\frac{x}{\lambda} \right)^\alpha} \right)} \right] \quad (20)$$

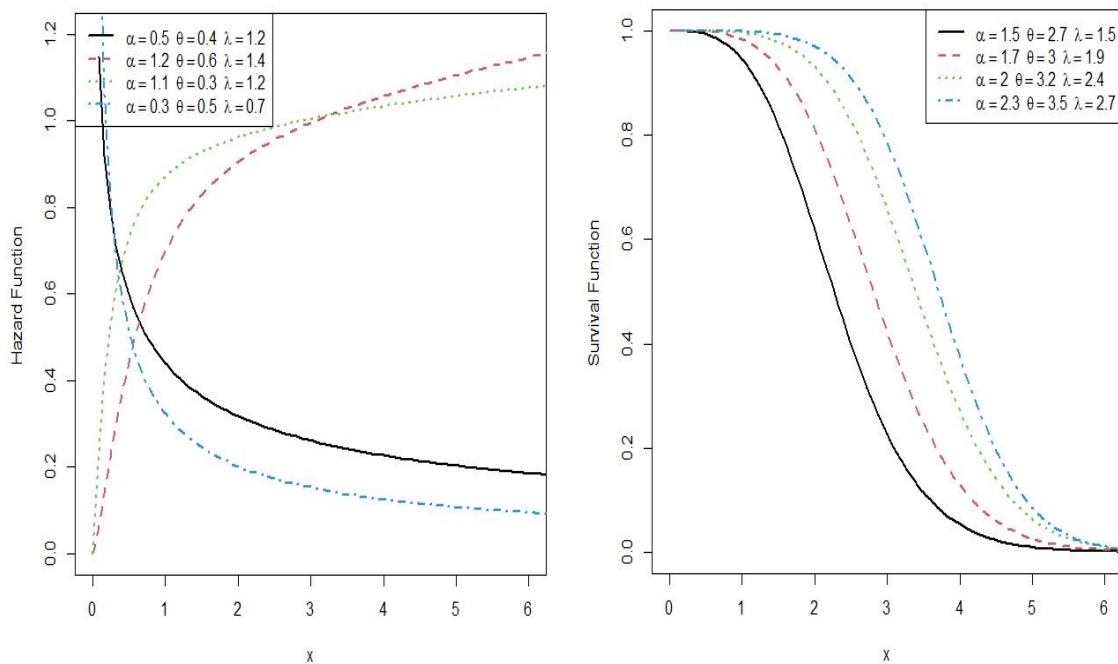


Figure 2: Plots of hazard and reliability function of CMO-W distribution

The figure 2 above reveals the shapes of the hazard and reliability function the hazard shapes obviously shows increasing and decreasing failure rate, and the reliability shapes shows a drop from one to zero with varying values of parameters

III. Results

3.1 Simulation study

In this section, we provide, we provides the simulation of parameters of the CMO-W distribution using Maximum products of spacing estimation method. Random numbers were systematically generated from fixed values of the parameters $\theta=0.5, \lambda=2, \alpha=1$, $\theta=0.7, \lambda=2.2, \alpha=1$ and $\theta=0.6, \lambda=2.3, \alpha=1$ and $\theta=0.8, \lambda=2.1, \alpha=1$ based on 10,000 replications. The sample sizes (n) considered are 20, 50, 100, 250, 500 and 1000. The result is displayed in Table 1 and Table 2

Table 1: The MPSs parameter estimates (Est. value), Biases and RMSEs of various parameters values

n	Parameters	Est. value	Bias	RMSE	Est. value	Bias	MSE
20	θ	0.6195	0.1195	0.4675	0.8346	0.1346	0.5855
	λ	2.1325	0.1325	0.8886	2.3199	0.1199	0.9108
	α	0.9536	-0.0464	0.2357	0.9563	-0.0437	0.2327
50	θ	0.5984	0.0984	0.3570	0.8354	0.1354	0.5037
	λ	2.0474	0.0474	0.7208	2.2352	0.0352	0.7881
	α	0.9581	-0.0419	0.1652	0.9608	-0.0392	0.1722
100	θ	0.5719	0.0719	0.2906	0.7981	0.0981	0.3977
	λ	2.0375	0.0375	0.6079	2.2232	0.0232	0.6503
	α	0.9703	-0.0297	0.1247	0.9712	-0.0288	0.1281
250	θ	0.5572	0.0572	0.2226	0.7788	0.0788	0.3004
	λ	2.0014	0.0014	0.4941	2.2114	0.0114	0.5370
	α	0.9786	-0.0214	0.0970	0.9787	-0.0213	0.1010
500	θ	0.5318	0.0318	0.1674	0.7592	0.0592	0.2422
	λ	2.0092	0.0092	0.3890	2.2032	0.0032	0.4317
	α	0.9873	-0.0127	0.0745	0.9832	-0.0168	0.0801
1000	θ	0.5197	0.0197	0.1254	0.7326	0.0326	0.1810
	λ	2.0041	0.0041	0.2914	2.2011	0.0010	0.3362
	α	0.9918	-0.0082	0.0547	0.9904	-0.0096	0.0608

Table 2: The MPSs parameter estimates (Est. value), Biases and RMSEs of various parameters values

n	Parameters	Est. value	Bias	RMSE	Est. value	Bias	RMSE
20	θ	0.7259	0.1259	0.5302	0.9459	0.1459	0.6535
	λ	2.4416	0.1416	0.9796	2.2125	0.1125	0.8606
	α	0.9567	-0.0433	0.2366	0.9607	-0.0393	0.2402
50	θ	0.7174	0.1174	0.4218	0.9525	0.1525	0.5519
	λ	2.3376	0.0376	0.8233	2.1128	0.0128	0.7333
	α	0.9593	-0.0407	0.1690	0.9588	-0.0412	0.1711
100	θ	0.6863	0.0863	0.3426	0.9240	0.0981	0.1240
	λ	2.3231	0.0231	0.6725	2.1063	0.0232	0.0063
	α	0.9703	-0.0297	0.1253	0.9698	-0.0288	-0.0302
250	θ	0.0722	0.0722	0.0722	0.9008	0.1008	0.3486
	λ	-0.0181	-0.0181	-0.0181	2.0640	-0.0360	0.4893
	α	-0.0226	-0.0226	-0.0226	0.9752	0.0248	0.1006
500	θ	0.6444	0.0444	0.2047	0.8628	0.0628	0.2716
	λ	2.2928	-0.0072	0.4473	2.0824	-0.0176	0.4125
	α	0.9851	-0.0149	0.0772	0.9840	-0.0160	0.0808
1000	θ	0.6300	0.0300	0.1558	0.8433	0.0433	0.2073
	λ	2.2845	-0.0155	0.3418	2.0783	-0.0217	0.3206
	α	0.9897	-0.0103	0.0585	0.9886	-0.0114	0.0624

3.2 Applications

Application of the CMO-W distribution to two real life data sets are provided and revealing its applicability in practice along with comparison with its comparators. The proposed Cosine Marshall-olkin-Weibull distribution (CMO-W) is compared with four other Cosine extended Weibull distributions, namely: Cosine Topp–Leone Weibull (CTL-W) distribution [36], Extended Cosine Weibull (ECS-W) distribution [37], New Alpha Power Cosine-Weibull (NACos-W) distribution [38] and Cosine Weibull (C-W) distribution [39].

The information criteria explored to investigate the goodness-of-fit of the distribution appropriate for the data are Akaike's Information Criterion (AIC), Consistent Akaike's Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan-Quinn Information Criterion (HQIC). The computation can be seen as follows

$$AIC = -2\ell + 2p,$$

$$CAIC = -2\ell + \frac{2np}{n-p-1},$$

$$BIC = -2\ell + p\log(n),$$

$$HQIC = -2\ell + 2p\log(\log(n)),$$

where ℓ is the maximized log likelihood of the parameter vector $\Omega = (\theta, \lambda, \alpha)$, p is the number of estimated parameters and n is the number of observations. The best fitted model is selected based on minimum value obtained through the information criteria measures.

Dataset 1:

"The data set shown below represents the civil engineering data with 85 hailing times, previously used by Kotz and Dorp (2004):"

4.79, 4.75, 5.40, 4.70, 6.50, 5.30, 6.00, 5.90, 4.80, 6.70, 6.00, 4.95, 7.90, 5.40, 3.50, 4.54, 6.90, 5.80, 5.40, 5.70, 8.00, 5.40, 5.60, 7.50, 7.00, 4.60, 3.20, 3.90, 5.90, 3.40, 5.20, 5.90, 4.40, 5.20, 7.40, 5.70, 6.00, 3.60, 6.20, 5.70, 5.80, 5.90, 6.00, 5.15, 6.00, 4.82, 5.90, 6.00, 7.30, 7.10, 4.73, 5.90, 3.60, 6.30, 7.00, 5.10, 6.00, 6.60, 4.40, 6.80, 5.60, 5.90, 5.90, 8.60, 6.00, 5.80, 5.40, 6.50, 4.80, 6.40, 4.15, 4.90, 6.50, 8.20, 7.00, 8.50, 5.90, 4.40, 5.80, 4.30, 5.10, 5.90, 4.70, 3.50, 6.80.

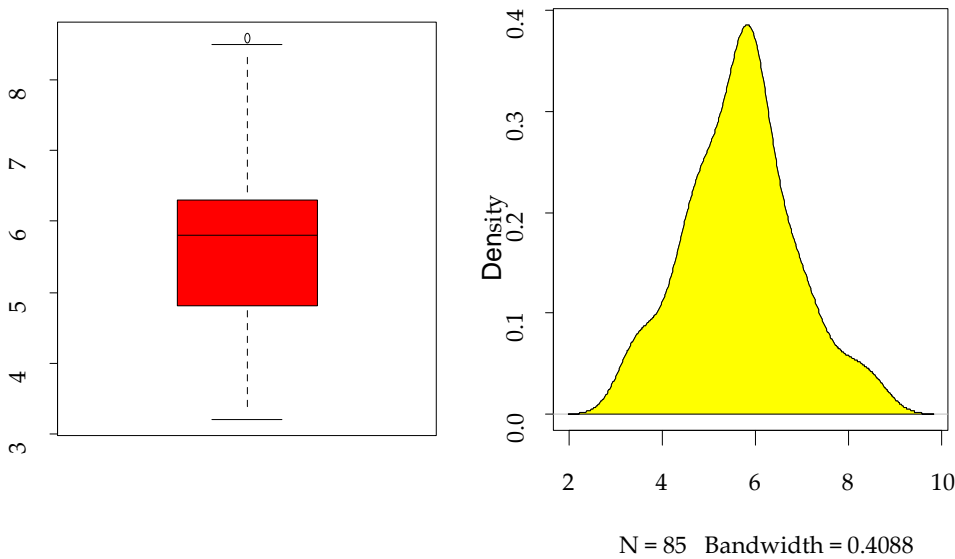


Figure 3: The boxplot and kernel density of the data set 1

Dataset 2:

"The data set shown below represents the strength of carbon fibers tested under tension at gauge lengths of 10mm, previously used Bi and Gui (2017):"

1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

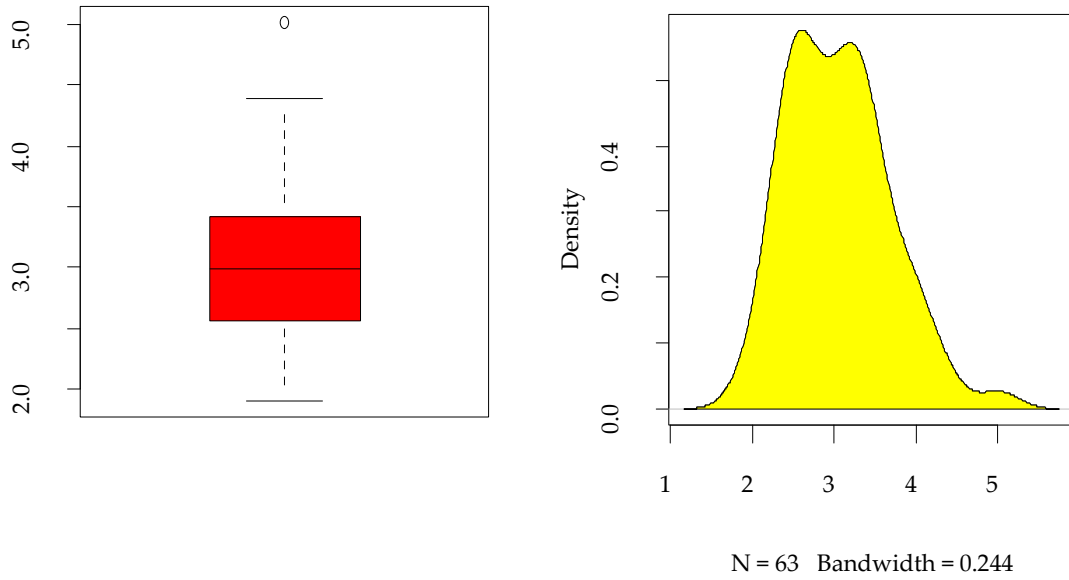


Figure 4: The boxplot and kernel density of the data set 2

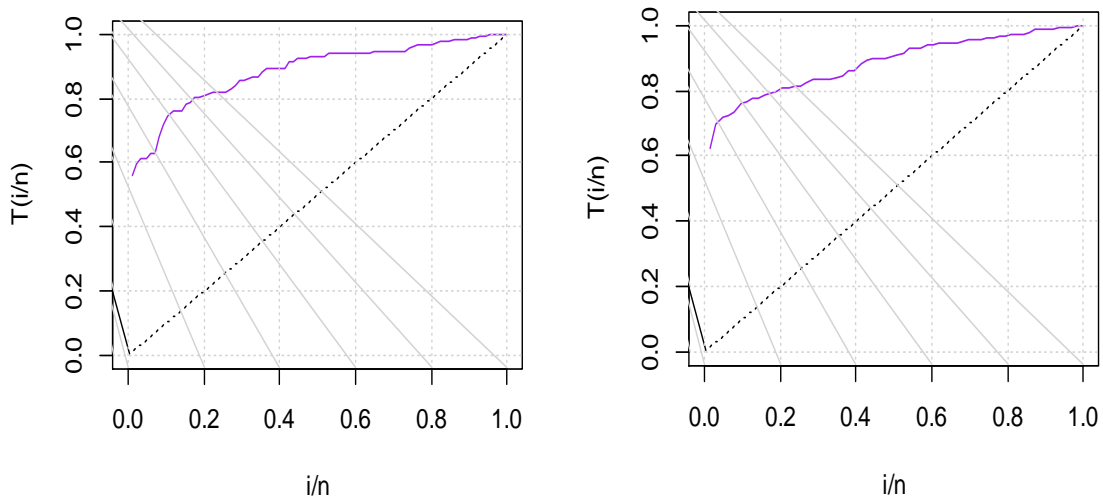


Figure 5: The TTT plot of data set 1 and 2

Table 3: MPSs, Log-likelihoods and Goodness of Fits Statistics for the Data Set 1

Distributions	λ	α	θ	β	LL	AIC	CAIC	BIC	HQIC
CMO-W	3.9229	2.2488	2.8840		-110.2788	226.5576	226.8539	233.8856	229.5051
CTL-W	0.0026	1.3745		3.2049	-132.9372	271.8744	272.1707	279.2024	274.8219
ECS-W	3.8988	0.0382	0.7231		-258.3757	522.7514	523.0477	530.0794	525.6989
NACos-W		4.8815	3.1622	0.0026	-193.4512	392.9024	393.1987	400.2304	395.8499
C-W		2.8953	0.0094	0.1471	-138.7963	283.5926	283.8889	290.9206	286.5401

Table 4: MPSs, Log-likelihoods and Goodness of Fits Statistics for the Data Set 2

Distributions	λ	α	θ	β	LL	AIC	CAIC	BIC	HQIC
MO-W	4.4779	6.1830	0.0415		-82.1587	170.3174	170.7242	176.7468	172.8461
CTL-W	0.3398	12.5851		1.4887	-86.6096	179.2192	179.6260	185.6486	181.7479
ECS-W	0.0027	0.9870	4.4644		-84.5562	175.1124	175.5192	181.5418	177.6411
NACos-W		8.1270	0.0128	2.6986	-85.2837	176.5674	176.9742	182.9968	179.0961
C-W		0.5119	6.9721	0.0020	-86.0652	178.1304	178.5372	184.5598	180.6591

IV. Discussion

We introduce a novel Cosine Marshall-Olkin family of distribution and its properties, therein, we extended the Weibull distribution to form a new sub-model known as Cosine Marshall-Olkin Weibull distribution. We conducted a comprehensive study of the new Cosine Marshall-Olkin Weibull distribution properties. Furthermore, we investigate the consistency and efficiency of the estimates obtained from the parameters of the novel distribution. We employ the maximum products of spacing estimation technique, which enabled us to access the values of the parameters effectively. To demonstrate the applicability of the proposed distribution, we provide insights on its performance using two real-life datasets. The analysis reveals that the new model outperforms other trigonometric family of distribution with the same baseline.

References

- [1] Bourguignon M, Silva R. B, Cordeiro G. M. (2014). The Weibull-G family of probability distributions. *J Data Sci.*12:1253–68.
- [2] Gupta R. D, Kundu D. Generalized exponential distribution. *Australian N Zeal J Stat.* 1999;41(2):173–88.
- [3] Al-Shomrani A, Arif O, Shawky K, Hanif S, Shahbaz M. Q. ToppLeone family of distributions: some properties and application. *Pak J Stat Oper Res.* 2016;12:443–51.
- [4] Cordeiro, G. M., Alizadeh, M. and Marinho, E. P. R. D. (2015). The Type I half logistic family of distributions. *Journal of Statistical Computation and Simulation*, 86, 707-728.
- [5] Bantan R. A, Jamal F, Chesneau C, Elgarhy M. (2019). A new power Topp-Leone generated family of distributions with applications. *Entropy.* 21(12)
- [6] Hassan, A. S., Elgarhy, M. and Shakil, M. (2017). Type II Half Logistic Family of Distributions with Applications. *Pakistan Journal of Statistics and Operation Research*, 13, 245-264.
- [7] Bantan R. A, Jamal F, Chesneau C, Elgarhy M. (2019). Truncated inverted Kumaraswamy generated family of distributions with applications. *Entropy.* 21(11):1–22.
- [8] Elbatal I, Ahmed Z, Elgarhy M, Almarashi A. M. (2019) A new alpha power transformed family of distributions: properties and applications to the Weibull model. *J Nonlinear Sci Appl.* 11:1099–112.

- [9] Ahmad Z, Elgarhy M, Hamedani GG, Butt N. S. (2020). Odd generalized N–H generated family of distributions with application to exponential model. *Pak J Stat Oper Res.*16:53–71.
- [10] Bantan R. A, Jamal F, Chesneau C, Elgarhy M. (2020). Type II power Topp-Leone generated family of distributions with applications. *Symmetry*, 12(1):1–22.
- [11] Suleiman, A.; Othman, M.; Ishaq, A.; Daud, H.; Indawati, R.; Abdullah, M.L.; Husin, A. (2023). The Odd Beta Prime-G Family of Probability Distributions: Properties and Applications. In *Proceedings of the 1st International Online Conference on Mathematics and Applications*, Online, 1–15 May 2023; MDPI: Basel, Switzerland.
- [12] Kumar D, Singh U, Singh S. K. (2015). A new distribution using sine function its application to bladder cancer patients data. *J Stat Appl Probability.*4(3):417–27.
- [13] Souza L. (2015) New trigonometric classes of probabilistic distributions, Thesis. Universidade Federal Rural de Pernambuco.
- [14] Isa A. M., Sule O. B., Ali B. A., Akeem A. A., and Ibrahim I. I. (2022). Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset. *UMYU Scientifica*, (1), 127 – 131.
- [15] Faruk M. U., Isa, A. M. Kaigama A. (2024) Sine-Weibull Distribution: Mathematical Properties and Application To Real Datasets. *Reliability: Theory & Applications.* 1(77): 65-72
- [16] Adepoju, A. A., Abdulkadir, S. S., & Jibasen, D. (2023). The Type I Half Logistics-Topp-Leone-G Distribution Family: Model, its Properties and Applications. *UMYU Scientifica*, 2(4), 09-22.
- [17] Bello O., A., Doguwa S., I., Yahaya A., Jibril H., M. (2021). A Type II Half Logistic Exponentiated-G Family of Distributions with Applications in Survival Analysis, *FUDMA Journal of Science*, 5(3):177-190.
- [18] Isa A. M., Kaigama A., Adepoju A. A., Bashiru S. O., Lehmann Type II-Lomax Distribution: Properties and Application to Real Data Set. (2023). *Communication in Physical Sciences*, 9(1):63 – 72.
- [19] Adepoju A. A., Abdulkadir S. S., Jibasen D, Olumoh J. S. (2024). A Type I Half Logistic Topp-Leone Inverse Lomax Distribution with Pplications In Skinfolds Analysis. *Reliability: Theory & Applications.* March 1(77): 618-630.
- [20] Osi, A. A., Doguwa, S. I., Abubakar , Y., Zakari, Y., & Abubakar , U. (2024). Development of Exponentiated Cosine Topp-Leone Generalized Family of Distributions and its Applications to Lifetime Data. *UMYU Scientifica*, 3(1), 157–167. <https://doi.org/10.56919/usci.2431.017>
- [21] Al-Babtain A. A., Elbatal I., Chesneau C., Elgarhy M. (2020). Sine Topp-Leone-G family of distributions: Theory and applications. *DE GRUYTER. Open Physics*; 18: 574–593.
- [22] Nanga S., Nasiru S., Diogbhan J. (2023). Cosine Topp–Leone family of distributions: Properties and Regression, *Research in Mathematics*, 10:1, 2208935, DOI: 10.1080/27684830.2023.2208935
- [23] Benchiha, S., Sapkota, L. P., Al Mutairi, A., Kumar, V., Khashab, R. H., Gemeay, A. M., & Nassr, S. G. (2023). A New Sine Family of Generalized Distributions: Statistical Inference with Applications. *Mathematical and Computational Applications*, 28(4), 1- 19.
- [24] Famoye, A., Algarni, A. & Almarashi, A. M. (2021). Sine Inverse Lomax Generated Family of Distributions with Applications. *Mathematical Problems in Engineering*, 1-11.
- [25] Muhammad, M., Alshanbari, H. M., Alanzi, A. R., Liu, L., Sami, W., Chesneau, C., Jamal, F. (2021). A new generator of probability models: the exponentiated sine-G family for lifetime studies. *Entropy*, 23(11), 1-30.
- [26] Sakthivel, K. M., & Rajkumar, J. (2021). Transmuted sine-G family of distributions: theory and applications. *Statistics and Applications*, 20(2), 73-92.
- [27] Nanga, S., Nasiru, S., Diogbhan, J. (2022). Tangent Topp-Leone Family of Distributions. *Scientific African*, 17, e01363.

- [28] Alkhairy, I., Nagy, M., Muse, A. H., Hussam, E. (2021). The Arctan-X family of distributions: Properties, Simulation, and Applications to Actuarial Sciences. *Complexity*, 1-14.
- [29] He, W., Ahmad, Z., Afify, A. Z. & Goual, H. (2020). The arcsine exponentiated-X family: validation and insurance application. *Complexity*, 1-18.
- [30] Jamal, F., Chesneau, C. & Aidi, K. (2021). The Sine Extended odd Frechet-G family of distribution with applications to Complete and Censored Data. *Mathematica Slovaca*, 71(4), 961-982.
- [31] Ferreira, T. A. (2021). Tan-G class of trigonometric distributions and its applications. *Cubo (Temuco)*, 23(1), 1-20.
- [32] Kumar, D., Singh, U. & Singh, S. K. (2015). A new distribution using Sine function its application to Bladder Cancer Patients" Data. *Journal of Statistics Applications and Probability*, 4(3), 417.
- [33] A. M. Isa, S. I. Doguwa , B. B. Alhaji , H. G. Dikko Sine Type II Topp-Leone G Family of Probability Distribution: Mathematical Properties and Application
- [34] A new method for adding a parameter to a family of distributions with application to the exponential and weibull families *Biometrika*, 84 (3) (1997), pp. 641-652
- [35] Souza, L., Junior, W. R., Brito, C. C., & Chesneau, C. (2019). General properties for the Cos-G class of distributions with applications. *Eurasian Bulletin of Mathematics*, 2(2), 63–79.
- [36] Nanga, S., Nasiru, S., & Diogban, J. (2023). Cosine Topp-Leone family of distributions: Properties and Regression. *Research in Mathematics*, 10(1), 2208935.
- [37] Muhammad, M., Bantan, R. A. R., Liu, L., Chesneau, C., Tahir, M. H., Jamal, F., & Elgarhy, M. (2021). A New Extended Cosine-G Distributions for Lifetime Studies. *Mathematics*, 9(21), 2758.
- [38] Alghamdi, A. S., & Abd El-Raouf, M. M. (2023). A New Alpha Power Cosine-Weibull Model with Applications to Hydrological and Engineering Data. *Mathematics*, 11(3), 673.
- [39] Souza, L., Junior, W. R., Brito, C. C., & Chesneau, C. (2019). General properties for the Cos-G class of distributions with applications. *Eurasian Bulletin of Mathematics*, 2(2), 63–79.

ANALYSIS OF MMAP/PH1,PH2/1 PREEMPTIVE PRIORITY INVENTORY RETRIAL QUEUEING SYSTEM WITH SINGLE VACATION, WORKING BREAKDOWN, REPAIR AND CLOSEDOWN

G. AYYAPPAN, S. MEENA



Department of Mathematics,
Puducherry Technological University,
Puducherry, India.

ayyappan@ptuniv.edu.in, meenasundar2296@gmail.com

Abstract

This paper analyzes preemptive priority inventory retrial queueing system with a single vacation, working breakdown, repair, and closedown. We assume that an arrival follows the Marked Markovian arrival process and that the server will provide them with phase-type services. The (s, S) policy to replenish the items and the replenishing duration follow an exponential distribution. In this paper, we consider two types of customers: high-priority(HP) customers and low-priority(LP) customers. Arriving HP customers should get the service if the server is idle and has a positive inventory level; otherwise, they should wait in front of the service station. Arriving LP customers get service only if there is a positive inventory level and there are no high-priority customers in the system; otherwise, go for the finite capacity size of the orbit. After the completion of service, if no one is present in the high-priority queue and orbit, the server will close down the system and then go on a single vacation. The server is idle when the vacation period ends. When the server breaks down, it only serves the present customer and operates in slow mode while it is being repaired. The number of high-priority customers in the system, the number of low-priority customers in the orbit, the inventory level, and server status may all be determined in a steady state. Numerous key performance indicators are defined, and a cost analysis is obtained. To make our mathematical concept clearer, a few numerical examples are provided.

Keywords: Queueing-inventory, (s, S) policy, Retrial, Preemptive Priority, Single Vacation, Working Breakdown, Repair, Closedown, Markovian Arrival Process, Phase-type distribution, Matrix Analytic Method.

AMS Subject Classification (2010): 60K30, 68M20, 90B05.

1. INTRODUCTION

The field of inventory retrial queueing systems has seen a rise in popularity in recent years due to developments in computer networking and communications technologies. In a queueing-inventory model, each client receives a product from the inventory upon completion of the service. Neuts [19] presented the modified Markovian point process for the first time. A number of well-known techniques fall under the large category of point processes known as MAP, including PH-renewal, Markov-modulated Poissons, and Poisson. The Markovian arrival process with several correlated and non-correlated arrival types, as well as the phase-type distribution, were both extensively clarified by Chakravarthy [8]. Neuts [20] investigated the methods used in matrix-analytic queueing theory.

Reordering products in a queueing order-demand inventory system is best done using the techniques described by Melikov and Molchanov [16]. A study by Berman et al. [6] examined a

system for inventory control for a service center that uses one inventory item for each service rendered. According to their assumptions, there must always be a shortage of items for the queue to form, and demand and service rates are predictable and steady. Berman and Kim [7] developed two types of queuing-inventory models with service facilities. The first had an infinite one, while the other had a finite capacity for queuing. In their evaluation, Yadavalli et al. [25] made the assumption that reorders are readily available and that requests belong to a renewal process. The inventory system included a service station and an indefinite waiting area. Amirthakodi and Sivakumar [2] looked at retrial inventory queuing, in which there is a finite orbit size, a single server, and customer feedback. Sanjukta and Nabendu [21] looked into a carbon tax and an inventory queueing system with a partial replenishment strategy and a limited shelf life for perishable commodities.

Most of the time, it is believed that inventory and queueing models have not failed at service stations. In actuality, we regularly come across circumstances where service station malfunctions could occur. A server interruption inventory retry queueing system was covered by Krishnamoorthy et al. [11]. In their model, they took into consideration a (s, S) replenishment policy where the lead time and service time follow an exponential distribution, while the arrival follows a Poisson distribution. The retrial inventory queueing system with server failure was examined by Ushakumari [24]. When the server is processing requests or is idle, it could malfunction. If a server failure results in service disruption, users are placed into an infinitely long orbit, obliged to retry service after an arbitrary period, and so on unless the server is rendered inoperable.

The server could simply quit waiting for customers and remain unreachable in a variety of situations. It might also be completing other duties, such as maintenance or servicing more clients. Krishnamoorthy and Narayanan [12] considered a manufacturing inventory system including server vacations. They held that the manufacturing process adhered to the Markovian manufacturing method and that the service times for each customer were dispersed in a phase-type manner. The inventory queue for retrials with several vacations was analyzed by Suganya and Sivakumar [23]. They took into account a pair of servers and a limited orbit size capacity in their model. The retrial queueing system incorporating a single server, Bernoulli feedback, and vacation has been examined by Ayyappan and Gowthami [4]. They took into account both the client's arrival based on MAP and the server's service delivery based on PH distribution. Melikov et al. [17] examined the retrial queueing system that incorporates Poisson arrival, exponential service time, and delayed feedback. For their investigation, they employed both the (s, S) and (s, Q) replenishing policies.

An inventory queueing approach with MAP arrivals, PH offerings, and perishable goods was examined by Manuel et al. [15]. Additionally, they take into account their model, in which a positive customer advances one regular customer to the front of the line while a negative customer pushes one regular customer back. An inventory retrial queueing system involving two commodities was presented by Anbazhagan and Jeganathan [3]. They think of their model as having a core item and a supplement item. Jeganathan and Selvakumar [9] examined a queueing system for inventory that employed a traditional retry rate. In their work, they present an optional oscillatory client arrival procedure that is subject to Bernoulli testing and can pass through a waiting room or an infinite orbit. A two-component demand inventory retrial queueing system was examined by Abdul Reiyas and Jeganathan [1]. They took into account the (s, Q) replenishment policy while placing the order. The retrial inventory queueing model was examined by Jeganathan et al. [10] with two different kinds of clients. Mustapha and Majid [18] developed a two-phase production period production inventory model for non-immediately degrading products. For mixed demand with trade credit programs, Manisha et al. [14] have created the ideal replacement and conservation investment strategy. Ayyappan and Archana [5] discussed the non-preemptive priority queueing model with optional service and single vacation.

2. DESCRIPTION OF THE MODEL

We take into consideration a single server with the preemptive priority inventory retrieval queueing system, featuring inventory with maximum storage capacity of S units. Customers arrive via the Marked Markovian Arrival Process (MMAP) with depictions (D_0, D_1, D_2) with order size k , where the matrix $D = D_0 + D_1 + D_2$. In the system, no arrival is governed by the square matrix D_0 , the arrival of high priority customers is governed by the square matrix D_1 and the arrival of low priority customers is governed by the square matrix D_2 . With π representing the probability vector of D , the mean arrival rate of HP customers is $\lambda_1 = \pi D_1 e_k$ and the mean arrival rate of LP customers is $\lambda_2 = \pi D_2 e_k$. The server provides high priority (HP) and low priority (LP) services that both follow a PH-distribution with depictions (γ, P) and (ν, M) of orders l_1 and l_2 , respectively. For HP and LP customers mean service rate is $\zeta_1 = [\gamma(-P)^{-1}e_{l_1}]^{-1}$ and $\zeta_2 = [\nu(-M)^{-1}e_{l_2}]^{-1}$.

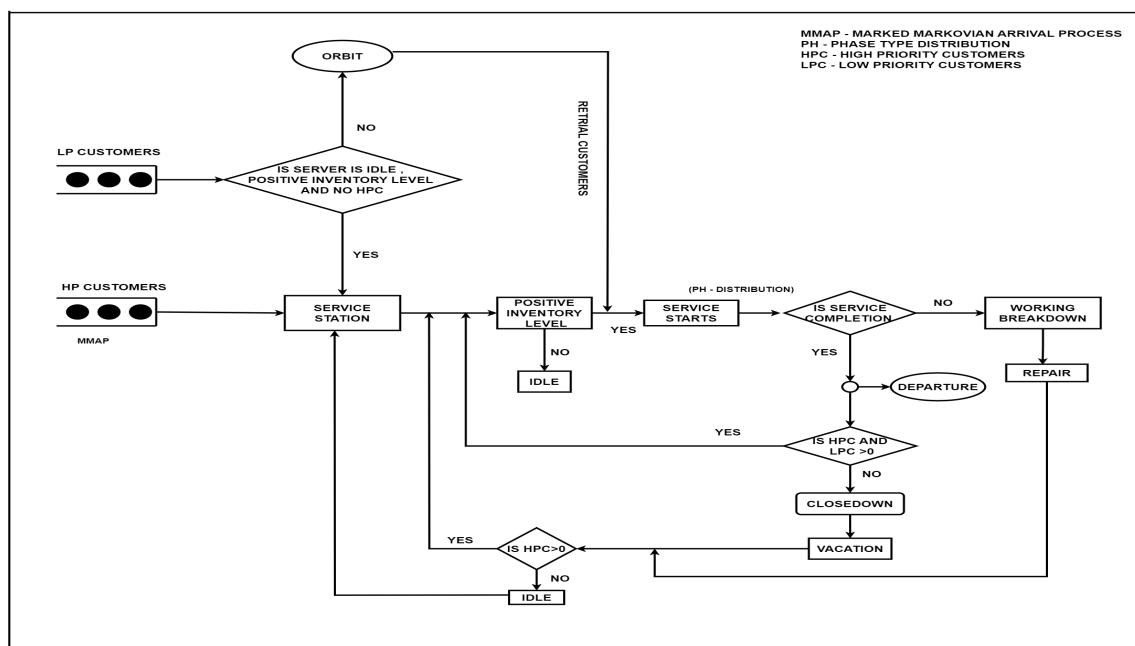


Figure 1: A pictorial illustration of the model

If the server breaks down while serving HP or LP customers, it will first offer a slow service mode to the impacted customers before beginning the repair procedure. The PH-distribution is followed by the slower service for HP and LP customers, together with a representation of order l_1 and l_2 , respectively, represented by $(\gamma_1, \theta P)$ and $(\nu_1, \theta M)$. The breakdown time has an exponential distribution with parameter σ , and the repair process has a PH-distribution with a depiction (α, U) of order m_2 . When HP customers arrive, they only interrupt their regular service if LP customer services are still in progress, and the server serves HP clients. In the event that there are no pending requests in the HP queue, the server will serve LP customers. Arriving HP customers should get the service if the server is idle and has a positive inventory level; otherwise, they should wait in front of the service station. Arriving LP customers get service only if there is a positive inventory level and there are no high-priority customers in the system; otherwise, go for the finite capacity size of the orbit, say N . After the completion of service, if no one is present in the high-priority queue and orbit, the server will close down the system and then go on a single vacation. After the completion of the vacation period, the server is idle. The closedown times follow an exponential distribution with parameter δ . The Vacation times follow the PH-distribution with depiction (β, W) of order m_1 . The LP customers retrying for their service after the fixed times, the constant retrial rate follow an exponential distribution with parameter χ .

The average rate of repair and vacation is given by η and ψ respectively.

3. THE QUASI BIRTH AND DEATH PROCESS FOR THE MATRIX GENERATIONS

We are going to discuss this part, which comprises the notation that forms the basis of the Quasi Birth and Death (QBD) process in our model.

- \otimes - Any two different order matrices can be multiplied to create a Kronecker product, and this can be founded on the research of Steeb and Hardy [22].
- \oplus - The Kronecker sum is the sum of any two of the different orders of matrices.
- I_k - The identity matrix has k dimensions.
- e - The column vector's appropriate dimension for each of its elements is 1.
- e_k - For every k elements in a column vector, the value is 1.
- $e_k(L)$ - The column vector with dimension L , where the k^{th} element is 1 and remaining elements are 0.
- $e_k'(L)$ - The transpose of $e_k(L)$.
- The arrival rate of HP and LP customers is represented by λ_i and described as $\lambda_i = \pi D_i e_k$, where $i=1,2$ respectively.
- The service rate for HP customers is represented by ζ_1 and described as $\zeta_1 = [\gamma(-P)^{-1}e_{l_1}]^{-1}$.
- The service rate for LP customers is represented by ζ_2 and described as $\zeta_2 = [v(-M)^{-1}e_{l_2}]^{-1}$.
- The vacation rate of the server is represented by ψ and described as $\psi = [\beta(-W)^{-1}e_{m_1}]^{-1}$.
- The server's rate of repair is represented by η and described as $\eta = [\alpha(-U)^{-1}e_{m_2}]^{-1}$.
- The number of HP customers in the system at time t can be represented by $N_1(t)$.
- The number of LP customers in the orbit at time t can be represented by $N_2(t)$.
- Let $V(t)$ be the state of the server at time t .

$$V(t) = \begin{cases} 0, & \text{the vacation state of the server,} \\ 1, & \text{the idle state of the server,} \\ 2, & \text{the server is offering service for HP customers,} \\ 3, & \text{the server is offering service for LP customers,} \\ 4, & \text{the server is offering slow service for HP customers,} \\ 5, & \text{the server is offering slow service for LP customers,} \\ 6, & \text{the server is under repair,} \\ 7, & \text{the server is under closedown process.} \end{cases}$$

- Let $I(t)$ be the level of inventory items at time t .
- $J_1(t)$ denotes the phases of the vacation process.
- $J_2(t)$ denotes the phases of the repair process.
- $S(t)$ denotes the phases of the service process.
- $M(t)$ denotes the phases of the arrival process.

Let $\{N_1(t), N_2(t), V(t), I(t), J_1(t), J_2(t), S(t), M(t) : t \geq 0\}$ indicate the Continuous Time Markov Chain (CTMC) with state-level independent QBD processes. The state space is as follows:

$$\Phi = I(0) \cup_{u_1=1}^{\infty} I(u_1),$$

where

$$\begin{aligned}
 l(0) = & \{(0, u_2, 0, j, a_1, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq a_1 \leq m_1, 1 \leq c \leq k\} \\
 & \cup \{(0, u_2, 1, j, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq c \leq k\} \\
 & \cup \{(0, u_2, 3, j, d_2, c) : 0 \leq u_2 \leq N, 1 \leq j \leq S, 1 \leq d_2 \leq l_2, 1 \leq c \leq k\} \\
 & \cup \{(0, u_2, 5, j, d_2, c) : 0 \leq u_2 \leq N, 1 \leq j \leq S, 1 \leq d_2 \leq l_2, 1 \leq c \leq k\} \\
 & \cup \{(0, u_2, 6, j, a_2, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq a_2 \leq m_2, 1 \leq c \leq k\} \\
 & \cup \{(0, u_2, 7, j, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq c \leq k\},
 \end{aligned}$$

for $u_1 \geq 1$,

$$\begin{aligned}
 l(u_1) = & \{(u_1, u_2, 0, j, a_1, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq a_1 \leq m_1, 1 \leq c \leq k\} \\
 & \cup \{(u_1, u_2, 1, 0, c) : 0 \leq u_2 \leq N, 1 \leq c \leq k\} \\
 & \cup \{(u_1, u_2, 2, j, d_1, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq d_1 \leq l_1, 1 \leq c \leq k\} \\
 & \cup \{(u_1, u_2, 4, j, d_1, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq d_1 \leq l_1, 1 \leq c \leq k\} \\
 & \cup \{(u_1, u_2, 5, j, d_2, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq d_2 \leq l_2, 1 \leq c \leq k\} \\
 & \cup \{(u_1, u_2, 6, j, a_2, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq a_2 \leq m_2, 1 \leq c \leq k\} \\
 & \cup \{(u_1, u_2, 7, j, c) : 0 \leq u_2 \leq N, 0 \leq j \leq S, 1 \leq c \leq k\}.
 \end{aligned}$$

The QBD procedure generates an infinitesimal matrix, as provided by

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & \dots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & 0 & 0 & \dots \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

The entries in Q's block matrices are specified as follows:

$$B_{00} = \begin{bmatrix} B_{00}^{11} & B_{00}^{12} & 0 & 0 & 0 & 0 \\ 0 & B_{00}^{22} & B_{00}^{23} & 0 & 0 & 0 \\ 0 & B_{00}^{32} & B_{00}^{33} & B_{00}^{34} & 0 & B_{00}^{36} \\ 0 & 0 & 0 & B_{00}^{44} & B_{00}^{45} & 0 \\ 0 & B_{00}^{52} & 0 & 0 & B_{00}^{55} & 0 \\ B_{00}^{61} & 0 & 0 & 0 & 0 & B_{00}^{66} \end{bmatrix},$$

where

$$\begin{aligned}
 B_{00}^{11} = & \begin{bmatrix} C_{001} & C_{002} & 0 & \dots & 0 & 0 \\ 0 & C_{001} & C_{002} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{001} & C_{002} \\ 0 & 0 & 0 & \dots & 0 & C_{001} + C_{002} \end{bmatrix}, \\
 C_{001} = & \begin{bmatrix} J_1 & 0 & \dots & 0 & 0 & \dots & J_3 \\ 0 & J_1 & \dots & 0 & 0 & \dots & J_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_1 & 0 & \dots & J_3 \\ 0 & 0 & \dots & 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_2 \end{bmatrix}, \quad C_{002} = I_{S+1} \otimes I_{m_1} \otimes D_2,
 \end{aligned}$$

$$B_{00}^{12} = I_{N+1} \otimes I_{S+1} \otimes W^0 \otimes I_k,$$

where $J_1 = W \oplus D_0 - \tau I_{m_1k}$, $J_2 = W \oplus D_0$, $J_3 = \tau I_{m_1k}$,

$$B_{00}^{22} = \begin{bmatrix} C_{003} & C_{004} & 0 & \dots & 0 & 0 \\ 0 & C_{005} & C_{004} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{005} & C_{004} \\ 0 & 0 & 0 & \dots & 0 & C_{006} \end{bmatrix},$$

$$C_{003} = \begin{bmatrix} J_4 & 0 & \dots & 0 & 0 & \dots & J_6 \\ 0 & J_4 & \dots & 0 & 0 & \dots & J_6 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_4 & 0 & \dots & J_6 \\ 0 & 0 & \dots & 0 & J_5 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_5 \end{bmatrix},$$

$$C_{005} = \begin{bmatrix} J_4 & 0 & \dots & 0 & 0 & \dots & J_6 \\ 0 & J_7 & \dots & 0 & 0 & \dots & J_6 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_7 & 0 & \dots & J_6 \\ 0 & 0 & \dots & 0 & J_8 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_8 \end{bmatrix},$$

$$C_{006} = \begin{bmatrix} J_9 & 0 & \dots & 0 & 0 & \dots & J_6 \\ 0 & J_{10} & \dots & 0 & 0 & \dots & J_6 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{10} & 0 & \dots & J_6 \\ 0 & 0 & \dots & 0 & J_{11} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{11} \end{bmatrix},$$

$$C_{004} = \begin{bmatrix} e_1'(S+1) \otimes D_2 \\ 0 \end{bmatrix},$$

where $J_4 = D_0 - \tau I_k$, $J_5 = D_0$, $J_6 = \tau I_k$, $J_7 = D_0 - (\chi + \tau)I_k$, $J_8 = D_0 - \chi I_k$,

$J_9 = (D_0 + D_2) - \tau I_k$, $J_{10} = (D_0 + D_2) - (\chi + \tau)I_k$, $J_{11} = (D_0 + D_2) - \chi I_k$,

$$B_{00}^{23} = \begin{bmatrix} C_{007} & 0 & \dots & 0 & 0 & 0 \\ C_{008} & C_{007} & \dots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & C_{008} & C_{007} & 0 \\ 0 & 0 & \dots & 0 & C_{008} & 0 \end{bmatrix},$$

$$C_{007} = \begin{bmatrix} 0 \\ I_S \otimes \nu \otimes D_2 \end{bmatrix}, \quad C_{008} = \begin{bmatrix} 0 \\ I_S \otimes \nu \otimes \chi I_m \end{bmatrix},$$

$$B_{00}^{33} = \begin{bmatrix} C_{009} & C_{0010} & 0 & \dots & 0 & 0 \\ 0 & C_{009} & C_{0010} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{009} & C_{0010} \\ 0 & 0 & 0 & \dots & 0 & C_{009} + C_{0010} \end{bmatrix},$$

$$C_{009} = \begin{bmatrix} J_{12} & 0 & \dots & 0 & 0 & \dots & J_{14} \\ 0 & J_{12} & \dots & 0 & 0 & \dots & J_{14} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{12} & 0 & \dots & J_{14} \\ 0 & 0 & \dots & 0 & J_{13} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{13} \end{bmatrix}, C_{0010} = I_S \otimes I_{l_2} \otimes D_2,$$

where $J_{12} = M \oplus D_0 - (\sigma + \tau)I_{l_2k}$, $J_{13} = M \oplus D_0 - \sigma I_{l_2k}$, $J_{14} = \tau I_{l_2k}$,

$$B_{00}^{34} = I_{N+1} \otimes I_S \otimes e_{l_2} \otimes v_1 \sigma I_k, \quad B_{00}^{32} = \begin{bmatrix} 0 & 0 \\ 0 & I_N \otimes C_{0011} \end{bmatrix},$$

$$B_{00}^{36} = [e_1(N+1) \otimes C_{0011} \quad 0], \quad C_{0011} = [I_S \otimes M^0 \otimes I_k \quad 0],$$

$$B_{00}^{44} = \begin{bmatrix} C_{0012} & C_{0010} & 0 & \dots & 0 & 0 \\ 0 & C_{0012} & C_{0010} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{0012} & C_{0010} \\ 0 & 0 & 0 & \dots & 0 & C_{0012} + C_{0010} \end{bmatrix},$$

$$C_{0012} = \begin{bmatrix} J_{15} & 0 & \dots & 0 & 0 & \dots & J_{14} \\ 0 & J_{15} & \dots & 0 & 0 & \dots & J_{14} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{15} & 0 & \dots & J_{14} \\ 0 & 0 & \dots & 0 & J_{16} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{16} \end{bmatrix},$$

where $J_{15} = \theta M \oplus D_0 - \tau I_{l_2k}$, $J_{16} = \theta M \oplus D_0$,

$$B_{00}^{45} = I_{N+1} \otimes C_{0013}, \quad C_{0013} = [I_S \otimes \theta M^0 \alpha \otimes I_k \quad 0],$$

$$B_{00}^{55} = \begin{bmatrix} C_{0014} & C_{0015} & 0 & \dots & 0 & 0 \\ 0 & C_{0014} & C_{0015} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{0014} & C_{0015} \\ 0 & 0 & 0 & \dots & 0 & C_{0014} + C_{0015} \end{bmatrix},$$

$$B_{00}^{52} = I_{N+1} \otimes I_{S+1} \otimes U^0 \otimes I_k,$$

$$C_{0014} = \begin{bmatrix} J_{17} & 0 & \dots & 0 & 0 & \dots & J_{19} \\ 0 & J_{17} & \dots & 0 & 0 & \dots & J_{19} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{17} & 0 & \dots & J_{19} \\ 0 & 0 & \dots & 0 & J_{18} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{18} \end{bmatrix}, \quad C_{0015} = I_{S+1} \otimes I_{m_2} \otimes D_2,$$

where $e J_{17} = U \oplus D_0 - \tau I_{m_2 k}$, $J_{18} = U \oplus D_0$, $J_{19} = \tau I_{m_2 k}$,

$$B_{00}^{66} = \begin{bmatrix} C_{0016} & C_{0017} & 0 & \dots & 0 & 0 \\ 0 & C_{0016} & C_{0017} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{0016} & C_{0017} \\ 0 & 0 & 0 & \dots & 0 & C_{0016} + C_{0017} \end{bmatrix},$$

$B_{00}^{61} = I_{N+1} \otimes I_{S+1} \otimes \beta \otimes \delta I_k$,

$$C_{0016} = \begin{bmatrix} J_{20} & 0 & \dots & 0 & 0 & \dots & J_6 \\ 0 & J_{20} & \dots & 0 & 0 & \dots & J_6 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{20} & 0 & \dots & J_6 \\ 0 & 0 & \dots & 0 & J_{21} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{21} \end{bmatrix}, \quad C_{0017} = I_{S+1} \otimes D_2,$$

where $e J_{20} = D_0 - (\delta + \tau) I_k$, $J_{21} = D_0 - \delta I_k$,

$$B_{01} = \begin{bmatrix} B_{01}^{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & B_{01}^{22} & B_{01}^{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & B_{01}^{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{01}^{45} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{01}^{56} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & B_{01}^{67} \end{bmatrix},$$

where e

$B_{01}^{11} = I_{N+1} \otimes I_{S+1} \otimes I_{m_1} \otimes D_1$, $B_{01}^{22} = I_{N+1} \otimes e_1(S+1) \otimes D_1$,

$B_{01}^{23} = I_{N+1} \otimes C_{011}$, $C_{011} = \begin{bmatrix} 0 \\ I_S \otimes \gamma \otimes D_1 \end{bmatrix}$,

$B_{01}^{33} = I_{N+1} \otimes I_S \otimes e_{l_2} \otimes \gamma \otimes D_1$, $B_{01}^{45} = I_{N+1} \otimes I_S \otimes I_{l_2} \otimes D_1$,

$B_{01}^{56} = I_{N+1} \otimes I_{S+1} \otimes I_{m_2} \otimes D_1$, $B_{01}^{67} = I_{N+1} \otimes I_{S+1} \otimes D_1$,

$$B_{10} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & B_{10}^{32} & 0 & 0 & 0 & B_{10}^{36} \\ 0 & 0 & 0 & 0 & B_{10}^{45} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where

$$B_{10}^{32} = \begin{bmatrix} 0 & 0 \\ 0 & I_N \otimes C_{101} \end{bmatrix}, B_{10}^{36} = \begin{bmatrix} C_{101} & 0 \\ 0 & 0 \end{bmatrix}, C_{101} = [I_S \otimes P^0 \otimes I_k \quad 0],$$

$$B_{10}^{45} = I_{N+1} \otimes C_{102}, C_{102} = [I_S \otimes \theta P^0 \alpha \otimes I_k \quad 0],$$

$$A_1 = \begin{bmatrix} A_1^{11} & A_1^{12} & A_1^{13} & 0 & 0 & 0 & 0 \\ 0 & A_1^{22} & A_1^{23} & 0 & 0 & 0 & 0 \\ 0 & 0 & A_1^{33} & A_1^{34} & 0 & 0 & 0 \\ 0 & 0 & 0 & A_1^{44} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_1^{55} & A_1^{56} & 0 \\ 0 & A_1^{62} & A_1^{63} & 0 & 0 & A_1^{66} & 0 \\ A_1^{71} & 0 & 0 & 0 & 0 & 0 & A_1^{77} \end{bmatrix},$$

where

$$A_1^{11} = B_{00}^{11}, A_1^{12} = I_{N+1} \otimes e_1(S+1) \otimes W^0 \otimes I_k,$$

$$A_1^{13} = I_{N+1} \otimes C_{111}, C_{111} = \begin{bmatrix} 0 \\ I_S \otimes W^0 \gamma \otimes I_k \end{bmatrix},$$

$$A_1^{22} = \begin{bmatrix} D_0 - \tau I_k & D_2 & 0 & \dots & 0 & 0 \\ 0 & D_0 - \tau I_k & D_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & D_0 - \tau I_k & D_2 \\ 0 & 0 & 0 & \dots & 0 & (D_0 + D_2) - \tau I_k \end{bmatrix},$$

$$A_1^{23} = I_{N+1} \otimes e'_S(S) \otimes \gamma \otimes \tau I_k, A_1^{34} = I_{N+1} \otimes I_S \otimes e_{l_1} \otimes \gamma_1 \otimes \sigma I_k,$$

$$A_1^{33} = \begin{bmatrix} C_{112} & C_{113} & 0 & \dots & 0 & 0 \\ 0 & C_{112} & C_{113} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{112} & C_{113} \\ 0 & 0 & 0 & \dots & 0 & C_{112} + C_{113} \end{bmatrix}, C_{113} = I_S \otimes I_{l_1} \otimes D_2,$$

$$C_{112} = \begin{bmatrix} J_{22} & 0 & \dots & 0 & 0 & \dots & J_{24} \\ 0 & J_{22} & \dots & 0 & 0 & \dots & J_{24} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{22} & 0 & \dots & J_{24} \\ 0 & 0 & \dots & 0 & J_{23} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{23} \end{bmatrix},$$

where $J_{22} = P \oplus D_0 - (\sigma + \tau)I_{l_1 k}$, $J_{23} = P \oplus D_0 - \sigma I_{l_1 k}$, $J_{24} = \tau I_{l_1 k}$,

$$A_1^{44} = \begin{bmatrix} C_{114} & C_{113} & 0 & \dots & 0 & 0 \\ 0 & C_{114} & C_{113} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & C_{114} & C_{113} \\ 0 & 0 & 0 & \dots & 0 & C_{114} + C_{113} \end{bmatrix},$$

$$C_{114} = \begin{bmatrix} J_{25} & 0 & \dots & 0 & 0 & \dots & J_{24} \\ 0 & J_{25} & \dots & 0 & 0 & \dots & J_{24} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & J_{25} & 0 & \dots & J_{24} \\ 0 & 0 & \dots & 0 & J_{26} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & J_{26} \end{bmatrix},$$

where $e J_{25} = \theta P \oplus D_0 - \tau I_{1k}$, $J_{26} = \theta P \oplus D_0$,

$$A_1^{55} = B_{00}^{44}, \quad A_1^{56} = B_{00}^{45}, \quad A_1^{66} = B_{00}^{55},$$

$$A_1^{62} = I_{N+1} \otimes e_1(S+1) \otimes U^0 \otimes I_k, \quad A_1^{63} = I_{N+1} \otimes C_{115},$$

$$C_{115} = \begin{bmatrix} 0 \\ I_S \otimes U^0 \gamma \otimes I_k \end{bmatrix},$$

$$A_1^{77} = B_{00}^{66}, \quad A_1^{71} = B_{00}^{61},$$

$$A_0 = \begin{bmatrix} A_0^{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_0^{22} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_0^{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_0^{44} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_0^{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_0^{66} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_0^{77} \end{bmatrix},$$

where $e A_0^{11} = B_{01}^{11}$, $A_0^{22} = I_{N+1} \otimes D_1$, $A_0^{33} = I_{N+1} \otimes I_S \otimes I_1 \otimes D_1$,

$$A_0^{44} = A_0^{33}, \quad A_0^{55} = B_{01}^{45}, \quad A_0^{66} = B_{01}^{56}, \quad A_0^{77} = B_{01}^{67},$$

$$A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_2^{32} & A_2^{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_2^{46} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where $e A_2^{32} = I_{N+1} \otimes e_1(S) \otimes P^0 \otimes I_k$, $A_2^{33} = I_{N+1} \otimes C_{211}$,

$$C_{211} = \begin{bmatrix} 0 & 0 \\ I_{S-1} \otimes P^0 \gamma \otimes I_k & 0 \end{bmatrix}, \quad A_2^{46} = B_{10}^{45}.$$

4. STATIONARY ANALYSIS

We analyze our model in a few consistent system configurations.

4.1. Criteria for stability

Let us define the matrix A as follows: $A = A_0 + A_1 + A_2$, signifying that it is an irreducible infinitesimal generator matrix with dimensions of $((N+1)(S+1)m_1k + (N+1)k + 2(N+1)Sl_1k + (N+1)Sl_2k + (N+1)(S+1)m_2k + (N+1)(S+1)k)$.

The vector π represents the stationary probability vector of A that achieving the criteria $\pi A = 0$ and $\pi e = 1$. The vector π is divided by $\pi = (\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6) = (\pi_{000}, \pi_{001}, \dots, \pi_{00S}, \pi_{010}, \pi_{011}, \dots, \pi_{01S}, \dots, \pi_{0N0}, \pi_{0N1}, \dots, \pi_{0NS}, \pi_{100}, \pi_{110}, \dots, \pi_{0N0}, \pi_{201}, \dots, \pi_{20S}, \pi_{211}, \dots, \pi_{21S}, \dots, \pi_{2N1}, \dots, \pi_{2NS}, \pi_{301}, \dots, \pi_{30S}, \pi_{311}, \dots, \pi_{31S}, \dots, \pi_{3N1}, \dots, \pi_{3NS}, \pi_{401}, \dots, \pi_{40S}, \pi_{411}, \dots, \pi_{41S}, \dots, \pi_{4N1}, \dots, \pi_{4NS}, \pi_{500}, \pi_{501}, \dots, \pi_{50S}, \pi_{510}, \pi_{511}, \dots, \pi_{51S}, \dots, \pi_{5N0}, \pi_{5N1}, \dots, \pi_{5NS}, \pi_{600}, \pi_{601}, \dots, \pi_{60S}, \pi_{610}, \pi_{611}, \dots, \pi_{61S}, \dots, \pi_{6N0}, \pi_{6N1}, \dots, \pi_{6NS})$, where e π_0 has a dimension of $(N + 1)(S + 1)m_1k$, π_1 has a dimension of $(N + 1)k$, π_2 has a dimension of $(N + 1)Sl_1k$, π_3 has a dimension of $(N + 1)Sl_1k$, π_4 has a dimension of $(N + 1)Sl_2k$, π_5 has a dimension of $(N + 1)(S + 1)m_2k$ and π_6 has a dimension of $(N + 1)(S + 1)k$. When examining the Markov process within the framework of QBD, our model's stability should satisfy the essential and sufficient requirements of $\pi A_0 e < \pi A_2 e$. Upon performing certain algebraic reductions, the stability condition $\pi A_0 e < \pi A_2 e$ is determined to be

$$\begin{aligned} & \sum_{u_2=0}^N \sum_{j=0}^S \pi_{0u_2j} (e_{m_1} \otimes D_1 e_k) + \sum_{u_2=0}^N \pi_{1u_20} (D_1 e_k) + \sum_{u_2=0}^N \sum_{j=1}^S \pi_{2u_2j} (e_{l_1} \otimes D_1 e_k) \\ & + \sum_{u_2=0}^N \sum_{j=1}^S \pi_{3u_2j} (e_{l_1} \otimes D_1 e_k) + \sum_{u_2=0}^N \sum_{j=1}^S \pi_{4u_2j} (e_{l_2} \otimes D_1 e_k) + \sum_{u_2=0}^N \sum_{j=0}^S \pi_{5u_2j} (e_{m_2} \otimes D_1 e_k) \\ & + \sum_{u_2=0}^N \sum_{j=0}^S \pi_{6u_2j} (D_1 e_k) < \sum_{u_2=0}^N \sum_{j=1}^N \pi_{2u_2j} (P^0 \otimes e_k) + \sum_{u_2=0}^N \sum_{j=1}^N \pi_{3u_2j} (\theta P^0 \otimes e_k). \end{aligned}$$

4.2. Analysis of Stationary Probability Vector

Let ϕ represent the stationary probability vector for Q, and this is divided as $\phi = (\phi_0, \phi_1, \phi_2, \dots)$. Mention that ϕ_0 has a dimension of $(N + 1)(S + 1)m_1k + 2(N + 1)(S + 1)k + 2(N + 1)Sl_2k + (N + 1)(S + 1)m_2k$ and ϕ_1, ϕ_2, \dots have a dimension of $(N + 1)(S + 1)m_1k + (N + 1)k + 2(N + 1)Sl_1k + (N + 1)Sl_2k + (N + 1)(S + 1)m_2k + (N + 1)(S + 1)k$ and the vector ϕ satisfies $\phi Q = 0$ and $\phi e = 1$.

Additionally, after the stability requirement of the model is met, the stationary probability vector ϕ can be obtained by applying the following equation:

$$\phi_{u_1} = \phi_1 R^{u_1 - 1}, \quad u_1 \geq 1.$$

The matrix quadratic equation $R^2 A_2 + R A_1 + A_0 = 0$ is satisfied by the minimal non-negative solution R based on Neuts [20]. The matrix quadratic equation yields the rate matrix. The order of the rate matrix R is given by $((N + 1)(S + 1)m_1k + (N + 1)k + 2(N + 1)Sl_1k + (N + 1)Sl_2k + (N + 1)(S + 1)m_2k + (N + 1)(S + 1)k)$ and it fulfills the condition $R A_2 e = A_0 e$.

By solving the following equations, the sub vectors ϕ_0 and ϕ_1 can be determined.

$$\phi_0 B_{00} + \phi_1 B_{10} = 0,$$

$$\phi_0 B_{01} + \phi_1 (A_1 + R A_2) = 0,$$

Subject to the normalizing condition

$$\phi_0 e_0 + \phi_1 (I - R)^{-1} e_1 = 1,$$

where $e_0 = e_{(N+1)(S+1)m_1k+2(N+1)(S+1)k+2(N+1)Sl_2k+(N+1)(S+1)m_2k}$ and $e_1 = e_{(N+1)(S+1)m_1k+(N+1)k+2(N+1)Sl_1k+(N+1)Sl_2k+(N+1)(S+1)m_2k+(N+1)(S+1)k}$.

According to Latouche and Ramaswami [13], by utilizing important stages in the logarithmic reduction process, the R matrix can be produced analytically.

5. MEASURES OF SYSTEM PERFORMANCE

- Average size of the HP customers in the system

$$E_{sys} = \sum_{u_1=1}^{\infty} u_1 \phi_{u_1} e.$$

- Average size of the LP customers in the orbit

$$\begin{aligned} E_{orb} = & \sum_{u_1=0}^{\infty} \sum_{u_2=1}^N \sum_{j=0}^S \sum_{a_1=1}^{m_1} \sum_{c=1}^k u_2 \phi_{u_1 u_2 0 j a_1 c} + \sum_{u_2=1}^N \sum_{j=0}^S \sum_{c=1}^k u_2 \phi_{0 u_2 1 j c} \\ & + \sum_{u_1=1}^{\infty} \sum_{u_2=1}^N \sum_{c=1}^k u_2 \phi_{u_1 u_2 1 0 c} + \sum_{u_1=1}^{\infty} \sum_{u_2=1}^N \sum_{j=1}^S \sum_{d_1=1}^{l_1} \sum_{c=1}^k u_2 \phi_{u_1 u_2 2 j d_1 c} \\ & + \sum_{u_2=1}^N \sum_{j=1}^S \sum_{d_2=1}^{l_2} \sum_{c=1}^k u_2 \phi_{0 u_2 3 j d_2 c} + \sum_{u_1=1}^{\infty} \sum_{u_2=1}^N \sum_{j=1}^S \sum_{d_1=1}^{l_1} \sum_{c=1}^k u_2 \phi_{u_1 u_2 4 j d_1 c} \\ & + \sum_{u_1=0}^{\infty} \sum_{u_2=1}^N \sum_{j=1}^S \sum_{d_2=1}^{l_2} \sum_{c=1}^k u_2 \phi_{u_1 u_2 5 j d_2 c} + \sum_{u_1=0}^{\infty} \sum_{u_2=1}^N \sum_{j=0}^S \sum_{a_2=1}^{m_2} \sum_{c=1}^k u_2 \phi_{u_1 u_2 6 j a_2 c} \\ & + \sum_{u_1=0}^{\infty} \sum_{u_2=1}^N \sum_{j=0}^S \sum_{c=1}^k u_2 \phi_{u_1 u_2 7 j c}. \end{aligned}$$

- Expected size of the inventory items

$$\begin{aligned} E_{inv} = & \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{a_1=1}^{m_1} \sum_{c=1}^k j \phi_{u_1 u_2 0 j a_1 c} + \sum_{u_2=0}^N \sum_{j=1}^S \sum_{c=1}^k j \phi_{0 u_2 1 j c} \\ & + \sum_{u_1=1}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_1=1}^{l_1} \sum_{c=1}^k j \phi_{u_1 u_2 2 j d_1 c} + \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_2=1}^{l_2} \sum_{c=1}^k j \phi_{0 u_2 3 j d_2 c} \\ & + \sum_{u_1=1}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_1=1}^{l_1} \sum_{c=1}^k j \phi_{u_1 u_2 4 j d_1 c} + \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_2=1}^{l_2} \sum_{c=1}^k j \phi_{u_1 u_2 5 j d_2 c} \\ & + \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{a_2=1}^{m_2} \sum_{c=1}^k j \phi_{u_1 u_2 6 j a_2 c} + \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{c=1}^k j \phi_{u_1 u_2 7 j c}. \end{aligned}$$

- Expected reorder rate

$$\begin{aligned} E_R = & \sum_{u_2=0}^N \sum_{d_2=1}^{l_2} \sum_{c=1}^k \phi_{0 u_2 3(s+1) d_2 c} (M^0 \otimes I_k) e + \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{d_2=1}^{l_2} \phi_{u_1 u_2 5(s+1) d_2 c} (\theta M^0 \alpha \otimes I_k) e \\ & + \sum_{u_2=0}^N \sum_{d_1=1}^{l_1} \sum_{c=1}^k \phi_{1 u_2 2(s+1) d_1 c} (P^0 \otimes I_k) e + \sum_{u_1=2}^{\infty} \sum_{u_2=0}^N \sum_{d_1=1}^{l_1} \sum_{c=1}^k \phi_{u_1 u_2 2(s+1) d_1 c} (P^0 \gamma \otimes I_k) e \\ & + \sum_{u_1=1}^{\infty} \sum_{u_2=0}^N \sum_{d_1=1}^{l_1} \sum_{c=1}^k \phi_{u_1 u_2 2(s+1) d_1 c} (\theta P^0 \alpha \otimes I_k) e. \end{aligned}$$

- Probability for the vacation state of the server

$$P_{vac} = \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=0}^S \sum_{a_1=1}^{m_1} \sum_{c=1}^k \phi_{u_1 u_2 0 j a_1 c}.$$

- Probability for the idle state of the server

$$P_{idle} = \sum_{u_2=0}^N \sum_{j=0}^S \sum_{c=1}^k \phi_{0 u_2 1 j c} + \sum_{u_1=1}^{\infty} \sum_{u_2=0}^N \sum_{c=1}^k \phi_{u_1 u_2 1 0 c}.$$

- The probability that HP customers receive normal mode service from the server

$$P_{HNB} = \sum_{u_1=1}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_1=1}^{l_1} \sum_{c=1}^k \phi_{u_1 u_2 2 j d_1 c}.$$

- The probability that LP customers receive normal mode service from the server

$$P_{LNB} = \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_2=1}^{l_2} \sum_{c=1}^k \phi_{0 u_2 3 j d_2 c}.$$

- The probability that HP customers receive slow mode service from the server

$$P_{HSB} = \sum_{u_1=1}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_1=1}^{l_1} \sum_{c=1}^k \phi_{u_1 u_2 4 j d_1 c}.$$

- The probability that LP customers receive slow mode service from the server

$$P_{LSB} = \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=1}^S \sum_{d_2=1}^{l_2} \sum_{c=1}^k \phi_{u_1 u_2 5 j d_2 c}.$$

- Probability of the server is in repair process

$$P_{rep} = \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=0}^S \sum_{a_2=1}^{m_2} \sum_{c=1}^k \phi_{u_1 u_2 6 j a_2 c}.$$

- Probability of the server is in closedown process

$$P_{cd} = \sum_{u_1=0}^{\infty} \sum_{u_2=0}^N \sum_{j=0}^S \sum_{c=1}^k \phi_{u_1 u_2 7 j c}.$$

- The rate of effective retrials

$$\mathfrak{R} = \chi \sum_{u_2=1}^N \sum_{j=1}^S \sum_{c=1}^k \phi_{0 u_2 1 j c}$$

6. ANALYSIS OF COST FUNCTION

We have assumed that every cost factor (per unit of time) correlates to a distinct system measure while developing the expense function for our model.

- E_I - The cost of inventory for retaining each unit of goods.
- E_{H_1} - Keeping a HP customer's cost in the system for each unit of time.
- E_{H_2} - Keeping a LP customer's cost in the system for each unit of time.
- E_S - Initial costs for each order.

$$TC(s, S) = E_I E_{inv} + E_{H_1} E_{sys} + E_{H_2} E_{orb} + E_S E_R$$

7. NUMERICAL RESULTS

Using both numerical and graphical illustrations, we will be studying the behavior of the models in the section that follows. The next three are various MAP representations with the same mean value of 1 across all arrival processes. Chakra varthy [8] used these three arrival value sets as input data in their literature.

- **A-ER(Arrival in Erlang):**

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 \\ 1.2 & 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 & 0 \\ 0.8 & 0 \end{bmatrix}.$$

- **A-EX(Arrival in Exponential):**

$$D_0 = [-1], \quad D_1 = [0.6], \quad D_2 = [0.4].$$

- **A-HE(Arrival in Hyper-exponential):**

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1.026 & 0.114 \\ 0.1026 & 0.0114 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0.684 & 0.076 \\ 0.0684 & 0.0076 \end{bmatrix}.$$

The service, vacation, and repair processes each have three distinct phase-type distributions that we should take into consideration. We will use the notations X-ER, X-EX and X-HE respectively for Erlang, exponential and hyper-exponential cases dealing with X-type distribution where X = S, V, R depending on whether the services, vacations or repairs are under consideration.

- **X-ER(Erlang):**

$$\gamma = \nu = \beta = \alpha = (1, 0), \quad P = M = W = U = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}.$$

- **X-EX(Exponential):**

$$\gamma = \nu = \beta = \alpha = [1], \quad P = M = W = U = [1].$$

- **X-HE(Hyper-exponential):**

$$\gamma = \nu = \beta = \alpha = (0.8, 0.2), \quad P = M = W = U = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}.$$

7.1. Illustrative Example 1

We explored the effects of repair rate (η) versus the average size of HP customers in the system (E_{sys}). In order to attain system stability, we fix $\lambda = 1, \zeta_1 = 10, \zeta_2 = 8, \psi = 8, \sigma = 1, \tau = 5, \chi = 4, \delta = 4, \theta = 0.6, s = 3, S = 6, N = 5$.

- We combine the arrival and service time categories in Tables 1 through 3 to investigate the repair rate versus the average size of HP customers in the system.
- When the repair rate (η) rises, the corresponding the average size of HP customers in the system (E_{sys}) reduces.
- When comparing arrival times to all other arrivals, the E_{sys} drops quickly for hyper-exponential arrivals, and slowly for Erlang arrivals. Similarly, for service durations, the E_{sys} decreases more slowly in Erlang services than it does with hyper-exponential services.

7.2. Illustrative Example 2

We explored the effects of HP service rate (ζ_1) versus the Total Cost (TC) of the system. In order to attain system stability, we fix $\lambda = 1, \zeta_2 = 8, \psi = 8, \eta = 6, \sigma = 1, \tau = 5, \chi = 4, \delta = 4, \theta = 0.6, s = 3, S = 6, N = 5, E_I = 50, E_{H_1} = 200, E_{H_2} = 180, E_R = 220$.

- We combine the arrival and service time categories in Tables 4 through 6 to investigate the HP service rate versus the total cost of the system
- When the HP service rate (ζ_1) rises, the corresponding the total cost of the system (TC) reduces.
- When comparing arrival times to all other arrivals, the TC drops quickly for hyper-exponential arrivals, and slowly for Erlang arrivals. Similarly, for service durations, the TC decreases more slowly in Erlang services than it does with hyper-exponential services.

7.3. Illustrative Example 3

We explored the effects of retrial rate (χ) versus the average size of LP customers in the orbit (E_{orb}). In order to attain system stability, we fix $\lambda = 1, \zeta_1 = 10, \zeta_2 = 8, \psi = 8, \sigma = 1, \tau = 5, \eta = 6, \delta = 4,$

$\theta = 0.6, s = 3, S = 6, N = 5.$

- We combine the arrival and service time categories to investigate the rate of retrial versus the average size of LP customers in the orbit, using Figures 2 through 4.
- When the retrial rate (χ) rises, the corresponding average size of LP customers in the orbit (E_{orb}) reduces.
- When comparing arrival times to all other arrivals, the E_{orbit} drops quickly for hyper-exponential arrivals, and slowly for Erlang arrivals.

Table 1: Repair rate(η) vs E_{sys} - X-ER

X-ER			
η	A-ER	A-EX	A-HE
6.0	0.10813513	0.12190924	0.14286803
6.5	0.10786026	0.12152199	0.14226519
7.0	0.10763579	0.12120252	0.14176864
7.5	0.10744949	0.12093492	0.14135332
8.0	0.10729269	0.12070781	0.14100134
8.5	0.10715912	0.12051288	0.14069960
9.0	0.10704411	0.12034388	0.14043834
9.5	0.10694416	0.12019608	0.14021010
10.0	0.10685657	0.12006582	0.14000914
10.5	0.10677924	0.11995019	0.13983096

Table 2: Repair rate(η) vs E_{sys} - X-EX

X-EX			
η	A-ER	A-EX	A-HE
6.0	0.10901178	0.12413720	0.14698476
6.5	0.10867336	0.12369032	0.14629182
7.0	0.10839815	0.12332339	0.14572339
7.5	0.10817062	0.12301736	0.14524978
8.0	0.10797985	0.12275870	0.14484984
8.5	0.10781790	0.12253750	0.14450816
9.0	0.10767894	0.12234641	0.14421327
9.5	0.10755857	0.12217984	0.14395643
10.0	0.10745340	0.12203347	0.14373094
10.5	0.10736082	0.12190393	0.14353155

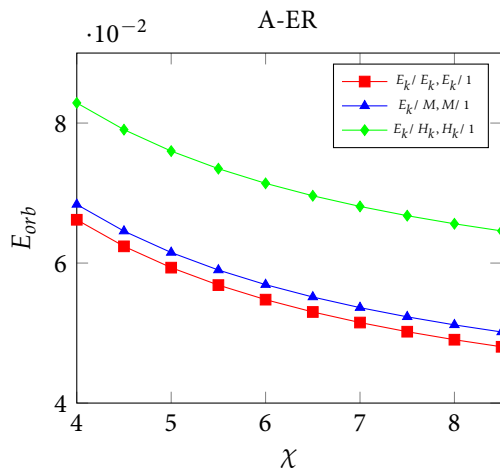


Figure 2: Retrial rate(χ) vs E_{orb} - A-ER

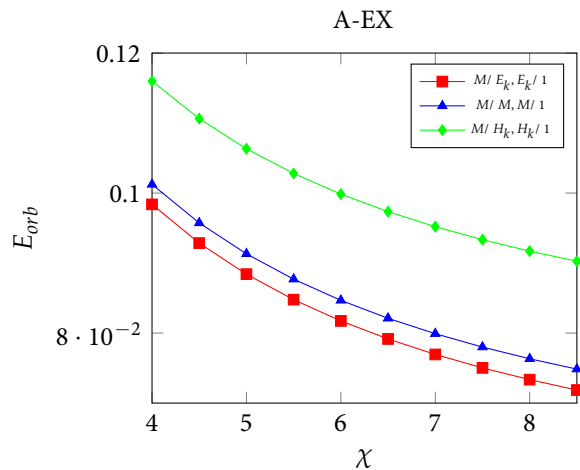


Figure 3: Retrial rate(χ) vs E_{orb} - A-EX

Table 3: Repair rate(η) vs E_{sys} - X-HE

η	X-HE		
	A-ER	A-EX	A-HE
6.0	0.11918175	0.13742625	0.16861900
6.5	0.11850292	0.13665603	0.16748394
7.0	0.11795708	0.13603251	0.16656203
7.5	0.11751074	0.13551942	0.16580136
8.0	0.11714044	0.13509118	0.16516511
8.5	0.11682931	0.13472934	0.16462655
9.0	0.11656496	0.13442027	0.16416588
9.5	0.11633814	0.13415374	0.16376814
10.0	0.11614181	0.13392192	0.16342189
10.5	0.11597051	0.13371874	0.16311819

Table 4: HP service rate(ζ_1) vs TC - X-ER

ζ_1	X-ER		
	A-ER	A-EX	A-HE
10	287.97326332	298.59933362	316.43100759
11	287.74939522	298.08848910	315.15693318
12	287.58577777	297.70317492	314.18547489
13	287.46248681	297.40469160	313.42429235
14	287.36716153	297.16827941	312.81429994
15	287.29182156	296.97748060	312.31615614
16	287.23113722	296.82099764	311.90277542
17	287.18144377	296.69085871	311.55497162
18	287.14015651	296.58130309	311.25881669
19	287.10541130	296.48808035	311.00398146

Table 5: HP service rate(ζ_1) vs TC - X-EX

ζ_1	X-EX		
	A-ER	A-EX	A-HE
10	288.98571417	299.98654358	318.45145147
11	288.70159130	299.40909492	317.07970441
12	288.49311376	298.97229601	316.03161949
13	288.33556222	298.63313169	315.20927312
14	288.21349742	298.36397016	314.54968265
15	288.11689338	298.14638167	314.01074509
16	288.03902055	297.96767769	313.56338397
17	287.97523158	297.81888191	313.18695269
18	287.92223650	297.69349308	312.86643407
19	287.87765496	297.58670481	312.59067408

Table 6: HP service rate(ζ_1) vs TC - X-HE

ζ_1	X-HE		
	A-ER	A-EX	A-HE
10	295.87738733	307.40357235	328.41423547
11	295.31869946	306.58861243	326.77655347
12	294.89650447	305.96029062	325.50031090
13	294.56917993	305.46410267	324.48262461
14	294.30990735	305.06435501	323.65537379
15	294.10075941	304.73681683	322.97187661
16	293.92937439	304.46453218	322.39919670
17	293.78699896	304.23532420	321.91350697
18	293.66729077	304.04024987	321.49719185
19	293.56556118	303.87261147	321.13697796

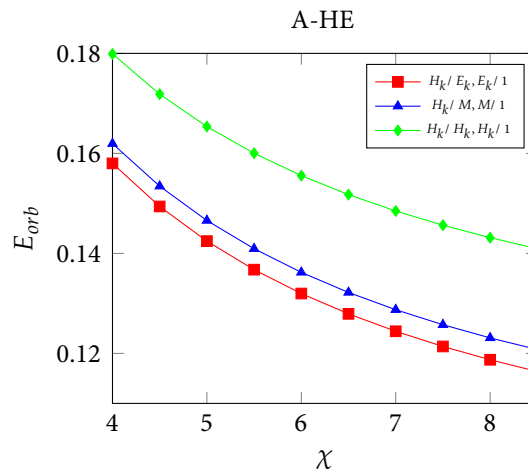


Figure 4: Retrial rate(χ) vs E_{orb} - A-HE

8. CONCLUSION

The present study investigated a retrial inventory queuing system that incorporates MMAP arrivals for HP and LP customers, services, vacations, and repairs, all of which follow phase-type distribution, (s, S) replenishment inventory policy, working breakdown, and closedown. We examined the system's stability criteria as well as the invariant probability vector. We analyzed the active period and also offered cost evaluations and system performance measures. Employing numeric values of arrivals and services in this model, we computed the average size of HP customers in the system for different values of repair rate and the total cost of the system for different values of service rate. The two-dimensional plots show the average size of LP customers in the orbit for different values of retrial rate. The average size of HP customers in the system for various values of vacation and service rates is depicted in the three-dimensional graphs. Every table and graph shows the stability of the system. We also expand our research to include multi-server systems with two commodity inventory queueing systems.

REFERENCES

- [1] Abdul Reiyas, M. and Jeganathan, K. (2023). A classical retrial queueing inventory system with two component demand rate, *International Journal of Operational Research*, 47(4):508–533.
- [2] Amirthakodi, M. and Sivakumar, B. (2015). An inventory system with service facility and finite orbit size for feedback customers, *OPSEARCH*, 52(2):225–255.
- [3] Anbazhagan, N. and Jeganathan, K. (2013). Two-commodity Markovian inventory system with compliment and retrial demand, *British Journal of Mathematics and Computer Science*, 3(2):115–134.
- [4] Ayyappan, G. and Gowthami, R. (2019). Analysis of MAP/PH/1 retrial queue with constant retrial rate, Bernoulli schedule vacation, Bernoulli feedback, breakdown and repair, *Reliability: Theory and Applications*, 14(2):86–103.
- [5] Ayyappan, G. and Archana Gurulakshmi, G. (2023). Analysis of MMAP/PH/1 Classical Retrial Queue with Non-preemptive priority, Second optional service, Differentiate breakdowns, Phase type repair, Single vacation, Emergency vacation, Closedown, Setup and Discouragement, *Reliability: Theory and Applications*, 18(3):528:551.
- [6] Berman, O., Kaplan, E. H. and Shimshak, D. G. (1993). Deterministic approximations for inventory management at service facilities, *IIE Transactions*, 25(5):98–104.
- [7] Berman, O. and Kim, E. (1999). Stochastic models for inventory management at service facility, *Communications in Statistics. Stochastic Models*, 15(4):695–718.
- [8] Chakra varthy, S. R. (2010). Markovian arrival process, Wiley Encyclopaedia of Operations Research and Management, John Wiley and Sons, Inc., USA.
- [9] Jeganathan, K. and Selvakumar, S. (2022). An optional arrival process of the oscillator y demands in the inventory queueing system with queue dependent service rate, *International Journal of Mathematics in Operational Research*, 22(2):162–194.
- [10] Jeganathan, K., Vidhya, S., Hema vathy, R., Anbazhagan, N., Joshi, G. P., Kang, C. and Seo, C. (2022). Analysis of M/M/1/N Stochastic Queueing- Inventory System with Discretionary Priority Service and Retrial Facility, *Sustainability*, 14(10), 6370.
- [11] Krishnamoorthy, A., Nair, S. S. and Narayanan, V. C. (2011). An inventory model with server interruptions and retrials, *Operational Research*, 12(22):151–171.
- [12] Krishnamoorthy, A. and Narayanan, V. C. (2011). Production inventory with service time and vacation to the server, *IMA Journal of Management Mathematics*, 22(1):33–45.
- [13] Latouche, G. and Ramaswami, V. (1999). Introduction of Matrix Analytic Methods in Stochastic Modeling, Society for Industrial and Applied Mathematics, Philadelphia.
- [14] Manisha Pant, Seema Sharma and Anand Chauhan (2022). Optimal replenishment and preservation investment policy for hybrid demand with trade credit schemes, *International Journal of Mathematics in Operational Research*, 23(2):232–258.

- [15] Manuel, P., Sivakumar, B. and Arivarignan, G. (2007). A perishable inventory system with service facilities, MAP arrivals and PH-Service times, *Journal of Systems Science and Systems Engineering*, 16(1):62–73.
- [16] Melikov, A. Z. and Molchanov, A. A. (1992). Stock optimization in transportation/storage systems, *Cybernetics and System Analysis*, 28(1):484–487.
- [17] Melikov, A., Aliyeva, S., Nair, S. S. and Kumar, B. K. (2022). Retrial Queuing-Inventory Systems with Delayed Feedback and Instantaneous Damaging of Items, *Axioms*, 11(5), 241.
- [18] Mustapha Lawal Malumfashi and Majid Khan Majahar Ali (2023). A production inventory model for non-instantaneous deteriorating items with two-phase production period, stock-dependent demand and shortages, *International Journal of Mathematics in Operational Research*, 24(2):173–193.
- [19] Neuts, M. F. (1979). A versatile Markovian point process, *Journal of Applied Probability*, 16(4):764–779.
- [20] Neuts, M. F. (1984). Matrix-analytic methods in queueing theory, *European Journal of Operational Research*, 15(1):2–12.
- [21] Sanjukta Malakar and Nabendu Sen (2023). An inventory model to study partial replacement policy and finite shelf life for deteriorating items with carbon tax, *International Journal of Mathematics in Operational Research*, 24(2):286–299.
- [22] Steeb, W. H. and Hardy, Y. (2011). Matrix Calculus and Kronecker Product: A Practical Approach to Linear and Multilinear Algebra, World Scientific Publishing, Singapore.
- [23] Suganya, C. and Sivakumar, B. (2019). MAP/PH(1), PH(2)/2 finite retrial inventory system with service facility, multiple vacations for servers, *International Journal of Mathematics in Operational Research*, 15(3):265–295.
- [24] Ushakumari, P. V. (2017). A retrial inventory system with an unreliable server, *International Journal of Mathematics in Operational Research*, 10(2):190–210.
- [25] Yada valli, V. S. S., Sivakumar, B. and Arivarignan, G. (2008). Inventory system with renewal demands at service facilities, *International Journal of Production Economics*, 114(1):252–264.

OPTIMIZATION OF AN INVENTORY MODEL FOR DETERIORATING ITEMS ASSUMING DETERIORATION DURING CARRYING WITH TWO-WAREHOUSE FACILITY

KRISHAN KUMAR YADAV¹, AJAY SINGH YADAV^{1,*}, SHIKHA BANSAL¹



¹Department of Mathematics, SRM Institute of Science and Technology,
Delhi-NCR Campus, Ghaziabad, India, 201204
krishankumaryadav222@gmail.com, ajaysiny@srmist.edu.in, shikhab@srmist.edu.in

*Corresponding author

Abstract

A common topic in the context of its application in today's business contexts is inventory modelling and management. It is well-known that deterioration has a big impact on inventory management. One of the most frequent supply chain concerns is the deterioration of items during transit from a supplier's storehouse to a retailer's storehouse. In light of this, a two-level supply chain inventory model for decaying goods is developed with two warehouse (storehouse) facilities for retailers, namely Owned Warehouse (OW) and Rented Warehouse (RW), assuming deterioration both during carrying from a supplier's storehouse to a retailer's storehouses and in the retailer's storehouses themselves. Also, we are assuming the selling price and time sensitive demand. We are developed this model under inflation. Shortages are not allowed. The main objective of this study is to determine the optimal ordering policy in order to maximize the retailer's profit per unit of time. The applicability of our suggested model is investigated using a numerical example and with the support of MATLAB programming software (version: R2021b). Sensitivity analysis is used to examine the effects of changing the values of system parameters. Graphical representations are also shown in this paper.

Keywords: Two-warehouse, Demand based on timing and selling price, Inflation, Deterioration during carrying and Optimization.

1. INTRODUCTION AND LITERATURE SURVEY

Design and production operations plays an important role in supply chain inventory management. Two-warehouse inventory management is a useful for optimising discrete item design and production operations. It enables producers to customise their manufacturing and distribution processes based on unique product quality, demand patterns, and lead times, resulting in increased operational efficiency and customer satisfaction. Deterioration is a key factor in both deterministic and probabilistic inventory models of the classical type. Profit changes anti-proportionally to the decline rate, meaning that if the deterioration rate rises, the retailer's profit falls, and if the deterioration rate falls, the retailer's profit rises. In the current analysis of an inventory model, the rate of deterioration cannot be disregarded. Deterioration is defined as the loss of the initial product's marginal values as well as damage, decay, disappearance, obsolescence and harm to utility. Poswal, P., et al. [47] and Mahata, S. and Debnath, B.K. [51] are also constructing an

inventory model based on certain novel assumptions. Kumar, A., et al. [49], Kumar, K., et al. [50], Kundu, T. and Islam, S. [52], Yusuf T.I., et al. [53] and Kumar, P., et al. [54] have also implemented optimisation approaches in various domains.

Inventory management is essential for preventing waste, maintaining product quality, and guaranteeing timely component delivery in the mechanical and electrical industries when it comes to perishable or deteriorating products. Here are a few specific applications for deteriorating goods in these industries: Temperature-controlled storage, management of humidity, First In First Out (FIFO), tracking of expiration dates, real-time monitoring, frequent quality checks, appropriate packaging, cooperation with suppliers, shortened storage durations, customised storage solutions, emergency response plan, waste reduction techniques, and continuous improvement. A study on the application of manufacturing in the Malaysian electrical and electronics industries was conducted by Wong, Y.C., et al. [55]. Basdere, B., et al. [56] Electronic and electrical product disassembly factories to recover resources in material and product cycles. Colledani, M., et al. [57] Manufacturing system design and management for superior product quality. Yusuf T., et al. [53] studied about analysis of the parameters relating to manufacturing flexibility and efficient performance.

Several kinds of realistic assumptions are taken into account when developing this work. Deterioration during carrying is one such sensible supposition. A portion of the entire order spoils during carrying for a variety of causes. Long distances travelled by the carrying vehicle (such kind of justification is appropriate for medicine, blood, radioactive elements, vaccine, fruits, vegetables, etc.), weather conditions while carrying (such kind of justification is appropriate for sugar, salt, vegetables, fruit, fish, meat, eggs, etc.), carelessness while loading and unloading (such kind of justification is the primary cause of an untrained labour force, and such kind of justification is appropriate for any kind of product), etc. are some possible reasons. A lot of study has already been done in inventory control and inventory management systems that take deterioration into consideration as a crucial factor. Many researchers in the past, including Ghare and Schrader [1] and Aggrawal and Jaggi [5], accepted that once things are received, they begin to deteriorate. The analyses of the development of the deteriorating inventory literature were given by some researchers, including Bakker et al. [12] and Yadav et al. [10].

Another significant factor related to an inventory system is the item's demand. One of the modelling community's major concerns has been it. In order to reflect practical scenarios, a variety of inventory models have been built and explored over time for various item types, taking into account various demand patterns. The demand is influenced by a wide range of variables, including quality, stock, various promotional deals, service quality, etc. One of these crucial factors that greatly influences customer's demand is selling price. The majority of the products are evidently price dependent. Some goods are extremely sensitive, while others are not. As a consequence, when an item's selling price increases, demand for that item declines, and when it decreases, demand for that item increases. Demand obviously declines as the selling price rises. On the other hand, a cheap selling price for some goods might make consumers wonder about their freshness and quality. The demand is also influenced by time. According to some researchers, demand can be a time-based function, while others contend that a quadratic function of time would be more suitable. The market demand in the earlier instance varies dramatically over time, while the market demand in the later case changes gradually over time. In especially for products like vegetables, fruits, sweets, etc., these situations rarely correspond to actual market scenarios. Thus, it appears that a demand function with a linear time dependence is more accurate and a better representation of the changing market needs over time. (An in-depth analysis of the time-dependent linear demand rate pattern is provided in [45]). A large number of inventory management models are informed by the realistic feature that are lower selling prices result in higher sales for many decaying products. Mondal et al. [7] established a selling price-dependent inventory model based on customer's demand. You [8] optimized the product's selling price to maximize the average profit of a manufacturing company. Maihami and Kamalabadi [11] investigated the impact of time and selling price on customer demands in an inventory system. Sarkar et al. [17] considered price and time-based demand in their manufacturing inventory

system under reliability and inflation. Manna et al. [29] conducted additional research on the impact of advertising and selling price on the rate of demand in a manufacturing inventory model. In the past few years, Kumar et al. [36], Yadav et al. [32], Yadav and Swami [33], Aditi and Jaggi [37], Gautam et al. [34] and Yadav and Swami [26] are used price sensitive demand or time sensitive demand in their research.

Another crucial factor of an inventory system is inflation. The total price of products and services rises as a result of inflation over time. When prices rise overall, each unit of currency may purchase fewer products and services. Therefore, inflation denotes a decline in the buying power of money, or a loss of real value in the internal medium of exchange and unit of account of the economy. In 1975, Buzacott [2] constructed the first Economic Order Quantity (EOQ) model that took inflationary effects into account, and it was at this time that he first introduced inventory models with inflation. The inventory model with inflation is then suggested by Harold and Thomas [4]. In the past few years, Inventory models Kausar et al. [38], Tiwari et al. [25], Yadav et al. [23], Yadav and Swami [27] and the Yadav et al. [24] have been proposed in an inflationary context.

Another crucial component of inventory management is deciding where to store the goods. The development of traditional inventory models took into account only one storage facility with infinite capacity. Typically, this storing space is referred to as an owned warehouse. (OW). In actual market situations, however, a merchant may choose to buy more goods than his storage capacity at once due to a price reduction offered for bulk purchases, a high reordering cost, high demand for a product, or seasonal products. Consequently, a second storage space is leased in order to store extra items. The owned warehouse (i.e., OW) is nearby and is known as the rented warehouse (i.e., RW), which we typically presume to have unlimited capacity. Typically, the carrying cost in RW is greater than that in OW. So, in order to lower inventory expenses, RW items are released first, followed by OW items. The two-warehouse inventory approach was first presented by Hartley [3]. A two-storage stock model for degradation goods with time-sensitive demand was then put forth by Bhunia and Maiti [6]. Yang [9] provided some consideration to the two-storage model with incomplete backlogs for deteriorating products and a constant demand rate under inflation. In the recent few years, Swati et al. [13], Chaman and Singh [14], Yadav et al. [21], Hatibaruah and Saha [46], Yadav and Swami [30], Nath and Sen [39], Yadav et al. [43], Vandana and Das [44] and Aarya, D.D., et al. [48] are developed inventory models under two storage facility.

Table 1: *The comparison of our current work with previously published work*

Source	Demand	Deterioration	Warehouses	Preservation Technology	Inflation	Deterioration during carrying
Ghiami et al. [15]	Stock dependent	Yes	Single	No	No	No
Rizwanullah et al. [35]	Stock dependent	Yes	Two	No	Yes	No
Tayal et al. [16]	Selling price & time dependent	Yes	Single	Yes	Yes	No
Tiwari et al. [41]	Constant	Yes	Two	No	No	No
Saha and Chakrabarti [28]	Stock and advertisement dependent	Yes	Single	No	No	No
Momeni et al. [31]	Stock and selling price dependent	Yes	Single	No	No	No
Mahata and Debnath [42]	Selling price dependent	Yes	Single	Yes	No	Yes
Bhunia et al. [19]	Time, selling price and advertisement dependent	Yes	Two	No	No	No
Palanivel et al. [20]	Stock dependent	Yes	Two	No	Yes	No
Jiangtao et al. [18]	Stock dependent	Yes	Single	No	No	No
Huang et al. [40]	Selling price & stock dependent	Yes	Single	No	Yes	No
Akhtar et al. [45]	Selling price & time dependent	Yes	Single	No	No	No
Present paper	Selling price & time dependent	Yes	Two	No	Yes	Yes

2. PRESUMPTIONS AND NOTATIONS

2.1. Presumptions

The following presumptions were used in the formulation of the mathematical model.

1. We know that selling price and time have an impact on market demand. As a result of this finding, we share Akhtar, et al. [45] and Arya, D.D., et al. [48] view that demand is a function of both time and selling price-dependent.
2. Demand rate function is $f(p, t) = a - bp + ct$, where $a, b, c > 0$ are constants.
3. We consider the deterioration during carrying same as Mahata and Debnath [42].
4. We are assuming constant rate of deterioration, which is $\theta(0 < \theta \ll 1)$ in supplier's storehouse and $\gamma(0 < \gamma \ll 1)$ in retailer's storehouses..
5. Planning horizon is infinite.
6. There is no lead time.
7. In retailer's storehouses, OW capacity is limited but RW capacity is deemed boundless same as Vandana and Das [44] and Arya, D.D., et al. [48].
8. We are assuming constant holding cost in both the storehouses.
9. Both the time and the expense of transportation are insignificant.
10. We consider the inflation for developed this model same as Palanivel et al. [20] and Huang et al. [40].
11. There is no item replacement or repair.
12. Stock out are not allowed.

2.2. Notations

Table 2 is provided a description of the notations utilised for the constructed mathematical model.

Table 2: *Notations*

Notation	Units	Description
a	Constant	Coefficient of demand function
b	Constant	Coefficient of demand function
c	Constant	Coefficient of demand function
θ	Constant	Rate of deterioration during carrying.
γ	Constant	Rate of deterioration in retailer's warehouses.
Q	Units	The quantity of orders made each cycle
M	Units	Deteriorating items quantity due to carrying
W	Units	Retailer's Owned Warehouse capacity
$S - W$	Units	Retailer's Rented Warehouse
t_1	Weeks	The stock arrived in retailer's warehouses at this time.
$I(t)$	Units	Inventory level at time t .
h	\$/Unit	Holding cost per unit..
β	\$/Units	Deterioration cost per unit.
r	Constants	Inflation rate.
α	\$/Units	Cost of purchasing per unit.
$TAIPF$	\$/Cycle	The total average inventory profit function

Table 3: *Decision-making parameters*

Notation	Units	Description
p	\$/Units	Selling price of each product, where $p > \alpha$.
T	Weeks	Length of the cycle..

3. MATHEMATICAL MODEL FORMULATION

In the starting, Q units of deteriorating goods were ordered by the retailer from supplier. Thus, Q represents the inventory quantity at time zero. The inventory level steadily drops M at the time $t = t_1$ due to the deterioration rate θ while during carrying from the supplier's storehouse to retailer's storehouses (i.e., RW and OW). At the time interval $t \in [t_1, t_2]$, the joint impact of deterioration and demand decreases the inventory/stock level in RW of the retailer's storehouse to drop until it reaches zero. Also, at the time interval $t \in [t_1, t_2]$, only the effect of deterioration decreases the inventory level in OW of the retailer's storehouse. Again, at the time interval $t \in [t_2, T]$, the joint impact of deterioration and demand decreases the inventory/stock level in OW of the retailer's storehouse to drop until it reaches zero (See fig.1).

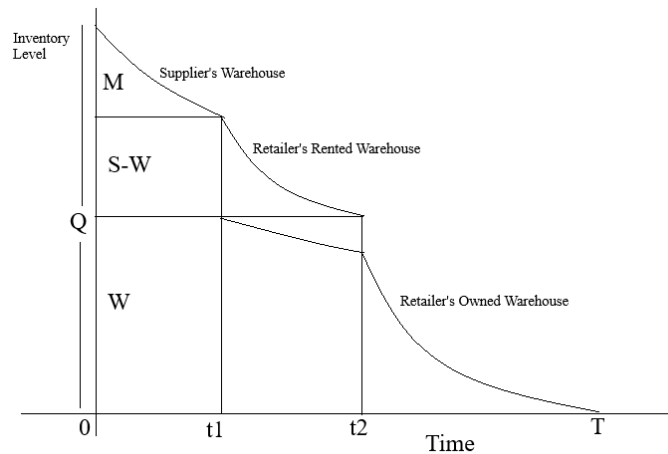


Figure 1: A graphical representation of a deteriorated inventory model with two warehouses

The stock level at $t = 0$ to $t = T$ is characterised in the differential equations as follows:

$$\frac{dI_1(t)}{dt} + \theta I_1(t) = -f(p, t); \quad t \in [0, t_1] \quad (1)$$

with the boundary conditions (B.C.) $I_1(t_1) = S$ and $I_1(0) = Q$.

$$\frac{dI_2(t)}{dt} + \gamma I_2(t) = -f(p, t); \quad t \in [t_1, t_2] \quad (2)$$

with the boundary conditions (B.C.) $I_2(t_2) = 0$ and $I_2(t_1) = S - W$.

$$\frac{dI_3(t)}{dt} + \gamma I_3(t) = 0; \quad t \in [t_1, t_2] \quad (3)$$

with the boundary conditions (B.C.) $I_3(t_1) = W$.

$$\frac{dI_4(t)}{dt} + \gamma I_4(t) = -f(p, t); \quad t \in [t_2, T] \quad (4)$$

with the boundary conditions (B.C.) $I_4(T) = 0$.

The equations (5), (6), (7) and (8) are the solutions of equations (1), (2), (3) and (4), respectively:

$$I_1(t) = \frac{1}{\theta^2} \left[c + \theta^2 e^{(t_1-t)\theta} \left(S + \frac{(a - bp + ct_1)}{\theta} - \frac{c}{\theta^2} \right) \right] - \frac{(a - bp + ct)}{\theta} \quad (5)$$

$$I_2(t) = \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(t_2-t)\gamma} \left(\frac{(a - bp + ct_2)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a - bp + ct)}{\gamma} \quad (6)$$

$$I_3(t) = We^{(t_1-t)\gamma} \quad (7)$$

$$I_4(t) = \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(T-t)\gamma} \left(\frac{(a-bp+cT)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct)}{\gamma} \quad (8)$$

From the equations (7) and (8), using the continuity at $t = t_2$, we get

$$W = e^{(t_2-t_1)\gamma} \left\{ \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(T-t_2)\gamma} \left(\frac{(a-bp+cT)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_2)}{\gamma} \right\} \quad (9)$$

From the equations (7) and (9), we get

$$I_3(t) = e^{(t_2-t_1)\gamma} \left\{ \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(T-t_2)\gamma} \left(\frac{(a-bp+cT)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_2)}{\gamma} \right\} e^{(t_1-t)\gamma} \quad (10)$$

Using $I_2(t) = S - W$ in equation (6), we get

$$S = W + \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(t_2-t_1)\gamma} \left(\frac{(a-bp+ct_2)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_1)}{\gamma} \quad (11)$$

From the equations (9) and (11), we get

$$S = e^{(t_2-t_1)\gamma} \left\{ \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(T-t_2)\gamma} \left(\frac{(a-bp+cT)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_2)}{\gamma} \right\} + \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(t_2-t_1)\gamma} \left(\frac{(a-bp+ct_2)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_1)}{\gamma} \quad (12)$$

Using $I_1(0) = Q$ in equation (5), we get

$$Q = \frac{1}{\theta^2} \left[c + \theta^2 e^{t_1\theta} \left(S + \frac{(a-bp+ct_1)}{\theta} - \frac{c}{\theta^2} \right) \right] - \frac{(a-bp)}{\theta} \quad (13)$$

Since $M = Q - S$, So from equations (12) and (13), we get

$$M = \frac{1}{\theta^2} \left[c + \theta^2 e^{t_1\theta} \left(S + \frac{(a-bp+ct_1)}{\theta} - \frac{c}{\theta^2} \right) \right] - \frac{(a-bp)}{\theta} - \left\{ e^{(t_2-t_1)\gamma} \left\{ \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(T-t_2)\gamma} \left(\frac{(a-bp+cT)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_2)}{\gamma} \right\} + \frac{1}{\gamma^2} \left[c + \gamma^2 e^{(t_2-t_1)\gamma} \left(\frac{(a-bp+ct_2)}{\gamma} - \frac{c}{\gamma^2} \right) \right] - \frac{(a-bp+ct_1)}{\gamma} \right\} \quad (14)$$

Next, we compute the associated costs and profit as follows:

1. The Ordering cost:

$$OC = A \quad (15)$$

2. The Holding cost:

$$HC = h \left[\int_{t_1}^{t_2} I_2(t) e^{-rt} dt + \int_{t_1}^{t_2} I_3(t) e^{-rt} dt + \int_{t_2}^T I_4(t) e^{-rt} dt \right]$$

$$\begin{aligned}
 HC = & \left\{ h \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1+rT)}{r^2 e^{rT}} - \frac{(1+rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) \right. \right. \\
 & - \left(\frac{ae^{T\gamma} e^{-T(r+\gamma)} - ae^{T\gamma} e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{T\gamma} \left(\frac{ce^{-T(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \\
 & - \left(\frac{pbe^{-rT} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{Tce^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \left. \right\} \\
 & - h \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1+rt_1)}{r^2 e^{rt_1}} - \frac{(1+rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rt_1}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rt_1}} - \frac{1}{e^{rt_2}} \right) \right. \\
 & - \left(\frac{ae^{t_2\gamma} e^{-t_1(r+\gamma)} - ae^{t_2\gamma} e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{t_2\gamma} \left(\frac{ce^{-t_1(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \\
 & - \left(\frac{pbe^{-rt_1} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{t_1(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{t_2 ce^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{t_1(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \left. \right\} \\
 & + \frac{h}{\gamma^2(r+\gamma)} \left[\left(\frac{1}{e^{t_2(r+\gamma)}} - \frac{1}{e^{t_1(r+\gamma)}} \right) \left(c(e^{T\gamma} - e^{t_2\gamma}) - a\gamma(e^{T\gamma} - e^{t_2\gamma}) \right. \right. \\
 & \left. \left. + bp\gamma(e^{T\gamma} - e^{t_2\gamma}) + (c\gamma t_2 e^{t_2\gamma} - Tc\gamma e^{T\gamma}) \right) \right] \left. \right\} \quad (16)
 \end{aligned}$$

3. The Deterioration cost:

$$DC = \gamma\beta \left[\int_{t_1}^{t_2} I_2(t)e^{-rt} dt + \int_{t_1}^{t_2} I_3(t)e^{-rt} dt + \int_{t_2}^T I_4(t)e^{-rt} dt \right]$$

$$\begin{aligned}
 DC = & \left\{ \gamma\beta \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1+rT)}{r^2 e^{rT}} - \frac{(1+rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) \right. \right. \\
 & - \left(\frac{ae^{T\gamma} e^{-T(r+\gamma)} - ae^{T\gamma} e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{T\gamma} \left(\frac{ce^{-T(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \\
 & - \left(\frac{pbe^{-rT} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{Tce^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \left. \right\} \\
 & - \gamma\beta \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1+rt_1)}{r^2 e^{rt_1}} - \frac{(1+rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rt_1}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rt_1}} - \frac{1}{e^{rt_2}} \right) \right. \\
 & - \left(\frac{ae^{t_2\gamma} e^{-t_1(r+\gamma)} - ae^{t_2\gamma} e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{t_2\gamma} \left(\frac{ce^{-t_1(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \\
 & - \left(\frac{pbe^{-rt_1} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{t_1(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{t_2 ce^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{t_1(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \left. \right\} \\
 & + \frac{\beta}{\gamma^2(r+\gamma)} \left[\left(\frac{1}{e^{t_2(r+\gamma)}} - \frac{1}{e^{t_1(r+\gamma)}} \right) \left(c(e^{T\gamma} - e^{t_2\gamma}) - a\gamma(e^{T\gamma} - e^{t_2\gamma}) \right. \right. \\
 & \left. \left. + bp\gamma(e^{T\gamma} - e^{t_2\gamma}) + (c\gamma t_2 e^{t_2\gamma} - Tc\gamma e^{T\gamma}) \right) \right] \left. \right\} \quad (17)
 \end{aligned}$$

4. The Purchasing cost:

$$PC = \alpha \cdot Q = \alpha \cdot \frac{1}{\theta^2} \left[c + \theta^2 e^{t_1\theta} \left(S + \frac{(a - bp + ct_1)}{\theta} - \frac{c}{\theta^2} \right) \right] - \frac{(a - bp)}{\theta} \quad (18)$$

5. The Sales Revenue:

$$SR = p \int_{t_1}^T f(p, t) dt = \frac{p(2a + Tc - 2bp + ct_1)(T - t_1)}{2} \quad (19)$$

Thus, we compute the total profit per unit time by the following equation:

$$TAIPF(p, T) = \frac{1}{T}[SR - OC - HC - DC - PC]$$

$$\begin{aligned} TAIPF(p, T) = & \frac{1}{T} \left\{ \frac{p(2a + Tc - 2bp + ct_1)(T - t_1)}{2} - A \right. \\ & - \left\{ (h + \gamma\beta) \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1 + rT)}{r^2 e^{rT}} - \frac{(1 + rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) \right. \right. \\ & \quad - \left. \left(\frac{ae^{T\gamma}e^{-T(r+\gamma)} - ae^{T\gamma}e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{T\gamma} \left(\frac{ce^{-T(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \right. \\ & \quad \left. - \left(\frac{pbe^{-rT} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{Tce^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \right\} \\ & - (h + \gamma\beta) \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1 + rt_1)}{r^2 e^{rt_1}} - \frac{(1 + rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rt_1}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rt_1}} - \frac{1}{e^{rt_2}} \right) \right. \\ & \quad - \left. \left(\frac{ae^{t_2\gamma}e^{-t_1(r+\gamma)} - ae^{t_2\gamma}e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{t_2\gamma} \left(\frac{ce^{-t_1(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \right. \\ & \quad \left. - \left(\frac{pbe^{-rt_1} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{t_1(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{t_2ce^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{t_1(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \right\} \\ & \quad + \frac{h + \beta}{\gamma^2(r+\gamma)} \left[\left(\frac{1}{e^{t_2(r+\gamma)}} - \frac{1}{e^{t_1(r+\gamma)}} \right) \left(ce^{T\gamma} - e^{t_2\gamma} \right) - a\gamma(e^{T\gamma} - e^{t_2\gamma}) \right. \\ & \quad \left. + bp\gamma(e^{T\gamma} - e^{t_2\gamma}) + (c\gamma t_2 e^{t_2\gamma} - Tc\gamma e^{T\gamma}) \right] \left. \right\} \\ & - \alpha \cdot \frac{1}{\theta^2} \left[c + \theta^2 e^{t_1\theta} \left(S + \frac{(a - bp + ct_1)}{\theta} - \frac{c}{\theta^2} \right) \right] - \frac{(a - bp)}{\theta} \left. \right\} \quad (20) \end{aligned}$$

Let $t_1 = kt_2$, $0 < k < 1$, then we get equation (21) from equation (20).

$$\begin{aligned} TAIPF(p, T) = & \frac{1}{T} \left\{ \frac{p(2a + Tc - 2bp + ckt_2)(T - kt_2)}{2} - A - \left\{ (h + \gamma\beta) \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1 + rT)}{r^2 e^{rT}} - \frac{(1 + rt_2)}{r^2 e^{rt_2}} \right) \right] \right. \right. \right. \\ & + \frac{a}{r\gamma} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rT}} - \frac{1}{e^{rt_2}} \right) - \left. \left(\frac{ae^{T\gamma}e^{-T(r+\gamma)} - ae^{T\gamma}e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{T\gamma} \left(\frac{ce^{-T(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) \right. \\ & \quad \left. - \left(\frac{pbe^{-rT} - pbe^{-rt_2}}{r\gamma} \right) + \frac{bpe^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{Tce^{\gamma T}}{\gamma(r+\gamma)} \left(\frac{1}{e^{T(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \right\} \\ & - (h + \gamma\beta) \left\{ \frac{1}{\gamma} \left[c \left(\frac{(1 + rkt_2)}{r^2 e^{rkt_2}} - \frac{(1 + rt_2)}{r^2 e^{rt_2}} \right) \right] + \frac{a}{r\gamma} \left(\frac{1}{e^{rkt_2}} - \frac{1}{e^{rt_2}} \right) - \frac{c}{r\gamma^2} \left(\frac{1}{e^{rkt_2}} - \frac{1}{e^{rt_2}} \right) \right. \\ & \quad - \left. \left(\frac{ae^{t_2\gamma}e^{-kt_2(r+\gamma)} - ae^{t_2\gamma}e^{-t_2(r+\gamma)}}{\gamma(r+\gamma)} \right) + e^{t_2\gamma} \left(\frac{ce^{-kt_2(r+\gamma)} - ce^{-t_2(r+\gamma)}}{\gamma^2(r+\gamma)} \right) - \left(\frac{pbe^{-rkt_2} - pbe^{-rt_2}}{r\gamma} \right) \right. \\ & \quad \left. + \frac{bpe^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{kt_2(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) - \frac{t_2ce^{\gamma t_2}}{\gamma(r+\gamma)} \left(\frac{1}{e^{kt_2(r+\gamma)}} - \frac{1}{e^{t_2(r+\gamma)}} \right) \right\} \\ & \quad + \frac{h + \beta}{\gamma^2(r+\gamma)} \left[\left(\frac{1}{e^{t_2(r+\gamma)}} - \frac{1}{e^{kt_2(r+\gamma)}} \right) \left(ce^{T\gamma} - e^{t_2\gamma} \right) - a\gamma(e^{T\gamma} - e^{t_2\gamma}) \right. \\ & \quad \left. + bp\gamma(e^{T\gamma} - e^{t_2\gamma}) + (c\gamma t_2 e^{t_2\gamma} - Tc\gamma e^{T\gamma}) \right] \left. \right\} - \alpha \cdot \frac{1}{\theta^2} \left[c + \theta^2 e^{kt_2\theta} \left(S + \frac{(a - bp + ckt_2)}{\theta} - \frac{c}{\theta^2} \right) \right] - \frac{(a - bp)}{\theta} \left. \right\} \quad (21) \end{aligned}$$

4. SOLUTION PROCEDURE

In this section, we explore the concavity of the objective function. Tiwari et al. [22] is also used the following technique for optimization in their research article. The necessary criteria must be satisfied in order to attain a maximum total profit:

$$\frac{\partial TAI PF(p, T)}{\partial p} = 0, \quad \frac{\partial TAI PF(p, T)}{\partial T} = 0 \quad (22)$$

Two equations are derived from equation (22). The optimal values of p and T (namely, p^* and T^*) are determined by solving these two equations, which include two unknown variables, p and T , the subsequent sufficient conditions are also satisfied by it.

The conditions as mention below that must be satisfied in order to maximize $TAI PF(p, T)$ using the hessian matrix HM , a matrix of 2^{nd} order partial derivatives:

$$\mathcal{H} = \begin{bmatrix} \frac{\partial^2 TAI PF(p, T)}{\partial p^2} & \frac{\partial^2 TAI PF(p, T)}{\partial p \partial T} \\ \frac{\partial^2 TAI PF(p, T)}{\partial T \partial p} & \frac{\partial^2 TAI PF(p, T)}{\partial T^2} \end{bmatrix},$$

$$D_{11} = \frac{\partial^2 TAI PF(p, T)}{\partial p^2} < 0, \quad D_{22} = \det(HM) > 0.$$

Here, D_{11} and D_{22} are the minors of the hessian matrix HM . Figure 2 displays the whole solution strategy for our proposed model, which was derived using MATLAB software (version: R2021b).

5. NUMERICAL ILLUSTRATION

To maximize the total average profit function $TAI PF(p, T)$, the present model seeks to identify the p and T optimal values. Since $TAI PF(p, T)$ generated in equation (21) is to find the best values for the decision variables p and T , it is extremely difficult to calculate a complex function analytically. The model is solved with the help of the following algorithm.

5.1. Algorithm

1. Fix the values of the parameters $a, b, c, \theta, t_2, \gamma, S, \alpha, \beta, h, r, A, k$.
2. Build the function $TAI PF(p, T)$ given by equation (21).
3. Maximize $TAI PF(p, T)$ subject to the constraints $0 < t_1 < t_2 < T$ and $0 < \alpha < p$.
4. Calculate the optimum values p^*, T^*, t_1^*, Q^* and $TAI PF^*$.

5.2. Example

A numerical example is given in this section to show how the model works. The following parameters values are used as input:

$$a = 75, b = 1.5, c = 2.1, \theta = 0.46, t_2 = 4.0 \text{ weeks}, \gamma = 0.44, S = 900 \text{ units},$$

$$\alpha = \$1.1/\text{unit}, \beta = \$1.8/\text{unit}, h = \$3.6/\text{unit}, r = 0.85, A = \$110, k = 0.55.$$

The optimal solution obtained is as below:

$$p^* = \$32.4827/\text{unit}, T^* = 10.6011 \text{ weeks}, t_1^* = 2.2000 \text{ weeks},$$

$$Q^* = 2586.3 \text{ units}, TAI PF^* = \$550.0893/\text{cycle}.$$

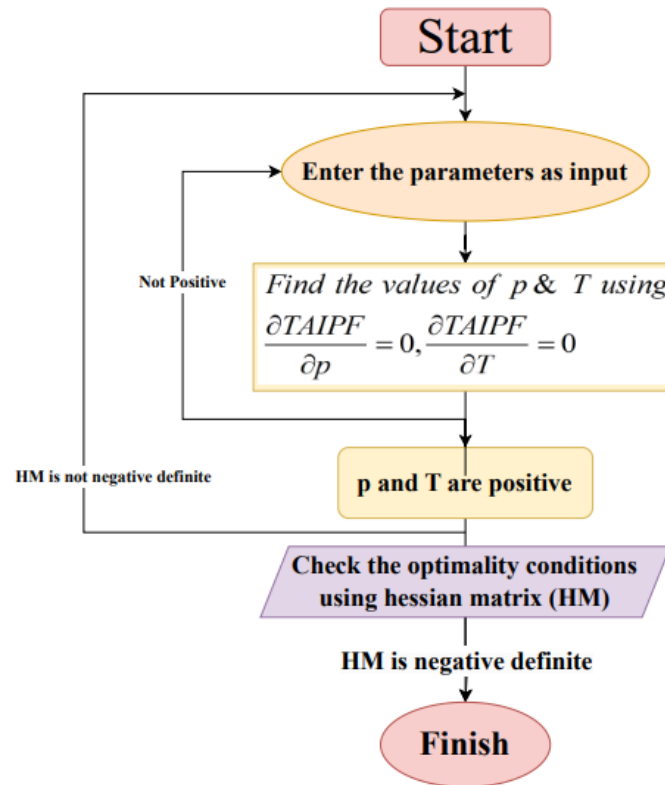


Figure 2: A flowchart depicting our established model solving process.

6. SENSITIVITY ANALYSIS

Sensitivity analysis is used in this part to investigate how changes in the parameter's values affect the optimum values. In order to do these studies, one parameter was changed by $\pm 5\%$ and $\pm 10\%$ at a time while remaining at its original value for the other parameters. The following table 3 is demonstrate the outcomes of the sensitivity analysis.

Table 4: Sensitivity analysis with respect to above example

Parameters	%	% change in optimal value					
		Change	p^*	T^*	t_1^*	Q^*	$TAIPF^*$
a	-10%		29.9669	10.5885	2.2000	2572.1	400.1475
	-5%		31.2235	10.5938	2.2000	2579.2	473.2606
	+5%		33.7443	10.6103	2.2000	2593.3	630.6352
	+10%		35.0079	10.6211	2.2000	2600.4	714.9006
b	-10%		35.9273	10.7299	2.2000	2587.1	664.1233
	-5%		34.1133	10.6628	2.2000	2586.7	603.9984
	+5%		31.0093	10.5443	2.2000	2585.8	501.4878
	+10%		29.6713	10.4918	2.2000	2585.4	457.4537
c	-10%		31.9951	10.5684	2.2000	2588.0	523.7440
	-5%		32.2385	10.5847	2.2000	2587.1	536.8524
	+5%		32.7277	10.6176	2.2000	2585.4	563.4554
	+10%		32.9733	10.6341	2.2000	2584.5	576.9513

Table 5: Sensitivity analysis with respect to above example (Continue)

Parameters	%	% change in optimal value					
		Change	p^*	T^*	t_1^*	Q^*	$TAIPF^*$
θ	-10%		32.3843	10.5333	2.2000	2342.1	575.5644
	-5%		32.4326	10.5667	2.2000	2461.1	563.1258
	+5%		32.5348	10.6366	2.2000	2717.9	536.4270
	+10%		32.5889	10.6731	2.2000	2856.2	522.1102
t_2	-10%		32.2278	9.9178	1.9800	2331.5	542.8833
	-5%		32.3541	10.2592	2.0900	2455.7	546.6066
	+5%		32.6138	10.9435	2.3100	2723.6	553.3205
	+10%		553.3205	11.2863	2.4200	2868.0	556.2894
γ	-10%		32.7035	11.4222	2.2000	2585.0	606.0342
	-5%		32.5851	10.9900	2.2000	2585.7	577.2419
	+5%		32.3940	10.2492	2.2000	2586.8	524.3797
	+10%		32.3168	9.9289	2.2000	2587.2	499.9510
α	-10%		32.3676	10.5285	2.2000	2586.9	577.0201
	-5%		32.4255	10.5652	2.2000	2586.6	563.5307
	+5%		32.5394	10.6363	2.2000	2585.9	536.6945
	+10%		32.5955	10.6708	2.2000	2585.6	523.3453
β	-10%		32.4765	10.6361	2.2000	2586.3	553.6050
	-5%		32.4796	10.6185	2.2000	2586.3	551.8408
	+5%		32.4859	10.5839	2.2000	2586.2	548.3502
	+10%		32.4891	10.5668	2.2000	2586.2	546.6234
h	-10%		32.4551	10.7659	2.2000	2586.4	566.4975
	-5%		32.4687	10.6816	2.2000	2586.3	558.1543
	+5%		32.4973	10.5241	2.2000	2586.2	542.2818
	+10%		32.5122	10.4503	2.2000	2586.1	534.7136
r	-10%		32.5857	10.1147	2.2000	2585.7	499.6508
	-5%		32.5293	10.3579	2.2000	2586.0	525.3019
	+5%		32.4454	10.8445	2.2000	2586.5	574.0624
	+10%		32.4164	11.0879	2.2000	2586.6	597.2693
A	-10%		32.4794	10.5985	2.2000	2586.3	551.1270
	-5%		32.4811	10.5998	2.2000	2586.3	550.6081
	+5%		32.4844	10.6025	2.2000	2586.3	549.5705
	+10%		32.4861	10.6038	2.2000	2586.2	549.0518
k	-10%		32.2278	9.9178	1.9800	2331.5	542.8833
	-5%		32.3541	10.2592	2.0900	2455.7	546.6066
	+5%		32.6138	10.9435	2.3100	2723.6	553.3205
	+10%		32.7473	11.2863	2.4200	2868.0	556.2894

Following are a few insights drawn from the sensitivity analysis's observations.

1. If we increase in a , then the total average inventory profit $TAIPF$ is increases, because demand is increases. Simultaneously selling price, total cycle length T and total quantity Q are increases (See figure 7).
2. If we increase in b , then the total average inventory profit $TAIPF$ is decreases, because demand is decreases. Simultaneously selling price, total cycle length T and total quantity Q

are decreases.

3. If we increase in c , then the total average inventory profit $TAIPF$ is increases, because demand is increases. Simultaneously selling price and total cycle length T are increases, and total quantity Q is decreases.
4. If we increase in θ (deterioration during carrying), then the total average inventory profit $TAIPF$ is decreases, because the associated cost is increases. Simultaneously selling price, total cycle length T and total quantity Q are increases (See figure 4).
5. If we increase in γ , then the total average inventory profit $TAIPF$ is decreases, because the deterioration in retailer's warehouses are increases, so that associated inventory cost is increases. Simultaneously selling price, total cycle length T are decreases, and total quantity Q is increases.
6. If we increase in r , then the total average inventory profit $TAIPF$ is increases, because selling price will be decreases, so that demand will be increases. If demand will be increase, then the total average inventory profit is increases. Simultaneously the total cycle length T and total quantity Q are increases.
7. If we increase in α, β and h , then the total average inventory profit $TAIPF$ is decreases, because associated inventory costs are increases. Simultaneously, selling price p is increases and the total quantity Q is decreases.

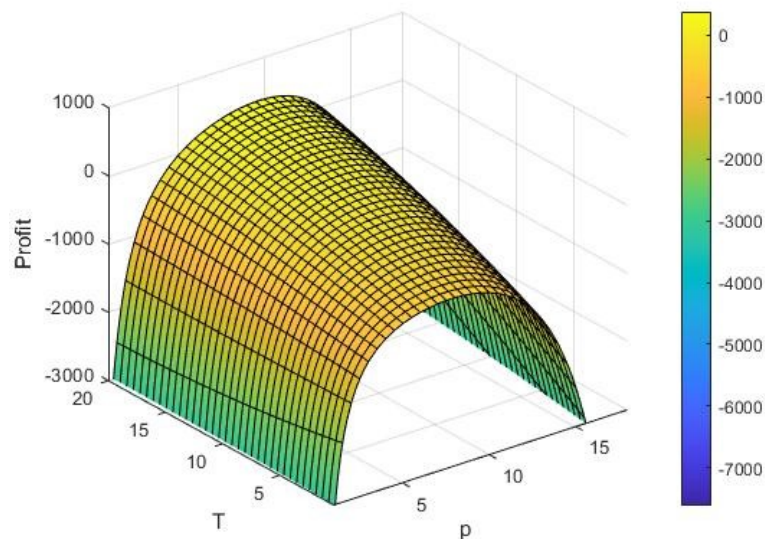


Figure 3: Concavity of $TAIPF(p, T)$ with respect to p and T

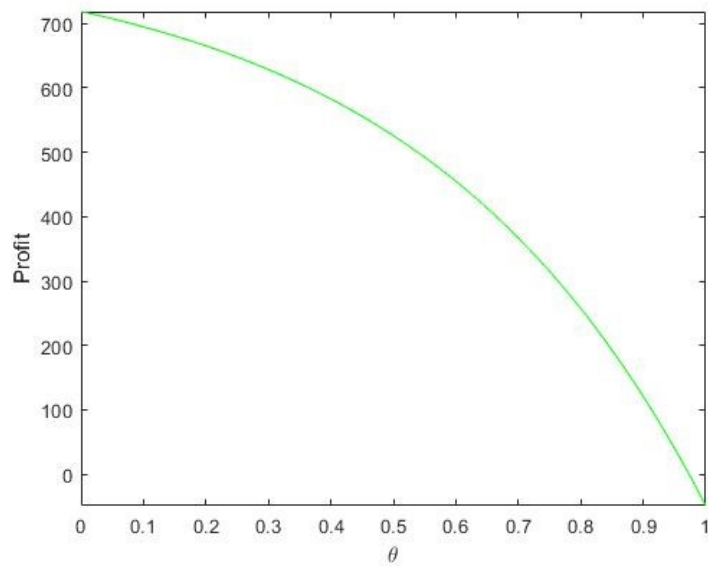


Figure 4: Variation between Profit vs. Deterioration during carrying (θ)

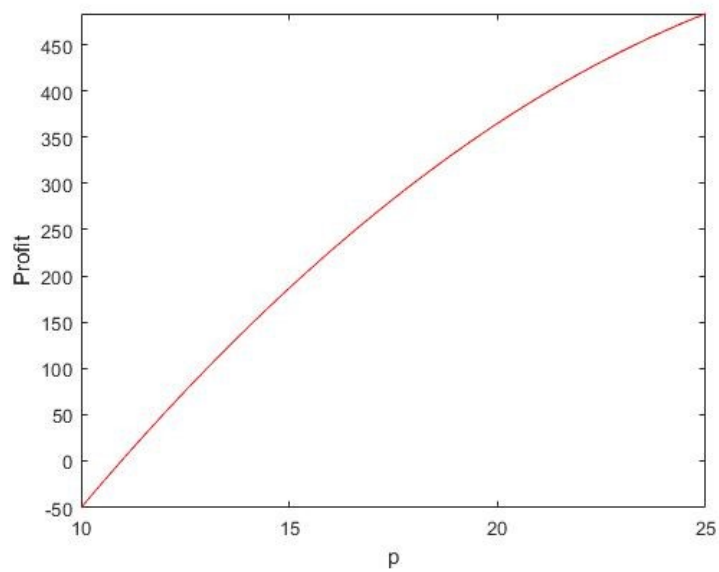


Figure 5: Variation between Profit vs. Selling price (p)

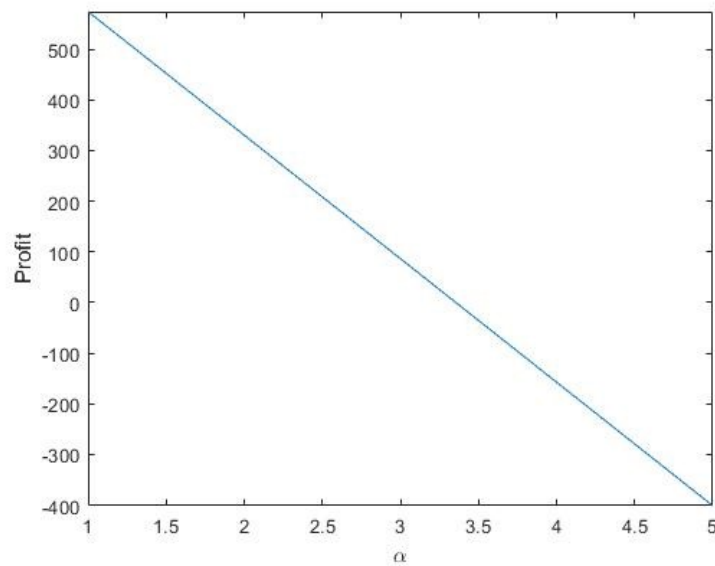


Figure 6: Variation between Profit vs. Purchasing cost (α)

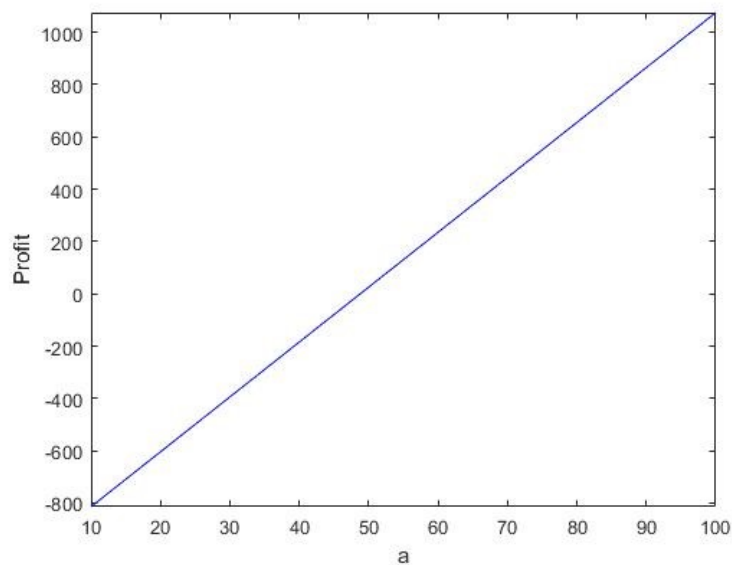


Figure 7: Variation between Profit vs. Constant part of demand function (a)

7. CONCLUSION AND FUTURE DIRECTIONS

In this study, the carrying of decaying goods from the supplierTMs storehouse to the retailerTMs storehouses and deterioration in the retailerTMs storehouses with time and selling price incumbent demand under inflation have both been addressed in a two-level supply chain inventory model. We are presuming that retailers have two warehouses, named as OW and RW. We are taking a steady rate of decline into account. In such a scenario, shortages are not permitted. This study aims to find the optimal selling price and cycle length by implementing an algorithm that maximizes the total average inventory profit per unit of time. Considering deterioration during carrying is the most important part of this article. Many researchers have not yet taken into account that part of this article. Finally, the applicability of the proposed model is demonstrated

with a numerical example and pictorial representation. A sensitivity analysis of important parameters is provided with the help of MATLAB software (version: R2021b). The proposed model can also be modified to take into consideration different types of variable demands. It is also possible to recommend future research, such as investigating payment policies and preservation technologies.

REFERENCES

- [1] Ghare, P. M., and G. F. Schrader. "An inventory model for exponentially deteriorating items". *Journal of Industrial Engineering* 14.2 (1963): 238-243.
- [2] Buzacott, J. A. "Economic order quantities with inflation." *Journal of the Operational Research Society* 26.3 (1975): 553-558.
- [3] Hartley, R. V. "Operations research: a managerial emphasis (Vol. 976)." (1976).
- [4] Bierman Jr, Harold, and Joseph Thomas. "Inventory decisions under inflationary conditions." *Decision Sciences* 8.1 (1977): 151-155.
- [5] Aggarwal, S. P., and C. K. Jaggi. "Ordering policies of deteriorating items under permissible delay in payments." *Journal of the operational Research Society* 46.5 (1995): 658-662.
- [6] Bhunia, Asoke Kumar, and Manoranjan Maiti. "A two-warehouse inventory model for deteriorating items with a linear trend in demand and shortages." *Journal of the Operational Research Society* 49.3 (1998): 287-292.
- [7] Mondal, Biswajit, Asoke Kumar Bhunia, and Manoranjan Maiti. "An inventory system of ameliorating items for price dependent demand rate." *Computers & industrial engineering* 45.3 (2003): 443-456.
- [8] You, Peng-Sheng. "Ordering and pricing of service products in an advance sales system with price-dependent demand." *European Journal of Operational Research* 170.1 (2006): 57-71.
- [9] Yang, Hui-Ling. "Two-warehouse partial backlogging inventory models for deteriorating items under inflation." *International Journal of Production Economics* 103.1 (2006): 362-370.
- [10] Singh, Ajay Kumar, S. Yadav, and S. R. Singh. "A two-warehouse inventory model for a deteriorating item with exponential demand rate, partially backlogged shortages." *International Transactions in Applied Sciences* 2.4 (2010).
- [11] Maihmi, Reza, and Isa Nakhai Kamalabadi. "Joint pricing and inventory control for non-instantaneous deteriorating items with partial backlogging and time and price dependent demand." *International Journal of Production Economics* 136.1 (2012): 116-122.
- [12] Bakker, Monique, Jan Riezebos, and Ruud H. Teunter. "Review of inventory systems with deterioration since 2001." *European Journal of Operational Research* 221.2 (2012): 275-284.
- [13] Agrawal, Swati, Snigdha Banerjee, and Sotirios Papachristos. "Inventory model with deteriorating items, ramp type demand and partially backlogged shortages for a two-warehouse system." *Applied Mathematical Modelling* 37.20-21 (2013): 8912-8929.
- [14] Singh, Chaman, and S. R. Singh. "Optimal ordering policy for deteriorating items with power-form stock dependent demand under two-warehouse storage facility." *Opsearch* 50 (2013): 182-196.
- [15] Ghiami, Yousef, Terry Williams, and Yue Wu. "A two-echelon inventory model for a deteriorating item with stock dependent demand, partial backlogging and capacity constraints." *European Journal of Operational Research* 231.3 (2013): 587-597.
- [16] Tayal, Shilpy, et al. "Two echelon supply chain model for deteriorating items with effective investment in preservation technology." *International Journal of Mathematics in Operational Research* 6.1 (2014): 84-105.
- [17] Sarkar, Biswajit, Papiya Mandal, and Sumon Sarkar. "An EMQ model with price and time dependent demand under the effect of reliability and inflation." *Applied Mathematics and Computation* 231 (2014): 414-421.
- [18] Jiangtao, Mo, et al. "Optimal ordering policies for perishable multi-item under stock-dependent demand and two-level trade credit." *Applied Mathematical Modelling* 38.9-10 (2014): 2522-2532.

- [19] Bhunia, Asoke Kumar, et al. "A two-storage inventory model for deteriorating items with variable demand and partial backlogging." *Journal of Industrial and Production Engineering* 32.4 (2015): 263-272.
- [20] Palanivel, M., R. Sundararajan, and R. Uthayakumar. "Two-warehouse inventory model with non-instantaneously deteriorating items, stock-dependent demand, shortages and inflation." *Journal of Management Analytics* 3.2 (2016): 152-173.
- [21] Yadav, A. S., et al. "Multi objective optimization for electronic component inventory model deteriorating items with two-warehouse using genetic algorithm." *International Journal of Control Theory and applications* 9.2 (2016): 15-35.
- [22] Tiwari, S., et al. "Impact of trade credit and inflation on retailerTMs ordering policies for non-instantaneous deteriorating items in a two-warehouse environment." *International Journal of Production Economics* 176 (2016): 154-169.
- [23] Yadav, Ajay Singh, et al. "Effect of inflation on a two-warehouse inventory model for deteriorating items with time varying demand and shortages." *International Journal of Procurement Management* 10.6 (2017): 761-775.
- [24] Yadav, A. S., et al. "An inflationary inventory model for deteriorating items under two storage systems." *International Journal of Economic Research* 14.9 (2017): 29-40.
- [25] Tiwari, Sunil, et al. "Two-warehouse inventory model for non-instantaneous deteriorating items with stock dependent demand and inflation using particle swarm optimization." *Annals of Operations Research* 254 (2017): 401-423.
- [26] Yadav, Ajay Singh, and Anupam Swami. "A partial backlogging production inventory lot-size model with time varying holding cost and weibull deterioration." *International Journal of Procurement Management* 11.5 (2018): 639-649.
- [27] Yadav, Ajay Singh, and Anupam Swami. "Integrated supply chain model for deteriorating items with linear stock dependent demand under imprecise and inflationary environment." *International Journal of Procurement Management* 11.6 (2018): 684-704.
- [28] Saha, Sujata, and Tripti Chakrabarti. "Two-echelon supply chain model for deteriorating items in an imperfect production system with advertisement and stock dependent demand under trade credit." *International Journal of Supply and Operations Management* 5.3 (2018): 207-217.
- [29] Manna, Amalesh Kumar, et al. "An EPQ model with promotional demand in random planning horizon: population varying genetic algorithm approach." *Journal of Intelligent Manufacturing* 29 (2018): 1515-1531.
- [30] Yadav, Ajay Singh, and Anupam Swami. "A volume flexible two-warehouse model with fluctuating demand and holding cost under inflation." *International Journal of Procurement Management* 12.4 (2019): 441-456.
- [31] Arab Momeni, Mojtaba, Saeed Yaghoubi, and Mohammad Reza Mohammad Aliha. "A cost sharing-based coordination mechanism for multiple deteriorating items in a one manufacture-one retailer supply chain." *Journal of Industrial Engineering and Management Studies* 6.1 (2019): 79-110.
- [32] Yadav, Ajay Singh, et al. "Supply chain inventory model for deteriorating item with warehouse & distribution centres under inflation." *International Journal of Engineering and Advanced Technology* 8.2 (2019): 7-13.
- [33] Yadav, Ajay Singh, and Anupam Swami. "An inventory model for noninstantaneous deteriorating items with variable holding cost under two-storage." *International journal of procurement management* 12.6 (2019): 690-710.
- [34] Gautam, Prerna, Aditi Khanna, and Chandra K. Jaggi. "Preservation technology investment for an inventory system with variable deterioration rate under expiration dates and price sensitive demand." *Yugoslav Journal of Operations Research* 30.3 (2020): 289-305.
- [35] Rizwanullah, Mohd, Sachin Kumar Verma, and Maqsood Hussain Junnaidi. "Supply-chain two-warehouse inventory model for deteriorating items on exponential time function with shortage and partial backlogging in inflationary environment." *Journal of Physics: Conference Series*. Vol. 1913. No. 1. IOP Publishing, 2021.

- [36] Kumar, Satish, et al. "An inventory model with price dependent demand rates as power law form using ant colony optimization." *Journal of Management Information and Decision Sciences* 24.6 (2021): 1-10.
- [37] Khanna, Aditi, and Chandra K. Jaggi. "An inventory model under price and stock dependent demand for controllable deterioration rate with shortages and preservation technology investment: revisited." *Opsearch* 58.1 (2021): 181-202.
- [38] Kausar, Amrina, et al. "Dual Warehouse Inventory Management of Deteriorating Items Under Inflationary Condition." *Advances in Interdisciplinary Research in Engineering and Business Management* (2021): 39-53.
- [39] Nath, Biman Kanti, and Nabendu Sen. "A Completely backlogged two warehouse inventory model for non-instantaneous deteriorating items with time and selling price dependent demand." *International Journal of Applied and Computational Mathematics* 7.4 (2021): 145.
- [40] Huang, Xiangmeng, Shuai Yang, and Zhanyu Wang. "Optimal pricing and replenishment policy for perishable food supply chain under inflation." *Computers & Industrial Engineering* 158 (2021): 107433.
- [41] Tiwari, Sunil, et al. "RetailerTMs credit and inventory decisions for imperfect quality and deteriorating items under two-level trade credit." *Computers & Operations Research* 138 (2022): 105617.
- [42] Mahata, Sourav, and Bijoy Krishna Debnath. "A profit maximization single item inventory problem considering deterioration during carrying for price dependent demand and preservation technology investment." *RAIRO Operations Research* 56.3 (2022): 1841-1856.
- [43] Yadav, A.S., et al. "Corona virus vaccine supply chain: green inventory due to covid-19 for storage of vaccine waste items and environmental pollution removal using optimized flower pollination services." *Journal of Tianjin University Science and Technology* (2022): 544-561.
- [44] Vandana and Das, A. K. "Two-warehouse supply chain model under preservation technology and stochastic demand with shortages." *OPSEARCH* (2022): 1-26.
- [45] Akhtar, Md, Amalesh Kumar Manna, and Asoke Kumar Bhunia. "Optimization of a non-instantaneous deteriorating inventory problem with time and price dependent demand over finite time horizon via hybrid DESGO algorithm." *Expert Systems with Applications* 211 (2023): 118676.
- [46] Hatibaruah, Ajoy, and Sumit Saha. "An inventory model for two-parameter Weibull distributed ameliorating and deteriorating items with stock and advertisement frequency dependent demand under trade credit and preservation technology." *OPSEARCH* (2023): 1-52.
- [47] Poswal, Preety, et al. "Investigation and analysis of fuzzy EOQ model for price sensitive and stock dependent demand under shortages." *Materials Today: Proceedings* 56 (2022): 542-548.
- [48] Arya, Deo Datta, et al. "Selling price, time dependent demand and variable holding cost inventory model with two storage facilities." *Materials Today: Proceedings* 56 (2022): 245-251.
- [49] Kumar, Amit, et al. "Stochastic Petri nets modelling for performance assessment of a manufacturing unit." *Materials Today: Proceedings* 56 (2022): 215-219.
- [50] Kumar, Kaushal, Ajay Kumar, and Vinay Singh. "Optimization of process parameters for erosion wear in slurry pipeline." *Advances in Engineering Design: Select Proceedings of FLAME 2018*. Springer Singapore, 2019.
- [51] Mahata, Sourav, and Bijoy Krishna Debnath. "The impact of RD expenditures and screening in an economic production rate (EPR) inventory model for a flawed production system with imperfect screening under an interval-valued environment." *Journal of Computational Science* 69 (2023): 102027.
- [52] Kundu, Tanmay, and Sahidul Islam. "A new interactive approach to solve entropy based fuzzy reliability optimization model." *International Journal on Interactive Design and Manufacturing (IJIDeM)* 13 (2019): 137-146.
- [53] °S, Yusuf Tansel, et al. "Analysis of the manufacturing flexibility parameters with effective performance metrics: A new interactive approach based on modified TOPSIS-Taguchi

- method." *International Journal on Interactive Design and Manufacturing (IJIDeM)* 16.1 (2022): 197-225.
- [54] Kumar, Pankaj, et al. "Optimization of cycle time assembly line for mass manufacturing." *International Journal on Interactive Design and Manufacturing (IJIDeM)* (2023): 1-12.
- [55] Wong, Yu Cheng, Kuan Yew Wong, and Anwar Ali. "A study on lean manufacturing implementation in the Malaysian electrical and electronics industry." *European Journal of Scientific Research* 38.4 (2009): 521-535.
- [56] Basdere, Bahadir, and Guenther Seliger. "Disassembly factories for electrical and electronic products to recover resources in product and material cycles." *Environmental science & technology* 37.23 (2003): 5354-5362.
- [57] Colledani, Marcello, et al. "Design and management of manufacturing systems for production quality." *Cirp Annals* 63.2 (2014): 773-796.

ON MODELING OF BIOMEDICAL DATA WITH EXPONENTIATED GOMPERTZ INVERSE RAYLEIGH DISTRIBUTION

Sule Omeiza Bashiru¹, Alaa Abdulrahman Khalaf², Alhaji Modu Isa³ and Aishatu
Kaigama⁴

¹Department of Mathematics and Statistics, Confluence University of Science and Technology,
Osara, Kogi State, Nigeria.

²Diyala Education Directorate, Diyala, Iraq.

^{3,4}Department of Mathematics and Computer Science, Borno State University, Nigeria.

Email: ¹bash0140@gmail.com; ²alaa.a.khalaf35510@st.tu.edu.iq; ³alhajimoduisa@bosu.edu.ng;
⁴a.kaigama@bosu.edu.ng

Abstract

This paper introduces and thoroughly examines the Exponentiated Gompertz Inverse Rayleigh (EtGoIr) Distribution, a four-parameter extension of the Gompertz Inverse Rayleigh distribution. The primary focus is on its application to biomedical datasets, shedding light on its mathematical and statistical properties. Some properties of the distribution that were derived include the quantile function, median, moments, incomplete moments, Rényi entropy, and probability weighted moments. The model parameters were estimated using the method of maximum likelihood. A simulation study was conducted to investigate the consistency of the proposed model. The outcome of the investigation revealed that the model demonstrates consistency, as evidenced by the reduction in both root mean square error (RMSE) and bias as sample sizes increase. To showcase the practical relevance of the EtGoIr distribution, the paper applies the model to three distinct biomedical datasets. The results highlight its enhanced flexibility, demonstrating superior fit compared to its counterpart.

Keywords: Exponentiated G, MLE, Moment, Renyi Entropy, Biomedical

I. Introduction

Statistical theory continually evolves to meet the demands of modeling complex natural phenomena effectively. Traditional probability distributions have long served as foundational tools, yet the complexities of modern biomedical datasets often necessitate the development of novel models to extract deeper insights. This necessity is particularly pronounced in biomedical research, where conventional distributions struggle to capture the intricacies of physiological measurements, disease outcomes, and survival times across various medical conditions. Recent advancements in distribution theory have underscored the importance of innovative models for accommodating the skewness prevalent in the aforementioned datasets. This skewness poses a significant challenge to conventional distributions, prompting researchers to explore extensions of established models to better capture these complexities. Notable among these extensions are the works of [1] – [9].

In this study, we focus on extending the Gompertz inverse Rayleigh (GoIR) distribution, introduced by [10], to create a more adaptable model. We investigate the exponentiated (Et) family of distributions, as proposed by [11], to achieve this extension. By combining the GoIR distribution with the Et family, our aim is to develop a versatile model capable of accurately fitting real-world datasets, particularly in biomedical science applications.

The cumulative distribution function (cdf) and probability density function (pdf) of the Et family are given respectively as:

$$F(x) = [G(x)]^\theta ; \tag{1}$$

$$f(x) = \theta g(x)[G(x)]^{\theta-1} : \theta > 0 \tag{2}$$

where $G(x)$ and $g(x)$ are the cdf and pdf of the baseline distribution.

The cdf and pdf of GoIR distribution taken as baseline are given as:

$$G(x) = 1 - e^{-\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \tag{3}$$

and

$$g(x) = 2\beta\gamma^2 x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} e^{-\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} ; \theta > 0, \beta > 0, \gamma > 0, \sigma > 0 \tag{4}$$

The motivation for this research arises from the recognition that traditional distributions often fall short in accommodating the complexities of biomedical datasets, especially those exhibiting skewness. By extending the GoIR distribution, we seek to contribute to the development of hybrid distributions that better reflect the intricacies of real-world data.

II. Methods

2.1 Derivation of Exponentiated Gompertz Inverse Rayleigh (EtGoIR) Distribution

This section introduces a new model called the EtGoIR distribution. The cdf of the EtGoIR distribution is derived by substituting equation (3) into equation (1), as follows:

$$F(x) = \left(1 - e^{-\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^\theta \tag{5}$$

On differentiating equation (5) with respect to x , we obtain pdf of EtGoIR distribution given as:

$$f(x) = 2\theta\beta\gamma^2 x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} e^{-\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \left(1 - e^{-\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{\theta-1} \tag{6}$$

The pdf plot of the EtGoIR distribution is given in Figure 1 below.

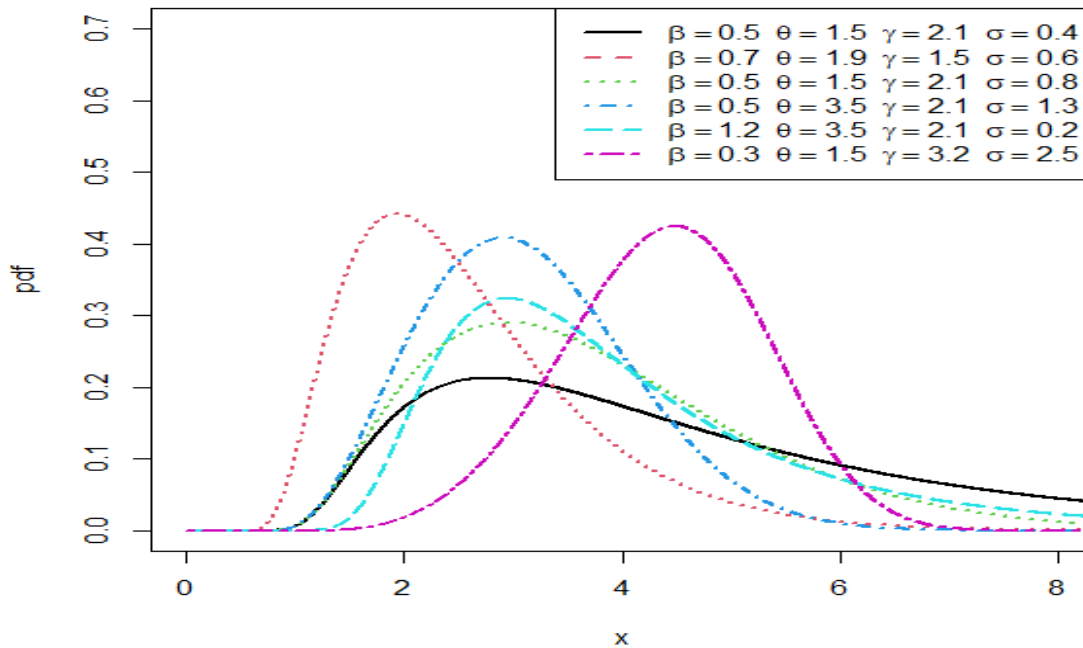


Figure 1: pdf plot of EtGolR distribution

2.2 Expansion of Density

Using the generalized binomial expansion given as

$$(1 - y)^{\rho-1} = \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(\rho)}{i! \Gamma(\rho - i)} y^i \quad (7)$$

Applying equation (7) on the last term in equation (6), we have

$$\left(1 - e^{\left(\frac{\beta}{\sigma}\right) \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}} \right)^{\theta-1} = \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(\rho)}{i! \Gamma(\rho - i)} e^{\left(\frac{\beta}{\sigma}\right) \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}}^i$$

Where

$$e^{-vy^n} = \sum_{j=0}^{\infty} \frac{(-1)^j v^j}{j!} y^{vj} \quad (8)$$

and

$$e^{-vy^n} = \sum_{j=0}^{\infty} \frac{v^j}{j!} y^{vj} \quad (9)$$

Therefore,

$$e^{\left(\frac{\beta}{\sigma}\right) y^i} = \sum_{j=0}^{\infty} \frac{\left(\frac{\beta}{\sigma}\right)^j}{j!} y^j$$

where

$$y^j = \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}^j = \sum_{k=0}^{\infty} \binom{j}{k} (-1)^k \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma k}$$

Substituting back all the expansions into equation (6), we have

$$f(x) = 2\beta\theta\gamma^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{i+k+l} \left(\frac{\beta}{\sigma}\right)^j \binom{j+1}{k} \Gamma(-\sigma(K+1))\Gamma(\theta)}{i!j!l!\Gamma(-\sigma(K+1)-m)\Gamma(\theta-i)} x^{-3} \quad (10)$$

$$f(x) = \psi x^{-3} \left[e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{m+1}$$

where

$$\psi = 2\beta\theta\gamma^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{i+k+l} \left(\frac{\beta}{\sigma}\right)^j \binom{j+1}{k} \Gamma(-\sigma(K+1))\Gamma(\theta)}{i!j!l!\Gamma(-\sigma(K+1)-m)\Gamma(\theta-i)}$$

2.3 Properties of EtGoIR Distribution

This section derives some statistical properties of the EGILx distribution including moments, survival function, hazard function, quantile functions, and order statistics.

2.3.1 Quantile function

The quantile function is the inverse of the cdf of a distribution and is used in simulation studies. It is also applied as a measure of the spread of a distribution. The quantile function is obtained using:

$$Q(u) = F^{-1}(u) \quad (11)$$

Applying equation (13) to the cdf of the new model, we have the quantile function given as

$$x = \gamma \left\{ -\log \left[1 - \left(1 - \frac{\sigma \log \left(1 - u^{\frac{1}{\theta}} \right)}{\beta} \right)^{-\left(\frac{1}{\sigma}\right) - \left(\frac{1}{2}\right)} \right] \right\} \quad (12)$$

2.3.2 Median

Median of EtGoIR distribution is obtained by setting $u=0.5$ in equation (12) and it is given as

$$x_{median} = \gamma \left\{ -\log \left[1 - \left(1 - \frac{\sigma \log \left(1 - 0.5^{\frac{1}{\theta}} \right)}{\beta} \right)^{-\left(\frac{1}{\sigma}\right) - \left(\frac{1}{2}\right)} \right] \right\} \quad (13)$$

2.3.3 Moments

$$E(X^r) = \int_0^{\infty} x^r f(y) dx \quad (14)$$

$$E(X^r) = \psi \int_0^{\infty} x^{r-3} \left[e^{-\left(\frac{\gamma}{x}\right)^2} \right] dx \quad (15)$$

On solving the integral part in equation (15), we have

$$E(X^r) = \psi \left[\frac{\gamma^r \Gamma\left(1 - \frac{r}{2}\right)}{(k+1)^{1 - \frac{r}{2}}} \right] \quad (16)$$

When $r=1$ in equation (16), we have mean of EtGoIR distribution

2.3.4 Incomplete Moments

The r^{th} ($r > 0$) incomplete moments for the EtGoIR distributions follow from equation (10) as

$$\dot{\mu}_r(u) = \int_0^u \psi x^{r-3} e^{-(m+1)\left(\frac{\gamma}{x}\right)^2} dx$$

$$\text{Let } t = (m+1)\left(\frac{\gamma}{x}\right)^2 \Rightarrow x = \left(\frac{(m+1)\gamma^2}{t}\right)^{\frac{1}{2}}$$

$$\text{When } x = 0 \Rightarrow t = 0, \text{ and if } x = u \Rightarrow t = (m+1)\left(\frac{\gamma}{u}\right)^2$$

Then

$$\dot{\mu}_r(u) = \frac{\psi}{2(m+1)\gamma^2} \gamma \left(1 - \frac{r}{2}, (m+1)\gamma^2\right) \quad (17)$$

2.3.5 Rényi Entropy

Define the Rényi entropy of the EtGoIR distributions with the following formula [12]

$$T_R(\eta) = \frac{1}{1-\eta} \log \int_0^\infty f^\eta(x) dx, \quad \eta > 0, \eta \neq 1$$

By equation (6) we find $f(x)^\eta$:

$$f(x)^\eta = 2\theta^\eta \beta^\eta \gamma^{2\eta} x^{-3\eta} e^{-\left(\frac{\gamma}{x}\right)^2} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-(\sigma+1)\eta} e^{\frac{\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)} \left(1 - e^{\frac{\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)}\right)^{\eta(\theta-1)}$$

By generalized binomial series and exponential expansion, we get

$$\left(1 - e^{\frac{\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)}\right)^{\eta(\theta-1)} = \sum_{i=0}^{\infty} \frac{\Gamma(\eta(\theta-1) + i)}{i! \Gamma(\eta(\theta-1))} e^{\frac{\beta i}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)}$$

And

$$e^{\frac{\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)} = \sum_{z=0}^{\infty} \frac{j^z \beta^z (\eta + i)^z}{z! \sigma^z} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)^z$$

Then

$$f(x)^\eta = \sum_{i=z=0}^{\infty} \frac{\Gamma(\eta(\theta-1) + i) j^z \beta^z (\eta + i)^z}{i! \Gamma(\eta(\theta-1)) z! \sigma^z} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^\sigma\right)^z$$

Again using a generalized binomial, we get

$$f(x)^\eta = \mathbb{W} x^{-3\eta} e^{-\left(\frac{\gamma}{x}\right)^2} (\eta + q) \quad (18)$$

$$\text{Where } \mathbb{W} = \sum_{i=0}^{\infty} \frac{2\theta^\eta \beta^\eta \gamma^{2\eta} \Gamma(\eta(\theta-1) + i) j^z \beta^z (\eta + i)^z (-1)^p \Gamma((\sigma+1)\eta + \sigma p + q)}{i! \Gamma(\eta(\theta-1)) z! \sigma^z q! \Gamma((\sigma+1)\eta + \sigma p)} \binom{z}{p}$$

By substituting equation (18) into the equation above, we get:

$$T_R(\eta) = \frac{1}{1-\eta} \log \int_0^\infty \mathbb{W} x^{-3\eta} e^{-\left(\frac{\gamma}{x}\right)^2} (\eta + q) dx$$

The last integral, we get

$$T_R(\eta) = \frac{1}{1-\eta} \log \left(\frac{\eta \Gamma(\frac{3}{2}(\eta-1)+1)}{2((\eta+q)\gamma^2)^{\frac{1}{2}(3\eta+1)}} \right) \quad (19)$$

2.3.6 Probability Weighted Moments

The probabilistic weighted moment $(\varphi, \eta)^{th}$ for EtGoIR distributions can be expressed as follows:

$$\rho_{(\varphi, \eta)} = E \left(X^\varphi (F^\eta(X)) \right) = \int_{-\infty}^{\infty} x^\varphi F^\eta(x) f(x) dx$$

By equation (5), we can find $F^\eta(x)$:

$$F(x)^\eta = 1 - e^{-\frac{\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)^{\theta \eta}}$$

By generalized binomial series:

$$1 - e^{-\frac{\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)^{\theta \eta}} = \sum_{j=0}^{\infty} (-1)^j \binom{\theta \eta}{j} e^{-\frac{j\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)}$$

And using exponential expansion

$$e^{-\frac{j\beta}{\sigma} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)} = \sum_{r=0}^{\infty} \frac{j^r \beta^r}{r! \sigma^r} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)^r$$

Then

$$F(x)^\eta = \sum_{j=r=0}^{\infty} \frac{(-1)^j j^r \beta^r}{r! \sigma^r} \binom{\theta \eta}{j} \left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)^r$$

And using generalized binomial

$$\left(1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right)^r = \sum_{s=0}^{\infty} (-1)^s \binom{r}{s} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-r\sigma}$$

$$\text{And } \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-r\sigma} = \sum_{w=0}^{\infty} \frac{\Gamma(r\sigma + w)}{w! \Gamma(r\sigma)} e^{-\left(\frac{\gamma}{x}\right)^2 w}$$

Then

$$F(x)^\eta = \mathbb{K} e^{-\left(\frac{\gamma}{x}\right)^2 w} \quad (20)$$

$$\text{Where } \mathbb{K} = \sum_{j=r=s=w=0}^{\infty} \frac{(-1)^{j+s} j^r \beta^r \Gamma(r\sigma + w)}{r! \sigma^r w! \Gamma(r\sigma)} \binom{\theta \eta}{j} \binom{r}{s}$$

By substituting equation (20) into the equation above, we get:

$$\rho_{(\varphi, \eta)} = \mathbb{K} \psi \int_{-\infty}^{\infty} x^{\varphi-3} e^{-(w+m+1)\left(\frac{\gamma}{x}\right)^2} dx$$

Let $u = (w + m + 1) \left(\frac{\gamma}{x}\right)^2$ then

$$\rho_{(\varphi, \eta)} = \frac{\mathbb{K} \psi \Gamma\left(1 - \frac{\varphi}{2}\right)}{2((w+m+1)\gamma^2)^{1 - \frac{\varphi}{2}}} \quad (21)$$

2.3.7 Survival function

$$S(x) = 1 - F(x) \tag{22}$$

$$S(x) = 1 - \left[1 - e^{-\left(\frac{\beta}{\sigma}\right) \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}^\theta} \right] \tag{23}$$

2.3.8 Hazard function

$$H(x) = \frac{f(x)}{S(x)} \tag{24}$$

$$H(x) = \frac{2\theta\beta\gamma^2 x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma-1} \left(\frac{\beta}{\sigma}\right) \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}^{\theta-1}}{\left[1 - e^{-\left(\frac{\beta}{\sigma}\right) \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}^\theta} \right]} \tag{25}$$

The hazard plot of the EtGoIR distribution is given in Figure 2 below.

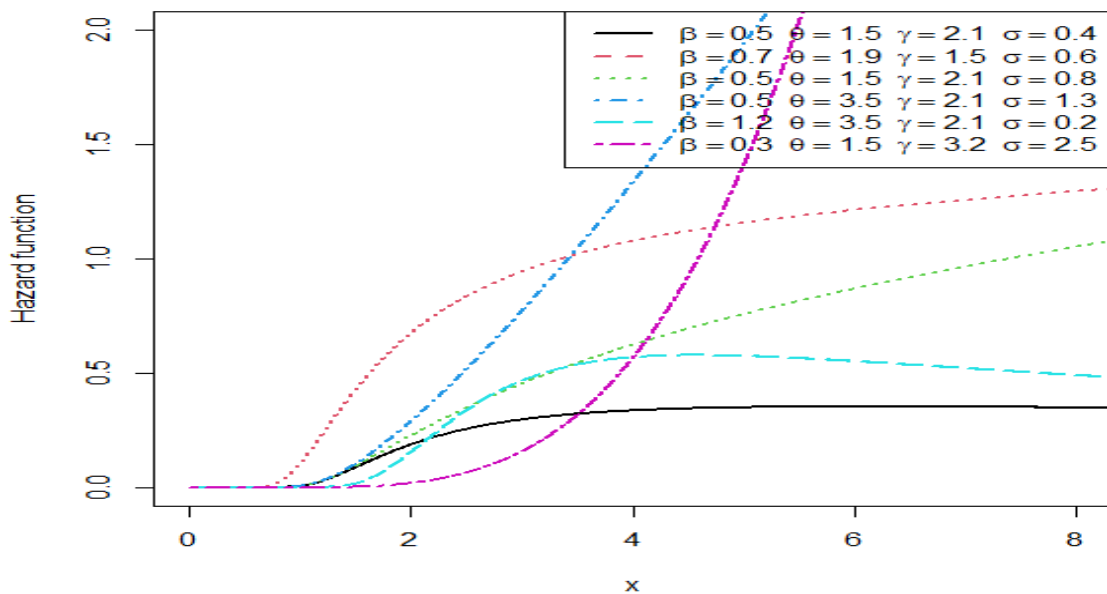


Figure 2: plot of hazard function of EtGoIR distribution

2.3.9 Cumulative hazard function

$$C(x) = -\log[S(x)] \tag{26}$$

$$C(x) = -\log \left[1 - \left[1 - e^{-\left(\frac{\beta}{\sigma}\right) \left\{ 1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right)^{-\sigma} \right\}^\theta} \right] \right] \tag{27}$$

2.3.10 Reverse hazard function

$$R(x) = \frac{f(x)}{F(x)} \tag{28}$$

$$R(x) = \frac{2\theta\beta\gamma^2 x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}}}{\left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma-1} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}}}\right)^{\theta-1}} \tag{29}$$

2.3.11 Order Statistics

The pdf of the r^{th} order statistics of $X_{r:n}$ is given as:

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} \sum_{i=0}^{n-r} (-1)^i [F(x)]^{r+i-1} f(x) \tag{30}$$

Inserting equation (5) and equation (6) into equation (30), we have

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} \sum_{i=0}^{n-r} (-1)^i \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma-1} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{\theta-1} 2\theta\beta\gamma^2 x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{\theta-1}$$

On bringing the like terms together, we have

$$f_{r:n}(x) = \frac{2\theta\beta\gamma^2}{B(r, n-r+1)} \sum_{i=0}^{n-r} (-1)^i x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{\theta(r+i)-1} \tag{31}$$

Using the generalized binomial expansion on the last term in equation (31), we have

$$\left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{\theta(r+i)-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\theta(r+i))}{j! \Gamma(\theta(r+i)-j)} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^j$$

Substituting back into equation (31), we have

$$f_{r:n}(x) = \frac{2\theta\beta\gamma^2}{B(r, n-r+1)} \sum_{i=0}^{n-r} \sum_{j=0}^{\infty} \frac{(-1)^{i+j} \Gamma(\theta(r+i))}{j! \Gamma(\theta(r+i)-j)} x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{j+1} \tag{32}$$

Also, expanding the last term in equation (32), we have

$$\left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma} e^{\left(\frac{\beta}{\sigma}\right)\left\{1 - \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma}\right\}} \right)^{j+1} = \sum_{k=0}^{\infty} (-1)^k \binom{j+1}{k} \left(1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right)^{-\sigma k}$$

$$f_{r:n}(x) = \frac{2\theta\beta\gamma^2}{B(r, n-r+1)} \sum_{i=0}^{n-r} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^{i+j+1} \binom{j+1}{k} \Gamma(\theta(r+i))}{j! \Gamma(\theta(r+i)-j)} x^{-3} e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma(k+1)-1} \quad (33)$$

$$\left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma(k+1)-1} = \sum_{l=0}^{\infty} \frac{(-1)^l \Gamma(-\sigma(k+1))}{l! \Gamma(-\sigma(k+1)-l)} \left[e^{-\left(\frac{\gamma}{x}\right)^2}\right]^l$$

Putting all the expansions together, we have the r^{th} order statistics of EtGoIR distribution given as:

$$f_{r:n}(x) = \frac{2\theta\beta\gamma^2}{B(r, n-r+1)} \sum_{i=0}^{n-r} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{i+j+k+1} \binom{j+1}{k} \Gamma(-\sigma(k+1)) \Gamma(\theta(r+i))}{j! l! \Gamma(\theta(r+i)-j) \Gamma(-\sigma(k+1)-l)} x^{-3} \left[e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{i+1} \quad (34)$$

To obtain minimum order statistics for EtGoIR distribution, we set $r=1$ in equation (34) to get

$$f_{1:n}(x) = 2n\theta\beta\gamma^2 \sum_{i=0}^{n-1} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{i+j+k+1} \binom{j+1}{k} \Gamma(-\sigma(k+1)) \Gamma(\theta(1+i))}{j! l! \Gamma(\theta(1+i)-j) \Gamma(-\sigma(k+1)-l)} x^{-3} \left[e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{i+1} \quad (35)$$

To obtain maximum order statistics for EtGoIR distribution, we set $r=n$ in equation (34) to get

$$f_{n:n}(x) = 2n\theta\beta\gamma^2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^{j+k+1} \binom{j+1}{k} \Gamma(-\sigma(k+1)) \Gamma(\theta(n))}{j! l! \Gamma(\theta(n)-j) \Gamma(-\sigma(k+1)-l)} x^{-3} \left[e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{i+1} \quad (36)$$

2.4 Maximum Likelihood Estimation (MLE)

Given some observed data, a method known as maximum likelihood estimation (MLE) can be used to estimate a probability distribution's parameters. This is accomplished by maximizing a likelihood function to make the observed data as probable as possible given the assumed statistical model. The log-likelihood function of EtGoIR is given as

$$\begin{aligned} \log L = & n \log(2) + n \log(\theta) + n \log(\beta) + 2n \log(\gamma) - 3 \sum_{i=1}^n \log(x) - \gamma^2 \sum_{i=1}^n \left(\frac{1}{x}\right)^2 - (\sigma + \\ & 1) \sum_{i=1}^n \log \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right] + \frac{\beta}{\sigma} \sum_{i=1}^n \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma} \right] + (\theta - 1) \sum_{i=1}^n \left[1 - e^{-\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma} \right]} \right] \end{aligned} \quad (37)$$

The maximum likelihood estimate is the location in the parameter space where the likelihood function is maximized. The maximum likelihood estimates of θ , β , γ and σ are the values that maximize the likelihood function. We can find these values by taking the partial derivatives of the likelihood function with respect to θ , β , γ , σ and setting them equal to zero. This gives us the following equations:

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log \left[1 - e^{-\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma} \right]} \right] = 0 \quad (38)$$

$$\frac{\partial \log L}{\partial \beta} = \frac{n}{\beta} + \frac{1}{\sigma} \sum_{i=1}^n \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma} \right] - (\theta - 1) \sum_{i=1}^n \frac{e^{-\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma} \right]}}{\left[1 - e^{-\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2} \right]^{-\sigma} \right]} \right]} = 0 \quad (39)$$

$$\frac{\partial \log l}{\partial \gamma} = \frac{2n}{\gamma} + 2\gamma \sum_{i=1}^n \left(\frac{1}{x}\right)^2 + (\sigma + 1) \sum_{i=1}^n \frac{\gamma e^{-\left(\frac{\gamma}{x}\right)^2}}{1 - e^{-\left(\frac{\gamma}{x}\right)^2}} + \frac{\beta}{\sigma} \sum_{i=1}^n 2\sigma \gamma e^{-\left(\frac{\gamma}{x}\right)^2} \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma-1} + (\theta -$$

$$1) \sum_{i=1}^n \frac{e^{\left(\frac{\beta}{\sigma}\right) 2\sigma \gamma e^{-\left(\frac{\gamma}{x}\right)^2}} e^{\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right]}{1 - e^{\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right}}} = 0 \quad (40)$$

$$\frac{\partial \log l}{\partial \sigma} = \frac{\beta}{\sigma} \sum_{i=1}^n \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma} \log \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right] + \frac{\beta}{\sigma^2} \sum_{i=1}^n \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right] \sum_{i=1}^n \log \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right] + (\theta -$$

$$1) \sum_{i=1}^n \frac{e^{-\left(\frac{\beta}{\sigma}\right) \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}} \log \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right] - e^{-\left(\frac{\beta}{\sigma^2}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right]}}{1 - e^{\left(\frac{\beta}{\sigma}\right) \left[1 - \left[1 - e^{-\left(\frac{\gamma}{x}\right)^2}\right]^{-\sigma}\right}}} = 0 \quad (41)$$

Since equations (38), (39), (40) and (41) are non-linear in parameters, techniques such as Newton-Raphson method in R-software can be used to accomplish the task of estimating the parameters from equations (38), (39), (40) and (41).

III. Results

3.1 Simulation

In this section, we conduct a simulation study to assess the performance of the Maximum Likelihood Estimation (MLE) for the EtGoIR distribution. We generate random numbers using the quantile function (qf) of the distribution. Specifically, if U is a uniform random variable on the interval (0, 1), then x follows the EtGoIR distribution. We generated a total of n = 10000 samples, with each sample having sizes n=20, 50, 100, 250, 500, and 1000. These samples were drawn from the EtGoIR distribution using its quantile function. Subsequently, we calculated the empirical means, biases, and root mean squared errors (RMSE) of the MLE.

Table.1 MLEs, biases and RMSE for some values of parameters

n	Parameters	(0.5,0.1,0.1,0.5)			(2,1,3,2.5)		
		Estimated Values	Bias	RMSE	Estimated Values	Bias	RMSE
20	θ	0.4548	-0.0452	0.1484	2.2647	0.2647	0.9692
	β	0.1266	0.0266	0.0976	1.0825	0.0825	0.5743
	γ	0.1262	0.0262	0.0579	3.0253	0.0253	0.2659
	σ	0.5770	0.0770	0.1908	2.7354	0.2354	0.9156
50	θ	0.4737	-0.0263	0.1151	2.1251	0.1251	0.6905
	β	0.1075	0.0075	0.0503	1.0966	0.0966	0.4272
	γ	0.1110	0.0110	0.0310	3.0438	0.0438	0.1938
	σ	0.5366	0.0366	0.1216	2.5940	0.0940	0.6200

100	θ	0.4890	-0.0110	0.0903	2.0670	0.0670	0.4628
	β	0.1035	0.0035	0.0338	1.0951	0.0951	0.3017
	γ	0.1054	0.0054	0.0193	3.0519	0.0519	0.1539
	σ	0.5185	0.0185	0.0889	2.5425	0.0425	0.4413
250	θ	0.4972	-0.0028	0.0665	2.0166	0.0166	0.2872
	β	0.1006	0.0006	0.0227	1.0665	0.0665	0.2210
	γ	0.1019	0.0019	0.0126	3.0435	0.0435	0.1115
	σ	0.5097	0.0097	0.0630	2.5241	0.0241	0.2873
500	θ	0.5017	0.0017	0.0511	2.0052	0.0052	0.1923
	β	0.1012	0.0012	0.0160	1.0511	0.0511	0.1606
	γ	0.1006	0.0006	0.0091	3.0318	0.0318	0.0825
	σ	0.5012	0.0012	0.0415	2.5051	0.0051	0.1930
1000	θ	0.5028	0.0028	0.0370	2.0010	0.0010	0.1367
	β	0.1010	0.0010	0.0105	1.0434	0.0434	0.1208
	γ	0.1002	0.0002	0.0064	3.0288	0.0288	0.0727
	σ	0.5000	0.0001	0.0289	2.5048	0.0048	0.1400

Table 1 presents the simulation outcomes corresponding to the EtGoIR distribution. It is observed that as the sample size increases, the Root Mean Square Error (RMSE) and bias associated with the parameter estimators consistently decreases. The outcome suggest that the model is consistent.

3.2 Applications

This section demonstrates the practical application of the EtGoIR distribution by utilizing it to model biomedical datasets. We compare its performance in providing a robust parametric fit to the datasets with that of the Gompertz Inverse Rayleigh (GoIR) distribution, the generalized Gompertz (GGo) distribution, the exponentiated exponential (EtEx) distribution, and the inverse Rayleigh (IR) distribution. Metrics such as the log likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) are employed for this comparison. To discern the most suitable model, computations of the log likelihood, AIC, and BIC values are carried out for both the proposed EtGoIR model and the alternative models used for comparison. The model exhibiting the lowest log likelihood, AIC, and BIC values is deemed the most appropriate match for the provided datasets. For this analytical endeavor, the R software is employed, facilitating the necessary calculations and comparisons.

Data set 1 has been utilized by [13] and [14]. The dataset comprises the summation of skinfold measurements from 202 athletes at the Australian Institute of Sports. It consists of the following values:

28.0, 98, 89.0, 68.9, 69.9, 109.0, 52.3, 52.8, 46.7, 82.7, 42.3, 109.1, 96.8, 98.3, 103.6, 110.2, 98.1, 57.0, 43.1, 71.1, 29.7, 96.3, 102.8, 80.3, 122.1, 71.3, 200.8, 80.6, 65.3, 78.0, 65.9, 38.9, 56.5, 104.6, 74.9, 90.4, 54.6, 131.9, 68.3, 52.0, 40.8, 34.3, 44.8, 105.7, 126.4, 83.0, 106.9, 88.2, 33.8, 47.6, 42.7, 41.5, 34.6, 30.9, 100.7, 80.3, 91.0, 156.6, 95.4, 43.5, 61.9, 35.2, 50.9, 31.8, 44.0, 56.8, 75.2, 76.2, 101.1, 47.5, 46.2, 38.2, 49.2, 49.6, 34.5, 37.5, 75.9, 87.2, 52.6, 126.4, 55.6, 73.9, 43.5, 61.8, 88.9, 31.0, 37.6, 52.8, 97.9, 111.1, 114.0, 62.9, 36.8, 56.8, 46.5, 48.3, 32.6, 31.7, 47.8, 75.1, 110.7, 70.0, 52.5, 67, 41.6, 34.8, 61.8, 31.5, 36.6, 76.0, 65.1, 74.7, 77.0, 62.6, 41.1, 58.9, 60.2, 43.0, 32.6, 48, 61.2, 171.1, 113.5, 148.9, 49.9, 59.4, 44.5, 48.1, 61.1, 31.0, 41.9, 75.6, 76.8, 99.8, 80.1, 57.9, 48.4, 41.8, 44.5, 43.8, 33.7, 30.9, 43.3, 117.8, 80.3, 156.6, 109.6, 50.0, 33.7, 54.0, 54.2, 30.3, 52.8, 49.5, 90.2, 109.5, 115.9, 98.5, 54.6, 50.9, 44.7, 41.8, 38.0, 43.2, 70.0, 97.2, 123.6, 181.7, 136.3, 42.3, 40.5, 64.9, 34.1, 55.7, 113.5, 75.7, 99.9, 91.2, 71.6, 103.6, 46.1, 51.2, 43.8, 30.5, 37.5, 96.9, 57.7, 125.9, 49.0, 143.5, 102.8, 46.3, 54.4, 58.3, 34.0, 112.5, 49.3, 67.2, 56.5, 47.6, 60.4, 34.9.

Data set 2, encompassing the remission times (in months) of a randomized collection of one hundred and twenty-eight (128) bladder cancer patients, has been utilized by [15] and [14]. The dataset comprises the following values:

0.08, 0.20, 0.40, 0.50, 0.51, 0.81, 0.90, 1.05, 1.19, 1.26, 1.35, 1.40, 1.46, 1.76, 2.02, 2.02, 2.07, 2.09, 2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.25, 3.31, 3.36, 3.36, 3.48, 3.52, 3.57, 3.64, 3.70, 3.82, 3.88, 4.18, 4.23, 4.26, 4.33, 4.34, 4.40, 4.50, 4.51, 4.87, 4.98, 5.06, 5.09, 5.17, 5.32, 5.32, 5.34, 5.41, 5.41, 5.49, 5.62, 5.71, 5.85, 6.25, 6.54, 6.76, 6.93, 6.94, 7.09, 7.26, 7.28, 7.32, 7.39, 7.59, 7.62, 7.63, 7.66, 7.87, 7.93, 8.26, 8.37, 8.53, 8.65, 8.66, 9.02, 9.22, 9.47, 10.06, 10.34, 10.66, 10.75, 11.25, 11.64, 11.79, 11.98, 12.02, 12.03, 12.07, 12.63, 13.11, 13.29, 13.80, 14.24, 14.76, 14.77, 14.83, 14.83, 15.96, 16.62, 17.12, 17.14, 17.36, 17.36, 18.10, 19.13, 20.28, 21.73, 22.69, 23.63, 25.74, 25.82, 26.31, 32.15, 34.26, 36.66, 43.01, 46.12, 79.05.

Data set 3, representing the survival times of one hundred and twenty-one (121) patients with breast cancer obtained from a large hospital during the period from 1929 to 1938, was obtained from [17]. The dataset is outlined as follows:

0.3, 0.3, 1.0, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0, 154.0.

Table 2: Summary Statistics of data

	N	Min.	Max.	Q1	Q2	Mean	Q3	Var.	SD	Ku	Sk
Data1	202	28.00	200.80	43.85	58.60	69.02	90.35	1060.501	32.565	4.365	1.175
Data2	128	0.080	79.050	3.348	6.395	9.366	11.838	110.425	10.508	18.485	3.286
Data3	121	0.30	154.00	17.30	40.00	46.08	60.00	1259.567	35.490	3.372	1.029

Table 2 demonstrate that the three datasets exhibit a high degree of skewness.

Table 3: The models' MLEs and performance requirements based on data set 1

Models	$\hat{\beta}$	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$	ll	AIC	BIC
EtGoIR	0.1985	369.5184	0.3036	0.1799	-953.632	1915.2650	1928.4980
GoIR	0.0031	-	0.0000	0.8601	-987.520	1981.0410	1990.9660
GGo	-0.0052	15.4031	-	0.0597	-956.086	1918.1730	1928.9200
EtEx	0.0406	8.5786	-	-	-958.006	1920.0130	1926.6300
IR	52.6054	-	-	-	-966.462	1934.9250	1938.2330

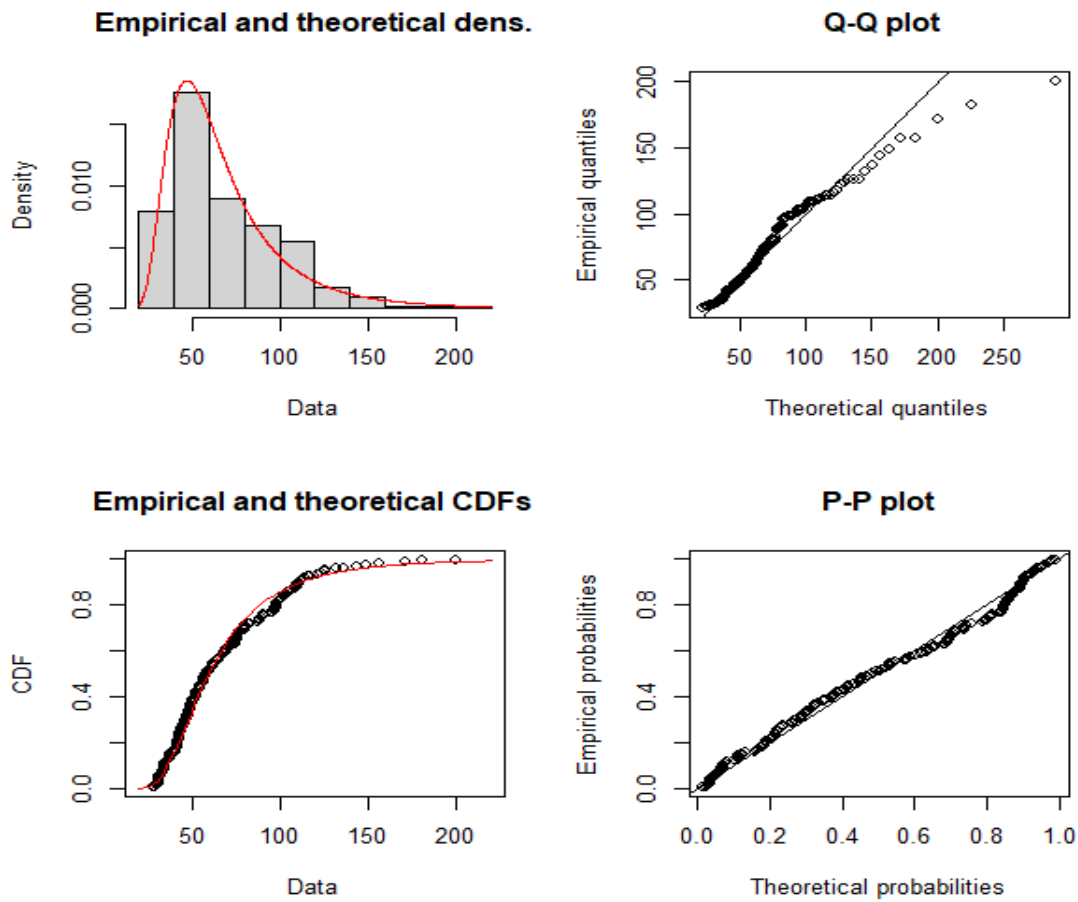


Figure 3: Density plots for data set 1

Table 4: The models' MLEs and performance requirements based on data set 2

Models	$\hat{\beta}$	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$	ll	AIC	BIC
EtGoIR	0.0003	2.5796	0.0001	0.3400	-410.704	829.4088	834.1479
GoIR	0.0839	-	0.0041	0.5129	-413.575	833.1505	836.1377
GGo	-0.0224	1.5034	-	0.1678	-413.183	832.3668	835.3539
EtEx	0.1213	1.2180	-	-	-413.077	830.1552	834.8592
IR	2.2612	-	-	-	-774.341	1550.683	1553.535

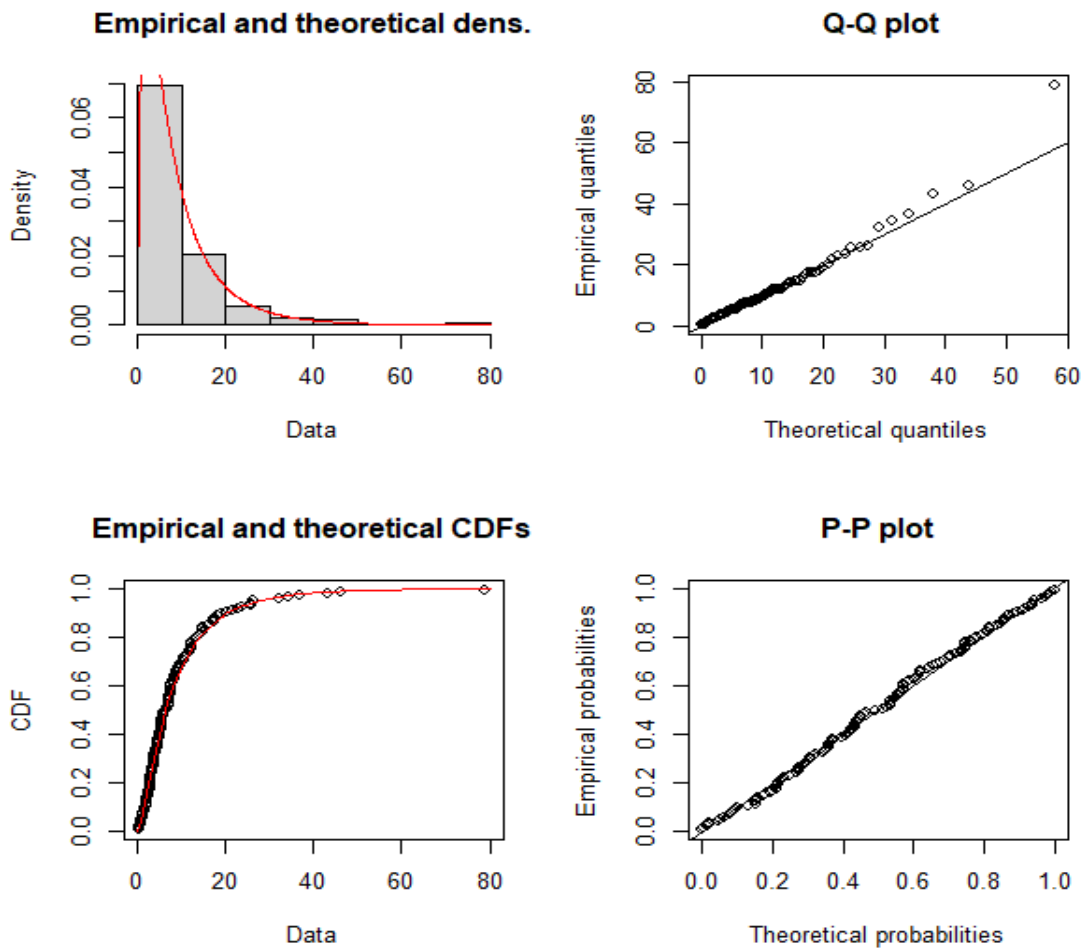


Figure 4: Density plots for data set 2

Table 5: The models' MLEs and performance requirements based on data set 3.

Models	$\hat{\beta}$	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\sigma}$	ll	AIC	BIC
EtGoIR	0.0000	0.5664	0.4016	0.9033	-578.7145	1165.4290	1176.6120
GoIR	0.0933	-	0.0002	0.6341	-579.9791	1165.9580	1176.7450
GGo	0.0066	1.1485	-	0.0182	-579.9435	1165.9371	1176.7274
EtEx	0.0269	1.4244	-	-	-581.7091	1167.4182	1168.2120
IR	2.2612	-	-	-	-1087.464	2176.9290	2179.7240

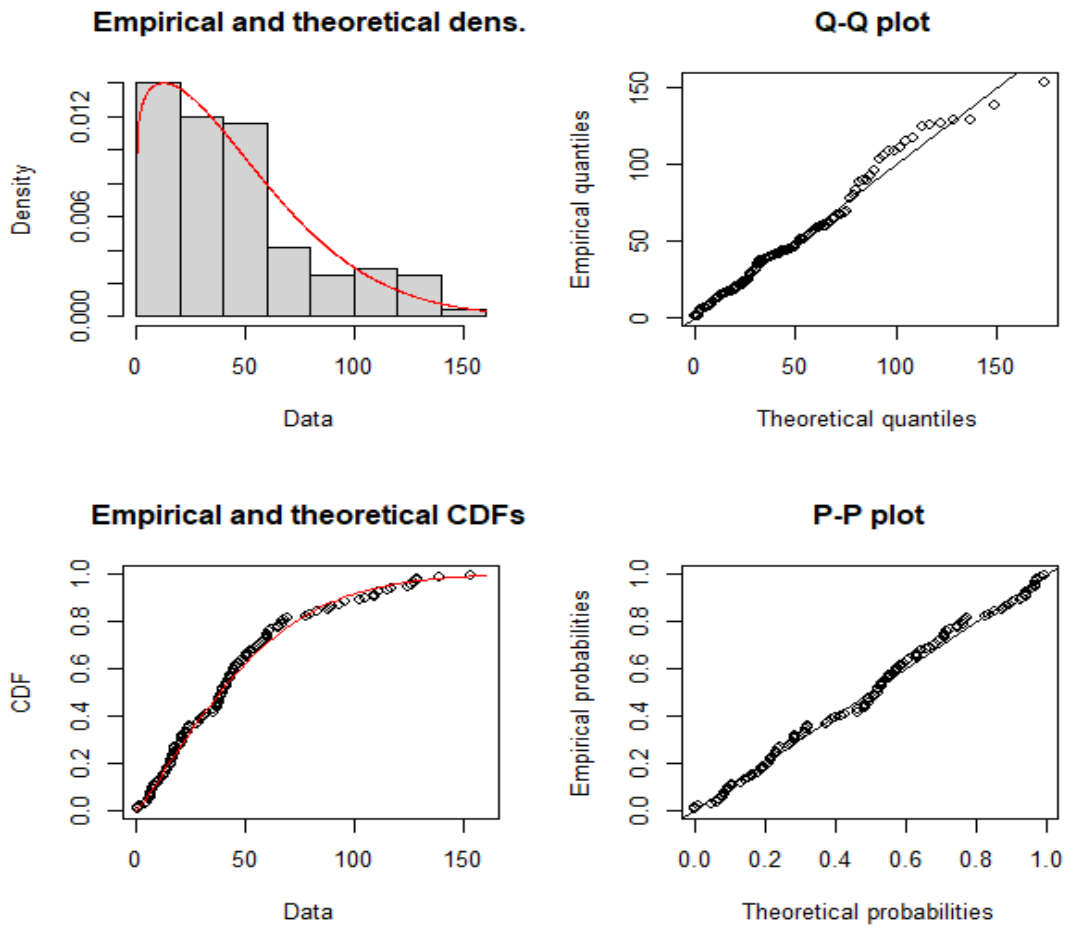


Figure 5: Density plots for data set 3.

Tables 3 to 5 showcase the superior ability of the proposed model to effectively fit the highly skewed datasets compared to the competing models, as indicated by the evaluation metrics employed. Figures 3 to 5 also showed that the proposed model fits the data set adequately.

IV. Discussion

This paper introduces a novel distribution termed the Exponentiated Gompertz Inverse Rayleigh (EtGoIR) distribution, extending the framework of the Gompertz Inverse Rayleigh (GoIR) distribution. The introduction of a new parameter enhances the distribution's adaptability in capturing various nuances present in biomedical datasets. The paper extensively examines the properties of the EtGoIR distribution, effectively demonstrating its practical applicability to real-life scenarios through the implementation of Maximum Likelihood Estimation (MLE). The empirical findings consistently substantiate that the proposed EtGoIR model outperforms the alternative distribution models under consideration in accurately fitting the provided datasets.

References

- [1] Isah A.M., Sule O.B., Kaigama A. and Khalaf A.A. (2024). Topp-Leone Exponentiated Burr XII distribution: Theory and Application to Real-life Data sets. *Iraqi Statisticians Journal*. 1(1): 63 – 72.
- [2] Sule O.B. (2024). A Study on Topp-Leone Kumaraswamy Fréchet Distribution with Applications: Methodological Study. *Turkish Journal of Biostatistics*. 16(1): 1 – 15.
- [3] Kumar, K. and Kumari A. (2023). Inferences for Two Inverse Rayleigh Populations Based on Joint Progressively Type-II Censored Data. *Computational Intelligence in Sustainable Reliability Engineering*, 159-179.
- [4] Leao, J., Saulo, H., Bourguignon, M., Cintra, R. J., Rego, L. C. and Cordeiro, G. M. (2022). On some properties of the beta inverse Rayleigh distribution. *arXiv preprint*,
- [5] Sule O.B., Ibrahim I.I. and Isa A.M. (2023). On the properties of generalized Rayleigh distribution with applications. *Reliability: theory & applications*. 18(3): 374-386.
- [6] Sule O.B. and Halid O.Y (2023). On the Properties and Applications of Topp-Leone Gompertz Inverse Rayleigh distribution. *Reliability: theory & applications*. 18(4): 1032 – 1045.
- [7] Sule O.B. and Halid O.Y (2023). On Gompertz Exponentiated Inverse Rayleigh Distribution. *Reliability: Theory & Applications*. 18(1): 412-424.
- [8] Adegoke T.M., Oladoja O.M., Sule O.B., Mustapha A.A., Aderupatan D.E and Nzei L.C. (2023). Topp-Leone Inverse Gompertz Distribution: Properties and different estimations techniques and Applications. *Pakistan Journal of Statistics*. 39(4): 433 – 456.
- [9] Sule O.B. (2021). A New Extended Generalized Inverse Exponential Distribution: Properties and Applications. *Asian Journal of Probability and Statistics*. 11(2): 30 – 46.
- [10] Halid, O. Y and Sule B.O. (2022). A classical and Bayesian estimation techniques for Gompertz inverse Rayleigh distribution: properties and application, *Pakistan Journal of Statistics*, 38(1): 49-76.
- [11] Mudholkar, G.S. and Srivastava, D.K. (1993) Exponentiated Weibull Family for Analyzing Bathtub Failure-Rate Data. *IEEE Transactions on Reliability*, 42, 299-302
- [12] Khalaf, A. A. and khaleel, M. A. (2022). [0, 1] Truncated exponentiated exponential Gompertz distribution: Properties and applications. In AIP Conference Proceedings. 1(23): 070035.
- [13] Sule, O. B. (2023). A Study On The Properties of a New Exponentiated Extended Inverse Exponential Distribution with Applications. *Reliability: Theory & Applications*. 18(3 (74)): 59-72.
- [14] Almarashi, A. M., Elgarhy, M., Jamal, F. and Chesneau, C. (2022). The Exponentiated truncated inverse Weibull-generated family of distributions with applications. *Symmetry*. 12(4): 650.
- [15] Alotaibi, R., Baharith, L. A., Almetwally, E. M., Khalifa, M., Ghosh, I. and Rezk, H. (2022). Statistical Inference on a Finite Mixture of Exponentiated Kumaraswamy-G Distributions with Progressive Type II Censoring Using Bladder Cancer Data. *Mathematics*.10(15): 2800.
- [16] Chaudhary, A. K., and Kumar, V. (2020). The Logistic NHE Distribution with Properties and Applications. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*. 8(12): 591-603.
- [17] Lee, E. T. (1986). Statistical methods for survival data analysis. *IEEE Transactions on Reliability*, 35(1): 123-123.

BEHAVIOR ANALYSIS PRESENTED SYSTEM WITH FAILURE AND MAINTENANCE RATE WITH USING DEEP LEARNING ALGORITHMS

Shakuntla Singla ¹, Shilpa Rani ^{2*}, Diksha Mangla ³, Umar Muhammad Modibbo ⁴

•

^{1,2,3} Department of Mathematics and Humanities, MMEC, Maharishi Markandeshwar
(Deemed to be University), Mullana, Ambala, India

⁴Department of Statistics and Operations Research, School of Physical Sciences, Modibbo Adama
University of Technology, YOLA-NIGERIA

¹shakus25@gmail.com, ^{2*}gargshilpa46@gmail.com, ³dikshamangla1995@gmail.com,
⁴umarmodibbo@mau.edu.ng

Abstract

The paper discusses the behavioral analysis and dependability of a three-unit system utilizing RPGT for system parameters. Since all three units P, Q and R include parallel subcomponents, in the event that one of them fails, the system continues to operate although at a reduced capacity, but it is not profitable to run the system when two units are in reduced state hence considered failed state. The rates of failures are exponentially distributed, but the rates of repair are generalized, independent, and differ based on the operational unit. Fuzzy concept is used to declare/ determine whether the system is in failed/ reduced/ failed state. Graphs and tables are drawn to compare failure/repair effect on the parameters values. The system parameters are modelled using Regenerative Point graphical Technique (RPGT) and optimized using Deep learning methods such as Adam, SGD, RMS prop. The results of the optimization may be used to validate and challenge existing models and assumptions about the systems.

Keywords: MTSF, RPGT, Deep learning, Adaptive Moment Estimation, Stochastic Gradient Descent, RMS prop

I. Introduction

The paper analyzes system parameter reliability and behavioral analysis of three units using deep learning. Because the three units are all parallel subcomponents, the system may continue to function at a reduced capacity in the event that one or more fail. However, when two units work at a reduced capacity, the system is not profitable and is therefore deemed to be in a failed condition. The rates of failures exhibit exponential distribution, but the rates of repair are general, autonomous, and variable between different operational units. Units have varying capacity. The repairs are flawless. When two units are in a reduced state or any one unit is failing, the system is down. As in the system three are three units, P, Q, R all of which have parallel sub components initially when all the three units are good the system. The state S_0 [PQR], so upon their partial failures to states \bar{P} , \bar{Q} , \bar{R} the failure rates for which are λ_1 , λ_2 , λ_3 the system enters the reduced states S_1 [\bar{P} QR], S_2 [P \bar{Q} R], S_3 [PQ \bar{R}].

There is a single repairman available in the system that can repair all the three type of units from partial as well as from full failed states so from partially failed states upon repair of partial failed units

the system reenters the state so from state S_1 at repair rate w_1 of unit A, at a repair rate λ_3 of unit Q at a rate w_5 of unit R, from the partially states S_1, S_3, S_6 these units may fail further to full/ partial states. The system is considered to be in a failed state if one unit is in a failed state or if two or more units are in a reduced the state Fuzzy concept is used to declare/ determine whether the system is in failed/ reduced/ failed state. In state S_1 if unit P fails fully then the system enters the failure state S_2 [pQR]. if the unit Bar R fail then the system enters the failed states S_9 , or S_7 . In reduced state S_3 if units P, Q, R fail partially/ fully the system enters the states S_9, S_4, S_5 respectively from the failed state S_9 and S_4 upon repair of the units the system enters the state S_3 from partially reduced state S_6 upon failure of units P, Q, R the system enters the failed states S_7, S_5, S_8 respectively from these failed states as the repairman is free to repair the failed units, so as repair of these units the system again enters the state S_6 when more than one unit fail, the system is in failed state in these states the priority order of repairs is $\bar{P} > \bar{Q} > \bar{R}$. Taking the transition failure and repair rates the system may be stable in the states S_i ($0 \leq i \leq 9$) as shown in the figure 1.

Hsieh et al. [1] has discussed Reliability of two dimensions consecutive lower bounds system. John et al. [2] has study reliability multi hardware and software system multi-hardware–software system interaction failure less attention. Kumar [3] study investigated help of mathematical modelling find out of reliability. Kumar et al. [4] has study minimizing the risk of machine failure urea fertilizer plant. Kim. H.K et al. [5] discussed demonstrate transparent or flexible capacitive designed multi touch screen. Khan. M. F et al. [6] has study three stage mathematical formulation computational procedures, numerical has two distinct approaches. Singla et al. [7] have discussed comparison of availability of a pipe and sub system of independent failure. Raghav et al. [8] has study maximize the availability and minimize the cost of function with help of PSO. Singha. A. K. [9] has done the study of x -rays and computed tomography scans images of corona virus. Kumari et al. [10] have discussed with help of RPGT profit analysis of thresher plant three sub system blower, concave, hopper more result. Singla et al. [11] has study polytube manufacturing plant solve by using of RK method. Saliva et al. [12] has study failure probability comparison with usual 1- dimension model. Singla et al. [13] with mathematical model find availability under the reduces capacity using chapman Kolmogorov method. Singla et al. [14] all three units of different capacities in working in parallel in which two or three unit in full working. Singla et al. [15] with help of GA mathematical model depend availability with working time. Singla et al. [16] has study3 out of 4 good system optimizations modelled and analysis. An analysis on reliability parameters using an algorithm ABC, has been discussed by Ahmadini et al. [17]. Singla et al. [18] studied the two unit repairable system under the concept of fuzzy linguistic and discussed the overall availability.

The total of five sections are included in this study. The 2nd section includes model description with assumption used and different mathematical values used in the study. The methodology is covered in section 3rd. The results and conclusion is studied in section 4th and 5th respectively.

II. Assumption, Notation and Transformation Diagram

- The repair procedure arises soon after a unit flops.
- Repaired unit is as if a new one.
- Failure/repair rates of units are exponential.
- Server facility is 24x7 hours.
- $S_0 = PQR, S_1 = \bar{P}QR, S_2 = pQR, S_3 = P\bar{Q}R, S_4 = PqR,$
- $S_5 = P\bar{Q}\bar{R}, S_6 = P\bar{Q}\bar{r}, S_7 = \bar{P}Q\bar{R}, S_8 = PQR, S_9 = \bar{P}\bar{Q}R$

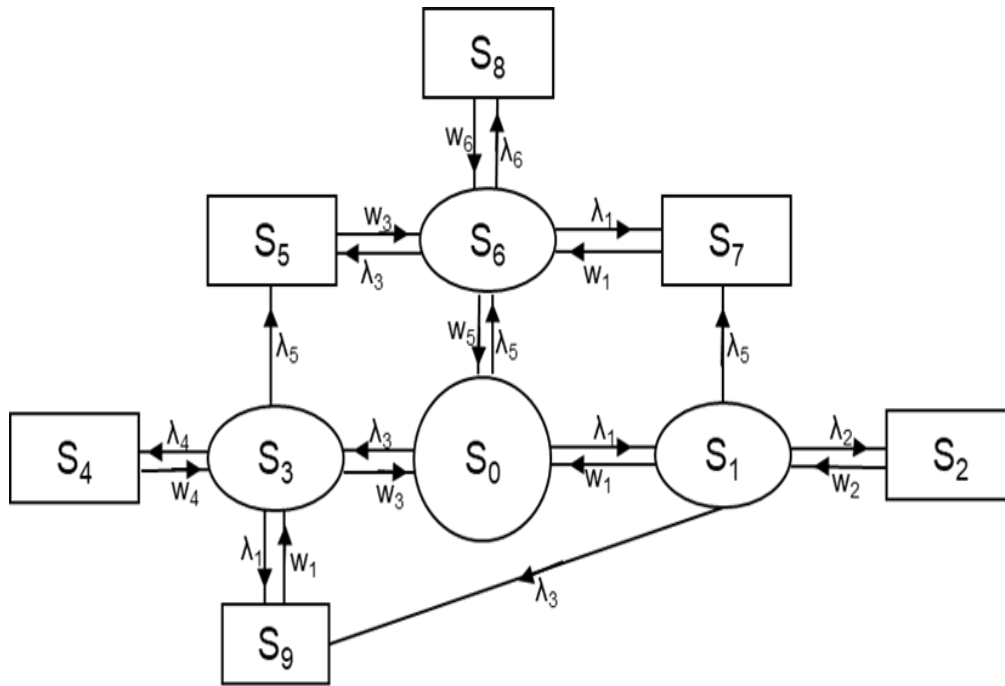


Figure 1: Transformation Diagram

2.1 Probability Density function (\$q_{ij}^{(t)}\$)

The probability density function associated with the transformation diagram from different states to other is given below.

$$\begin{aligned}
 q_{0,1} &= \lambda_1 e^{-(\lambda_1 + \lambda_5 + \lambda_3)t} \\
 q_{0,3} &= \lambda_3 e^{-(\lambda_1 + \lambda_5 + \lambda_3)t} \\
 q_{0,6} &= \lambda_5 e^{-(\lambda_1 + \lambda_5 + \lambda_3)t} \\
 q_{1,0} &= w_1 e^{-(\lambda_5 + \lambda_2 + \lambda_3 + w_1)t} \\
 q_{1,2} &= \lambda_2 e^{-(\lambda_5 + \lambda_2 + \lambda_3 + w_1)t} \\
 q_{1,7} &= \lambda_5 e^{-(\lambda_5 + \lambda_2 + w_1 + \lambda_3)t} \\
 q_{1,9} &= \lambda_3 e^{-(\lambda_5 + \lambda_2 + \lambda_3 + w_1)t} \\
 q_{2,1} &= w_2 e^{-w_2 t} \\
 q_{3,0} &= w_3 e^{-(w_3 + \lambda_1 + \lambda_5 + \lambda_4)t} \\
 q_{3,4} &= \lambda_4 e^{-(w_3 + \lambda_1 + \lambda_5 + \lambda_4)t} \\
 q_{3,5} &= \lambda_5 e^{-(w_3 + \lambda_1 + \lambda_5 + \lambda_4)t} \\
 q_{3,9} &= \lambda_1 e^{-(w_3 + \lambda_1 + \lambda_5 + \lambda_4)t} \\
 q_{4,3} &= w_4 e^{-w_4 t} \\
 q_{5,6} &= w_3 e^{-w_3 t} \\
 q_{6,0} &= w_5 e^{-(\lambda_3 + \lambda_6 + \lambda_1 + w_5)t} \\
 q_{6,5} &= \lambda_3 e^{-(\lambda_3 + \lambda_6 + \lambda_1 + w_5)t} \\
 q_{6,7} &= \lambda_1 e^{-(\lambda_3 + \lambda_6 + \lambda_1 + w_5)t} \\
 q_{6,8} &= \lambda_6 e^{-(\lambda_3 + \lambda_6 + \lambda_1 + w_5)t} \\
 q_{7,6} &= w_1 e^{-w_1 t} \\
 q_{8,6} &= w_6 e^{-w_6 t} \\
 q_{9,3} &= w_1 e^{-w_1 t}
 \end{aligned}$$

2.2 Cumulative probability density

Cumulative probability density functions in moving from state 'i' to state 'j' by taking Laplace Transforms of above function for infinite time interval is given as under.

$P_{ij} = q^*_{i,j}(t)$, i.e.

$$p_{0,1} = \lambda_1 / (\lambda_1 + \lambda_5 + \lambda_3)$$

$$p_{0,3} = \lambda_3 / (\lambda_1 + \lambda_5 + \lambda_3)$$

$$p_{0,6} = \lambda_5 / (\lambda_1 + \lambda_5 + \lambda_3)$$

$$p_{1,0} = w_1 / (\lambda_5 + \lambda_2 + \lambda_3 + w_1)$$

$$p_{1,2} = \lambda_2 / (\lambda_5 + \lambda_2 + \lambda_3 + w_1)$$

$$p_{1,7} = \lambda_5 / (\lambda_5 + \lambda_2 + w_1 + \lambda_3)$$

$$p_{1,9} = \lambda_3 / (\lambda_5 + \lambda_2 + \lambda_3 + w_1)$$

$$p_{2,1} = w_2 / w_2 = 1$$

$$p_{3,0} = w_3 / (w_3 + \lambda_1 + \lambda_5 + \lambda_4)$$

$$p_{3,4} = \lambda_4 / (w_3 + \lambda_1 + \lambda_5 + \lambda_4)$$

$$p_{3,5} = \lambda_5 / (w_3 + \lambda_1 + \lambda_5 + \lambda_4)$$

$$p_{3,9} = \lambda_1 / (w_3 + \lambda_1 + \lambda_5 + \lambda_4)$$

$$p_{4,3} = w_4 / w_4 = 1$$

$$p_{5,6} = w_3 / w_3 = 1$$

$$p_{6,0} = w_5 / (\lambda_3 + \lambda_6 + \lambda_1 + w_5)$$

$$p_{6,5} = \lambda_3 / (\lambda_3 + \lambda_6 + \lambda_1 + w_5)$$

$$p_{6,7} = \lambda_1 / (\lambda_3 + \lambda_6 + \lambda_1 + w_5)$$

$$p_{6,8} = \lambda_6 / (\lambda_3 + \lambda_6 + \lambda_1 + w_5)$$

$$p_{7,6} = w_1 / w_1 = 1$$

$$p_{8,6} = w_6 / w_6 = 1$$

$$p_{9,3} = w_1 / w_1 = 1$$

2.3 Mean Sojourn Transition rate ($R_i(t)$) for different states are

$$R_0^{(t)} = e^{-(\lambda_1 + \lambda_5 + \lambda_3)t}$$

$$R_1^{(t)} = e^{-(\lambda_5 + \lambda_2 + \lambda_3 + w_1)t}$$

$$R_2^{(t)} = e^{-w_2 t}$$

$$R_3^{(t)} = e^{-(w_3 + \lambda_1 + \lambda_5 + \lambda_4)t}$$

$$R_4^{(t)} = e^{-w_4 t}$$

$$R_5^{(t)} = e^{-w_3 t}$$

$$R_6^{(t)} = e^{-(\lambda_3 + \lambda_6 + \lambda_1 + w_5)t}$$

$$R_7^{(t)} = e^{-w_1 t}$$

$$R_8^{(t)} = e^{-w_6 t}$$

$$R_9^{(t)} = e^{-w_1 t}$$

2.4 Mean Sojourn Time ($\mu_i = R_i^*(0)$) for different states are

$$\mu_0 = 1 / (\lambda_1 + \lambda_5 + \lambda_3)$$

$$\mu_1 = 1 / (\lambda_5 + \lambda_2 + \lambda_3 + w_1)$$

$$\mu_2 = 1 / w_2$$

$$\mu_3 = 1 / (w_3 + \lambda_1 + \lambda_5 + \lambda_4)$$

$$\mu_4 = 1 / w_4$$

$$\mu_5 = 1 / w_3$$

$$\mu_6 = 1 / (\lambda_3 + \lambda_6 + \lambda_1 + w_5)$$

$$\mu_7 = 1 / w_1$$

$$\mu_8 = 1 / w_6$$

$$\mu_9 = 1 / w_1$$

2.5 Transition Probability

$$V_{0,0} = 1 \text{ (Verified)}$$

$$\begin{aligned} V_{0,1} &= (0,1)/1-(1,2,1) \\ &= p_{0,1}/(1-p_{1,2}p_{2,1}) \\ &= (\lambda_1/\lambda_5+\lambda_2+\lambda_3+w_1)/(\lambda_1+\lambda_5+\lambda_3) (\lambda_5+\lambda_3+w_1) \end{aligned}$$

$$\begin{aligned} V_{0,2} &= (0,1,2)/1-(1,2,1) \\ &= (p_{0,1}p_{1,2}/1-p_{1,2}p_{2,1}) \\ &= \lambda_1\lambda_2/(\lambda_1+\lambda_5+\lambda_3) (\lambda_5+\lambda_3+w_1) \end{aligned}$$

$$\begin{aligned} V_{0,3} &= (0,3)/1-(3,4,3)1-(3,9,3) + (0,1,9,3)/1-(1,2,1) \{1-(9,3,9)/1-(3,4,3)\}1-(3,4,3) \\ &= p_{0,3}/(1-p_{3,4}p_{4,3}) (1-p_{3,9}p_{9,3}) + p_{0,1}p_{1,9}p_{9,3}/(1-p_{1,2}p_{2,1}) \{(1-p_{3,4}p_{4,3}-p_{9,3}p_{3,9})/(1-p_{3,4}p_{4,3})\} \\ &\quad (1-p_{3,4}p_{4,3}) \\ &= \lambda_3(\lambda_1+\lambda_4+\lambda_5+w_3)^2/(\lambda_1+\lambda_5+w_3) (\lambda_1+\lambda_3+\lambda_5) \{(\lambda_3\lambda_5+\lambda_3w_3+\lambda_5^2+\lambda_5w_3+w_1w_3+\lambda_1w_3+\lambda_1\lambda_4 \\ &\quad +\lambda_1\lambda_5)/(w_3+\lambda_4+\lambda_5) (\lambda_3+\lambda_5+w_7) (\lambda_5+w_3)\} \end{aligned}$$

$$\begin{aligned} V_{0,4} &= (0,3,4)/1-(3,4,3)1-(3,9,3) + (0,1,9,3,4)/1-(1,2,1) \{1-(9,3,9)/1-(3,4,3)\}1-(3,4,3) \\ &= p_{0,3}p_{3,4}/(1-p_{3,4}p_{4,3}) (1-p_{3,9}p_{9,3}) + p_{0,1}p_{1,9}p_{9,3}p_{3,4}/(1-p_{1,2}p_{2,1})(1-p_{3,4}p_{4,3}-p_{9,3}p_{3,9}) \\ &= \lambda_3\lambda_4/(\lambda_1+\lambda_5+\lambda_3)\{(w_3+\lambda_1+\lambda_5+\lambda_4)(\lambda_5+\lambda_3+w_1)(w_3+\lambda_5)+\lambda_1(w_3+\lambda_1+\lambda_5)(w_3+\lambda_4+\lambda_5)/ \\ &\quad (w_3+\lambda_1+\lambda_5)(w_3+\lambda_4+\lambda_5)(\lambda_5+\lambda_3+w_1)(\lambda_5+w_3)\} \end{aligned}$$

III. Methodology

3.1 MTSF(T₀)

Initial state '0', before joining down state are: 'i' = 0,1,3,6 taking initial state 'ξ' = '0'

$$MTSF(T_0) = \left[\sum_{i,sr} \left\{ \frac{\left\{ \text{pr} \left(\xi^{sr(sff)} \right) \right\} \mu_i}{\prod_{m_1 \neq \xi} \{1-V_{m_1 m_1}\}} \right\} \right] \div \left[1 - \sum_{sr} \left\{ \frac{\left\{ \text{pr} \left(\xi^{sr(sff)} \right) \right\}}{\prod_{m_2 \neq \xi} \{1-V_{m_2 m_2}\}} \right\} \right]$$

3.2 Availability of the System

The regenerative states at which the system is available are 'j' = 0,1,3,6 and the regenerative states are 'i' = 0 to 9 taking 'ξ' = '0' the availability for which the system is available is given by

$$A_0 = \left[\sum_{j,sr} \left\{ \frac{\{ \text{pr}(\xi^{sr \rightarrow j}) \} f_j, \mu_j}{\prod_{m_1 \neq \xi} \{1-V_{m_1 m_1}\}} \right\} \right] \div \left[\sum_{i,sr} \left\{ \frac{\{ \text{pr}(\xi^{sr \rightarrow i}) \} \mu_i^1}{\prod_{m_2 \neq \xi} \{1-V_{m_2 m_2}\}} \right\} \right]$$

$$A_0 = \left[\sum_j V_{\xi,j}, f_j, \mu_j \right] \div \left[\sum_i V_{\xi,i}, f_j, \mu_i^1 \right]$$

3.3 Busy Period of the Server

The states where the server is busy for doing some job are 'i' = 1 to 9, taking 'ξ' = '0', using RPQT busy period is given as

$$B_0 = \left[\sum_{j,sr} \left\{ \frac{\{ \text{pr}(\xi^{sr \rightarrow j}) \} n_j}{\prod_{m_1 \neq \xi} \{1-V_{m_1 m_1}\}} \right\} \right] \div \left[\sum_{i,sr} \left\{ \frac{\{ \text{pr}(\xi^{sr \rightarrow i}) \} \mu_i^1}{\prod_{m_2 \neq \xi} \{1-V_{m_2 m_2}\}} \right\} \right]$$

$$B_0 = \left[\sum_j V_{\xi,j}, n_j \right] \div \left[\sum_i V_{\xi,i}, \mu_i^1 \right]$$

3.4 Expected Fractional Number of repairman's Visits

States 1, 3, and 6 are the regeneration states that the repairman visits first to complete this task. The repairman's visitation count is determined by

$$V_0 = \left[\sum_{j,sr} \left\{ \frac{\{ \text{pr}(\xi^{sr \rightarrow j}) \}}{\prod_{k_1 \neq \xi} \{1-V_{k_1 k_1}\}} \right\} \right] \div \left[\sum_{i,sr} \left\{ \frac{\{ \text{pr}(\xi^{sr \rightarrow i}) \} \mu_i^1}{\prod_{k_2 \neq \xi} \{1-V_{k_2 k_2}\}} \right\} \right] = \left[\sum_j V_{\xi,j} \right] \div \left[\sum_i V_{\xi,i}, \mu_i^1 \right]$$

3.5 Dataset: Behavior analysis Using Deep Learning Algorithms

To perform optimization using deep learning, you would need a dataset that contains information on the input parameters and the system's output [5, 6]. The input parameters could include factors such as the system's design, operating conditions, and maintenance schedule. The output could include metrics such as system availability, MTSF, and busy period.

Table 1: *Parameter*

$W(w_1, w_2, \dots, w_n)$ (0-100)	$\lambda(\lambda_1, \lambda_2, \dots, \lambda_n)$ (0-100)	$S(s_1, s_2, \dots, s_n)$ (0-100)	p (0-.68)
--------------------------------------	--	--------------------------------------	----------------

Table 2: *Performance of model*

Model	MTSF	Expected Number of Inspections by the repair man	Busy Period	Availability
Adam	0.928	.9067	0.8021	.9348
Sgd	.9128	.9007	.8123	.9128
MS prop	.9013	0.8710	.8101	0.9234

- **Collection of data:** Gather a dataset that contains information on the input parameters and the system's output. The input parameters could include factors such as the system's design, operating conditions, and maintenance schedule. The output could include metrics such as system availability, downtime, and failure rate in table 1 and table 2.
- **Preprocess data:** Clean and preprocess the dataset, splitting it into training, validation, and test sets.
- **Train the model:** Use a deep learning algorithm, such as a neural network, to model the connection among the input parameters and the output. Train the model using the training set and validate it using the set of values in table 1. You could use techniques such as early stopping and regularization to prevent over fitting.
- **Appraise the model:** After the model is proficient, appraise its performance by means of test set. Estimate metrics such as busy period.
- **Perform sensitivity analysis:** Using the trained model, vary the values of one parameter at a time while keeping the others constant. Record the effect on the system's output. Repeat this process for each input parameter, recording the impact of each parameter on the system's output. The output could include metrics such as system availability, MTSF, and busy period.
- **Once you have a dataset,** you could use a deep learning algorithm to model the relationship among the input parameters and the production. One approach could be to use a neural network, which can learn complex relationships between inputs and outputs. To perform optimization using a neural network, you could first train the network on the dataset, using a portion of the data for training and another portion for validation.

Optimization of a repairable system undertaken for analysis using deep learning typically involves the following steps:

- **Data collection:** Collect data on the input parameters and output metrics of the system. The input parameters could include factors such as the system's design, operating conditions, and maintenance schedule. The output metrics could include measures such as system availability, MTSF, and busy period in show table 2 included.
- **Data preprocessing:** Clean and preprocess the data, splitting it into training, validation, and test sets. Normalize the input variables to ensure that they are on the same scale.
- **Model selection:** Choose appropriate deep learning optimization techniques (Adam, SGD, RMS prop) for the sensitivity analysis. Some options contain feed forward neural systems, convolutional neural systems, and regular neural networks. Consider influences such as the size of the dataset, the difficulty of the input-output connection, and the computational capitals existing.
- **Model training:** Train the selected model on the training data. Use techniques such as stochastic gradient descent and back propagation to minimize the bust time. Monitor the performance of the model on the validation data, and adjust the hyper parameters as needed.
- **Model evaluation:** Evaluate the trained model on the test data. Calculate metrics such as mean absolute bust time and mean squared error to assess the model's performance of deep learning optimization in show table 1 and table 2.

IV. Results and discussion

The results and discussion of a Optimization of undertaken repairable system parameters using deep learning will depend on the specific system and dataset analyzed. However, here are general insights that could be gained from such an analysis:

- **Identification of critical system parameters:** The optimization could reveal which input parameters require the greatest effect on the output metric of interest. For example, it could show that system availability is most optimization to the frequency of care or the quality of the components used in the organization.
- **Understanding of the non-linear relationship amongst input strictures and output metrics:** The deep learning model used in the analysis can capture non-linear relationships amongst input restrictions and output metrics, which could not detect using traditional statistical methods. The optimization can provide insights into the shape and magnitude of these relationships.
- **Validation of existing models and assumptions:** The optimization's outcomes were used to support or refute preexisting theories and hypotheses about the system. The research may reveal, for instance, that a particular parameter significantly affects system performance more than previously believed.

Prediction of system behavior under different scenarios: The deep learning model applied to predict system performance under different setups, such as vagaries in operating conditions or maintenance schedules. This can support decision-makers assess the impact of changed strategies and style informed verdicts. Overall, Behavior analysis presented syysem with failure and maintenance rate Using Deep Learning Methods can provide valuable insights into the factors

that affect system performance, (MTSF), Expected Fractional Number of repairman's Visits Busy Period and Availability of the System are shown in figure 2, 3, 4 and 5. valuable insights into the factors that affect system performance, (MTSF), Expected Fractional Number of repairman's Visits Busy Period and Availability of the System are shown in figure 2, 3, 4 and 5.

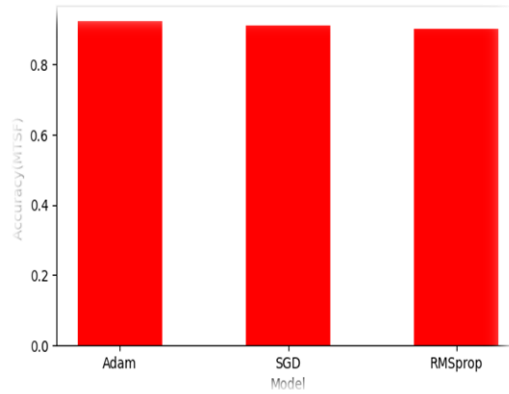


Figure 2: comparing between models according to MTSF

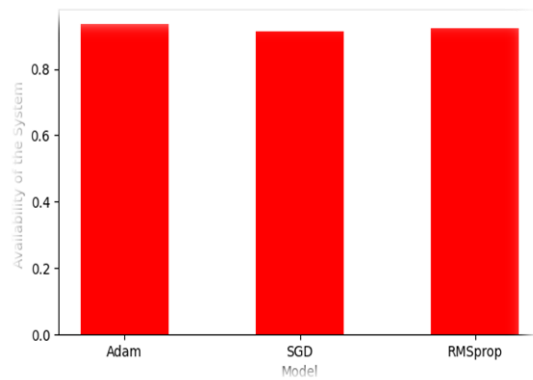


Figure 3: comparing between models according to Availability

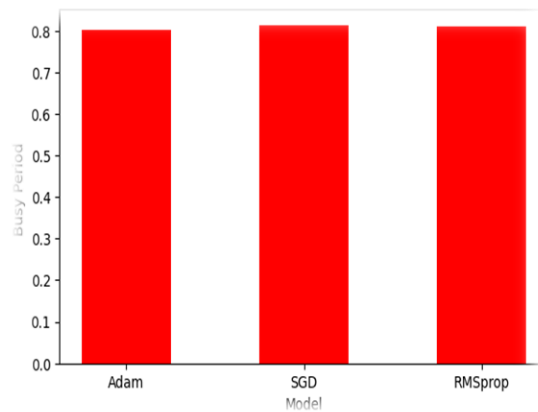


Figure 4: comparing between models according to busy period

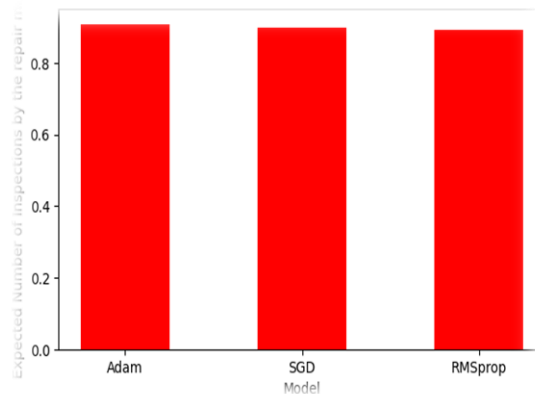


Figure 5: comparing between models according to Expected Fractional Number of repairman's Visit

V. Conclusion

In conclusion, the Behavior analysis presented system with failure and maintenance rate with using deep learning as the optimization tool has provided valuable insights into the dynamics and performance of the system across various operational scenarios. Through comprehensive experimentation and simulation, we have gained a deeper understanding of the system's response to different configurations, maintenance policies, and environmental factors. The results demonstrate the effectiveness of the deep learning guided approach in optimizing maintenance schedules and resource allocation to maximize system reliability and availability. By iteratively refining maintenance strategies, significant improvements in key performance metrics such as mean time to failure (MTTF), mean time to repair (MTTR), and overall system uptime have been achieved. This highlights the potential of deep learning to adaptively optimize complex systems in dynamic environments.

References

- [1] Hsieh, Y., C. and Chen, T.C. (2004). Reliability lower bounds for two-dimensional consecutive-k-out-of-n: F systems. *Computers & Operations Research* 31(8):1259–72
- [2] John, Y. M., Sanusi, A., Yusuf, I., and Modibbo, U., M. (2023). Reliability Analysis of Multi-Hardware–Software System with Failure Interaction. *Journal of Computational and Cognitive Engineering*, 2(1), 38-46.
- [3] Kumar, A. (2020). Reliability And Sensitivity Analysis of Linear Consecutive 2-out-of- 4: F System. *European Journal of Molecular & Clinical Medicine*, 7(07), 2020
- [4] Kumar, A. (2022). Sensitivity Analysis of Urea Fertilizer Plant. *Journal of Reliability Theory and Applications* Volume 17, RT&A No. 2 (68)
- [5] Kim, H.K., Lee, S. and Yun, K.S. (2011). Capacitive tactile sensor array for touch screen application. *Sensors and Actuators A: Physical* 165:2–7.
- [6] Khan, M. F., Modibbo, U. M., Ahmad, N., and Ali, I. (2022). Nonlinear optimization in bi-level selective e maintenance allocation problem. *Journal of King Saud University-Science*, 34(4), 101933.
- [7] Singla, S., Lal, A.K., and Bhatia, S.S (2011). Comparative study of the subsystems subjected to independent and simultaneous failure. *Eksploratacja I Niezawodnosc-Maintenance and Reliability*, 4, 63-71.
- [8] Raghav, Y. S., Varshney, R., Modibbo, U. M., Ahmadini, A. A. H., and Ali, I. (2022). Estimation and optimization for system availability under preventive maintenance. *IEEE Access*, 10, 94337-94353.

- [9] Singha, A. K., Pathak, N., Sharma, N., Gandhar, A., Urooj, S., Zubair, S., and Nagalaxmi, G. (2022). An Experimental Approach to Diagnose Covid-19 Using Optimized CNN. *Intelligent Automation & Soft Computing*, 34(2).
- [10] Singla, S., and Kumari, S. (2022). Behavior and profit analysis of a thresher plant under steady state B *International Journal of System Assurance Engineering and Management* 13, 166–171.
- [11] Singla, S., Lal, A. K., and Bhatia, S.S. (2021). Reliability analysis of poly tube industry using supplementary variable Technique *Applied Mathematics and Computation* 281, 3981–3992.
- [12] Salvia, A. A. and Lasher, W.C. (1990). 2-dimensional consecutive-k-out-of-n: F models. *IEEE Transactions on Reliability* 39(3):382–5
- [13] Singla, S., Modibbo, U. M., Mijinyawa, M., Malik, S., Verma, S., and Khurana, P. (2022). Mathematical Model for Analysing Availability of Threshing Combine Machine Under Reduced Capacity. *Yugoslav Journal of Operations Research*, 32(4), 425-437.
- [14] Singla, S., Rani, S., Modibbo, U. M. and Ali, I. (2023). Optimization of System Parameters of 2:3 Good Serial System using Deep Learning. *Reliability Theory and Applications*, 670-679.
- [15] Singla, S., Mangla, D., Panwar, P. and Taj, S. Z. (2024). Reliability Optimization of a Degraded System under Preventive Maintenance using Genetic Algorithm. *Journal of Mechanics of Continua and Mathematical Sciences*, 1-14.
- [16] Singla, S., Rani, S. (2023). Performance optimization of 3:4 good system. *IEEE second international conference* 979-8-3503-4383-0.
- [17] Ahmadini, A.A.H., Singla, S., Mangla, D., Modibbo, U.M., Rani, S. (2024). Reliability Assessment and Profit Optimization of Multi-unit Mixed Configured System using ABC Algorithm under Preventive Maintenance. *IEEE Access*.
- [18] Singla, S., Mangla, D., Dhawan, P., Ram, G. (2024). Reliability analysis of a two-unit repairable system using fuzzy linguistic approach. *Applications of fuzzy theory in applied sciences and computer applications*, 45-57.

OPTIMIZING INVENTORY CONTROL THROUGH A GRADIENT-BASED MULTILEVEL APPROACH IN THE FACE OF DEMAND AND LEAD TIME UNCERTAINTIES

MURAGESH MATH^{1,4},D.GOPINATH^{2,*},B. S.BIRADAR³

^{1,2} Chaitanya(Deemed To Be University),India

³University of Mysore,India

⁴ Bharati Vidyapeeth (Deemed To Be University) Medical College Sangli, India

¹*murageshmathapati5@gmail.com*,³*biradarbs1@gmail.com*

^{2,*}*Corresponding Author : drgopinathduggi@gmail.com*

Abstract

Systems of two-level assembly with unknown timing of leads are taken into consideration while arranging supplies. Probably, the final product's demand and its deadline are known. When all required parts are on hand, each level's assembly process gets underway. To address these problems, we have developed a model for the control of inventories for an uncapitated warehousing space in a manufacturing plant with unpredictable demand and lead times. The goal is to choose orders in a way that minimizes the overall system's cost. We present a multilevel optimization model including a rotating horizon that utilizes gradients to handle unknown lead time and demand, irrespective of the distributions at the core of them. Furthermore, a precise algorithm is created to solve the model. In a case study, we compare our approach with the current model. Our computational results indicate that while the new gradient-based multi-level optimization model nearly continuously yields the least expensive overall across all parameter settings. These models' performances are either systematically worse or extremely sensitive to cost parameters (holding cost, shortfall cost, etc.).

Keywords: multilevel optimization model; rolling horizon; uncertain demand; uncertain lead time.

1. INTRODUCTION

Supply chain management is a top priority for businesses in the modern global marketplace because of the fierce competition and elevated customer expectations. A special focus is on supply chain network architecture because it is recognized that efficient management of the supply chain a critical role in lowering costs and improving service levels. A skilled network design approach can significantly reduce expenses for a company by as much as 60% [1-3]. Historically, the planning and placement of facilities have been considered a strategic choice in supply chain network architecture. That being said, sub-optimality may result from the traditional method of making tactical inventory decisions after site selections. Inventory costs are heavily impacted by strategic location choices, highlighting the necessity of incorporating inventory concerns into strategic network design models [4-6]. As a result, there has been a significant push in recent years for the creation of inventory models that integrate tactical and strategic decisions. Several reasons, including changes in consumer needs and the arrival of raw materials, cause uncertainty to become a ubiquitous feature in supply chain networks. Three

main areas of uncertainty are identified: customers, manufacturing, and suppliers. Unexpected expenses can arise from supplier uncertainty, which introduces unpredictability in lead time, and customer uncertainty, which shows up as variances in order time or quantity. To improve overall operational efficiency and optimize supply chains, businesses must acknowledge and manage these uncertainties [7-8]. This research presents a novel Model of inventory control that considers demand variability and lead time uncertainty. Stochastic programming is accepted as a useful technique when the randomness's probabilistic description is provided; but, in practice, this information is not always available. The suggested model attempts to strike a balance between the curse of dimensionality and solution resilience when addressing multi-period decision-making situations with uncertainty [9–10]. The inventory control model is intended for usage in a manufacturing facility warehouse, where a single product is produced from an ordered part. Even though in reality several items are manufactured using different parts, in some situations it is appropriate to assume that there is only one product, particularly when production lines are independent and separate for different products. The goal of the study is to specify an order strategy that reduces system expenses [11–12]. The paper analyzes two ambiguous parameters: lead time and demand with unknown likelihood distributions and assumes that these quantities are independent random variables within given intervals. The assumption is consistent with empirical observations of dynamic lead time and demand dynamics, where trends in the past are not indicative of the future. The model also allows for shortages, which have a backlog that is entirely unfilled. To minimize the entire rate, which includes order, inventory holding, and shortfall expenses, the goal is to ascertain the timing and size of orders. Because demand is unpredictable at each stage and lead times are realized only after orders are placed, the intrinsic dimensionality curse affects multi-stage decision-making problems. [13–14]. This study presents three distinct contributions. Unlike earlier models, it first takes supply and demand uncertainties into account. Second, it presents a novel multilevel inventory control optimization model that approximates the multi-stage computational tractability decision-making problem. Thirdly, the work creates a precise algorithm for the multilevel optimization framework based on gradients, which makes it possible to explore the scenario space and worst-case situations with efficiency. The following is the arrangement of the paper's succeeding sections: Section 2 discuss about existing relevant works. Section 3 goes into a thorough discussion of the problem formulation and algorithm. Experimental results and sensitivity analysis are detailed in Section 4. Lastly, Section 5 concludes the paper, summarizing key findings.

2. LITERATURE SURVEY

This literature survey delves into various topics related to inventory control, including lost sales inventory systems, perishable inventory systems, and supply chain management. These papers can provide valuable insights into the challenges and solutions related to inventory control and supply chain management, which can help in developing an effective inventory control policy for the given research problem. Hansen et al. [15] presented a perishable product inventory control policy for business-to-consumer retail that takes lead time and demand volatility into account. To reduce expenses associated with stockouts and excess inventory, the study concentrated on handling perishable inventory concerns. The established replenishment approach provided efficient management of perishable inventory by balancing holding costs against lost sales expenses. The idea outperformed traditional approaches, as shown by mathematical models, offering shops a useful resolution. Dey et al. [2] investigated controlled lead time and adaptability in astute supply chain management, highlighting the advantages of shortening lead time for diverse elements. Instead of using projected formulae for total costs, the study presented an accurate overall cost calculation that took backorder relationships and on-hand inventory into account. The study's use of marginal value analysis showed that the total expense of the supply chain is convex in terms of lead time and volatility. An intelligent manufacturing procedure that addresses stochastic demand and variable production rates was presented by the researcher. It was validated by numerical examples and classical optimization, and the model validation was further

strengthened by sensitivity analysis and graphical representations. Zhang et al. [16] presented a learning algorithm that takes into account missed sales, positive lead times, and suppressed demand for a single-product inventory system assessment regularly. It tackled the problem of adaptive inventory ordering depending primarily on previous sales information. A random cycle-updating rule with essential components including withheld on-hand inventory and estimate of double-phase cycle gradients was introduced by the nonparametric simulated cycle-update policy. The study demonstrated efficacy in managing intricate system dynamics by establishing a square root convergence rate as a lower constraint for learning algorithms through regret analysis. The discovered methods reduced the cost differences between practical learning algorithms and clairvoyance benchmarks by enabling adaptive inventory decisions based on historical sales. Das et al. [17] addressed inventory control with partial backlog, price-dependent demand, and preservation technology applied to non-instantaneously decaying commodities. With the inclusion of a Weibull distribution with three parameters for deterioration and preservation, the model took into account the effects of preservation technology, price-dependent demand, and deterioration. The extremely nonlinear optimization issues were solved using quantum-behaved particle swarm optimization (QPSO) techniques. By comparing findings with several QPSO variations, numerical examples were used to validate the proposed model. Sensitivity analysis looked into how changing a parameter would affect the best course of action. Sarkar et al. [18] discussed a collaborative advertising strategy for supply chain management in ambiguous circumstances. Equations [34–37] restrictions were used to examine the model's goal. Equation (24) provided the overall cost of the supply chain under a cooperative advertising collaboration policy with ambiguous conditions. Equation (15) provided the supplier's total cost after modeling each of the separate charges related to the supplier under the cooperative advertising collaboration. While creating the model, the paper took into account a few suppositions. Transchel et al. [19] addressed considering lead time unpredictability and service level limitations, inventory management and supply planning are implemented for perishable goods. A dynamic inventory control policy was implemented, taking into account a specified service level and an unpredictable lead time for replenishments. With a first-in-first-out (FIFO) inventory system and a non-stationary demand process, the study concentrated on a business-to-business (B2B) setting. Through simulation-based optimization, the authors addressed the influence on service levels and waste rates while offering analytical insights into the ideal replacement quantity under lead-time uncertainty. Goli et al. [20] addressed the arc-routing problem of sustainable periodic garbage pickup. To solve the garbage collection problem, the research presented a hybrid multi-objective optimization strategy. The issue was formulated as an arc-routing problem in the paper, which made use of multi-objective optimization to take into account numerous objectives at once and the creation of a hybrid algorithm to discover the best possible answers. Our goal is to critically assess how these studies contribute to inventory control advancements, particularly focusing on their effectiveness in handling uncertainty and several other issues.

3. METHODOLOGY

3.1. Problem statement

We examine a manufacturing facility's vacant warehouse for a single item. Both the lead time and the demand are ambiguous. To reduce the costs associated with orders, inventory, and shortages, Choices are made across an arbitrary distinct time frame. We execute the assumption that the management has complete knowledge of the level of demand and inventories at the moment shortfall, and order arrival status before deciding which order to place during that time. We also assume that the shortfall is fully backlogged, and order volume and demand arrive at the start of the time frame for decisions. The present period is designated as period 1 for modeling reasons, and a finite planning horizon $n \{1, 2... T\}$ is enforced. The model has been resolved using updated data for every decision-making window, and alone the sequence choice for the present period is carried out when using the rolling horizon method of this model 105's

answer [23–24] The planning horizon of the $P(\tau)$ decision-making model spans the time from τ to $\tau + T - 1$. Following the resolution of the model $P(\tau)$ for decision-making, the policy of order was established, We separate the horizon of planning decision into two sections: The choice reached in the first era, τ , and in the second subsequent periods, $(\tau + 1, \dots, \tau + T - 1)$. The order guidelines for period τ are put into effect, and τ is raised by one are repeated after updating the model's initial settings for the upcoming planning horizon. As a result, the choice about periods $(\tau + 1 \dots \tau + T - 1)$ may need to be rescheduled for the following planning span.

The planned horizon measure T has a significant impact on how accurate the previously mentioned planning model is. Because $T \geq 3$ are multi-phase models of decision-making that are free from the well-acknowledged dimensionality curse. Infamously difficult to solve from the standpoint of computational tractability. However, from an application standpoint, models with this little planning horizon are essentially blind and could produce overly narrow-minded results.

3.2. Deterministic model

Consider an examination of a condensed form of the inventory control model, which takes lead time and demand into account to be known and constant for every period. As a result, a Model for deterministic single-stage optimization is all that remains of the multi-phase issue requiring decision-making. It is important to note that a set of binary parameters $\delta_{k,t} \forall k, t$, which indicates whether or not the order placed in period k arrives by period t , constitutes the random lead time. For instance, if an order placed during period 3 has a lead time of 4, then $\delta_{3,4} = \delta_{3,5} = \delta_{3,6} = 0$ and $\delta_{3,t} = 1, \forall t \in \{7, 8, \dots, T\}$. The model for deterministic inventory management is presented in (1a)–(1d). The model aims to reduce the overall cost along the horizon of planning. The inventory holding cost, shortfall cost, and both fixed and variable order costs are the four cost terms in (1a), in that order. After period t , the inventory level is determined by equation (1b). The quantity of deficiency at period t , the total number of ordered products that reach by time t , the initial inventory at period 0, and the entire quantity of the demand that is satisfied among periods 1 and t consist of the four terms listed on Constraint (1b)'s right side. By requiring the ordering of at least one item within that time frame, constraint (1c) guarantees the imposition of a fixed order fee. Constraint (1d) defines the decision variable supports. article amsmath

$$\min \quad \zeta = (c\mu) \sum_{t=1}^T q_t + f \sum_{t=1}^T v_t + h \sum_{t=1}^T I_t + p \sum_{t=1}^T g_t \quad (1a)$$

$$\text{s.t.} \quad I_t = I_0 + \sum_{k=1-K}^{t-1} \mu q_k \widehat{\delta}_{k,t} + g_t - \sum_{i=1}^t \widehat{d}_i \quad t \in \{1, 2, \dots, T\} \quad (1b)$$

$$q_t \leq Mv_t \quad t \in \{1, 2, \dots, T\} \quad (1c)$$

$$q_t, I_t, g_t \in \mathbb{Z}^+; \quad v_t \in \{0, 1\} \quad t \in \{1, 2, \dots, T\} \quad (1d)$$

3.3. Proposed method: Gradient computation in the approximated problem

The estimated problem for multilevel optimization is asymptotically convergent; hence we suggest using a projected gradient approach. The formula for $\nabla_{x_1} \widetilde{F}_1(x_1)$ is derived in this section, and the projected gradient method's local and global convergence is verified [18-24].

(i) Gradient of the objective function in the approximated problem

A formula for computing is given by the following theorem $\nabla_{x_1} \widetilde{F}_1(x_1)$.

Theorem 1: Formula for (gradients in n-level optimization problems). The expression for the

gradient $\nabla_{x_1} \tilde{F}_1(x_1)$ is as follows. article amsmath

$$\begin{aligned} \nabla_{x_1} \tilde{F}_1(x_1) &= \nabla_{x_1} f_1^s(x_1, x_2^{(T_2)}, \dots, x_n^{(T_n)}) \\ &\quad + \sum_{i=2}^n Z_i \nabla_{x_1} f_1(x_1, x_2^{(T_2)}, \dots, x_n^{(T_n)}), \\ Z_i &= \sum_{t=1}^{T_i} \left(\sum_{j=2}^{i-1} Z_j C_{ij}^{(t)} + B_i^{(t)} \right) \prod_{s=t+1}^{T_i} A_i^{(s)}, \\ A_i^{(t)} &= \nabla_{x_i} \Phi_i^{(t)}(x_1, x_2^{(T_2)}, \dots, x_{i-1}^{(T_{i-1})}, x_i^{(t-1)}), \\ B_i^{(t)} &= \nabla_{x_1} \Phi_i^{(t)}(x_1, x_2^{(T_2)}, \dots, x_{i-1}^{(T_{i-1})}, x_i^{(t-1)}), \\ C_{ij}^{(t)} &= \nabla_{x_j} \Phi_i^{(t)}(x_1, x_2^{(T_2)}, \dots, x_{i-1}^{(T_{i-1})}, x_i^{(t-1)}) \end{aligned}$$

For any $i = 2, \dots, n; t = 1, \dots, T_i; \text{ and } j = 2, \dots, i-1$, wherever we define $\prod_{s=t+1}^{T_i} A_i^{(s)} := A_i^{(t+1)} A_i^{(t+2)} \dots A_i^{(T_i)}$ for $t < T_i$ and $\prod_{s=T_i+1}^{T_i} A_i^{(s)} = I$.

We consider computing $\nabla_{x_1} \tilde{F}_1(x_1)$ using **Theorem 1**. Observe that computing is simple for us $Z_2 = \sum_{t=1}^{T_2} B_2^{(t)} \prod_{s=t+1}^{T_2} A_2^{(s)}$. For $i = 3, \dots, n$, when we have Z_2, \dots, Z_{i-1} , we can calculate Z_i . We present a computation technique that $\nabla_{x_1} \tilde{F}_1(x_1)$ by computing Z_2, \dots, Z_n in this sequence within Algorithm 1 article amsmath algorithm algpseudocode

Algorithm 1: Calculation of $\nabla_{x_1} \hat{F}_1(x_1)$.

Input: $x_1 =$ existing first-level variable value. $\{x_i^{(0)}\}_{i=2}^n =$ the lower-level iteration's initial values.
Output: The exact value of $\nabla_{x_1} \tilde{F}_1(x_1)$.

Algorithm 1: Calculation of $\nabla_{x_1} \hat{F}_1(x_1)$.

Input: $x_1 =$ existing first-level variable value. $\{x_i^{(0)}\}_{i=2}^n =$ the lower-level iteration's initial values.

Output: The exact value of $\nabla_{x_1} \tilde{F}_1(x_1)$. $g := (0, \dots, 0)^\top$.

for $i := 2, \dots, n$ do

$Z_i := O$.

for $t := 1, \dots, T_i$ do

$x_i^{(t)} := \Phi_i^{(t)}(x_1, x_2^{(T_2)}, \dots, x_{i-1}^{(T_{i-1})}, x_i^{(t-1)})$.

$\hat{B}_i^{(t)} := \sum_{l=2}^{i-1} Z_l C_{il}^{(t)} + B_i^{(t)}$.

$Z_i := Z_i A_i^{(t)} + \hat{B}_i^{(t)}$.

for $i = 2, \dots, n$ do

$g := g + Z_i \nabla_{x_1} f_1$.

$g := g + \nabla_{x_1} f_1$.

return g

For $i = 2, \dots, n$ and $t = 1, \dots, T_i, \Phi_i^{(t)}$, It is mentioned in Algorithm 1's fifth line is the revised formula that takes the gradient into account. $\nabla_{x_1} \tilde{F}_i(x_1, \dots, x_i)$ of the i th function of the level objective. $\nabla_{x_i} \tilde{F}_i(x_1, \dots, x_i)$ is calculated by utilizing Algorithm 1 on the $(n - i + 1)$ -a level optimization issue using objective functions $\tilde{F}_i, \dots, \tilde{F}_n$. Therefore, in the computation, call Algorithm 1 recursively of $\phi_i^{(t)}$, we can compute $\nabla_{x_1} \hat{F}_1(x_1)$.

3.4 Complexity of the gradient computation

By repeatedly running Algorithm 1, we examine the computational complexity of $\nabla_{x_1} \hat{F}_1(x_1)$. An asymptotic large symbol for O notation is $O(*)$ in the following theorem.

Theorem 2. Let c_i and s_i be the complexity of time and space, respectively, for computing $\nabla_{x_1} \tilde{F}_i(x_i)$. To calculate $\nabla_{x_1} \hat{F}_i(x_1, \dots, x_i)$, we recursively call the algorithm $\Phi_i^{(t)}$ based on $\nabla_{x_1} \tilde{F}_i(x_i)$. Moreover, we use the following assumptions:

- The intricacy of space and time in assessing $\Phi_i^{(t_i)}$ are $O(c_i)$ and $O(s_i)$.
- The time and space complexity of Algorithm 1 pales in contrast to the for loops in lines 2–9. of $\nabla_{x_1} f_1$ and $\nabla_{x_1} f_1$ for $i = 1, \dots, n$ are reduced in terms of order.
- Then, the whole intricacy of time c_1 and space intricacy s_1 for calculating can $\nabla_{x_1} \widehat{F}_i(x_1, \dots, x_i)$ be expressed as

$$c_1 = O\left(p^n n! c_n \prod_{i=1}^{n-1} (T_{i+1} d_i)\right),$$

$$s_1 = O(q^n s_n).$$

Correspondingly, for a fixed $p, q > 1$.

3.4. Global convergence of the projected gradient method

Here, we investigate using the projected gradient approach to solve the problem. In this method, the revised point is projected on S_1 in each iteration and computes the gradient vector using Algorithm [1]. We may determine when all lower-level updates are made using the steepest descent method, the Lipschitz continuity of the objective function's gradient. Therefore, by choosing a short enough step size, we may ensure both local and global projected gradient method convergence.

Theorem 3. Assume that $\Phi_i^{(t)}(x_1, \dots, x_{i-1}, x_i^{(t-1)}) = x_i^{(t-1)} - \alpha_i^{(t-1)} \nabla_{x_1} \widetilde{F}_i(x_1, \dots, x_{i-1}, x_i^{(t-1)})$ for all $i = 2, \dots, n$ and $t_i = 1, \dots, T_i$ where parameters for $\alpha_i^{(t-1)}$ and $x_i^{(0)}$ are provided for each i and t . Let us assume that $\nabla_{x_j} f_i$ is limited and Lipschitz continuous for all $i = 1, \dots, n$ and $j = 1, \dots, n$; furthermore, $\nabla_{x_1} \Phi_i^{(t)}$, $\nabla_{x_1} \Phi_i^{(t)}$, and $\nabla_{x_j} \Phi_i^{(t)}$ are bounded and Lipschitz continuous for all $i = 2, \dots, n$; $j = 2, \dots, i - 1$; $t = 1, \dots, T_i$. In such cases, $\nabla_{x_1} \widehat{F}_1$ is Lipschitz continuous with $\nabla_{(x_1)}$.

This theorem is proven; see Supplementary material A.5. Assume the same premises as Theorem 3 in the following corollary. Assume that the set S_1 is compact and convex. Assume that L is the Lipschitz constant for $\nabla_{x_1} \widehat{F}_1$. In the case of Problem (6), a sequence $\{x_1^{(t)}\}$ produced by the projected gradient technique with a compact, steady-step size (e.g., smaller than $2/L$) converges to a stationary point at a convergence rate via a convergent subsequence starting at any initial point of $O(1/\sqrt{t})$.

Proof. The gradient of the problem's L -Lipschitz is the objective function continuous, according to **Theorem 3**. The gradient mapping related to \widetilde{F}_1 , the indicator function of S_1 , and the fixed step magnitude $\alpha_1^{(t)}$ with fulfilling $0 < \alpha_1^{(t)} < 2/L$ for all t should be represented by $G : \text{int}(\text{dom}(\widetilde{F}_1)) \rightarrow \mathbb{R}^{d_1}$. Keep in mind that $\epsilon'G(x_1)\epsilon=0$ only in the event that x_1 is a stationary point. With \bar{x}_1 as a limit point of $\{x_1^{(t)}\}$, we obtain $\min_{s=0}^t \|G(x_1^{(s)})\| \leq O(1/\sqrt{t})$ and $\|G(\bar{x}_1)\| = 0$.

4. COMPUTATIONAL EXPERIMENTS

To evaluate and contrast the capabilities of novel gradient based multilevel optimization model with existing stochastic programming, and pessimistic decision-making models, we ran an experiment. The five patterns of demand utilized in the research are based on actual request information from the Census Bureau of the United States Department of Commerce and the Federal Reserve Bank of St. We carried a total of five sets of tests for each of the six models and five cases, with $h/p = \{0.1, 0.3, 0.5, 0.7, 0.9\}$ for $h = 5$ and $c = 1$. Since the ratio of h/p is important in inventory control models. Given the importance of the h/p ratio in models of inventory control,

we ran a total of five sets of experiments for each of the six models and five occurrences, with $h/p = \{0.1, 0.3, 0.5, 0.7, \text{ and } 0.9\}$ when $h = 5$ and $c = 1$. We took the demand data some of which are expressed in millions of dollars and created random lead-time estimates. We divided the values by the price of one unit to translate them to the number of units and assure uniformity. It's also believed that period 1 will not see any orders placed, but that period 1's initial inventory will be enough to meet demand for the first two periods. It is significant to highlight that the computer system using MATLAB (R2022a) software was able to generate the numerical results and tables following the specifications.

Simulation results

According to simulation data, for various h/p ratios, the gradient-based multilevel model has a lower total cost on average less optimistic than other models of decision-making, such as the stochastic programming model. We performed five instances of a sensitivity analysis using various cost parameter choices for each model, and the results are displayed in Figure 1. The combination of the cost factors h and p is shown in this figure. One instance is represented by each graph column, and the first row displays the pattern of demand for every case. The cost parameter configurations for various h/p ratios are shown in the following graph rows. The difference between each model and the ideal model is shown by the bars' vertical axis, which is computed by dividing each model's answer by the perfect model's solution minus one. Consequently, a smaller value implies the model's output is more similar to the perfect model. For instance, in the graph where $h/p = 0.5$, the multilevel optimization model's solution is $1.39e7$, whereas the perfect model's solution is $0.85e7$. Consequently, the multilevel model's value of the bar chart on this graph is $(1.39e7/0.85e7) - 1 = 64\%$. Overall parameter settings and instances, the average difference between our model and the ideal model is 71.4%, 84.7%, and 79.5%, respectively.

To demonstrate the extent to which the multilevel model's total cost is superior to or inferior to other models, we also provide the numerical experiment's results from a different angle. The multilevel model's relative performance is assessed using the ratio $R = 100$ in comparison to the stochastic programming and pessimistic techniques for making decisions. The expression "multi" denotes the average overall price of the multilevel model, whereas "mdl" represents the average overall expense of the pessimistic models or stochastic programming, across five instances. In Figure 2, we compare and display the outcomes of the overall cost-performance ratio. The multilevel model performs better than the comparison model if there is a good performance ratio; hence, a greater percentage indicates an improved comparative performance of the multilevel model. For example, the multilevel model's average overall cost is 10.6% lower than that of the stochastic programming and pessimistic models when $h/p = 0.3$. For example, for $h/p = 0.3$, the multilevel model's average total cost beats the probabilistic and stochastic programming models by 2.6% and 10.6%, respectively. When h/p is raised, the pessimistic model performs worse. By projecting it usually has a larger inventory level to minimize lead times and future needs; hence, raising h/p raises the model's overall cost.

Several performance metrics in use now have different service levels. We examined the single service level—the multilevel, pessimistic models, and stochastic programming fill rate. The proportion of client orders that are fulfilled right away from available stock is known as the fill rate. Reducing the h/p ratio generally improves it. Figure 3 shows the average percentage of fills of five instances for each model. The pessimistic model's fill rates are 97% in all possible combos of h and p . The fill rates of the stochastic programming model and the multilevel optimization model are nearly equal at 98% when the expense of a shortage is extremely significant, that is, $h/p = 0.1$. Nevertheless, the multilevel optimization model's fill rate decreases by only 2%, while the stochastic programming model's fill rate falls to 85% when the shortage cost is reduced.

While the multilevel optimization model responds to adjustments to shortages cost extra subtly and effectively than the stochastic model, figure 3 shows the outcomes suggest that the multilevel optimization model is not the stochastic model, sensitive to scarcity cost. We divided the overall cost into the costs associated with shortages and inventory holding to explain this finding. Table 1 summarizes the percentage shifts in shortage and inventory levels for each of the five cases when the h/p ratio rises from 0.1 to 0.9. Percentages that are positive or negative

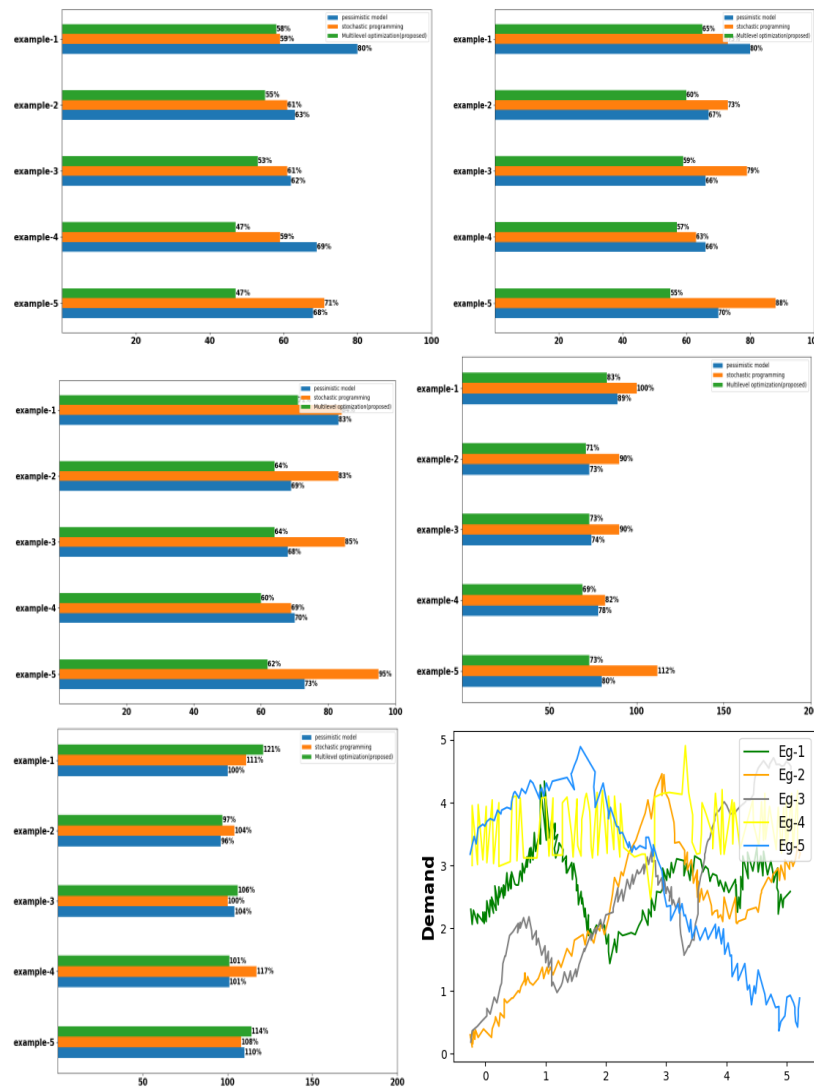


Figure 1: The comparison of multilevel optimization, stochastic programming, Pessimistic models under $h/p = \{0.9, 0.7, 0.5, 0.3, 0.1\}$

denote growth or decrease, accordingly. The multilevel model decreased the average inventory level each period (and the related inventory cost) by 22% to get the benefits of the decreased shortfall cost when the shortage cost decreased (the h/p ratio increased from 0.1 to 0.9) in Example 1 in Table 1. The fill rate decreased by 2% as a result, and the average shortage level rose by 160% (from 0.34 to 0.90 units each period). However, because of the sharp decline in shortfall cost per unit, the shortfall cost was drastically reduced by 71%. However, there was a more notable reaction from the stochastic model to the lower shortage cost. 51% less inventory meant a 14% lower fillrate, an increase in the scarcity level of 961% (from 0.35 to 3.70 units each period), and an 18% increase in the cost of shortages. These adjustments result in a 36% and 31% reduction in the multilevel optimization model's total amount of inventory held and shortfall cost, respectively, and the stochastic programming model. The average percentages for five cases are reported in the final two rows of Table 1.

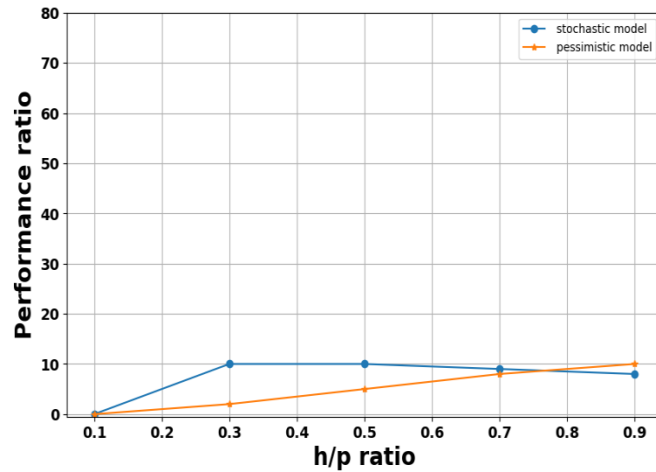


Figure 2: Retaining and scarcity costs' effects on the multi-level model's (multi) relative performance ratio concerning other models (model). * (mdl - multi)/mdl is the performance ratio, or R = 100.

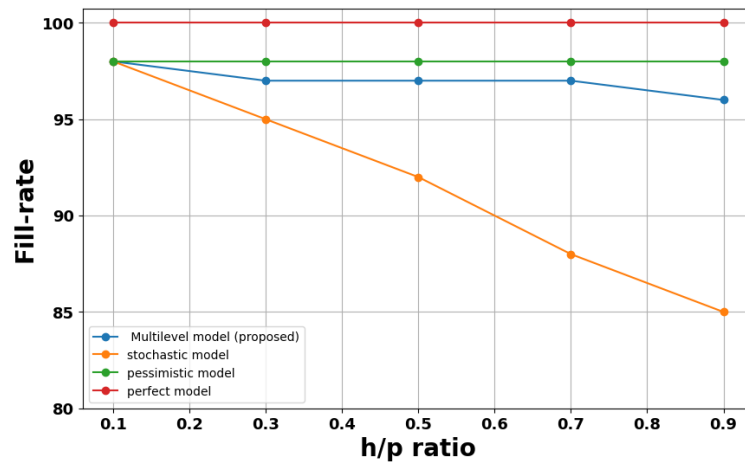


Figure 3: Effects of holding and shortage expenses on fill rate

Table 1: Inventory, shortage level, cost, and fill rate change when the h / p ratio rises from 0.1 to 0.9.

Example	Type	Inventory Cost	Shortage Cost	Inventory Level	Shortage Level	Fill-rate	Total Inventory and Shortage Cost
1	Multi-level	-22%	-71%	-22%	160%	-2%	-36%
	Stochastic	-51%	18%	-51%	961%	-14%	-31%
2	Multi-level	-14%	-69%	-14%	175%	-2%	-27%
	Stochastic	-50%	26%	-50%	1032%	-15%	-27%
3	Multi-level	-14%	-75%	-14%	126%	-2%	-32%
	Stochastic	-47%	28%	-47%	1056%	-14%	-25%
4	Multi-level	-26%	-67%	-26%	200%	-3%	-35%
	Stochastic	-41%	-19%	-41%	631%	-13%	-32%
5	Multi-level	-17%	-83%	-17%	51%	-1%	-38%
	Stochastic	-30%	-4%	-30%	762%	-9%	-22%
Average	Multi-level	-19%	-73%	-19%	143%	-2%	-34%
	Stochastic	-44%	10%	-44%	888%	-13%	-27%

In conclusion, the thorough comparison of the multilevel model with other methods shows its reliable performance, resilience, and sophisticated reaction to changing circumstances. These results further our knowledge of efficient inventory control techniques and offer practitioners and researchers insightful information for enhancing supply chain management in practical settings.

5. CONCLUSION

In this research, we suggest an innovative method for dealing with unpredictability in a manufacturing facility that places fresh orders in response to demand. There is a backlog of shortages, and the lead time and demand are unreliable metrics. Selecting orders in a way that minimizes the overall cost is the goal. Three new insights are added to the literature by this work. Initially, we consider two distinct types of uncertainty arising from lead time and demand. The majority of previously suggested models concentrated on one of these two, but they are still highly ambiguous due to the interactions between the two sources and the other. Second, as a trade-off between computational tractability and accurately representing the multiple-phase decision-making process under uncertainty, we suggest a multilevel Inventory control problem optimization model. Third, we design a precise algorithm for the multilevel optimization model that effectively searches in the worst instance, without counting all possible outcomes in the vast scenario space using Bender's decomposition foundation. The results imply that in reaction to the variety of cost factors, the multilevel optimization model operates more adaptable. The performance of the model of stochastic programming is primarily dependent on distributional knowledge and historical data; however, it attempts to strike a trade-off between holding costs and shortage. Concerning the cost parameters, the suggested multilevel optimization model automatically modifies its optimal ordering methods to produce the lowest (or nearly the lowest) total cost across all parameter configurations. Furthermore, the outcomes demonstrate that in terms of fill rate and total cost, the multilevel optimization model performs better than the stochastic programming model under various cost parameter values. The moderate model is nearly always in the middle of the results of the pessimistic and optimistic models, which are dependent on the cost factors. On the other hand, the multilevel optimization model finds the lowest (or nearly the lowest) total expense for every parameter configuration by automatically modifying its optimal ordering methods based on the cost parameters. There are various limitations to this study that point to potential areas for future investigation. In the suggested model, for instance, a single item created from a single component is assumed. If this supposition were to be relaxed, a more intricate model reflecting the ambiguity and interdependency of several components on the supply and demand sides would be necessary. Moreover, the decision-maker may choose to ship particular components or goods as a batch to reduce transportation costs by incorporating fixed and variable transportation costs into the model.

REFERENCES

- [1] Belanche, D., Casal, L. V., & Flavin, C. (2021). Frontline robots in tourism and hospitality: service enhancement or cost reduction?. *Electronic Markets*, 31(3), 477-492.
- [2] Dey, B. K., Bhuniya, S., & Sarkar, B. (2021). Involvement of controllable lead time and variable demand for a smart manufacturing system under a supply chain management. *Electronic Markets Expert Systems with Applications*, 184, 115464.
- [3] Babagolzadeh, M., Shrestha, A., Abbasi, B., Zhang, Y., Woodhead, A., & Zhang, A. (2020). Sustainable cold supply chain management under demand uncertainty and carbon tax regulation. *Transportation Research Part D: Transport and Environment*, 80, 102245.
- [4] Golpira, H. (2020). Optimal integration of the facility location problem into the multi-project multi-supplier multi-resource Construction Supply Chain network design under the vendor managed inventory strategy *Expert Systems with Applications*, 139, 112841.
- [5] Ye, Y., Jiao, W., & Yan, H. (2020). Managing relief inventories responding to natural disasters: Gaps between practice and literature. *Production and Operations Management*, 29(4), 807-832.

- [6] De Giovanni, P. (2021). Smart Supply Chains with vendor managed inventory, coordination, and environmental performance. *European Journal of Operational Research*, 292(2), 515-531.
- [7] Kim, T., Kim, Y. W., & Cho, H. (2020). Dynamic production scheduling model under due date uncertainty in precast concrete construction. *Journal of Cleaner Production*, 257, 120527.
- [8] Alzoubi, H. M., Elrehail, H., Hanaysha, J. R., Al-Gasaymeh, A., & Al-Adaileh, R. (2022). The role of supply chain integration and agile practices in improving lead time during the COVID-19 crisis. *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*, 13(1), 1-11.
- [9] Li, X., Uysal, A. S., & Mulvey, J. M. (2022). Multi-period portfolio optimization using model predictive control with mean-variance and risk parity frameworks. *European Journal of Operational Research*, 299(3), 1158-1176.
- [10] Lan, G., Lee, S., & Zhou, Y. (2020). Communication-efficient algorithms for decentralized and stochastic optimization. *Mathematical Programming*, 180(1-2), 237-284.
- [11] Lyu, Z., Lin, P., Guo, D., & Huang, G. Q. (2020). Towards zero-warehousing smart manufacturing from zero-inventory just-in-time production. *Robotics and Computer-Integrated Manufacturing*, 64, 101932.
- [12] Liu, A., Zhu, Q., Xu, L., Lu, Q., & Fan, Y. (2021). Sustainable supply chain management for perishable products in emerging markets: An integrated location-inventory-routing model. *Transportation Research Part E: Logistics and Transportation Review*, 150, 102319.
- [13] Ponte, B., Framinan, J. M., Cannella, S., & Dominguez, R. (2020). Quantifying the Bullwhip Effect in closed-loop supply chains: The interplay of information transparencies, return rates, and lead times. *International Journal of Production Economics*, 230, 107798.
- [14] Ghadge, A., Er, M., Ivanov, D., & Chaudhuri, A. (2022). Visualisation of ripple effect in supply chains under long-term, simultaneous disruptions: a system dynamics approach. *International Journal of Production Research*, 60(20), 6173-6186.
- [15] Hansen, O., Transchel, S., & Friedrich, H. (2023). Replenishment strategies for lost sales inventory systems of perishables under demand and lead time uncertainty. *European Journal of Operational Research*, 308(2), 661-675.
- [16] Zhang, H., Chao, X., & Shi, C. (2020). Closing the gap: A learning algorithm for lost-sales inventory systems with lead times. *Management Science*, 66(5), 1962-1980.
- [17] Das, S. C., Zidan, A. M., Manna, A. K., Shaikh, A. A., & Bhunia, A. K. (2020). An application of preservation technology in inventory control system with price dependent demand and partial backlogging. *Alexandria Engineering Journal*, 59(3), 1359-1369.
- [18] Sarkar, B., Omair, M., & Kim, N. (2020). A cooperative advertising collaboration policy in supply chain management under uncertain conditions. *Applied Soft Computing*, 88, 105948.
- [19] Transchel, S., & Hansen, O. (2019). Supply planning and inventory control of perishable products under lead-time uncertainty and service level constraints. *Procedia Manufacturing*, 39, 1666-1672.
- [20] Goli, A., Tirkolaee, E. B., & Weber, G. W. (2020). A perishable product sustainable supply chain network design problem with lead time and customer satisfaction using a hybrid whale-genetic algorithm. *Logistics operations and management for recycling and reuse*, 99-124.
- [21] Zaneti, L. A., Arias, N. B., de Almeida, M. C., & Rider, M. J. (2022). Sustainable charging schedule of electric buses in a University Campus: A rolling horizon approach. *Renewable and Sustainable Energy Reviews*, 161, 112276.
- [22] Cuisinier, P., Lemaire, P., Penz, B., Ruby, A., & Bourasseau, C. (2022). Sustainable charging schedule of electric buses in a New rolling horizon optimization approaches to balance short-term and long-term decisions: An application to energy planning. *Energy*, 245, 122773.
- [23] Wunnava, A., Naik, M. K., Panda, R., Jena, B., & Abraham, A. (2020). An adaptive Harris hawks optimization technique for two dimensional grey gradient based multilevel image thresholding. *Applied Soft Computing*, 95, 106526.
- [24] Ahmadianfar, I., Bozorg-Haddad, O., & Chu, X. (2020). Gradient-based optimizer: A new metaheuristic optimization algorithm. *Information Sciences*, 540, 131-159.

THE EXPONENTIATED SKEW LAPLACE DISTRIBUTION: PROPERTIES AND APPLICATIONS

Timothy Kayode Samson¹

Statistics Programme, College of Agriculture, Engineering and Science, Bowen University, Iwo,
Nigeria

kayode.samson@bowen.edu.ng¹



Christian Elendu Onwukwe², Ekaette Inyang Enang³

Department of Statistics, Faculty of Physical Sciences, University of Calabar, Calabar, Nigeria

christianonwukwe@unical.edu.ng², ekaettenang@unical.edu.ng³

Abstract

In this paper, a 4-parameter Exponentiated Skew Laplace distribution is defined and studied. Various statistical properties including its moment generating function, characteristics function, hazard function, and reliability function of the proposed ESLD were derived. The estimation of its parameters was carried out using the maximum likelihood method of estimation. The performance of the proposed ESLD compared with other similar distributions was demonstrated empirically with daily returns of S & P 500 between 2/02/24 and 28/03/2024 and daily returns of Bitcoin between 2/02/24 and 1/04/24 as obtained from Yahoo Finance. The fitness performance of the proposed distribution was evaluated based on log-likelihood, AIC, and BIC. Results obtained show that the proposed ESLD reported the highest log likelihood as well as the lowest AIC and BIC in the two data sets. This study therefore underscores the superiority of the proposed distribution over the some of the similar existing distributions.

Keywords: Skew Laplace, Exponentiated, Distribution, Reliability function, Maximum Likelihood

I. Introduction

The Laplace distribution also commonly known as the double exponential distribution is named after Pierre Simon Laplace. The growing popularity of the Laplace-based models as described by Lakshmi and Sebastian [1] is due to the properties of the sharp peak at the mode, heavier than normal distribution making it an asymmetric distribution Laplace distribution as described by Nadarajah and Kotz [2] is a tractable lifetime model with applications in various areas including telecommunication, biological sciences, engineering, and life testing among other areas of human endeavours. Despite the significance of this distribution in probabilistic modelling, the Laplace distribution lacks a skewness parameter and hence not be able to account for skewness in data. To overcome this challenge, the Laplace distribution as noted by Kotz *et al.*, [3] and Safavinejad *et al.*, [4] was extended by adding a skewness parameter. This extension therefore gave rise to the skew Laplace distribution.

The use Skew-Laplace distribution in Economics, Finance and Engineering has been emphasized by Puig and Stephens [5] while Julia and Vives-Rego [6] used Skew Laplace distribution to analyze bacterial sizes in axenic cultures. The skew Laplace distribution has been applied to different areas of research. This is due to its flexibility to model real data that exhibit skewness. Skew Laplace distribution has also been used in probabilistic modelling of financial data. For instance, Jing *et al.* [7] applied asymmetric Laplace to currency exchange rates in Australia, Canada, European and United Kingdom. Similarly, Shi *et al.* [8] used the asymmetric Laplace distribution in portfolio selection. The use of Skew Laplace distribution can also be extended to cryptocurrencies and stock data because of its ability to capture skewed data as cryptocurrencies and stocks have some stylized properties which include departure from normality, price jumps, and high volatility patterns.

In this study, a new form of skew Laplace distribution of Safavinejad *et al.* [4] is proposed, and some of the statistical properties are proposed with application to real data. The fitness performance of the proposed ESLD compared to some existing distributions will also be investigated in this study. This study adopts the method of exponentiation introduced by Gupta *et al.* [9] in generating the proposed Exponentiated Skew Laplace Distribution (ESLD). In the exponentiation method, only one shape parameters is introduced into the parent distribution. This method has been used by Agboola *et al.*, [10], Oguntunde *et al.*, [11], Nadarajah and Bakar [12], Datta and Datta [13], Andrade *et al.*, [14], Adubisi *et al.*, [15] among others to generate more flexible distribution.

II. Methods

Exponentiated Skew Laplace Distribution (ESLD)

The probability density function (pdf) and the Cumulative Density Function (CDF) of the Skewed Laplace Distribution (SLD) are defined in (1) and (2) as follows:

$$f(x; \theta, \varepsilon, \delta) = \begin{cases} \frac{1}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right), & x < \delta \\ \frac{1}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right), & x \geq \delta \end{cases} \quad (1)$$

Safavinejad *et al.* [4]

$$F(x; \theta, \varepsilon, \delta) = \begin{cases} \frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right), & x < \delta \\ 1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right), & x \geq \delta \end{cases} \quad (2)$$

$-\infty < x < \infty, \lambda > 0, \theta, \varepsilon, > 0, -\infty < \delta < \infty$

Safavinejad *et al.* [4]

where, θ is a scale parameter, δ location parameter and skewness parameter.

The probability density function of an Exponentiated generation distribution is given as:

$$g(x) = \lambda F(x)^{\lambda-1} f(x), \lambda > 0 \quad (3)$$

The Cumulative Density Function is given as:

$$G(x) = [F(x)]^\lambda \quad (4)$$

Where, $f(x)$ is the density function of the Skewed Laplace Distribution (SLD) and $F(x)$ is its corresponding cumulative density function.

Substituting for $f(x)$ in (3) and $F(x)$ in (4) give the pdf and the CDF of the proposed Exponentiated Skew Laplace Distribution.

$$g(x) = \begin{cases} \frac{\lambda}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}, & x < \delta \\ \frac{\lambda}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}, & x \geq \delta \end{cases} \quad (5)$$

$\theta, \varepsilon > 0, \lambda > 0, -\infty < x < \infty,$

$$G(x) = \begin{cases} \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}, & x < \delta \\ \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}, & x \geq \delta \end{cases} \quad (6)$$

Theorem 1

A random variable X is said to follow an Exponentiated Skew Laplace Distribution (ESLD), if its probability density function can be expressed as:

$$g(x) = \begin{cases} \frac{\lambda}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}, & x < \delta \\ \frac{\lambda}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}, & x \geq \delta \end{cases} \quad (7)$$

$\theta, \varepsilon > 0, \lambda > 0, -\infty < x < \infty,$

Proof

The purpose of this proof is to show that the proposed ESLD is a probability density function. Since, it a continuous distribution function, then

$$\int_{-\infty}^{\infty} g(x) dx = 1 \quad (8)$$

$$\int_{-\infty}^{\delta} \frac{\lambda}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1} dx + \int_{\delta}^{\infty} \left(\frac{\lambda}{\theta + \varepsilon}\right) \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1} dx \quad (9)$$

Splitting the expression in equation (9) into two parts as follows:

$$\text{Let } g_1 = \int_{-\infty}^{\delta} \frac{\lambda}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1} dx \quad (10)$$

$$g_2 = \frac{\lambda}{\theta + \varepsilon} \int_{\delta}^{\infty} \exp\left\{-\frac{x - \delta}{\varepsilon}\right\} \left\{1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right\}^{\lambda - 1} dx \quad (11)$$

Solving for g_1 .

$$\text{Let } z = \frac{\theta}{\theta + \varepsilon} e^{\frac{x - \delta}{\theta}} \Rightarrow \frac{dz}{dx} = \frac{1}{\theta + \varepsilon} e^{\frac{x - \delta}{\theta}} \quad (12)$$

$$\text{when, } x = -\infty, z = 0 \text{ and when, } x = \delta, z = \frac{\theta}{\theta + \varepsilon} \quad (13)$$

Then

$$= g_1 = \left(\frac{\theta}{\theta + \varepsilon}\right)^\lambda - 0 = \left(\frac{\theta}{\theta + \varepsilon}\right)^\lambda \quad (14)$$

Similarly,

$$g_2 = \frac{\lambda}{\theta + \varepsilon} \int_{\delta}^{\infty} \exp\left\{-\frac{x - \delta}{\varepsilon}\right\} \left\{1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\theta}\right)\right\}^{\lambda - 1} dx = 1 - \left(\frac{\theta}{\theta + \varepsilon}\right)^\lambda \quad (15)$$

Therefore,

$$\int_{-\infty}^{\infty} g(x) dx = \left(\frac{\theta}{\theta + \varepsilon}\right)^\lambda + 1 - \left(\frac{\theta}{\theta + \varepsilon}\right)^\lambda = 1. \quad (16)$$

This implies that the proposed ESLD is a probability density function.

When $\lambda = 1$, the ESLD in (7) reduces to Skew Laplace distribution

When $\lambda = \theta = 1$, the ESLD in (5) reduces to Laplace distribution.

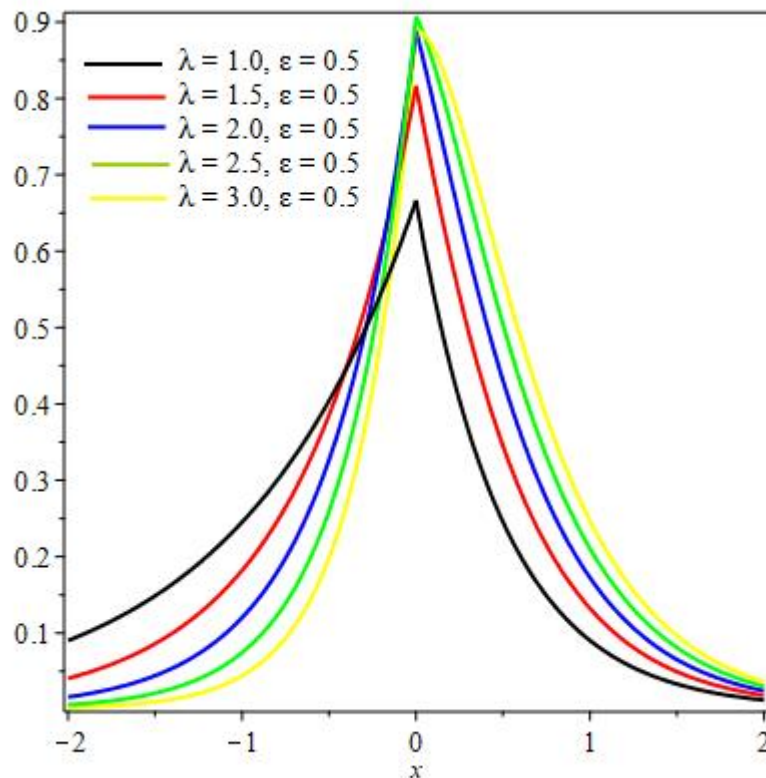


Figure 1: Plot of pdf of the proposed Exponentiated Skewed Laplace distribution (ESLD).

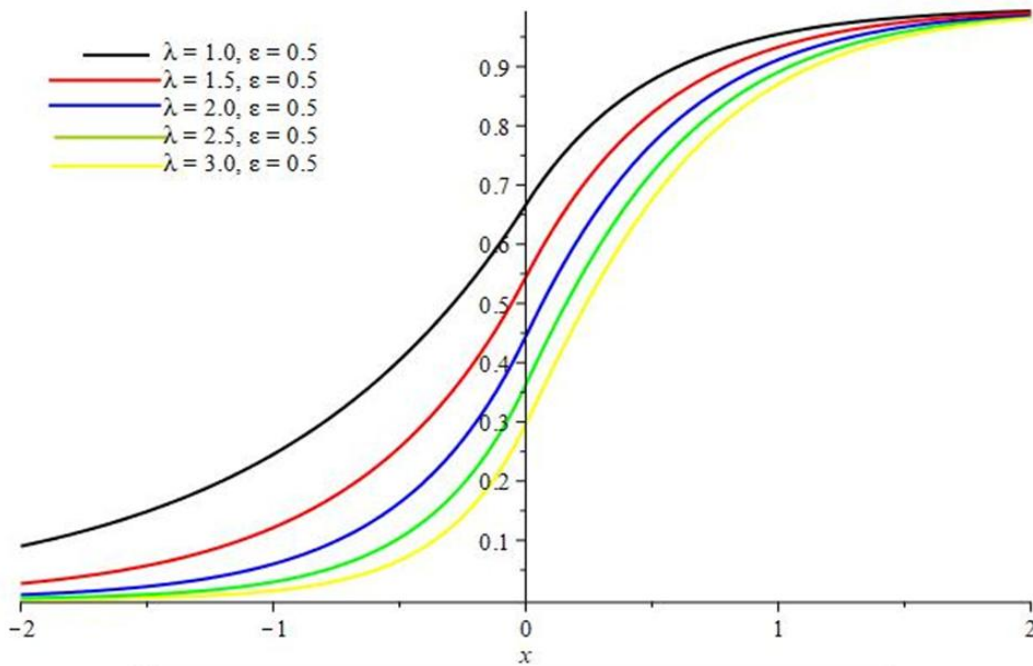


Figure 2: Plot of the CDF of the proposed distribution ESLD.

Asymptotic Behaviour of the Proposed ESLD

Theorem 2

If the random variable $X \sim \text{ESLD}(\theta, \delta, \varepsilon, \lambda)$, $\lim_{x \rightarrow -\infty} g(x) = 0$ and $\lim_{x \rightarrow \infty} g(x) = 0$.

Proof

$$g(x) = \begin{cases} \frac{\lambda}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}, & x < \delta \\ \frac{\lambda}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}, & x \geq \delta \end{cases}$$

$$\text{For, } \lim_{x \rightarrow -\infty} g(x) = \left(\frac{\lambda}{\theta + \varepsilon}\right) \exp\left(\frac{-\infty - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{-\infty - \delta}{\theta}\right)\right]^{\lambda - 1} = 0 \quad (17)$$

Similarly,

$$\lim_{x \rightarrow \infty} g(x) = \left(\frac{\lambda}{\theta + \varepsilon}\right) \exp\left\{-\frac{\infty - \delta}{\varepsilon}\right\} \left\{1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{\infty - \delta}{\varepsilon}\right)\right\}^{\lambda - 1} = 0 \quad (18)$$

Hence, $\lim_{x \rightarrow -\infty} g(x) = 0$ and $\lim_{x \rightarrow \infty} g(x) = 0$

Theorem 4

If, $X \sim \text{ESLD}(\theta, \delta, \varepsilon, \lambda)$, $\lim_{x \rightarrow -\infty} G(x) = 0$ and $\lim_{x \rightarrow \infty} G(x) = 1$.

Proof

$$\lim_{x \rightarrow -\infty} G(x) = \lim_{x \rightarrow -\infty} \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda} \quad (19)$$

$$= \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{-\infty - \delta}{\theta}\right) \right]^\lambda = 0 \tag{20}$$

Similarly, $\lim_{x \rightarrow \infty} G(x) = \lim_{x \rightarrow \infty} \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \right]^\lambda$

$$\left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{\infty - \delta}{\varepsilon}\right) \right]^\lambda = 1 - 0 = 1 \tag{21}$$

Hence, $\lim_{x \rightarrow -\infty} G(x) = 0$ and $\lim_{x \rightarrow \infty} G(x) = 1$

The proof of the Theorems 3 and 4 indicate that the proposed distribution satisfied the property of the limiting property a probability density function and cumulative density function respectively.

Statistical Properties of the proposed ESLD

Moment Generating Function of the Proposed ESLD

Suppose that X is a random variable that follows ESLD with parameters $(\theta, \delta, \varepsilon, \lambda)$, the moment generation function is given as:

$$M_x(t) = \frac{\lambda}{\theta + \varepsilon} \exp(t\delta) \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^j \left[\frac{\theta}{t\theta + 1 + j} - \frac{\varepsilon}{-t\varepsilon + 1 + j} \right] \tag{12}$$

Proof

By definition of the moment generating function:

$$M_x(t) = \int_{-\infty}^{\infty} e^{tx} g(x) dx \tag{22}$$

$$\frac{\lambda}{\theta + \varepsilon} \left[\int_{-\infty}^{\delta} \exp(tx) \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \right]^{\lambda-1} dx + \int_{\delta}^{\infty} \exp(tx) \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \right]^{\lambda-1} dx \right] \tag{23}$$

Let, $P_1 = \int_{-\infty}^{\delta} \frac{\lambda}{\theta + \varepsilon} \exp(tx) \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \right]^{\lambda-1} dx$ and

$$P_2 = \int_{\delta}^{\infty} \frac{\lambda}{\theta + \varepsilon} \exp(tx) \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \right]^{\lambda-1} dx \tag{24}$$

Let $w = x - \delta$, $\Rightarrow x = \delta + w$, if $x = -\infty$, $w = -\infty - \delta = -\infty$
 $x = \delta$, $w = \delta - \delta = 0$

$$P_1 = \int_{-\infty}^0 \exp(tw + t\delta) \exp\left(\frac{w}{\theta}\right) \left\{ \frac{\theta}{\theta + \varepsilon} \exp\left(\frac{w}{\theta}\right) \right\}^{\lambda-1} dw \tag{25}$$

$$P_2 = \int_{-\infty}^0 \exp(tw + t\delta) \exp\left(\frac{w}{\theta}\right) \left\{ 1 - \left(1 - \frac{\theta}{\theta + \varepsilon}\right) \exp\left(\frac{w}{\theta}\right) \right\}^{\lambda-1} dw \tag{26}$$

$$P_1 = \int_{-\infty}^0 \exp(tw + t\delta) \exp\left(\frac{w}{\theta}\right) \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \exp\left[\frac{jw}{\theta}\right] dw \tag{27}$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \int_{-\infty}^0 \exp(tw + t\delta) \exp\left(\frac{w}{\theta}\right) \exp\left[\frac{jw}{\theta}\right] dw \quad (28)$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \int_{-\infty}^0 \exp\left(tw + t\delta + \frac{w}{\theta} + \frac{jw}{\theta}\right) dw \quad (29)$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \exp(t\delta) \int_{-\infty}^0 \exp\left(tw + \frac{w}{\theta} + \frac{jw}{\theta}\right) dw \quad (30)$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \exp(t\delta) \int_{-\infty}^0 \exp\left[w\left(t + \frac{1}{\theta} + \frac{j}{\theta}\right)\right] dw \quad (31)$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \exp(t\delta) \frac{1}{t + \frac{1}{\theta} + \frac{j}{\theta}} \exp\left[w\left(t + \frac{1}{\theta} + \frac{j}{\theta}\right)\right]_{-\infty}^0 \quad (32)$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \frac{\exp(t\delta)}{t\theta + 1 + j} [1 - 0] \quad (33)$$

$$P_1 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta + \varepsilon}\right)^j \frac{\exp(t\delta)}{t\theta + 1 + j} \quad (34)$$

Similarly,

$$P_2 = \int_{-\infty}^{\delta} \frac{\lambda}{\theta + \varepsilon} \exp(tx) \exp\left(-\frac{x-\delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x-\delta}{\varepsilon}\right)\right]^{\lambda-1} \quad (35)$$

When $x = \delta$, since $w = x - \delta$ $w = 0$, When $x = \infty$, $u = \infty$

$$P_2 = \int_0^{\infty} \exp(tw + t\delta) \exp\left(-\frac{w}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{w}{\varepsilon}\right)\right]^{\lambda-1} dw \quad (36)$$

$$P_2 = \int_0^{\infty} \exp(tw + t\delta) \exp\left(-\frac{w}{\varepsilon}\right) \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^j \exp\left(-\frac{jw}{\varepsilon}\right) dw \quad (37)$$

$$P_2 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^j \int_0^{\infty} \exp(tw + t\delta) \exp\left(-\frac{w}{\varepsilon} - \frac{jw}{\varepsilon}\right) dw \quad (38)$$

$$P_2 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^j \exp(t\delta) \int_0^{\infty} \exp\left(tw - \frac{w}{\varepsilon} - \frac{jw}{\varepsilon}\right) dw \quad (39)$$

$$P_2 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^j \exp(t\delta) \int_0^{\infty} \exp\left[w\left(t - \frac{1}{\varepsilon} - \frac{j}{\varepsilon}\right)\right] dw \quad (40)$$

$$P_2 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^j \exp(t\delta) \int_0^{\infty} \exp\left[-w\left(-t + \frac{1}{\varepsilon} + \frac{j}{\varepsilon}\right)\right] dw \quad (41)$$

$$P_2 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \exp(t\delta) \frac{1}{-t + \frac{1}{\varepsilon} + \frac{j}{\varepsilon}} \exp \left[-w \left(-t + \frac{1}{\varepsilon} + \frac{j}{\varepsilon} \right) \right]_0^{\infty} \quad (42)$$

$$P_2 = \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \frac{\exp(t\delta)\varepsilon}{-t\varepsilon+1+j} [0-1] \quad (43)$$

$$P_2 = - \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \frac{\varepsilon \exp(t\delta)}{-t\varepsilon+1+j} \quad (44)$$

$$M_x(t) = \frac{\lambda}{\theta+\varepsilon} (P_1 + P_2) \quad (45)$$

$$M_x(t) = \frac{\lambda}{\theta+\varepsilon} \left[\sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(1 - \frac{\theta}{\theta+\varepsilon} \right)^j \frac{\theta \exp(t\delta)}{t\theta+1+j} - \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \frac{\varepsilon \exp(t\delta)}{-t\varepsilon+1+j} \right] \quad (46)$$

Note that, $1 - \frac{\theta}{\theta+\varepsilon} = \frac{\theta+\varepsilon-\theta}{\theta+\varepsilon} = \frac{\varepsilon}{\theta+\varepsilon}$ (47)

$$M_x(t) = \frac{\lambda}{\theta+\varepsilon} \exp(t\delta) \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \left[\frac{\theta}{t\theta+1+j} - \frac{\varepsilon}{-t\varepsilon+1+j} \right] \quad (48)$$

Characteristic Function of the Proposed ESLD

Suppose that X is a random variable that follows ESLD with parameters $(\theta, \delta, \varepsilon, a)$, the characteristic function, $\varphi_x(t)$ is given by:

$$\varphi_x(t) = \frac{\lambda}{\theta+\varepsilon} \exp(it\delta) \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \left[\frac{\theta}{it\theta+1+j} - \frac{\varepsilon}{-it\varepsilon+1+j} \right] \quad (49)$$

Proof

By definition,

$$\varphi_x(t) = E[\exp(itx)] = \int_{-\infty}^{\infty} \exp(itx) g(x) dx \quad (50)$$

Since the mgf of the ESLD has been obtained and presented in (40), then, it is easy to obtain its characteristic function and it is given by:

$$\varphi_x(t) = \frac{\lambda}{\theta+\varepsilon} \exp(it\delta) \sum_{j=0}^{\infty} \binom{\lambda-1}{j} (-1)^j \left(\frac{\varepsilon}{\theta+\varepsilon} \right)^j \left[\frac{\theta}{it\theta+1+j} - \frac{\varepsilon}{-it\varepsilon+1+j} \right] \quad (51)$$

Hazard Function of ESLD

The Hazard function by definition is given as:

$$H(x) = \frac{g(x)}{1-G(x)} \quad (52)$$

Where, $H(x)$ is the hazard function while $g(x)$ and $G(x)$ are the probability density function and cumulative density function respectively.

Hence, the Hazard function of the proposed ESLD can be expressed as:

$$H(x) = \begin{cases} \frac{\frac{\lambda}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right) \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}}{\left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}}, & x < \delta \\ \frac{\frac{\lambda}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right) \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}}{\left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}}, & x \geq \delta \end{cases} \quad (53)$$

Reliability Function

The reliability function of the ESLD is given by:

$$R(x) = 1 - G(x) \quad (54)$$

Hence,

$$R(x) = 1 - \begin{cases} \left[\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x - \delta}{\theta}\right)\right]^{\lambda - 1}, & x < \delta \\ \left[1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x - \delta}{\varepsilon}\right)\right]^{\lambda - 1}, & x \geq \delta \end{cases} \quad (55)$$

Parameter Estimation for the ESLD.

The estimation of the parameter of the ESLD defined in (5) will be estimated using the Method of Maximum Likelihood derived below:

Let X_1, X_2, \dots, X_n be a random sample of size n from Exponentiated Skew Laplace Distribution (ESLD). Then the likelihood function is given by:

For, $x < \delta$

$$L(x) = \prod_{i=1}^n \left(\frac{\lambda}{\theta + \varepsilon}\right) \exp\left(\frac{x_i - \delta}{\theta}\right) \left\{\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x_i - \delta}{\theta}\right)\right\}^{\lambda - 1} \quad (56)$$

$$L(x) = \left(\frac{\lambda}{\theta + \varepsilon}\right)^n \prod_{i=1}^n \exp\left(\frac{x_i - \delta}{\theta}\right) \left\{\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x_i - \delta}{\theta}\right)\right\}^{\lambda - 1} \quad (57)$$

Taking the natural log of both sides.

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \log \left[\prod_{i=1}^n \exp\left(\frac{x_i - \delta}{\theta}\right) \left\{\frac{\theta}{\theta + \varepsilon} \exp\left(\frac{x_i - \delta}{\theta}\right)\right\} \right] \quad (58)$$

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \left[\log \prod_{i=1}^n \frac{\theta}{\theta + \varepsilon} \exp\left(\frac{2x_i - 2\delta}{\theta}\right) \right] \quad (59)$$

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \log\left[\left(\frac{\theta}{\theta + \varepsilon}\right)^n \exp\left(\frac{2 \sum_{i=1}^n (x_i - \delta)}{\theta}\right)\right] \quad (60)$$

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \left\{ n \log\left(\frac{\theta}{\theta + \varepsilon}\right) + \frac{2 \sum_{i=1}^n (x_i - \delta)}{\theta} \right\} \quad (61)$$

$x \geq \delta$

$$L(x) = \prod_{i=1}^n \left(\frac{\lambda}{\theta + \varepsilon}\right) \exp\left(-\frac{x_i - \delta}{\varepsilon}\right) \left\{ 1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x_i - \delta}{\varepsilon}\right) \right\}^{\lambda - 1} \quad (62)$$

$$L(x) = \left(\frac{\lambda}{\theta + \varepsilon}\right)^n \prod_{i=1}^n \exp\left(-\frac{x_i - \delta}{\varepsilon}\right) \left\{ 1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x_i - \delta}{\varepsilon}\right) \right\}^{\lambda - 1} \quad (63)$$

Taking the natural log of both sides.

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \log\left[\prod_{i=1}^n \exp\left(-\frac{x_i - \delta}{\varepsilon}\right) \left\{ 1 - \frac{\varepsilon}{\theta + \varepsilon} \exp\left(-\frac{x_i - \delta}{\varepsilon}\right) \right\}\right] \quad (64)$$

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \log\left[e^{-\frac{1}{\varepsilon} \sum_{i=1}^n (x_i - \delta)} - \left(\frac{\varepsilon}{\theta + \varepsilon}\right)^n e^{-\frac{2}{\varepsilon} \sum_{i=1}^n (x_i - \delta)}\right] \quad (65)$$

$$\log L(x) = n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \left[-\frac{1}{\varepsilon} \sum_{i=1}^n (x_i - \delta) - n \log\left(\frac{\varepsilon}{\theta + \varepsilon}\right) - \frac{2}{\varepsilon} \sum_{i=1}^n (x_i - \delta) \right] \quad (66)$$

Hence, the Log Likelihood of the proposed Exponentiated Skew Laplace Distribution (ESLD) is given as:

$$\log L(x) = \begin{cases} n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \left[n \log\left(\frac{\theta}{\theta + \varepsilon}\right) + \frac{2}{\theta} \sum_{i=1}^n (x_i - \delta) \right], & x < \delta \\ n \log\left(\frac{\lambda}{\theta + \varepsilon}\right) + (\lambda - 1) \left[-\frac{1}{\varepsilon} \sum_{i=1}^n (x_i - \delta) - n \log\left(\frac{\varepsilon}{\theta + \varepsilon}\right) - \frac{2}{\varepsilon} \sum_{i=1}^n (x_i - \delta) \right], & x \geq \delta \end{cases} \quad (67)$$

Differentiating $\log L(x)$ in (57) with respect to each of the parameter in the model and setting them to zero give the estimate of each of the parameter. This was done in R software using the appropriate optimization function.

III. Results

Applications to Real Data

The empirical application of the proposed ESLD was carried out using two sets of real life data. The first data is the daily returns of S& P 500 between 2/02/2024 and 28/03/2024 as obtained from Yahoo finance. The second data set is on the daily returns of Bitcoin between 2/02/2024 and 01/04/2024 which was also obtained from the Yahoo finance website (www.yahoofinance.com). Daily closing returns series were computed from daily closing prices using the formula below:

$$DCR_t = \log\left(\frac{DCP_t}{DCP_{t-1}}\right) \times 100 \quad (68)$$

Where, DCR_t is the closing returns at the day t while DCP_t and DCP_{t-1} are the closing prices at the present day and previous day respectively.

The fitness performance of the proposed ESLD was compared with that of similar distributions such as the skew Laplace distribution and Laplace distribution using Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) as defined below:

$$AIC = 2k - 2\ln(LL) \quad (69)$$

$$BIC = -2LL + k \log(n) \quad (70)$$

Where, n is the number of observations, k is the number of parameters and LL is the log-likelihood.

Data set 1

0.46	-0.14	0.10	0.36	0.02	0.25	-0.04	-0.60	0.41	0.25
-0.21	-0.26	0.05	0.91	0.02	-0.16	0.07	-0.07	0.23	0.35
-0.05	-0.44	0.22	0.45	-0.28	-0.05	0.48	-0.08	-0.12	-0.28
0.27	0.24	0.38	0.14	-0.06	-0.13	-0.12	0.37	0.05	

Data set 2

0.11	-0.20	-0.42	0.08	0.43	1.23	0.95	1.73	0.57	0.47
1.47	-0.19	1.78	0.09	0.19	-0.42	0.39	-0.29	0.42	-0.37
-0.45	-0.49	0.71	0.14	2.28	2.00	3.94	-0.92	0.87	-0.29
0.79	3.41	-2.98	1.54	0.54	0.88	0.13	0.33	1.91	-0.39
0.96	-1.01	-1.23	-2.64	2.00	-0.54	-3.78	4.02	-1.58	-1.15
0.19	2.10	1.73	0.02	-0.33	0.80	-0.53	-0.15	1.04	-1.01

Table 1: Summary statistics for the data set

Data set	n	Min.	Max.	Mean	SD	Skewness
Data set 1 (S&P 500)	39	-0.60	0.91	.0763	.29451	.268
Data set II (Bitcoin)	60	-3.78	4.02	.3484	1.42987	.013

SD- Standard deviation.

Table 2: Results of the estimated parameters as well as the goodness fit for the proposed distribution and other similar distribution

Dataset	Distributions	δ	θ	ε	λ	LL	AIC	BIC
I	ESLD	-0.060	0.134	0.368	1.052	-5.431	18.862	17.975
	SLD	-0.061	0.102	0.348	-	-7.938	21.876	20.649
	LD	0.048	0.233	-	-	-9.247	22.494	21.677
II	ESLD	0.103	0.431	0.462	1.203	-98.346	204.691	203.804
	SLD	0.111	0.505	0.442	-	-103.041	212.082	211.417
	LD	0.189	1.034	-	-	-103.620	211.240	210.796

ESLD- Exponentiated skew Laplace Distribution, SLD- Skew Laplace Distribution, LD- Laplace Distribution, SGED- Skew Generalized Error Distribution.

Result presented in Table 3 show the parameter estimates of the proposed ESLD and compared to other related probability distribution (Skew Laplace distribution and Laplace distribution) as well as their fitness performance. Result of the LogLikelihood shows that among these distributions, the proposed ESLD reported the highest LogLikelihood in the two real data compared with other competing distributions. Similarly, the AIC and BIC reported by the proposed ESLD is lower than that of Skew Laplace distribution and Laplace distribution for both the two data set. These result show better fitness performance of the proposed ESLD than both Skew Laplace distribution and Laplace distribution in modelling financial data.

IV. Discussion

This study improve on the robustness of the Skew Laplace distribution introduced by Safavinejad *et al.* [4] by introducing additional shape parameter using the method of exponentiation. We derived some of the statistical properties of the proposed Skew Laplace Distribution (ESLD) after ensuring that the proposed ESLD satisfied the properties of a statistical distribution. Some of the statistical properties of the proposed ESLD derived include: moment generating function, characteristic function, hazard function and reliability function. The estimation of its parameters was carried out using the maximum likelihood. The performance of the proposed ESLD compared with other similar distributions were demonstrated empirically with returns from S & P 500 (Dataset 1) and returns from Bitcoin (Dataset 2). The findings suggest that the new distribution outperforms the existing models considered, indicating its better representation and flexibility when compared with some existing models. This findings show that the ELSD outperformed Skew Laplace Distribution (SLD) which indicates that the use of the method of exponentiation improve the performance of distribution. This is corroborated by that of other studies, Agboola *et al.*, [10], Oguntunde *et al.*, [11], Nadarajah and Bakar [12], Datta and Datta [13], Andrade *et al.*, [14] and Adubisi *et al.*, [15] which also found that exponentiated distribution performed better than their parent distributions.

References

- [1] Lekshmi, S. and Sebastian, S. (2014). A skewed generalized discrete Laplace distribution. *International Journal of Mathematics and Statistics Invention*, 2: 95-102.
- [2] Nadarajah, S. and Kotz, S. (2003) The Exponentiated Frechet Distribution. *Interstat Lectronic Journal*,14:1-7.
- [3] Kotz, S., Kozubowski, T. J. and Podgorski, K. The Laplace Distribution and Generalizations. Birkh ¨auser, Berlin, 2001.
- [4] Safavinejad, M., Jomhoori, S. and Noughabi, H.A. (2016). Testing skew-Laplace distribution using density-based empirical likelihood approach. *Journal of Statistical Research*, 13: 1–24.
- [5] Puig, P. and Stephens, A.M. (2007). Goodness of fit tests for the skew-Laplace distribution. *SORT*, 31 : 45-54.
- [6] Julia, O. and Vives-Rego, J. (2005). Skew-Laplace Distribution in Gram-negative Bacterial Axenic cultures: New Insights into Intrinsic Cellular Heterogeneity. *Microbiology*, 151:749- 755.
- [7] Jing, H.; Liu, Y.; Zhao, J. (2022). Asymmetric Laplace Distribution Models for Financial Data: VaR and CVaR . *Symmetry*, 14 :807.
- [8] Shi, Y.,Ng, C.T.and Yiu.K.C. (2018). Portfolio selection based on asymmetric Laplace distribution, coherent risk measure, and expectation-maximization estimation. *Quantitative Finance and Economics*, 2: 776-797.
- [9] Gupta, R. C., Gupta, P. L. and Gupta, R. D (1998). Modeling failure time data by Lehman alternatives, *Communications in Statistics, Theory and Methods*, 27: 887–904.
- [10] Agboola, S., Dikko, H.G. and Asiribo, O.E. (2018). On A New Exponentiated Error Innovation Distributions: Evidence of Nigeria Stock Exchange. *Journal of Statistics Applications and Probability Letters*, 7: 321-331.
- [11] Oguntunde, P. E., Adejumo A. O. and Balogun O. S. (2014). Statistical Properties of the Exponentiated Generalized Inverted Exponential Distribution. *Applied Mathematics* , 4: 47-55.
- [12] Nadarajah, S. and Bakar,S. A. A. (2016). An exponentiated geometric distribution. *Applied Mathematical Modelling*, 40:6775-6787.
- [13] Datta, D. and Datta, D. (2013). Comparison of Weibull distribution and exponentiated Weibull distribution based estimation of mean and variance of wind data.*International Journal of Energy, Information and Communications*, 4: 1-11.
- [14] Andrade, T.A. , Rodrigues, H., Bourguignon, M. and Cordeiro, G. (2015). The exponentiated generalized Gumbel distribution. *Revista Colombiana de Estadística*, 38: 123-143.
- [15] Adubisi, O.D, Abdulkadir, A.,Farouk, U.A., and Chiroma, H. (2022). The exponentiated half logistic skew-t distribution with GARCH-type volatility models. *Scientific African*, 16 : e01253.

A NOVEL ASYMMETRIC COMPOUND CLASS OF DISTRIBUTIONS WITH ESTIMATION AND APPLICATION

A.G. Al-Kilany¹, Amal S. Hassan^{2*} , L.S. Diab³, and E.S. El-Atfy¹

¹Faculty of Science for (girls), Al-Azhar University, Nasr City, 11884, Egypt,

^{2*}Faculty of Graduate Studies for Statistical Research, Cairo University, 12613, Giza, Egypt
amal52_soliman@cu.edu.eg

³Department of Mathematics and Statistics, Faculty of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia
*Correspondence: Email amal52_soliman@cu.edu.eg

Abstract

This paper introduces and discusses the novel asymmetric class of distributions that have the name inverse power Lomax power series (IPLPS). This class of distributions is produced by combining the inverse power Lomax with the power series distributions. This combined approach provides an opportunity for the creation of flexible distributions with significant physical implications in many fields, like biology and engineering. The IPLPS distributions encompass several new compound distributions as sub-models along with a new class of compound distributions. Many statistical features, including moments, quantile function, conditional moments, inverse moments, uncertainty measures, and probability-weighted moments, are obtained. As a special model of the generated class, the parameters of the inverse power Lomax Poisson distribution are estimated by different methods, including least squares, Cramér von Mises, maximum likelihood, and weighted least squares. Through an extensive simulation analysis, the execution of different parameter estimation techniques for the inverse power Lomax Poisson model is performed to show its validity based on its mean squared error and absolute bias. Two real datasets are utilized to show the practicality of the newly generated model. Results show that the inverse power Lomax Poisson distribution provides the most fitted model for these datasets in comparison to other distributions such as power Lomax, Marshall-Olkin power Lomax, power Lomax Poisson, and Topp-Leone Lomax distributions.

Keywords: Power series distributions, inverse power Lomax distribution, moments, compounding, Havrda and Charvat measure, Cramér von Mises.

1. Introduction

Recent academic focus has shifted towards the creation of new univariate distributions. Univariate distributions, whether for theoretical, practical, or combined purposes, hold significant importance in statistical and related fields. Analyzing the reliability of experimental failure components is a primary objective. It's often assumed that these failures occur due to certain processes, yet a thorough investigation into the causes of component failure seems lacking; see Barreto-Souza et al. [1]. Consider a system's lifetime composed of N components, and N is the discrete random variable that follows geometric, Poisson, logarithmic, or binomial distributions.

Power series (PS) is the general form of these chosen distributions. For further information on the PS class of distributions, refer to Noack [2]. Suppose that X denotes the continuous random variable for each component. Consequently, the random variable $X = \text{Min}(X_1, X_2, \dots, X_N)$ or $X = \text{Max}(X_1, X_2, \dots, X_N)$ signifies any component lifetimes depending on whether they are arranged in a series or in parallel structure, respectively.

Suppose that the random variable N associated with the PS class of distributions, characterized by a probability mass function, is given by:

$$P(N = n) = \frac{a_n \theta^n}{C(\theta)}, \quad n = 1, 2, \dots$$

where $a_n \geq 0$ only dependent on n , and $C(\theta) = \sum_{n=1}^{\infty} a_n \theta^n$ is finite.

Several compound lifetime models have been created by combining several lifetime distributions with the PS class of distributions. For instance, the exponential PS [3], the Weibull-PS [4], Lindley-PS [5], exponential Pareto-PS [6], Burr XII-PS [7], exponentiated power Lindley-PS [8], generalized Burr XII-PS [9], odd log-logistic-PS [10], Topp-Leone generalized exponential-PS [11], power function-PS [12], inverse gamma-PS [13], inverse exponentiated Lomax-PS [14], beta exponential-PS [15], unit exponentiated half logistic-PS [16], inverted Nadarajah-Haghighi-PS [17], power quasi Lindley-PS [18], unit Burr XII-PS [19], unit Gompertz-PS [20], log-logistic modified Weibull-PS [21], power inverted Topp-Leone-PS [22] distributions, among others.

Numerous writers have highlighted the significance and usefulness of inverted distributions in many fields, including engineering, economics, and medicine. In this work, the inverse power Lomax (IPL) distribution with three parameters, which was recently presented by Hassan and Abd-Allah [23], attracts our attention. The probability density function (PDF) and the cumulative distribution function (CDF) of the IPL distribution, having $\alpha, \beta > 0$, as shape parameters, and its scale parameter $\lambda > 0$, is defined, respectively, as follows:

$$g(x) = \frac{\alpha\beta}{\lambda} x^{-(\beta+1)} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-(\alpha+1)}; \quad x > 0, \quad (1)$$

and,

$$G(x) = \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha}; \quad x > 0. \quad (2)$$

Due to the IPL distribution's non-monotonic failure rate, it offers greater flexibility, making it more appropriate for various practical data modeling and analytic applications. Hassan and Abd-Allah [23] looked at a few statistical characteristics and provided estimators of the parameters in censored samples. Shi and Shi [24] studied how to statistically estimate parameters of the IPL distribution when employing progressive first-failure censoring. The inference of the IEL distribution based on generalized order statistics was discussed by Nassr et al. [25].

This paper's primary objective is to create a novel asymmetric compound class of distributions that is produced by combining the IPL and PS distributions to analyze a system with parallel components; this system is known as the inverse power Lomax power series (IPLPS). We are introducing this class due to the following:

- To design several distinct models with different symmetric and asymmetric density and hazard rate functions (HRFs) shapes.
- To go over a few of its statistical characteristics, including moments, quantile function (QF), conditional moments, uncertainty measures, inverse moments, and probability-weighted moments (PWMs).

- To estimate the IPLPS class of distribution parameters, some estimation techniques are taken into consideration, such as weighted least squares (WLS), maximum likelihood (ML), Cramér von Mises (CM), and least squares (LS).
- To evaluate the effectiveness of various estimates using specific metrics, a dedicated simulation study is conducted for one special model, namely the IPL Poisson (IPLP) distribution.
- The IPLP distribution, as a sub-model within this class, demonstrates superiority over certain other distributions, as revealed through an analysis of two real-data applications.

This paper’s contents are arranged as follows. The IPLPS distributions are introduced in Section 2. Many structural properties of the class are provided in Section 3. Section 4 provides certain examples of the suggested distributions. Parameter estimators for the IPLPS class using different classical methods are shown in Section 5, while Section 6 provides simulation studies. Section 7 presents the application of the suggested distribution’s particular case, whereas Section 8 offers concluding findings.

2. Construction of the IPLPS Class

The IPLPS class is introduced in this section. This class of distributions is motivated by a key assumption that renders it suitable for application in each survival and reliability study. Specifically, it assumes that a device’s failure arises from the presence of an unspecified number of initial faults, denoted as N , of the same type. These faults remain undetected until they lead to failure and are subsequently fixed.

If we consider $X_i, i=1, \dots, N$ to represent the time until device failure caused by the i th defect supposing that these X_i ’s are independent and identically distributed (iid) IPL random variables, independent of N , then a truncated PS random variable, a distribution within the IPLPS class, can be utilized to model the time until the last failure. This proposed class of distributions can effectively model systems with parallel components, as many biological and industrial applications frequently do. Currently, let us explore a parallel of N iid random variables from the IPL distribution, denoted as X_i , where $i=1, \dots, N$.

Assuming that $X = \max \{X_i\}_{i=1}^N$ be iid breakdown times of N items connected at a parallel structure, then the conditional CDF of $X | N$ is introduced as:

$$F_{X|N=n}(x) = [G(x)]^n = \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-an},$$

where $G(\cdot)$ is the CDF (2) of the IPL distribution. The joint CDF is given as follows:

$$P(X \leq x, N = n) = P(N = n)F_{X|N=n}(x) = \frac{a_n \theta^n}{C(\theta)} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-an}.$$

Hence, the IPLPS class is represented by the marginal CDF of X , which takes the form:

$$F(x; \psi) = \sum_{n=1}^{\infty} \frac{a_n \theta^n}{C(\theta)} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-an} = \frac{1}{C(\theta)} C \left(\theta \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha} \right), \tag{3}$$

where, $\psi \equiv (\alpha, \beta, \lambda, \theta)$ denotes the set of parameters, $\lambda, \theta > 0$, represent scale parameters and, $\beta, \alpha > 0$ indicate shape parameters. Another simplified form for (3) is as follows:

$$F(x; \psi) = \frac{1}{C(\theta)} C[k(x; \psi)]; \quad x > 0, \tag{4}$$

where, $k(x;\psi) = \theta \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha}$. Also, the PDF of the IPLPS class of distributions can be introduced by:

$$f(x;\psi) = \frac{\alpha\beta\theta}{\lambda C(\theta)} x^{-\beta-1} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha-1} C'[k(x;\psi)]; \quad x > 0. \quad (5)$$

The survival function and HRF associated to the IPLPS distribution are expressed as follows, respectively:

$$\bar{F}(x;\psi) = 1 - \frac{C[k(x;\psi)]}{C(\theta)},$$

and,

$$h(x;\psi) = \frac{\alpha\beta\theta x^{-\beta-1} C'[k(x;\psi)]}{\lambda [C(\theta) - C[k(x;\psi)]]} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha-1}.$$

Proposition: When θ approaches zero, the IPL distribution appears as a limiting special case of the IPLPS distributions

Proof: If θ approaches zero, then

$$\lim_{\theta \rightarrow 0^+} F(x;\psi) = \frac{\left[\left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha} \right] + a_1^{-1} \lim_{\theta \rightarrow 0^+} \sum_{n=2}^{\infty} na_n \theta^{n-1} \left[\left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha} \right]^{n-1}}{1 + a_1^{-1} \lim_{\theta \rightarrow 0^+} \sum_{n=2}^{\infty} na_n \theta^{n-1}} = \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha},$$

which represents the CDF (2) of the IPL distribution.

Lemma 1: For the IPLPS class of distributions, the density function can be expressed as an infinite mixture of IPL distributions with parameters $(n\alpha, \beta, \lambda)$,

$$f(x;\psi) = \sum_{n=1}^{\infty} P(N = n) g(x; n\alpha, \beta, \lambda).$$

Proof: The following is an alternative form for the PDF given in Equation (5):

$$\begin{aligned} f(x;\psi) &= \sum_{n=1}^{\infty} \frac{n\alpha\beta\theta^n a_n x^{-(\beta+1)}}{\lambda C(\theta)} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-(\alpha n+1)} \\ &= \sum_{n=1}^{\infty} P(N = n) g(x; n\alpha, \beta, \lambda), \end{aligned} \quad (6)$$

where, $g(x; n\alpha, \beta, \lambda)$ refers to the IPL distribution's density function (1) with parameters $(n\alpha, \beta, \lambda)$.

3. Some Statistical Properties

Here, several distinct statistical features of the IPLPS distributions are derived, which may include the quantile function, r th moment and inverse moment, PWMs, conditional moments, and entropy measures.

3.1 Quantile Function

The QF of the IPLPS class of X , denoted by $x_u = Q(u) = F^{-1}(u)$, is represented as follows:

$$x_u = \left\{ \lambda \left[\left(\frac{C^{-1}[uC(\theta)]}{\theta} \right)^{-1/\alpha} - 1 \right] \right\}^{-1/\beta} \quad (7)$$

Particularly, the median, denoted by m , of the IPLPS distribution, is derived by letting $u = 0.5$ in Equation (7).

3.2 Moments and Inverse Moments

Most important properties for any distribution are concluded using ordinary moments. The r th moment of X can be introduced by using Equation (6) as follows:

$$\mu'_r = \sum_{n=1}^{\infty} P(N=n) \int_0^{\infty} \frac{n\alpha\beta}{\lambda} x^{r-(\beta+1)} \left(1 + \frac{x^{-\beta}}{\lambda} \right)^{-(\alpha n+1)} dx.$$

After simplification, the r th moment of the IPLPS distribution, can be written as:

$$\mu'_r = \sum_{n=1}^{\infty} n\alpha\lambda \frac{-r}{\beta} P(N=n) B\left(1 - \frac{r}{\beta}, \alpha n + \frac{r}{\beta} \right), \quad \beta > r, \quad r = 1, 2, \dots \quad (8)$$

where, $B(\dots)$ denotes the beta function. For $r=1$ in Equation (8), the mean of the IPLPS distribution is given. Also, we can get the IPLPS moment generating function from the moments by the following equation:

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu'_r = \sum_{r=0}^{\infty} \sum_{n=1}^{\infty} \frac{n\alpha\lambda}{r!} \frac{-r}{\beta} t^r P(N=n) B\left(1 - \frac{r}{\beta}, \alpha n + \frac{r}{\beta} \right).$$

Furthermore, the r th inverse moment for the IPLPS distribution is derived using Equation (6) which leads to:

$$\delta_r = \sum_{n=1}^{\infty} n\alpha\lambda \frac{r}{\beta} P(N=n) B\left(1 + \frac{r}{\beta}, \alpha n - \frac{r}{\beta} \right), \quad r = 1, 2, \dots$$

3.3 Conditional Moments

Studying the conditional moments is very important in lifetime models. The conditional moments of the IPLPS distribution, defined by $E(X^r | X > t)$, can be introduced by the following lemma.

Lemma 3.1: Supposing that X has the IPLPS $(x; \psi)$, the r th conditional moment of X , is obtained such that:

$$M_r = \sum_{n=1}^{\infty} \frac{n\alpha P(N=n)}{\lambda^{r/\beta} \bar{F}(t; \psi)} B\left(1 - \frac{r}{\beta}, \alpha n + \frac{r}{\beta}, \left(1 + \frac{t^{-\beta}}{\lambda} \right)^{-1} \right),$$

where $B(\dots)$ refers to the incomplete beta function.

Proof: Since

$$M_r = E(X^r | X > t) = \frac{1}{\bar{F}(t; \psi)} \int_t^{\infty} x^r f(x; \psi) dx.$$

Hence, by inserting the PDF (4) in M_r then

$$M_r = \sum_{n=1}^{\infty} \frac{n\alpha\beta P(N=n)}{\lambda \bar{F}(t;\psi)} \int_t^{\infty} x^{r-\beta-1} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-(\alpha n+1)} dx.$$

By simplifying, then the r th conditional moment of the IPLPS class of distributions can be rewritten as,

$$M_r = \sum_{n=1}^{\infty} \frac{n\alpha P(N=n)}{\lambda^{r/\beta} \bar{F}(t;\psi)} \int_0^{(1+t\beta/\lambda)^{-1}} (1-z)^{-\frac{r}{\beta}} z^{\alpha n + \frac{r}{\beta} - 1} dz = \sum_{n=1}^{\infty} \frac{n\alpha P(N=n)}{\lambda^{r/\beta} \bar{F}(t;\psi)} B\left(1 - \frac{r}{\beta}, \alpha n + \frac{r}{\beta}, \left(1 + \frac{t\beta}{\lambda}\right)^{-1}\right),$$

where $B(\cdot, \cdot, x)$ is the incomplete beta function and $\bar{F}(x;\psi)$ represents the IPLPS survival function.

3.4 Probability-Weighted Moments

Greenwood et al. [26] were the first to propose the PWM approach, with the main goal being the derivation of quantiles and parameter estimators for several generalized distributions that are only analytically represented in reverse form. Eventually, for a random variable X , the PWM is expressed by the following equation:

$$\phi_{s,r} = E[X^s F(x)^r] = \int_{-\infty}^{\infty} x^s f(x) (F(x))^r dx. \tag{9}$$

Substituting Equations (4) and (5) in Equation (9) we get:

$$\phi_{s,r} = \int_0^{\infty} x^s \frac{\alpha\beta\theta}{\lambda} x^{-\beta-1} \left(1 + \frac{x^{-\beta}}{\lambda}\right)^{-\alpha-1} \frac{C'[k(x;\psi)]}{\{C(\theta)\}^{r+1}} (C[k(x;\psi)])^r dx. \tag{10}$$

An expansion for $(C[k(x;\psi)])^r$ can be written as follows:

$$\begin{aligned} [C(k(x;\psi))]^r &= \left\{ \sum_{n=1}^{\infty} a_n [k(x;\psi)]^n \right\}^r = (a_1)^r (k(x;\psi))^r \\ &\times \left\{ 1 + \frac{a_2}{a_1} [k(x;\psi)] + \frac{a_3}{a_1} [k(x;\psi)]^2 + \dots \right\}^r \\ &= [a_1 k(x;\psi)]^r \left\{ \sum_{m=0}^{\infty} c_m [k(x;\psi)]^m \right\}^r, \quad c_m = \frac{a_{m+1}}{a_1}, m = 1, 2, 3, \dots \end{aligned} \tag{11}$$

After that, using the Gradshteyn and Ryzhik [27] relation, which states that; for any positive integer m , the following expansion, for a positive integer r , is used:

$$\left(\sum_{m=0}^{\infty} c_m w^m \right)^r = \sum_{m=0}^{\infty} d_{r,m} w^m, \tag{12}$$

where $d_{r,0} = 1, t \geq 1$ and the coefficients $d_{r,t} = t^{-1} \sum_{m=1}^t (m(r+1) - t) c_m d_{r,t-m}$. Then, using expansion

(12) in (11) provides the following:

$$[C(k(x;\psi))]^r = a_1^r \sum_{m=0}^{\infty} d_{r,m} [k(x;\psi)]^{m+r}. \tag{13}$$

In addition,

$$C'[k(x;\psi)] = \sum_{n=1}^{\infty} n a_n [k(x;\psi)]^{n-1}. \tag{14}$$

Assuming that $z = n - 1$, then Equation (14) is rewritten in this form:

$$C'[k(x; \psi)] = a_1 \sum_{z=0}^{\infty} b_z(z+1)[k(x; \psi)]^z, \quad b_z = \frac{a_{z+1}}{a_1}. \quad (15)$$

Hence, the PWM of the IPLPS class of distributions is represented by placing Equations (13) and (15) into (10) and after some simplification,

$$\phi_{s,r} = \frac{\alpha\beta\theta^{r+m+z+1}}{\{\lambda C(\theta)\}^{r+1}} \sum_{m,z=0}^{\infty} a_1^{r+1} d_{r,m} b_z(z+1) \int_0^{\infty} x^{s-\beta-1} \left(1 + \frac{x-\beta}{\lambda}\right)^{-\alpha(r+m+z+1)-1} dx.$$

Hence,

$$\phi_{s,r} = \sum_{m,z=0}^{\infty} A^* B\left(1 - \frac{s}{\beta}, \alpha(r+m+z+1) + \frac{s}{\beta}\right), \quad s < \beta.$$

where, $A^* = \frac{\alpha\theta^{r+m+z+1} a_1^{r+1} d_{r,m} b_z(z+1) \lambda^{-\frac{s}{\beta}}}{\{C(\theta)\}^{r+1}}$, and $B(.,.)$ is the beta function.

3.6 Entropy Measures

Entropy serves as a metric for quantifying the uncertainty within data and finds applications across diverse fields such as science, physics, and engineering. Essentially, higher entropy values indicate greater uncertainty within the data. In this sub-section, expressions for certain entropy measures within the IPLPS class are derived. Let X refers to random variable drawn from IPLPS distributions, so the Rényi entropy (RE) can be represented by the following equation:

$$I_R = \frac{1}{\delta-1} \log \left[\int_0^{\infty} f^{\delta}(x; \psi) dx \right], \quad \delta \neq 1, \delta > 0. \quad (16)$$

Suppose $IP = \int_0^{\infty} (f(x; \psi))^{\delta} dx$, then by using PDF (5) and expansion (14) in integral IP, we have

$$IP = \frac{(\alpha\beta\theta)^{\delta}}{\{\lambda C(\theta)\}^{\delta}} \int_0^{\infty} x^{-\delta(\beta+1)} \left(1 + \frac{x-\beta}{\lambda}\right)^{-\delta(\alpha+1)} \left[\sum_{n=1}^{\infty} na_n [k(x; \psi)]^{n-1} \right]^{\delta} dx, \quad (17)$$

But $\left[\sum_{n=1}^{\infty} na_n [k(x; \psi)]^{n-1} \right]^{\delta} = a_1^{\delta} \left[\sum_{m=0}^{\infty} c_m^{\bullet} [k(x; \psi)]^m \right]^{\delta}$, $c_m^{\bullet} = \frac{a_{m+1}}{a_1} (m+1)$, $m = 1, 2, \dots$

According to Ref. [27], the previous equation can be expressed as:

$$\left[\sum_{n=1}^{\infty} na_n [k(x; \psi)]^{n-1} \right]^{\delta} = a_1^{\delta} \sum_{m=0}^{\infty} d_{\delta,m} [k(x; \psi)]^m. \quad (18)$$

By using Equation (18) in the last term in (17), then

$$IP = \sum_{m=0}^{\infty} \Lambda_m B\left(\frac{\delta(\beta+1)-1}{\beta}, \delta\alpha + \alpha m + \delta - \frac{\delta(\beta+1)-1}{\beta}\right), \quad (19)$$

$$\Lambda_m = \frac{(\alpha\theta a_1)^{\delta} \beta^{\delta-1} d_{\delta,m} \theta^m \lambda^{\frac{\delta-1}{\beta}}}{(C(\theta))^{\delta}}.$$

Hence, by substituting (19) in Equation (16), the RE of the IPLPS class of distributions takes the following form:

$$I_R(\delta) = \frac{1}{1-\delta} \log \left[\sum_{m=0}^{\infty} \Lambda_m B \left(\frac{\delta(\beta+1)-1}{\beta}, \delta\alpha + \alpha m + \delta - \frac{\delta(\beta+1)-1}{\beta} \right) \right].$$

Tsallis entropy (TE), introduced by Tsallis [28] as a thermodynamic measure, has a wide application across various real-world domains. Generally, TE offers intriguing explanations in physical, chemical, and biological phenomena. The TE measure is represented as:

$$I_{TE} = \frac{1}{\delta-1} \left[1 - \int_0^{\infty} f \delta(x) dx \right], \delta \neq 1, \delta > 0.$$

Using the similar procedure discussed above, the TE is given by:

$$I_{TE} = \frac{1}{\delta-1} \left[1 - \left(\sum_{m=0}^{\infty} \Lambda_m B \left(\frac{\delta(\beta+1)-1}{\beta}, \delta(\alpha+1) + \alpha m - \frac{\delta(\beta+1)-1}{\beta} \right) \right) \right].$$

4. Special Sub-Models

Here, certain special cases of this class are introduced. Graphs depicting the PDF and HRF are presented to showcase the IPLP distribution' flexibility for some chosen values for the parameters.

- If $\beta = 1$, the IPLPS class offers the inverse Lomax PS class of distributions (new-class).
- Letting $C(\theta) = e^\theta - 1$, the IPLPS distribution turns to the IPLP distribution.
- Supposing that $\beta = 1$, $C(\theta) = e^\theta - 1$ and the IPLPS distribution provides the IL Poisson (ILP) distribution (new).
- Setting $C(\theta) = -\log(1-\theta)$, the IPLPS class becomes the IPL logarithmic (IPLL) distribution (new).
- By putting $\beta = 1$, and $C(\theta) = -\log(1-\theta)$, the IPLPS distribution provides the IL logarithmic (ILL) distribution (Buzaridah et al. [29]).
- Considering that $C(\theta) = \theta(1-\theta)^{-1}$, the IPLPS distribution introduces the IPL geometric (IPLG) distribution (new).
- By letting $\beta = 1$, and $C(\theta) = \theta(1-\theta)^{-1}$, the IPLPS distribution presents the IL geometric (ILG) distribution (new).
- Substituting $C(\theta) = (1-\theta)^m - 1$, that yields the IPL binomial (IPLB) distribution (new).
- Taking $\beta = 1$, besides $C(\theta) = (1-\theta)^m - 1$, it gives the IL binomial (ILB) distribution.

The IPLP Distribution

By setting $C(\theta) = e^\theta - 1$, and $C'(\theta) = e^\theta$ in (4) and (5), the PDF and CDF of the IPLP distribution is obtained by:

$$f_1(x; \psi) = \frac{\alpha\beta\theta}{\lambda(e^\theta - 1)} x^{-\beta-1} \left(1 + \frac{x^{-\beta}}{\lambda} \right)^{-\alpha-1} e^{k(x; \psi)}; \quad x > 0,$$

$$F_1(x; \psi) = \frac{e^{k(x; \psi)} - 1}{e^\theta - 1}; \quad x > 0,$$

where, α, β denote the shape parameters and θ, λ refer to the scale parameters. The HRF of the IPLP distribution is given as follows:

$$H_1(x; \psi) = \frac{\alpha\beta\theta x^{-\beta-1} e^{k(x; \psi)}}{\lambda [e^\theta - e^{k(x; \psi)}]} \left(1 + \frac{x^{-\beta}}{\lambda} \right)^{-\alpha-1}; \quad x > 0.$$

The PDF and HRF plots for the IPLP distribution are given in Figure 1.

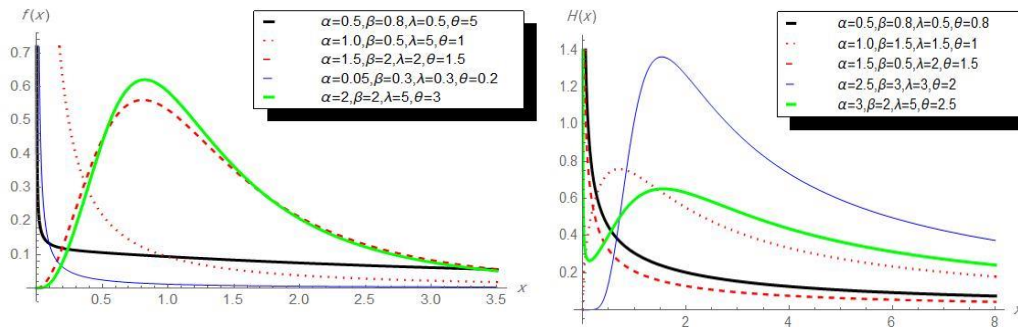


Figure 1: PDF and HRF plots for specific parameter values of the IPLP distribution.

Figure 1 indicates that the IPLP distribution's density may exhibit reversed-J, skewed to the right, or unimodal shapes. Moreover, the HRF can take on increasing, decreasing, upside down, or reversed J-shaped forms at different parameter values. This suggests that the IPLP distribution is versatile for fitting datasets with diverse shapes.

5. Parameter Estimation

Here, the parameter estimation for the IPLPS distributions is discussed by applying the ML, LS, WLS, and CM methods.

5.1 Maximum Likelihood Estimators

Let x_1, x_2, \dots, x_n be a simple random sample from the IPLPS class of distributions with a set of parameters $\psi = (\alpha, \beta, \lambda, \theta)^T$. The likelihood function of this sample, denoted by L_n based on the observed random sample of size n from density (5) is given by:

$$L_n = \left(\frac{\alpha\beta\theta}{\lambda C(\theta)} \right)^n \prod_{i=1}^n x_i^{-\beta-1} \left(1 + \frac{x_i^{-\beta}}{\lambda} \right)^{-\alpha-1} C'(k(x_i; \psi)).$$

The log-likelihood, say $\log L_n$, can be expressed as:

$$\begin{aligned} \log L_n = & n \log(\alpha\beta\theta) - n \log(\lambda C(\theta)) - \sum_{i=1}^n (\beta+1) \log x_i - (\alpha+1) \sum_{i=1}^n \log \left(1 + \frac{x_i^{-\beta}}{\lambda} \right) \\ & + \sum_{i=1}^n \log(C'(k(x_i; \psi))), \end{aligned} \tag{20}$$

Hence, by differentiating (20) with respect to α, β, λ and θ , respectively, yields

$$\begin{aligned} \frac{\partial \log L_n}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^n \log \left(1 + \frac{x_i^{-\beta}}{\lambda} \right) - \sum_{i=1}^n \frac{\theta C''(k(x_i; \psi))}{C'(k(x_i; \psi))} \left(1 + \frac{x_i^{-\beta}}{\lambda} \right)^{-\alpha} \log \left(1 + \frac{x_i^{-\beta}}{\lambda} \right), \\ \frac{\partial \log L_n}{\partial \beta} &= \frac{n}{\beta} - \sum_{i=1}^n \log x_i - \sum_{i=1}^n \frac{(\alpha+1)x_i^{-\beta} \log x_i}{\lambda + x_i^{-\beta}} + \sum_{i=1}^n \frac{\alpha \theta x_i^{-\beta} \log x_i C''(k(x_i; \psi))}{\lambda C'(k(x_i; \psi))} \left(1 + \frac{x_i^{-\beta}}{\lambda} \right)^{-\alpha-1}, \\ \frac{\partial \log L_n}{\partial \lambda} &= -\frac{n}{\lambda} + (\alpha+1) \sum_{i=1}^n \frac{x_i^{-\beta}}{\lambda^2 + \lambda x_i^{-\beta}} + \sum_{i=1}^n \frac{\theta \alpha C''(k(x_i; \psi))}{\lambda^2 C'(k(x_i; \psi))} \left(1 + \frac{x_i^{-\beta}}{\lambda} \right)^{-\alpha-1}, \end{aligned}$$

and,

$$\frac{\partial \log L_n}{\partial \theta} = \frac{n}{\theta} - \frac{nC'(\theta)}{C(\theta)} + \sum_{i=1}^n \frac{C''(k(x_i; \psi))}{C'(k(x_i; \psi))} \left(1 + \frac{x_i^{-\beta}}{\lambda} \right)^{-\alpha}.$$

Then the ML estimates (MLEs) for the parameters α, β, λ and θ , denoted by $\hat{\alpha}, \hat{\beta}, \hat{\lambda}$ and $\hat{\theta}$, can be derived by setting $(\partial L_n / \partial \alpha), (\partial L_n / \partial \beta), (\partial L_n / \partial \lambda)$ and $(\partial L_n / \partial \theta)$ to be zero and solving these equations numerically.

5.2 Least Squares and Weighted Least Squares Estimators

Consider $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ refers to an observed ordered sample and x_1, x_2, \dots, x_n represents n random samples from the IPLPS distribution. Johnson et al. [30] claimed that the distribution's expectation and variance are determined independently of the unknown parameter by

$$E(F(X_{(i)})) = \frac{i}{n+1}, \text{ and, } \text{Var}(F(X_{(i)})) = \frac{i(n-i+1)}{(n+1)^2(n+2)},$$

where $F(X_{(i)})$ indicates the CDF of any given distribution and $X_{(i)}$ denotes the statistic of order i . So, the LS estimates (LSEs) and WLS estimates (WLSEs) can be given by the minimization of the sum of all squared errors

$$H(\psi) = \sum_{i=1}^n v_i \left[F(x_{(i)}; \psi) - E(F(x_{(i)}; \psi)) \right]^2.$$

The LSEs and WLSEs of α, β, λ and θ , are produced by the minimization of the preceding function

$$H(\psi) = \sum_{i=1}^n v_i \left[\left(\frac{C[k(x_{(i)}; \psi)]}{C(\theta)} \right) - \frac{i}{n+1} \right]^2. \tag{21}$$

Based on Equation (21), the LSEs $\hat{\alpha}_1, \hat{\beta}_1, \hat{\lambda}_1$ and $\hat{\theta}_1$ are provided by using $v_i = 1$, while the WLSE $\hat{\alpha}_2, \hat{\beta}_2, \hat{\lambda}_2$ and $\hat{\theta}_2$ are obtained by putting $v_i = \frac{(n+1)^2(n+2)}{i(n-i+1)}$.

These estimates can be given by solving each of the following equations numerically.

$$\begin{aligned} \frac{\partial H(\psi)}{\partial \alpha} &= \sum_{i=1}^n v_i \left[\frac{C(k(x_{(i)}; \psi))}{C(\theta)} - \frac{i}{n+1} \right] \varphi_1(x_{(i)}, \psi) = 0, \\ \frac{\partial H(\psi)}{\partial \beta} &= \sum_{i=1}^n v_i \left[\frac{C(k(x_{(i)}; \psi))}{C(\theta)} - \frac{i}{n+1} \right] \varphi_2(x_{(i)}, \psi) = 0, \\ \frac{\partial H(\psi)}{\partial \lambda} &= \sum_{i=1}^n v_i \left[\frac{C(k(x_{(i)}; \psi))}{C(\theta)} - \frac{i}{n+1} \right] \varphi_3(x_{(i)}, \psi) = 0, \\ \frac{\partial H(\psi)}{\partial \theta} &= \sum_{i=1}^n v_i \left[\frac{C(k(x_{(i)}; \psi))}{C(\theta)} - \frac{i}{n+1} \right] \varphi_4(x_{(i)}, \psi) = 0, \end{aligned}$$

where,

$$\begin{aligned} \varphi_1(x_{(i)}, \psi) &= -\frac{\theta}{C(\theta)} \left(1 + \frac{x_{(i)}^{-\beta}}{\lambda} \right)^{-\alpha} \log \left(1 + \frac{x_{(i)}^{-\beta}}{\lambda} \right) C'(x_{(i)}, \psi), \\ \varphi_2(x_{(i)}, \psi) &= \frac{\alpha \theta}{\lambda C(\theta)} \left(1 + \frac{x_{(i)}^{-\beta}}{\lambda} \right)^{-\alpha-1} x_{(i)}^{-\beta} \ln x_{(i)} C'(x_{(i)}, \psi), \end{aligned}$$

$$\varphi_3(x(i), \psi) = \frac{\alpha \theta x(i)^{-\beta}}{\lambda^2 C(\theta)} \left(1 + \frac{x(i)^{-\beta}}{\lambda} \right)^{-\alpha-1} C'(x(i), \psi),$$

and,

$$\varphi_4(x(i), \psi) = \frac{C'(k(x; \psi))}{C'(\theta)} \left(1 + \frac{x(i)^{-\beta}}{\lambda} \right)^{-\alpha} - \frac{C(k(x(i), \psi)) C'(\theta)}{(C'(\theta))^2}.$$

5.3 Cramèr –von-Mises Estimators

This method can be defined as a type of estimator that relies on minimal distance principles since it relies on the disparity between the empirical distribution function and the CDF estimate. According to Macdonald [31], in this method, the CM estimator's are presented as the minimization of the given equation with respect to α, β, λ , and θ , respectively,

$$H(\psi) = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{C(k(x(i), \psi))}{C(\theta)} - \frac{2i-1}{2n} \right]^2.$$

The CM estimates (CMEs) $\hat{\alpha}_3, \hat{\beta}_3, \hat{\lambda}_3$, and $\hat{\theta}_3$ can be obtained by differentiating the previous equation with respect to $\alpha, \beta, \lambda, \theta$, respectively, and equating it to zero.

6. Simulation Study

For each estimation problem, the investigation of the estimator's properties is very important. Analytical study of the obtained expressions for the estimators can't be effective due to their complexity. As a result, a numerical study will be established, handling the estimates' sampling distribution independently. This estimation is conducted in order to assess the estimators presented at the preceding section. All calculations are produced by using the *Mathematica11.3 program*. The performances of the different estimates will be compared according to their absolute bias (AB) and mean squared error (MSE). These numerical procedures will be shown by steps below:

Step 1: 1000 random samples given the sizes of 50, 100, 150, and 200 are conducted from the inverse power Lomax Poisson distribution.

Step 2: Four cases of parameter values have been selected such that:

Case 1 $\equiv (\alpha = 0.2, \beta = 0.5, \lambda = 0.5, \theta = 0.5)$, Case 2 $\equiv (\alpha = 0.1, \beta = 0.7, \lambda = 0.5, \theta = 0.5)$,

Case 3 $\equiv (\alpha = 0.35, \beta = 0.75, \lambda = 0.5, \theta = 0.5)$, Case 4 $\equiv (\alpha = 0.7, \beta = 0.25, \lambda = 0.5, \theta = 0.5)$.

Step 3: The MLEs, LSEs, WLSEs, and CMEs are derived for each unknown parameter.

Step 4: The ABs and MSEs of different estimates of unknown parameters are calculated. The results are written down in Tables A.1 to A.4 (Appendix A). By the help of these tables, the following conclusions can be concluded to predict the performance for all these different estimates

- For fixed value of $\lambda = 0.5$ and $\theta = 0.5$, ABs and MSEs for each α estimates and β estimate values in the MLEs decrease while sample size increases (see Table A.1).
- For fixed values of λ and θ , the MSEs of CMEs for α and β are decreasing and the sample size will be increasing in the same time (see Table A.3).
- For $\beta = 0.75$, and for fixed values of λ and θ , the MSEs of the WLSEs increase as the sample size increases (see Table A.2).
- By increasing the sample size, the ABs of MLEs at $\alpha = 0.35$ and $\beta = 0.75$ decrease consistently, for fixed values for λ and θ as shown in Table A.3.

- At $n = 50$ and $\lambda = 0.5$, the MSEs have the smallest values for all different sets of parameters at $\alpha = 0.35$ and $\beta = 0.75$, as indicated in Table A.4.
- As the sample size increases, it is evident through all estimation methods that both MSEs and ABs decrease, as demonstrated in Table A.1 for instance.
- Almost in all cases, the estimated MSEs of the MLEs are the smallest compared to other estimation methods across all parameter values.

7. Data Analysis

This section presents the application of the IPLP model on two real data sets, illustrating its practical adaptability and utility. The IPLP distribution is contrasted with alternative models including the power Lomax (PL) [32], PL Poisson (PLP) [33], Topp-Leone Lomax (TLLO) [34], and Marshall Olkins PL (MOPL) [35] distributions for two real datasets.

The first dataset has been introduced by Murthy et al. [36], represents 84 observations recording the failure time for specific aircraft windshield model. The dataset is as follows:

0.04	1.866	2.385	3.443	0.301	1.876	2.481	3.467	0.309	1.899
2.61	3.478	0.557	1.911	2.625	4.57	1.652	2.3	3.344	4.602
1.757	3.578	0.943	1.912	2.632	3.595	1.07	1.914	2.646	3.699
1.124	1.981	2.661	3.779	1.248	2.01	2.224	3.117	4.485	1.652
2.229	3.166	2.688	3.924	1.281	2.038	2.823	4.035	1.281	2.085
2.89	4.121	1.303	2.089	2.902	4.167	1.432	4.376	1.615	2.223
3.114	4.449	1.619	2.097	2.934	4.24	1.48	2.135	2.962	4.255
1.505	2.154	2.964	4.278	1.506	2.19	3	4.305	1.568	2.194
3.103	2.324	3.376	4.663						

To examine the utility of the proposed models, various criteria measures, including -2Log-likelihood (L^*), Akaike information criterion (A^*), Bayesian information criterion (B^*), consistent Akaike information criterion (C^*), the Kolmogorov-Smirnov distance (K^*) and its p-value (K^* -PV), and CM statistics (W^*) are evaluated. In general, the smaller the value of these statistics, a better fit model for the data will be found. Table 1 offers MLEs for all models that are suggested, and Table 2 lists several goodness of fitting metrics.

Table 1: MLEs for all parameters of the models fitted to first dataset

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$
IPLP	0.1987	4.4769	0.011	3.9019
PL	22.4127	2.3992	270.085	---
PLP	121.997	1.6083	332.485	3.1412
TLLO	3.7045	4.1343	0.1044	---
MOPL	7.5277	1.417	7.5784	18.0908

Table 2: Statistical metrics for all models according to the first dataset

Model	L^*	A^*	B^*	C^*	K^*	W^*	K^* -PV
IPLP	310.726	318.726	319.232	328.449	0.06679	0.05164	0.823436
PL	524.280	530.279	530.579	537.572	0.0717773	0.06906	0.752356
PLP	544.968	552.968	553.475	562.692	0.0713355	0.05521	0.758912
TLLO	464.11	470.11	470.41	477.403	0.132497	0.38564	0.095490
MOPL	616.978	624.978	625.484	634.701	0.0706109	0.05595	0.769575

Table 2 clearly indicates that among all the models fitted, the IPLP model exhibits the lowest values for statistical measures. Hence, it could be regarded as the best model. Figure 2 illustrates non-parametric plots for the first dataset, encompassing total time on test (TTT), box plot, and percentile-percentile (PP) plots. Furthermore, Figure 3 presents the estimated cumulative and density functions

for the fitted models.

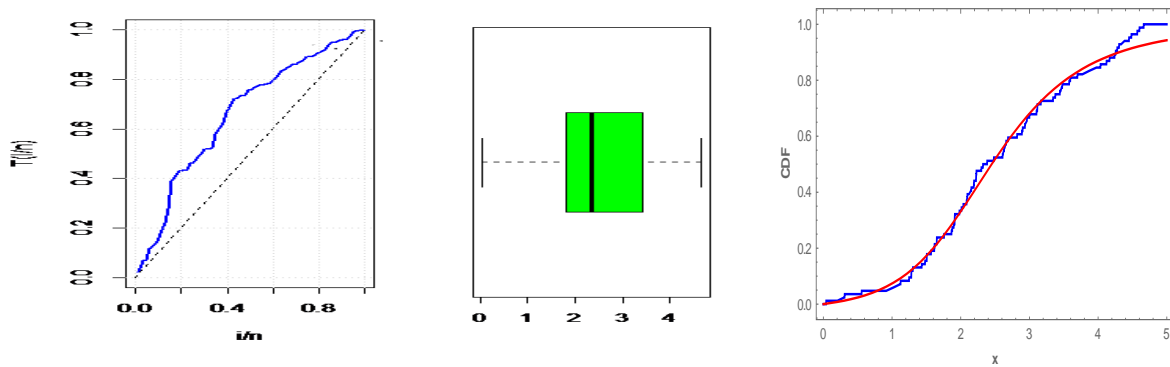


Figure 2: The TTT plot, Box plot, and PP-plot for first data

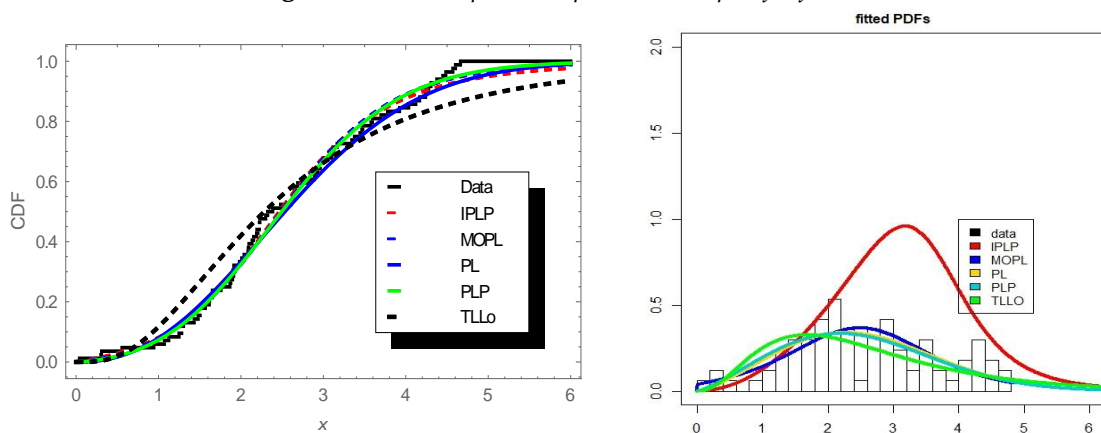


Figure 3: Estimated CDF and PDF for the models fitted to the first dataset.

Depending on Figure 3, the IPLP distribution provides the closest fit to the provided data, and then it is the best model among the other models to analyze these data.

Data 2: This dataset represents 63 aircraft windshield service times, presented by Murthy et al. [36]. The data can be shown as follows:

0.046	1.436	2.592	0.14	1.492	2.6	0.15	1.58	2.67	0.248
1.719	2.717	0.28	1.794	2.819	0.313	1.915	2.82	0.389	1.92
2.878	0.487	1.963	2.95	0.622	1.978	3.003	0.9	2.053	3.102
0.952	2.065	3.304	0.996	2.117	3.483	1.003	2.137	3.5	1.01
2.141	3.622	1.085	2.163	3.665	1.092	2.183	3.695	1.152	2.24
4.015	1.183	2.341	4.628	1.244	2.435	4.806	1.249	2.464	4.881
1.262	2.543	5.14							

Table 3 lists the MLEs for all models that are suggested, while Table 4 gives the numerical values of the statistical metrics.

Table 3: MLEs for the unknown parameters of the models fitted to the second dataset

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$
IPLP	0.2238	3.8211	0.0233	2.0577
PL	108.647	1.6327	422.985	_____
PLP	131.468	1.3335	256.861	1.8047
TLLO	1.9449	4.5615	0.0834	_____
MOPL	0.7228	3.1157	0.0161	69.0443

Table 4: Statistical metrics for the proposed models according to the second dataset

Model	L*	A*	B*	C*	K*	W*	K*-PV
IPLP	536.838	544.839	545.529	553.411	0.0684147	0.056686	0.909991
PL	633.596	639.596	640.002	646.025	0.109307	0.0945244	0.409648
PLP	545.928	553.928	554.617	562.5	0.0898046	0.0571536	0.656499
TLLo	569.62	575.62	576.026	582.049	0.145844	0.273631	0.123926
MOPL	542.74	550.74	551.429	559.312	0.137781	0.204885	0.166429

Results in Table 4 show the utility of the IPLP model as it has the lowest L*, A*, B*, C*, K*, and W* values and has the greatest K*-PV compared to the others, which indicates that the IPLP distribution is the best model. In addition, Figures 4 and 5 give TTT Plot, box plot, and PP-plot, along with the estimated cumulative and densities of the fitted models plot as well, respectively, for the data.

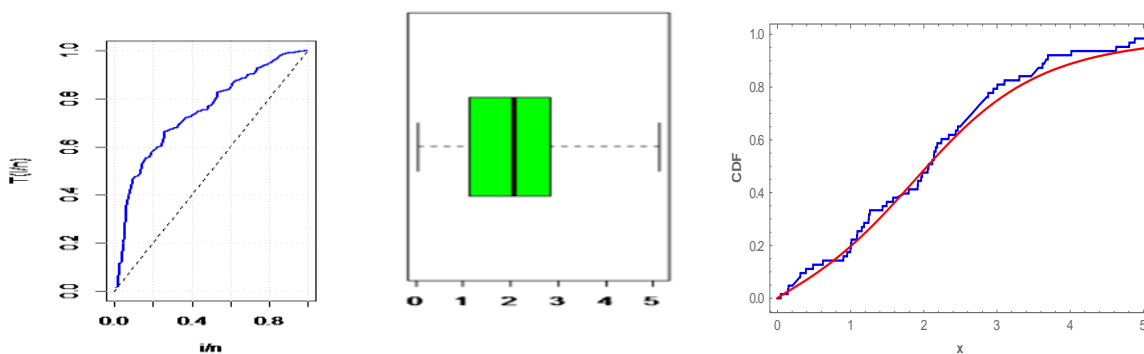


Figure 4: The TTT plot, box plot and PP-plot for second data

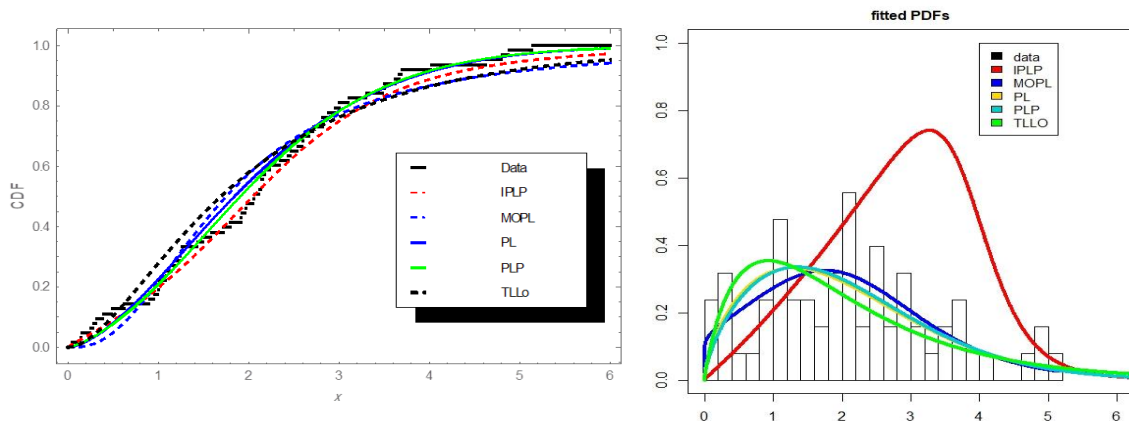


Figure 5: The estimated CDF and PDF for the models fitted to the second dataset

Figure 5 demonstrates that the IPLP distribution closely aligns with the histogram, indicating its superiority over other models for analyzing this data.

8. Concluding Remarks

A novel asymmetric four-parameter IPLPS class of distributions formed by combining the inverse power Lomax and power series distributions is introduced in this paper. This blending technique enables the creation of adaptable distributions with significant implications across diverse fields such as engineering and biology. The IPLPS class includes a new compound class and many novel

compound distributions, which come as new sub-models. Expressions for the QF, conditional moments, inverse moments, PWMs, and uncertainty measures are constructed. Estimation of model parameters is carried out using WLS, ML, CVM, and LS techniques. We assess and compare several parameter estimators for the IPLP distribution using an in-depth simulation study. Additionally, we demonstrate the efficacy of the proposed model using two real datasets, where it exhibits superior fit compared to alternative models.

Appendix A: Tables

Table A.1: Results of simulation study of different estimates for the IPLP distribution: Case 1

Case 1 ($\alpha = 0.2, \beta = 0.5, \lambda = 0.5, \theta = 0.5$)							
n	Method	Measure	α	β	λ	θ	
50	ML	AB	0.006745	0.077041	0.029852	0.115995	
		MSE	0.007617	0.041746	0.120200	0.197901	
	LS	AB	0.430544	0.487070	0.144346	0.113605	
		MSE	0.269951	0.238003	0.140357	0.146791	
	WLS	AB	0.288091	0.441809	0.096579	0.384180	
		MSE	0.149819	0.215351	0.123369	0.789747	
	CM	AB	0.339837	0.487085	0.131450	0.304000	
		MSE	0.174023	0.238043	0.114371	0.620196	
	100	ML	AB	0.000455	0.057409	0.001818	0.076642
			MSE	0.004357	0.023858	0.083599	0.185695
			SE	0.000066	0.000143	0.000289	0.000424
		LS	AB	0.447667	0.491493	0.156331	0.123996
MSE			0.281292	0.241848	0.137407	0.143057	
SE			0.000284	0.000017	0.000336	0.000357	
WLS		AB	0.279684	0.461081	0.073950	0.399205	
		MSE	0.145609	0.221363	0.119117	0.831226	
		SE	0.000259	0.000094	0.000337	0.000819	
CM		AB	0.334056	0.493331	0.132073	0.361127	
		MSE	0.162428	0.243538	0.102497	0.692470	
150		ML	AB	0.000429	0.034402	0.004477	0.047673
	MSE		0.002593	0.013023	0.065901	0.178032	
	SE		0.000051	0.000109	0.000257	0.000419	
	LS	AB	0.444868	0.494006	0.155821	0.116442	
		MSE	0.270110	0.244186	0.128611	0.130407	
	WLS	AB	0.296139	0.462616	0.085070	0.348997	
		MSE	0.152255	0.222201	0.115882	0.733505	
	CM	AB	0.343748	0.493717	0.146006	0.373102	
		MSE	0.166483	0.244084	0.097878	0.750284	
	200	ML	AB	0.001551	0.023436	0.012966	0.031401
			MSE	0.002002	0.009062	0.056420	0.170785
		LS	AB	0.452990	0.494879	0.168071	0.114122
MSE			0.275720	0.244991	0.131350	0.134244	
WLS		AB	0.300224	0.468995	0.087634	0.360930	
		MSE	0.155801	0.226458	0.117244	0.801838	
CM		AB	0.338092	0.495897	0.128463	0.359770	
		MSE	0.163050	0.245969	0.094788	0.693901	

Table A.2: Results of simulation study of different estimates for the IPLP distribution: Case 2

Case 2 ($\alpha = 0.1, \beta = 0.7, \lambda = 0.5, \theta = 0.5$)						
n	Method	Measure	α	β	λ	θ
50	ML	AB	0.002242	0.091436	0.052315	0.088154
		MSE	0.001661	0.047267	0.112424	0.190250
	LS	AB	0.514374	0.688831	0.134479	0.098571
		MSE	0.348814	0.475298	0.140524	0.136474
	WLS	AB	0.382270	0.650093	0.080623	0.348311
		MSE	0.215249	0.443329	0.119549	0.738434
	CM	AB	0.408666	0.690469	0.107712	0.382826
		MSE	0.221501	0.477186	0.104083	0.756821
100	ML	AB	0.000734	0.064783	0.027290	0.064352
		MSE	0.000925	0.032769	0.087944	0.176691
	LS	AB	0.549471	0.692815	0.168821	0.119962
		MSE	0.383591	0.480338	0.142214	0.142375
	WLS	AB	0.356645	0.660133	0.060534	0.444925
		MSE	0.191185	0.447833	0.117684	0.891369
	CM	AB	0.425101	0.694151	0.110837	0.335976
		MSE	0.234942	0.481978	0.100780	0.717249
150	ML	AB	0.001048	0.047594	0.030172	0.060538
		MSE	0.000739	0.026150	0.076517	0.168294
	LS	AB	0.548601	0.694918	0.164571	0.123721
		MSE	0.378798	0.483023	0.138253	0.138424
	WLS	AB	0.363182	0.662171	0.067058	0.387271
		MSE	0.191053	0.449580	0.114219	0.801316
	CM	AB	0.433434	0.695742	0.138198	0.388423
		MSE	0.236302	0.484116	0.098432	0.783373
200	ML	AB	0.000355	0.035961	0.024822	0.018055
		MSE	0.000581	0.021184	0.065165	0.162195
	LS	AB	0.557505	0.695863	0.168104	0.133845
		MSE	0.390652	0.484309	0.138774	0.144132
	WLS	AB	0.358781	0.670564	0.063388	0.438549
		MSE	0.190053	0.454457	0.111519	0.828456
	CM	AB	0.421589	0.696359	0.125404	0.424671
		MSE	0.228737	0.484958	0.102214	0.859722

Table A.3: Results of simulation study of different estimates for the IPLP distribution: Case 3

Case 3 ($\alpha = 0.35, \beta = 0.75, \lambda = 0.5, \theta = 0.5$)						
n	Method	Measure	α	β	λ	θ
50	ML	AB	0.028882	0.038793	0.041519	0.097016
		MSE	0.021385	0.024007	0.093766	0.205833
	LS	AB	0.330981	0.720645	0.172635	0.214191
		MSE	0.200888	0.522177	0.156952	0.153552
	WLS	AB	0.171788	0.664002	0.111915	0.315342
		MSE	0.094188	0.468671	0.125394	0.676175
	CM	AB	0.200807	0.723881	0.139624	0.297306
		MSE	0.094502	0.526193	0.108873	0.592616
100	ML	AB	0.011510	0.032625	0.034154	0.045939
		MSE	0.010988	0.017202	0.073926	0.191679
	LS	AB	0.343760	0.730998	0.176993	0.239639
		MSE	0.204578	0.535951	0.150556	0.161608
	WLS	AB	0.183232	0.681750	0.121003	0.315958
		MSE	0.096159	0.485995	0.114875	0.638105
	CM	AB	0.196654	0.733830	0.131675	0.316076
		MSE	0.088577	0.539333	0.100398	0.589016
150	ML	AB	0.009429	0.021253	0.021613	0.045372
		MSE	0.008488	0.012985	0.061332	0.185365
	LS	AB	0.347205	0.736845	0.183141	0.228068
		MSE	0.198331	0.543845	0.142716	0.156601
	WLS	AB	0.158364	0.678758	0.094665	0.387196
		MSE	0.091379	0.481589	0.118032	0.745790
	CM	AB	0.197847	0.737241	0.133215	0.322106
		MSE	0.088204	0.544097	0.096084	0.576261
200	ML	AB	0.005265	0.018039	0.032937	0.005198
		MSE	0.005849	0.010298	0.052733	0.183501
	LS	AB	0.330284	0.738295	0.161452	0.216034
		MSE	0.183964	0.545553	0.130130	0.155453
	WLS	AB	0.181529	0.692042	0.112438	0.315812
		MSE	0.096914	0.492804	0.114267	0.617612
	CM	AB	0.185704	0.739790	0.125346	0.377463
		MSE	0.083501	0.541593	0.094968	0.691230

Table A.4: Results of simulation study of different estimates for the IPLP distribution: Case 4

Case 4 ($\alpha = 0.7, \beta = 0.25, \lambda = 0.5, \theta = 0.5$)						
n	Method	Measure	α	β	λ	θ
50	ML	AB	0.002412	0.026550	0.017636	0.062220
		MSE	0.056895	0.005150	0.097176	0.176375
	LS	AB	0.004458	0.236684	0.175355	0.196397
		MSE	0.090972	0.056577	0.154987	0.154997
	WLS	AB	0.118599	0.213013	0.132290	0.293232
		MSE	0.095884	0.049828	0.129123	0.667014
	CM	AB	0.100852	0.237882	0.168351	0.251028
		MSE	0.074475	0.057127	0.116832	0.562473
100	ML	AB	0.011686	0.012692	0.031510	0.009029
		MSE	0.396155	0.002067	0.076041	0.170751
	LS	AB	0.019738	0.241765	0.204124	0.214818
		MSE	0.087939	0.058649	0.158975	0.161972
	WLS	AB	0.115818	0.222881	0.142248	0.273083
		MSE	0.085862	0.051960	0.121264	0.574088
	CM	AB	0.131302	0.241384	0.149405	0.309364
		MSE	0.070562	0.058503	0.105365	0.588599
150	ML	AB	0.019717	0.005788	0.035701	0.010170
		MSE	0.030271	0.001102	0.058667	0.153535
	LS	AB	0.012704	0.244189	0.198267	0.207426
		MSE	0.082305	0.059729	0.150544	0.160624
	WLS	AB	0.114129	0.224038	0.144723	0.293595
		MSE	0.086913	0.052262	0.124999	0.661949
	CM	AB	0.133418	0.243985	0.142315	0.290336
		MSE	0.070266	0.059684	0.103061	0.557025
200	ML	AB	0.012894	0.005315	0.039573	0.030536
		MSE	0.025921	0.000908	0.053184	0.153984
	LS	AB	0.027965	0.245094	0.218293	0.217001
		MSE	0.078031	0.060166	0.152014	0.158960
	WLS	AB	0.103315	0.226621	0.161816	0.278613
		MSE	0.085451	0.052928	0.130257	0.629098
	CM	AB	0.117752	0.244956	0.169220	0.286765
		MSE	0.064152	0.060075	0.104847	0.525627

References

[1] Barreto-Souza, W., de Morais, A. L., and Cordeiro, G. M. (2011). The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, 81(5):645–657.

[2] Noack, A. (1950). A class of random variables with discrete distributions. *The Annals of Mathematical Statistics*, 21(1):127–132.

[3] Chahkandi, M., and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate, *Computational Statistics and Data Analysis*, 53:4433–4440

[4] Morais, A. L., and Barreto-Souza, W., (2011). A compound class of Weibull and power series distributions. *Computational Statistics and Data Analysis*, 55:1410–1425.

- [5] Warahena-Liyanage, G., and Pararai, M. (2015). The Lindley power series class of distributions: model, properties and applications. *Journal of Computations & Modelling*, 5(3): 35–80.
- [6] Elbatal, I., Zayed, M., Rasekhi, M., and Butt, N. (2017). The exponential Pareto power series distribution: *Theory and applications*. *Pakistan Journal of Statistics and Operation Research*, 13(3): 603–615, <https://doi.org/10.18187/pjsor.v13i3.2072>
- [7] Silva, R. B. and Cordeiro, G. M. (2015). The Burr XII power series distributions: A new compounding family. *Brazilian Journal of Probability and Statistics*, 29(3): 565–589
- [8] Alizadeh, M., Bagheri, S. F., Samani, E. B., Ghobadi, S., and Nadarajah, S. (2018). Exponentiated power Lindley power series class of distributions: Theory and applications. *Communications in Statistics-Simulation and Computation*, 47(9):2499–2531. <https://doi.org/10.1080/03610918.2017.1350270>
- [9] Elbatal, I., Altun, E., Afify, A. Z., and Ozel, G. (2018). The generalized Burr XII power series distributions with properties and applications. *Annals of Data Science*, 6:571–597 <https://doi.org/10.1007/s40745-018-0171-2>.
- [10] Goldoust, M., Rezaei, S., Alizadeh, M., Nadarajah, S. (2019). The odd log-logistic power series family of distributions: Properties and applications. *Statistica*, 79(1):77–107. <https://doi.org/10.6092/issn.1973-2201/8115>
- [11] Kunjiratanachot, N., Bodhisuwan, W. and Volodin, A. (2018). The Topp-Leone generalized exponential power series distribution with applications. *Journal of Probability and Statistical Science*, 16(2): 197–208.
- [12] Hassan, A. S., and Assar, S. M. (2021). A new class of power function distribution: Properties and applications. *Annals of Data Science*, 8: 205–225, <https://doi.org/10.1007/s40745-019-00195-7>
- [13] Rivera, P. A., Calderín-Ojeda, E., Gallardo, D. I., and Gómez, H. W. (2021). A compound class of the inverse gamma and power series distributions. *Symmetry*, 13, 1328.
- [14] Hassan, A. S., Almetwally, E. M., Gamoura, S. C, Metwally, A. S. M. (2022). Inverse exponentiated Lomax power series distribution: Model, estimation, and application. *Journal of Mathematics*, 2022, 1998653, <https://doi.org/10.1155/2022/1998653>
- [15] Khojastehbakht, N., Ghatari, A., and Samani, E. B. (2023). The beta exponential power series distribution. *Annals of Data Science*, 10(5):1157–1178, <https://doi.org/10.1007/s40745-022-00414-8>
- [16] Alghamdi, S. M., Shrahili, M., Hassan, A. S., Mohamed, R.E., Elbatal, I., and Elgarhy, M. (2023). Analysis of milk Production and failure data: Using unit exponentiated half logistic power series class of distributions. *Symmetry*, 15, 714, <https://doi.org/10.3390/sym15030714>
- [17] Ul-Haq, M, A., Shahzad, M, K., and Tariq, S. (2023). The inverted Nadarajah–Haghighi power series distributions. *International Journal of Applied and Computational Mathematics*, 10(11), <https://doi.org/10.1007/s40819-023-01551-1>.
- [18] Hassan, A. S., and Abd-Allah, M., (2023). Power quasi-Lindley power series class of distributions: Theory and applications. *Thailand Statistician*, 21(2): 314–336
- [19] Zayed, M. A., Hassan, A. S., Almetwally, E. M., Aboalkhair, A. M., Al-Nefae, A. H., and Almonry, H. M. (2023). A compound class of unit Burr XII model: Theory, estimation, fuzzy, and application. *Scientific Programming*, <https://doi.org/10.1155/2023/4509889>
- [20] Yousef, M. M., Hassan, A. S., and Almetwally, E. M. (2024). Statistical inference for the unit Gompertz power series distribution using ranked set sampling with applications. *Assiut University Journal of Multidisciplinary Scientific Research*, 53(1):154–189.
- [21] Oluyede, B., Dingalo, N. and Chipepa, F. (2024). A new and generalized class of log-logistic modified Weibull power series distributions with applications. *Thailand Statistician* 22(2), 237–273

- [22] El-Saeed, A.R., Hassan, A.S., Elharoun, N.M., Al Mutairi, A., Khashab, R.H., Nassr, S.G. (2023). A class of power inverted Topp-Leone distribution: Properties, different estimation methods & applications. *Journal of Radiation Research and Applied Sciences*, 16 (2023) 100643, <https://doi.org/10.1016/j.jrras.2023.100643>
- [23] Hassan, A. S., and Abd-Allah, M. (2019). On the inverse power Lomax distribution. *Annals of Data Science*, 6: 259–278, <https://doi.org/10.1007/s40745-018-0183-y>
- [24] Shi, X., and Shi, Y. (2021). Inference for inverse power Lomax distribution with progressive first-failure censoring. *Entropy*, 23, 1099. <https://doi.org/10.3390/e23091099>
- [25] Nassr, S. G., Hassan, A. S., Almetwally, E. M., Al Mutairi, A., Khashab, R. H. and ElHaroun, N. M. (2023). Statistical inference of the inverted exponentiated Lomax distribution using generalized order statistics with application to COVID-19. *AIP Advances*, 13, 105118; <https://doi.org/10.1063/5.0174540>.
- [26] Greenwood, J. A., Landwehr, J. M., Wallis, J. R., and Matala, N. C. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15:1094–1054
- [27] Gradshteyn, I. S.; and Ryzhik, I. M. Table of Integrals, Series and Products; Academic Press: San Diego, CA, USA, 2000.
- [28] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52(1):479–487
- [29] Buzaridah, M., Ramadan, D. A, and El-Desouky, B. (2021). Flexible reduced logarithmic-inverse Lomax distribution with application for bladder cancer. *Open Journal of Modelling and Simulation*. 09(04):351–369. <https://doi.org/10.4236/ojmsi.2021.94023>.
- [30] Johnson, N. L., Kemp, A. W., and Kotz, S. Univariate Discrete Distributions. Wiley-Interscience, third edition, 2005.
- [31] MacDonald, P. D. M. (1971). Comment on “an estimation procedure for mixtures of distributions” by Choi and Bulgren. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(2): 326–329.
- [32] Rady, E. A., Hassanein, W. A., and Elhaddad, T. A. (2016). The power Lomax distribution with an application to bladder cancer data, *Springer Plus*, <https://www.ncbi.nlm.nih.gov/pubmed/27818876>.
- [33] Hassan, A. S., and Nassr, S. (2018). Power Lomax Poisson distribution: Properties and estimation. *Journal of Data Science*, 18, 105–128. [https://doi.org/10.6339/JDS.201801_16\(1\).0007](https://doi.org/10.6339/JDS.201801_16(1).0007)
- [34] Oguntunde, P., Khaleel, M., Okagbue, H., and Odetunmbi, O. (2019). The Topp–Leone Lomax (TLLo) distribution with applications to airborne communication transceiver dataset. *Wireless Personal Communications*. 109. <https://doi.org/10.1007/s11277-019-06568-8>.
- [35] Ul-Haq, M. A., Hamedani, G. G., Elgarhy, M., and Ramos, P. L. (2020). Marshall-Olkin power Lomax distribution: Properties and estimation based on complete and censored samples. *International Journal of Statistics and Probability, Canadian Center of Science and Education*, 9(1): 1–48.
- [36] Murthy, D. N. P., Xie, M. and Jiang, R. Weibull Models. Wiley, 2004.

EXPLORING LENGTH BIASED QUASI SUJA DISTRIBUTION: PROPERTIES AND APPLICATIONS

Vidya Yerneni¹ and Aafaq A. Rather^{2*}

•

^{1,2}Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India

¹Email: vidyay@yahoo.com, ^{2*}Email: aafaq7741@gmail.com

Abstract

This paper introduces a new statistical distribution called length biased quasi suja distribution (LBQS). It explores its properties, including moments, moment generating function(MGF), characteristic function(CF), harmonic mean, reliability, hazard rate and reverse hazard rate. Order statistics of the above distribution is obtained. Furthermore, the paper also examines various entropy which measures the randomness of system, like Renyi entropy and Tsalli's entropy. It also evaluates Bonferroni and Lorenz curves which are useful in measuring the inequality. It also discusses parameter estimation techniques specifically maximum likelihood estimation and likelihood ratio testing. Moreover, a simulation study has been conducted to demonstrate how well the distribution would perform in real-life situation. The validity of the distribution is also demonstrated with real-world data example of failure data, highlighting its potential for practical applications in data analysis.

Key words: Length biased quasi suja distribution, Moments, Entropy, Estimation, Simulation

1. Introduction

The weighted distributions are applied in various research areas related to biomedicine, reliability, ecology and branching processes. A number of continuous distributions are used like Weibull, lindley, exponential, lognormal and gamma for modelling this type of data. If x is the original observation with its pdf $f(x)$, then in case of any bias in sampling appropriate weighted function, say $w(x)$ which is a function of random variable will be introduced to model the situation. This concept of weighted distributions was given by Fisher [7] to model the ascertainment bias. Later Rao [13] developed this concept in a unified manner while modelling the statistical data when the standard distributions were not appropriate to record these observations with equal probabilities. As a result, weighted models were formulated in such situations to record the observations according to some weighted function. The weighted distribution reduces to length biased distribution when the weight function considers only the length of the units. The concept of length biased sampling was first introduced by Cox [5]. Weighted distributions are applied in various research areas related to reliability, biomedicine, ecology and branching processes. Dey et al [6] discussed weighted exponential distribution with its properties and different methods of estimation. Kilany [12] have obtained the weighted version of lomax distribution. Ahmad et al [1] have obtained the length biased weighted version of lomax distribution with properties and applications. Khan et al. [11] discussed the weighted modified weibull distribution. Rather and Subramanian [17] discussed the characterization and estimation of length biased weighted generalized uniform distribution. Recently Rather and Subramanian [19] also discussed on weighted

Sushila distribution with properties and applications. Ganaie R.A et. al [9] studied about how Uma distribution is applicable in engineering sciences. Saraja V D et al. [20] explored the length biased Tornumonkpe distribution, their properties, estimations and practical applications. Rashid Ganaie A et.al [16] formulated exponentiated ADYA distribution and studied their properties and applications. A new generalization of Akshaya distribution with applications in engineering science was studied by Rather and Subramanian [18]. Rather and Ozel [14] modelled Weighted Power Lindley distribution and its application on Life time data.

Recently, Shanker [21] proposed a one parameter distribution suja distribution and studied its statistical properties, estimation of parameter using method of moment and method of maximum likelihood and applications to some real lifetime data and observed that Suja distribution gives much closer fit than several one parameter lifetime distributions. Recently, Al Omari and Alsmairan [2] obtained length-biased Suja distribution and studied its statistical properties and applications. Al-Omari et al [3] proposed power length-biased suja distribution and discussed its properties and applications. Alsmairan I. K [4] derived weighted suja distribution and discussed its statistical properties and applications to ball bearings data in safety engineering. Todoka et al [22] have studied on the cdf of various modifications of suja distribution and discussed their applications in the field of analysis of computer- virus propagation and debugging theory.

In this paper, we proposed length biased quasi suja distribution. The quasi suja distribution, introduced by Shanker et al in [15], is a recently obtained two-parameter model for extreme right skewed data which contains suja distribution as particular case designed for various applications in engineering and medical sciences. Validity and significance of proposed model in modelling lifetime data was better than quasi suja distribution, exponential distribution, Lindley and erlang truncated exponential distribution.

The paper is classified into following sections: Section 2 defines the proposed length biased quasi suja distribution and reliability. Some structural properties are discussed in Section 3. The likelihood ratio test is given in Section 4. Then, Renyi and Tsalli’s entropy measures of the LBQS distribution are obtained in Section 5. Section 6 describes the method of obtaining order statistics. Income distribution curve and estimation of parameters is discussed in section 7 and 8 respectively. Simulation study is shown in section 9. Finally, fitted the distribution to the real-life data and found to be fitting good compared to various other models.

2. Length biased quasi suja distribution (LBQS)

2.1 Density and cumulative density functions

The probability density function (pdf) and cumulative distribution function (cdf) of quasi suja distribution with parameters α and θ is defined by

$$f(x; \theta, \alpha) = \frac{\theta^4}{\alpha\theta^3 + 24} (\alpha + \theta x^4) e^{-\theta x} \quad x > 0, \theta > 0, \alpha > 0 \quad (1)$$

$$F(x; \theta, \alpha) = 1 - \left[1 + \frac{\theta^4 x^4 + 4\theta^3 x^3 + 12\theta^2 x^2 + 24\theta x}{\alpha\theta^3 + 24} \right] e^{-\theta x} \quad x > 0, \theta > 0, \alpha > 0$$

Suppose X is a non-negative random variable with pdf $f(x)$. Let $w(x)$ be the non-negative weight function, then the pdf of the weighted random variable X_w is given by

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))}, \quad x > 0$$

where $w(x)$ is a non-negative weight function and $E(w(x)) = \int w(x)f(x)dx$

For different weighted models, we have different choices of the weight function $w(x)$. when $w(x) = x^c$, the resulting distribution is termed as weighted distribution. In this paper, we are finding the length biased version of quasi suja distribution, so we will take $c = 1$ in weights x^c , in order to get the length biased quasi suja distribution and its probability density function (pdf) is given by:

$$f_1(x) = \frac{xf(x)}{E(x)} \tag{2}$$

using (1) we get

$$\begin{aligned} E(x) &= \int_0^{\infty} xf(x) dx \\ &= \frac{\alpha\theta^3 + 120}{\theta(\alpha\theta^3 + 24)} \end{aligned} \tag{3}$$

substituting equations (1) and (3) in (2) we obtain the density function of length biased quasi suja distribution as follows

$$f_1(x; \theta, \alpha) = \frac{\theta^5 x(\alpha + \theta x^4)e^{-\theta x}}{\alpha\theta^3 + 120} \quad x > 0, \theta > 0 \tag{4}$$

and the cumulative distribution function (cdf) of LBQS distribution is obtained by

$$\begin{aligned} F_1(x) &= \int_0^x f_1(x; \alpha, \theta) dx \\ &= \int_0^x \frac{\theta^5 x(\alpha + \theta x^4)e^{-\theta x}}{\alpha\theta^3 + 120} dx \end{aligned}$$

on simplification, the cdf of LBQS distribution is given by

$$F_1(x; \theta, \alpha) = \frac{\alpha\theta^3\gamma(2, \theta x) + \gamma(6, \theta x)}{\alpha\theta^3 + 120} \tag{5}$$

Graphs for the pdf and cdf of the LBQS distribution for several values of parameters are showed in Fig. 1 and Fig. 2.

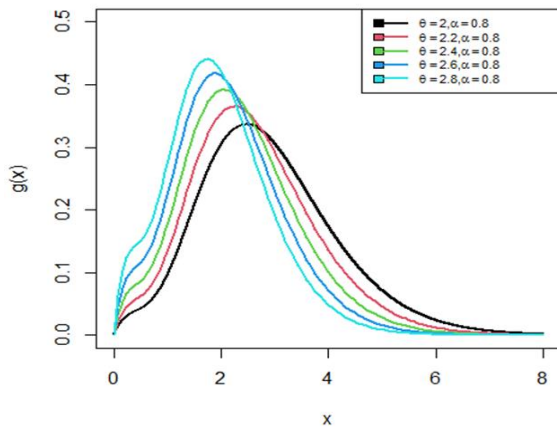


Fig.1: pdf plot of LBQS distribution

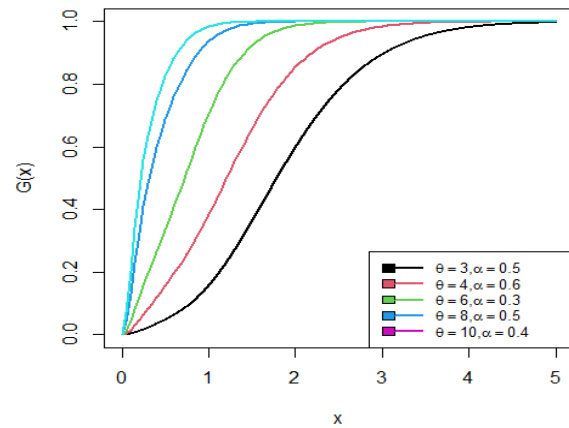


Fig.2: cdf plot of LBQS distribution

2.2 Survival, Hazard and Reversed hazard functions

In this section, we discuss about the survival function, hazard and reverse hazard functions of the LBQS distribution.

The survival function or the reliability function of the LBQS distribution is given by

$$S(x) = 1 - F_1(x)$$

$$S(x) = 1 - \left(\frac{\alpha\theta^3\gamma(2, \theta x) + \gamma(6, \theta x)}{\alpha\theta^3 + 120} \right)$$

$$= \frac{\alpha\theta^3(1 - \gamma(2, \theta x)) + (120 - \gamma(6, \theta x))}{\alpha\theta^3 + 120}$$

The hazard function is also known as the hazard rate function, instantaneous failure rate or force of mortality and is given for LBQS distribution as

$$h(x) = \frac{f_1(x)}{S(x)}$$

$$= \frac{x\theta^5(\alpha + \theta x^4)e^{-\theta x}}{\alpha\theta^3(1 - \gamma(2, \theta x)) + (120 - \gamma(6, \theta x))}$$

The reverse hazard function of the LBQS distribution is given by

$$h_{rl}(x) = \frac{f_1(x)}{F_1(x)}$$

$$= \frac{x\theta^5(\alpha + \theta x^4)e^{-\theta x}}{\alpha\theta^3\gamma(2, \theta x) + \gamma(6, \theta x)}$$

Fig. 3 and Fig. 4 represent graphs for the Survival function and Hazard rate function respectively of the LBQS distribution for several values of parameters.

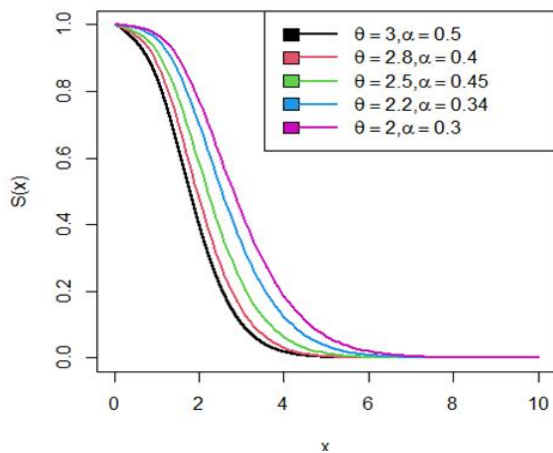


Fig.3: Survival plot of LBQS distribution

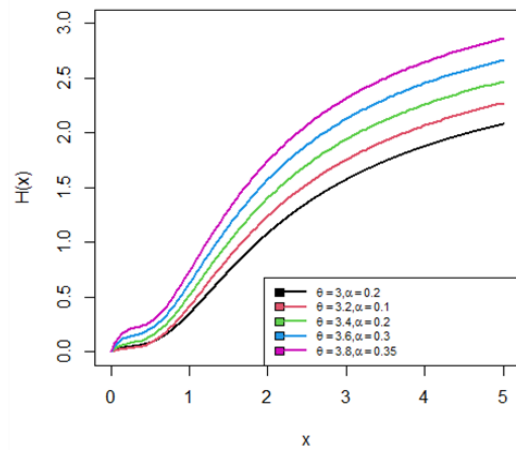


Fig.4: Hazard rate plot of LBQS distribution

3. Structural Properties

In this section, we investigate various structural properties of the LBQS distribution. Let X denotes the random variable of LBQS distribution with parameters α, θ and then its r^{th} order moment $E(X^r)$ about origin is given by

$$E(X^r) = \mu_r' = \int_0^{\infty} x^r f_1(x) dx$$

$$= \int_0^{\infty} \frac{\theta^5 x^{r+1} (\alpha + \theta x^4) e^{-\theta x}}{\alpha \theta^3 + 120} dx$$

After simplifying the expression, we get

$$E(X^r) = \frac{\alpha \theta^3 \Gamma(2+r) + \Gamma(6+r)}{\theta^r (\alpha \theta^3 + 120)} \tag{6}$$

Putting $r=1$, we get the expected value of LBQS distribution as follows

$$E(X) = \frac{2\alpha\theta^3 + 720}{\theta(\alpha\theta^3 + 120)}$$

and putting $r=2$, we get the second moment as follows

$$E(X^2) = \frac{6\alpha\theta^3 + 5040}{\theta^2(\alpha\theta^3 + 120)}$$

therefore, the variance of LBQS distribution is given by

$$V(X) = \frac{6\alpha\theta^3 + 5040}{\theta^2(\alpha\theta^3 + 120)} - \left(\frac{2\alpha\theta^3 + 720}{\theta(\alpha\theta^3 + 120)} \right)^2$$

3.1 Harmonic mean

The harmonic mean of LBQS distributed random variable X can be written as

$$\begin{aligned}
 H &= E\left(\frac{1}{X}\right) = \int_0^{\infty} \frac{1}{x} f_1(x) dx \\
 &= \int_0^{\infty} \frac{\theta^5 (\alpha + \theta x^4) e^{-\theta x}}{\alpha \theta^3 + 120} dx
 \end{aligned}$$

after simplifying the expression, we get

$$H = \frac{\alpha \theta^3 + 24}{\theta(\alpha \theta^3 + 120)}$$

3.2 Moment generating function (MGF) and Characteristic function

Let X have a LBQS distribution, then the MGF of X is obtained as

$$M_X(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} f_1(x) dx$$

According to Taylor's series, we obtain

$$\begin{aligned}
 M_X(t) &= E(e^{tx}) = \int_0^{\infty} \left(1 + tx + \frac{(tx)^2}{2!} + \dots \right) f_1(x) dx \\
 &= \sum_{j=0}^{\infty} \frac{t^j}{j!} E(X^j) \\
 &= \int_0^{\infty} \sum_{j=0}^{\infty} \frac{t^j}{j!} x^j f_1(x) dx \\
 &= \sum_{j=0}^{\infty} \frac{t^j}{j!} \left(\frac{\alpha \theta^3 \Gamma(j+2) + \Gamma(j+6)}{\theta^j (\alpha \theta^3 + 120)} \right)
 \end{aligned}$$

Similarly, we obtained characteristic function of the LBQS distribution as follows

$$\phi_X(t) = M_X(it) = \sum_{j=0}^{\infty} \frac{(it)^j}{j!} \left(\frac{\alpha \theta^3 \Gamma(j+2) + \Gamma(j+6)}{\theta^j (\alpha \theta^3 + 120)} \right)$$

4. Likelihood Ratio Test

Suppose X be a random sample from the LBQS distribution. We use the following hypothesis

$$H_0: f(x) = f(x; \alpha, \theta) \text{ against } H_1: f(x) = f_1(x; \alpha, \theta)$$

to test whether the random sample of size n comes from the quasi suja (QS) distribution or the LBQS distribution. The test statistic used is

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^n \frac{f_1(x; \alpha, \theta)}{f(x; \alpha, \theta)}$$

$$= \left(\frac{(\alpha\theta^3 + 24)\theta}{\alpha\theta^3 + 120} \right)^n \prod_{i=1}^n x_i$$

We reject null hypothesis, if

$$\Delta > k$$

$$\left(\frac{(\alpha\theta^3 + 24)\theta}{\alpha\theta^3 + 120} \right)^n \prod_{i=1}^n x_i > k$$

$$\prod_{i=1}^n x_i > k \left(\frac{\alpha\theta^3 + 120}{(\alpha\theta^3 + 24)\theta} \right)^n$$

$$\text{or } \Delta^* = \prod_{i=1}^n x_i > k^* \text{ where } k^* = k \left(\frac{\alpha\theta^3 + 120}{(\alpha\theta^3 + 24)\theta} \right)^n$$

5. Entropy Measures

The idea of entropy is important in various areas such as probability and statistics, physics, communication theory and economics. Entropy measures quantify the diversity, uncertainty, or randomness of a system. Entropy of a random variable X is a measure of variation of the uncertainty.

5.1 Renyi Entropy

The Renyi entropy is important in ecology and statistics as index of diversity. It was proposed by Renyi. The Renyi entropy of order β for a random variable X is given by

$$e(\beta) = \frac{1}{1-\beta} \log \left(\int_0^{\infty} f^\beta(x) dx \right)$$

Where $\beta > 0$ and $\beta \neq 1$. So, we get

$$e_1(\beta) = \frac{1}{1-\beta} \log \int_0^{\infty} \left(\frac{\theta^5 x (\alpha + \theta x^4) e^{-\theta x}}{\alpha\theta^3 + 120} \right)^\beta dx$$

$$= \frac{1}{1-\beta} \log \left(\left(\frac{\theta^5}{\alpha\theta^3 + 120} \right)^\beta \int_0^{\infty} \left(x e^{-\theta x} (\alpha + \theta x^4) \right)^\beta dx \right)$$

After simplification we get

$$= \frac{1}{1-\beta} \log \left(\left(\frac{\theta^5}{\alpha\theta^3 + 120} \right)^\beta \sum_{j=0}^{\infty} \beta C_j \alpha^{\beta-j\theta} \frac{\Gamma(\beta + 4j + 1)}{(\theta\beta)^{\beta+4j+1}} \right)$$

5.2 Tsalli's Entropy

$$\begin{aligned}
 S_\lambda &= \frac{1}{\lambda-1} \left(1 - \int_0^\infty f^\lambda(x) dx \right) \\
 &= \frac{1}{\lambda-1} \left(1 - \int_0^\infty \left(\frac{\theta^5 x (\alpha + \theta x^4) e^{-\theta x}}{\alpha \theta^3 + 120} \right)^\lambda dx \right) \\
 &= \frac{1}{\lambda-1} \left(1 - \left(\frac{\theta^5}{\alpha \theta^3 + 120} \right)^\lambda \sum_{j=0}^\infty {}^\lambda C_j \alpha^{\lambda-j} \theta^j \frac{\Gamma(\lambda + 4j + 1)}{(\theta \lambda)^{\lambda+4j+1}} \right)
 \end{aligned}$$

6. Order Statistics

Consider $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the order statistics of a random sample X_1, X_2, \dots, X_n drawn from the continuous population with pdf $f(x)$ and cdf $F_X(x)$, then the pdf of r^{th} order statistic $X_{(r)}$ is given by

$$f_{X_{(r)}}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x) [F_X(x)]^{r-1} [1-F_X(x)]^{n-r} \quad (7)$$

Using Equations (4) and (5) in Equation (7), the pdf of r^{th} order statistic $X_{(r)}$ of the LBQS distribution is given by

$$\begin{aligned}
 f_{X_{(r)}}(x) &= \frac{n!}{(r-1)!(n-r)!} \left(\frac{\theta^5 x (\alpha + \theta x^4) e^{-\theta x}}{\alpha \theta^3 + 120} \right) \times \left[\frac{\alpha \theta^3 \gamma(2, \theta x) + \gamma(6, \theta x)}{\alpha \theta^3 + 120} \right]^{r-1} \\
 &\quad \times \left[1 - \frac{\alpha \theta^3 \gamma(2, \theta x) + \gamma(6, \theta x)}{\alpha \theta^3 + 120} \right]^{n-r}
 \end{aligned}$$

pdf of higher order statistics is given by

$$f_{X_{(n)}}(x) = n \left(\frac{\theta^5 x (\alpha + \theta x^4) e^{-\theta x}}{\alpha \theta^3 + 120} \right) \times \left[\frac{\alpha \theta^3 \gamma(2, \theta x) + \gamma(6, \theta x)}{\alpha \theta^3 + 120} \right]^{n-1}$$

pdf of first order statistics is given by

$$f_{X_{(1)}}(x) = n \left(\frac{\theta^5 x (\alpha + \theta x^4) e^{-\theta x}}{\alpha \theta^3 + 120} \right) \times \left[1 - \frac{\alpha \theta^3 \gamma(2, \theta x) + \gamma(6, \theta x)}{\alpha \theta^3 + 120} \right]^{n-1}$$

7. Income Distribution Curve

The bonferroni and the Lorenz curves are not only used in economics in order to study the income and poverty, but it is also being used in other fields like reliability, medicine, insurance and demography. The bonferroni and lorenz curves are given by

$$B(p) = \frac{1}{p\mu_1} \int_0^q x f(x) dx$$

and

$$L(p) = pB(p) = \frac{1}{\mu_1'} \int_0^q xf(x)dx$$

From (6)

$$\mu_1' = \frac{2\alpha\theta^3 + 720}{\theta(\alpha\theta^3 + 120)}$$

and $q = F^{-1}(p)$. Then, we have

$$B(p) = \frac{1}{p\mu_1'} \int_0^q \frac{\theta^5 x^2 (\alpha + \theta x^4) e^{-\theta x}}{\alpha\theta^3 + 120} dx$$

after simplification, we get

$$B(p) = \frac{\alpha\theta^3\gamma(3, \theta q) + \gamma(7, \theta q)}{p(2\alpha\theta^3 + 720)}$$

similarly, lorenz curve is obtained as

$$L(p) = \frac{\alpha\theta^3\gamma(3, \theta q) + \gamma(7, \theta q)}{(2\alpha\theta^3 + 720)}$$

8. Estimation

In this section, we will discuss the maximum likelihood estimators (MLE's) of the parameters of the LBQS distribution. Consider X_1, X_2, \dots, X_n be the random sample of size n from the LBQS distribution, then the likelihood function is given by

$$L(x; \alpha, \theta) = \left(\frac{\theta^5}{\alpha\theta^3 + 120} \right)^n \prod_{i=1}^n \left(x_i (\alpha + \theta x_i^4) e^{-\theta x_i} \right)$$

The log likelihood is obtained as

$$\begin{aligned} \log L &= 5n \log \theta - n \log(\alpha\theta^3 + 120) \\ &+ \sum_{i=1}^n \log x_i + \sum_{i=1}^n \log(\alpha + \theta x_i^4) - n\theta \sum_{i=1}^n x_i \end{aligned} \tag{8}$$

Differentiating (8) w.r.t θ and α we get the following likelihood equations

$$\frac{\partial \log L}{\partial \theta} = \frac{5n}{\theta} - \frac{n}{\alpha\theta^3 + 120} (3\alpha\theta^2) + \sum_{i=1}^n \frac{x_i^4}{\alpha + \theta x_i^4} - n \sum_{i=1}^n x_i = 0$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{-n\theta^3}{\alpha\theta^3 + 120} + \sum_{i=1}^n \frac{1}{\alpha + \theta x_i^4} = 0$$

Because of the complicated form of the likelihood equations, algebraically it is very difficult to solve the system of nonlinear equations. Therefore, we use R-software for estimating the required parameters.

9. Simulation Study

In this section, a simulation study is conducted to examine the performance of maximum likelihood estimators of the LBQS distributions using R-software. We generated random number for different sample sizes and different parameter values. We examined the mean estimates, biases, mean square errors and variances of the MLE's. The simulation results for different parameter values of LBQS distribution is presented in table 1. It reveals from the table that as the sample size increases, biases, MSEs and variances of the MLE's of the parameters become smaller respectively. Fig. 5, Fig. 6, Fig.7, Fig. 8, Fig.9, Fig.10, Fig.11 and Fig. 12 shows histogram graphs of simulated data.

Table 1: The biases variances and MSEs of LBQS distribution for different parameter values

n	$\theta=2$			$\alpha=8$		
	Bias	Variance	MSE	Bias	Variance	MSE
20	-0.1663918	0.07275417	0.1004404	9.1904	113.4698	197.9333
30	-0.02764637	0.09921434	0.09997866	4.44103	79.39125	99.114
50	-0.07826967	0.04104919	0.04717534	5.231117	48.07655	75.44114
80	-	0.0192557	0.01925604	1.946187	26.27957	30.06722
	0.0005882063					
100	-0.06645069	0.0126943	0.01710999	2.637763	18.40008	25.35787
200	0.0317005	0.00394192	0.004946842	-0.8056873	7.551154	8.200286
300	-0.003812567	0.004438482	0.004453018	0.7725432	4.295084	4.891907
	$\theta=3$			$\alpha=0.8$		
	Bias	Variance	MSE	Bias	Variance	MSE
20	0.1077014	0.1079198	0.1195194	0.10619	0.7403965	0.7516728
30	0.05250122	0.08043579	0.08319217	-0.095104	0.3896097	0.3896097
50	0.01232364	0.07130409	0.07145596	-0.1934582	0.256577	0.2940031
80	0.07141307	0.03529702	0.04039685	-0.1773731	0.2315595	0.2630207
100	0.03355058	0.01024581	0.01137146	0.01653806	0.05744206	0.05771557
200	-0.02356278	0.01017962	0.01073482	0.009685886	0.05365214	0.05374595
300	0.01235768	0.00395297	0.004105682	-0.03921203	0.0146893	0.01622688
	$\theta=6$			$\alpha=0.3$		
	Bias	Variance	MSE	Bias	Variance	MSE
20	-0.2409736	0.8117087	0.869777	0.4400109	0.6451856	0.8387952
30	0.2179345	0.4019304	0.4494258	-0.03356909	0.04080886	0.04193574
50	0.2340156	0.3274527	0.382216	0.00622955	0.03390483	0.03394364
80	0.15726	0.08666046	0.1113912	-0.02508142	0.01842258	0.01905166
100	-0.1231709	0.06968093	0.08485199	0.02899184	0.01667758	0.01751811
200	0.07676134	0.06580578	0.07169809	0.0002268879	0.007942669	0.007942721
300	0.023855	0.0234289	0.02399796	0.009857188	0.005338747	0.005435911
	$\theta=8$			$\alpha=0.5$		
	Bias	Variance	MSE	Bias	Variance	MSE
20	1.798051	6.027411	9.260398	0.01659618	0.3433603	0.3436357
30	0.5363827	2.157112	2.444818	0.04257179	0.2365076	0.23832

50	0.7529427	1.423766	1.990688	-0.1185865	0.09495143	0.1090142
80	0.6645098	0.729458	1.171031	-0.1246661	0.04489143	0.06043306
100	0.1758446	0.4423283	0.4732496	-0.02355057	0.05949109	0.06004572
200	0.1146044	0.1470872	0.1602214	-0.03848292	0.01408208	0.01556301
300	0.1408959	0.06484056	0.08469221	-0.04973074	0.007839619	0.01031277

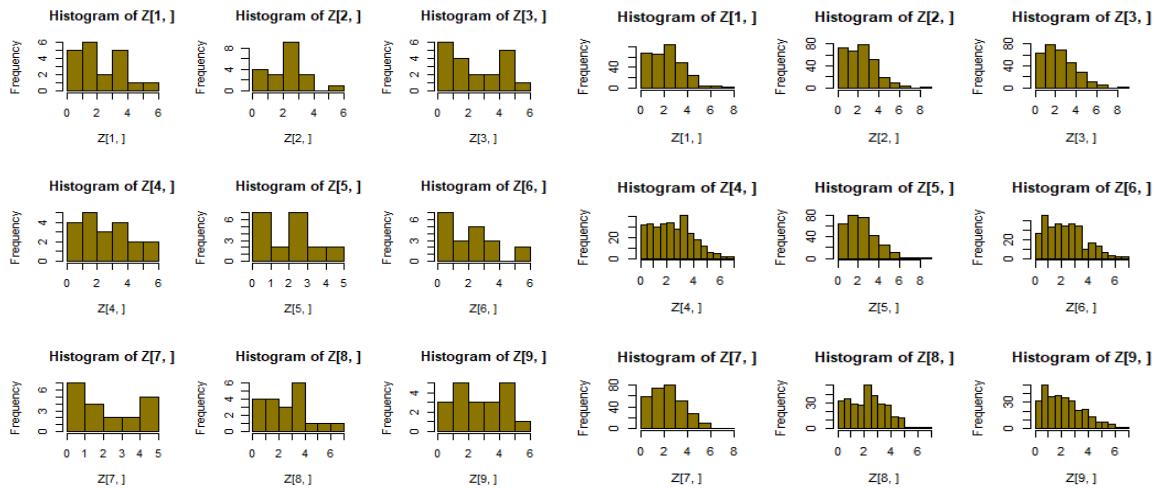


Fig.5: Simulation histogram when $n=20, \theta=2, \alpha=8$

Fig.6: Simulation histogram when $n=300, \theta=2, \alpha=8$

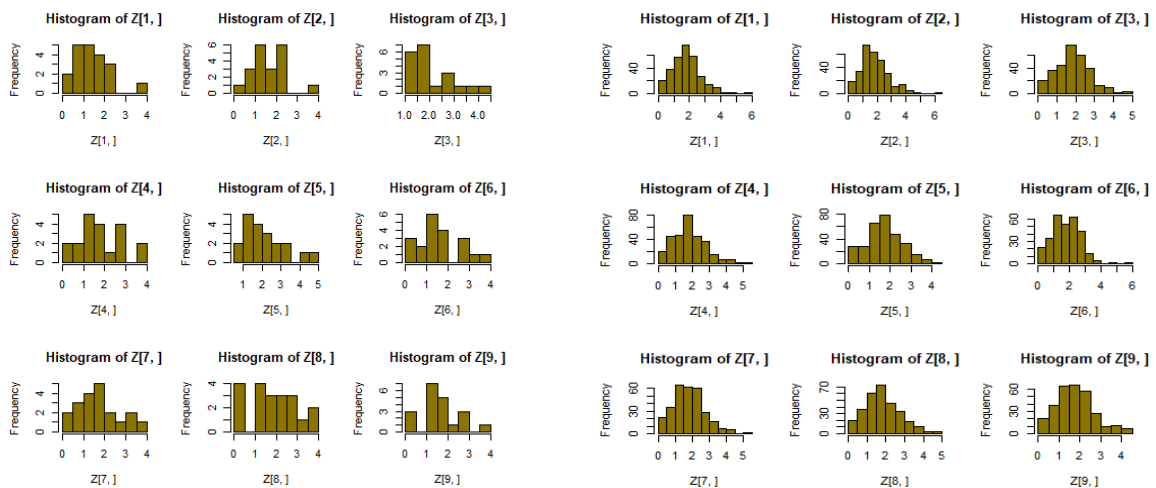


Fig.7: Simulation histogram when $n=20, \theta=3, \alpha=0.8$

Fig. 8: Simulation histogram when $n=300, \theta=3, \alpha=0.8$

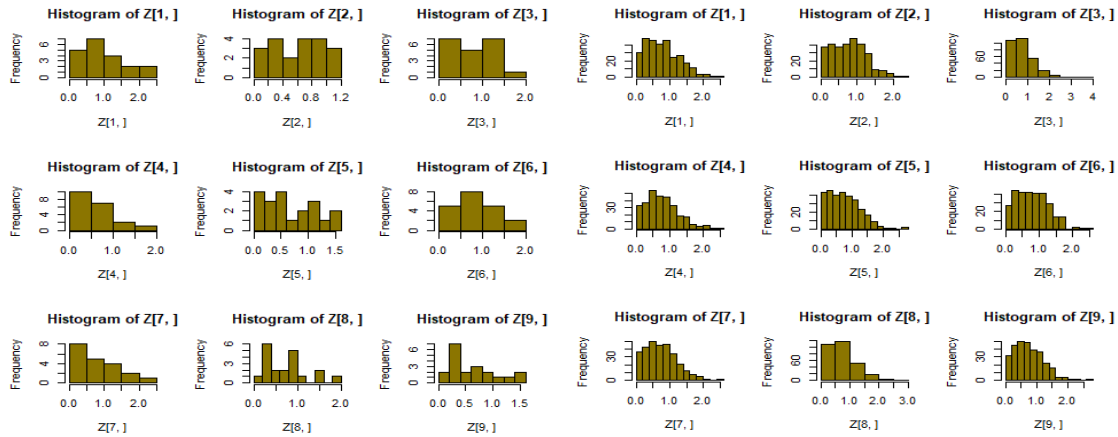


Fig.9: Simulation histogram when $n=20, \theta=6, \alpha=0.3$

Fig. 10: Simulation histogram when $n=300, \theta=6, \alpha=0.3$

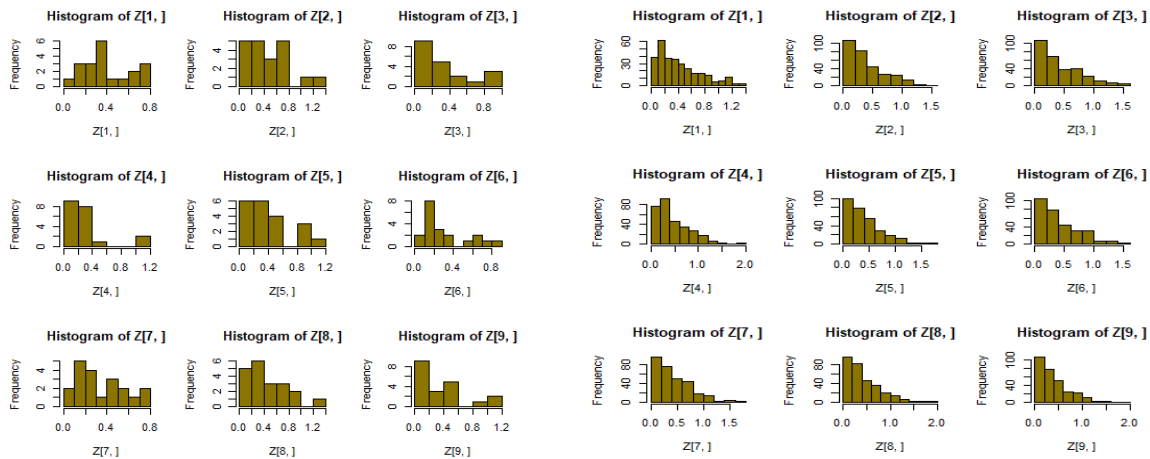


Fig.11: Simulation histogram when $n=20, \theta=8, \alpha=0.5$

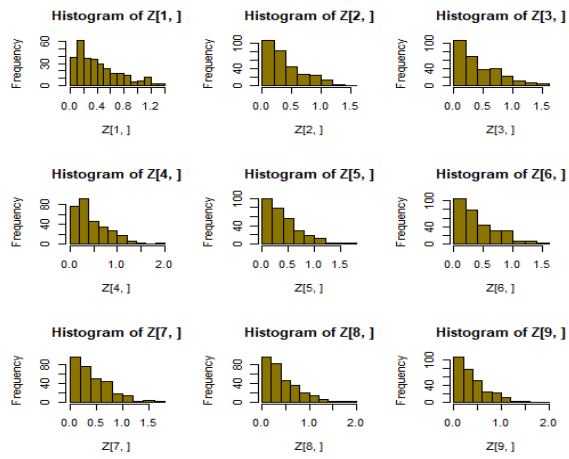


Fig.12: Simulation histogram when $n=300, \theta=8, \alpha=0.5$

10. Application

In this section we fitted LBQS distribution on a real-life time data set and compared the model with various distributions namely, Quasi Suja (QS), Exponential, Lindley, Erlang truncated exponential (ETE) distribution.

The data set is the period between failures for three repairable objects. This data set is provided in Hassan A.S [10] and explored later by Gadde S. R [8]

1.43, 1.23, 1.46, 0.11, 0.94, 0.30, 0.71, 4.36, 1.82, 0.77, 0.40, 2.37, 2.63, 1.74, 0.63, 1.49, 4.73, 1.23, 3.46, 2.23, 1.24, 2.46, 0.45, 1.97, 0.59, 0.70, 1.86, 0.74, 1.06, 1.17.

In order to compare LBQS distribution with the above-mentioned distributions, we consider the criteria like Bayesian information criterion (BIC), Akaike information criterion (AIC), Akaike information criterion corrected (AICC) and $-2 \log L$. Better distribution is said to be the one which has lower values of AIC, BIC, AICC and $-2 \log L$. These criteria can be calculated by using the following formulae.

$$AIC = 2k - 2 \log L, \quad BIC = k \log n - 2 \log L, \quad AICC = AIC + \frac{2k(k+1)}{n-k-1}$$

k is the number of parameters; n is the sample size and $-2\log L$ is the maximized value of log likelihood

Data set	Distributions	Parameters	MLE (Standard error)	$-2\log L$	AIC	BIC	AICC
1	LBQS	α	3.06130 (2.37265)	79.26209	83.26209	86.06448	83.706
		θ	1.29691 (.16734)				
	QS	α	0.29151 (0.29806)	86.74344	90.74344	93.54584	91.632
		θ	2.72816 (0.34681)				
	ETE	β	1.02149 (78.67718)	86.01075	90.01075	92.81315	90.455
		θ	1.00673 (133.7592)				
	Exponential	θ	0.64822 (0.11835)	86.01075	88.01075	89.41195	88.153
	Lindley	θ	0.9762395 (0.1345043)	83.09456	85.09456	86.49576	85.238

function. They are calculated for above mentioned data set and showed in table 2.

Table 2: Parameter estimations and goodness of fit test statistics

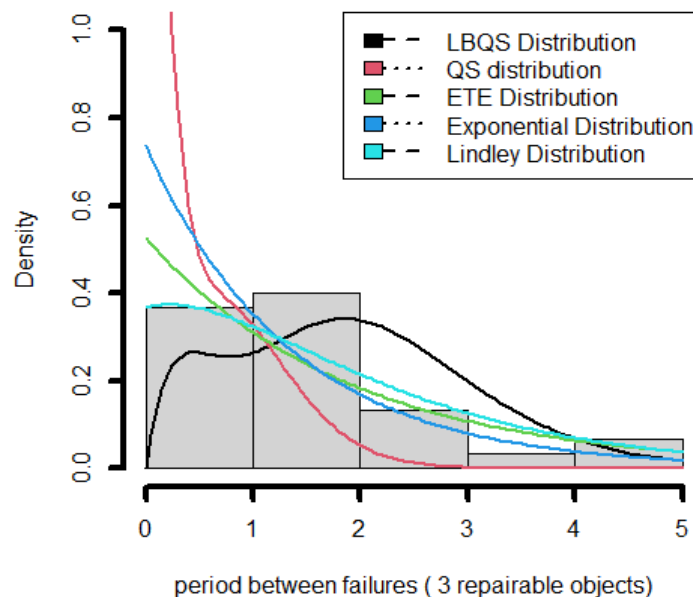


Fig. 13: Density curve of data set

From table 2 and density curve of data shown in Fig.13, it is evident that the LBQS distribution leads to better fit than the Quasi Suja, Exponential, Lindley, ETE distributions.

11. Conclusion

In this paper, a new modification of quasi suja distribution is executed namely length biased quasi suja distribution with two parameters and its different statistical properties are discussed and investigated. The distribution is generated by using the length biased technique and taking the two-parameter quasi suja distribution as the base distribution. The parameters of the executed distribution are obtained by using the maximum likelihood estimator. Finally, the usefulness of newly introduced distribution is discussed by applying the to real life data set and the result of the data set witnessed that the length biased quasi suja distribution fits better than the quasi suja, exponential, erlang truncated exponential, and lindley distributions.

References

- [1] Ahmad, A., Ahmad, S. P. and Ahmed, A. (2016). Length-Biased weighted Lomax distribution: Statistical properties and applications. *Pak.j.Stat.Oper.res*, 12:245-255.
- [2] Al-Omari, A. I, and Alsmairan, I. K. (2019). Length-biased Suja distribution and Its application. *Journal of Applied Probability and Statistics*, 14:95-116.
- [3] Al-Omari, A. I., Alhyasat, I. K. and Abu, B.M.A. (2019). Power Length-biased Suja distribution properties and application. *Electronic Journal of Applied Statistical Analysis*, 12:429-452.
- [4] Alsmairan, I. K., and Al-Omari, A. I. (2020). Weighted Suja distribution with application to ball bearings data. *Life Cycle Reliability and Safety Engineering*, 9:195-211.
- [5] Cox, D. R. (1969). Some sampling problems in technology, In New Development in Survey Sampling. *New York Wiley- Interscience*, 506-527.
- [6] Dey, S., Ali, S., and Park, C. (2015). Weighted exponential distribution: properties and different methods of estimation. *Journal of Statistical Computation and Simulation*, 85:3641-3661
- [7] Fisher, R. A. (1934), The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6:13-25.
- [8] Gadde, S. R., and Al-Omari, A. I., (2022). "Attribute control charts based on TLT for length-biased weighted Lomax distribution". *Journal of Mathematics*.
- [9] Ganaie, R. A, Subramanian, C., Soumya, V. P, Shenbagaraja, R., Mahfooz Alam, Vedavathi Saraja, D., Rushika Kinjawadekar, Rather, A. A. and Showkat Dar, A. (2023). Enhancing engineering sciences with Uma distribution: A perfect fit and valuable contributions. *Reliability: Theory & Applications*, 18:99-111.
- [10] Hassan, A. S., and Assar, S. M., (2017). The exponentiated Weibull-power function distribution. *Journal of Data Sciences*, 15:589-614.
- [11] Khan, M. N., Saeed, A. and Alzaatreh, A. (2018). Weighted Modified Weibull distribution. *Journal of Testing and Evaluation*, 47:170-370.
- [12] Kilany, N. M. (2016). Weighted Lomax distribution. *SpringerPlus*, 5.
- [13] Rao, C. R (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete. *Pergamum Press and Statistical publishing Society, Calcutta*, 320-332.
- [14] Rather, A. A. and Gamze Ozel (2020). The weighted power lindley distribution with applications on the life time data. *Pakistan Journal of Statistics and operation research*, 16:225-237.
- [15] Rama Shanker, Reshma Upadhyay and Kamlesh Kumar Shukla (2022). A Quasi Suja Distribution. *RT&A*, 17:69.

[16] Rashid Ganaie, A., Rather, A. A., Vedavathi Saraja, D., Mahfooz Alam, Aijaz Ahmad, Arif Muhammad Tali, Rifat Nisa, Berihan, R. and Elemary (2023). Exponentiated ADYA distribution: properties and applications. *Reliability; Theory & Applications*, 18:39-49.

[17] Rather, A. A. and Subramanian, C. (2018). Characterization and Estimation of Length Biased Weighted Generalized Uniform Distribution. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 5:72-76.

[18] Rather, A. A. and Subramanian, C. (2019). A new generalization of Akshaya distribution with applications in engineering science. *International Journal of Management, Technology and Engineering*, 11:2175-2184.

[19] Rather, A. A. and Subramanian, C. (2019). On weighted Sushila distribution with properties and its applications. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 6:105-117.

[20] Vedavathi Saraja, D., Jayakumar, B., Madhulika Mishra, Deshpande Priya, Subramanian, C, Rashid Ganaie, A, Rather, A. A, Bilal Ahmad Bhat. (2023). Exploring the length biased Tornumonkpe distribution: Properties estimations and practical applications. *Reliability: Theory & Applications*, 18:87-98.

[21] Shanker, R. (2017). Suja distribution and its application. *International Journal of Probability and Statistics*, 6:11-19.

[22] Todorka, T., Anton, I., Asen, R., and Nokolay, K. (2020). Comments on some modification of Suja Cumulative Functions with applications to the theory of computer viruses propagation. *International Journal of Differential Equations and Applications*, VIII, 19:83-95.

A NEW GENERALIZATION OF SABUR DISTRIBUTION

Suvarna Ranade¹, Aafaq A. Rather^{2,*}

•

^{1,2}Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India

¹Dr. Vishwanath Karad MIT World Peace University, Pune

¹suvarnamacs@gmail.com, ^{2,*}Corresponding Author: aafaq7741@gmail.com

Abstract

When the weight function depends on the lengths of the units of interest, the resulting distribution is called length biased. Length biased distribution is thus a special case of the more general form, known as weighted distribution. In this study, we introduce a novel probability distribution named the Length-Biased Sabur distribution (LBSD). This new distribution enhances the traditional Sabur distribution by incorporating a weighted transformation approach. The paper investigates the probability density function (pdf) and the cumulative distribution function (cdf) associated with the LBSD. A thorough examination of the distinctive structural properties of the proposed model is conducted, covering the survival function, conditional survival function, hazard function, cumulative hazard function, mean residual life, moments, moment generating function, characteristic function, likelihood ratio test, ordered statistics, entropy measures, and Bonferroni and Lorenz curve.

Key words: Sabur distribution, Length biased, Weighted transformation, Reliability analysis, Maximum likelihood estimator, Ordered statistics

1. Introduction

Weighted distributions occur when observations from a stochastic process are recorded with unequal probabilities, determined by a specific weighting function. When the weight function depends on the lengths of the units of interest, the resulting distribution is called length biased. Length biased distribution is thus a special case of the more general form, known as weighted distribution. The concept of length-biased distribution finds various applications in biomedical area such as family history and disease survival and intermediate events and latency period of AIDS due to blood transfusion [2]. The study of human families and wildlife populations was the subject of an article developed by Patil and Rao [8]. Patil, et al. [9] presented a list of the most common forms of the weight function useful in scientific and statistical literature as well as some basic theorems for weighted distributions and length-biased as special case. They arrived at the conclusion that the length biased version of some mixture of discrete distributions arises as a mixture of the length biased version of these distributions. Gupta R.D and Kundu D. [3] studied a new class of weighted exponential distribution which has applications in many fields such as: ecology, social and behavioural sciences and species abundance studies. Gupta R.C and Kirmani S. [2], studied the role of weighted distributions in stochastic modelling. Much work was done to characterize relationships between original distributions and their length biased version. A table for some basic distributions and their length biased forms is given by Patil and Rao [8] such as lognormal, Gamma, Pareto, Beta distribution. Khatree [4] presented a useful result by giving a relationship between the original random variable X and its length biased version Y . Recently Mudasar and S.P. Ahmad [5] studied the length biased Nakagami distribution. In subsequent years, Rather and Subramanian [10] explored the length-biased Erlang

truncated exponential distribution, highlighting its practical applications, Rather and Ozel [11] introduced a new length-biased power Lindley distribution with applications.

2. Probability density function and cumulative distribution function

The probability density function (pdf) of the Sabur distribution with two parameters α and β is defined as

$$f(x, \alpha, \beta) = \frac{\beta^2}{\alpha\beta + \beta^2 + 1} \left(\alpha + \beta + \frac{\beta}{2} x^2 \right) e^{-\beta x} \quad x > 0, \alpha, \beta > 0 \quad (1)$$

Suppose X is a non-negative random variable with pdf $f(x)$. Let $w(x)$ be the non-negative weight function, then the pdf of the weighted random variable X_w is given by

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))}, \quad x > 0 \quad (2)$$

Where $w(x)$ is a non-negative weight function and

$$E(w(x)) = \int w(x)f(x)dx \quad \text{and} \quad w(x) = x^c$$

In this paper, we will consider the weight function was $w(x) = x$, where $c=1$ and using the definition of weighted distribution, the pdf of the LBSD on is given as

$$f_w(x) = \frac{x f(x)}{E(w(x))} \quad (3)$$

Expected value is defined as

$$E(x) = \int_0^{\infty} x f(x) dx$$

$$E(x) = \frac{(\alpha\beta + \beta^2 + 3)}{\beta(\alpha\beta + \beta^2 + 1)} \quad (4)$$

Substituting equation (1) and (3) in equation (2) we obtain the density function of LBSD as follows

$$f_w(x, \beta, \alpha) = \frac{x\beta^3 \left(\alpha + \beta + \frac{\beta}{2} x^2 \right) e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)} \quad (5)$$

And the cumulative density function (cdf) of LBSD is obtained by

$$F_w(x) = \int_0^x f_w(x) dx$$

$$F_w(x) = \int_0^x \frac{x\beta^3 \left(\alpha + \beta + \frac{\beta}{2} x^2 \right) e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)} dx \quad (6)$$

After simplification, the cdf of the LBSD is given by

$$F_w(x) = \frac{2\beta(\alpha + \beta)\gamma(2, \beta x) + \gamma(4, \beta x)}{2(\alpha\beta + \beta^2 + 3)} \quad (7)$$

Fig. 1 and Fig. 2 visually illustrates the pdf and cdf of LBSD.

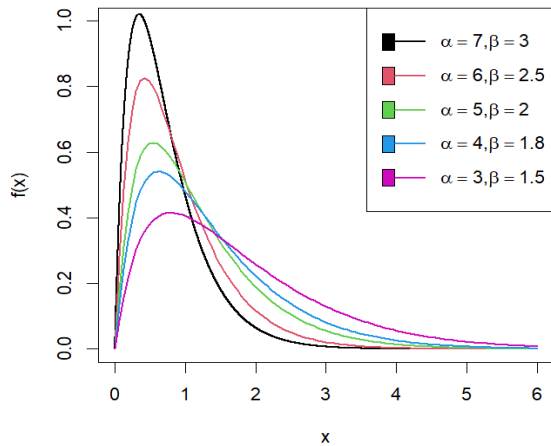


Fig. 1 : pdf plot of LBSD

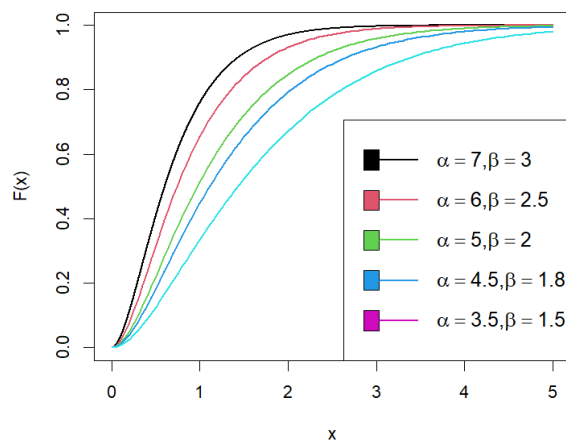


Fig. 2 : cdf plot of LBSD

3. Survival, Hazard and Reversed Hazard Functions

In this section we discuss about the survival function, hazard and reverse hazard functions of the LBSD. The survival function or the reliability function of is given by

$$S(x) = 1 - F_w(x) \tag{8}$$

$$S(x) = 1 - \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x) + \gamma(4,\beta x)}{2(\alpha\beta + \beta^2 + 3)} \right) \tag{9}$$

The hazard function is also known as the hazard rate function, instantaneous failure rate or force of mortality and is given for the LBSD as

$$h(x) = \frac{f_w(x)}{s(x)} \tag{10}$$

$$h(x) = \frac{\frac{x\beta^3(\alpha+\beta+\frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)}}{1 - \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x) + \gamma(4,\beta x)}{2(\alpha\beta + \beta^2 + 3)} \right)} \tag{11}$$

$$h(x) = \frac{2x\beta^3(\alpha+\beta+\frac{\beta}{2}x^2)e^{-\beta x}}{2(\alpha\beta + \beta^2 + 3) - 2\beta(\alpha+\beta)\gamma(2,\beta x) + \gamma(4,\beta x)} \tag{12}$$

The reverse hazard function of the LBSD is given by

$$h_r(x) = \frac{f_w(x)}{F_w(x)} \tag{13}$$

$$h_r(x) = \frac{2x\beta^3(\alpha+\beta+\frac{\beta}{2}x^2)e^{-\beta x}}{2\beta(\alpha+\beta)\gamma(2,\beta x) + \gamma(4,\beta x)} \tag{14}$$

Fig. 3 and Fig. 4 depicts the graphical survival function and Hazard function plot of LBSD.

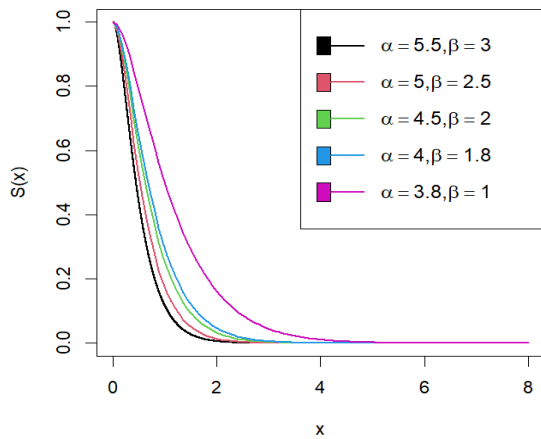


Fig. 3: Survival function plot of LBSD

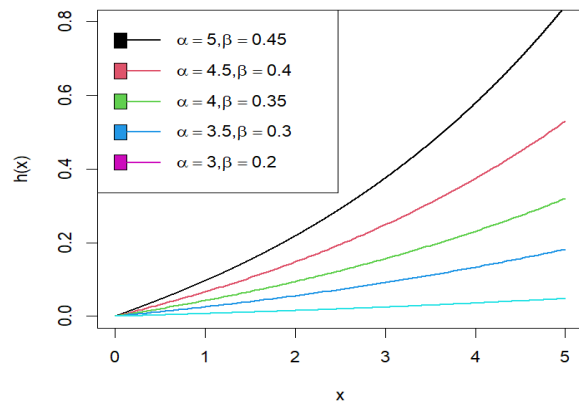


Fig. 4: Hazard function plot of LBSD

4. Structural properties

In this section we investigate various structural properties of the LBSD

Let X denote the random variable of LBSD with parameters α, β , then its r^{th} order moment $E(x^r)$ about origin is given by

$$E(x^r) = \mu_r' = \int_0^\infty x^r f_w(x) dx \tag{15}$$

$$E(x^r) = \int_0^\infty x^r \frac{x\beta^3(\alpha+\beta+\frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta+\beta^2+3)} dx \tag{16}$$

After simplifying the expression, we get

$$E(x^r) = \frac{2\beta(\alpha+\beta)\Gamma(r+2)+\Gamma(r+4)}{2\beta^r(\alpha\beta+\beta^2+3)} \tag{17}$$

Putting $r=1$, we get the expected value of LBSD as follows

$$E(x) = \frac{2\beta(\alpha+\beta)+12}{\beta(\alpha\beta+\beta^2+3)} \tag{18}$$

Put $r=2$, we obtained second moment as

$$E(x^2) = \frac{6\beta(\alpha+\beta)+60}{\beta^2(\alpha\beta+\beta^2+3)} \tag{19}$$

The variance of LBSD is calculated as

$$V(x) = E(x^2) - [E(x)]^2$$

$$V(x) = \frac{6\beta(\alpha+\beta)+60}{\beta^2(\alpha\beta+\beta^2+3)} - \left[\frac{2\beta(\alpha+\beta)+12}{\beta(\alpha\beta+\beta^2+3)}\right]^2 \tag{20}$$

4.1 Harmonic mean

The harmonic mean of the LBSD of random variable X can be written as

$$H = E\left(\frac{1}{x}\right) = \int_0^{\infty} \frac{1}{x} f_w(x) dx$$

$$H = \int_0^{\infty} \frac{1}{x} \frac{x\beta^3(\alpha+\beta+\frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta+\beta^2+3)} dx \tag{21}$$

After simplification we get

$$H = \frac{\beta(\beta^2+\alpha\beta+1)}{(\beta^2+\alpha\beta+3)} \tag{22}$$

4.2 Moment generating function and characteristic function

Let X have a LBSD, then the Moment generating function of X is obtained as

$$M_X(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} f_w(x) dx$$

Using Taylor's series, we obtain

$$M_X(t) = E(e^{tx}) = \int_0^{\infty} \left(1 + tx + \frac{(tx)^2}{2!} + \dots\right) f_w(x) dx$$

$$M_X(t) = \int_0^{\infty} \sum_{i=0}^{\infty} \frac{t^i}{i!} x^i f_w(x) dx$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} E(x^j) dx$$

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!} \frac{2\beta(\alpha+\beta)\Gamma j+2 + \Gamma j+4}{2\beta^j(\alpha\beta+\beta^2+3)} \tag{23}$$

Similarly, the characteristic function of LBSD of random variable X can obtained as

$$\Phi_X(t) = M_X(it) = \sum_{j=0}^{\infty} \frac{(it)^j}{j!} \frac{2\beta(\alpha+\beta)(\Gamma j+2) + (\Gamma j+4)}{2\beta^j(\alpha\beta+\beta^2+3)} \tag{24}$$

5. Likelihood Ratio Test

Let X1, X2, X3....be a random sample from the LBSD, we use the hypothesis

$$H_0 : f(x) = f(x; \alpha, \beta) \text{ against } H_1 : f(x) = f_w(x; \alpha, \beta, 1)$$

In order to test whether the random sample of size n comes from the Sabur distribution or weighted Sabur distribution, we will use following statistics

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^n \frac{f_w(x; \alpha, \beta, 1)}{f(x; \alpha, \beta)}$$

$$\Delta = \prod_{i=1}^n \frac{x\beta(\alpha\beta+\beta^2+1)}{(\alpha\beta+\beta^2+3)} \tag{25}$$

$$\Delta = A^n \prod_{i=1}^n x_i \quad \text{where}$$

$$A = \frac{\beta(\alpha\beta+\beta^2+1)}{(\alpha\beta+\beta^2+3)} \tag{26}$$

We reject the null hypothesis, if

$$\Delta = A^n \prod_{i=1}^n x_i > k$$

$$\Delta^* = \prod_{i=1}^n x_i > k A^n$$

For large sample size n, $2\log \Delta$ is distributed as chi square distribution with one degree of freedom and also p-value is obtained from the chi-square distribution. Thus we reject the null hypothesis, when the probability value is given by

$$P(\Delta^* > a^*)$$

Where a^* is less than a specified level of significance and $\prod_{i=1}^n x_i$ is the observed value of the statistics Δ^* .

6. Entropy Measures

The concept of entropy is important in different areas such as probability and statistics, physics, communication theory and economics. Entropy measures quantify the diversity, uncertainty or randomness of a system. Entropy of a random variable X is measure of variation of the uncertainty.

6.1 Renyi Entropy

It was proposed by Renyi(1957). The Renyi entropy of order ξ for a random variable X is given by

$$e(\xi) = \frac{1}{1-\xi} \log \left(\int_0^\infty f^\xi(x) dx \right) \quad \text{where } \xi > 0 \text{ and } \xi \neq 1$$

$$e(\xi) = \frac{1}{1-\xi} \log \left(\int_0^\infty \left(\frac{x\beta^3(\alpha+\beta+\frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta+\beta^2+3)} \right)^\xi dx \right) \tag{27}$$

After simplifying the equation we get

$$e(\xi) = \frac{1}{1-\xi} \log \left(\left(\frac{\beta^3}{(\alpha\beta+\beta^2+3)} \right)^\xi \sum_{i=0}^\infty \binom{\xi}{i} (\alpha + \beta)^{\xi-i} \left(\frac{\beta}{2} \right)^i \frac{\Gamma(\xi+2i+1)}{\beta^\xi \xi^{2i+1}} \right) \tag{28}$$

6.2 Tsallis Entropy

A generalization of Boltzman-Gibbs(B-G) statistical mechanics initiated by Tsallis has focussed a great deal to attention. This generalization of B-G statistics was proposed firstly by introducing the mathematical expression of Tsallis entropy (Tsallis, 1988) for a continuous random variable. Tsallis entropy of order λ of the weighted Sabur distribution is given by

$$S_\lambda = \frac{1}{\lambda - 1} \left(1 - \int_0^\infty f^\lambda(x) dx \right)$$

$$S_\lambda = \frac{1}{\lambda - 1} \left(1 - \int_0^\infty \left(\frac{x\beta^3(\alpha + \beta + \frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)} \right)^\lambda dx \right) \tag{29}$$

After simplifying the expression, we get

$$S_\lambda = \frac{1}{\lambda - 1} \left[\left(1 - \left(\frac{\beta^3}{(\alpha\beta + \beta^2 + 3)} \right)^\lambda \right) \sum_{i=0}^\infty \binom{\lambda}{i} (\alpha + \beta)^{\lambda - i} \left(\frac{\beta}{2} \right)^i \frac{\Gamma(\lambda + 2i + 1)}{\beta^\lambda (\lambda + 2i + 1)} \right] \tag{30}$$

7. Order Statistics

Let $X_{(1)}, X_{(2)}, X_{(3)}, \dots, X_{(n)}$ be the order statistics of a random sample $X_1, X_2, X_3, \dots, X_n$ drawn from the continuous population with pdf $f_x(x)$ and cdf $F_x(x)$ then the pdf of r th order statistic $X(r)$ is given by

$$f_{x(r)}(x) = \frac{n!}{(r - 1)!(n - r)!} f_x(x) [F_x(x)]^{r-1} [1 - F_x(x)]^{n-r}$$

Substituting equation (4) and (5) in equation (6), the pdf of order statistics $X(r)$ of the weighted Sabur distribution is given by

$$f_{x(r)}(x) = \frac{n!}{(r - 1)!(n - r)!} \left(\frac{x\beta^3(\alpha + \beta + \frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)} \right) \times \left(\frac{2\beta(\alpha + \beta)\gamma(2, \beta x) + \gamma(4, \beta x)}{2(\alpha\beta + \beta^2 + 3)} \right)^{r-1} \times \left(1 - \left(\frac{2\beta(\alpha + \beta)\gamma(2, \beta x) + \gamma(4, \beta x)}{2(\alpha\beta + \beta^2 + 3)} \right) \right)^{n-r} \tag{31}$$

Therefore, the pdf of the higher order statistics $X(n)$ can be obtained as

$$f_{x(n)}(x) = n \left(\frac{x\beta^3(\alpha + \beta + \frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)} \right) \times \left(\frac{2\beta(\alpha + \beta)\gamma(2, \beta x) + \gamma(4, \beta x)}{2(\alpha\beta + \beta^2 + 3)} \right)^{n-1} \tag{32}$$

And the pdf of the first order statistics $X_{(1)}$ can be obtained as

$$f_{x(1)}(x) = n \left(\frac{x\beta^3(\alpha + \beta + \frac{\beta}{2}x^2)e^{-\beta x}}{(\alpha\beta + \beta^2 + 3)} \right) \times \left(1 - \left(\frac{2\beta(\alpha + \beta)\gamma(2, \beta x) + \gamma(4, \beta x)}{2(\alpha\beta + \beta^2 + 3)} \right) \right)^{n-1} \tag{33}$$

8. Income Distribution Curve

The Bonferroni and the Lorenz curves are not only used in economics in order to study the income and poverty, but it is also being used in other fields like reliability, medicine and demography. The Bonferroni and Lorenz curves are given by

$$B(p) = \frac{1}{p\mu_1} \int_0^q x f(x) dx \quad \text{and}$$

$$L(p) = PB(p) = \frac{1}{\mu'_1} \int_0^q x f(x) dx$$

Here, we define the first raw moments as

$$\mu'_1 = \frac{2\beta(\alpha+\beta)+12}{\beta(\alpha\beta+\beta^2+3)} \tag{34}$$

And $q = F^{-1}(p)$, Then we have

$$B(p) = \frac{2\beta(\alpha+\beta)\gamma(3,\beta q)+\gamma(5,\beta q)}{2p(2\beta(\alpha+\beta)+12)} \tag{35}$$

$$L(p) = PB(p) = \frac{2\beta(\alpha+\beta)\gamma(3,\beta q)+\gamma(5,\beta q)}{2(2\beta(\alpha+\beta)+12)} \tag{36}$$

9. Estimation

We will discuss the maximum likelihood estimators (MLEs) of the LBSD. Consider $X_1, X_2, X_3, \dots, X_n$ be the random sample of size n from the LBSD, then the likelihood function is given by

$$L(x; \alpha, \beta) = \prod_{i=1}^n x_i \frac{\beta^3(\alpha+\beta+\frac{\beta}{2}x_i^2)e^{-\beta x_i}}{(\alpha\beta+\beta^2+3)} \tag{37}$$

$$L(x; \alpha, \beta) = \frac{\beta^{3n}}{(\alpha\beta+\beta^2+3)^n} \prod_{i=1}^n x_i \left(\alpha + \beta + \frac{\beta}{2} x_i^2 \right) e^{-\beta x_i} \tag{38}$$

The loglikelihood function is obtained as

$$\text{Log } L = 3n \log \beta - n \log(\alpha\beta + \beta^2 + 3) + \log \sum x_i + \sum \log(\alpha + \beta + \frac{\beta}{2} x_i^2) - \beta \sum x_i \tag{39}$$

The MLEs of α, β can be obtained by differentiating Log L with respect to α, β and must satisfy the normal equation.

$$\frac{\partial \log L}{\partial \beta} = -\frac{3n}{\beta} - \frac{n(\alpha+2\beta)}{(\alpha\beta+\beta^2+3)} + \sum_{i=1}^n \frac{\left(1+\frac{x_i^2}{2}\right)}{(\alpha+\beta+\frac{\beta}{2}x_i^2)} - \sum x_i = 0 \tag{40}$$

$$\frac{\partial \log L}{\partial \alpha} = \left[-\frac{n\beta}{(\alpha\beta+\beta^2+3)} + \sum_{i=1}^n \frac{1}{\alpha+\beta+\frac{\beta}{2}x_i^2} \right] = 0 \tag{41}$$

To obtain confidence interval we use the asymptotic normality results. We have that, if $\hat{\lambda} = (\hat{\alpha}, \hat{\beta}, \hat{c})$ denotes the MLE of $\lambda = (\alpha, \beta, c)$ we can state the results as follows

$$(\hat{\lambda} - \lambda) \rightarrow N_3(0, I^{-1}(\lambda))$$

Where $I(\lambda)$ is Fisher's Information matrix given by

$$I(\lambda) = -\frac{1}{n} \begin{pmatrix} E\left(\frac{\partial^2 \log l}{\partial \alpha^2}\right) & E\left(\frac{\partial \log l}{\partial \beta \partial \alpha}\right) \\ E\left(\frac{\partial \log l}{\partial \beta \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial \beta^2}\right) \end{pmatrix} \tag{42}$$

Here we define

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{3n}{\beta^2} - n \left(\frac{-\alpha^2 - 2\alpha\beta - 2\beta^2 + 6}{(\alpha\beta + \beta^2 + 3)^2} \right) - \frac{\left(1 + \frac{1}{2}\sum x_i^2\right)^2}{\left(\alpha + \beta + \frac{\beta}{2}\sum x_i^2\right)^2} \quad (43)$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{n\beta}{(\alpha\beta + \beta^2 + 3)^2} - \frac{n}{\left(\alpha + \beta + \frac{\beta}{2}\sum x_i^2\right)^2} \quad (44)$$

$$\frac{\partial^2 \log L}{\partial \beta \partial \alpha} = -n \left(\frac{-\beta(\alpha + 2\beta) - 1}{(\alpha\beta + \beta^2 + 3)^2} \right) \quad (45)$$

10. Conclusion

In this paper, we introduce a novel extension of the Sabur distribution by incorporating a weighted transformation approach. This extension builds upon the existing two-parameter Sabur distribution, resulting in a three-parameter model known as the Length-Biased Sabur distribution. We conduct a comprehensive analysis of this new distribution, exploring its mathematical formulation and statistical properties in detail. Parameter estimation is performed using maximum likelihood estimation techniques.

References

- [1] Alzaatreh, A., Famoye, F. and Lee, C. (2013). Weibull-pareto distribution and its applications, *Communications in Statistics-Theory and Methods*. 42(9): 1673-1691.
- [2] Gupta, R.C. and Kirmani, S. (1990). The role of weighted distributions in stochastic modelling, *Communications in Statistics, Theory and Methods*. 19, 3147- 3162.
- [3] Gupta, R.D. and Kundu, D. (2009). A new class of weighted exponential distribution, *Statistics*. 43, 621 - 634.
- [4] Khatree, R. (1989). Characterization of Inverse-Gaussian and Gamma distributions through their length biased distribution. *IEEE Transactions on Reliability*. 38 (1) 610-611.
- [5] Mudasir, S., & Ahmad, S. P. (2015). Structural Properties of Length-Biased Nakagami Distribution. *International Journal of Modern Mathematical Sciences*. 13, 217-227.
- [6] Maxwell O, Chukwudike NC, and Bright OC (2019). Modelling lifetime data with the odd generalized exponentiated inverse Lomax distribution. *Biom Biostat Int J*. 8(2), 39–42.
- [7] Oguntunde, P. E., Adejumo, A. O., Okagbue, H. I. and Rastogi M. K. (2016). Statistical properties and applications of a new lindley exponential distribution, *Gazi University Journal of Science*, 29(4): 831-838
- [8] Patil, G.P. & Rao, C.R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families, *Biometrics*. 34, 179– 184.
- [9] Patil, G. P., and Ord, J. K. (1976). On size-Biased Sampling and Related Form Invariant Weighted distribution, *Sankhya: The Indian Journal of Statistics*. 38, Series B, 48-61.
- [10] Rather, A. A. & Subramanian, C. (2019), The Length-Biased Erlang–Truncated Exponential Distribution with Life Time Data, *Journal of Information and Computational Science*, vol-9, Issue 8, pp 340-355.
- [11] Rather, A. A. & Ozel G., (2021): A new length-biased power Lindley distribution with properties and its applications, *Journal of Statistics and Management Systems*, DOI: 10.1080/09720510.2021.1920665.

MODELING THE INTERCONNECTED OPERATION OF ENERGY SYSTEMS FOR ENERGY SECURITY STUDY IN TODAY'S CONTEXT

Dmitry Krupenev, Natalia Pyatkova

•
Energy Security Department

Melentiev Energy Systems Institute SB RAS

Irkutsk, Russia Lermontov str. 130

Corresponding author: Dmitry Krupenev krupenev@isem.irk.ru

Abstract

The paper shows the need for comprehensive research into energy security problems to assess the possibilities of interconnected operation of all energy industries with the view to identifying the implications for consumers of energy resources in the event of emergencies in one or several industries at the same time. The paper presents a methodological framework and features of modeling the interrelated operation of the industries in current context and a model developed for these studies. The results of experimental studies using the developed methodology are shown through the analysis of several critical situations (threats to energy security) of various nature.

Keywords: energy security; critical situations; economic and mathematical model.

I. Introduction

Ensuring energy security and maintaining reliability of fuel and energy supply are crucial in today's social and economic landscape. Energy security (ES) is of utmost importance for protecting the citizens, society, the state, and the economy from the threat of a shortage when meeting their energy needs with economically affordable energy resources of high quality and from the threats of potential disruption of constant energy supply [1, 2]. Essentially, we are discussing the importance of maintaining the balance between the supply and demand of different kinds of fuel and energy when various kinds of threats that affect the energy sector and lead to a decrease in the supply of consumers with energy resources come into being [3].

Energy systems form the energy sector of a country or individual regions. The fuel and energy sector of the country is one of the largest intersectoral complexes, including formally independent industries, such as electric power, thermal power, coal, oil, oil refining and gas industries, which are united technologically and territorially [4,5, 6]. At the present level of consideration, it is crucial to highlight and incorporate renewable energy, cooling energy, water management system, and water supply system into the structure of the energy sector.

The primary objectives of the energy security research are to predict the conditions for the operation and expansion of fuel and energy systems and the development of the energy sector as a whole in the context of possible critical situations and emergencies; to provide state estimation under these conditions and to identify "bottlenecks" in fuel and energy systems, and in the energy

supply to the consumer; to choose possible alternatives and specific measures to prevent critical situations and emergencies in these systems and at energy facilities, or reduce their negative impact to an acceptable level [7, 8]. In the current context, it is imperative to enhance the existing methodological, modeling, and software tools and develop new ones for conducting such studies, since the concept of risk of critical and emergency situations and their consequences is becoming increasingly prominent. Of particular importance is analysis of potential threats, the development of disturbance scenarios (critical situations and emergencies) based on the analysis results, and the related modeling problems [8].

Assessment of the energy security level normally rests on two methods: an indicator-based method, which is the primary and simplest way of assessing the level of energy security, and a method based on mathematical modeling of the interconnected operation of energy systems. The second method is used to investigate and assess the impact of energy security threats on the reliability of energy supply to consumers with adequate mathematical models of energy components and systems.

The indicator-based method has found wide application in the analysis of the energy security level in many countries, since the indicators obtained are the most understandable and minimal effort is required to determine them. Despite its simplicity, this method is quite effective for energy security assessment.

In Russia, since the 1990s, the indicator-based method of analysis has been widely developed in the works performed by the research team of L.A. Melentiev Energy Systems Institute of the Siberian Branch of the Russian Academy of Sciences, which focused on the energy security studies for the country and its regions [9-14]. These works offered a wide set of parameters (indicators) to characterize the state and conditions for the expansion and operation of individual energy systems and the levels of threats to energy security [10]. The developed system of indicators was applied to assess the state of energy security of Russia's regions [11, 12]. A methodology developed for the indicator-based analysis determines:

- a) a set of energy security indicators at the federal level;
- b) a procedure for obtaining the numerical value of each indicator from this set;
- c) a procedure for determining threshold values for each indicator at issue [13].

The impact of energy security threats on the reliability of fuel and energy supply to consumers was also assessed using the method of indicator-based analysis [14].

Researchers in other countries use the method of indicator-based analysis to assess the level of energy security based on the indicators established for their country. Eighteen indicators are used for China [15]. They are categorized into three areas: energy supply, economic-technical, and environmental. The studies carried out for six countries of the Caspian Sea basin (Azerbaijan, Iran, Kazakhstan, Uzbekistan, Turkmenistan, Russia) aim to analyze the level of energy security [16] for each country individually and collectively based on three dimensions: resources and dependency, intensity and sustainability, cost and poverty.

At the same time, the indicator-based analysis is static and yields an assessment of the indicators for a specific time span, neglecting the influence of the dynamics of processes in the energy sector and the influence of the system-wide effect from the mutually coordinated operation of energy systems. Therefore, studies of the interrelated operation of energy industries and their modeling to assess the impact of energy security threats on reliability of energy supply are of greatest interest.

Modeling the interconnected operation of industry systems can be divided into two categories: modeling of concentrated nodes, this modeling concept is also called the concept of an energy hub, and detailed modeling of energy systems, which either partially or fully considers the transmission links between individual nodes.

Studies based on the concept of energy hubs [17-19] normally present an energy hub model involving interaction of three types of energy resources (heat, electricity, and gas). These studies also

take into account renewable generation and the possibility of using energy storage systems (ESS). Particular interactions of energy flows are considered, for example, modeling the interdependence of the gas and electric power industries [20]. The main difference between the models lies in the description of the production capacities of the region for the processing and storage of energy resources in the form of energy nodes or hubs. For example, regional condensing power plants (CPPs), combined heat and power plants (CHPPs) and boiler houses are combined into an energy hub to generate heat and electricity. This method of description offers a convenient scheme for presenting the initial data and simplifies the architecture of the corresponding software, which provides a gain in time during distributed computing. In this case, the focus is on simplification of calculations but without sufficient consideration of the links between industries as an interconnected system.

The second approach to modeling the interrelated operation of various sectoral systems is more preferable for investigating energy security problems and analyzing threats to the normal functioning and expansion of sectoral energy systems. The analysis of the current state of research in this area helped identify the following areas: 1) research on the conceptual area of the energy security problem in terms of reliable energy supply and 2) research on modeling energy systems and energy sector.

The papers [21-26] present the results of studies that involve assessing the possibilities of network flow modeling of resource supply processes with a sharp deterioration in the properties of the energy infrastructure. These works examine power supply options in geographically distributed systems after destructive impacts. The studies focus on the conditions of regulatory restrictions on accidents, in contrast to the research presented in this paper, which considers all possible emergency situations, analyzes threats and their manifestations, evaluates the operating conditions of sectoral systems in this context, assesses the reliability of supplying the consumer with energy resources (in the form of undersupplied amounts).

The studies presented in [27] are also close in their formulation of the problem. They are aimed at solving the security problem for a group of target objects that receive energy resources from the network infrastructure in case of a negative impact on the network components. The problem solved however belongs to the industry level of the hierarchy. At the same time, a comprehensive assessment of energy facilities with their mutually coordinated operation within a single system is not carried out.

Of some interest are studies related to the energy independence of countries importing energy resources, and the related issues of creating reserves of these energy resources and diversifying them in the event of supply disruptions. In these studies [28], a quantitative model of the energy security of the country (China) is proposed to calculate the optimal scale of strategic reserves of oil and alternative fuels (for example, coal-fired methanol). The developed model mainly focuses on cost-benefit analysis, including the economic costs of creating such reserves and reserves of alternative fuels, and thereby increasing the energy security of the country. The results of experimental calculations for China show that the creation of strategic oil reserves and alternative fuels have a positive effect on energy security, and alternative fuel is an additional way to reduce the negative consequences of rising oil prices and related economic losses. This approach can be used when creating structural redundancy in case of emergencies affecting the energy supply of consumers.

There are important developments in modeling energy systems and energy sector, with two classes of models to be distinguished: 1) simulation ones that optimize the technological structure of the energy industry; 2) economic or macroeconomic ones, in which energy is presented as a sector of the economy as a whole.

The main representatives of simulation class models are MARKAL (MARKet ALlocation), EFOM (Energy Flow Optimization Model), MESSAGE (Model for Energy Supply Strategy Alternatives and their General Environmental Impacts), and others [29-32].

The MESSAGE model is the most widespread. It was developed by the International Institute for Applied Systems Analysis (IIASA) with the view to planning and projecting the expansion of energy systems. This modeling system is intended for medium- and long-term expansion planning of energy systems, energy policy analysis, and design of development scenarios. This tool allows comparing alternative energy technologies and build the most appropriate scenario for the development of the energy system. It also enables comparison of alternative scenarios for the development of the energy system in terms of environmental impact. The linear programming method is used to find the optimal solution in MESSAGE. The selection criterion is the minimum of the reduced system costs. According to the objectives to be accomplished, this model tool is focused on a time interval under average statistical conditions of development. Therefore, it cannot be used to study the response of sectoral systems to disturbing influences when threats to the normal functioning of industry systems and consumer systems materialize, which is proposed to be used in this article.

The main representative of the second class is the National Energy Modeling System (NEMS) developed by the US Department of Energy [33].

This system models the US economy with the allocation of energy as a separate sector. This model makes projections for energy production, import, processing, consumption and prices, given the macroeconomic and financial indices of the world market for energy resources; supports the choice of certain technologies, their quantitative and qualitative characteristics. Adaptation of this modeling system to assessing the behavior of the energy industries and the consumer sector in critical situations is not possible for the following reasons:

- The considered time intervals do not coincide;
- The insufficiently detailed representation of the energy industries in the complex;
- The impossibility of determining the shortage of energy resources in the event of emergencies in the energy industries.

The most representative domestic development in the field of energy system modeling is the SCANNER model-information system developed at the Energy Research Institute of the Russian Academy of Sciences [34]. This is a unique tool for systems research into the development of Russia's energy sector as an important part of the national economy, and global energy markets in the medium and long term (until 2030-2050). SCANNER combines powerful analytical tools, about twenty mathematical models for comprehensive projection and optimization of the national and global energy development for the main stages of energy conversion - from production (about 20 types of primary energy resources) to consumer use (10 main energy carriers). The analyzed tool is aimed at studying the prospects for the development of the country's energy economy under normal operating conditions, considering the influence of global factors. At the same time, the reliability of power supply is taken into account according to average statistical standards, leaving extreme situations out of consideration. It is problematic to use it to investigate the mutually coordinated operation of energy industries in critical situations, since this tool does not determine the impact of energy shortages on the reliability of fuel and energy supply to the consumer.

The presented developments and models focus on solving the problems of long-term planning of the power industry under normal operating conditions with a horizon of up to 15-20 years. The studies conducted and described in this paper stand out by their emphasis on assessing the behavior of energy systems in the face of energy security threats and optimizing the interrelated operation of energy systems in the event of emergencies to provide reliable energy supply to consumers.

Modern conditions for the advancement of information technologies, the emergence of high-performance computing tools, as well as the intelligentization of energy systems and the need for their functioning in a digital economy, on the one hand, impose special requirements on the modeling and computing tools to be used. On the other hand, they provide opportunities to enhance the adequacy and correctness of modeling the real-world systems by considering the inertia of

processes, the dynamics of unfolding critical situations in the models designed to optimize energy systems within the energy sector; by taking into consideration non-linearity in terms of the adequacy of the representation of processes in energy systems to improve the accuracy of decisions made.

This paper aims to present a methodological tool for modeling the interlinked operation of energy systems within the energy sector for examining energy security problems and illustrating the outcomes achieved by applying the developed models.

II. Methodology for integrated modeling of energy systems to solve the energy security problems.

To investigate the problems of energy security in today's context, it is proposed to develop new and improve (adapt) existing mathematical models and methods of the interrelated operation of large energy systems within energy sector under various operating conditions. The use of an enhanced modeling system will make it possible to assess the possibilities of providing consumers with energy resources when various threats to energy security materialize.

The general scheme of tasks to be accomplished when assessing the effect of threats on the state of energy security is shown in Figure 1.

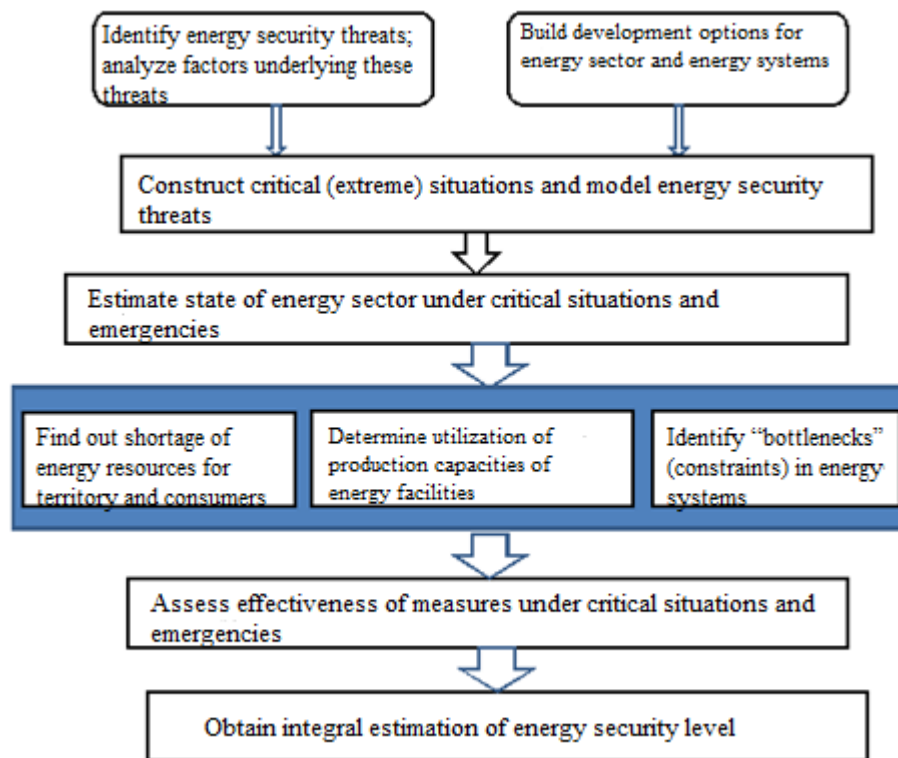


Figure 1. General scheme of energy security research

The initial basis for the research is the technical and economic characteristics of energy facilities and reporting data on the state of energy systems, the findings of the energy development research providing the rationale for the choice of a long-term strategy and the formulation of an energy policy. Based on the adopted socio-economic program for the future development of the national economy, which determines the demand for fuel and energy resources, an analysis and assessment of energy consumption levels is made considering energy conservation.

Following the above characteristics and analysis of energy security threats, the design

conditions are established for a computational experiment, which is carried out using models of energy systems.

Models of energy systems represent *a system of economic and mathematical models for assessing the territorial and production structure of the energy sector, in terms of energy security requirements* [35]. These models can be used in two modes:

- in the mode of determining the optimal development of energy technologies (given structural redundancy in the form of capacity reserves, fuel reserves, interchangeability of energy resources) and the optimal distribution of consumed energy resources,

- in the mode of identifying undersupply of energy resources (shortage in fuel and energy resources) in the country as a whole and in individual regions.

The structure of the energy sector is shown in Figure 2. Technologically, it consists of the modules of industry-specific subsystems of the energy sector (gas, coal, oil refining (in terms of fuel oil supply), electricity, and thermal power industries), and a module of consumers (consumption of energy resources at various types of power plants and boiler houses for generating electricity and heat, and other consumers, with separately allocated export consumers). This version does not include energy storage systems, water management and water supply systems.

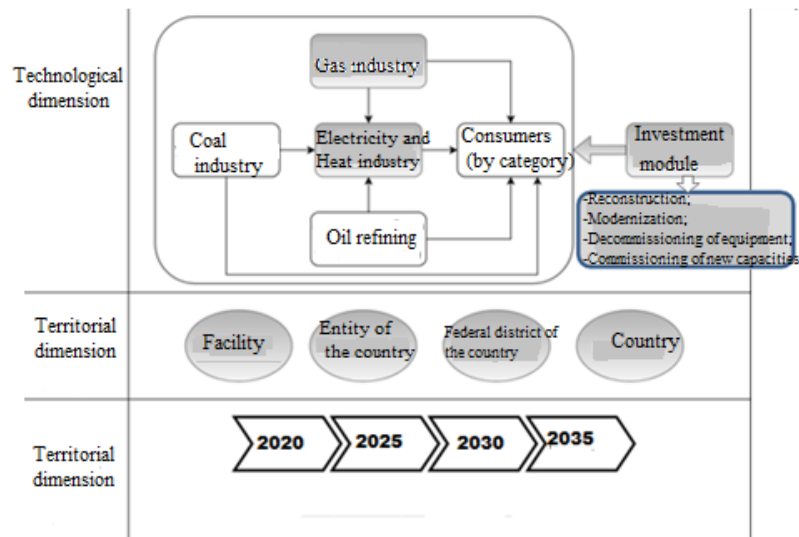


Figure 2. Territorial, temporal and technological structure of models

The mathematical description of the model is represented by balance equations and constraints on variables and the corresponding objective function.

$$AX - \sum_{t=1}^T y^t = 0, \quad (1)$$

$$0 \leq X \leq D, \quad (2)$$

$$0 \leq Y^t \leq R^t, \quad (3)$$

where t is consumer category; X is the desired vector, with the components characterizing the intensity of using technological methods for operation of energy facilities (extraction, processing, conversion and transportation of energy resources); Y^t is the desired vector with the components characterizing the volumes of individual types of fuel and energy consumed by certain categories of consumers (t); A is matrix of input-output ratios of production (extraction, processing, conversion) and transportation of individual types of fuel and energy (inputs - output); D is a vector that determines the technically possible intensity of using individual technological and production

methods; R^t is a vector with components equal to the volumes of specified consumption of individual types of fuel and energy by individual categories of consumers.

The objective function has the following form:

$$(C, X) + \sum_{t=1}^T (r^t, g^t) \rightarrow \min. \quad (4)$$

The first component of such an objective function reflects the costs associated with the operation of the industries within the energy sector, its constituent energy systems and subsystems, and capital investments for their development. Here C is the vector of unit costs for individual technological methods of operation of existing, reconstructed or modernized, as well as newly constructed energy facilities.

The second component is the damage from shortages for each type of fuel and energy for each of the selected consumer categories. The magnitude of the energy resource shortage (g^t) for consumers of category t corresponds to the difference ($R^t - Y^t$). Vector r^t consists of components conventionally called "specific damages." The cost assessment of the real (full) amount of damage caused by a shortage poses certain difficulties due to the various manifestations of the consequences of a shortage of energy resources, which cannot always be identified and quantified. In this case, this difficulty is (rather conventionally) overcome by introducing a scale of priorities in meeting the demand of the consumer of the categories at issue for certain fuel and energy types.

The final implementation of the models includes a financial module that describes the investment costs for reconstruction, modernization of existing facilities, decommissioning of obsolete equipment, and commissioning of new facilities at energy facilities. With these models, it is also possible to take account of the development dynamics, which allows tracking such features of the multi-stage development process of the energy sector as:

- commissioning of new production facilities;
- dismantling and conservation of old facilities,
- reconstruction of facilities with a change in the flow diagram.

Consideration of dynamics is implemented in the form of T independent static modules, each of which describes all the territorial and technological links of the energy sector for stage t of the considered period. Dynamic connections between modules are built using equations that formulate for all x_i facilities of the energy sector the condition for the continuity of their productive capacities at various stages of the considered period. This condition for the first stage is written as

$$x_{i1}^o + x_{i1}^c + x_{i1}^d = P_{i0}, \quad (5)$$

and for subsequent stages, it is written in the form of equations

$$x_{it-1}^o + x_{it-1}^c + x_{it-1}^n = x_{it}^o + x_{it}^c + x_{it}^d, \quad (6)$$

where P_{i0} is productive capacity of technology (facility i) by the beginning of the considered period,

x_{it-1}^n is productive capacity of a new part of technology (facility i) in stage $t-1$,

x_{it}^o is productive capacity of the operating part of the technology (facility i) in stage t ,

x_{it}^c is conservation of part of facility i in stage t ,

x_{it}^d is liquidation of part of facility i in stage t .

For the convenience of building the connections, equation (6) is divided into two parts

$$\begin{aligned} -x_{it-1}^o - x_{it-1}^c - x_{it-1}^n + Z_{it-1} &= 0, \\ -Z_{it-1} + x_{it}^o + x_{it}^c + x_{it}^d &= 0, \end{aligned}$$

where Z_{it-1} is an intermediate variable characterizing the overall performance of facility i at the beginning of stage t . It takes into account the retirement of capacities in stage t and the

introduction of new capacities in the stage $t+1$.

In general, these models are used to determine the following characteristics (indices):

- the size of undersupply (shortage) in certain types of energy resources for the categories of consumers at issue, selected territorial entities and the entire country, as the value of discrepancy between set demand and feasibility of producing this type of energy resource (considering such factors as the reserves, the possibilities of replacing this type of energy resource in other consumers, etc.);

- changes in the capacity of transportation links, which are determined by comparing the relevant indicators of the considered option with the original one;

- rational use of production capacities of energy facilities, and the distribution of certain types of energy resources according to selected categories of consumers.

The backbone module in the energy sector model is the one of electricity and heat industry, therefore, correct modeling of the facilities that constitute the system boosts the adequacy of the model.

The developed modeling system consists of industry models connected by information flows and an integrating model of the energy sector as a whole. This system makes it possible to enhance the existing practice of assessing the materialization of energy security threats through:

- identifying the mutual influence of energy systems on each other and comprehensively assessing the impact of energy security threats;

- considering the dynamics of the development of critical situations in the models of optimization of energy systems within the energy sector;

- taking into account the load curve, which imposes requirements of the consumer on industry systems;

- taking into consideration nonlinearity in terms of the adequacy of the representation of processes in energy systems to improve the accuracy of decisions made (in models of industry systems);

- allowing for natural factors in terms of their impact on the operation of renewable energy sources (periods of low water for hydropower plants, cloudy days for solar power plants, low-wind periods for wind turbines);

- considering the influence of inertia in the gas, coal, and oil refining industries on the development of emergencies in them and relate them with their unfolding in the electricity industry, which is a backbone for the energy sector.

- taking into account the specific features of the mutual influence of the gas and electricity industries when meeting the conditions for the reliability of gas and electricity supply to the production facilities of these industries;

- taking into consideration the specific features of the mutual influence of the oil/oil refining and electricity industries when meeting the conditions for the reliability of electricity and oil products supply to the production facilities of these industries.

III. Case study. Assessment of a shortage caused by critical situations in energy systems.

There are many energy security threats of which the natural and technology-driven are the most common ones. We will analyze the materialization of several of them in a real energy sector.

For countries with harsh climatic conditions, a particularly urgent threat is the threat of a sharp cooling. Critical situations in the fuel and energy supply arise from the rapid and widespread cooling that can envelop vast areas of the country. At the same time, depending on the climatic conditions of a particular region and the type of consumers, the maximum seasonal heating loads

can deviate from the average annual values by a significant amount, up to 20-30%. The threat of a cold snap is extremely relevant for Russia. Modeling and analysis of the threat of a sharp cooling and its implications will be made on the example of the Russian energy sector. The calculations assumed a decrease in the average outdoor air temperature during one quarter of the heating season in the European part of Russia by 2°C versus the long-term average, which will lead to an increase in the demand for boiler and furnace fuel by about 8% (Critical situation (CS)1).

One of the most dangerous (in terms of consequences) situations in the gas supply system is the possibility of damage to transcontinental gas pipelines. This threat is especially acute with a large share of gas in the production of electricity and heat, which is typical of some regions of Russia. The possibility of failure of one of the main gas pipeline sections was considered as the design conditions for a critical situation in the gas supply system. This situation, given the restoration work leads to a decline in gas supply to some regions of Russia by 5% per quarter (CS 2).

In the oil system, it is important to analyze the impact of a decrease in the fuel oil supply from regions where large oil refining capacities are concentrated and possible complications of a various nature, including socio-political ones. At the same time, disturbances were introduced into the model in the form of a decrease in the fuel oil supply by 8% of its total production during the analyzed period (CS 3).

In coal supply, a potentially dangerous situation is when a high proportion of coal comes from one source. In this case, the design conditions provide for a 30% reduction in coal supplies to power plants in one of Russia's regions and a similar reduction in coal supplies from another region (CS 4).

Currently, one of the dangerous factors for reliable fuel and energy supply is the imbalance of some regional electric power systems. Therefore, consideration was given to the consequences of the rupture of backbone ties in the electric power industry. In addition to this, a possible reduction of 30% in the capacity of nuclear power plants in one of the energy systems was introduced (CS 5).

Furthermore, the issue of potential overlap between the aforementioned disruptions (CS 6) and the potential utilization of additional reserves of fuel oil and coal equivalent to a 10-day demand (CS 7) was considered. While it is highly improbable for all critical situations to overlap, this circumstance provides a valuable opportunity to evaluate the limitations of the energy sector in meeting the fuel and energy demands of consumers. In addition, it emphasizes the importance of mutual reservation of energy systems and regions in the face of a global deterioration in energy conditions. The diagram for the formation of emergency situations is shown in Fig. 3.

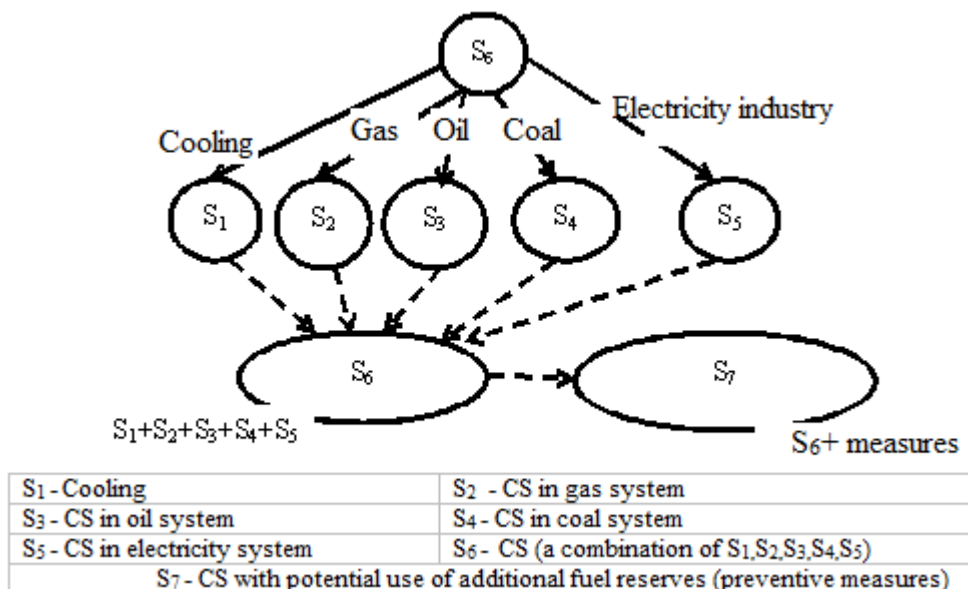


Figure 3. Scheme of formation of critical situations.

An analysis of calculations based on the generated scenarios of critical situations indicates that the most significant situations in terms of their consequences for consumers of energy resources proved to be those with cooling (CS 1) and mutual overlap of critical situations (CS 6). A good example is the shortage observed in the case of a decrease in fuel supplies and other disturbances (CS 1-CS7) (Fig. 4).

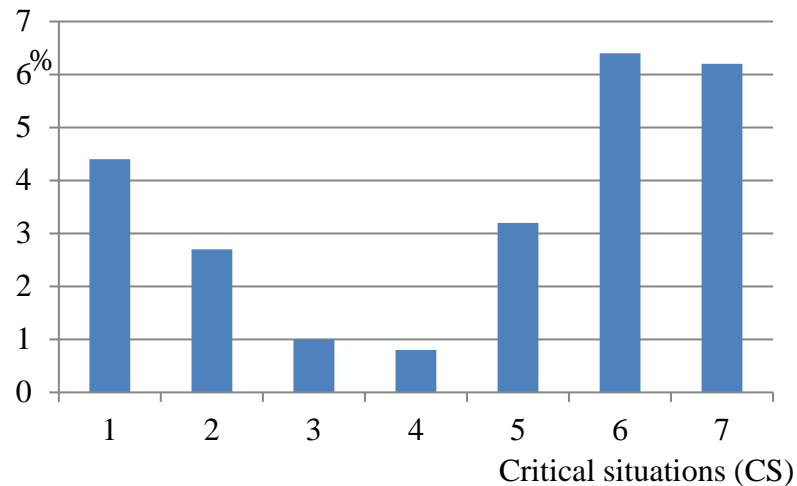


Figure 4. Electricity shortage in critical situations №№ 1-7, % of the needs.

Disturbing impacts caused by a combination of critical situations (CS6) had a greater impact on the systems of coal, electricity, and heat supply. The overall shortage of coal in the country amounted to about 11% of its total consumption, electricity and heat shortage was about 6% (Fig.5).

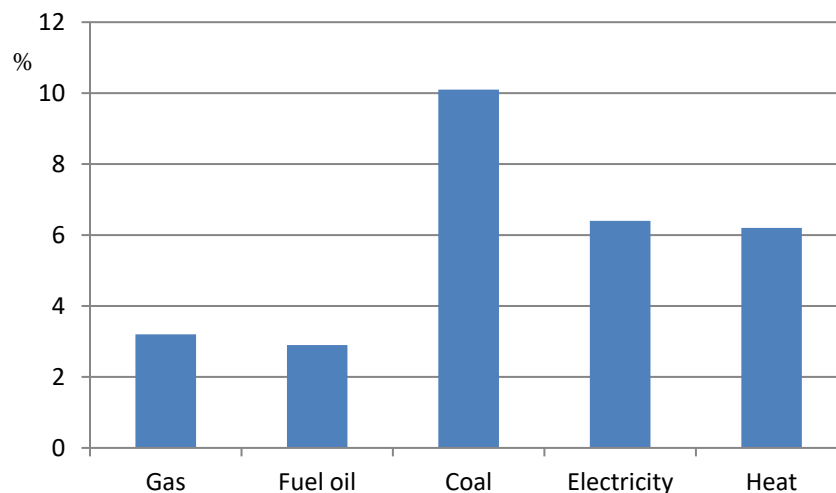


Figure 5. Shortage of energy resources in CS, % of the needs.

There is no shortage of resources in gas and fuel oil systems, but there has been a decrease in gas and fuel oil consumption at power plants and boiler houses. This led to a greater consumption of coal for the production of heat and electricity, resulting in a significant shortage of this fuel. This is explained by the fact that the closing type of fuel in this implementation of the model is coal.

The cooling and rupture of links in the electric power system and the reduction in the power of nuclear power plants, assumed in the calculations (CS 5), resulted in a shortage of electricity throughout country by about 7% (with the full use of the available backup generating capacities of

thermal power plants).

The introduction of additional fuel reserves (coal and fuel oil) into the model in the amount of a 10-day demand (CS 7) reduced the shortage of coal to 10%. This was achieved by redistributing electricity generation between gas-oil and coal-fired thermal power plants, i.e., additional fuel oil resources were used at thermal power plants, which freed up coal to partially compensate for the coal shortage of consumers in non-energy industries.

The involvement of additional fuel reserves in the supply of electricity and heat did not cause changes and the shortage amounted to the same value. This is explained by the fact that all the reserves of thermal power plants were used to compensate for the growth in demand for energy resources and the decrease in the power of nuclear power plants. In this case, the increase in fuel resources without additional commissioning of generating capacities of power plants did not lead to additional generation of electrical and thermal energy.

In the context of Critical Situation 6 (CS 6), coal shortage was observed in almost all regions. However, after the creation of 10-day fuel reserves, this shortage has decreased in many regions.

The performed experimental calculations have demonstrated the effectiveness of energy security research based on modeling of the energy security threats. A preliminary assessment of the fuel and energy supply to consumers under various critical situations showed a rather high degree of sensitivity of the model to changes in parameters and the possibility of its effective use for this kind of research.

IV. Conclusion

Assessing the energy security and reliability of fuel and energy supply in the context of current social and economic development in different countries is crucial and highly relevant. This is because the functioning of all life-sustaining systems and structures relies heavily on a reliable energy supply.

The energy security research requires an adequate dedicated modeling system for analyzing the interrelated operation of energy systems in the event of materialization of energy security threats. The paper presents a methodological framework of such a system and the methodological features of modeling the interlinked work of industries in today's context. The developed system is designed to conduct experimental research to find ways to provide consumers with energy resources without a shortage when functioning under normal conditions and in critical situations.

The experimental part illustrates the results of case studies on the application of the presented methodological tool to real energy systems with the view to analyzing several different critical situations (energy security threats). These are a severe drop in temperature, damage to transcontinental gas pipelines, a decrease in the supply of fuel oil, the termination of coal supply from the dominant source, rupture of backbone connections in the electric power system, overlap of all the listed situations and overlap with one compensatory measure. The findings of all cases show the high performance and efficiency of the proposed methodological tool for the energy security research.

Acknowledgment. The research was carried out under State Assignment Project no. FWEU-2021-0003 (reg. no. AAAA-A21-121012090014-5) of the Fundamental Research Program of Russian Federation 2021-2030

References

[1] Bushuev V.V., Voropai N.I., Mastepanov A.M. et al. Energy security of Russia. Novosibirsk: Nauka, Siberian publ. comp. of RAS 1998.

- [2]. N.I. Pyatkova, V.I. Rabchuk, S.M. Senderov et al. Energy security of Russia: problems and solutions. Novosibirsk: Nauka 2011.
- [3]. The Routledge Handbook of Energy Security. Edited by Benjamin K. Sovacool. Taylor & Francis Group. London and New York 2011.
- [4]. Senderov S.M. et al. Reliability of fuel and energy supply to consumers from the energy security perspective. Ed. by N.I. Voropai; ESI SB RAS. Novosibirsk: SB RAS 2022.
- [5]. Systems research in the energy sector: Retrospective of scientific areas of SEI-ESI. Ed. by Voropai N.I.. Novosibirsk: Nauka 2010.
- [6]. Belyaev L.S., Voropai N.I., Kononov Yu.D. et al. Methods of studying and managing energy systems. Novosibirsk: Nauka 1987.
- [7]. Antonov G. N., Cherkesov G. N., Krivorutsky L. D. et al. Methods and models for assessing the survivability of energy systems. Novosibirsk: Nauka 1990.
- [8]. Hierarchical modeling of energy systems. Ed. by Voropai N. I., Stennikov V. A.. Elsevier Inc. 2023. ISBN: 978-0-443-13917-8. Available at: <https://doi.org/10.1016/B978-0-44-313917-8.00011-0> (accessed 22.04.2013).
- [9]. Voropai N.I., Klimenko S.M., Krivorutsky L.D., Massel L.V., Pyatkova N.I., Senderov S.M., et al. Energy security of Russia (introduction to the problem): Preprint No. 1. Irkutsk: SEI SB RAS 1997.
- [10]. Voropai N.I., Klimenko S.M., Kovalev G.F., Krivorutsky L.D., Pyatkova N.I., Senderov S.M., et al. Basic provisions and methodology for monitoring and indicator-based analysis of the energy security of Russia and its regions: Preprint No. 4. - Irkutsk: ESI SB RAS 1998.
- [11]. Senderov S.M., Pyatkova N.I. et al. Methods for monitoring the state of Russia's energy security at the regional level. Preprint No. 4. - Irkutsk: ESI SB RAS 2014.
- [12]. Pyatkova N.I., Rabchuk V.I., Senderov S.M. et al. Energy security of Russia: problems and solutions. Novosibirsk: Nauka 2011.
- [13]. Senderov S.M., Rabchuk V.I., Pyatkova N.I., Vorobyov S.V.. Ensuring the energy security of Russia: the choice of priorities. Ed. by Senderov S.M.. Melentiev Energy Systems Institute SB RAS. Novosibirsk 2017.
- [14]. Senderov S.M. and others. Reliability of fuel and energy supply to consumers from the energy security perspective. ESI SB RAS. Novosibirsk: SB RAS 2022.
- [15]. Song Y., Zhang M., Sun R. Using a new aggregated indicator to evaluate China's energy security. Energy Policy 2019, 132. <https://doi.org/10.1016/j.enpol.2019.05.036>
- [16]. Karatayev M., Hall S. Establishing and comparing energy security trends in resource-rich exporting nations (Russia and the Caspian Sea region). Resources Policy, 2020, 68, 101746. <https://doi.org/10.1016/j.resourpol.2020.101746>
- [17]. Peng Li, Zixuan Wang, Weihong Yang, Haitao Liu, Yunxing Yin, Jiahao Wang et al.. Hierarchically partitioned coordinated operation of distributed integrated energy system based on a master-slave game. Energy 2021. <https://doi.org/10.1016/j.energy.2020.119006>
- [18]. Y. Chen, W. Wei, F. Liu, Q. Wu, S. Mei. Analyzing and validating the economic efficiency of managing a cluster of energy hubs in multi-carrier energy systems. Applied Energy 2018. 230, 403-416.
- [19]. Suslov K., Piskunova V., Gerasimov D. et al.. Development of the methodological basis of the simulation modelling of the multi-energy systems. E3S Web of Conferences 124, 01049 (2019). <https://doi.org/10.1051/e3sconf/201912401049>
- [20]. Dokic S.B., Rajakovic N.L. Security modelling of integrated gas and electrical power systems by analyzing critical situations and potentials for performance optimization. Energy 2019; 184, pp. 141–150, Available at: <https://doi.org/10.1016/j.energy.2018.04.165> (accessed 22.04.2013).

- [21]. Kozlov M.V., Malashenko Yu.E., Nazarova I.A. Management of fuel and energy system in case of large-scale damage. I. Network model and software implementation. Proceedings of the Russian Academy of Sciences. Theory and control systems 2017. No 6. pp. 50–73.
- [22]. Malashenko Yu. E., Nazarova I. A., Novikova N. M. Management of fuel and energy system in case of large-scale damage. II. Statements of optimization problems. Proceedings of the Russian Academy of Sciences. Theory and control systems 2018; No. 2, pp. 39-51.
- [23]. Malashenko Yu.E., Nazarova I.A., Novikova N.M. A method for analyzing the functional vulnerability of flow network systems. Informatics and its applications 2017; Vol. 11, No. 4, pp. 47-54.
- [24]. Malashenko Yu.E., Nazarova I.A., Novikova N.M. Vulnerability diagrams of streaming network systems. Informatics and its Applications 2018; Vol. 12, No. 1, pp. 11-17.
- [25]. Malashenko Y.E., Nazarova I.A., Novikova N.M. Fuel and energy system control at large-scale damage. IV. A Prior estimates of structural and functional vulnerability. Journal of Computer and Systems Sciences International 2018; Vol. 57. No.6, pp. 907-920.
- [26]. Malashenko Y.E., Nazarova I.A. [Analysis of critical damage in the communication network. IV: Multicriteria estimations of cluster vulnerability. Journal of Computer and Systems Sciences International](#) 2021; Vol. 60. No.6. pp. 956-965.
- [27]. Grebenyuk G.G., Nikishov S.M. Blocking energy and resource supplies to target objects in network infrastructures. Control Sciences 2016; No. 4, pp. 52-57.
- [28]. Dan Gao, Zheng Li, Pei Liu, Jiazhu Zhao, Yuning Zhang, Canbing Li. A coordinated energy security model taking strategic petroleum reserve and alternative fuels into consideration. Energy 2018; pp. 171-181. Available at: <https://doi.org/10.1016/j.energy.2017.11.097> (accessed 22.04.2013).
- [29]. Seebregts Ad J., Goldstein Gary A., Smekens Koen. Energy/Environmental Modeling with the MARKAL Family of Models [On-line] Proc. Int. Conf. on Operations Research (OR 2001), Energy and Environment Session, September 3-5, 2001, Duisburg, Germany. Available at: <ftp://ftp.ecn.nl/pub/www/library/report/2001/rx01039.pdf> (accessed 22.04.2013).
- [30]. MARKAL [online] IEA-ETSAP: The Energy Technology Systems Analysis Programme (ETSAP). 2011. Available at: <http://www.iea-etsap.org/web/Markal.asp> (accessed 22.04.2013).
- [31]. MESSAGE (Model for Energy Supply Strategy Alternatives and their General Environmental Impact) [Online] International Institute for Applied Systems Analysis (IIASA) [Laxenburg, Austria]. Available at: <http://www.iiasa.ac.at/web/home/research/researchPrograms/Energy/MESSAGE.en.html> (accessed 22.04.2013).
- [32]. Fedorova E.V., Zorina T.G.. Interregional Energy Exchange Modeling by Using MESSAGE Cod. Communications of Higher Schools. Nuclear Power Engineering. 2004; No.4, pp. 3-11.
- [33]. The National Energy Modeling System: An Overview 2009 [online] –Available at: <http://www.eia.doe.gov/oiaf/aeo/overview/index.html> (accessed 22.04.2013).
- [34]. The Energy Research Institute of the Russian Academy of Sciences is 25 years old/Ed. by Makarov A.A.. M.: ERI RAS. 2010.
- [35]. Pyatkova N.I., Senderov S.M., Pyatkova E.V. Methodological features of the study on energy security problems at the present stage. Proceedings of the Russian Academy of Sciences. Power Engineering 2014; No. 2, pp.81-87.

ESTIMATION OF PARAMETERS FOR KUMARASWAMY EXPONENTIAL DISTRIBUTION BASED ON PROGRESSIVE TYPE-I INTERVAL CENSORED SAMPLE

*MANOJ CHACKO AND SHILPA S DEV

Department of Statistics, University of Kerala, Trivandrum-695581, India
*manojchacko02@gmail.com; sdevshilpa0246@gamil.com

Abstract

In this paper, we consider the problem of estimation of parameters of the Kumaraswamy exponential distribution using progressive type-I interval censored data. The maximum likelihood estimators (MLEs) of the parameters are obtained. As it is observed that there is no closed-form solutions for the MLEs, we implement the Expectation-Maximization (EM) algorithm for the computation of MLEs. Bayes estimators are also obtained using different loss functions such as the squared error loss function and the LINEX loss function. For the Bayesian estimation, Lindley's approximation method has been applied. To evaluate the performance of the various estimators developed, we conduct an extensive simulation study. The different estimators and censoring schemes are compared based on average bias and mean squared error. A real data set is also taken into consideration for illustration.

Keywords: Maximum likelihood estimate, EM algorithm, Bayesian inference, Lindley's approximation

1. INTRODUCTION

In life testing experiment and survival analysis, the test units may leave the experiment before failure due to restriction of time, budget cost or accidental breakage. A censored sample refers to data that was gathered from such cases but may not be complete. Over the last few decades, a number of censoring methodologies have been developed for the analysis of such situations. In the exiting literature, two commonly used traditional censoring schemes are type-I and type-II, in which experiment is terminated after a prescribed time point and number of failures, respectively. However, neither of these two censoring strategies permit the experimenter to remove live units from the experiment prior to its termination time. To remove the units in between the experiments, the idea of progressive censoring was developed by [7]. It is further observed that, in many practical situations it is not possible for the experimenter to continuously observe the life test units to observe the precise failure lifetimes. For example, in medical and clinical trials, specific information regarding the patient survival lifetime for those diagnosed with a particular treatment may not be available. In such cases, the failure lifetimes are often observed in the intervals, known as interval censoring. However, this censoring does not allow to remove the units in between the experiments. The concept of progressive type-I interval censoring, incorporating the principles of type-I, progressive, and interval censoring schemes, was introduced by [2]. In this type of censoring, items can be withdrawn between two successive time points that have been prescheduled.

The progressive type-I interval censored sample is gathered in the following manner. Assume that n units are placed on a life test at the time $t_0 = 0$. Units are inspected at m predefined times t_1, t_2, \dots, t_m , with t_m being the experiment's scheduled finish time. At the i th inspection

time $t_i, i = 1, \dots, m$, the number X_i , of failures within $(t_{i-1}, t_i]$ is recorded and R_i surviving units are randomly removed from the life test. The number of surviving units at time t_1, \dots, t_m is a random variable, hence the number of removals R_1, \dots, R_m can be estimated as a percentage of the remaining surviving units. Specifically, $\lfloor q_i \times (\text{number of surviving units at time } t_i) \rfloor$ remaining surviving units are eliminated from the life test with pre-specified values of q_1, \dots, q_{m-1} and $q_m = 1$, where $\lfloor w \rfloor$ = the largest integer less than or equal to w . Alternatively, R_1, R_2, \dots, R_m can be pre-specified non-negative integers, with $R_i^{obs} = \min(R_i, \text{number of surviving units at time } t_i), i = 1, 2, \dots, m - 1$, and $R_m^{obs} = \text{number of surviving units at time } t_m$. Data observed under this censoring scheme can be represented as $(X_i, R_i, t_i)_{i=1}^m$. If $F(x, \theta)$ is the cumulative distribution function (cdf) of the population from which the progressive type-I censored sample is taken, then the likelihood function of θ can be constructed as follows (see, [2])

$$L(\theta) \propto \prod_{i=1}^m [F(t_i, \theta) - F(t_{i-1}, \theta)]^{X_i} [1 - F(t_i, \theta)]^{R_i} , \quad (1)$$

where $t_0 = 0$.

In the recent past, several authors studied progressive type-I interval censored sampling schemes under various circumstances. The maximum likelihood estimates of the parameters of the exponentiated Weibull family and their asymptotic variances were obtained by [4]. Optimally spaced inspection times for the log-normal distribution were determined by [12], while different estimation methods based on progressive type-I interval censoring were considered for the Weibull distribution by [17] and for the Generalized exponential distribution by [6]. The statistical inference under this censoring for Inverse Weibull distribution was further discussed by [19]. Bayesian inference under this censoring has been discussed by [3] for Dagum distribution.

In this paper, we consider progressive type-I interval censored sample taken from a Kumaraswamy exponential (KE) distribution with probability density function (pdf) given by

$$f(x) = \beta \lambda e^{-x} (1 - e^{-x})^{\beta-1} (1 - (1 - e^{-x})^\beta)^{\lambda-1}, x > 0 . \quad (2)$$

The cdf corresponding to the above pdf is given by

$$F(x) = 1 - (1 - (1 - e^{-x})^\beta)^\lambda, x > 0 , \quad (3)$$

where $\beta > 0, \lambda > 0$ are two shape parameters. Through out the paper, we use the notation $KE(\beta, \lambda)$ to denote Kumaraswamy exponential distribution with shape parameters β and λ . The KE distribution is a generalisation of the exponential distribution that was created as a model for issues in environmental studies and survival analysis. Several studies on Kumaraswamy distribution and its generalisations have been published in recent years. An exponentiated Kumaraswamy distribution and its properties were considered and discussed by [11]. The Kumaraswamy linear exponential distribution with four parameters was introduced by [9], who also derived some of its mathematical properties. The maximum likelihood estimation of the unknown parameters for the Kumaraswamy exponential distribution was considered by [1]. The exponentiated Kumaraswamy exponential distribution and its characterization properties were introduced by [18]. The estimation of parameters for the Kumaraswamy exponential distribution under a progressive type-II censored scheme was considered by [5].

The structure of this paper is outlined as follows. The maximum likelihood estimators of $KE(\beta, \lambda)$ parameters are obtained in Section 2. In this section, estimators are also obtained using EM algorithm. In Section 3, Bayes estimates for β and λ are obtained for different loss functions such as squared error and LINEX. Here, Lindley's approximation method is used to evaluate these Bayes estimates. In Section 4, a simulation study is carried out for analysing the properties of various estimators developed in this paper. In Section 5, a real data is considered for illustration. Finally, in Section 6, we present some concluding remarks.

2. MAXIMUM LIKELIHOOD ESTIMATION

Let $(X_i, R_i, t_i), i = 1, 2, \dots, n$ be a progressively type-I interval censored sample taken from the $KE(\beta, \lambda)$ distribution defined in (2), then by using (1), the likelihood function is given by

$$L(\beta, \lambda) \propto \sum_{i=1}^m \left[\left[1 - (1 - e^{-t_{i-1}})\beta \right]^\lambda - \left[1 - (1 - e^{-t_i})\beta \right]^\lambda \right]^{X_i} \left[\left[1 - (1 - e^{-t_i})\beta \right]^\lambda \right]^{R_i}. \quad (4)$$

Then the log-likelihood function is given by

$$l(\beta, \lambda) = \ln L(\beta, \lambda) = \sum_{i=1}^m X_i \ln \left[\left[1 - (1 - e^{-t_{i-1}})\beta \right]^\lambda - \left[1 - (1 - e^{-t_i})\beta \right]^\lambda \right] + \sum_{i=1}^m R_i \ln \left[\left[1 - (1 - e^{-t_i})\beta \right]^\lambda \right]. \quad (5)$$

The MLEs of β and λ are the solutions to the following normal equations

$$\sum_{i=1}^m \frac{R_i \lambda [1 - Z_i^\beta]^\lambda Z_i^\beta \ln Z_i}{[1 - Z_i^\beta]^\lambda} = - \sum_{i=1}^m \frac{X_i \left[\lambda [1 - Z_i^\beta]^{\lambda-1} Z_i^\beta \ln Z_i - \lambda [1 - Z_{i-1}^\beta]^{\lambda-1} Z_{i-1}^\beta \ln Z_{i-1} \right]}{[(1 - Z_i^\beta)^\lambda - (1 - Z_{i-1}^\beta)^\lambda]} \quad (6)$$

and

$$\sum_{i=1}^m \frac{R_i \lambda [1 - Z_i^\beta]^\lambda \ln(1 - Z_i^\beta)}{[1 - Z_i^\beta]^\lambda} = - \sum_{i=1}^m \frac{X_i \left[(1 - Z_i^\beta)^\lambda \ln(1 - Z_i^\beta) - (1 - Z_{i-1}^\beta)^\lambda \ln(1 - Z_{i-1}^\beta) \right]}{[(1 - Z_i^\beta)^\lambda - (1 - Z_{i-1}^\beta)^\lambda]}, \quad (7)$$

where $Z_i^\beta = (1 - e^{-t_i})$.

As the above equations have no closed form solutions, the MLEs can be obtained through an iterative numerical methods such as Newton-Raphson method. Since the MLEs are obtained using numerical method, in the following subsection, the EM algorithm is used to find the MLEs of β and λ .

2.1. EM Algorithm

The Expectation-Maximization (EM) algorithm is a broadly applicable method of iterative computing of maximum likelihood estimates and useful in a variety of incomplete-data scenarios where methods like the Newton-Raphson method may prove to be more difficult. The expectation step, also known as the E-step, and the maximisation step, often known as the M-step, are two steps that comprise each iteration of the EM algorithm. Therefore, the algorithm is known as the EM algorithm, and its detailed development can be found in [8]. The EM algorithm for finding MLEs of the parameter of the two-parameter Kumaraswamy exponential distribution is as follows.

Let $\psi_{i,j}, j = 1, 2, \dots, X_i$, be the survival times of the units failed within subinterval $(t_{i-1}, t_i]$ and $\psi_{i,j}^*, j = 1, 2, \dots, R_i$ be the durations of survival for those units withdrawn at t_i for $i = 1, 2, 3, \dots, m$, then the log likelihood function, $\ln(L^c)$, based on the lifetimes of all n items (complete sample) from the two-parameter $KE(\beta, \lambda)$ distribution is given by

$$\ln(L^c) = \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \log(f(\psi_{i,j}, \theta)) + \sum_{j=1}^{R_i} \log(f(\psi_{i,j}^*, \theta)) \right],$$

$$\begin{aligned} \ln(L^c) = & [\ln(\beta) + \ln(\lambda)] \sum_{i=1}^m [X_i + R_i] - \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \psi_{i,j} + \sum_{j=1}^{R_i} \psi_{i,j}^* \right] + \\ & (\beta - 1) \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \ln(1 - e^{-\psi_{i,j}}) + \sum_{j=1}^{R_i} \ln(1 - e^{-\psi_{i,j}^*}) \right] + \\ & (\lambda - 1) \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \ln \left[1 - (1 - e^{-\psi_{i,j}})^\beta \right] + \sum_{j=1}^{R_i} \ln \left[1 - (1 - e^{-\psi_{i,j}^*})^\beta \right] \right], \end{aligned} \quad (8)$$

where $\sum_{i=1}^m (X_i + R_i) = n$

Taking the derivatives of (8) with respect to β and λ , respectively, the following normal equations are obtained:

$$\begin{aligned} \frac{n}{\beta} = & (\lambda - 1) \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \frac{(1 - e^{-\psi_{i,j}})^\beta \ln(1 - e^{-\psi_{i,j}})}{[1 - (1 - e^{-\psi_{i,j}})^\beta]} \right] + \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \frac{(1 - e^{-\psi_{i,j}^*})^\beta \ln(1 - e^{-\psi_{i,j}^*})}{[1 - (1 - e^{-\psi_{i,j}^*})^\beta]} \right] \\ & - \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \ln(1 - e^{-\psi_{i,j}}) + \sum_{j=1}^{R_i} \ln(1 - e^{-\psi_{i,j}^*}) \right] \end{aligned} \quad (9)$$

and

$$\frac{n}{\lambda} = - \sum_{i=1}^m \left[\sum_{j=1}^{X_i} \ln \left[1 - (1 - e^{-\psi_{i,j}})^\beta \right] + \sum_{j=1}^{R_i} \ln \left[1 - (1 - e^{-\psi_{i,j}^*})^\beta \right] \right]. \quad (10)$$

The lifetimes of X_i failures in the i^{th} interval $(t_{i-1}, t_i]$ are independent and follow a doubly truncated Kumaraswamy exponential distribution from left at t_{i-1} and right at t_i , while the lifetimes of the R_i censored items at the time t_i are independent and follow a truncated Kumaraswamy exponential distribution from the left at $t_i, i = 1, 2, \dots, m$.

For the EM algorithm, the following expected values of a doubly truncated Kumaraswamy exponential random variable Y , from a on the left and b on the right with $0 < a < b \leq \infty$ are needed.

$$E_{\beta,\lambda} \left[\ln(1 - e^{-Y}) | Y \in [a, b] \right] = \int_a^b \frac{\ln(1 - e^{-y}) f(y; \beta, \lambda) dy}{F(b; \beta, \lambda) - F(a; \beta, \lambda)},$$

$$E_{\beta,\lambda} \left[\ln \left[1 - (1 - e^{-Y})^\beta \right] | Y \in [a, b] \right] = \int_a^b \frac{\ln \left[1 - (1 - e^{-y})^\beta \right] f(y; \beta, \lambda) dy}{F(b; \beta, \lambda) - F(a; \beta, \lambda)}$$

and

$$E_{\beta,\lambda} \left[\frac{(1 - e^{-Y})^\beta \ln(1 - e^{-Y})}{[1 - (1 - e^{-Y})^\beta]} | Y \in [a, b] \right] = \int_a^b \frac{\frac{(1 - e^{-y})^\beta \ln(1 - e^{-y})}{[1 - (1 - e^{-y})^\beta]} f(y; \beta, \lambda) dy}{F(b; \beta, \lambda) - F(a; \beta, \lambda)}.$$

The iterative process that results in the EM algorithm is as follows:

Step 1: Given starting values of β and λ , say $\beta^{(0)}$ and $\lambda^{(0)}$ and set $k=0$.

Step 2: In the $(k + 1)^{th}$ iteration, the following conditional expectations are computed by the E-step. For $i = 1, 2, \dots, m$

$$\begin{aligned}
 E_{1i} &= E_{\beta^{(k)}, \lambda^{(k)}} \left[\ln(1 - e^{-Y}) | Y \in [t_{i-1}, t_i] \right], \\
 E_{2i} &= E_{\beta^{(k)}, \lambda^{(k)}} \left[\ln(1 - e^{-Y}) | Y \in [t_i, \infty) \right], \\
 E_{3i} &= E_{\beta^{(k)}, \lambda^{(k)}} \left[\ln[1 - (1 - e^{-Y})^{\hat{\beta}^{(k)}}] | Y \in [t_{i-1}, t_i] \right], \\
 E_{4i} &= E_{\beta^{(k)}, \lambda^{(k)}} \left[\ln[1 - (1 - e^{-Y})^{\hat{\beta}^{(k)}}] | Y \in [t_i, \infty) \right], \\
 E_{5i} &= E_{\beta^{(k)}, \lambda^{(k)}} \left[\frac{(1 - e^{-Y})^{\hat{\beta}^{(k)}} \ln(1 - e^{-Y})}{[1 - (1 - e^{-Y})^{\hat{\beta}^{(k)}}]} | Y \in [t_{i-1}, t_i] \right]
 \end{aligned}$$

and

$$E_{6i} = E_{\beta^{(k)}, \lambda^{(k)}} \left[\frac{(1 - e^{-Y})^{\hat{\beta}^{(k)}} \ln(1 - e^{-Y})}{[1 - (1 - e^{-Y})^{\hat{\beta}^{(k)}}]} | Y \in [t_i, \infty) \right].$$

Then, the likelihood equations (9) and (10) are respectively given by

$$\frac{n}{\beta} = (\lambda - 1) \sum_{i=1}^m [X_i E_{5i} + R_i E_{6i}] - \sum_{i=1}^m [X_i E_{1i} + R_i E_{2i}] \quad (11)$$

and

$$\frac{n}{\lambda} = - \sum_{i=1}^m [X_i E_{3i} + R_i E_{4i}]. \quad (12)$$

Step 3: The M-step requires to solve the equations (11) and (12) and obtains the next values, $\beta^{(k+1)}$ and $\lambda^{(k+1)}$, of β and λ , respectively, as follows:

$$\beta^{(k+1)} = \frac{n}{(\hat{\lambda}^{(k+1)} - 1) \sum_{i=1}^m [X_i E_{5i} + R_i E_{6i}] - \sum_{i=1}^m [X_i E_{1i} + R_i E_{2i}]}$$

and

$$\lambda^{(k+1)} = - \frac{n}{\sum_{i=1}^m [X_i E_{3i} + R_i E_{4i}]}.$$

Step 4: Checking for convergence; if convergence happens, then the current $\beta^{(k+1)}$ and $\lambda^{(k+1)}$ are the approximated maximum likelihood estimates of β and λ via EM algorithm. If the convergence doesn't happens, then set $k = k + 1$ and go to step 2.

3. BAYESIAN ESTIMATION

In this section, Bayesian estimation of parameters of $KE(\beta, \lambda)$ are obtained under both symmetric and assymetric loss functions.

The squared error is a symmetric loss function and is defined as

$$L_1(\delta, \hat{\delta}) = (\hat{\delta} - \delta)^2,$$

where $\hat{\delta}$ is the estimate of parameter δ .

An asymmetric loss function is the LINEX loss function, defined as

$$L_2(\delta, \hat{\delta}) \propto e^{h(\hat{\delta} - \delta)} - h(\hat{\delta} - \delta) - 1, \quad h \neq 0.$$

We assume that the prior distributions for β and λ follow independent gamma distributions given by

$$\pi_1(\beta|a, b) \propto \beta^{a-1} e^{-b\beta}, \quad \beta > 0, a > 0, b > 0,$$

and

$$\pi_2(\lambda|c, d) \propto \lambda^{c-1} e^{-d\lambda}, \quad \lambda > 0, c > 0, d > 0.$$

In addition, the hyper-parameters a, b, c , and d represent the prior knowledge of the unknown parameters.

The joint prior distribution of β and λ is of the form

$$\pi(\beta, \lambda) \propto \beta^{a-1} e^{-b\beta} \lambda^{c-1} e^{-d\lambda}, \quad \beta > 0, \lambda > 0. \quad (13)$$

Then, the posterior density of (β, λ) is given by

$$\pi^*(\beta, \lambda | \underline{x}) = \frac{L(\beta, \lambda | \underline{x}) \pi(\beta, \lambda)}{\int_0^\infty \int_0^\infty L(\beta, \lambda) \pi(\beta, \lambda | \underline{x}) d\lambda d\beta}. \quad (14)$$

The Bayes estimates of β and λ against the loss function L_1 are respectively obtained as

$$\hat{\beta}_{SB} = E(\beta | \underline{x}) = \frac{\int_\beta \int_\lambda \beta l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}{\int_\beta \int_\lambda l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta} \quad (15)$$

and

$$\hat{\lambda}_{SB} = E(\lambda | \underline{x}) = \frac{\int_\beta \int_\lambda \lambda l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}{\int_\beta \int_\lambda l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}. \quad (16)$$

The Bayes estimates of β under the loss function L_2 is obtained as

$$\hat{\beta}_{LB} = -\frac{1}{h} \log E(e^{-h\beta} | \underline{x}), \quad h \neq 0,$$

where

$$E(e^{-h\beta} | \underline{x}) = \frac{\int_\beta \int_\lambda e^{-h\beta} l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}{\int_\beta \int_\lambda l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}. \quad (17)$$

The Bayes estimates of λ under the loss function L_2 is obtained as

$$\hat{\lambda}_{LB} = -\frac{1}{h} \log E(e^{-h\lambda} | \underline{x}), \quad h \neq 0,$$

where

$$E(e^{-h\lambda} | \underline{x}) = \frac{\int_\beta \int_\lambda e^{-h\lambda} l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}{\int_\beta \int_\lambda l(\beta, \lambda) \pi(\beta, \lambda) d\lambda d\beta}. \quad (18)$$

The ratios of integrals given in equations (15), (16), (17) and (18) cannot be obtained in a closed form. Thus, [13] approximation method for evaluating the ratio of two integrals have been used. This has been adopted by several researchers, such as [10], [5], to obtain the approximate Bayes estimates.

3.1. Lindley approximation method

Since all estimates have the forms of ratios of two integrals, to obtain these estimates numerically, we use the Lindley's approximation method. Since Bayes estimates of β and λ depend on the ratio of two integrals, we define,

$$I(\underline{x}) = \frac{\int_0^\infty \int_0^\infty u(\beta, \lambda) e^{l(\beta, \lambda | \underline{x}) + \rho(\beta, \lambda)} d\beta d\lambda}{\int_0^\infty \int_0^\infty e^{l(\beta, \lambda | \underline{x}) + \rho(\beta, \lambda)} d\beta d\lambda}, \quad (19)$$

where $u(\beta, \lambda)$ is function of β and λ only and $l(\beta, \lambda | \underline{x})$ is the same as $\log L(\beta, \lambda | \underline{x})$ and $\rho(\beta, \lambda) = \log \pi(\beta, \lambda)$. Then by Lindley's method, $I(\underline{x})$ can be approximated as

$$\begin{aligned} \hat{I}(\underline{x}) = & u(\hat{\beta}, \hat{\lambda}) + \frac{1}{2} [(u_{\hat{\beta}\beta} + 2u_{\hat{\beta}\hat{\rho}\beta}) \sigma_{\hat{\beta}\beta} + (u_{\hat{\lambda}\beta} + 2u_{\hat{\lambda}\hat{\rho}\beta}) \sigma_{\hat{\lambda}\beta} + \\ & (u_{\hat{\beta}\lambda} + 2u_{\hat{\beta}\hat{\rho}\lambda}) \sigma_{\hat{\beta}\lambda} + (u_{\hat{\lambda}\lambda} + 2u_{\hat{\lambda}\hat{\rho}\lambda}) \sigma_{\hat{\lambda}\lambda}] \\ & + \frac{1}{2} [(u_{\hat{\beta}\beta}\sigma_{\hat{\beta}\beta} + u_{\hat{\lambda}\beta}\sigma_{\hat{\beta}\lambda}) (l_{\hat{\beta}\beta\beta}\sigma_{\hat{\beta}\beta} + l_{\hat{\beta}\lambda\beta}\sigma_{\hat{\beta}\lambda} + l_{\hat{\lambda}\beta\beta}\sigma_{\hat{\lambda}\beta} + l_{\hat{\lambda}\lambda\beta}\sigma_{\hat{\lambda}\lambda}) \\ & + (u_{\hat{\beta}\lambda}\sigma_{\hat{\beta}\lambda} + u_{\hat{\lambda}\lambda}\sigma_{\hat{\lambda}\lambda}) (l_{\hat{\lambda}\beta\beta}\sigma_{\hat{\beta}\beta} + l_{\hat{\beta}\lambda\lambda}\sigma_{\hat{\beta}\lambda} + l_{\hat{\lambda}\beta\lambda}\sigma_{\hat{\lambda}\beta} + l_{\hat{\lambda}\lambda\lambda}\sigma_{\hat{\lambda}\lambda})] , \end{aligned} \quad (20)$$

where $\hat{\beta}$ and $\hat{\lambda}$ are the ML estimators of β and λ , respectively. Also, $u_{\beta\beta}$ is the second derivative of the function $u(\beta, \lambda)$ with respect to β and $u_{\hat{\beta}\beta}$ is the same expression evaluated at $(\hat{\beta}, \hat{\lambda})$. Other expressions are given by

$$\begin{aligned} l_{\beta\beta} &= \frac{\partial^2 l(\beta, \lambda)}{\partial \beta^2} \\ &= \sum_{i=1}^m \left[Xi \left(\frac{\left(\frac{\partial^2 F_i}{\partial \beta^2} - \frac{\partial^2 F_{i-1}}{\partial \beta^2} \right)}{(F_i - F_{i-1})} - \frac{\left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)^2}{(F_i - F_{i-1})^2} \right) + Ri \left(\frac{-\frac{\partial^3 F_i}{\partial \beta^3}}{(1 - F_i)} - \frac{\left(\frac{\partial F_i}{\partial \beta} \right)^2}{(1 - F_i)^2} \right) \right] \\ l_{\lambda\lambda} &= \frac{\partial^2 l(\beta, \lambda)}{\partial \lambda^2} \\ &= \sum_{i=1}^m \left[Xi \left(\frac{\left(\frac{\partial^2 F_i}{\partial \lambda^2} - \frac{\partial^2 F_{i-1}}{\partial \lambda^2} \right)}{(F_i - F_{i-1})} - \frac{\left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)^2}{(F_i - F_{i-1})^2} \right) + Ri \left(\frac{-\frac{\partial^3 F_i}{\partial \lambda^3}}{(1 - F_i)} - \frac{\left(\frac{\partial F_i}{\partial \lambda} \right)^2}{(1 - F_i)^2} \right) \right] \\ l_{\beta\lambda} &= \frac{\partial^2 l(\beta, \lambda)}{\partial \lambda \partial \beta} \\ &= \sum_{i=1}^m \left[Xi \left(\frac{\left(\frac{\partial^2 F_i}{\partial \beta \partial \lambda} - \frac{\partial^2 F_{i-1}}{\partial \beta \partial \lambda} \right)}{(F_i - F_{i-1})} - \frac{\left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right) \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} \right) + \right. \\ & \left. Ri \left(\frac{-\frac{\partial^2 F_i}{\partial \beta \partial \lambda}}{(1 - F_i)} - \frac{\left(\frac{\partial F_i}{\partial \beta} \right) \left(\frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^2} \right) \right] \end{aligned} \quad (21)$$

$$\begin{aligned}
 l_{\lambda\beta} &= \frac{\partial^2 l(\beta, \lambda)}{\partial \lambda^2} \\
 &= \sum_{i=1}^m \left[Xi \left(\frac{\left(\frac{\partial^2 F_i}{\partial \beta \partial \lambda} - \frac{\partial^2 F_{i-1}}{\partial \beta \partial \lambda} \right)}{(F_i - F_{i-1})} - \frac{\left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right) \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} \right) + \right. \\
 &\quad \left. R_i \left(\frac{-\frac{\partial^2 F_i}{\partial \beta \partial \lambda}}{(1 - F_i)} - \frac{\left(\frac{\partial F_i}{\partial \beta} \right) \left(\frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^2} \right) \right]. \tag{22}
 \end{aligned}$$

From equations (21) and (22), we have

$$\begin{aligned}
 l_{\beta\beta\beta} &= \frac{\partial^3 l(\beta, \lambda)}{\partial \beta^3} \\
 &= \sum_{i=1}^m \left[Xi \left(\frac{\frac{\partial^3 F_i}{\partial \beta^3} - \frac{\partial^3 F_{i-1}}{\partial \beta^3}}{F_i - F_{i-1}} - \frac{3 \left(\frac{\partial^2 F_i}{\partial \beta^2} - \frac{\partial^2 F_{i-1}}{\partial \beta^2} \right) \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} + \frac{2 \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)^3}{(F_i - F_{i-1})^3} \right) \right. \\
 &\quad \left. - R_i \left(\frac{\frac{\partial^3 F_i}{\partial \beta^3}}{1 - F_i} + \frac{3 \left(\frac{\partial^2 F_i}{\partial \beta^2} \right) \left(\frac{\partial F_i}{\partial \beta} \right)}{(1 - F_i)^2} + \frac{2 \left(\frac{\partial F_i}{\partial \beta} \right)^3}{(1 - F_i)^3} \right) \right],
 \end{aligned}$$

$$\begin{aligned}
 l_{\lambda\lambda\lambda} &= \frac{\partial^3 l(\beta, \lambda)}{\partial \lambda^3} \\
 &= \sum_{i=1}^m \left[Xi \left(\frac{\frac{\partial^3 F_i}{\partial \lambda^3} - \frac{\partial^3 F_{i-1}}{\partial \lambda^3}}{F_i - F_{i-1}} - \frac{3 \left(\frac{\partial^2 F_i}{\partial \lambda^2} - \frac{\partial^2 F_{i-1}}{\partial \lambda^2} \right) \left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)}{(F_i - F_{i-1})^2} + \frac{2 \left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)^3}{(F_i - F_{i-1})^3} \right) \right. \\
 &\quad \left. - R_i \left(\frac{\frac{\partial^3 F_i}{\partial \lambda^3}}{1 - F_i} + \frac{3 \left(\frac{\partial^2 F_i}{\partial \lambda^2} \right) \left(\frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^2} + \frac{2 \left(\frac{\partial F_i}{\partial \lambda} \right)^3}{(1 - F_i)^3} \right) \right],
 \end{aligned}$$

$$\begin{aligned}
 l_{\lambda\beta\beta} &= \frac{\partial^3 l(\beta, \lambda)}{\partial \lambda \partial \beta^2} \\
 &= \sum_{i=1}^m \left[Xi \left(\frac{\frac{\partial^3 F_i}{\partial \lambda \partial \beta^2} - \frac{\partial^3 F_{i-1}}{\partial \lambda \partial \beta^2}}{F_i - F_{i-1}} - \frac{\left(\frac{\partial^2 F_i}{\partial \beta^2} - \frac{\partial^2 F_{i-1}}{\partial \beta^2} \right) \left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)}{(F_i - F_{i-1})^2} \right. \right. \\
 &\quad \left. \left. - \frac{2 \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right) \left(\frac{\partial^2 F_i}{\partial \lambda \partial \beta} - \frac{\partial^2 F_{i-1}}{\partial \lambda \partial \beta} \right)}{(F_i - F_{i-1})^2} + \frac{2 \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)^2 \left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right)}{(F_i - F_{i-1})^3} \right) \right. \\
 &\quad \left. - R_i \left(\frac{\frac{\partial^3 F_i}{\partial \lambda \partial \beta^2}}{1 - F_i} + \frac{\left(\frac{\partial^2 F_i}{\partial \beta^2} \right) \left(\frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^2} + 2 \frac{\left(\frac{\partial F_i}{\partial \beta} \right) \left(\frac{\partial^2 F_i}{\partial \lambda \partial \beta} \right)}{(1 - F_i)^3} + \frac{2 \left(\frac{\partial F_i}{\partial \beta} \right)^2 \left(\frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^3} \right) \right],
 \end{aligned}$$

$$\begin{aligned}
 l_{\beta\lambda\beta} &= \frac{\partial^3 l(\beta, \lambda)}{\partial \lambda \partial \beta^2} \\
 &= \sum_{i=1}^m \left[X_i \left(\frac{\frac{\partial^3 F_i}{\partial \beta \partial \lambda \partial \beta} - \frac{\partial^3 F_{i-1}}{\partial \beta \partial \lambda \partial \beta}}{F_i - F_{i-1}} - \frac{\left(\frac{\partial^2 F_i}{\partial \beta \partial \lambda} - \frac{\partial^2 F_{i-1}}{\partial \beta \partial \lambda} \right) \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} \right. \right. \\
 &\quad \left. \left. - \left(\frac{\left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right) \left(\frac{\partial^2 F_i}{\partial \beta^2} - \frac{\partial^2 F_{i-1}}{\partial \beta^2} \right) + \left(\frac{\partial^2 F_i}{\partial \lambda \partial \beta} - \frac{\partial^2 F_{i-1}}{\partial \lambda \partial \beta} \right) \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)}{(F_i - F_{i-1})^2} \right) \right. \right. \\
 &\quad \left. \left. - \frac{2 \left(\frac{\partial F_i}{\partial \lambda} - \frac{\partial F_{i-1}}{\partial \lambda} \right) \left(\frac{\partial F_i}{\partial \beta} - \frac{\partial F_{i-1}}{\partial \beta} \right)^2}{(F_i - F_{i-1})^2} \right) \right. \\
 &\quad \left. + R_i \left(\left(-\frac{\frac{\partial^3 F_i}{\partial \beta \partial \lambda \partial \beta^2}}{1 - F_i} + \frac{\left(\frac{\partial^2 F_i}{\partial \beta \partial \lambda} \right) \left(\frac{\partial F_i}{\partial \beta} \right)}{(1 - F_i)^2} \right) - \left(\frac{\left(\frac{\partial F_i}{\partial \beta} \right) \left(\frac{\partial^2 F_i}{\partial \lambda \partial \beta} \right)}{(1 - F_i)^3} + \frac{\left(\frac{\partial F_i}{\partial \beta} \right) \left(\frac{\partial F_i}{\partial \lambda} \right)}{(1 - F_i)^2} \right) \right) \right. \\
 &\quad \left. \left. + \frac{2 \left(\frac{\partial F_i}{\partial \lambda} \right) \left(\frac{\partial F_i}{\partial \lambda} \right)^2}{(1 - F_i^3)} \right) \right].
 \end{aligned}$$

Let $u_i = 1 - e^{-t_i}$. Then

$$\begin{aligned}
 \frac{\partial F_i}{\partial \beta} &= \lambda [1 - u_i^\beta]^{\lambda-1} u_i^\beta \ln u_i \\
 \frac{\partial^2 F_i}{\partial \beta^2} &= \lambda [1 - u_i^\beta]^{\lambda-1} u_i^\beta (\ln u_i)^2 - \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} (u_i^\beta \ln u_i)^2 \\
 \frac{\partial^3 F_i}{\partial \beta^3} &= \lambda [1 - u_i^\beta]^{\lambda-1} u_i^\beta (\ln u_i)^3 + u_i^\beta (\ln u_i)^2 \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} \\
 \frac{\partial F_i}{\partial \lambda} &= -\lambda [1 - u_i^\beta]^{\lambda-1} \\
 \frac{\partial^2 F_i}{\partial \lambda^2} &= -\lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} \\
 \frac{\partial^3 F_i}{\partial \lambda^3} &= -\lambda(\lambda - 1)(\lambda - 2) [1 - u_i^\beta]^{\lambda-3} \\
 &\quad - 2\lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} u_i^\beta (\ln u_i)^3 + \lambda(\lambda - 1)(\lambda - 2) [1 - u_i^\beta]^{\lambda-3} (u_i^\beta \ln u_i)^3 \\
 \frac{\partial^2 F_i}{\partial \beta \partial \lambda} &= \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} u_i^\beta \ln u_i \\
 \frac{\partial^3 F_i}{\partial \beta^2 \partial \lambda} &= \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} u_i^\beta (\ln u_i)^2 - \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-3} (u_i^\beta \ln u_i)^2 \\
 \frac{\partial^3 F_i}{\partial \beta \partial \lambda^2} &= \lambda(\lambda - 1)(\lambda - 2) [1 - u_i^\beta]^{\lambda-3} u_i^\beta \ln u_i \\
 \frac{\partial^2 F_i}{\partial \lambda \partial \beta} &= \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} u_i^\beta \ln u_i \\
 \frac{\partial^3 F_i}{\partial \beta \partial \lambda \partial \beta} &= \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} u_i^\beta (\ln u_i)^2 - \lambda(\lambda - 1)(\lambda - 2) [1 - u_i^\beta]^{\lambda-3} u_i^\beta \ln u_i \\
 \frac{\partial^3 F_i}{\partial \lambda \partial \beta^2} &= \lambda(\lambda - 1) [1 - u_i^\beta]^{\lambda-2} u_i^\beta (\ln u_i)^2 - \lambda(\lambda - 1)(\lambda - 2) [1 - u_i^\beta]^{\lambda-3} (u_i^\beta \ln u_i)^2
 \end{aligned}$$

Also,

$$\rho(\beta, \lambda) \propto (c - 1)\log\beta - d\beta + (b - 1)\log\lambda - a\lambda. \quad (23)$$

Thus,

$$\hat{\rho}_\beta = \frac{c - 1}{\hat{\beta}} - d$$

and

$$\hat{\rho}_\lambda = \frac{b - 1}{\hat{\lambda}} - a.$$

Here

$$\begin{pmatrix} \hat{\sigma}_{\beta\beta} & \hat{\sigma}_{\beta\lambda} \\ \hat{\sigma}_{\lambda\beta} & \hat{\sigma}_{\lambda\lambda} \end{pmatrix} = - \begin{pmatrix} \hat{I}_{\beta\beta} & \hat{I}_{\beta\lambda} \\ \hat{I}_{\lambda\beta} & \hat{I}_{\lambda\lambda} \end{pmatrix}^{-1}$$

We now determine the approximate Bayes estimates of β and λ under various loss functions using the above-mentioned equations. First, we derive the Bayes estimates for β and λ under the squared error loss function L_1 . For estimating β , we take $u(\beta, \lambda) = \beta$. Therefore $u_\beta = 1$ and $u_{\beta\beta} = u_\lambda = u_{\lambda\lambda} = u_{\beta\lambda} = u_{\lambda\beta} = 0$. Then the Bayes estimate of β under the loss function L_1 is obtained as

$$\begin{aligned} \hat{\beta}_{SB} = & \hat{\beta} + 0.5[2\hat{\beta}_\beta\hat{\sigma}_{\beta\beta} + 2\hat{\beta}_\lambda\hat{\sigma}_{\beta\lambda} + \hat{\sigma}_{\beta\beta}\hat{I}_{\beta\beta\beta} + \\ & \hat{\sigma}_{\beta\beta}\hat{\sigma}_{\beta\lambda}\hat{I}_{\beta\lambda\beta} + 2\hat{\sigma}_{\beta\beta}\hat{\sigma}_{\lambda\beta}\hat{I}_{\lambda\beta\beta} + \hat{\sigma}_{\lambda\beta}\hat{\sigma}_{\lambda\lambda}\hat{I}_{\lambda\lambda\lambda}]. \end{aligned}$$

To estimate λ , we take $u(\beta, \lambda) = \lambda$. Thus $u_\lambda = 1$ and $u_\beta = u_{\beta\beta} = u_{\lambda\lambda} = u_{\lambda\beta} = u_{\beta\lambda} = 0$. Then the Bayes estimate of λ under the loss function L_1 can be determined as

$$\begin{aligned} \hat{\lambda}_{SB} = & \hat{\lambda} + 0.5[2\hat{\beta}_\beta\hat{\sigma}_{\lambda\beta} + 2\hat{\beta}_\lambda\hat{\sigma}_{\lambda\lambda} + \hat{\sigma}_{\beta\beta}\hat{\sigma}_{\beta\lambda}\hat{I}_{\beta\beta\beta} + \\ & \hat{\sigma}_{\beta\lambda}^2\hat{I}_{\beta\lambda\beta} + \hat{\sigma}_{\beta\lambda}\hat{\sigma}_{\lambda\beta}\hat{I}_{\lambda\beta\beta} + \hat{\sigma}_{\beta\beta}\hat{\sigma}_{\lambda\lambda}\hat{I}_{\lambda\beta\beta} + \hat{\sigma}_{\lambda\lambda}^2\hat{I}_{\lambda\lambda\lambda}]. \end{aligned}$$

Now, we obtain the Bayes estimates of β and λ under LINEX loss function L_2 . For estimating β , we take $u(\beta, \lambda) = e^{-h\beta}$. Thus $u_\beta = -he^{-h\beta}$, $u_{\beta\beta} = h^2e^{-h\beta}$ and $u_\lambda = u_{\lambda\lambda} = u_{\lambda\beta} = u_{\beta\lambda} = 0$. Therefore, Bayes estimate of β under the loss function L_2 is obtained as

$$\hat{\beta}_{LB} = -\frac{1}{h}\log[E(e^{-h\beta}|\underline{x})], \quad (24)$$

where

$$\begin{aligned} E(e^{-h\beta}|\underline{x}) = & e^{-h\hat{\beta}} + 0.5[\hat{u}_{\beta\beta}\hat{\sigma}_{\beta\beta} + \hat{u}_\lambda(2\hat{\beta}_\lambda\hat{\sigma}_{\beta\beta} + 2\hat{\beta}_\lambda\hat{\sigma}_\lambda + \hat{\sigma}_{\beta\beta}^2\hat{I}_{\beta\beta\beta} + \\ & \hat{\sigma}_{\beta\beta}\hat{\sigma}_{\beta\lambda}\hat{I}_{\beta\lambda\beta} + 2\hat{\sigma}_{\beta\beta}\hat{\sigma}_{\lambda\beta}\hat{I}_{\lambda\beta\beta})], \end{aligned} \quad (25)$$

To estimate λ , we take $u(\beta, \lambda) = e^{-h\lambda}$. Thus, $u_\lambda = -he^{-h\lambda}$, $u_{\lambda\lambda} = h^2e^{-h\lambda}$ and $u_\beta = u_{\beta\beta} = u_{\lambda\beta} = u_{\beta\lambda} = 0$. Therefore, the Bayes estimate of λ under the loss function L_2 is obtained as

$$\hat{\lambda}_{LB} = -\frac{1}{h}\log[E(e^{-h\lambda}|\underline{x})], \quad (26)$$

where

$$\begin{aligned} E(e^{-h\lambda}|\underline{x}) = & e^{-h\hat{\lambda}} + 0.5 \left[\hat{u}_{\lambda\lambda}\hat{\sigma}_{\lambda\lambda} + \hat{u}_\lambda \left(2\hat{\beta}_\beta\hat{\sigma}_{\lambda\beta} + 2\hat{\beta}_\lambda\hat{\sigma}_{\lambda\lambda} + \hat{\sigma}_{\beta\beta}\hat{\sigma}_{\beta\lambda}\hat{I}_{\beta\beta\beta} + \right. \right. \\ & \left. \left. \hat{\sigma}_{\beta\lambda}^2\hat{I}_{\beta\lambda\beta} + \hat{\sigma}_{\beta\lambda}\hat{\sigma}_{\lambda\beta}\hat{I}_{\lambda\beta\beta} + \hat{\sigma}_{\lambda\lambda}\hat{\sigma}_{\beta\beta}\hat{I}_{\lambda\beta\beta} + \hat{\sigma}_{\lambda\lambda}^2\hat{I}_{\lambda\lambda\lambda} \right) \right]. \end{aligned} \quad (27)$$

We use $h = 1$ to evaluate Bayes estimators under the LINEX loss function L_2 . In each case, we have assessed bias and MSE based on 500 iterations. We repeat the simulation study for various values of β and λ also. The bias and MSE for the estimate of β for both informative and non-informative priors are given in Table 2. Table 3 provides the bias and MSE for the estimate of λ for both informative and non-informative priors.

The tabulated values shows that all of the estimates do improve with a higher value of n . From Table 1, regarding MSE and Bias, we found that the estimates based on the censoring schemes 2 and 5 give the better estimates of β and λ , followed by the scheme 4. In case of maximum likelihood estimation given in table 1, as n increases the MSE of estimates decrease as expected. Also, we can see that bias and MSEs of the estimates of β and λ via EM algorithm are smaller than bias and MSEs of the corresponding MLEs. Also the Bayes estimators based on informative prior perform much better than the MLEs in terms of biases and MSEs. From the tables 1, 2 and 3, it is clear that the bias and MSE of Bayes estimators under informative prior are smaller than those of MLE's.

As expected, the Bayes estimators based on informative prior perform much better than the Bayes estimators based on non-informative prior in terms of biases and MSEs. From Tables 2 and 3, one can see that for β and λ , estimators based on informative priors perform better to those of non-informative priors in terms of bias and MSE. Also, among the Bayes estimators of β , the estimator under the LINEX loss function performs better. Again, when compared to squared error loss functions, estimators of λ under the LINEX loss function have the least bias and MSE.

Table 1: Bias and MSE of parameters under different censoring schemes for different values of β and λ

				$\hat{\beta}$				$\hat{\lambda}$			
(β, λ)	n	m	c.s	MLE		EM		MLE		EM	
				Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
(1.25,1.5)	75	10	1	-0.3509	0.1807	-0.3055	0.0956	-0.4086	0.3751	-0.3131	0.3318
		12	2	-0.2609	0.1073	-0.1063	0.0268	-0.2298	0.1260	-0.1192	0.1057
		15	3	-0.3948	0.2090	-0.3129	0.0985	-0.6539	0.7903	-0.2564	0.7338
		20	4	-0.2811	0.1276	-0.2259	0.0631	-0.2734	0.2140	-0.1871	0.2175
		25	5	-0.2725	0.1174	-0.2195	0.0495	-0.2567	0.1782	-0.1691	0.1555
(1.75,2)	100	10	1	-0.3487	0.1360	-0.3160	0.0772	-0.1784	0.2405	-0.1686	0.1903
		12	2	-0.0542	0.0394	-0.0528	0.0901	-0.1275	0.1113	-0.1136	0.0511
		15	3	-0.3526	0.1757	-0.3016	0.0912	-0.3358	0.3049	-0.2373	0.2762
		20	4	-0.1386	0.0538	-0.1029	0.0558	-0.1628	0.1496	-0.1421	0.1040
		25	5	-0.0982	0.0501	-0.0898	0.0242	-0.1422	0.1386	-0.1315	0.0757

Table 2: Bias and MSE of the Bayes estimates of β under squared error loss function, $\hat{\beta}_{SB}$, and LINEX loss function, $\hat{\beta}_{LB}$, for different values of β and λ

β	λ	n	m	c.s	Informative Prior 1						Informative Prior 2						Non-Informative Prior					
					Bias	MSE	$\hat{\beta}_{SB}$	Bias	MSE	$\hat{\beta}_{LB}$	Bias	MSE	$\hat{\beta}_{SB}$	Bias	MSE	$\hat{\beta}_{LB}$	Bias	MSE	$\hat{\beta}_{SB}$	Bias	MSE	$\hat{\beta}_{LB}$
1.25	1.5	75	10	1	-0.3408	0.1736	-0.3378	0.1717	-0.3324	0.1602	-0.3282	0.1580	-0.5498	0.3766	-0.5468	0.3733						
			12	2	-0.2562	0.1043	-0.2540	0.1033	-0.2159	0.0878	-0.2141	0.0872	-0.2804	0.1246	-0.2776	0.1183						
			15	3	-0.3821	0.1986	-0.3786	0.1961	-0.1147	0.1000	-0.1151	0.0802	-0.4722	0.2432	-0.4592	0.2268						
		100	20	4	-0.2477	0.0860	-0.2468	0.0856	-0.2338	0.0770	-0.2322	0.0764	-0.2708	0.1103	-0.2696	0.1032						
			25	5	-0.2670	0.0943	-0.2664	0.0940	-0.1117	0.0375	-0.1109	0.0373	-0.2694	0.0980	-0.2682	0.0975						
	2	75	10	1	-0.2720	0.1067	-0.2711	0.1063	-0.2338	0.0991	-0.2318	0.9014	-0.3475	0.1689	-0.3473	0.1688						
			12	2	-0.0533	0.0368	-0.0530	0.0361	-0.1740	0.1046	-0.1733	0.1043	-0.0656	0.0306	-0.0654	0.0304						
			15	3	-0.3469	0.1683	-0.3459	0.1677	-0.1537	0.0748	-0.1542	0.0747	-0.3824	0.1997	-0.3790	0.1974						
		100	20	4	-0.0995	0.0537	-0.0986	0.0536	-0.1106	0.0389	-0.1163	0.0389	-0.2372	0.0930	-0.2363	0.0927						
			25	5	-0.0929	0.0499	-0.0921	0.0498	-0.0894	0.0541	-0.0886	0.0540	-0.1559	0.0635	-0.1552	0.0634						

Table 3: Bias and MSE of the Bayes estimates of λ under squared error loss function, $\hat{\lambda}_{SB}$, and LINEX loss function, $\hat{\lambda}_{LB}$, for different values of β and λ .

β	λ	n	m	c.s	Informative Prior 1						Informative Prior 2						Non-Informative Prior					
					$\hat{\lambda}_{SB}$		$\hat{\lambda}_{LB}$		$\hat{\lambda}_{SB}$		$\hat{\lambda}_{LB}$		$\hat{\lambda}_{SB}$		$\hat{\lambda}_{LB}$		$\hat{\lambda}_{SB}$		$\hat{\lambda}_{LB}$			
					Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE		
1.25	1.5	75	10	1	-0.4077	0.3745	-0.4076	0.3744	-0.4548	0.4265	-0.4546	0.4264	-0.6067	0.6583	-0.6066	0.6581	-0.6066	0.6581				
					-0.2519	0.1573	-0.2518	0.1573	-0.1998	0.1438	-0.1998	0.1438	-0.3475	0.3341	-0.3473	0.3340	-0.3473	0.3340				
					-0.6532	0.7865	-0.6529	0.7862	-0.2819	0.3809	-0.2818	0.3809	-1.4426	2.6977	-1.4425	2.6973	-1.4425	2.6973				
					-0.2291	0.1237	-0.2291	0.1237	-0.2376	0.1202	-0.2376	0.1202	-0.2761	0.1893	-0.2760	0.1892	-0.2760	0.1892				
					-0.2283	0.1258	-0.2283	0.1258	-0.1312	0.0791	-0.1312	0.0790	-0.2436	0.1420	-0.2435	0.1419	-0.2435	0.1419				
	1.75	2	75	10	1	-0.1742	0.1449	-0.1742	0.1449	-0.1293	0.2154	-0.1293	0.2154	-0.3583	0.2185	-0.3583	0.2185	-0.3583	0.2185			
						-0.1348	0.1299	-0.1346	0.1297	-0.1774	0.2478	-0.1771	0.2477	-0.2544	0.1782	-0.2544	0.1782	-0.2544	0.1782			
						-0.3182	0.2985	-0.3182	0.2985	0.3172	0.1586	0.3172	0.1586	-0.4004	0.3473	-0.4004	0.3473	-0.4004	0.3473			
						-0.1550	0.1466	-0.1550	0.1466	-0.1297	0.1313	-0.1297	0.1313	-0.2658	0.1912	-0.2657	0.1912	-0.2657	0.1912			
						-0.1440	0.1479	-0.1440	0.1479	-0.1242	0.1357	-0.1242	0.1357	-0.2098	0.1581	-0.2098	0.1581	-0.2098	0.1581			

5. ILLUSTRATIONS USING REAL DATA

In this section, a real-life data is utilised to demonstrate the inference methods proposed in this paper. The data was previously studied by [14] and [15].

The data shows the running and failure times for a sample of devices from the larger system's eld-tracking research. The failure times are:

2.75, 0.13, 1.47, 0.23, 1.81, 0.30, 0.65, 0.10, 3.00, 1.73, 1.06, 3.00, 3.00, 2.12, 3.00, 3.00, 3.00, 0.02, 2.61, 2.93, 0.88, 2.47, 0.28, 1.43, 3.00, 0.23, 3.00, 0.80, 2.45, 2.66.

The data was also previously considered by [16] and fitted for $KE(\beta, \lambda)$ distribution. For evaluating the goodness of fit, they used the Anderson-Darling test. The Anderson-Darling test statistic has a value of 2.00757 and the related P-value is 0.0913729. Based on the aforementioned estimation procedures, we have obtained the estimates of β and λ , which is included in Table 4.

Table 4: Estimates of β and λ for the real data

<i>n</i>	<i>m</i>	<i>Censoring scheme</i>		<i>Bayes</i>			
				<i>MLE</i>	<i>EM</i>	<i>SE</i>	<i>LINEX</i>
30	5	$q = (0.25, 0.25, 0.5, 0.5, 1)$ $X_i = (8, 3, 4, 1, 4)$ $R_i = (2, 1, 3, 4, 0)$	β	1.2857	1.5756	1.5875	1.5173
			λ	0.5192	0.5739	0.5324	0.5338
30	7	$q = (0.5, 0, 0, 0, 0, 0, 1)$ $X_i = (7, 3, 3, 2, 3, 5, 0)$ $R_i = (3, 0, 0, 0, 0, 4, 0)$	β	1.1000	1.3050	0.9374	0.9214
			λ	0.5355	0.5784	0.5875	0.5866
30	12	$q = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1)$ $X_i = (5, 2, 1, 2, 1, 2, 1, 1, 1, 2, 3, 4)$ $R_i = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 5)$	β	0.8453	1.1082	0.7944	0.7866
			λ	0.4445	0.4961	0.4869	0.4346

6. CONCLUSION

In this paper, we considered the problem of estimation of parameters of Kumaraswamy-exponential distribution based on progressive type-I interval censored sample. The maximum likelihood estimators of the parameters β and λ were obtained. Since the MLEs of the unknown parameters of the distribution does not admit closed form, we employed the EM algorithm approach. The Bayes estimators were also obtained using different loss functions such as squared error loss function and LINEX loss function. To evaluate the Bayes estimators, Lindley's approximation method was applied. Based on simulation study, we have the following conclusions. We observed that the performance of EM algorithm was quite satisfactory. In addition, it was found that for both β and λ , the bias and MSE of the Bayes estimators under an informative prior are smaller than those of MLEs. The performance of Informative prior was better than the Non-informative prior both β and λ in terms of bias and MSE values. For both β and λ , Bayes estimators under LINEX loss function perform better with regard to bias and MSE. The estimation methods employed in this paper were also illustrated using real data sets.

REFERENCES

- [1] Adepoju, K. and Chukwu, O. (2015). Maximum likelihood estimation of the Kumaraswamy exponential distribution with applications. *Journal of Modern Applied Statistical Methods*, 14(1):18.
- [2] Aggarwala, R. (2001). Progressive interval censoring: some mathematical results with applications to inference. *Communications in Statistics-Theory and Methods*, 30(8- 9):1921–1935.
- [3] Alotaibi, R., Rezk, H., Dey, S., and Okasha, H. (2021). Bayesian estimation for dagum distribution based on progressive type-I interval censoring. *Plos One*, 16(6):e0252556
- [4] Ashour, S. and Afify, W. (2007). Statistical analysis of exponentiated Weibull family under type-I progressive interval censoring with random removals. *Journal of Applied Sciences Research*, 3(12):1851–1863.
- [5] Chacko, M. and Mohan, R. (2017). Estimation of parameters of Kumaraswamy-exponential distribution under progressive type-II censoring. *Journal of Statistical Computation and Simulation*, 87(10):1951–1963.
- [6] Chen, D. and Lio, Y. (2010). Parameter estimations for generalized exponential distribution under progressive type-I interval censoring. *Computational Statistics & Data Analysis*, 54(6):1581–1591.
- [7] Cohen, A. C. (1963). Progressively censored samples in life testing. *Technometrics*, 5(3):327–339.
- [8] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22.
- [9] Elbatal, I. (2013). Kumaraswamy linear exponential distribution. *Pioneer J Theor Appl Statist*, 5:59–73.
- [10] Kundu, D. and Gupta, R. D. (2008). Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis*, 52(4):1873–1883.
- [11] Lemonte, A. J., Barreto-Souza, W., and Cordeiro, G. M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27(1), 31-53.
- [12] Lin, C.-T., Wu, S. J., and Balakrishnan, N. (2009). Planning life tests with progressively type-I interval censored data from the lognormal distribution. *Journal of Statistical Planning and Inference*, 139(1):54–61.
- [13] Lindley, D. V. (1980). Approximate bayesian methods. *Trabajos de estadística y de investigación operativa*, 31:223–245.
- [14] Meeker, W. and Escobar, L. (1998). Statistical methods for reliability data john wiley & sons new york. *New York*.
- [15] Merovci, F. and Elbatal, I. (2015). Weibull Rayleigh distribution: Theory and applications. *Appl. Math. Inf. Sci*, 9(5):1–11.
- [16] Mohan, R. and Chacko, M. (2021). Estimation of parameters of Kumaraswamy-exponential distribution based on adaptive type-II progressive censored schemes. *Journal of Statistical Computation and Simulation*, 91(1):81–107.
- [17] Ng, H. K. T. and Wang, Z. (2009). Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample. *Journal of Statistical Computation and Simulation*, 79(2):145–159.
- [18] Rodrigues, J. and Silva, A. (2015). The exponentiated Kumaraswamy-exponential distribution. *British Journal of Applied Science & Technology*, 10(1):12.
- [19] Singh, S. and Tripathi, Y. M. (2018). Estimating the parameters of an inverse Weibull distribution under progressive type-I interval censoring. *Statistical Papers*, 59:21–56.

ANALYSIS OF $M^X/G/1$ QUEUE WITH OPTIONAL SECOND SERVICE, FEEDBACK AND BERNOULLI VACATION

¹S. KARPAGAM, ²B. SOMASUNDARAM, ³A. KAVIN SAGANA MARY ⁴R. LOKESH,

^{1,2,3,4}Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,
Tamil Nadu, India.

¹karpagam19sks@gmail.com, ²somu.b92@gmail.com,

³kavinmary28@gmail.com, ⁴rlokesh1020@gmail.com

Abstract

In this article the single-server queue situation described with batch arrivals, a mandatory first service and a choice of second service are provided to the customers. A general distribution governs the service times, whereas a compound Poisson distribution follows customer arrivals. Although each new customer requests the first mandatory service, only some of them choose the optional second service. Customers who are dissatisfied with mandatory service are more likely to get the required services later on. After every service is finished, the server might choose to go on Bernoulli vacation. Time dependent probability generating functions are constructed in terms of Laplace transforms using the supplementary variable approach, and explicit results are obtained for the steady state. Additionally, mean waiting time and mean queue length expressions are examined. The graphical and numerical representations improve comprehension of the results even further.

Keywords: Optional second service, Feedback, Bernoulli vacation

1. INTRODUCTION

Queueing system is useful in a wide range of scenarios. Wireless networks, supermarkets, restaurants, hospitals, modulation lines, and communication systems all use queueing systems. There will always be lines. People are permitted to wait in line if a service time exceeds the pace of arrivals. Recently, there have been several contributions on taking the $M/G/1$ queue with second optional service into consideration. These kinds of queue situations arise in daily life. An optional second service is part of a single-server queueing system that has been studied by Madan [11]. Madheswari and Suganthi [17] have discussed an $M/G/1$ orbital queue with an optional service and beginning failures. The retrial time distribution is assumed to be exponential. Chowdary and Paul [7] have explained a customers arrival in batches with an optional second service within an N-policy framework. A batch arrival and single server with two service phases (the second of which is optional) and working breakdown has been developed by Somasundaram et al. [22] The system's performance, steady state outcomes, and optimization analysis are examined and also discusses how the concept to be applied to cellular networks and how crucial it is to take server failures into account while providing services.

Queueing systems with feedback is allow customers to return to the same server for re-service with certain probability. Li and Jinting [10] have suggested the implementation of a single server orbital queue featuring numerous optional services and feedback options. Varalakshmi et al. [24]

added to the idea of instant feedback in a unique way. After service completion, if a customer need further service, they will instantly get the another service. A two-phase $M/G/1$ queueing model with instantaneous feedback only for finite number of customers was studied by Kalidass and Kasturi [8]. Moreover, there are several real-world situations where an orbital queue with feedback arises.

Queueing models with vacations have been examined by several researchers. Maragathasundari and Srinivasan [15] investigated the $M/G/1$ queue models with a single vacation in a transient study. Madan [12] examined a single server queue that required mandatory server shutdowns. In the industrial business, Karpagam et al. [9] examined a bulk queueing system with rework that had a single vacation and beginning failure. Thangaraj and Vanitha [23] have studied two stages of service, with single server and single arrival subject to compulsory server vacation and random breakdown.

A batch arrival and single server queueing model with balking and vacation was studied by Charan et al. [21]. Customers may leave the system when server is busy or on vacation. Ayyappan and Deepa [3] discovered batch arrival bulk service queueing system with mandatory and optional repair process. They formulate for the number of performance metrics, including expected queue length, expected waiting time and idle duration. A single server retrial queue with working vacations under multiple vacation policy, vacation interruptions, breakdown and impatient of the customers was examined by Rajadurai et al. [18]. In this study, server breakdown due to the arrival of negative consumers. Shanmugasundaram and Sivaram [20] discussed a single server queue that includes feedback for a client and sever vacation. The steady-state probability and some importance measures were obtained in this investigation. A two-stage batch arrival queue system with reneging during vacation and breakdown times was presented by Baruah et al. [16].

Vacations can be categorized into various types, which are single vacations, multiple vacations, compulsory vacations, modified vacations, J-vacations, Bernoulli vacations, modified Bernoulli vacations, working vacations, multiple working vacations etc., Numerous of studies have used the Bernoulli vacation as a parameter while analyzing queueing models. Arivudainambi and Gowsalya [1] analysed an $M/G/1$ retrial queue with staring failure and vacation scheduled on Bernoulli type. An $M^{[X]}G/1$ retrial queue was studied by Madhu and Kaur [13], who combined Bernoulli feedback, optional service, and the Bernoulli vacation idea for an unreliable server. A retrial queue with batch arrivals was studied by Madhu et al. [14]. Bernoulli vacation was used to provide the server with the opportunity to take a rest during both service phases.

Arivudainambi and Gowsalya [2] developed a Bernoulli vacation schedule and two types of service for a retrial queueing system. The study covered the growing applications in teletraffic theory, client-server communication, etc. A discrete time retrial queue with Bernoulli vacation, preemptive resume priority, general Bernoulli feedback, and retrial periods was analysed by Chen et al. [6]. This study indicates that if the server becomes idle after service, it will either wait for a customer or initiate a single vacation. A repairable queue model with Bernoulli vacation and a two-phase service structure was created by Wang and Li [26]. An $M^{[X]}G/1$ feedback retrial queue with two-phase service, Bernoulli vacation, delayed repair, and orbit search was analyzed by Chandrasekaran et al. [5].

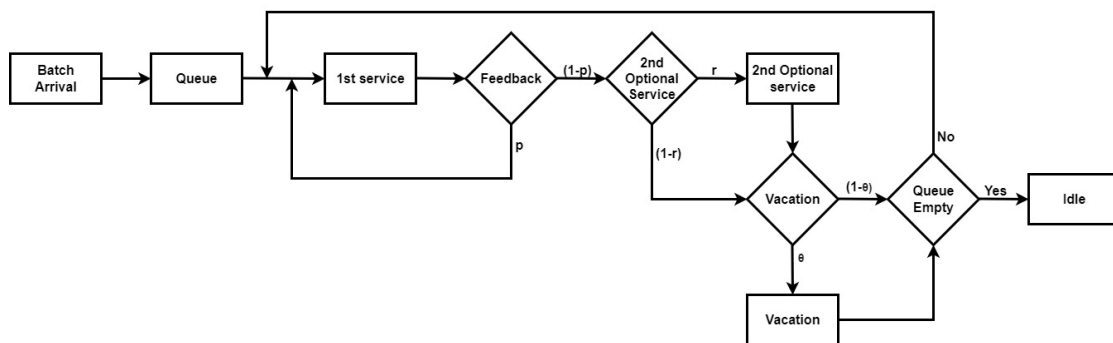
Ayyappan and Somasundaram [4] analyzed a two-stage retrial G-queue with Bernoulli vacation, working breakdown, and discretionary priority services. During the first stage of this investigation, incoming priority units are free to disrupt service; however, during the second stage, they are not allowed to do so. A modified Bernoulli vacation of batch arrival and retrial queue with balking customers due to beginning failure was communicated by Rajam and Uma [19]. There are several uses for this model in the fields of healthcare, mail, manufacturing, production lines, and communication networks. A retrial queue with feedback, working breakdowns, and Bernoulli vacation was examined by Varalakshmi and Rajadurai [25].

In this paper, we study an $M^{[X]}G/1$ queue with feedback, optional second service, and Bernoulli vacations where the second service stage is optional and the first service stage is mandatory. Customers might choose to try again if they are not satisfied with the service they received re-service after the first stage. The server can also take a Bernoulli vacation when each

service is finished.

The primary goal of this inquiry is to ascertain average customer wait times and queue lengths, which are crucial metrics for assessing the system effectiveness. The study comprehensive outline is given in the parts that follow. The mathematical model description, integrating the essential presumptions, is presented in Section 2. Section 3, explores real-world examples and applications. Section 4, presents a series of equations that define the model governing the system. Section 5, deduces time-dependent solutions, while Section 6 concentrates on figuring out steady-state outcomes. Some important performance metrics are calculated in Section 7. Section 8, follows provides graphical representations and numerical results.

2. METHODS



Diagrammatic representation of this model

- Customers use a compound Poisson process to enter the system in batches of different sizes. $\lambda C_k dt$ (where $k = 1, 2, 3, \dots$) represents the likelihood that a batch of consumers of size k would join the system during a short time interval $(t, t + dt)$. The probabilities are $\sum_{k=1}^{\infty} C_k = 1$ and $0 \leq C_k \leq 1$.
- The single server provides each customer with the first mandatory service. Let $r_1(V)$ stand for the first service time density function and $\mathfrak{R}_1(V)$ for the distribution function.
- Following the fulfillment of the mandatory service, the customer dissatisfaction of service, there is a probability 'p' that they will receive their standard service again. Alternatively, the customer may permanently exit the system with a probability 'q=1-p'.
- Customers might choose to proceed with a second optional service after completing their mandatory service. With the probability 'r', if they desire the optional service they will immediately get the service again or else with probability '1-r' the customer may leave the system.
- Consider $\mathfrak{R}_2(V)$ as the distribution function and let $r_2(V)$ represent the density function specifically associated with the second optional service.
- The likelihood that the i^{th} service will be completed within the time range $[\ell, \ell + d\ell]$, provided that ℓ has elapsed, is expressed by the equation $\mu_i(\ell)d\ell$.

$$\mu_i(\ell) = \frac{r_i(\ell)}{1 - \mathfrak{R}_i(\ell)} \quad i = 1, 2, \dots$$

and therefore

$$r_i(v) = \mu_i(v) e^{-\int_0^v \mu_i(\ell) d\ell} \quad i=1, 2, \dots$$

- Following the completion of each service, either the server may go for Bernoulli vacation with probability ' θ ' or stay in the system to serve the next customer with a probability of ' $1-\theta$ '. If the queue length is > 0 , the server will start the next service; else, it remains idle.
- The system follows first-come, first-served queue discipline.
- Let $v(\ell)$ be the vacation density function and $\mathcal{M}(\ell)$ be the distribution function.

$$\gamma(\ell) = \frac{v(\ell)}{1 - \mathcal{M}(\ell)}$$

and therefore

$$v(s) = \gamma(s)e^{-\int_0^\infty \gamma(\ell)d\ell}$$

- All stochastic processes involved are mutually independent.

3. APPLICATION OF THE MODEL

This model is useful in many different kinds of real-world situations. Take a scenario where a consumer goes to a bank to get money from the cashier. Should the client be dissatisfied with draw the amount, he might request another service. Upon receiving the money, the customer's choice is taken into account while making a passbook entry. Following the end of each customer's service, the cashier may choose to work on other duties (Bernoulli vacation).

4. SYSTEM-GOVERNING DEFINITIONS AND EQUATIONS.

$\mathcal{H}_n^{(1)}(\ell, t)$ and $\mathcal{H}_n^{(2)}(\ell, t)$ reflects the probability that, at time 't' with elapsed service time ' ℓ ', the server is active and occupied with service, excluding the one customer being serviced at the server station and ' n ' (≥ 0) customers in the queue.

$\mathcal{M}_n(\ell, t)$ reflects the probability that the server is on vacation with an elapsed vacation time ' ℓ ' while there are $n(\geq 0)$ customers in the queue.

$\mathfrak{J}(t)$ reflects the probability that the system has no customers at time 't' and the server is idle but available.

The following differential-difference equations are framed based on the model defined:

$$\frac{\partial}{\partial \ell} \mathcal{H}_n^{(1)}(\ell, t) + \frac{\partial}{\partial t} \mathcal{H}_n^{(1)}(\ell, t) = -(\lambda + \mu_1(\ell))\mathcal{H}_n^{(1)}(\ell, t) + \lambda(1 - \delta_{n0}) \sum_{k=1}^n c_k \mathcal{H}_{n-k}^{(1)}(\ell, t) \quad (1)$$

$$\frac{\partial}{\partial \ell} \mathcal{H}_n^{(2)}(\ell, t) + \frac{\partial}{\partial t} \mathcal{H}_n^{(2)}(\ell, t) = -(\lambda + \mu_2(\ell))\mathcal{H}_n^{(2)}(\ell, t) + \lambda(1 - \delta_{n0}) \sum_{k=1}^n c_k \mathcal{H}_{n-k}^{(2)}(\ell, t) \quad (2)$$

$$\frac{\partial}{\partial \ell} \mathcal{M}_n(\ell, t) + \frac{\partial}{\partial t} \mathcal{M}_n(\ell, t) = -(\lambda + \gamma(\ell))\mathcal{M}_n(\ell, t) + \lambda(1 - \delta_{n0}) \sum_{k=1}^n c_k \mathcal{M}_{n-k}(\ell, t) \quad (3)$$

$$\frac{d}{dt} \mathfrak{J}(t) = -\lambda \mathfrak{J}(t) + (1-r)(1-\theta) \int_0^\infty \mathcal{H}_0^{(1)}(\ell, t) \mu_1(\ell) d\ell + (1-\theta) \int_0^\infty \mathcal{H}_0^{(2)}(\ell, t) \mu_2(\ell) d\ell \quad (4)$$

Equations (1) to (4) must be solved under the following boundary conditions:

$$\begin{aligned} \mathcal{H}_n^{(1)}(0, \tau) = & \lambda c_{n+1} \mathfrak{J}(\tau) + p \int_0^\infty \mathcal{H}_n^{(1)}(\ell, \tau) \mu_1(\ell) d\ell + q(1-r)(1-\theta) \int_0^\infty \mathcal{H}_{n+1}^{(1)}(\ell, \tau) \mu_1(\ell) d\ell \\ & + (1-\theta) \int_0^\infty \mathcal{H}_{n+1}^{(2)}(\ell, \tau) \mu_2(\ell) d\ell + \int_0^\infty \mathcal{M}_{n+1}(\ell, \tau) \gamma(\ell) d\ell \end{aligned} \quad (5)$$

$$\mathcal{H}_n^{(2)}(0, \tau) = r q \int_0^\infty \mathcal{H}_n^{(1)}(\ell, \tau) \mu_1(\ell) d\ell \quad (6)$$

$$\mathcal{M}_n(0, \tau) = \theta \int_0^\infty \mathcal{H}_n^{(2)}(\ell, \tau) \mu_2(\ell) d\ell + \theta(1-r) q \int_0^\infty \mathcal{H}_n^{(1)}(\ell, \tau) \mu_1(\ell) d\ell \quad (7)$$

The initial conditions are

$$\mathfrak{J}(0) = 1, \mathcal{H}^{(1)}(0) = \mathcal{H}^{(2)}(0) = \mathcal{M}(0) = 0 \quad (8)$$

5. THE TIME-DEPENDENT SOLUTION FOR GENERATING QUEUE FUNCTIONS:

We establish the probability generating functions as follows:

$$\mathcal{H}_q^{(1)}(\ell, z, \tau) = \sum_{n=0}^{\infty} z^n \mathcal{H}_n^{(1)}(\ell, \tau) \tag{9}$$

$$\mathcal{H}_q^{(2)}(\ell, z, \tau) = \sum_{n=0}^{\infty} z^n \mathcal{H}_n^{(2)}(\ell, \tau) \tag{10}$$

$$\mathcal{M}_q(\ell, z, \tau) = \sum_{n=0}^{\infty} z^n \mathcal{M}_n(\ell, \tau) \tag{11}$$

$$C(z) = \sum_{k=1}^{\infty} z^k c_k(\tau) \tag{12}$$

By taking the Laplace transforms of Equations (1) through (7) and applying Equation (8), we derive:

$$\frac{\partial}{\partial \ell} \bar{\mathcal{H}}_n^{(1)}(\ell, \wp) + (\wp + \lambda + \gamma_1(\ell)) \bar{\mathcal{H}}_n^{(1)}(\ell, \wp) = \lambda(1 - \delta_{n0}) \sum_{k=1}^n c_k \bar{\mathcal{H}}_{n-k}^{(1)}(\ell, \wp) \tag{13}$$

$$\frac{\partial}{\partial \ell} \bar{\mathcal{H}}_n^{(2)}(\ell, \wp) + (\wp + \lambda + \gamma_2(\ell)) \bar{\mathcal{H}}_n^{(2)}(\ell, \wp) = \lambda(1 - \delta_{n0}) \sum_{k=1}^n c_k \bar{\mathcal{H}}_{n-k}^{(2)}(\ell, \wp) \tag{14}$$

$$\frac{\partial}{\partial \ell} \bar{\mathcal{M}}_n(\ell, \wp) + (\wp + \lambda + \gamma(\ell)) \bar{\mathcal{M}}_n(\ell, \wp) = \lambda(1 - \delta_{n0}) \sum_{k=1}^n c_k \bar{\mathcal{M}}_{n-k}(\ell, \wp) \tag{15}$$

$$(\wp + \lambda) \bar{\mathcal{J}}(\wp) = 1 + (1-p)(1-r)(1-\theta) \int_0^{\infty} \bar{\mathcal{H}}_0^{(1)}(\ell, \wp) \mu_1(\ell) d\ell + (1-\theta) \int_0^{\infty} \bar{\mathcal{H}}_0^{(2)}(\ell, \wp) \mu_2(\ell) d\ell + \int_0^{\infty} \bar{\mathcal{M}}_0(\ell, \wp) \gamma(\ell) d\ell \tag{16}$$

$$\bar{\mathcal{H}}_n^{(1)}(0, \wp) = \lambda c_{n+1} \bar{\mathcal{J}}(\wp) + p \int_0^{\infty} \bar{\mathcal{H}}_n^{(1)}(\ell, \wp) \mu_1(\ell) d\ell + q(1-r)(1-\theta) \int_0^{\infty} \bar{\mathcal{H}}_{n+1}^{(1)}(\ell, \wp) \mu_1(\ell) d\ell + (1-\theta) \int_0^{\infty} \bar{\mathcal{H}}_{n+1}^{(2)}(\ell, \wp) \mu_2(\ell) d\ell + \int_0^{\infty} \bar{\mathcal{M}}_{n+1}(\ell, \wp) \gamma(\ell) d\ell \tag{17}$$

$$\bar{\mathcal{H}}_n^{(2)}(0, \wp) = r q \int_0^{\infty} \bar{\mathcal{H}}_n^{(1)}(\ell, \wp) \mu_1(\ell) d\ell \tag{18}$$

$$\bar{\mathcal{M}}_n(0, \wp) = \theta \int_0^{\infty} \bar{\mathcal{H}}_n^{(2)}(\ell, \wp) \mu_2(\ell) d\ell + \theta(1-r) q \int_0^{\infty} \bar{\mathcal{H}}_n^{(1)}(\ell, \wp) \mu_1(\ell) d\ell \tag{19}$$

Equations (13) multiplied by suitable powers of z, summed over n, and simplified using (9). Following algebraic computations, we obtain:

$$\frac{\partial}{\partial \ell} \bar{\mathcal{H}}_q^{(1)}(\ell, z, \wp) + (\wp + [\lambda(1 - C(z)) + \mu_1(\ell)]) \bar{\mathcal{H}}_q^{(1)}(\ell, z, \wp) = 0 \tag{20}$$

By using (10) and carrying out similar procedures on (14), we derive:

$$\frac{\partial}{\partial \ell} \bar{\mathcal{H}}_q^{(2)}(\ell, z, \wp) + (\wp + [\lambda(1 - C(z)) + \mu_2(\ell)]) \bar{\mathcal{H}}_q^{(2)}(\ell, z, \wp) = 0 \tag{21}$$

By using (11) and carrying out similar procedures on (15), we derive:

$$\frac{\partial}{\partial \ell} \bar{\mathcal{M}}_q(\ell, z, \wp) + (\wp + [\lambda(1 - C(z)) + \gamma(\ell)]) \bar{\mathcal{M}}_q(\ell, z, \wp) = 0 \tag{22}$$

The two sides of equation (17) are then multiplied by z, the sum over n from 0 to ∞ , we get

$$\begin{aligned} \bar{H}_q^{(1)}(0, z, \varphi) = & \lambda C(z) \bar{J}(\varphi) + pz \int_0^\infty \bar{H}_q^{(1)}(\ell, z, \varphi) \mu_1(\ell) d\ell + (1-p)(1-r)(1-\theta) \\ & \left[\int_0^\infty \bar{H}_q^{(1)}(0, z, \varphi) \mu_1(\ell) d\ell - \int_0^\infty \bar{H}_0^{(1)}(\ell, \varphi) \mu_1(\ell) d\ell \right] \\ & + (1-\theta) \left[\int_0^\infty \bar{H}_q^{(2)}(0, z, \varphi) \mu_2(\ell) d\ell - \int_0^\infty \bar{H}_0^{(2)}(\ell, \varphi) \mu_2(\ell) d\ell \right] \\ & + \left[\int_0^\infty \bar{M}_q(0, z, \varphi) \gamma(\ell) d\ell - \int_0^\infty \bar{M}_0(\ell, \varphi) \gamma(\ell) d\ell \right] \end{aligned} \quad (23)$$

By carrying out analogous procedures on equations (18) and (19), we obtain

$$\bar{H}_q^{(2)}(0, z, s) = r(1-p) \int_0^\infty \bar{H}_q^{(1)}(\ell, z, s) \mu_1(\ell) d\ell \quad (24)$$

$$\bar{M}_q(0, z, \varphi) = \theta \int_0^\infty \bar{H}_q^{(2)}(\ell, z, \varphi) \mu_2(\ell) d\ell + \theta(1-r)(1-p) \int_0^\infty \bar{H}_q^{(1)}(\ell, z, \varphi) \mu_1(\ell) d\ell \quad (25)$$

Using equation (16) in (23), we get

$$\begin{aligned} z\bar{H}_q^{(1)}(0, z, \varphi) = & 1 - (\varphi + \lambda(1-C(z))\bar{J}(\varphi)) + pz \int_0^\infty \bar{H}_q^{(1)}(\ell, z, \varphi) \mu_1(\ell) d\ell \\ & + (1-p)(1-r)(1-\theta) \int_0^\infty \bar{H}_q^{(1)}(\ell, z, \varphi) \mu_1(\ell) d\ell \\ & + (1-\theta) \int_0^\infty \bar{H}_q^{(2)}(\ell, z, \varphi) \mu_2(\ell) d\ell + \int_0^\infty \bar{M}_q(\ell, z, \varphi) \gamma(\ell) d\ell \end{aligned} \quad (26)$$

Integrating equation (20), (21) and (22), from 0 to ℓ yields

$$\bar{H}_q^{(1)}(\ell, z, \varphi) = \bar{H}_q^{(1)}(0, z, \varphi) e^{-(\varphi + \lambda(1-C(z))\ell - \int_0^\ell \mu_1(t) dt} \quad (27)$$

$$\bar{H}_q^{(2)}(\ell, z, \varphi) = \bar{H}_q^{(2)}(0, z, \varphi) e^{-(s + \lambda(1-C(z))\ell - \int_0^\ell \mu_2(t) dt} \quad (28)$$

$$\bar{M}_q(\ell, z, \varphi) = \bar{M}_q(0, z, \varphi) e^{-(\varphi + \lambda(1-C(z))\ell - \int_0^\ell \gamma(t) dt} \quad (29)$$

Integrating equation (27) to (29) by parts with respect to x yields, we get

$$\bar{H}_q^{(1)}(z, \varphi) = \bar{H}_q^{(1)}(0, z, \varphi) \left[\frac{1 - \bar{\mathfrak{R}}_1[f(z, \varphi)]}{[f(z, \varphi)]} \right] \quad (30)$$

where

$$\bar{\mathfrak{R}}_1[f(z, \varphi)] = \int_0^\infty e^{-f(z, \varphi)\ell} d\mathfrak{R}_1(\ell)$$

$$\bar{H}_q^{(2)}(z, \varphi) = \bar{H}_q^{(2)}(0, z, \varphi) \left[\frac{1 - \bar{\mathfrak{R}}_2[f(z, \varphi)]}{[f(z, \varphi)]} \right] \quad (31)$$

where

$$\bar{\mathfrak{R}}_2[f(z, \varphi)] = \int_0^\infty e^{-f(z, \varphi)\ell} d\mathfrak{R}_2(\ell)$$

$$\bar{M}_q(z, \varphi) = \bar{M}_q(0, z, \varphi) \left[\frac{1 - \bar{\mathcal{M}}[f(z, \varphi)]}{[f(z, \varphi)]} \right] \quad (32)$$

where

$$\bar{\mathcal{M}}[f(z, \varphi)] = \int_0^\infty e^{-f(z, \varphi)\ell} d\mathcal{H}(\ell)$$

where $f(z, \varphi) = \varphi + \lambda(1-C(z))$

Now, by multiplying $\mu_1(\ell)$ by both sides of equation (27), $\mu_2(\ell)$ by (28), & $\gamma(\ell)$ by (29), & integrating over ℓ , we get

$$\int_0^\infty \bar{\mathcal{H}}_q^{(1)}(\ell, z, \varphi) \mu_1(\ell) d\ell = \bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] \quad (33)$$

$$\int_0^\infty \bar{\mathcal{H}}_q^{(2)}(\ell, z, \varphi) \mu_2(\ell) d\ell = \bar{\mathcal{H}}_q^{(2)}(0, z, \varphi) \bar{\mathfrak{R}}_2[f(z, \varphi)] \quad (34)$$

$$\int_0^\infty \bar{\mathcal{M}}_q(\ell, z, \varphi) \gamma(\ell) = \bar{\mathcal{M}}_q(0, z, \varphi) \bar{\mathcal{M}}[f(z, \varphi)] \quad (35)$$

Substituting equation (34) in (24), we obtain

$$\bar{\mathcal{H}}_q^{(2)}(0, z, \varphi) = r(1-p) \bar{\mathcal{H}}_q^{(2)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] \quad (36)$$

Similarly using equation (33) and (34) in (25), we get

$$\bar{\mathcal{M}}_q(0, z, \varphi) = \theta(1-r)q \bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] + \theta r \bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] \bar{\mathfrak{R}}_2[f(z, \varphi)] \quad (37)$$

Using equation (33), (34) and (35) in (26) and solving $\bar{\mathcal{H}}_q(0, z, \varphi)$

$$\bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) = \frac{f(z, \varphi) \bar{\mathfrak{J}}(\varphi)}{\left[z - \bar{\mathfrak{R}}_1[f(z, \varphi)]pz + q(1-r)(1-\theta) + (1-\theta)r(1-p) \bar{\mathfrak{R}}_2[f(z, \varphi)] \right] + \theta(1-r)(1-p) \bar{\mathcal{M}}[f(z, \varphi)] + r\theta(1-p) \bar{\mathfrak{R}}_2[f(z, \varphi)] \bar{\mathcal{M}}[f(z, \varphi)]} \quad (38)$$

Equation (30), (31) and (32) becomes

$$\bar{\mathcal{H}}_q^{(1)}(z, \varphi) = \bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) \left[\frac{1 - \bar{\mathfrak{R}}_1[f(z, \varphi)]}{f(z, \varphi)} \right] \quad (39)$$

$$\bar{\mathcal{H}}_q^{(2)}(z, \varphi) = r(1-p) \bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] \left[\frac{1 - \bar{\mathfrak{R}}_2[f(z, \varphi)]}{f(z, \varphi)} \right] \quad (40)$$

$$\begin{aligned} \bar{\mathcal{M}}_q(z, \varphi) = & \{ \theta(1-r)(1-p) \bar{\mathcal{M}}_q^{(1)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] \\ & + \theta r(1-p) \bar{\mathcal{H}}_q^{(1)}(0, z, \varphi) \bar{\mathfrak{R}}_1[f(z, \varphi)] \} \left[\frac{1 - \bar{\mathcal{M}}[f(z, \varphi)]}{f(z, \varphi)} \right] \end{aligned} \quad (41)$$

6. THE STEADY-STATE RESULTS

The steady-state probability distribution for our queueing model is what we want to achieve. We exclude the time argument 't' from the time-dependent analysis in order to get the steady-state probability. This can be made easier by applying the well-known Tauberian property.

$$\lim_{\varphi \rightarrow 0} \varphi \bar{f}(\varphi) = \lim_{t \rightarrow \infty} f(t) \quad (42)$$

We will use the normalizing condition

$$\bar{\mathcal{H}}_q^{(1)}(1) + \bar{\mathcal{H}}_q^{(2)}(1) + \bar{\mathcal{M}}_q(1) + \bar{\mathfrak{J}} = 1 \quad (43)$$

The probability generating function of the queue size $\mathcal{P}(z)$ irrespective of the state of the system.

$$\begin{aligned} \mathcal{P}(z) &= \mathcal{H}_q^{(1)}(z) + \mathcal{H}_q^{(2)}(z) + \mathcal{M}_q(z) \\ &= \mathcal{H}_q^{(1)}(0, z) \left[\frac{1 - \mathfrak{R}_1[f(z)]}{f(z)} \right] + \mathcal{H}_q^{(2)}(0, z) \left[\frac{1 - \mathfrak{R}_2[f(z)]}{f(z)} \right] + \mathcal{M}_q(0, z) \left[\frac{1 - \mathcal{M}[f(z)]}{(z)} \right] \end{aligned} \quad (44)$$

where $\mathcal{H}_q^{(1)}(z)$, $\mathcal{H}_q^{(2)}(z)$ and $\mathcal{M}_q(z)$ are given by the following equations.

$$\mathcal{H}_q^{(1)}(z) = \mathcal{H}_q^{(1)}(z, 0) \left[\frac{1 - \mathfrak{R}_1[f(z)]}{f(z)} \right] \quad (45)$$

$$\mathcal{H}_q^{(2)}(z) = r(1 - p)\mathcal{H}_q^{(1)}(z, 0)\mathfrak{R}_1[f(z)] \left[\frac{1 - \mathfrak{R}_2[f(z)]}{f(z)} \right] \quad (46)$$

$$\begin{aligned} \mathcal{M}_q(z) &= \mathcal{H}_q^{(1)}(z, 0) \left\{ \theta(1 - r)(1 - p)\mathfrak{R}_1[f(z)] \left[\frac{1 - \mathcal{M}[f(z)]}{f(z)} \right] \right. \\ &\quad \left. + \theta r(1 - p)\mathfrak{R}_1[f(z)]\mathfrak{R}_2[f(z)] \left[\frac{1 - \mathcal{M}[f(z)]}{f(z)} \right] \right\} \end{aligned} \quad (47)$$

$$\mathcal{P}(z) = \frac{\left[\mathfrak{J}\{[1 - \mathfrak{R}_1[f(z)]] + \Psi\theta r[\mathfrak{R}_2[f(z)]] [1 - \mathcal{M}[f(z)]]\} \right.}{\left[z - r\Psi(1 - \theta)\mathfrak{R}_2[f(z)] - \Psi\theta r\mathfrak{R}_2[f(z)]\mathcal{M}[f(z)] \right.} \quad (48)$$

$$\left. \left. + \theta(1 - r)\Psi[1 - \mathcal{M}[f(z)]] + \Psi r[1 - \mathfrak{R}_1[f(z)]] \right\} \right]$$

$$\left[-\mathfrak{R}_1[f(z)]pz - (1 - r)\Psi(1 - \theta) - \theta\Psi(1 - r)\mathcal{M}[f(z)] \right]$$

where

$$\Psi = (1 - p)\mathfrak{R}_1[f(z)]$$

Observing that for $z=1$, $\mathcal{P}(z)$ takes on an indeterminate form of $0/0$, we apply L'Hopital's rule on equation (44) using the fact $\mathfrak{R}_1(0) = 1$, $\mathfrak{R}_2(0) = 1$, $\mathcal{M}(0) = 1$, $-\mathcal{M}'(0) = E(V)$, $-\mathfrak{R}_i'(0) = E(\mathfrak{R}_i)$, and $\mathfrak{R}_i''(0) = E(\mathfrak{R}_i^2)$. We get,

$$\mathcal{P}(1) = \frac{\mathfrak{J}[\lambda[-E(X)]]\{E(\mathfrak{R}_1) + rE(\mathfrak{R}_2)(1 - p) + (1 - p)E(V)\theta\}}{-\lambda[E(X)] + (1 - p)\{E(\mathfrak{R}_1)p + (1 - p)E(\mathfrak{R}_1) + r(1 - p)E(\mathfrak{R}_2) + \theta(1 - p)E(V)\}} \quad (49)$$

Therefore adding \mathfrak{J} to equation (49), we get

$$\mathfrak{J} = \frac{-\lambda[E(X)] + (1 - p)\{E(\mathfrak{R}_1) + r(1 - p)E(\mathfrak{R}_2) + \theta(1 - p)E(V)\}}{-2\lambda[E(X)] + (1 - p)\{E(\mathfrak{R}_1) + r(1 - p)E(\mathfrak{R}_2) + \theta(1 - p)E(V)\}} \quad (50)$$

Consequently, The system's utilization factor is established by

$$\rho = 1 - \mathfrak{J}$$

where $\rho < 1$ is the stability condition under which steady state exists, for the model.

7. PERFORMANCE METRICS

Let L_q be the mean number of customers in the queue. Following this,

$$L_q = \left. \frac{d}{dz} \mathcal{P}(z) \right|_{z=1}$$

$$L_q = \lim_{z \rightarrow 1} \frac{d}{dz} \mathcal{P}(z)$$

$$\begin{aligned}
 &= \lim_{z \rightarrow 1} \frac{[\Re'(z)\Re''(z) - \Re'(z)\Re''(z)]}{2[\Re'(z)]^2} \\
 &= \frac{[\Re'(1)\Re''(1) - \Re'(1)\Re''(1)]}{2[\Re'(1)]^2} \tag{51}
 \end{aligned}$$

where,

$$\Re'(1) = I[\lambda[-E(X)][E(\Re_1) + rE(\Re_2)(1-p) + \theta(1-p)E(V)]]$$

$$\begin{aligned}
 \Re''(1) = & -I[\lambda^2[E(X)]^2[E(\Re_1)^2 + r(1-p)[2E(\Re_1)E(\Re_2) + E(\Re_2)^2] \\
 & + \theta(1-p)[2E(\Re_1)E(V) + E(V)^2] + \theta r(1-p)[2E(\Re_2)E(V)]] \\
 & - \lambda[E(X^2)][E(\Re_1) + r(1-p)E(\Re_2) + \theta(1-p)E(V)]
 \end{aligned}$$

$$\Re'(1) = -\lambda[E(X)] + (1-p)[E(\Re_1) + r(1-p)E(\Re_2) + \theta(1-p)E(V)]$$

$$\begin{aligned}
 \Re''(1) = & \lambda[E(X)][2E(\Re_1)p - \lambda^2[E(X)]^2[E(\Re_1)^2 + r(1-p)[E(\Re_2)^2 + 2E(\Re_1)E(\Re_2)]] \\
 & + \theta(1-p)[E(v)^2 + 2E(\Re_1)E(\Re_2)]] + \theta r(1-p)[2E(\Re_2)E(V)] \\
 & - \lambda[E(X^2)][E(\Re_1) + r(1-p)[E(\Re_2)] + \theta(1-p)E(V)]
 \end{aligned}$$

Let \mathfrak{W}_q be the mean time while customers in the line have been waiting.

By using Little's formula

$$\mathfrak{W}_q = \frac{\mathfrak{L}_q}{\lambda} \tag{52}$$

8. THE NUMERICAL RESULTS

In this section, we shows various factors affect system performance metrics using MATLAB. We assume that vacation and service time are exponentially distributed. All the parameter values are selected to satisfy the stability condition.

Table 1: Service rate effectiveness

μ_1	\mathfrak{I}	\mathfrak{L}_q	\mathfrak{W}_q
11	0.5118	0.0174	0.0348
12	0.5508	0.0167	0.0335
13	0.5838	0.0159	0.0318
14	0.6121	0.0150	0.0300
15	0.6367	0.0141	0.0282
16	0.6581	0.0132	0.0265

From table 1, shows that idle time increase, mean waiting time and mean queue length decreases while service rate increase.

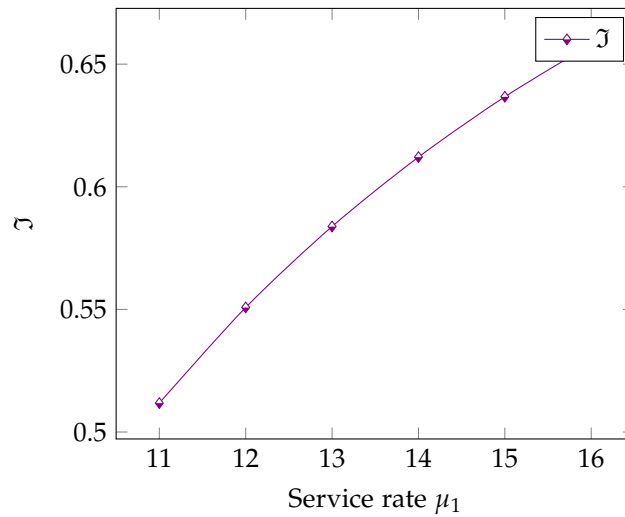


Figure 1: Service rate vs Idle

Figure 1 clearly demonstrates that as the mandatory service times increase, the idle time also increases.

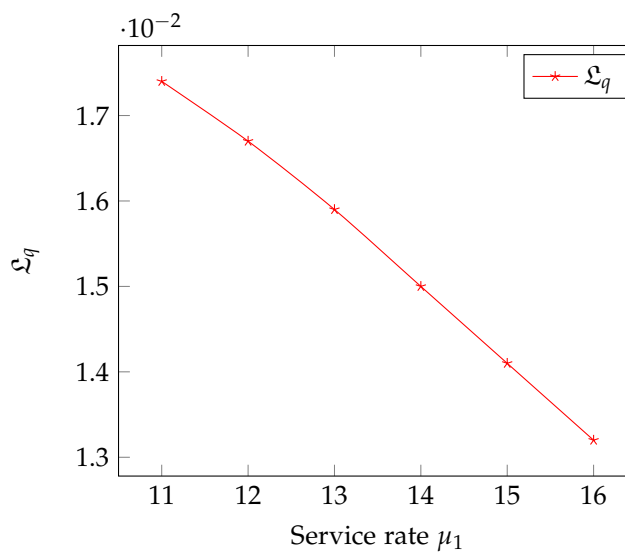


Figure 2: Service rate vs Queue length

Figure 2 clearly demonstrates that as the mandatory service times increase, the mean queue length decreases.

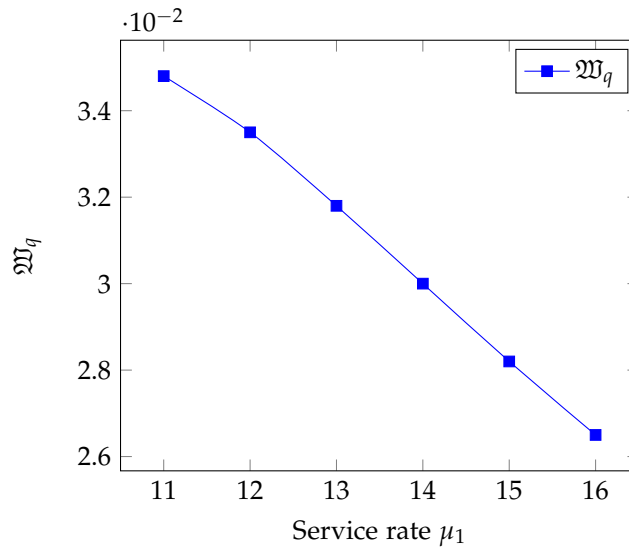


Figure 3: Service rate vs Waiting time

Figure 3 indicates that as the mandatory service times increase, the mean waiting time of customers in the queue decreases.

9. CONCLUSION

We analyzed an $M^X/G/1$ queueing system with optional second service, feedback, and Bernoulli vacation. Key performance indicators are obtained by applying the supplementary variable approach. Numerical outcomes validate using analytical results. Both the mean waiting time and mean queue length decrease while increase in service rate. For instance, if service rates rise in the banking industry, this model helps reduce customer mean waiting times and mean queue length.

REFERENCES

- [1] Arivudainambi, D. and Gowsalya, M.(2017). Analysis of an $M/G/1$ retrial queue with Bernoulli vacation, two types of service and starting failure, *International Journal of Artificial Intelligence and Soft Computing*, 6(3) : 222 – 249.
- [2] Arivudainambi, D. and Gowsalya, M.(2018). A single server non-Markovian retrial queue with two types of service and Bernoulli vacation. *International Journal of Operational Research*, 33(1) : 55 – 81.
- [3] Ayyappan, G. and Deepa, T. (2018). Analysis of batch arrival bulk service queue with multiple vacation closedown essential and optional repair, *Applications and Applied Mathematics*. 13(2) : 578 – 598.
- [4] Ayyappan,G. and Somasundaram, B.(2019).Analysis of Two Stage $M^{[X1]}, M^{[X2]}/G1, G2/1$ Retrial G-queue with Discretionary Priority Services, Working Breakdown, Bernoulli Vacation, Preferred and Impatient Units, *Applications and Applied Mathematics: An International Journal (AAM)*, 14(2) : 640 – 671.
- [5] Chandrasekaran, V M., Indhira, K., Rajadurai,P., and Saravanarajan MC.(2015). Analysis of an $M^{[X]}/G/1$ feedback retrial queue with two phase service, Bernoulli vacation, delaying repair and orbit search, *Advances in Physics Theories and Applications*, 40(2) : 668 – 673.
- [6] Chen, P., Zhou, Y. and Li, C.(2015). Discrete-time retrial queue with Bernoulli vacation, preemptive resume and feedback customers. *Journal of Industrial Engineering and Management*, 8(4) : 1236 – 1250.

- [7] Choudhury, G. and Paul, M. (2006). A Batch Arrival Queue with a Second Optional Service Channel Under N-Policy, *Stochastic Analysis and Applications*, 24(1) : 1 – 21.
- [8] Kalidass, K. and Kasturi, R. (2014). A two phase service $M/G/1$ queue with a finite number of immediate Bernoulli feedbacks. *OPSEARCH*, 51(2) : 201 – 218.
- [9] Karpagam, S., Ayyappan, G. and Somasundaram, B. (2020). A Bulk Queueing System with Rework in Manufacturing Industry with Starting Failure and Single Vacation, *International Journal of Applied and Computational Mathematics*, 6(6) : 174.
- [10] Li, J. and Wang, J. (2006). An $M/G/1$ retrial queue with second multi-optional service, feedback and unreliable server, *Applied Mathematics-A Journal of Chinese Universities*, 21(3) : 252 – 262.
- [11] Madan, K.C. (2000). An $M/G/1$ queue with second optional service, *Queueing Systems*, 34(1) : 37 – 46.
- [12] Madan, K. C. (1992). An $M/G/1$ queueing system with compulsory server vacations, *Trabajos de Investigacion Operativa*, 7(1) : 105 – 115.
- [13] Madhu, J. and Kaur, S. (2021). Bernoulli vacation model for $M^X/G/1$ unreliable server retrial queue with bernoulli feedback, balking and optional service. *RAIRO-Operation Research*, 55(2) : 2027 – S2053.
- [14] Madhu, J., Sharma, G.C., and Sharma, R. (2012). A batch arrival retrial queueing system for essential and optional services with server breakdown and Bernoulli vacation. *International Journal of Internet and Enterprise Management*, 8(1) 16 – 45.
- [15] Maragatha Sundari, S. and Srinivasan, S. (2012). Analysis of transient behaviour of $M/G/1$ queue with single vacation, *International Journal of Pure and Applied Mathematics*, 76(1) : 149 – 156.
- [16] Monita, B., Madan, K.C. and Eldabi, T. (2013). A Two Stage Batch Arrival Queue with Reneging during Vacation and Breakdown Periods, *American Journal of Operations Research*, 3(6) : 570 – 580.
- [17] Pavai Madheswari, S. and Suganthi, P. (2016). An $M/G/1$ Retrial Queue with Second Optional Service and Starting Failure under Modified Bernoulli Vacation, *Transylvanian Review*, 24(10) : 12 – 30.
- [18] Rajadurai, P., Saravananarajan, M.C. and Chandrasekaran, V.M. (2017). An $M/G/1$ retrial G-queue with optional re-service, impatient customers, multiple working vacations and vacation interruption, *International Journal on Operational Research*, 30(1) : 33 – 64.
- [19] Rajam, V. and Uma, S. (2021). Modified Bernoulli vacation batch arrival and retrial clients in a single server queueing model with server utilization, *Malaya Journal Of Matematik*, 9(1) : 46 – 51.
- [20] Shanmugasundaram, S. and Sivaram, G. (2020). $M/G/1$ Feedback Queue When Server is Off and on Vacation, *International Journal of Applied Engineering Research*, 15(10) : 1025 – 1028.
- [21] Singh, C.J. , Jain, M. and Kumar, B. (2014). Analysis of $M^X/G/1$ queueing model with balking and vacation, *International Journal of Operational Research*, 19(2) : 154 – 173.
- [22] Somasundaram, B., Karpagam, S., Lokesh, R. and Mary, A.K.S. (2023). An $M^X/G/1$ queue with optional service and working breakdown, *Ural Mathematical Journal*, 9(1) : 162 – 175.
- [23] Thangaraj, V. and Vanitha, S. (2010). $M/G/1$ queue with Two-Stage Heterogeneous Service Compulsory Server Vacation and Random Breakdowns, *International Journal of Contemporary Mathematical Sciences*, 5(7) : 307 – 322.
- [24] Varalakshmi, M., Chandrasekaran, V. M. and Saravananarajan M C. (2018). A Single Server Queue with Immediate Feedback, Working Vacation and Server Breakdown. *International Journal of Engineering & Technology*, 7(4.10):476-479.
- [25] Varalakshmi, M. and Rajadurai, P. (2021). A Priority Retrial queue with Bernoulli Vacation, Feedback and Working Breakdowns. *NOVYI MIR Research Journal*, 6(2):34-40.
- [26] Wang, J. and Li, J. (2008). A Repairable $M/G/1$ Retrial Queue with Bernoulli Vacation and Two-Phase Service. *Quality Technology & Quantitative Management*, 5(2): 179-192.

INVERTED DAGUM DISTRIBUTION: PROPERTIES AND APPLICATION TO LIFETIME DATASET.

ABDULHAMEED A. OSI*, SHAMSUDDEEN A. SABO AND IBRAHIM Z. MUSA

Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Nigeria.
abuammarosi@gmail.com, aaosi@kust.edu.ng

Abstract

This article presents the introduction of a novel univariate probability distribution termed the inverted Dagum distribution. Extensive analysis of the statistical properties of this distribution, including the hazard function, survival function, Renyi's entropy, quantile function, and the distribution of the order statistics, was conducted. Parameter estimation of the model was performed utilizing the maximum likelihood method, with the consistency of the estimates validated through Monte Carlo simulation. Furthermore, the applicability of the proposed distribution was demonstrated through the analysis of two real datasets.

Keywords: Dagum distribution, Inverse transformation, Maximum likelihood estimation, Monte Carlo Simulation, COVID-19 data

1. INTRODUCTION

Dagum distribution was named after Camilo Dagum who proposed it in the 1970s to fit wealth and income data as well as accommodate heavy-tailed models. Dagum distribution can be of two forms (the three-parameter and the four-parameter) respectively referred to as Dagum type I distribution [1] and Dagum type II distribution [2].

Definition: A random variable is said to have type I Dagum distribution with parameter if its cumulative density function and probability density function are given by:

$$F_{\alpha DD}(x; \theta, \alpha, \lambda) = [1 + \lambda x^{-\alpha}]^{-\theta} \quad x > 0 \quad (1)$$

$$f_{\alpha DD}(x; \theta, \alpha, \lambda) = \theta \alpha \lambda x^{-\alpha-1} (1 + \lambda x^{-\alpha})^{-\theta-1} \quad (2)$$

Where $\alpha, \theta, \lambda > 0$. With λ is the scale parameter and α, θ are the two shape parameters. The latter controls the tail weight and the former controls the size of the distribution. It can be observed that for $\theta = 1$, Eq. 1 becomes the log-logistic distribution proposed by [3]. It can also be observed that Eq. 1 is a Burr III distribution with an additional λ . One crucial characteristic possessed by Dagum distribution is that, in addition to its flexibility, its hazard function can be decreased, up-side down, bathtub, and then up-side bathtub shaped [4]. A lot of researchers utilized this behavior to study the Dagum distribution in several fields. An extensive review of the Dagum distribution and its application was detailed in [6] and [7]. The parameters of the Dagum distribution with censored samples was studied by [5] while [8] utilized TL-moments for similar purpose. Some classes of weighted Dagum and related distributions were proposed by [10] and the five-parameter beta-Dagum by [4]. Considering the properties of McDonald, Kumaraswamy and Dagum distribution, two new hybrid distributions called Mc-Dagum and Kum-Dagum

distributions were proposed by [10]. Numerous ways of extending well-known distributions were suggested by applied statisticians. An extended Dagum, distribution was proposed by [17]. The distribution of the reciprocal of any well-known distribution are their corresponding inverted distributions. This had received numerous attentions from researchers in the field of distribution theory. Few of these works are; inverted gamma by [12], inverted exponential by [19], inverted Weibull by [20] to mention a few. In this paper, we proposed an inverted Dagum distribution. Some statistical properties of the proposed distribution were derived. A simulation study is further performed to check the flexibility and usefulness of the proposed distribution.

2. METHODS

2.1. Proposed Distribution

Motivated by [10] and the literature cited therein, we proposed an Inverted Dagum Distribution (IDD). A random variable X is said to have an Inverted Dagum Distribution, if the following transformation is applicable $X = \frac{1}{Y}$. Where Y is the Dagum distribution random variable with pdf and cdf expressed respectively in Eq. 1 and Eq. 2. By applying cdf technique:

$$F_Y(y) = P(Y \leq y) = P\left(X > \frac{1}{y}\right) = 1 - P\left(X \leq \frac{1}{y}\right) \quad (3)$$

The pdf and cdf of the proposed distribution can be expressed as;

$$F_{IDD}(x; \theta, \alpha, \lambda) = 1 - [1 + \lambda x^\alpha]^{-\theta} \quad (4)$$

$$f_{IDD}(x; \theta, \alpha, \lambda) = \theta \alpha \lambda x^{\alpha-1} (1 + \lambda x^\alpha)^{-\theta-1} \quad x > 0 \quad (5)$$

2.2. Some Mathematical properties

Here, some important Mathematical and Statistical properties of the proposed inverted Dagum distribution like quantile function, hazard function, moments, moment generating function, Renyi's entropy were presented.

2.2.1 Quantile Function

Let X be a random variable with pdf given in Eq.(5). The quantile function of X (the proposed distribution) can be expressed as;

$$Q(u; \theta, \alpha, \lambda) = \left\{ \lambda^{-1} \left[(1-u)^{-\frac{1}{\theta}} - 1 \right] \right\}^{\frac{1}{\alpha}} \quad (6)$$

2.2.2 Moments

The moments of any distribution tell a lot about its features. Characteristics like tendency, skewness, dispersion, kurtosis etc can be observed through moments. If the random variable X has an inverted Dagum distribution, then its r^{th} moment about zero can be expressed as;

$$E(X^r) = \int_0^\infty x^r f(x) dx \quad (7)$$

$$= \theta \alpha \lambda^{2-r} \int_0^\infty (\lambda x^\alpha)^{r-1} (1 + \lambda x^\alpha)^{-((r+1)+(\theta-r))} dx \quad (8)$$

$$= \theta\alpha\lambda^{2-r}B(1+r, \theta-r) \tag{9}$$

From Eq. (13), the mean and the variance of IDD can be expressed respectively as;

$$E(X) = \theta\alpha\lambda B(2, \theta-1) \tag{10}$$

$$\text{Variance} = E(X^2) - [E(X)]^2 \tag{11}$$

$$= \theta\alpha B(3, \theta-2) - [\theta\alpha\lambda B(2, \theta-1)]^2 \tag{12}$$

2.2.3 Moment Generating Function

Many important features and characteristics of a distribution can be observed through its moment-generating function (mgf). Let X be a random variable having an inverted Dagum distribution with pdf given in Eq.(5), using the definition of moment generating function of X utilizing Eq. (5) we have;

$$\begin{aligned} M_X(t) &= E[e^{tx}] = \int_0^\infty e^{tx} f(x) dx \\ &= \theta\lambda\alpha \int_0^\infty e^{tx} x^{\alpha-1} (1 + \lambda x^\alpha)^{-\theta-1} dx \\ &= \theta\lambda\alpha \int_0^\infty \sum_{k=0}^\infty \frac{t^k x^k}{k!} x^{\alpha-1} (1 + \lambda x^\alpha)^{-\theta-1} dx \\ &= \theta\alpha \sum_{k=0}^\infty \frac{\lambda^{2-k} t^k}{k!} \int_0^\infty (\lambda x^\alpha)^{k-1} (1 + \lambda x^\alpha)^{-((r+1)+(\theta-r))} dx \end{aligned} \tag{13}$$

Eq. (13) can be expressed in a more compact form as;

$$M_X(t) = \theta\alpha M_k(\lambda) B((r+1), (\theta-r)) \tag{14}$$

Where; $M_k(\lambda) = \sum_{k=0}^\infty \frac{\lambda^{2-k} t^k}{k!}$

2.2.4 Survival Function

Let X be a random variable having an inverted Dagum distribution, its survival function is given by;

$$R_X(x; \theta, \alpha, \lambda) = 1 - F_{IDD}(x; \theta, \alpha, \lambda) \tag{15}$$

$$= [1 + \lambda x^\alpha]^{-\theta} \tag{16}$$

2.2.5 Hazard Function

The reliability characteristics of a system can be checked through its hazard rate function. The hazard rate function of the inverted Dagum distribution is given by;

$$h_X(x; \theta, \alpha, \lambda) = \frac{f(x; \theta, \alpha, \lambda)}{1 - F_X(x; \theta, \alpha, \lambda)} \tag{17}$$

$$= \theta\alpha\lambda x^{\alpha-1} (1 + \lambda x^\alpha)^{-1} \tag{18}$$

2.2.6 Renyi's Entropy

This is the measure of the variation of the uncertainty of the distribution. A large value of entropy indicates the greater uncertainty in the data. Renyi's Entropy is defined as:

$$\tau_R(\gamma) = \frac{1}{1-\gamma} \log \left(\int_0^{\infty} f^\gamma(x) dx \right) \quad (19)$$

From Eq. (5) we have:

$$f^\gamma(x) = (\theta\alpha)^\gamma \lambda^\gamma x^{\gamma(\alpha-1)} \left[(1 + \lambda x^\alpha)^{-\gamma(1+\theta)} \right] \quad (20)$$

2.2.7 Order Statistics

Suppose X_1, X_2, \dots, X_n are random samples having an inverted Dagum distribution. Let $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ denote the order statistics corresponding to the samples. The pdf and cdf of the order statistics are given as;

$$f_{r:n}(t) = \frac{\theta\alpha\lambda x^{\alpha-1} n!}{(r-1)!(n-r)!} \sum_{i=k}^n (-1)^n \binom{n-r}{u} (1 + \lambda x_i^\alpha)^{-\theta-1} [1 - (1 - \lambda x_i^\alpha)]^{r-1+u} \quad (21)$$

$$F_{r:n}(x) = \sum_{i=k}^n \sum_{u=0}^{n-r} (-1)^u \binom{n}{l} \binom{n-r}{u} \left[1 - (1 + \lambda x_i^\alpha)^{-\theta} \right]^{l+u} \quad (22)$$

2.2.8 Estimation

Here method of maximum likelihood method is used to estimate the parameters of the proposed inverted Dagum distribution. The likelihood function is given by;

$$L(\theta, \lambda, \alpha) = (\theta\lambda\alpha)^n \prod_{i=1}^n x_i^{\alpha-1} (1 + \lambda x_i^\alpha)^{-\theta-1} \quad (23)$$

The log-likelihood function is;

$$\ln L(\theta, \lambda, \alpha) = n \log \theta + n \log \alpha + n \log \lambda + (\alpha - 1) \sum_{i=1}^n \log x_i - (\theta + 1) \sum_{i=1}^n \log (1 + \lambda x_i^\alpha) \quad (24)$$

The log-likelihood function is maximized by differentiating Eq. (24) w.r.t to the parameters which yields;

$$\frac{\partial}{\partial \theta} \ln L(\theta, \lambda, \alpha) = \frac{n}{\theta} - \sum_{i=1}^n \ln (1 + \lambda x_i^\alpha) = 0 \quad (25)$$

$$\frac{\partial}{\partial \alpha} \ln L(\theta, \lambda, \alpha) = \frac{n}{\alpha} + \sum_{i=1}^n \ln x_i - (\theta + 1) \sum_{i=1}^n \frac{\alpha \lambda x_i^{\alpha-1}}{1 + \lambda x_i^\alpha} = 0 \quad (26)$$

$$\frac{\partial}{\partial \lambda} \ln L(\theta, \lambda, \alpha) = \frac{n}{\lambda} - (\theta + 1) \sum_{i=1}^n \frac{x_i^\alpha}{1 + \lambda x_i^\alpha} = 0 \quad (27)$$

The maximum likelihood estimates of the parameters can be obtained by solving the non-linear equations (25), (26), and (27) numerically.

Fig. 1 and Fig.3 display the density and CDF plots of IDD as it can be observed the density exhibits right-skewed and reverse J-shaped while Fig. 2 and Fig. 4 show the hazard function and survival function plots respectively. the hazard has increasing, decreasing, and bathtub shapes.

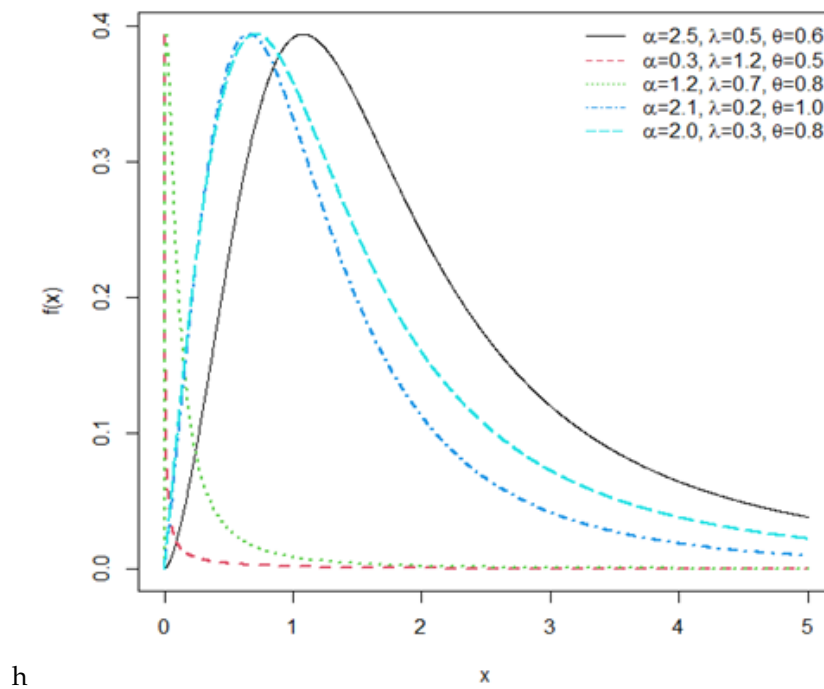


Figure 1: The IDD pdf plot

3. RESULTS AND DISCUSSION

3.1. Simulation

Here, we performed a numerical simulation study to assess the performance of the MLE procedure for estimating IDD parameters based on the Monte Carlo simulation method. A random sample of size $n = 50, 100, 250, 500$ and 1000 with 10000 replicate was generated from IDD using the quantile function $\tau = \left[\frac{1}{\lambda} (1 - u)^{-\theta} - 1 \right]^{\frac{1}{\alpha}}$, where u is uniform $(0, 1)$. Two sets of values of the parameters $\theta = 0.5, \alpha = 4, \lambda = 0.1$ and $\theta = 1.5, \alpha = 1, \lambda = 2$ were considered. In both cases, considered. The empirical results is presented in table 1 and table 2 which show that for both cases ML estimate are consistant as the root mean square error (RMSE) and the average bias (BIAS) decreases as the sample size increases. i.e the ML estimates converge to the true value of the parameter of IDD.

Table 1: MEANS, BIAS and RMSE of $\alpha = 4.0, \theta = 0.5$ and $\lambda = 0.1$

n	α			θ			λ		
	MEANS	BIAS	RMSE	MEANS	BIAS	RMSE	MEANS	BIAS	RMSE
20	4.5416	0.5416	1.5253	0.6184	0.1184	0.4057	0.1054	0.0054	0.1250
50	4.2997	0.2997	1.0773	0.5442	0.0442	0.2325	0.1004	0.0004	0.0549
100	4.1918	0.1918	0.7441	0.5125	0.1497	0.1497	0.0999	-0.0001	0.0375
250	4.0669	0.0669	0.4564	0.5069	0.0069	0.0942	0.0995	-0.0005	0.0224
500	4.0366	0.0366	0.3143	0.5018	0.0018	0.0664	0.1002	0.0002	0.0149
1000	4.0163	0.0163	0.2159	0.5014	0.0014	0.0462	0.1004	0.0004	0.0109

3.2. Applications

In this section, we apply IDD to two-lifetime datasets. We employed several information criteria and the goodness of fit statistics that allowed us to compare the fits of the IDD with seven different

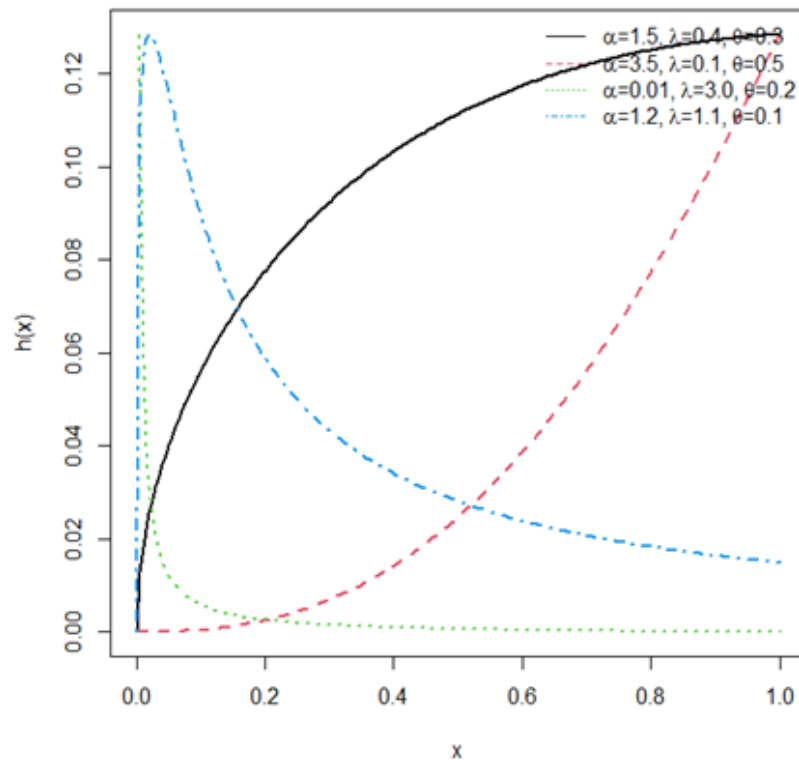
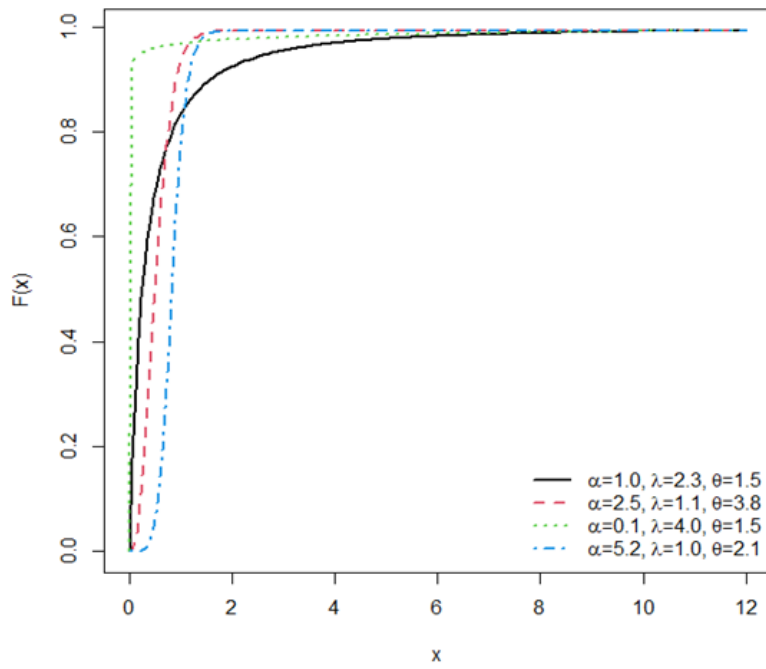


Figure 2: The IDD hazard function plot

existing distributions.

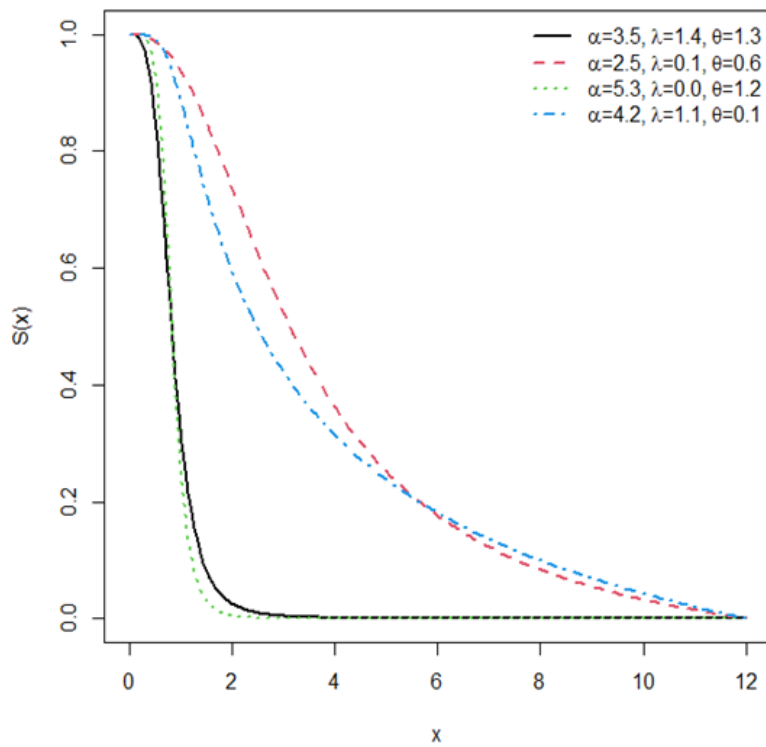
The first data set consists of the daily new fatalities caused by COVID-19 in New Jersey, USA, from March 12, 2020, to July 25, 2021, as retrieved from <https://www.worldometers.info/coronavirus/usa/new-jersey/>. The data are "1, 1, 2, 5, 2, 6, 5, 8, 19, 21, 22, 31, 37, 24, 42, 79, 100, 206, 124, 226, 81, 98, 259, 308, 222, 263, 284, 189, 106, 409, 398, 409, 365, 260, 150, 198, 425, 352, 413, 214, 279, 86, 121, 449, 372, 518, 352, 231, 162, 74, 386, 318, 297, 171, 150, 165, 88, 226, 210, 248, 231, 126, 121, 94, 161, 176, 120, 151, 111, 64, 18, 48, 162, 81, 141, 115, 84, 24, 58, 139, 114, 87, 73, 80, 87, 88, 112, 97, 56, 108, 42, 56, 63, 62, 41, 38, 29, 14, 36, 55, 52, 33, 45, 34, 27, 5, 17, 41, 52, 24, 22, 23, 50, 64, 106, 31, 50, 6, 30, 23, 43, 31, 20, 20, 5, 2, 2, 23, 33, 18, 13, 17, 16, 18, 15, 9, 10, 6, 3, 6, 3, 13, 3, 10, 11, 2, 5, 4, 15, 2, 31, 9, 7, 9, 1, 2, 2, 7, 2, 6, 8, 4, 3, 7, 5, 15, 9, 6, 5, 3, 22, 7, 11, 5, 9, 4, 4, 5, 8, 9, 4, 5, 4, 2, 1, 3, 9, 10, 3, 5, 4, 13, 7, 3, 3, 3, 4, 3, 7, 5, 7, 3, 7, 2, 1, 17, 12, 4, 4, 2, 4, 3, 18, 16, 17, 8, 12, 2, 10, 17, 15, 10, 6, 9, 1, 2, 16, 23, 16, 10, 12, 4, 10, 22, 12, 19, 25, 28, 15, 14, 38, 36, 36, 24, 32, 17, 16, 46, 59, 36, 27, 38, 11, 20, 84, 56, 62, 45, 42, 23, 15, 87, 111, 77, 59, 61, 29, 24, 86, 125, 66, 51, 50, 28, 26, 106, 139, 81, 47, 23, 19, 39, 115, 188, 82, 106, 32, 34, 38, 120, 124, 110, 99, 86, 173, 113, 34, 113, 87, 24, 16, 51, 136, 86, 109, 59, 17, 22, 132, 115, 81, 82, 72, 29, 30, 70, 109, 100, 93, 77, 24, 23, 93, 147, 79, 64, 47, 13, 13, 31, 135, 89, 62, 50, 24, 17, 104, 99, 69, 46, 46, 14, 21, 48, 128, 42, 30, 36, 16, 17, 45, 133, 46, 40, 34, 15, 22, 41, 79, 31, 27, 31, 40, 7, 61, 50, 38, 28, 24, 7, 15, 82, 75, 30, 24, 11, 9, 14, 51, 49, 34, 42, 33, 12, 25, 50, 59, 47, 43, 40, 9, 18, 45, 65, 30, 27, 39, 13, 19, 61, 49, 20, 25, 34, 12, 16, 42, 49, 33, 27, 22, 11, 10, 28, 43, 25, 26, 20, 9, 14, 23, 32, 23, 17, 14, 7, 9, 25, 34, 14, 12, 16, 6, 5, 7, 28, 6, 12, 8, 6, 5, 1, 31, 6, 2, 2, 3, 5, 11, 12, 7, 4, 4, 2, 3, 15, 18, 6, 12, 4, 3, 3, 6, 13, 5, 5, 1, 4, 13, 3, 3, 5, 4, 4, 7, 11, 2, 2, 5, 3, 6, 12, 5, 2, 4, 5, 2, 4, 6, 3, 2, 5, 7, 3, 1, 8, 11, 4, 7, 5, 5, 4, 9, 6, 7, 9".

Table 3 lists ML estimates of the parameters together with the values of information criteria; the Akaike Information Criteria (AIC), the Consistent Akaike Information Criteria, and the Bayesian



h

Figure 3: The IDD cdf Plot



h

Figure 4: The IDD Survival Function plot

Table 2: MEANS, BIAS and RMSE of $\alpha = 1.0$, $\theta = 1.5$ and $\lambda = 2.0$

n	α			θ			λ		
	MEANS	BIAS	RMSE	MEANS	BIAS	RMSE	MEANS	BIAS	RMSE
20	1.0820	0.0820	0.2427	1.8485	0.3485	0.9268	2.4032	0.4032	1.8064
50	1.0343	0.0343	0.1533	1.6827	0.1827	0.6381	2.3160	0.3160	1.5318
100	1.0189	0.0189	0.1066	1.5963	0.0963	0.4689	2.1882	0.1882	1.0169
250	1.0073	0.0073	0.0746	1.5592	0.0592	0.3407	2.0970	0.0970	0.7676
500	1.0036	0.0036	0.0553	1.5287	0.0287	0.2552	2.0712	0.0712	0.5735
1000	1.0015	0.0015	0.0397	1.5168	0.0168	0.1811	2.0389	0.0389	0.4127

Information Criteria. It can be observed that the IDD provides a better fit in comparison with seven competitors; the Weibull Weibull Distribution (WW), the Beta Exponentiated Exponential (BEE), the Beta Gamma Distribution (BGA), the Weibull Exponential (WE), the Inverse Weibull (IW), the Inverse Rayleigh (IR) and the Inverse Exponential (IE). Because the IDD has a lower value for all information criteria. Hence, we can conclude that the IDD gives a better fit to the COVID-19 dataset than the other probability distributions considered in the study. Table 4 displays the goodness of fit statistics; the Anderson Darling (A^*), the Cramer-Von Mises (W^*), and Kolmogorov-Smirnov (KS) statistics with its p-value. The proposed distribution appears to be a very competitive model for these data since the values of the investigated metrics are lower and the probability value of Kolmogorov-Smirnov statistics is larger than those of the other models.

Table 3: the model parameters MLE estimates and information criteria for the dataset two.

Model	α	θ	λ	β	AIC	CAIC	BIC
IDD	-	1.097	1.393	0.0213	4935.08	4935.12	4947.76
WW	-	3.7360	0.1490	0.3920	4980.01	4980.05	4992.69
BEE	0.2510	0.4350	0.0317	3.1370	4968.67	4968.75	4985.59
BGA	0.0602	3.5560	8.5420	0.0220	50.94.17	5094.25	5111.09
WE	-	-	0.5280	0.0090	5111.31	5119.76	5114.62
IW	-	-	6.0356	0.7537	4964.96	4964.99	4973.42
IR	-	-	24.1579	-	6463.86	6463.87	6468.08
IE	-	-	8.9379	-	5048.32	5048.33	5052.55

Table 4: The test results for the Goodness-of-fit for dataset two.

Model	A^*	W^*	KS	KS p-value
IDD	2.404	0.333	0.056	0.082
WW	5.835	0.858	0.081	0.003
BEE	4.811	0.677	0.084	0.002
BGA	11.688	1.761	0.177	<0.0001
WE	13.177	2.017	0.183	<0.0001
IW	5.771	0.936	0.082	0.002
IR	16.032	2.661	0.514	<0.0001
IE	8.041	1.316	0.193	0.081

The second dataset comprises 100 observations of the breaking stress of carbon fibers provided by [21]. the data are "0.920, 0.9280, 0.997, 0.9971, 1.0610, 1.117, 1.1620, 1.183, 1.187, 1.1920, 1.196, 1.2130, 1.215, 1.2199, 1.220, 1.2240, 1.225, 1.2280, 1.237, 1.240, 1.244, 1.259, 1.2610, 1.263, 1.276, 1.310, 1.3210, 1.3290, 1.3310, 1.337, 1.351, 1.359, 1.388, 1.4080, 1.449, 1.4497, 1.450, 1.459, 1.471, 1.475, 1.477, 1.480, 1.489, 1.501, 1.507, 1.515, 1.530, 1.5304, 1.533, 1.544, 1.5443, 1.552, 1.556, 1.5620, 1.566, 1.585, 1.586, 1.599, 1.602, 1.6140, 1.6160, 1.617, 1.6280, 1.6840, 1.7110, 1.7180, 1.733, 1.7380, 1.7430, 1.7590, 1.777, 1.7940, 1.799, 1.806, 1.814, 1.814, 1.8160, 1.8280, 1.830, 1.884, 1.892, 1.944,

1.972, 1.9840, 1.987, 2.02, 2.0304, 2.0290, 2.0350, 2.0370, 2.0430, 2.0460, 2.0590, 2.111, 2.165, 2.686, 2.778, 2.972, 3.504, 3.863, 5.3060".

Table 5 shows the statistics of AIC, CAIC, and BIC for all models investigated. When compared to the values of the other models, the IDD values are lower, demonstrating that this new distribution is a very competitive model for this data. Table 6 reveals that the IDD has the highest KS p-value and the lowest KS, A^* , and W^* values for the carbon fibers dataset. This demonstrates that the IDD distribution is superior at fitting this dataset.

Table 5: the model parameters MLE estimates and information criteria for the dataset two.

Model	α	θ	λ	β	AIC	CAIC	BIC
IDD	-	9.9065	0.5073	0.0577	111.88	112.13	119.72
WW	-	5.557	0.332	0.564	196.20	196.45	204.04
BEE	14.951	0.8590	4.0340	10.8910	119.40	119.81	129.86
BGA	10.6136	0.4324	5.6310	7.5055	122.65	123.07	133.11
WE	-	-	1.4008	0.3543	216.18	216.30	221.41
IW	-	-	4.3427	4.3754	112.54	112.66	117.77
IR	-	-	2.2095	-	183.67	183.71	186.29
IE	-	-	1.5317	-	304.42	304.03	307.03

Table 6: The test results for the Goodness-of-fit for dataset two.

Model	A^*	W^*	KS	KS p-value
IDD	0.497	0.067	0.070	0.704
WW	6.323	1.031	0.206	0.003
BEE	0.854	0.107	0.086	0.4444
BGA	1.022	0.123	0.097	0.298
WE	7.943	1.348	0.244	<0.0001
IW	0.805	0.116	0.089	0.399
IR	0.713	0.087	0.327	<0.0001
IE	0.986	0.121	0.447	<0.0001

4. CONCLUSION

In this work, we presented a new univariate continuous distribution called the inverted Dagum distribution by taking the transformation of the reciprocal of the random variable of the Dagum distribution on the ground of the type I Dagum distribution. After introducing the distribution, we obtained its basic properties like r^{th} moments, mean, moment generating function, quantile function, hazard function, survival function, Renyi entropy, and order statistics. The maximum likelihood method was used to provide the estimates of the model parameters. We also provided the density hazard rate plots of the distribution with some assumed values. Additionally, a simulation study was performed to show the ML estimates are consistence as the number as sample size increases the estimates converge to the actual values of the parameters. Our empirical analysis of two real datasets shows that the proposed distribution outperforms the well-known distribution based on goodness of fit metrics and information criteria.

REFERENCES

- [1] Dagum, C. (1977). A new model for personal income distribution: Specification and estimation. *Economic Applique*, 30(3): 413–437.
- [2] McDonald, B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3): 647–663.

- [3] McDonald, B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3): 647–663.
- [4] Domma, F., Condino, F. (2013). The Beta-Dagum distribution: definition and Properties. *Communications in Statistics Theory and Methods*, 42(22): 4070–4090.
- [5] Domma, F., Giordano, S. Zenga, M. M. (2011). Maximum likelihood estimation in Dagum distribution with censored sample. *Journal of Applied Statistics*, 38(12): 2971–2985.
- [6] Kleiber, C., Kotz, S. (2003). Statistical size distributions in economics and actuarial sciences. 470. *John Wiley Sons*.
- [7] Kleiber, C. (2008). A guide to the Dagum distributions. In *Modeling Income Distributions and Lorenz Curves*, , 97-117. Springer New York.
- [8] Shahzad, M. N., Asghar, Z. (2013). Comparing TL-Moments, L-Moments and Conventional Moments of Dagum Distribution by Simulated data. *Revista Colombiana de Estadística*, 36(1): 79–93.
- [9] Dey, S., Al-Zahrani, B., Basloom, S. (2017). Dagum distribution: properties and different method of estimation. *International journal of Statistics and probability*, 6(2): 74–92.
- [10] Oluyede, B.O., Rajasooriya, S. (2013). The MC-Dagum Distribution and its Statistical properties with applications. *Asian journal of Mathematics and applications*, 1(1): 1–16.
- [11] Rasheed, N. (2020). A new generalized-G class of distributions and its applications with Dagum distribution. *Research journal of Mathematics and Statistical Sciences*, 8(3): 1–13.
- [12] Abid, S.H., Al-Hassany, S. (2016). On the inverted Gamma distribution. *International journal of systems science and Applied Mathematics*, 1(3): 16–22.
- [13] Nanga, S., Nasiru, S. Diogban, J. (2023). Cosine Topp-Leone family of distributions: properties and regression. *Research in Mathematics*, 10(1): 1–17.
- [14] Souza, L., Rosa, W.O, Brito, C.R. (2019). On the Sin-G class of distributions: theory, model and application. *Journal of Mathematical Modelling*, 7(3): 357–379.
- [15] Al-Babtain, A.A., Elbatal, I. Chesneau, C. Elgarhy, M. (2020). Sine Topp-Leone-G family of distributions: theory and applications. *Open Physics*, 18: 574–593.
- [16] Muhammad, M., Alshanbari, H.M., Alanzi, R.A, Liu, L., Sami, W., Chesneau, C. Jamal, F. (2021). A new generator of probability models: the exponentiated Sine-G family for lifetime studies. *Entropy*, 23: 1–30.
- [17] Gomes-Silva, F., da Silva, R.V., Percontini, A., Ramos, M.W. Cordeiro, G.M. (2017). An extended Dagum distribution: properties and applications. *International journal of applied Mathematics and Statistics*, 56(1): 35–53.
- [18] Renyi, A. (1961). On measures of entropy and information. *Proceedings of the 4th Berkely Symposium on Mathematical Statistics and probability*, University of Callifonia Press, Berkely.
- [19] Abouammoh, A.M. Alshingiti, A.M. (2009). Reliability estimation of generalized inverted exponential distribution. *Journal of statistical computation and simulation*, 79(11): 1301–1315
- [20] Flaih, A., Elsallaoukh, H., Mendi, M.M. (2012). The exponentiated inverted Weibull distribution. *Appl. Math. Inf. Sci*, 6(2): 167–171
- [21] Nichols, Michele D and Padgett, WJ (2006). A bootstrap control chart for Weibull percentiles, *Quality and reliability engineering international*,

A NOVEL HYBRID DISTRIBUTED INNOVATION EGARCH MODEL FOR INVESTIGATING THE VOLATILITY OF THE STOCK MARKET

¹Mubarak M.T., ²Adubisi O.D., and ³Abbas U.F.

^{1,3}School of Science and Technology, Gombe State Polytechnic, Bajoga, Gombe State, Nigeria.

²Department of Mathematics and Statistics, Federal University Wukari, Nigeria.

mubarak@gspb.edu.ng¹, adubisiobinna@fuwukari.edu.ng², ufabbas@gspb.edu.ng³

Abstract

When calculating risk and making decisions, investors and financial institutions heavily rely on the modeling of asset return volatility. For the exponentiated generalized autoregressive conditional heteroscedasticity (EGARCH) model, we created a unique innovation distribution in this study called the type-II-Topp-Leone-exponentiated-Gumbel (TIIT_{LEGU}) distribution. The key mathematical characteristics of the distribution were determined, and Monte Carlo experiments were used to estimate the parameters of the novel distribution using maximum likelihood estimation (MLE) procedure. The performance of the EGARCH (1,1) model with TIIT_{LEGU} distributed innovation density in relation to other innovation densities in terms of volatility modeling is examined through applications using two Nigerian shock returns. The results of the diagnostic tests indicated that, with the exception of the EGARCH (1,1)-Johnson (SU) reparametrized (JSU) innovation density, the fitted models have been sufficiently specified. The parameters for the EGARCH (1,1) model with different innovation densities are significant at various levels. Furthermore, in out-of-sample prediction, the fitted EGARCH (1,1)-TIIT_{LEGU} innovation density performed better than the EGARCH (1,1)-existing innovation densities. As a result, it is decided that the EGARCH-TIIT_{LEGU} model is the most effective for analyzing Nigerian stock market volatility.

Keywords: EGARCH, Innovations density, Maximum likelihood estimation, Simulation.

I. Introduction

The Nigerian stock market has grown in terms of the number of stock exchanges and other financial intermediaries, the number of listed stocks, trading volumes, market capitalization, investor population, stock exchange turnover, and stock price indexes over time. The stock market's performance is a key measure of a country's progress and development. Because it reflects the potential viability and financial strength of corporate entities registered on the stock exchange, the stock market is one of the yardsticks for assessing an economy's growth and development. However, the state of investor confidence in different economic sectors is reflected in the stock market. It shows hopes for the stability of the financial system and reflects the strength of the producing sector [1-2]. A survey of pertinent literature reveals that while predicting stock market volatility, academics have neglected to account for the contributions of alternate innovation distributions. Since financial time series have leptokurtic and autocorrelation characteristics, mis-specification may result from applying the incorrect innovation density in the EGARCH volatility model. Additionally, an

erroneous innovation density specification can result in a significant loss of efficiency for the relevant estimators, an invalid risk assessment, incorrectly priced options, and an improper valuation of Value-at-Risk (VaR).

In making financial decisions such as portfolio selection, risk management, and option pricing, financial return volatility is an important metric to consider. Therefore, in order to model and anticipate the volatility of asset returns, it is imperative to create a model with strongly driven conditional innovation density. According to [3], a suitable volatility model is one that both accurately represents the disturbance term's heteroscedasticity and reflects the stylized facts that are present in stock return series. Financial institutions usually employ generalized autoregressive conditional heteroskedasticity (GARCH) models to predict return volatility for stocks, bonds, and market indexes, while they can also be used to analyze other types of financial data, such as macroeconomic data. In order to aid in their decisions about asset allocation, hedging, risk management, and portfolio optimization, they use the information that is produced to estimate the returns of current investments, help decide pricing, and assess which assets may yield larger returns [4]. However, these returns exhibit leverage effects, heavy tail, volatility clustering, significant skewness and excess leptokurtic behaviours which the symmetric GARCH model with normal distributed innovations in most cases fail to capture [5-7]. Despite the skewed form and appealing characteristics of the student-t in GARCH models, the tail behaviour is still too short to adequately describe skewness and fat tails in asset returns [8-9]. According to [10], volatility models that do not allow for conditional variance asymmetry typically result in inaccurate volatility estimations and projections. To address this flaw in GARCH's treatment of financial time series, the asymmetric exponential GARCH model was developed using the generalized error distributed innovation. Specifically, to accommodate asymmetric impacts between asset returns that are positive and negative [10-11]. Few studies had been done on returns volatility modeling utilizing the EGARCH model with common innovation densities, which raises questions about the choice of innovation densities [11-17].

Quite a few academics have focused on establishing novel distributions for volatility model innovations, and have conducted various studies in this area of altering the innovation density of the EGARCH volatility model. [18] advocated the use of the beta-student-t distribution for the EGARCH model in the estimation of volatility. [19] discovered that more flexible GARCH-type models are sufficiently acceptable in predicting volatility for all density assumptions. Using simulated and actual data, the Bayesian analysis of a stochastic model with generalized hyperbolic skew Student-t distribution. [20] proposed the EGARCH model with the beta-skew Student-t density for predicting daily volatility. [21] discovered that using leptokurtic distributions in GARCH-type models helped them produce more accurate volatility projections. [22-23] evaluated the daily volatility of stock index returns using a new generalization of the skew Student-t distribution, and demonstrated that it performed better than some innovation densities. [2] and [24] proposed the exponentiated half-logistic skew-t and generalized odd generalized skew-t densities for evaluating the daily volatility of bitcoin, Nigeria inflation and first bank Nigeria stock returns, and found that it performed better than other innovation densities in the GARCH-type models. [7] proposed the GARCH model with the exponentiated Gumbel density for predicting daily return volatility of the S&P 500 index. [9] proposed the odd generalized exponential Laplace density for evaluating the daily volatility of the Nigeria stock exchange, and found that it performed better than other innovation densities in the GARCH-type models. These models informed entrepreneur and investors on volatile nature of stock prices and bitcoin rates. However, developing robust distributions remains critical in improving the accuracy of the monetary risk system. Therefore, this research set out to propose a new innovation density for the asymmetric EGARCH model by introducing a novel distribution. This research is focused on the distinctiveness of the structural properties of the novel distribution and modification of the distributional assumption of the

innovations of the EGARCH volatility model.

The article is arranged as follows: Section 2 presents the developed novel distribution and its standardized form. Section 3 presents the properties of the novel distribution. Section 4 presents the maximum likelihood estimation procedure and Monte-Carlo simulation process. Section 5 presents the EGARCH model with the novel innovation density function including methods for selecting models and appraising predictions. Section 6 reports the empirical results of both estimation and forecast assessment, and conclusion in Section 7.

2. Distribution Genesis

The cumulative distribution function (cdf) of the Type II Topp Leone (TIITL-G) family of distributions developed by [25] is specified as

$$F(x) = 1 - [1 - G^2(x)]^\theta \quad (1)$$

and the corresponding probability density function (pdf) is

$$f(x) = 2\theta g(x)G(x)[1 - G^2(x)]^{\theta-1} \quad (2)$$

where $\theta > 0$ is the shape parameter, $G(x)$ and $g(x)$ are the baseline cdf and pdf, respectively. The cdf of the baseline distribution titled the exponentiated Gumbel (EGU) distribution introduced by [26] is specified as

$$G(x) = 1 - \left\{ 1 - e^{\left[-e^{\left(-\frac{x-\mu}{\sigma}\right)}\right]} \right\}^\alpha, \quad (3)$$

and the pdf is given as

$$g(x) = \frac{\alpha}{\sigma} \left\{ 1 - e^{\left[-e^{\left(-\frac{x-\mu}{\sigma}\right)}\right]} \right\}^{\alpha-1} e^{\left[-e^{\left(-\frac{x-\mu}{\sigma}\right)}\right]} e^{\left(-\frac{x-\mu}{\sigma}\right)}, \quad (4)$$

where $\alpha > 0$, $\sigma > 0$, $\mu \in \mathcal{R}$ are shape, scale and location parameters, and $x \in \mathcal{R}$.

The cdf and pdf of the type II Topp Leone exponentiated Gumbel (TIITL_{EGU}) model derived by inserting Equations (3) and (4) into Equations (1) and (2), respectively, are specified as

$$F(x; \alpha, \theta, \mu, \sigma) = 1 - \{1 - [\Theta(x)]^2\}^\theta, \quad (5)$$

$$f(x; \alpha, \theta, \mu, \sigma) = 2\theta\Phi(x)\{1 - [\Theta(x)]^2\}^{\theta-1}, \quad (6)$$

Hither, the pdf and cdf of the EGU are represented with $\Phi(x)$ and $\Theta(x)$. $\alpha > 0$, $\theta > 0$, $\sigma > 0$ are the shape and scale parameters, and $\mu \in \mathcal{R}$ is the location parameter. More so, the survival and hazard rate functions are specified as follows

$$S(x; \alpha, \theta, \mu, \sigma) = \{1 - [\Theta(x)]^2\}^\theta \quad (7)$$

$$h(x; \alpha, \theta, \mu, \sigma) = 2\theta\Phi(x)\Theta(x)\{1 - [\Theta(x)]^2\}^{-1}, \quad (8)$$

Figures 1 depicts that the TIITL_{EGU} density function can be very useful in describing symmetric, heavy-tailed, unimodality, leptokurtic and skew patterns of most data sets. Hence, a viable alternative innovation density for increasing the accuracy of the EGARCH volatility model prediction.

2.1 Standardized TIITL_{EGU} Model

The standardized TIITL_{EGU} model is obtained via the transformation $\varepsilon_t = z\sqrt{h_t^2}$, where $E(z_t) = 0$ and $var(z_t) = 1$. The random variable z_t can be expressed as $z_t = \frac{(x_t - \mu)}{\sqrt{h_t^2}} = \frac{\varepsilon_t}{\sqrt{h_t^2}}$, then $\frac{dz_t}{d\varepsilon_t} = \frac{1}{\sqrt{h_t^2}}$. Therefore,

the standardized TIITL_{EGU} density function is

$$f(\varepsilon_t; \alpha, \theta, h_t) = \frac{2\alpha\theta}{(h_t^2)^{\frac{1}{2}}} \eta^{\alpha-1} (1 - \eta) e^{\left(-\frac{\varepsilon_t}{h_t^2}\right)} (1 - \eta^\alpha) [1 - (1 - \eta^\alpha)^2]^{\theta-1} \frac{1}{(h_t^2)^{\frac{1}{2}}} \quad (9)$$

where $\eta = 1 - e^{-e^{\left(-\frac{\varepsilon_t}{h_t^2}\right)}}$, μ and $\sqrt{h_t^2}$ denote the mean and standard deviation. The standardized

TIITL_{EGU} is the novel hybrid distributed innovation function for the EGARCH volatility model.

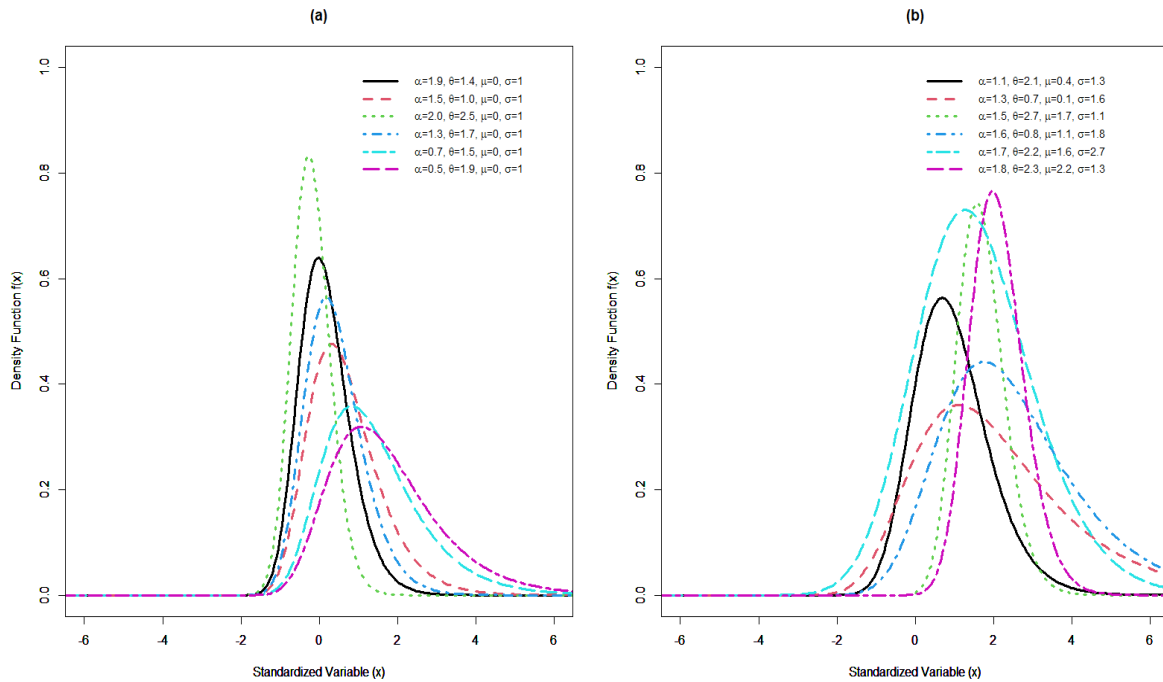


Figure 1: The TIITL_{EGU} density function (pdf) plots with selected parameter values.

3. Properties of the TIITL_{EGU} Model

3.1 Quantile and Median Functions

The quantile function (qf) of the TIITL_{EGU} model derived by inverting Eq. (5) is specified as

$$Q(u) = -\sigma \log \left[-\log \left(1 - \left\{ 1 - \left[1 - (1-u)^{\frac{1}{\theta}} \right]^{\frac{1}{2}} \right\}^{\frac{1}{\alpha}} \right) \right] + \mu \tag{10}$$

where $u \sim \text{Uniform}(0,1)$. The median (M) of the TIITL_{EGU} model is specified as

$$M = -\sigma \log \left[-\log \left(1 - \left\{ 1 - \left[1 - (0.5)^{\frac{1}{\theta}} \right]^{\frac{1}{2}} \right\}^{\frac{1}{\alpha}} \right) \right] + \mu \tag{11}$$

By means of the quantile function in Eq (10), various quantile measures can be estimated.

3.2 Raw-Moment and Moment-Generating-Function

The raw-moment (rm) of the TIITL_{EGU} model is derived using the expansion series approach. Let $X \sim \text{TIITL}_{EGU}(x; \alpha, \theta, \mu, \sigma)$, the rth raw-moment of X is specified as

$$\mu'_r = \int_0^{+\infty} x^r f(x; \alpha, \theta, \mu, \sigma) dx \tag{12}$$

The expanded form of the TIITL_{EGU} pdf using the series expansion is

$$f(x) = \sum_{m=0}^{\infty} \omega_m e^{-(m+1)\frac{x-\mu}{\sigma}} \tag{13}$$

where $\omega_m = \frac{2\alpha\theta}{\sigma} \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j+k+m} \binom{\theta-1}{i} \binom{2i-1}{j} \binom{\alpha(j+1)-1}{k}}{m!}$

By inserting the expanded form in Eq. (13) into Eq. (12), we have

$$\mu'_r = \sum_{m=0}^{\infty} \omega_m \int_0^{+\infty} x^r e^{-(m+1)\frac{x-\mu}{\sigma}} dx, \tag{14}$$

Let,

$$z = (m + 1)^{\frac{x-\mu}{\sigma}} \Rightarrow x = \frac{z}{\frac{\mu}{\sigma(m+1)}}$$

$$\frac{dx}{dz} = \frac{1}{-\frac{\mu}{\sigma(m+1)}} \Rightarrow dx = \frac{1}{-\frac{\mu}{\sigma(m+1)}} dz,$$

Hence, the rm of the TIITL_{EGU} is specified as

$$\mu_r = \sum_{m=0}^{\infty} \omega_m \int_0^{+\infty} \left(\frac{z}{\frac{\mu}{\sigma(m+1)}}\right)^r e^{-z} \frac{1}{-\frac{\mu}{\sigma(m+1)}} dz, \tag{15}$$

Simplifying Eq. (15) leads to

$$\mu_r = \sum_{m=0}^{\infty} \omega_m \left(-\frac{\mu}{\sigma}(m+1)\right)^{-r-1} \int_0^{+\infty} (z)^r e^{-z} dz, \tag{16}$$

Using the gamma integral representation $\Gamma(\alpha + 1) = \int_0^{+\infty} (y)^\alpha e^{-y} dy$. Therefore, the rm of the TIITL_{EGU} is specified as

$$\mu_r = \sum_{m=0}^{\infty} \omega_m \frac{\Gamma(r+1)}{\left(\frac{\mu}{\sigma}(m+1)\right)^{r+1}} \tag{17}$$

where $\omega_m = \frac{2\alpha\theta}{\sigma} \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j+k+m} (k+1)^m \binom{\theta-1}{i} \binom{2i-1}{j} \binom{\alpha(j+1)-1}{k}}{m!}$.

Table 1 reports the first four raw-moments, standard-deviation (SD), dispersion index (DI) and coefficient of variation (CV) for the TIITL_{EGU} model utilizing Eq. (17) with fixed $\mu = 0, \sigma = 1$.

Table 1: Summary statistics for the TIITL_{EGU} model

μ_r	$\alpha = 1.5, \theta = 1.5$	$\alpha = 2.5, \theta = 2.0$	$\alpha = 3.3, \theta = 3.5$	$\alpha = 3.7, \theta = 3.7$
μ_1	0.4810	0.0969	0.0057	0.0022
μ_2	0.5850	0.0577	0.0016	0.0006
μ_3	0.9590	0.0472	0.0006	0.0002
μ_4	1.9600	0.0483	0.0003	0.0000
SD	0.5950	0.2200	0.0396	0.0244
DI	0.7350	0.4990	0.2750	0.2710
CV	1.2400	2.2700	6.9500	11.1000

Additionally, the moment generating function (mgf) of the TIITL_{EGU} is specified as

$$M_r(t) = \sum_{r=0}^{\infty} \sum_{m=0}^{\infty} \omega_m \frac{\Gamma(r+1)t^r}{\left(-\frac{\mu}{\sigma}(m+1)\right)^{r+1}} \tag{18}$$

where $\omega_m = \frac{2\alpha\theta}{\sigma} \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j+k+m} (k+1)^m \binom{\theta-1}{i} \binom{2i-1}{j} \binom{\alpha(j+1)-1}{k}}{m!}$.

3.3 Order Statistics

Let x_1, x_2, \dots, x_n denote a random sample from the TIITL_{EGU} and $x_{1:n} < x_{2:n} < \dots < x_{n:n}$ be the order statistics obtained from the sample. The pth order statistics (os) is specified as

$$f_{p:n}(x) = \frac{f(x)}{B(p,n-p+1)} [F(x)]^{p-1} [1 - F(x)]^{n-p} \tag{19}$$

By inserting Eqs. (5) and (6) into Eq. (19), we have

$$f_{p:n}(x) = \frac{2\theta\Phi(x)\{1-[\Theta(x)]^2\}^{\theta-1}}{B(p,n-p+1)} [1 - \{1 - [\Theta(x)]^2\}^\theta]^{p-1} [1 - (1 - \{1 - [\Theta(x)]^2\}^\theta)]^{n-p} \tag{20}$$

Simplifying Eq. (20), the os of the TIITL_{EGU} is specified as

$$f_{p:n}(x) = \frac{2\theta\Phi(x)\{1-[\Theta(x)]^2\}^{\theta(n-p+1)-1}}{B(p,n-p+1)} [1 - \{1 - [\Theta(x)]^2\}^\theta]^{p-1} \tag{21}$$

The minimum and maximum order statistics is derived by inserting $p = 1$ and $p = n$ into Eq. (21).

4. Estimation and Simulation Study

4.1 Maximum Likelihood Estimation

Let x_1, x_2, \dots, x_n denote the observed random-values from the TIITL_{EGU}($x; \alpha, \theta, \mu, \sigma$). Assuming that

$\mu = 0, \sigma = 1$, without loss of generality, the log-likelihood-function (LL) is specified as

$$LL(\theta, \alpha) = n \log 2 + n \log \theta + n \log \alpha + (\alpha - 1) \sum_{i=1}^n \log\{\eta_i\} + \sum_{i=1}^n \log\{1 - \eta_i\} + \sum_{i=1}^n x_i + \sum_{i=1}^n \log(1 - \eta_i^\alpha) + (\theta - 1) \sum_{i=1}^n \log[1 - (1 - \eta_i^\alpha)^2] \tag{22}$$

where $\eta_i = 1 - e^{[-e^{(-x)}]}$. Differentiating Eq. (22) with respect to θ and α gives the follows:

$$\frac{\partial LL}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log[1 - (1 - \eta_i^\alpha)^2]$$

$$\frac{\partial LL}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \log\{\eta_i\} - \sum_{i=1}^n \frac{\eta_i^\alpha \log\{\eta_i\}}{1 - \eta_i^\alpha} + (\theta - 1) \sum_{i=1}^n \frac{2(1 - \eta_i^\alpha) \eta_i^\alpha \log\{\eta_i\}}{1 - (1 - \eta_i^\alpha)^2}$$

In this study, R-programming (optim function) is used in finding the maximum likelihood (ML) estimates $(\hat{\theta}, \hat{\alpha})$ of the TIITL_{EGU} parameters.

4.2 Simulation-Study

The simulation process for the ML is obtainable as follows: N = 10,000 samples (replicates) are generated from the TIITL_{EGU} model with sizes n' = 20,50, 150, 250, 500 and 1000 using R-programming. The precision of the ML estimates is evaluated via the mean estimates (MEs), absolute bias (Absbias), mean square errors (MSEs) and root mean square roots (RMSEs). The MLE is a suitable technique for estimating the TIITL_{EGU} parameters Based on the simulation study. The results reported in Table 2 indicate that the parameter estimates are quite stable and very close to the true parameter values for the various sample sizes. The ME tend to be closer to the true values of the parameter with minimum MSEs, and RMSEs values as the sample size increases.

Table 2: Numerical values of MEs, Absbias, MSEs and RMSEs

$\alpha = 1.5, \theta = 1.5$												$\alpha = 3.5, \theta = 3.3$											
n'	Par.	ME	Absbias	MSE	RMSE	n'	Par.	ME	Absbias	MSE	RMSE	n'	Par.	ME	Absbias	MSE	RMSE						
20	α	1.7238	0.2238	2.1842	1.4779	20	α	3.7210	0.4210	8.8988	2.9831	20	α	5.0225	0.5225	15.9168	3.9896						
	θ	3.1475	1.6475	10.9000	3.3015		θ	7.2017	3.7017	51.1596	7.1526		θ	8.4532	4.3532	69.0055	8.3070						
50	α	1.6239	0.1239	1.0649	1.0320	50	α	3.5275	0.2275	4.4887	2.1187	50	α	4.7706	0.2706	8.0638	2.8397						
	θ	2.4156	0.9156	4.5588	2.1351		θ	5.7094	2.2094	23.6443	4.8625		θ	6.8341	2.7341	35.0702	5.9220						
150	α	1.5312	0.0312	0.3634	0.6029	150	α	3.3703	0.0703	1.8493	1.3599	150	α	4.5916	0.0916	3.5820	1.8926						
	θ	1.8985	0.3985	1.3260	1.1515		θ	4.6138	1.1138	8.9932	2.9989		θ	5.5217	1.4217	13.8556	3.7223						
250	α	1.5156	0.0156	0.2146	0.4633	250	α	3.3349	0.0349	1.1736	1.0833	250	α	4.5407	0.0407	2.2462	1.4987						
	θ	1.7484	0.2484	0.7015	0.8375		θ	4.2532	0.7532	5.2725	2.2962		θ	5.0808	0.9808	8.5140	2.9179						
500	α	1.5089	0.0089	0.1105	0.3323	500	α	3.3155	0.0155	0.6437	0.8023	500	α	4.5209	0.0209	1.2639	1.1243						
	θ	1.6267	0.1267	0.3019	0.5495		θ	3.9323	0.4323	2.6383	1.6243		θ	4.6527	0.5527	4.0301	2.0075						
1000	α	1.5033	0.0033	0.0561	0.2368	1000	α	3.3028	0.0028	0.3303	0.5747	1000	α	4.5031	0.0031	0.6643	0.8150						
	θ	1.5653	0.0653	0.1336	0.3655		θ	3.7293	0.2293	1.1773	1.0850		θ	4.4001	0.3001	1.8519	1.3608						

5. EGARCH-Model

The asymmetric EGARCH model introduced by [10] is considered as a variate of the ARCH/GARCH model introduced by [3] and [27] for modeling time-varying volatility. The EGARCH differ from the symmetric GARCH variance structure given that the natural log variance is used in other for the parameters to be unrestricted and can take negative values while guaranteeing a positive conditional variance. Moreso, the EGARCH model includes the asymmetric impact of positive and negative shocks on volatility. The return of daily prices of assets is represented by r_t and the EGARCH (1,1) model is defined as:

$$\begin{aligned} r_t &= \mu + \varepsilon_t, \\ \varepsilon_t &= z_t \sqrt{h_t^2}, \quad z_t \sim i.i.d, \\ \log h_t^2 &= \lambda_0 + \beta_1 \log h_{t-1}^2 + \lambda_1 \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}^2}} + \gamma_1 \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}^2}} \right| \end{aligned} \quad (23)$$

where $\lambda_0 > 0$, $\lambda_1 > 0$, $\beta_1 > 0$ are the model parameters, z_t is the conditional innovation density with $E(z_t) = 0$ and $var(z_t) = 1$, μ_t is the conditional mean, γ_1 is the leverage parameter, $\log h_t^2$ is conditional log variance at present day t , ε_{t-1} and $\log h_{t-1}^2$ are the error and conditional log variance at preceding day $t - 1$, respectively. The commonly utilized conditional innovation densities are well described in the literature.

5.1 The Novel Conditional Innovation Density

The log-likelihood (LL) function of the standardized TIITLEGU model presented in Eq. (9), is specified as

$$\begin{aligned} LL(\vartheta) &= n \log 2 + n \log \theta + n \log \alpha - \frac{n}{2} \log h_t^2 + (\alpha - 1) \sum_{t=1}^n \log\{\eta_t\} + \sum_{t=1}^n \log\{1 - \eta_t\} + \\ &\sum_{t=1}^n \log \left[e^{\left(\frac{\varepsilon_t}{h_t^2}\right)} \right] + \sum_{t=1}^n \log(1 - \eta_t^\alpha) + (\theta - 1) \sum_{t=1}^n \log[1 - (1 - \eta_t^\alpha)^2] - \frac{n}{2} \log h_t^2 \end{aligned} \quad (24)$$

where $\vartheta = (\alpha, \theta, h_t)$, α, θ are the shape parameters and h_t is the EGARCH volatility model with

vector parameters and $\eta_t = 1 - e^{-e^{\left(\frac{\varepsilon_t}{h_t^2}\right)}}$.

5.2 Model Selection Criteria

The modified Akaike information criteria (AIC) and Bayesian information criteria (BIC) proposed by [28] are utilized in selecting the best model under the conditional innovation densities. The modified AIC and BIC criteria are given by

$$AIC = \frac{2k}{N} - \frac{2LL}{N} \quad (25)$$

$$BIC = \frac{k \log_e(N)}{N} - \frac{2LL}{N} \quad (26)$$

where k is the total number of estimated parameters, the estimated log-likelihood value and sample size are denoted by LL and N , respectively. The EGARCH model with the least AIC and BIC values is regarded as the most suitable model under the specified innovation density.

5.3 Forecasts Performance

The forecasts performance of the EGARCH models is appraised using the mean square error (MSE), root mean square root (RMSE), and mean absolute error (MAE). The performance measures for the volatility forecasts are given by

$$MSE = \frac{1}{N} \sum_{t=1}^N (\hat{h}_t - h_t)^2 \quad (27)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{h}_t - h_t)^2} \quad (28)$$

$$MSE = \frac{1}{N} \sum_{t=1}^N |\hat{h}_t - h_t| \quad (29)$$

Where \hat{h}_t and h_t represent the volatility forecast and realized volatility, and N is the sample size. The model with the least performance measures is regarded as the most appropriate for predicting the volatility of the daily log-returns.

6. Empirical Results

6.1 Data Report

To appraise the performance of the novel distributed innovation density in the EGARCH model, the United bank of Africa (UBA) and Total energies Nigeria (TEN) stock prices log-returns are utilized. The UBA dataset consists 5468 daily log-returns from 1/2/2000 to 5/1/2024 and TEN datasets consist 5510 daily returns from 2/1/2001 to 5/1/2024. The estimation process is executed using 5218 daily log-returns from 1/2/2000 to 30/12/2022 for the UBA while 5268 daily returns from 2/1/2001 to 30/12/2022 for the TEN. The forecast evaluation of the models is carried-out with 250 daily returns from 3/1/2023 to 5/1/2024 for both UBA and TEN. The summary statistics of the daily returns for the estimation processes are reported in Table 2. More so, the graphical plots of the daily returns, squared-returns and absolute-returns with their respective sample autocorrelation function (ACF) for both UBA and TEN are depicted in Figures 3 and 4.

Tables 2 reports positive skewness and high excess kurtosis, leading to large Jarque-Bera (JB) statistic ($p < 0.001$) signifying that the daily returns for the estimation process have non-normality characteristics. Figure 2 displays the density function of a normal distribution that has the same mean and standard deviation as those of the UBA and TEN return series. The plots provide a visual check of the normality assumption for the daily returns. The deviation between the solid (return series) and dashed line (Normal distribution) indicates that the daily returns are not normally distributed.

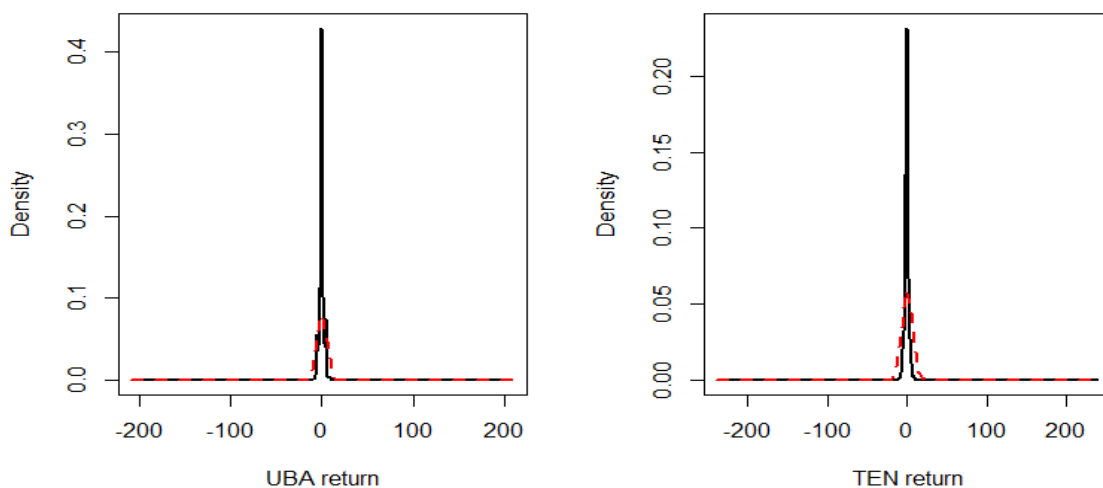


Figure 2: Empirical density function of the UBA and TEN daily returns.

Further, the ARCH Lagrange-multiplier (LM) and Ljung Box-Q tests at lag 10, indicates the incidence of conditional heteroscedasticity and autocorrelation in the returns while the Augmented Dickey-Fuller (ADF) test with its p-value indicates that the returns for the UBA and TEN are stationary.

Table 2: Summary statistics for the daily returns

UBA returns (Estimation process)					
Number of observations	Mean	Median	Minimum	Maximum	Std Dev.
5218	-0.004	0.000	-204.521	206.978	5.005
Skewness	Kurtosis	Jarque-Bera	ARCH (10)	Q (10)	ADF test
0.087	1096.432	26157 (< 0.0001)	894.89 (< 0.0001)	901.99 (< 0.0001)	-24.7 (< 0.01)

TEN returns (Estimation process)					
Number of observations	Mean	Median	Minimum	Maximum	Std Dev.
5268	0.022	0.000	-231.941	235.569	6.790
Skewness	Kurtosis	Jarque-Bera	ARCH (10)	Q (10)	ADF test
0.999	1027.144	23175 (< 0.0001)	107.55 (< 0.0001)	117.84 (< 0.0001)	-27.5 (< 0.01)

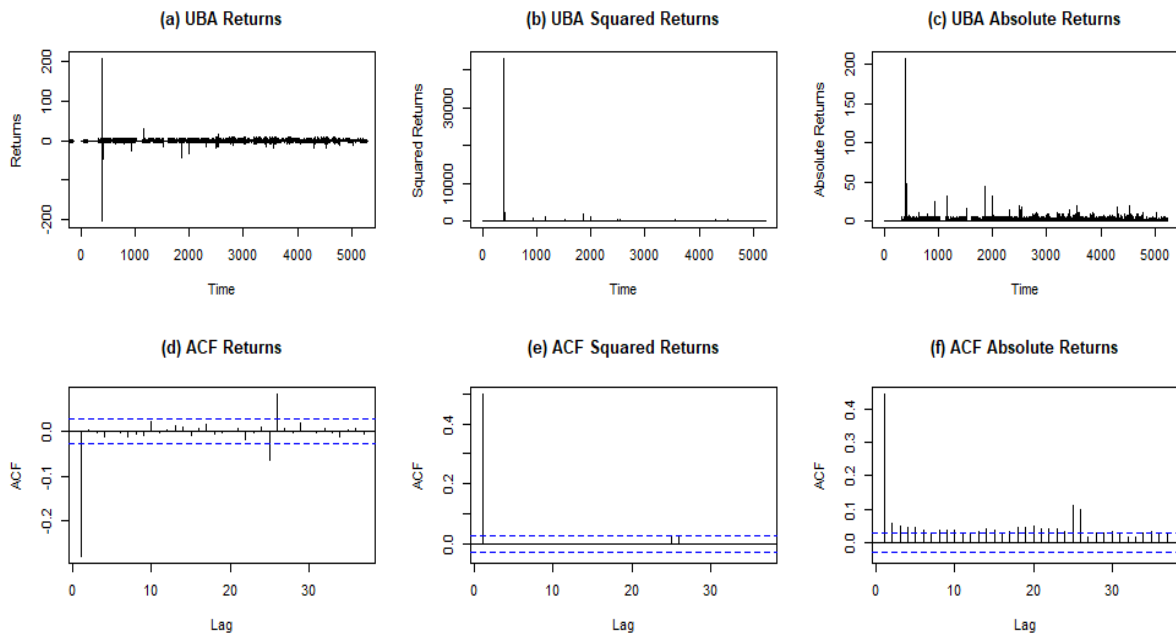


Figure 3: UBA daily returns, squared returns, absolute returns and sample autocorrelations.

6.2 Parameters Estimation of the EGARCH (1,1) Model

The EGARCH (1,1) model specified in Eq. (22) is estimated under ten different innovation densities: normal (NORM), student-t (ST), generalized error (GE), skew normal (SNORM), skew student-t (SST), skew generalized error (SGE), generalized hyperbolic (GHYP), Johnson (SU) reparametrized (JSU), Normal inverse Gaussian (NIG) and the novel type II Topp Leone exponentiated Gumbel (TIIT_{LEGU}). Tables 3 and 4 reports the estimated parameters of the EGARCH (1,1) models. The *rugarch* package in R-programming is used in estimating the parameters of the EGARCH-NORM, EGARCH-ST, EGARCH-GE, EGARCH-SNORM, EGARCH-SST, EGARCH-SGE, EGARCH-GHYP, EGARCH-JSU and EGARCH-NIG while the *Optim* function in R-programming is utilized to maximize the log-likelihood function of EGARCH-TIIT_{LEGU}. As reported in Table 5, the EGARCH-TIIT_{LEGU} model has the highest log-likelihood (LL) value and exhibits superior fit to the standardized residuals compare to others for both return series.

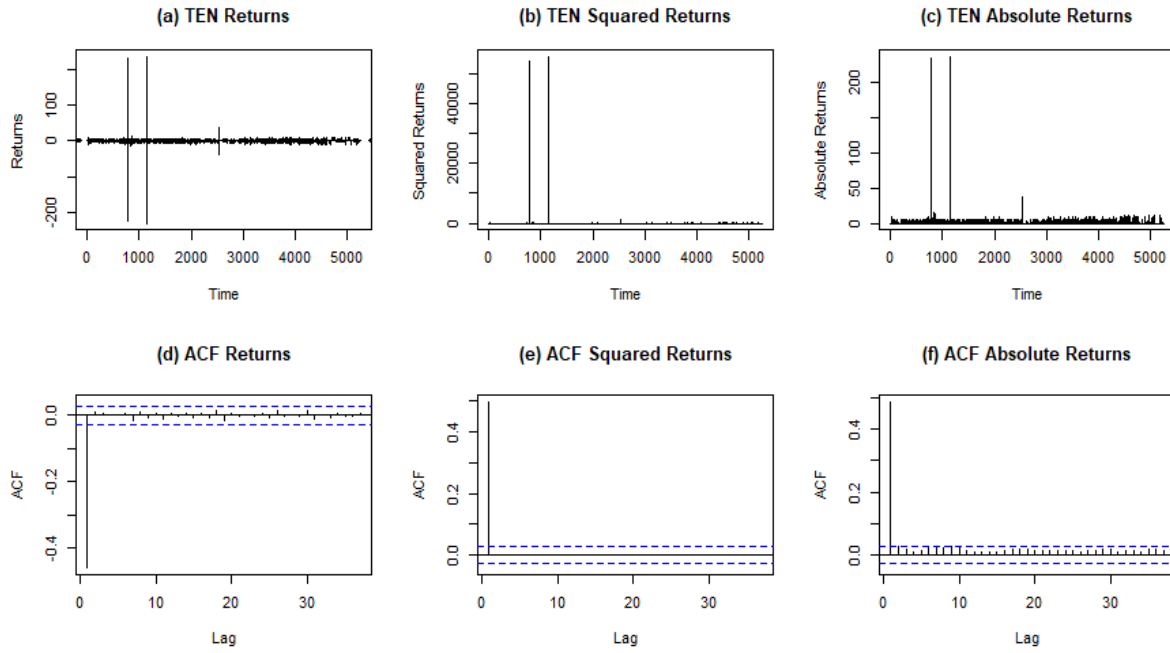


Figure 4: TEN daily returns, squared returns, absolute returns and sample autocorrelations.

Table 3: EGARCH Model Parameter Estimates with Innovation Densities (UBA returns).

Model	Cond. Distr.	λ_0	λ_1	β_1	γ_1	Ghlambda	Skew	Shape1	Shape2
EGARCH (1,1)	NORM	0.5707 ^{****}	-0.0581 ^{****}	0.7492 ^{****}	0.4881 ^{****}	-	-	-	-
	STD	0.4320 ^{****}	0.0960 ^{''}	0.888 ^{****}	2.1239 ^{****}	-	-	2.1000 ^{****}	-
	GED	8.5262 ^{****}	1.2433	0.1171	6.5026 ^{****}	-	-	0.1019 ^{****}	-
	SNORM	0.5680 ^{****}	-0.0488 ^{****}	0.7488 ^{****}	0.4867 ^{****}	-	0.9781 ^{****}	-	-
	SSTD	0.6725 ^{****}	0.2855 ^{''}	0.8914 ^{****}	6.3057 ^{****}	-	1.0003 ^{****}	2.0100 ^{****}	-
	SGED	1.6151 ^{****}	0.7086 ^{****}	0.3730 ^{****}	3.6658 ^{****}	-	1.0005 ^{****}	0.3956 ^{****}	-
	GHYP	0.0240 ^{****}	0.0524 ^{****}	0.9858 ^{****}	0.2100 ^{****}	-0.9677 ^{****}	-0.0547	0.2500 ^{****}	-
	JSU	0.2856 ^{''}	0.0571 ^{****}	0.8828 ^{****}	1.1730 ^{****}	-	0.0001	0.8269 ^{****}	-
	NIG	0.7949 ^{****}	0.3175 ^{****}	0.9532 ^{****}	4.3156 ^{****}	-	0.0172	0.0100 ^{****}	-
TIITLEGU	2.327e-10 ^{''}	3.929e-10 ^{''}	0.9828 ^{****}	0.7799 ^{****}	-	-	1.9433 ^{****}	3.5128 ^{****}	

Significance levels: 0^{****}, 0.001^{***}, 0.01^{**}, 0.05^{*}, 0.1^{''}, 1

Table 4: EGARCH Model Parameter Estimates with Innovation Densities (TEN returns).

Model	Cond. Distr.	λ_0	λ_1	β_1	γ_1	Ghlambda	Skew	Shape1	Shape2
EGARCH (1,1)	NORM	0.4283 ^{****}	-0.0869 ^{****}	0.7781 ^{****}	0.2932 ^{****}	-	-	-	-
	STD	-0.0587	0.2583 ^{****}	0.8234 ^{****}	0.6745 ^{****}	-	-	2.1000 ^{****}	-
	GED	1.3745 ^{****}	0.0974	-0.1137 ^{****}	1.1712 ^{****}	-	-	0.1024 ^{****}	-
	SNORM	0.4296 ^{****}	-0.0858 ^{****}	0.7765 ^{****}	0.3024 ^{****}	-	0.9552 ^{****}	-	-
	SSTD	0.2775 ^{****}	0.7821 ^{****}	0.8283 ^{****}	1.8532 ^{****}	-	1.0011 ^{****}	2.0100 ^{****}	-
	SGED	0.6894 ^{****}	0.0739 ^{****}	0.2451 ^{****}	0.0659 ^{****}	-	1.1175 ^{****}	0.3868 ^{****}	-
	GHYP	-0.0226 ^{****}	-0.0294 ^{****}	0.9880 ^{****}	0.0905 ^{****}	-0.2818 ^{****}	0.0025	0.2500 ^{****}	-
	JSU	0.4089 ^{****}	-0.0000	0.3658 ^{****}	-0.0000	-	0.2953 ^{****}	0.2616 ^{****}	-
	NIG	-0.0434 ^{****}	0.0071 ^{****}	1.0000 ^{****}	0.0177 ^{****}	-	-0.0071	0.0100 ^{****}	-
TIITLEGU	0.0552 ^{****}	0.0579 ^{****}	0.5195 ^{****}	1.5234 ^{****}	-	-	5.2086 ^{****}	9.8771 ^{****}	

Significance levels: 0^{****}, 0.001^{***}, 0.01^{**}, 0.05^{*}, 0.1^{''}, 1

Tables 3 and 4 reports that the parameter estimates of the EGARCH conditional variance specifications are highly statistically significant and γ_1 is highly significant which shows that the

daily log-returns have leverage effect. Therefore, the impact of the shocks is asymmetric in nature that is, the impact of positive shocks on volatility are higher than negative shocks of the similar size. The AIC and BIC values also reported in Table 5 suggest that the EGARCH-TIITL_{LEGU} model is best for investigating the volatility of the Nigerian stock market.

Table 5: Comparison of the Innovation Densities for estimated models.

Model	Innovation Dist.	(UBA Returns).			(TEN Returns).		
		Log-Likelihood	AIC	BIC	Log-Likelihood	AIC	BIC
EGARCH (1,1)	NORM	-12581.400	4.8242	4.8305	-11134.030	4.4296	4.461
	STD	-11070.810	4.2456	4.2532	-8291.421	3.3071	3.3149
	GED	-5747.029	2.2051	2.2126	19540.740	-7.7859	-7.7781
	SNORM	-12578.510	4.8235	4.8310	-11124.140	4.4361	4.4439
	SSTD	-11046.030	4.2365	4.2453	-8173.160	3.2603	3.2694
	SGED	-10738.730	4.1187	4.1275	-4397.584	1.7555	1.7646
	GHYP	-10923.350	4.1899	4.1999	-8299.449	3.3111	3.3215
	JSU	-11066.350	4.2443	4.2531	9042.654	-3.6013	-3.5922
	NIG	-10092	3.8709	3.8797	-4979.278	1.9874	1.9965
	TIITL_{LEGU}	13546.100	-5.1894	-5.1806	25184.930	-9.5588	-9.5501

Note(s): Bolded values indicate the highest log-likelihood value, and the least AIC and BIC values

Tables 6 and 7 reports the diagnostic tests results for the EGARCH (1,1) model under the various innovation densities. From Table 6, the Ljung Box-Q statistic ($p > 0.05$) specifies that the squared standardized residuals from the EGARCH-TIITL_{LEGU} model exhibit no sign of serial-correlation. Likewise, the ARCH-LM statistic ($p > 0.05$) indicates that the standardized residuals from the EGARCH-TIITL_{LEGU} model exhibit no additional conditional heteroscedasticity, that is, the conditional variance equation are correctly specified. Therefore, the results disclose that standardized TIITL_{LEGU} density is an improved distributed innovation function for the EGARCH (1,1) model.

Table 6: Estimated EGARCH (1,1) models diagnostic tests (UBA returns).

Model	Innovation Dist.	Ljung-Box Q-Statistic	p-value	ARCH-LM Statistic	p-value
EGARCH (1,1)	NORM	0.007	0.999	0.007	0.999
	STD	0.008	0.999	0.008	0.999
	GED	0.010	0.999	0.010	0.999
	SNORM	0.007	0.999	0.007	0.999
	SSTD	0.008	0.999	0.008	0.999
	SGED	0.007	0.999	0.007	0.999
	GHYP	0.066	0.999	0.067	0.999
	JSU	0.008	0.999	0.008	0.999
	NIG	0.021	0.999	0.021	0.999
	TIITL _{LEGU}	0.013	0.999	0.013	0.999

Note(s): Significance level: $\alpha = 0.05$.

Table 7: Estimated EGARCH (1,1) models diagnostic tests (TEN returns).

Model	Innovation Dist.	Ljung-Box Q-Statistic	p-value	ARCH-LM Statistic	p-value
EGARCH (1,1)	NORM	0.021	0.999	0.021	0.999
	STD	0.091	0.999	0.092	0.999
	GED	5.936	0.981	5.923	0.981
	SNORM	0.021	0.999	0.020	0.999
	SSTD	0.088	0.999	0.088	0.999
	SGED	0.010	0.999	0.010	0.999
	GHYP	0.045	0.999	0.046	0.999
	JSU	1251.5	2.2E-16	2317.7	2.2E-16
	NIG	0.477	0.999	0.474	0.999
	TIITL _{EGU}	0.039	0.999	0.039	0.999

Note(s): Significance level: $\alpha = 0.05$.

6.3 Forecasts Evaluation of the EGARCH Models

The forecast evaluation metrics for the out-of-sample are reported in Table 8, and the least MSE, RMSE and MAE values belong to the EGARCH-TIITL_{EGU} model. Therefore, the EGARCH-TIITL_{EGU} model is statistically efficient and displays superior capability in forecasting the volatility of the Nigerian stock market relative to other models.

Table 8: Forecasts evaluation metrics of the estimated EGARCH (1,1) models.

Model	Innovation Dist.	MSE	RMSE	MAE	MSE	RMSE	MAE
		(UBA returns)			(TEN returns)		
EGARCH (1,1)	NORM	8.155	2.856	1.752	1.731	1.316	0.213
	STD	8.151	2.855	1.751	1.733	1.316	0.197
	GED	8.151	2.855	1.751	1.733	1.316	0.197
	SNORM	8.121	2.849	1.752	1.730	1.315	0.267
	SSTD	8.151	2.855	1.751	1.733	1.316	0.197
	SGED	8.150	2.855	1.751	1.777	1.332	0.354
	GHYP	8.159	2.856	1.753	1.733	1.316	0.197
	JSU	8.151	2.855	1.751	1.733	1.316	0.198
	NIG	8.151	2.855	1.751	1.733	1.316	0.197
	TIITL_{EGU}	8.050	2.837	1.736	1.547	1.244	0.182

Note(s): Bolded values indicate the conditional distribution with the least MSE, RMSE and MAE.

7. Conclusion

The estimation of the TIITL_{EGU} model parameters using the MLE procedure, and introduction as a novel distributed innovation function for the EGARCH-volatility model is considered. The density and cumulative functions, failure rate function, quantile function, standardized density function and other mathematical properties are derived. Monte-Carlo experiments are carried out to study the performance of the MLE procedure. The experiments results indicate that the MLE is asymptotically unbiased and consistent given that the ME tend to be closer to the true values of the parameter with minimum MSEs, and RMSEs as the sample size increases.

Additionally, the standardized TIITL_{EGU} density is presented as a novel distributed innovation function for the EGARCH volatility model for investigating the volatility of the Nigerian

stock market via the UBA and TEN returns. The empirical findings showed that the EGARCH-TIITL_{EGU} model has the highest log-likelihood, and least AIC and BIC values. Equally, the EGARCH-TIITL_{EGU} model has the least forecast evaluation metrics among other models. In conclusion, the EGARCH-TIITL_{EGU} model is best for investigating the volatility of the Nigerian stock market.

References

- [1] Atoi, N. V. (2014). Testing volatility in Nigeria stock market using GARCH models. *CBN Journal of Applied Statistics*, 5:65–93.
- [2] Adubisi, O. D., Abdulkadir, A. and Adubisi, C. E. (2022b). A new hybrid form of the skew-t distribution: estimation methods comparison via Monte Carlo simulation and GARCH model application. *Data Science in Finance and Economics*, 2:54-79.
- [3] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with estimates of the variance of United Kingdom Inflation. *Econometrica*, 50:987-1008.
- [4] Investopedia (2021). Volatility definition. Retrieved December 15, 2023, from: <https://www.investopedia.com/terms/v/volatility.asp>.
- [5] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Econometrics and Statistics*, 69:542-547.
- [6] Adubisi, O. D., Abdulkadir, A. and Adashu, D. J. (2023a). Improved parameter estimators for the flexible extended skew-t model with extensive simulations, applications and volatility modeling. *Scientific African*, 19:e01443.
- [7] Olayemi, M. S., Olubiyi A., Olajide O. O. and Adigun, A. (2023). The use of standardized exponentiated Gumbel error innovation distribution to forecast volatility: A comparative study. *Journal of statistical modeling and analytics*, 5:23-31.
- [8] Altun, E. (2019). Two-sided exponential-geometric distribution: Inference and volatility modeling. *Computational Statistics*, 34:1215–1245.
- [9] Obalowu, J. and David, O. R. (2023). The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures. *Communication in Physical Sciences*, 9:572–584.
- [10] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A New Approach. *Econometrica*, 59:347-370.
- [11] Qiu, S. (2023). Research on the impact of crude oil price fluctuation based on EGARCH model. *Frontiers in Business, Economics and Management*, 10:27-33.
- [12] Yelamanchili, R. K. (2020). Modeling Stock Market Monthly Returns Volatility Using GARCH Models Under Different Distributions. *International Journal of Accounting and Finance Review*, 5:42-50.
- [13] Samson, T. K., Enang, E. T. and Onwukwe, C. E. (2020). Estimating the parameters of GARCH models and its extension: comparison between Gaussian and non-Gaussian innovation distributions. *Covenant Journal of Physical and Life Sciences (CJPL)*, 8:46-60.
- [14] Olanrewaju, R. O. and Oseni, E. (2021). GARCH and its variants models: An application of crude oil distributions in Nigeria. *International Journal of Accounting, Finance and Risk Management*, 6:25–35.
- [15] Husein, I. and Lubis A. I. D. (2022). EGARCH model prediction for sale stock price. *Jurnal Varian*, 6:49-60.
- [16] Adenomon, M. O. and Idowu, R. A. (2022). Modelling the impact of COVID-19 pandemic on some Nigerian sectorial stocks: Evidence from GARCH models with structural breaks. *FinTech, MDPI*, 2:1–20.
- [17] Yahaya, A., John, S. A., Adegoroye, A. and Olorunfemi, O. A. (2023). Stock market liquidity and volatility on the Nigerian exchange limited (NGX). *World Journal of Advanced Research and Reviews*, 20:147-156.
- [18] Harvey, A. C. and Chakravarty, T. Beta-t(E)GARCH. Cambridge working papers in

Economics, University of Cambridge, 2008.

[19] Liu, H. and Hung, J. (2010). Forecasting SandP-100 stock index volatility: The role volatility asymmetry and distributional assumption in GARCH models. *Expert systems with applications*, 37:4928-4934.

[20] Harvey, A. and Sucarrat, G. (2013). EGARCH models with fat tails, skewness and leverage. *Computational Statistics and Data Analysis*, 76:320-338.

[21] Feng, L. and Shi, Y. (2017). A simulation study on the distributions of disturbances in the GARCH model. *Cogent Economics and Finance*, 5:1-19.

[22] Agboola, S., Dikko, H. G. and Asiribo, O. E. (2018). On the exponentiated skewed student-t error distribution on some volatility models: Evidence of standard and poor-500 index return. *Asian Journal of Applied Sciences*, 11:38-45.

[23] Agboola, S., Dikko, H. G. and Asiribo, O. E. (2019). Some volatility modeling using new error innovation distribution. *International Research Journal of Applied Sciences*, 1:57-62.

[24] Adubisi, O. D., Abdulkadir, A., Farouk, U. A. and Chiroma, H. (2022a). The exponentiated half logistic skew-t distribution with GARCH-type volatility models. *Scientific African*, 16: e01253.

[25] Elgarhy, M., Nasir, A., Jamal, F. and Ozel, G. (2018). The type II Topp Leone generated family of distributions: Properties and applications. *Journal of Statistics and Management Systems*, 21:1529-1551.

[26] Nadarajah, S. (2006). The exponentiated Gumbel distribution with climate application. *Environmetrics*, 17:13-23.

[27] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307-327.

[28] Brooks, C. and Burke, S. P. (2003). Information criteria for GARCH model selection. *The European Journal of Finance*, 9:557-580.

SINGLE AND DOUBLE ACCEPTANCE SAMPLING PLAN FOR TRUNCATED LIFE TESTS BASED ON GAMMA LINDLEY DISTRIBUTION

Sriramachandran G. V.



Associate Professor, Dr. NGP Institute of Technology, Coimbatore, India
gvsriramachandran72@gmail.com

Abstract

For time-truncated life tests, this work defines single acceptance and double acceptance sampling plans assuming that the product's lifespan follows the Gamma Lindley distribution. The minimum sample size needed in a single acceptance sampling plan for lot approval is calculated for a range of parameter combinations and a fixed test termination time. This ensures the given average product life and the corresponding number of failures. Operational characteristic and producer risk values are also tabulated for these parameter values. Using a double acceptance sampling plan, the best first and second samples are obtained to ensure that the products specified average with a certain level of customer trust. Finally, under the same conditions, the minimum sample size obtained using these strategies are compared with other acceptance sampling plans.

Keywords: Acceptance sampling plan, consumer's risk, producer's risk, time truncated life test, Gamma Lindley distribution.

I. Introduction

A decision is made based on the quality of the samples after samples are removed from the lot and examined to determine whether the lot should be accepted. When inspecting the entire lot, it is usually not a good idea because it takes a long time and is very expensive. So, in order to minimize both producer and consumer risk, decisions are made, inspections are carried out, and samples are taken from the lot. The term acceptance sampling plan based on time truncated life tests refers to the fact that engineers typically incorporate the time parameter into the plan. There exist multiple methods for executing acceptance sampling plans that include time-truncated life tests. The two most widely used plans among the many that are available are the single acceptance and double acceptance sampling plans (SASP and DASP). Because of this, the two acceptance sampling plans are the primary subject of this work.

It is known as a SASP when only one sample is selected for decision-making out of the entire lot. This plan involves running the test for a predetermined amount of time and counting the number of failures that occur during that time. For a lot to be accepted, the number of failures observed at the end of the predetermined time must be less than the acceptance number; if it is greater than the acceptance number, the lot is rejected. In order to determine the consumer risk for an acceptance number and a fixed time t , it is therefore necessary to determine the minimum

sample size required to guarantee the average life of the product at different probability values. A minimum sample size must be determined for both samples in a DASP, where the acceptance numbers for the first and second samples are fixed at zero and one, respectively. Numerous researchers have examined the SASP and DASP for a range of lifetime models whose lives follow a specific probability distribution. Among the works are those by Mahendra Saha et. Al. [1] who created SASP and DASP for time-truncated life tests based on transmuted Rayleigh distribution (TRD), Amjad and Amer [2] who created an SASP based on time-truncated life tests for an exponentiated Frechet distribution; Sridhar Babu et al., [3] who examined a DASP for an exponentiated Frechet distribution with known shape parameters; Lio et. Al. [4] examined SASP for time-truncated life tests for Birnbaum-Saunders distribution for percentiles; Min and Gui [5] investigated SASP based on time-truncated life tests for Burr type X distribution; Balakrishnan et al., [6] investigated the generalized Birnbaum-Saunders model's SASP based on time-truncated life tests; Amjad and Obeidat [7] created an SASP using time-truncated life tests based on the Tsallis-q-exponential distribution. Further, Sriramachandran [8] studied sampling plan for truncated life test for Logistic family of distributions.

In order to analyse the lifetime data, Lindley [9] introduced a probability distribution that alternates between the gamma and exponential distributions in the appropriate ratio. The Lindley distribution has not been investigated since exponential and gamma distributions are applied independently in numerous statistical domains and are the subject of extensive research by researchers. Ghitany et al., [10] was the first to study the properties of Lindley distribution, which led to its popularity in data modelling. Numerous authors have also applied it to a variety of disciplines. Shanker et al., [11] also transformed the one parameter Lindley distribution into the two parameter Lindley distribution. SASP was created by Wu [12] for two parameter Lindley distribution, Deniz [13] described the discrete form of Lindley distribution and stated its properties, Ghitany et al [14] stated the two-parameter weighted Lindley distribution and its applications, Sihem [15] and Zeghdoudi [16] defined the Gamma Lindley distribution and studied its properties. The article's goal is to create single acceptance and double acceptance sampling plans that take into account the Gamma Lindley distribution (GLD) for the average life of the product being tested. For a given consumer risk probability, the minimum sample size needed to guarantee a given average life is obtained is known as the single acceptance sampling plan. For this plan, the production risk and the OC values are also obtained. The zero-one model is used in the double acceptance sampling plan to determine the minimum sample size for the first and second samples for different probabilities in order to guarantee the consumer risk. The structure of this article is as follows: Section 2 elaborates on the GLD. The average life of the product is assumed to follow GLD in Section 3 when designing and deriving the SASP for time-truncated life tests. In Section 4, a DASP for time-truncated life tests is designed and derived with the assumption that the product lifetime follows GLD. The minimum sample size required to be inspected under the SASP is presented along with the OC value and producer risk. In Section 5, the plan proposed in this article is compared with other suitable acceptance sampling plan based on the minimum sample size to be inspected. Section 6 submits the conclusions from this article.

II. Gamma Lindley Distribution Model

In lifetime data modelling, Lindley distribution is the special case of two parameter Lindley distribution (TPLD) studied by [11]. The probability density function (pdf) of TPLD is stated as:

$$f(\epsilon; \tau, \omega) = \frac{\tau(\omega + \tau\epsilon)}{\omega + 1} e^{-\tau\epsilon} \quad (1)$$

Hence, the corresponding cumulative distribution function (cdf) and the hazard rate function (hrf) of TPLD are stated as:

$$F(\epsilon; \tau, \omega) = 1 - \frac{1+\omega+\tau\epsilon}{\omega+1} e^{-\tau\epsilon} \tag{2}$$

$$h(\epsilon; \tau, \omega) = \frac{\tau(\omega+\tau\epsilon)}{1+\omega+\tau\epsilon}, \epsilon > 0, \tau > 0, \omega > -1, \omega > -\tau\epsilon \tag{3}$$

[16] mixes the gamma distribution with two parameter Lindley distribution and formed a new distribution called the GLD and studied all its properties.

The pdf of GLD is

$$f(\epsilon; \tau, \omega) = \begin{cases} \frac{\tau^2((\omega+\omega\tau-\tau)\epsilon+1)}{\omega(\tau+1)} e^{-\tau\epsilon}, \epsilon, \tau, \omega > 0 \\ 0, \text{ otherwise} \end{cases} \tag{4}$$

The corresponding cdf and hrf are defined as

$$F(\epsilon; \tau, \omega) = 1 - \frac{((\tau\omega+\omega-\tau)(\tau\epsilon+1)+\tau)}{\omega(\tau+1)} e^{-\tau\epsilon} \tag{5}$$

$$h(\epsilon; \tau, \omega) = \frac{\tau^2((\omega+\omega\tau-\tau)\epsilon+1)}{((\tau\omega+\omega-\tau)(\tau\epsilon+1)+\tau)} \tag{6}$$

Table 1: Minimum sample size for different parameter value

Parameter value	P*	c	$\frac{\epsilon}{u_0}$				
			0.628	0.942	1.257	2.356	3.141
$\tau = 0.5,$ $\omega = 1$	0.95	0	6	4	3	2	1
		1	10	7	5	3	2
		2	14	9	7	4	4
	0.99	0	9	6	4	2	2
		1	14	9	7	4	3
		2	18	11	9	5	4
$\tau = 1,$ $\omega = 1$	0.95	0	6	4	3	2	1
		1	10	6	5	3	3
		2	13	9	7	4	4
	0.99	0	9	6	4	2	2
		1	13	9	7	4	3
		2	17	11	9	5	4
$\tau = 1,$ $\omega = 0.5$	0.95	0	5	4	3	2	1
		1	9	6	5	3	3
		2	12	8	7	4	4
	0.99	0	8	5	4	2	2
		1	12	8	6	4	3
		2	15	11	8	5	5

The expected value and variance of GLD are derived as

$$E(x) = \frac{2\omega(1+\tau)-\tau}{\tau\omega(1+\tau)} \tag{7}$$

$$Var(x) = \frac{-(-2\omega\tau + \tau)^2 + (2+6\tau)\omega^2 - 2\omega(\omega\tau - 3\omega\tau^2 + 2\tau^2)}{\omega^2\tau^2(1+\tau)^2} \quad (8)$$

The cdf, pdf and hrf of GLDs for different parameter values are presented in Figure 1 and 2.

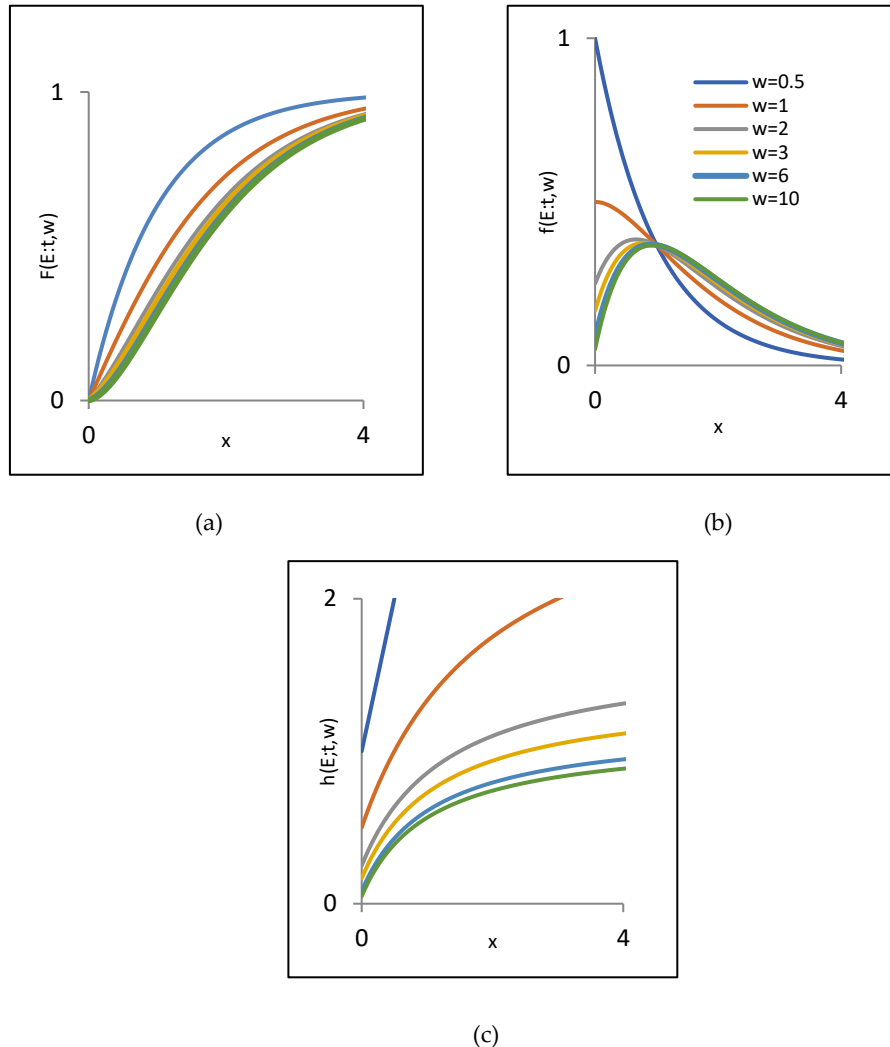


Figure 1: (a) cdf, (b) pdf and (c) hrf of GLDs with parameters $\tau = 1$ and $\omega = 0.5, 1, 2, 3, 6$ and 10 .

III. Design of Single Acceptance Sampling Plan (SASP) Based on Mean

When creating the SASP with time-truncated life tests, it was assumed that the products being studied would have an average lifespan of v . The above assumption is tested using the hypothesis test, where the product's precise average life, denoted by v_0 and $H_0: v \geq v_0$, alternative hypothesis $H_1: v < v_0$. The test's level of significance is defined as ρ^* represents the consumer's confidence level, the value $1 - \rho^*$ is called the level of significance for the test. The application of the binomial distribution in this case is noteworthy because the sample size is noticeably large. As proposed by the plan; to locate the minimum sample size, we have to iterate the inequality

$$\sum_{i=0}^c \binom{n}{i} p^i (1-p)^{n-i} \leq 1 - \rho^* \quad (9)$$

where $\rho = F(\epsilon; \tau, \omega)$ as in (5).

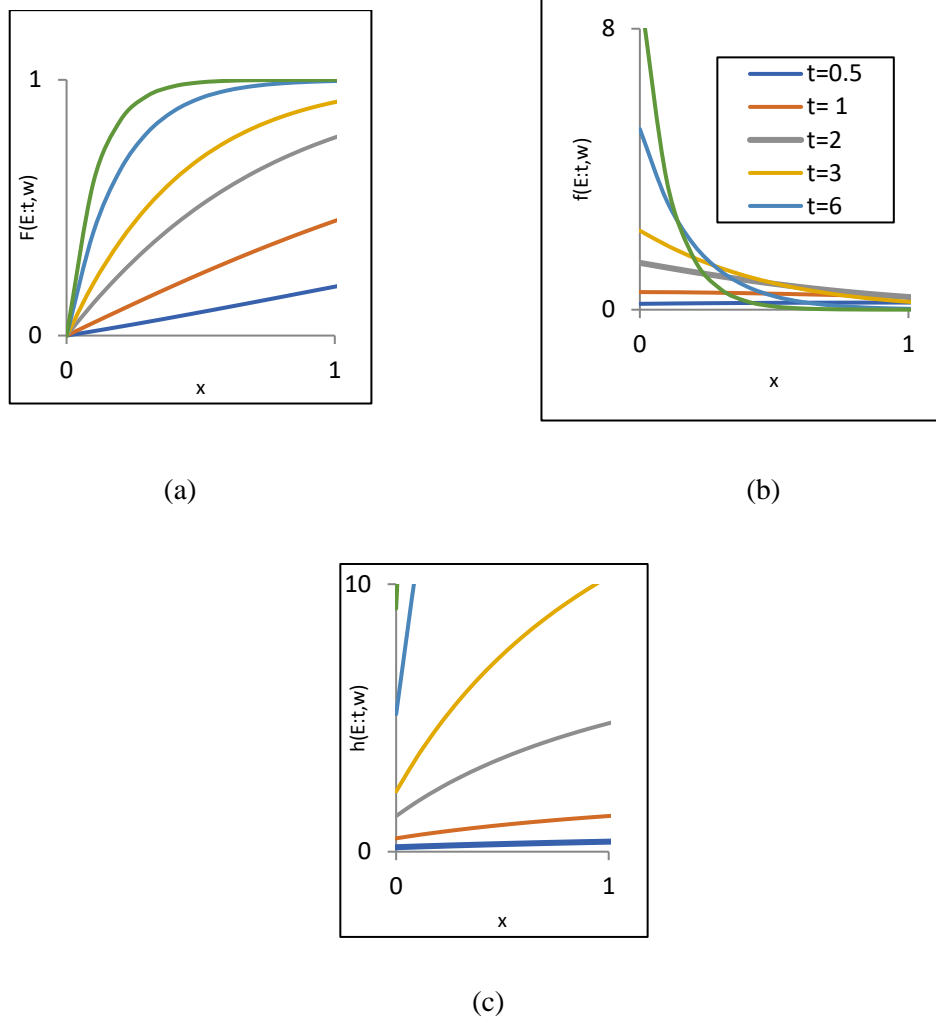


Figure 2: (a) *cdf*, (b) *pdf* and (c) *hrf* of GLDs with parameters $\omega = 1$ and $\tau = 0.5, 1, 2, 3, 6$ and 10

Minimum sample size for $\epsilon/v_0 = (0.628, 0.942, 1.257, 2.356$ and $3.141)$, $\rho^* = 0.95$ and 0.99 and for the parameters $\tau = 0.5, \omega = 1, \tau = 1, \omega = 1$ and $\tau = 1, \omega = 0.5$ are calculated and tabulated in the Table 1. The OC values of the plan give the probability of acceptance of the lot and it is stated as

$$L(p) = \sum_{i=0}^c \binom{n}{i} p^i (1-p)^{n-i} \tag{10}$$

where $\rho = F(\epsilon; \tau, \omega)$ in (5). For the fixed acceptance number i.e. $c=2$ the OC values are calculated and the values are tabulated in the Table 2 for the parameters $\tau = 0.5, \omega = 1, \tau = 1, \omega = 1$ and $\tau = 1, \omega = 0.5$ respectively. Next, it is necessary to reduce the producer's risk and we should guarantee that rejection of good lot is minimum. By fixing the PR as 95%, the value of ϵ/v_0 are obtained for the plan. The value obtained is the least number for which $\rho = F(\epsilon; \tau, \omega)$ as in (5) satisfies the inequality

$$\sum_{i=0}^c \binom{n}{i} p^i (1-p)^{n-i} \geq 0.95 \tag{11}$$

The least values of ϵ/v_0 which satisfies the equation (11) are determined and presented for the proposed plan and presented in Table 3.

Table 2: OC values for the parameter value when $c=2$

Parameter value	P*	$\frac{\varepsilon}{v_0}$	n	$\frac{v}{v_0}$					
				2	4	6	8	10	12
$\tau = 0.5,$ $\omega = 1$	0.95	0.628	14	0.445	0.855	0.947	0.976	0.987	0.992
		0.942	9	0.449	0.864	0.952	0.979	0.989	0.993
		1.257	7	0.411	0.850	0.948	0.977	0.988	0.993
		2.356	4	0.391	0.843	0.947	0.977	0.989	0.993
		3.141	4	0.181	0.696	0.884	0.947	0.973	0.984
	0.99	0.628	18	0.269	0.754	0.902	0.952	0.974	0.984
		0.942	11	0.300	0.785	0.919	0.962	0.979	0.988
		1.257	9	0.224	0.736	0.898	0.952	0.974	0.985
		2.356	5	0.196	0.716	0.893	0.951	0.974	0.985
		3.141	4	0.181	0.696	0.884	0.947	0.973	0.984
$\tau = 1,$ $\omega = 1$	0.95	0.628	13	0.401	0.812	0.924	0.963	0.979	0.987
		0.942	9	0.380	0.805	0.922	0.962	0.979	0.987
		1.257	7	0.363	0.799	0.920	0.961	0.978	0.987
		2.356	4	0.387	0.816	0.929	0.967	0.982	0.989
		3.141	4	0.191	0.671	0.859	0.929	0.960	0.976
	0.99	0.628	17	0.218	0.681	0.857	0.926	0.957	0.973
		0.942	11	0.236	0.704	0.871	0.934	0.962	0.977
		1.257	9	0.185	0.660	0.849	0.922	0.955	0.972
		2.356	5	0.193	0.676	0.860	0.930	0.960	0.976
		3.141	4	0.191	0.671	0.859	0.930	0.960	0.976
$\tau = 1,$ $\omega = 0.5$	0.95	0.628	12	0.333	0.752	0.891	0.943	0.967	0.979
		0.942	8	0.368	0.775	0.902	0.950	0.971	0.982
		1.257	7	0.285	0.715	0.870	0.932	0.960	0.975
		2.356	4	0.362	0.765	0.896	0.946	0.969	0.980
		3.141	4	0.193	0.619	0.812	0.896	0.937	0.960
	0.99	0.628	15	0.188	0.625	0.818	0.901	0.941	0.962
		0.942	11	0.155	0.585	0.793	0.885	0.931	0.951
		1.257	8	0.193	0.629	0.821	0.902	0.942	0.962
		2.356	5	0.174	0.602	0.803	0.891	0.935	0.957
		3.141	5	0.064	0.418	0.669	0.803	0.875	0.916

IV. Design of Double Acceptance Sampling Plan (DASP)

DASP, or two-stage acceptance sampling, is the preferred method for providing greater protection to producers and consumers. Because a second sample is tested before a final decision is made on the lot, producers are given double protection, as the name suggests. As a result, it diminishes producer risk and offers complete protection to producers. The plan's parameters are as follows: the testing time t , the first sample size ($n1$), its acceptance number ($c1$), and the second sample size ($n2$). In order to approve the lot, the sample must substantiate the hypothesis that the sample mean exceeds the given mean. The lot will be turned down otherwise. The consumer's risk is now fixed at no more than $(1-P^*)$, where P^* is the confidence level.

Then probability of acceptance of the lot is

$$PA = \sum_{i=0}^{c_1} \binom{n_1}{i} p^i (1-p)^{n_1-i} + \sum_{x=c_1+1}^{c_2} \binom{n_1}{x} p^x (1-p)^{n_1-x} \sum_{j=0}^{c_2} \binom{n_2}{j} p^j (1-p)^{n_2-j} \tag{12}$$

where p is defined in Equation (5) and depends on ratio $\frac{\epsilon}{v_0}$. As we are considering only zero-one failure form i.e., $c_1=0$ and $c_2=1$, the above Equation (13) for the considered GLD is given by

Table 3: Minimum ratio when producer's risk=95%

Parameter value	P*	c	$\frac{\epsilon}{v_0}$				
			0.628	0.942	1.257	2.356	3.141
$\tau = 0.5,$ $\omega = 1$	0.95	0	41.49	41.82	42.17	53.46	37.04
		1	9.94	10.44	9.91	10.87	9.03
		2	6.12	5.90	6.08	6.12	8.15
	0.99	0	61.90	62.23	55.80	53.46	71.
		1	13.89	13.42	13.93	14.77	14.49
		2	7.85	7.22	7.87	7.93	8.15
$\tau = 1,$ $\omega = 1$	0.95	0	55.33	55.44	55.60	69.76	47.04
		1	12.80	11.21	12.29	12.91	17.22
		2	7.11	7.19	7.25	6.86	9.15
	0.99	0	82.88	82.99	73.98	69.76	93.00
		1	16.78	17.20	17.63	18.00	17.22
		2	9.43	8.93	9.59	9.14	9.15
$\tau = 1,$ $\omega = 0.5$	0.95	0	61.22	73.46	73.52	91.86	61.24
		1	14.99	14.51	15.81	16.20	21.60
		2	8.42	8.00	9.12	8.24	10.99
	0.99	0	97.95	91.83	98.02	91.86	122.47
		1	20.30	19.83	19.37	22.94	21.60
		2	10.73	11.47	10.67	11.23	14.97

$$PA = (1-p)^{n_1} [1 + n_1 p (1-p)^{n_2-1}] \tag{13}$$

where p is given by the equation (5). Our aim is to find the minimum sample size for the plan, for this we have to minimize the following equation:

$$(1-p)^{n_1} [1 + n_1 p (1-p)^{n_2-1}] \leq 1 - P^* \tag{14}$$

Now, for the given consumer's confidence level P*, the minimum sample size for both the samples n_1 and n_2 , which ensure $v \geq v_0$, can be found by the solution of the following optimization problem, given as:

$$\begin{aligned} \text{Min ASN} &= n_1 + n_1 n_2 p (1-p)^{n_1-1} \\ \text{subject to} & (1-p)^{n_1} [1 + n_1 p (1-p)^{n_2-1}] \leq 1 - P^*, \\ & 1 \leq n_2 \leq n_1 \\ & n_1, n_2 \text{ are integers} \end{aligned} \tag{15}$$

While solving the above optimization problem, it provides many solutions for both n_1 and n_2 . We take the solution which minimizes our objective function i.e. our ASN as our best solution. Minimum sample size obtained for $P^* = 0.90, 0.95$ and 0.99 and $\frac{\epsilon}{v_0} = 0.628, 0.942, 1.257$ and 2.356 are presented in Table 4.

V. Comparison of GLD With Other Models

Planning an acceptance sampling plan with a minimum sample size is considered more efficient because it will cut down on the amount of time and money spent on inspection, as desired by quality engineers. This section compares and provides the following Figures 3 and 4 for the GLD's CDF and PDF, as well as the two-parameter Lindley distribution (TPLD) and transmuted Rayleigh distribution (TRD). In addition, the following tables 11 and 12 list the minimal sample sizes that were obtained from this distribution under comparable conditions. The following observations about the sample size n are found in Table 1 to 3.

- There is little variation in sample size n value when τ increases from 0.5 to 1, keeping $\omega = 1$.
- There is a marginal shift in the sample size n value when $\tau = 1$ and ω increases from 0.5 to 1
- Table 1 show that an increase in the value of c corresponds to an increase in the value of n .
- Table 4 indicates that the minimum sample size in a DASP when, $\tau = 1, \omega = 0.5$.
- When comparing GLD to TPLD and TRD, Table 5 indicates that the former requires the smallest sample size.
- According to Table 6, the GLD for DASP requires the smallest sample size when compared to TPLD and TRD.

Table 4: Minimum sample size for plan

Parameter value	P*	$\frac{\epsilon}{v_0}$							
		0.628		0.942		1.257		2.356	
		(n ₁ , n ₂)	ASN	(n ₁ , n ₂)	ASN	(n ₁ , n ₂)	ASN	(n ₁ , n ₂)	ASN
$\tau = 0.5$ $\omega = 1$	0.90	(6,3)	6.55	(4,2)	4.35	(3,1)	3.18	(2,1)	2.13
	0.95	(7,4)	7.51	(4,3)	4.53	(3,2)	3.37	(2,1)	2.13
	0.99	(10,5)	10.19	(6,3)	6.14	(4,3)	4.53	(3,1)	3.18
$\tau = 1$ $\omega = 1$	0.90	(5,4)	5.92	(4,1)	4.17	(3,1)	3.18	(2,1)	2.13
	0.95	(6,5)	6.79	(4,3)	4.53	(3,2)	3.37	(2,1)	2.13
	0.99	(9,5)	9.22	(6,3)	6.14	(4,3)	4.53	(3,1)	3.18
$\tau = 1$ $\omega = 0.5$	0.90	(5,2)	5.38	(4,1)	4.14	(3,1)	3.17	(2,1)	2.17
	0.95	(6,3)	6.36	(4,2)	4.29	(3,2)	3.35	(2,1)	2.17
	0.99	(8,5)	8.23	(5,4)	5.25	(4,3)	4.20	(3,1)	3.02

Therefore, if quality engineers wish to check the lot with a minimum sample size, the GLD of the SASP and DASP are commendable.

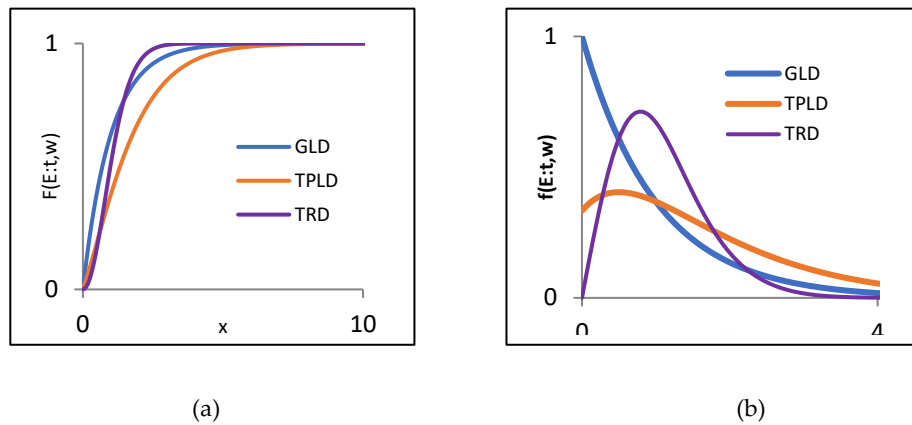


Figure 3: Comparison of (a) cdf and (b) pdf of GLD, TPLD and TRD when $\tau = 1$ and $\omega = 0.5$

Table 5: Comparison of single acceptance sampling plan when $\tau = 0.5, \omega = 1, P^*=0.99$ and $c=0$

Distributions	$\frac{\varepsilon}{v_0}$		
	0.628	0.942	1.257
GLD	9	6	4
TPLD	12	7	5
TRD	16	8	3

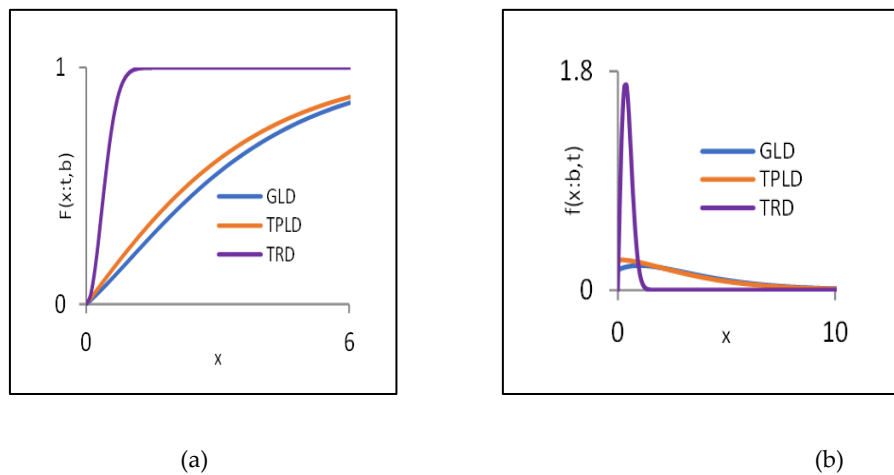


Figure 4: Comparison of (a) cdf and (b) pdf of GLD, TPLD and TRD when $\tau = 0.5$ and $\omega = 1$

Table 6: Comparison of double acceptance sampling plan when $\tau = 0.5, \omega = 1, P^*=0.99$ and $c_1=0$ and $c_2=1$

Distributions	$\frac{\varepsilon}{v_0}$		
	0.628	0.942	1.257
GLD	(10,5)	(6,3)	(4,3)
TPLD	(12,6)	(7,4)	(5,3)
TRD	(16,17)	(8,4)	(3,2)

VI. Conclusion

Within this paper, time-truncated life tests based on GLD were suggested for SASP and DASP. We consider the zero-one model for the DASP. Different GLD parameter values are used to calculate the minimum sample size needed to ensure a given average product life, the minimum producer's risk, and the OC values. Tables for both plans display the results that were obtained. It was discovered that the suggested plan required the smallest sample size for both single and double acceptance sampling plans when it came to minimum sample size comparisons with TPLD and TRD. Based on how this suggested plan is used, future work may expand the scope to include additional plans.

References

- [1] Mahendra Saha, Harsh Tripathi and Sanku Dey. (2021). Single and double acceptance sampling plans for truncated life tests based on transmuted Rayleigh distribution, *Journal of Industrial and Production Engineering*, 38(5):356-368.
- [2] Amjad D. Al-Nasser & Amer I. Al-Omari, (2013). Acceptance sampling plan based on truncated life tests for exponentiated Fréchet distribution, *Journal of Statistics and Management Systems*, 16 (1).
- [3] Sridhar Babu, M, Srinivasa Rao G, and Rosaiah. K, (2021). Double Acceptance Sampling Plan for Exponentiated Fréchet Distribution with Known Shape Parameters, *Mathematical Problems in Engg.*, 9.
- [4] Lio. Y. L., Tzong-Ru Tsai, & Shuo-Jye Wu, (2009). Acceptance Sampling Plans from Truncated Life Tests Based on the Birnbaum–Saunders Distribution for Percentiles, *Communications in Statistics Simulation and Computation*, 39 (1): 119-136.
- [5] Min Hu & Wenhao Gui, (2018). Acceptance sampling plans based on truncated life tests for Burr type X distribution, *Journal of Statistics and Management Systems*, 21 (3): 323-336.
- [6] Balakrishnan. N., Víctor Leiva and Jorge López, (2007). Acceptance Sampling Plans from Truncated Life Tests Based on the Generalized Birnbaum–Saunders Distribution, *Communications in Statistics - Simulation and Computation*, 36 (3): 643-656.
- [7] Amjad D. Al-Nasser and Mohammed Obeidat, (2020). Acceptance sampling plans from truncated life test based on Tsallis q-exponential distribution, *Journal of Applied Statistics*, 47 (4): 685–697.
- [8] Sriramachandran, G.V. (2023), Sampling plans based on truncated life test for logistic family of distributions, *Reliability: Theory and Applications*, 18 (4): 382-390.
- [9] Lindley. D. V., (1958). Fiducial distributions and Bayes theorem, *J. Roy. Soc. Ser. B*, 20 :102-107.
- [10] Ghitany. M. E., Atieh. B. and Nadarajah. S., (2008). Lindley distribution and its application, *Mathematics and Computers in Simulation*, 78(4): 493-506.
- [11] Shanker, R. Shambhu Sharma and Ravi Shanker, (2013). A Two-Parameter Lindley Distribution for Modelling Waiting and Survival Times Data, *Applied Mathematics*, 4 (2).
- [12] Chien-Wei Wu, Ming-Hung Shu and Nien-Yun Wua, (2021). Acceptance sampling schemes for two-parameter Lindley lifetime products under a truncated life test, *Quality Technology & Quantitative Management*, 18 (3): 382-395.
- [13] Deniz E. G. and Ojeda E. C., (2011). The Discrete Lindley Distribution-Properties and Applications, *Journal of Statistical Computation and Simulation*, 81(11): 1405-1416.
- [14] Ghitany. M. E., Alqallaf. F, Al-Mutairi, D. K and Hussain H. A., (2011). A Two Parameter Weighted Lindley Distribution and Its Applications to Survival Data, *Mathematics and Computers in Simulation*, 81(6):1190-1201.

[15] Sihem Nedjar and Halim Zeghdoudi, (2016). On gamma Lindley distribution: Properties and simulations, *Journal of Computational and Applied Mathematics*, 298: 167-174.

[16] Zeghdoudi. H and Nedjar, S (2015). Gamma Lindley distribution and its application, *Journal of Applied Probability. Statistics*, 11(1).

FUZZY VARIABLE LINEAR PROGRAMMING PROBLEMS USING A FUZZY DUAL SIMPLEX ALGORITHM

Srinivasa Rao Kolli¹, U.V. Adinarayana Rao², & Taviti Naidu Gongada³

Dept. of Operations, GITAM School of Business, GITAM (Deemed to be University),
Visakhapatnam, AP, India^{1,2,3}

srinivaskolli4@gmail.com¹

auppu@gitam.edu²

dr.tavitinaidugongada@gmail.com³

Abstract

In modern research, several brilliant minds investigate linear programming problems involving fuzzy variable quantities. Many researchers have turned to linear programming by fuzzy variables to address this problem. Various fuzzy simplex approaches have been developed, using ranking functions to handle fuzzy numbers. Results from this research suggest that linear ranking functions can provide a straightforward interpretation of problems involving linear programming by fuzzy variable quantities. To solve these types of problems, the Fuzzy Dual Simplex Tableau method is often applied, which proves useful for sensitivity analysis when modifications are made to the activity vectors of the fundamental columns. In this study, a numerical case is presented to demonstrate the potential benefits of this approach for future technologies.

Keywords: Linear programming by fuzzy variable quantity, fuzzy numbers, ranking functions, fuzzy simplex and duality algorithms, and trapezoidal fuzzy numbers.

I. Introduction

The fuzzy set model has proven useful in several areas, including simulation, artificial intelligence, control systems and organisational skills, scientific modelling, operations research, and manufacturing applications. The results of fuzzy mathematical programming problems have been the subject of extensive theoretical and computational research in recent years. On a more global scale, this concept was initially developed by [2] and presented the first-ever construction of fuzzy linear programming (FLP). Since then, various authors have investigated different FLP challenges and proposed various approaches to addressing them [3-6]. To address issues with fuzzy linear programming, several authors have turned to the method of fuzzy-number comparison. Methods that employ the idea of comparing fuzzy numbers with rank-order functions [2, 3, 6] have proven to be the most useful. Researchers have developed ranking functions to suit their purposes, but it is possible that these proposals are not necessarily based on industry-standard practices Fuzzy mathematical programming is reviewed in [7]. While there is a wealth of literature on modelling, publications on duality still need to be made available, leading to issues with a linear program in a fuzzy setting [7,8, 9] suggesting a primal-dual strategy utilizing linear ranking functions for resolving

linear programming with fuzzy variable quantities. Based on the duality outcomes published [4], this research provides an innovative fuzzy dual simplex strategy for solving fuzzy number linear programming problems. This method will benefit sensitivity analysis when the recent set of primary vectors rejection extended method is fundamental after difference due to changes in the elementary column. We base our method for solving FNLP problems on the linear rank functions initially introduced by [11] and exploited by in [4-6]. They point out that while trapezoidal fuzzy numbers were used to construct the procedure described in this research, linear ranking functions are only some of the options available.

Introduced a duality model to address issues with FVLP [4]. This inquiry emphasizes the problems associated with FVLPs. We, therefore, analyze the initial reactions to the theory and determine how people respond to the contrast presented by fuzzy numbers using a linear ranking function. Furthermore, a fuzzy elementary achievable resolution is suggested for FVLP problems and related optimality requirements, and a fuzzy dual simplex technique is projected to be useful in tackling FVLP concerns.

The procedures to be followed in writing this paper include providing an overview of the basic concepts of fuzzy established models in Part 2, presenting a brief overview of linear programming using fuzzy numbers, discussing the issue and definition of the consistent dual dilemma in Section II&III, advancing a fuzzy dual technique based on the fuzzy simplex tableau for solving linear programming problems using fuzzy numbers in section IV, and presenting the final results in Section V.

II. Preliminaries and fundamental concepts

The fundamental terms and symbols of the fuzzy established model are briefly covered (taken from [4]).

I. Definition Assume \mathbb{R} as a collective set. \bar{a} are termed as fuzzy set of X if \bar{a} is a set of well-organized sets $\bar{a} = \{(x, \mu_{\bar{a}}(x)) / x \in \mathbb{R}\}$, wherever $\mu_{\bar{a}}(x)$ is relationship function corresponding to \bar{a} , $\mu_{\bar{a}}(x) \in [0,1]$.

II. Definition The α -level related to a fuzzy set \bar{a} are termed for a normal set $[\bar{a}]_{\alpha}$ for whom the Maximum power of its relationship function tops the level α , $[\bar{a}]_{\alpha} = \{x \in \mathbb{R} | \mu_{\bar{a}}(x) \geq \alpha\}$.

III. Definition Support for fuzzy set $\bar{a} \in \mathbb{R}$ for $\mu_{\bar{a}}(x)$ are positive, $\text{supp } \bar{a} = \{x \in \mathbb{R} | \mu_{\bar{a}}(x) > 0\}$.

IV. Definition Fuzzy set \bar{a} will be referred to as convex only in the case of $x, y \in \mathbb{R}$, $\rho \in [0,1]$, $\mu_{\bar{a}}(\rho x + (1 - \rho)y) \geq \min\{\mu_{\bar{a}}(x), \mu_{\bar{a}}(y)\}$.

V. Definition A LR-type fixed fuzzy number [1] can be represented by $\bar{a} = (a^L, a^U, \alpha, \beta)_{LR}$, only for

$$\mu_{\bar{a}}(x) = \begin{cases} L\left(\frac{a^L - x}{\alpha}\right) & \text{for } a^L - \alpha \leq x \leq a^L \\ 1 & \text{for } a^L \leq x \leq a^U \\ R\left(\frac{x - a^L}{\beta}\right) & \text{for } a^U \leq x \leq a^U + \beta \\ 0 & \text{else} \end{cases} \quad (1)$$

$$\mu_{\bar{a}}(x) = \begin{cases} \frac{x-(a^L-\alpha)}{\alpha} & \text{for } a^L - \alpha \leq x \leq a^L \\ 1 & \text{for } a^L \leq x \leq a^U \\ \frac{(a^L-\beta)-x}{\beta} & \text{for } a^U \leq x \leq a^U + \beta \\ 0 & \text{else} \end{cases} \quad (2)$$

I. Remark in case of $a^L = a^U = a$ in TRFN $\bar{a} = (a^L, a^U, \alpha, \beta)_{LR}$, a three-cornered fuzzy number (TNF) can be accomplished, and denoted by $\bar{a} = (a, \alpha, \beta)$.

Let $\bar{a} = (a^L, a^U, \alpha, \beta)$ and $\bar{b} = (b^L, b^U, \gamma, \theta)$

Define,

$$\begin{aligned} x > 0, x \in \mathbb{R}; \quad x\bar{a} &= (xa^L, xa^U, x\alpha, x\beta) \\ x < 0, x \in \mathbb{R}; \quad x\bar{a} &= (xa^U, xa^L, -x\beta, -x\alpha) \\ \bar{a} + \bar{b} &= (a^L + b^L, a^U + b^U, \alpha + \gamma, \beta + \theta) \end{aligned}$$

I. Ranking functions

When faced with a fuzzy linear programming problem, one effective strategy is to make use of ranking functions that convey the impression of contrast between fuzzy numbers [4-7]). It is possible to take a functional approach to the set of $F_{\mathbb{R}}$ components by constructing a ranking function $\mathbb{R}: F_{\mathbb{R}} \rightarrow \mathbb{R}$

$$\bar{a} \geq \bar{b} \text{ if and only if } \mathbb{R}(\bar{a}) \geq \mathbb{R}(\bar{b}) \quad (3)$$

$$\bar{a} > \bar{b} \text{ if and only if } \mathbb{R}(\bar{a}) > \mathbb{R}(\bar{b}) \quad (4)$$

$$\bar{a} \cong \bar{b} \text{ if and only if } \mathbb{R}(\bar{a}) = \mathbb{R}(\bar{b}) \quad (5)$$

$$\bar{a} \text{ and } \bar{b} \in F_{\mathbb{R}}, \bar{a} \leq \bar{b} \leftrightarrow \bar{b} \leq \bar{a}$$

$$\mathbb{R}(k\bar{a} + \bar{b}) = k\mathbb{R}(\bar{a}) + \mathbb{R}(\bar{b}) \quad (6)$$

$$\bar{a} \text{ and } \bar{b} \in F_{\mathbb{R}}, k \in \mathbb{R}.$$

$$\mathbb{R}(\bar{a}) = c_L a^L + c_U a^U + c_{\alpha} \alpha + c_{\beta} \beta \quad (7)$$

Wherever $\bar{a} = (a^L, a^U, \alpha, \beta)$, and $c_L, c_U, c_{\alpha}, c_{\beta}$ is a set of permanent numbers, minimum one is non-zero. However, only the computation of a tiny portion of the available ranking functions for TFFVs for the sake of illustration. Listed below are second rank function proposed by [11].

$$\begin{aligned} Y_2(\bar{a}) &= \frac{1}{2} \int_0^1 \{inf(\bar{a})_{\alpha} + sup(\bar{a})_{\alpha}\} d\alpha \\ Y_2(\bar{a}) &= \frac{1}{2} \left\{ a^L + a^U + \frac{\beta - \alpha}{2} \right\} \end{aligned} \quad (8)$$

In their seminal work [6] the two authors developed a ranking function that looks like this:

$$\begin{aligned} CM_1^{\rho}(\bar{a}) &= \int_0^1 \{\rho inf(\bar{a})_{\alpha} + (1 - \rho) sup(\bar{a})_{\alpha}\} d\alpha \\ CM_1^{\rho}(\bar{a}) &= a^L + \rho \left\{ (a^U - a^L) + \frac{\alpha + \beta}{2} \right\} - \frac{\alpha}{2} \end{aligned} \quad (9)$$

This is a representation of a 2nd rank function proposed by [6]

$$\begin{aligned} CM_2^{\rho}(\bar{a}) &= \int_0^1 \alpha \{\rho inf(\bar{a})_{\alpha} + (1 - \rho) sup(\bar{a})_{\alpha}\} d\alpha \\ CM_2^{\rho}(\bar{a}) &= a^L + \rho \left\{ (a^U - a^L) + \frac{\alpha + \beta}{3} \right\} - \frac{\alpha}{3} \end{aligned} \quad (10)$$

Fuzzy numbers in the context of probability assumption were ranked by [7] To grow this, assume there are two fuzzy numbers, \bar{a}, \bar{b} . So, by expansion value given by Zadeh, crisp $x \leq y$

$$\mathbb{R}(\bar{a} \leq \bar{b}) = \sup_{x \leq y} \{ \min(\mu_{\bar{a}}(x), \mu_{\bar{b}}(y)) \}$$

The truthfulness worth $\otimes(\bar{a} \leq \bar{b})$ which is additionally termed as ranking the option as the domination of term \bar{b} on \bar{a} and is symbolized as $P(\bar{a} \leq \bar{b})$. Next, describe $\bar{a} \leq \bar{b} \leftrightarrow P(\bar{a} \leq \bar{b}) \leq P(\bar{a} \leq \bar{a})$.

Conclusion: When ranking fuzzy variables, we adopt the first class since any FVLP illustrations can be transformed into a crisp linear programming illustration by employing a linear ranking algorithm. For FLP resolution, we can also use existing dual simplex techniques, which we can apply in this context.

III. Linear Programming with Fuzzy Variables

I. Definition According to research [5] following descriptions of FVLPs are correct:

$$\begin{aligned} \text{Min } \bar{z} &\cong c\bar{x} \\ \text{s.t. } A\bar{x} &\geq \bar{b} \\ \bar{x} &\geq 0 \\ c \in \mathbb{R}^n, \bar{x} &\in (F(\mathbb{R}))^n, A \in \mathbb{R}^{m \times n}, \bar{b} \in (F(\mathbb{R}))^m. \end{aligned} \tag{11}$$

II. Definition: A vector $\bar{x} \in (F(\mathbb{R}))^n$ provides practical result for Equivalence (11) $\leftrightarrow \bar{x}$ fulfils the problem's restraints.

III. Definition: An Variable result \bar{x}_* such as optimum result of Equation (5) only for, \forall such as possible results \bar{x} of the Equation (11), $c\bar{x}_* \leq c\bar{x}$.

IV. Definition: The FVLP problem:

$$\begin{aligned} \text{Min } \bar{z} &= c\bar{x} \\ \text{s.t. } A\bar{x} &= \bar{b} \\ \bar{x} &\geq 0 \end{aligned} \tag{12}$$

where the problem's limitations are those specified by the (6)

V. Definition: Let $A = [a_{(i,j)}]_{m \times n}$ Assume $rank(A) = m$. Divider A as $[B \ N]$, where $B, m \times n$ is Plural. Clearly stated that $rank(B) = m$. assume y_j as result of $By = a_j$

$$\bar{x}_B = (x_{B_1}, x_{B_2}, \dots, x_{B_m})' \cong B^{-1}\bar{b}, \bar{x}_N = \bar{0} \tag{13}$$

is a solution of $A\bar{x} = \bar{b}$. $\bar{X} = (\bar{x}'_B, \bar{x}'_N)'$. If $\bar{x}_B \geq \bar{0}$, $\bar{z} = c_B\bar{x}_B$, where $C_B = (C_{B_1}, C_{B_2}, \dots, C_{B_m})$ The need for conformity with non-elementary variables has $\bar{x}_j, 1 \leq j \leq n, j \neq B_i, i = 1, 2, \dots, m$,

$$z_j = C_B y_j = C_B B^{-1} a_j.$$

In this section, we detail some of the most salient findings concerning the development of a workable solution, the existence of unbounded conditions, and the presence of optimality situations [5]

I. Theorem in case we obtain an fuzzy elementary viable result related with an fuzzy objective value $\bar{z} \ni z_k - c_k > 0$ for some non-primary variable \bar{x}_k and $y_k \leq 0$, then it is probable to attain an original fuzzy elementary feasible result consisting of an initial fuzzy objective value \bar{z} and $\bar{z} \leq \bar{z}$.

II. Theorem A result for illustration with a fuzzy elementary feasible is $z_k - c_k < 0$ with a few complex variable quantity \bar{x}_k and $y_k \leq 0$, then the solution to the illustration (Equation 10) is limitless.

III. Theorem Optimality restrictions. If an elementary solution $\bar{x}_B = B^{-1}\bar{b}, \bar{x}_N = \bar{0}$ are possible for Equation (10), $z_j - c_j$ for the whole of $j, 1 \leq j \leq n$

I. Duality

VI. Definition At (Mahdavi-Amiri and Nasser, 2007), the FVLP illustrations' dual (equation 9) is described in the following,

$$\begin{aligned} \text{Max } \bar{z} &= w\bar{b} \\ \text{s.t. } wA &\leq c \end{aligned} \tag{14}$$

$w \geq 0$
 $w = (w_1, w_2, \dots, w_m) \in \mathbb{R}^n$, includes a crisp variable quantity conforming to the restrictions of Equation (9).

Without going into proofs, we will list several important results about the FVLP problem and its dual, the DFVLP problem, as proven by [4]

IV. Theorem If x_0, w_0 act as practical results for FVLP, DFVLP illustrations, correspondingly, $\bar{c} x_0 \geq w_0 \bar{b}$.

I. Corollary If x_0, w_0 act as practical results for FVLP, DFVLP illustrations, correspondingly, and $\bar{c} x_0 \cong w_0 \bar{b}$, x_0 and w_0 are optimal solutions.

II. Corollary If any DFVLP problems are boundless, therefore, additional issue has zero practical result.

V. Theorem In case of FVLP, DFVLP illustrations can be solved optimally, then the two problems can be solved optimally.

IV. A Fuzzy Dual Simplex Method

I. Primal optimality and dual feasibility

$$\begin{aligned} & \text{FNLP} \\ & \text{Min } \bar{z} \cong \bar{c}x \\ & \text{Subject to restrictions } \bar{A}x \geq \bar{b} \\ & x \geq 0 \end{aligned} \tag{15}$$

Where $A = \gamma(\bar{a}_{(i,j)})$ and $b = \gamma(\bar{b}) = \gamma(\bar{b}_i)$, All $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.
 nonnegative slack variables s_i is $i = 1, 2, \dots, m$ We can reformat (16) as follows to account for the i th constraint.

$$\begin{aligned} & \text{Min } \bar{z} \cong \bar{c}x + \bar{0}s \\ & \text{Subject to restrictions } Ax - s_i = b \\ & x \geq 0, s \geq 0 \end{aligned} \tag{16}$$

Where $s = (s_1, s_2, \dots, s_m)'$

Define $\bar{x} \in \mathbb{R}^{n+m}$ and $\bar{c} \in (F(\mathbb{R}))^{n+m}$ as

$$\bar{x}_j = \begin{cases} x_j, & j=1,2,\dots,n \\ s_{j-n}, & j = n + 1, \dots, n + m, \end{cases} \tag{17}$$

$$\bar{c}_j = \begin{cases} \bar{c}_j, & j=1,2,\dots,n \\ \bar{0}, & j = n + 1, \dots, n + m, \end{cases} \tag{18}$$

Let's pretend that there is a simple answer to (17): $\bar{x}_B = B^{-1}b$, let $z_j = \bar{c}_B B^{-1}a_j$ where $\bar{c}_B = (\gamma(\bar{c}))_B = (\gamma(\bar{c}))_{B_1}, \dots, (\gamma(\bar{c}))_{B_m}$. and a_j as j^{th} line as constant matrix $[A, I] = [\gamma(\bar{A}) \quad I]$.

Table 4.1, wherever $(\bar{x}_B)_r$ as $x^{r^{\text{th}}}$ elementary variable, $y_j = B^{-1}a_j$, and $\gamma(\bar{z}_j - \bar{c}_j) = \gamma(\bar{c}_B)B^{-1}a_j - \gamma(\bar{c}_j) = \gamma(\bar{c}_B)y_j - \gamma(\bar{c}_j)$ is the real number consistent to $\bar{z}_j - \bar{c}_j \cong \bar{c}_B y_j - \bar{c}_j$. Supposing as $j = 1, \dots, n + m$, thus, we get $\bar{z}_j - \bar{c}_j \leq 0$ $w = \bar{c}_B B^{-1}$, where $w = (w_1, w_1, \dots, w_m)$. for $j = 1, \dots, n + m$

we have $\bar{z}_j - \bar{c}_j \leq 0$ for $j = 1, \dots, n$, $w a_j - c_j \leq 0$. then, $w \geq 0$ $\bar{z}_j - \bar{c}_j \leq 0$ for $j = 1, \dots, n + m \rightarrow wA \leq c$ and $w \geq 0$, where $w = \bar{c}_B B^{-1}$. the simple FVLP problem.

$$\bar{c}\bar{x} = \bar{c}_B \bar{b} = \bar{c}_B B^{-1} \bar{b} = w\bar{b} = w\gamma(\bar{b})$$

II. Corollary proves x and was optimal results as the FVL, DFVLP illustrations, correspondingly.

Table 1: An example of a fuzzy dual practical simplex tableau

Basis	\bar{x}_1	...	\bar{x}_k	...	\bar{x}_{n+m}	$\gamma(R.H.S)$
$\gamma(\bar{z})$	$\gamma(\bar{z}_1 - \bar{c}_1)$...	$\gamma(\bar{z}_k - \bar{c}_k)$...	$\gamma(\bar{z}_{n+m} - \bar{c}_{n+m})$	$\gamma(\bar{c}_B b)$
\bar{z}	$(\bar{z}_1 - \bar{c}_1)$...	$(\bar{z}_k - \bar{c}_k)$...	$(\bar{z}_{n+m} - \bar{c}_{n+m})$	$(\bar{c}_B b)$
$(\bar{x}_B)_1$	$y_{(1,1)}$...	$y_{(1,k)}$...	$y_{(1,n+m)}$	$\gamma(\bar{b}_1)$
\vdots	\vdots					\vdots
$(\bar{x}_B)_r$	$y_{(r,1)}$	$y_{(r,k)}$	$y_{(r,n+m)}$	$\gamma(\bar{b}_2)$
\vdots	\vdots					\vdots
$(\bar{x}_B)_m$	$y_{(m,1)}$	$y_{(m,k)}$	$y_{(m,n+m)}$	$\gamma(\bar{b}_m)$

I. Lemma If there is a r such that in a fuzzy dual viable simplex tableau, then.

$\gamma(\bar{b}_r) < 0$, as well as a non-primitive index number j exists $\exists y_{(r,j)} < 0$, the objective value corresponding to fuzzy dual feasible tableau will not decrease when the turning column k is set to $y_{(r,k)}$.

Proof. A criteria for picking an complex variable for the input the base is required to guarantee for news fuzzy simplex tableau retains dual viability, that the newest objective value is not reducing. Let's pretend, for the moment, that column k is the crucial one. We obtain the following new zero rows after turning on the pivot $y_{(r,k)}$:

$$(\bar{z}_j - \bar{c}_j)_{New} = (\bar{z}_j - \bar{c}_j) - \frac{(\bar{z}_k - \bar{c}_k)}{y_{(r,k)}} y_{(r,j)}, j = 1, \dots, n + m, j \neq B_i \quad (19)$$

$$(\bar{z}_j - \bar{c}_j)_{New} \leq \bar{0}, j \neq B_i \quad (20)$$

Which, using (19), results in

$$\frac{(\bar{z}_k - \bar{c}_k)}{y_{(r,k)}} \leq \frac{(\bar{z}_k - \bar{c}_k)}{y_{(r,j)}}, \text{ for all } j \neq B_i \quad (21)$$

To satisfy (20), it is sufficient to let.

$$\frac{\gamma(\bar{z}_k - \bar{c}_k)}{y_{(r,k)}} = \min_{j \neq B_i} \left\{ \frac{\gamma(\bar{z}_k - \bar{c}_k)}{y_{(r,j)}} \mid y_{(r,j)} < 0 \right\} \quad (22)$$

$$(w\gamma(\bar{b}))_{New} = w\gamma(\bar{b}) - \frac{\gamma(\bar{z}_k - \bar{c}_k)}{y_{(r,k)}} \gamma(\bar{b}_r) \geq w\gamma(\bar{b}),$$

$$\frac{\gamma(\bar{z}_k - \bar{c}_k)}{y_{(r,k)}} \gamma(\bar{b}_r) \leq 0.$$

I. Algorithm The fuzzy dual simplex method

- Assigned a base B to FNLP illustration for $\bar{z}_j - \bar{c}_j \leq \bar{0} \forall j$
- If $\bar{b} \geq 0$, if the current solution is best, then do nothing; if not, pick row r from the pivot table using $\bar{b}_r < 0$ (that is, r so that $(\bar{b}_r) < 0$).
- If $y_{(r,j)} \geq 0$ for all j , if the minimal ratio test fails (the FNLP problem is infeasible), then stop; otherwise, choose column k as the pivot.

$$\frac{\gamma(\bar{z}_k - \bar{c}_k)}{y_{(r,k)}} = \min_{j \in B_i} \left\{ \frac{\gamma(\bar{z}_k - \bar{c}_k)}{y_{(r,j)}} \mid y_{(r,j)} < 0 \right\}.$$

- Pivot on $y_{(r,k)}$ and proceed to (2).

I. Remark In (2), r so that is one possible proposition for a special of r .

$$\gamma(\bar{b}_r) = \min_{1 \leq i \leq m} \{ \gamma(\bar{b}_i) \}.$$

I. Numerical example

We examine the seeing instance, which may be explained using Maleki's method, to further our understanding of the aforesaid strategy [5, 7]. Please take note that the linear ranking function on $F(\mathbb{R})$ is modelled after the one proposed by Yager in this section.

$$\begin{aligned} \text{Minimum } z &\cong (1,3,1,1)x_1 + (2,4,1,1)x_2 + (3,5,1,1)x_3 \\ &\quad \left(0, 2, \frac{1}{2}, \frac{1}{2} \right) x_1 + (1,3,1,1)x_2 + \left(0, 2, \frac{1}{2}, \frac{1}{2} \right) x_3 \geq (2,4,1,1) \end{aligned}$$

Subject to restrictions

$$\left(1, 3, 1, 1 \right) x_1 - \left(0, 2, \frac{1}{2}, \frac{1}{2} \right) x_2 + (2,4,1,1)x_3 \geq (3,5,1,1)$$

$$\left(\begin{array}{l} x_1, x_2, x_3 \geq 0 \end{array} \right)$$

The fuzzy coefficient matrix is ranked using the Yager function \bar{A} as well as fuzzy right-hand-cross vector \bar{b} . Therefore, establishing slack of variable quantity x_4 and x_5 the problem decreases to:

$$\text{Minimum } z \cong (1,3,1,1)x_1 + (2,4,1,1)x_2 + (3,5,1,1)x_3$$

$$\text{Subject to restrictions } \left\{ \begin{array}{l} x_1 + 2x_2 + x_3 - x_4 = 3 \\ 2x_1 - x_2 + 3x_3 - x_5 = 4 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{array} \right.$$

Now, x_1 is an incoming variable, and x_5 is an exit variable quantity. Then by rotating on $y_{21} = -2$, we attain the next tableau as

Table 2: The first iteration

Basic	x_1	x_2	x_3	x_4	x_5	R.H. S
$\gamma(\bar{z})$	0	-4	-1	0	-1	4
\bar{z}	$\bar{0}$	$(-4, -4, \frac{3}{2}, \frac{3}{2})$	$(-3, 1, 2, 2)$	$\bar{0}$	$(-1, -1, 0, 0)$	$(0, 0, 0, 0)$
x_4	0	$-\frac{5}{2}$	$\frac{1}{2}$	1	$-\frac{1}{2}$	-1
x_1	1	$-\frac{1}{2}$	$\frac{3}{2}$	0	$-\frac{1}{2}$	2

Table 3: The second iteration

Basic	x_1	x_2	x_3	x_4	x_5	R.H.S
$\gamma(\bar{z})$	0	0	$-\frac{9}{5}$	$-\frac{8}{5}$	$-\frac{1}{5}$	$\frac{28}{5}$
\bar{z}	$\bar{0}$	$\bar{0}$	$(-\frac{31}{5}, \frac{13}{5}, \frac{18}{5}, \frac{18}{5})$	$(-\frac{8}{5}, -\frac{8}{5}, 0, 0)$	$(-\frac{1}{5}, -\frac{1}{5}, 0, 0)$	$(\frac{11}{5}, 9, \frac{17}{5}, \frac{17}{5})$
x_4	0	1	$-\frac{1}{5}$	$-\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$
x_1	1	0	$\frac{7}{5}$	$-\frac{1}{5}$	$-\frac{2}{5}$	$\frac{11}{5}$

To clarify, x_2 , an input variable x_4 as an output variable in this scenario. Updated data can be seen in Table 3 below.

Hence the optimum solution attained by the fuzzy dual simplex is $x_1 = \frac{11}{5}$, $x_2 = \frac{2}{5}$, also the fuzzy optimum estimate of the objective function is $\bar{z} \cong \left(\frac{11}{5}, 9, \frac{17}{5}, \frac{17}{5}\right)$.

V. Conclusion

In this study, we introduce the idea of a duality outcome in fuzzy variable linear programming problems using fuzzy dual simplex algorithms with generic linear ranking functions on fuzzy variables. In particular, we emphasize the proposed method for directly addressing fuzzy variable linear programming problems with the assistance of linear ranking functions. This method is applicable in linear programming and will prove useful in sensitivity analysis for basic column activity vectors. We use the fuzzy dual simplex approach to perform dual pivots and reach feasibility. Interesting studies will be possible in the future when the variables are also fuzzy integers.

References

- [1] H. Tanaka, T. Okuda, and K. Asai, On fuzzy mathematical programming, *J. Cybernetics* 3 (1974), 37-46.
- [2] H. J. Zimmermann, Fuzzy programming and linear programming with several objective functions, *Fuzzy Sets and Systems* 1 (1978), 45-55.
- [3] N. Mahdavi-Amiri et al., Duality in fuzzy number linear programming by use of a specific linear ranking function, *Appl. Math. Comput.* 180 (2006), 206-216.
- [4] N. Mahdavi-Amiri et al., Duality results and a dual simplex method for linear programming problems with trapezoidal fuzzy variables, *Fuzzy Sets and Systems* 158 (2007), 1961-1978.
- [5] N. Mahdavi-Amiri et al., Fuzzy primal simplex algorithms for solving fuzzy linear programming problems, *Iran. J. Oper. Res.* 2 (2009), 68-84.
- [6] H. R. Maleki, M. Tata and M. Mashinchi, Linear programming with fuzzy variables, *Fuzzy Sets and Systems* 109 (2000), 21-33.
- [7] P. Fortemps and M. Roubens, Ranking and defuzzification methods based on area compensation, *Fuzzy Sets and Systems* 82 (1996), 319-330.
- [8] Y. J. Lai and C. L. Hwang, *Fuzzy Mathematical Programming Methods and Applications*, Springer, Berlin, 1992.
- [9] A. Ebrahimnejad, S. H. Nasseri, F. Hosseinzadeh Lotfi and M. Soltanifar, A primal-dual method for linear programming problems with fuzzy variables, *European J. Industrial Engineering*, in press.
- [10] H. J. Zimmermann, *Fuzzy Set Theory and its Applications*, 2nd ed., Kluwer Academic Publishers, Dordrecht, 1991.
- [11] R. R. Yager, A procedure for ordering fuzzy subsets of the unit interval, *Inform. Sci.* 24 (1981), 143-161.

A MODIFIED AILAMUJIA DISTRIBUTION: PROPERTIES AND APPLICATION

DAVID Ikwuoche John, OKEKE Evelyn Nkiru and FRANKLIN Lilian

Department of Mathematics and Statistics, Federal University Wukari, Nigeria
davidij@fuwukari.edu.ng, okekeevelyn@fuwukari.edu.ng, lilianfranklin1@gmail.com
<https://orcid.org/0000-0002-7100-5357>

Abstract

This study presents a modified one-parameter Ailamujia distribution called the Entropy Transformed Ailamujia distribution (ETAD) is introduced to handle both symmetric and asymmetric lifetime data sets. The ETAD properties like order and reliability statistics, entropy, moment and moment generating function, quantile function, and its variability measures were derived. The maximum likelihood estimation (MLE) method was used in estimating the parameter of ETAD and through simulation at different sample sizes, the MLE was found to be consistent, efficient, and unbiased for estimating the ETAD parameter. The flexibility of ETAD was shown by fitting it to six different real lifetime data sets and compared it alongside seven competing one-parameter distributions. The goodness of fit (GOF) results from Akaike information criteria, Bayesian information criteria, corrected Akaike information criteria, and Hannan-Quinn information criteria show that the ETAD was the best fit amongst all the seven competing distributions across all the six data sets.

Keywords: Ailamujia distribution, Entropy Transformation, Maximum likelihood estimation, Goodness of fit, Information criteria

I. Introduction

Modelling, organizing, as well as analyzing real life phenomenon is essential in the field of applied sciences such as engineering, medical sciences, finance, architecture, amongst others. Recently, progressive hybrid censoring scheme is becoming popular in a life testing problems and reliability analysis. Since there are so many data generation processes characterized under different systems and environments, no single probability model can perfectly be used to describe and model all the phenomena. Therefore, the consistency and accuracy of statistical analysis is highly influenced by the probability distribution or model adopted {[1], [2], [3], and [4]}. One remarkable method is accomplished by adding parameter(s) to an existing/traditional distribution [5]. Also, [6] introduced a new dimension of distribution, which discussed a new family of distribution, namely the Exponentiated Exponential distribution. Kumaraswamy-G proposed by [7] while [8], [9], and [10] introduced Logistic-X, New Sine Inverse Rayleigh, and New Sine Inverted Exponential distributions, respectively. In each of the methods listed above, at least one parameter is added to the existing distribution which thus adds to the complexity in obtaining the estimates of the parameters for the new model. Considering the associated complexities, transformation techniques were recommended which do not require the addition of parameters to the baseline distribution to modify its flexibility and robustness [11]. In line with this system of modification, [12] proposed a

probability distribution called the Entropy Transformed Weibull distribution and Entropy Transformed Rayleigh distribution was derived by [13]. In the same system of transformation, this paper introduced an Entropy Transformed Ailamujia Distribution (ETAD) which is a modification of the Ailamujia Distribution introduced by [14].

II. Methods

In this section, the ETAD and its application are presented. The Probability Density Function (PDF), Cumulative Distribution Function (CDF), reliability functions, variability measures and distribution shape, measures of uncertainty, and the estimation of the ETAD parameter using the maximum likelihood estimation technique are presented. Simulation and six (6) real life data sets are fitted to the ETAD and compared to seven (7) other competing one parameter distributions using the technique of information criteria.

I. Entropy Transformed Ailamujia Distribution (ETAD)

Soleha & Sewilam [15] considered a random variable X which represents the appropriate runtime for any component, and they introduced the following expression

$$g(x) = F(x) + R(x) \ln R(x), \quad x \geq 0 \quad (1)$$

where, $F(x)$ and $R(x)$ are the cumulative and reliability function, respectively, of a positive continuous random variable X . They called $g(x)$ an "Entropy Transformation" and this possibility is because the term $R(x) \ln R(x)$ is similar to the associated entropy expression for the density function of a continuous random variable X

$$H(f) = - \int_0^{\infty} f(x) \ln f(x) dx. \quad (2)$$

Differentiating equation (1) with respect to x the PDF is obtained, its integral gives the CDF, and fixing the range validates the PDF:

$$\int_0^{\infty} g'(x) dx = \int_0^{\infty} R'(x) \ln R(x) dx = 1 \quad (3)$$

The CDF of the random variable X of Ailamujia distribution (AD) is given by [14]:

$$F(x) = 1 - (1 + 2\theta x) e^{-2\theta x}, \quad x > 0, \theta > 1 \quad (4)$$

The reliability function $R(x)$ of AD is given as

$$R(x) = 1 - F(x) = (1 + 2\theta x) e^{-2\theta x}, \quad x > 0, \theta > 1. \quad (5)$$

Substituting equation (4) and (5) into equation (1) and take the derivative w.r.t x gives

$$g(x) = 1 - (1 + 2\theta x) e^{-2\theta x} + (1 + 2\theta x) e^{-2\theta x} \ln((1 + 2\theta x) e^{-2\theta x}) \quad (6)$$

However, from equation (3) it can be easily seen that the differentiation of equation (6) can be easily obtained as

$$\begin{aligned} g'(x) &= \left(\frac{d}{dx} [(1 + 2\theta x) e^{-2\theta x}] \right) \ln((1 + 2\theta x) e^{-2\theta x}) = [2\theta e^{-2\theta x} + (1 + 2\theta x) e^{-2\theta x} (-2\theta)] \ln((1 + 2\theta x) e^{-2\theta x}) \\ &= (e^{-2\theta x} [2\theta - 2\theta - 4\theta^2 x]) \ln((1 + 2\theta x) e^{-2\theta x}) = -4\theta^2 x e^{-2\theta x} \ln((1 + 2\theta x) e^{-2\theta x}) \end{aligned} \quad (7)$$

Let $f(x)$ be $g'(x)$ then

$$f(x) = -4\theta^2 x e^{-2\theta x} \ln((1 + 2\theta x) e^{-2\theta x}) \quad x > 0, \theta > 1$$

is the PDF of the ETAD. The corresponding CDF after integrating (7) is obtained as

$$F(x) = (2\theta x + 1) e^{-2\theta x} (\ln(2\theta x + 1) - 2\theta x - 1). \quad (8)$$

The PDF and CDF plots at different values of θ are presented in Figure 1 and Figure 2, respectively.

II. Linear Representation of ETAD

The ETAD is linearly represented for simplicity using the Taylor series expansion theorem at order four (4). Expanding the \ln part of $f(x)$ and let it be represented by $t(x)$, thus

$$\frac{d}{dx} [t(x)] = t'(x) = \frac{d}{dx} \left[\ln((2\theta x + 1)e^{-2\theta x}) \right] = \frac{1}{(2\theta x + 1)e^{-2\theta x}} \frac{d}{dx} [(2\theta x + 1)e^{-2\theta x}] \quad (9)$$

$$= \frac{(2\theta e^{-2\theta x} - 2(2\theta x + 1)\theta e^{-2\theta x} \cdot 1)e^{2\theta x}}{(2\theta x + 1)e^{-2\theta x}} = \frac{(2\theta e^{-2\theta x} - 2\theta(2\theta x + 1)e^{-2\theta x})}{2\theta x + 1} \quad (10)$$

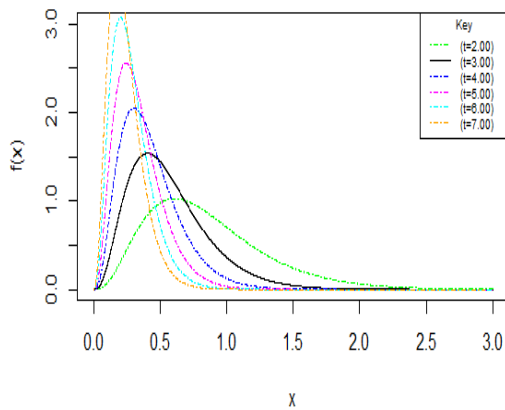


Figure 1: PDF Plots of ETAD

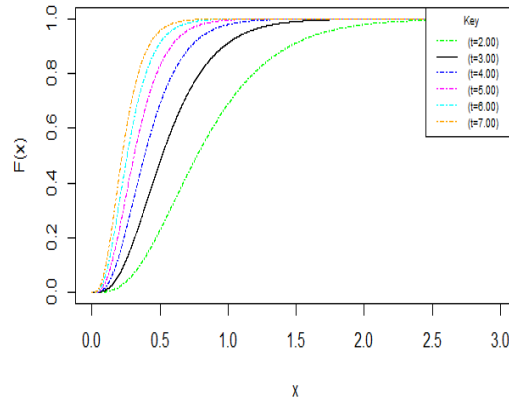


Figure 2: CDF Plots of ETAD

By expansion equation (10) becomes

$$t'(x) = -\frac{4\theta^2 x}{2\theta x + 1} = 0 \quad (11)$$

$$\frac{d^2}{dx^2} = t''(x) = \frac{d}{dx} \left[-\frac{4\theta^2 x}{2\theta x + 1} \right] = -4\theta^2 \frac{d}{dx} \left[\frac{x}{2\theta x + 1} \right] = \frac{4\theta^2 ((2\theta x + 1) - 2(\theta))}{(2\theta x + 1)^2} \quad (12)$$

By expansion equation (12) becomes

$$t''(0) = \frac{4\theta^2}{(2\theta x + 1)^2} = 4\theta^2 \quad (13)$$

$$\frac{d^3}{dx^3} = t'''(x) = \frac{d}{dx} \left[\frac{-4\theta^2}{(2\theta x + 1)^2} \right] = -4\theta^2 \frac{d}{dx} \left[\frac{1}{(2\theta x + 1)^2} \right] = -4\theta^2 (-2)(2\theta x + 1)^{-3} 2\theta \quad (14)$$

$$= \frac{16\theta^3}{(2\theta x + 1)^3} \quad (15)$$

By expansion equation (15) becomes

$$t'''(0) = 16\theta^3 \quad (16)$$

$$\frac{d^4}{dx^4} = t''''(x) = \frac{d}{dx} \left[\frac{16\theta^3}{(2\theta x + 1)^3} \right] = 16\theta^3 \frac{d}{dx} \left[\frac{1}{(2\theta x + 1)^3} \right] \quad (17)$$

$$= 16\theta^3 (-3)(2\theta x + 1)^{-4} 2\theta = -\frac{96\theta^4}{(2\theta x + 1)^4} \quad (18)$$

By expansion equation (18) becomes

$$t''''(\theta) = 96\theta^4 \tag{19}$$

Therefore,

$$t(0) = \frac{0x^0}{0!} + \frac{0x^1}{1!!} - 4\theta^2 \frac{x^2}{2!} + 16\theta^3 \frac{x^3}{3!} - 96\theta^4 \frac{x^4}{4!} + \dots$$

$$\sum_{n=2}^{\infty} t(0) = -2\theta^2 x^2 + 2 \cdot 67\theta^3 x^3 - 4\theta^4 x^4 + \dots \tag{20}$$

The approximate and n th series expansion of (20) are given respectively as

$$\sum_{n=2}^{\infty} t(0) = -2\theta^2 x^2 + 3\theta^3 x^3 - 4\theta^4 x^4 + \dots$$

and

$$\sum_{n=2}^{\infty} t(0) = \sum_{n=2}^{\infty} (2)(1.5)^{(n-2)} (-1)^{(n+1)} \theta^n x^n .$$

Therefore, the linear representation of ETAD is given as

$$f(x) = -4\theta^2 x e^{-2\theta x} \left[\sum_{n=2}^{\infty} (2)(1.5)^{(n-2)} (-1)^{(n+1)} \theta^n x^n \right] = -4\theta^2 x e^{-2\theta x} [G_n \theta^n x^n] = -4\theta^{2+n} G_n x^{1+n} e^{-2\theta x} \tag{21}$$

where, $G_n = \sum_{n=2}^{\infty} (2)(1.5)^{(n-2)} (-1)^{(n+1)}$.

III. Some Properties of ETAD

The survival $[S(x)]$, hazard $[h(x)]$, and odd $[O(x)]$ functions are presented as well as their respective plots follows.

$$S(x) = 1 - F(X \leq x) = 1 - (2\theta x + 1)e^{-2\theta x} (\ln(2\theta x + 1) - 2\theta x - 1) \tag{22}$$

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)} = \frac{-4\theta^2 x [\ln(2\theta x + 1)e^{-2\theta x}]}{1 - (2\theta x + 1)e^{-2\theta x} (\ln(2\theta x + 1) - 2\theta x - 1)} \tag{23}$$

$$\text{Odds function: } O(x) = \frac{F(x)}{1 - F(x)} = \frac{F(x)}{S(x)} = \frac{(2\theta x + 1)e^{-2\theta x} (\ln(2\theta x + 1) - 2\theta x - 1)}{1 - (2\theta x + 1)e^{-2\theta x} (\ln(2\theta x + 1) - 2\theta x - 1)} \tag{24}$$

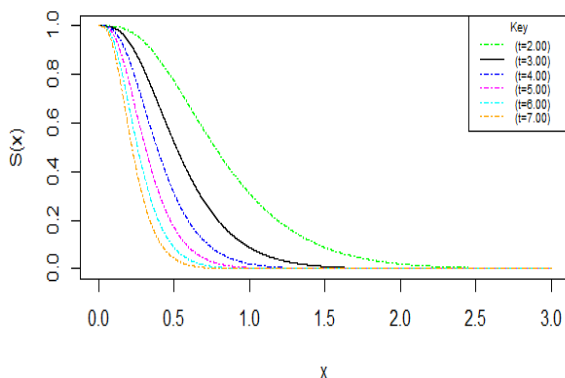


Figure 3: Plot of $S(x)$ at Different Parameter Values

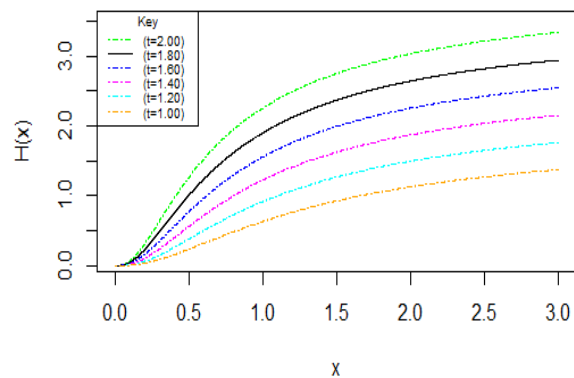


Figure 4: Plot of $h(x)$ at Different Parameter Values

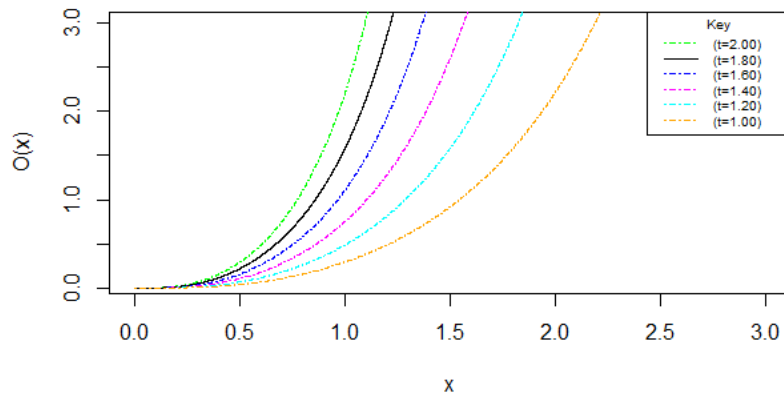


Figure 5: Plot of $O(x)$ at Different Parameter Values

Quantile Function: the inverse of the CDF of ETAD gives the quantile function of ETAD, through the Lambert W function transformation it is expressed as

$$x = \frac{-W_{-1} \left[-10W_0 \left(\frac{u}{10} \right) e^{-1} \right] - 1}{2\theta} \quad (25)$$

IV. Moments and Moment Generating Function

Moments: This is an important property of a distribution that is used to getting some measures comprising of the mean, variance, etc. Suppose the random variable $X \sim \text{ETAD}(x, \theta)$, the r^{th} moment (μ_r), can be obtained as

$$\mu^r = E(x^r) = \int_0^{\infty} x^r f(x, \theta) dx \quad (26)$$

$$f(x) = -4\theta^{2+n} G_n x^{1+n} e^{-2\theta x}$$

$$\mu_r = \int_0^{\infty} x^r \cdot -4\theta^{2+n} x^{n+1} G_n e^{-2\theta x} dx = -4\theta^{2+n} G_n \int_0^{\infty} x^{n+1+r} e^{-2\theta x} dx$$

Let $u = 2\theta x$; $\frac{du}{dx} = 2\theta \Rightarrow du = \frac{1}{2\theta} du$; $dx = \frac{du}{2\theta}$. Where $x = 0, u = 2\theta \times 0 = 0$; As $x \rightarrow \infty, u \rightarrow \infty$.

By substitution the following is obtained

$$\begin{aligned} \mu_r &= -4\theta^{2+n} G_n \int_0^{\infty} \left(\frac{u}{2\theta} \right)^{n+1+r} e^{-u} \frac{du}{2\theta} = -4G_n \int_0^{\infty} \theta^{2+n} (2\theta)^{-(n+1+r+1)} u^{n+1+1} e^{-u} du \\ &= -4G_n \theta^{2+n} (2\theta)^{-(n+1+r+2)} \int_0^{\infty} u^{n+r+1} e^{-u} du = -4G_n \theta^{2+n} (2\theta)^{-(n+1+r+2)} \int_0^{\infty} u^{n+r+1+1-1} e^{-u} du \end{aligned}$$

$$\mu_r = -4G_n \theta^{2+n} [2\theta]^{-(n+r+2)} \Gamma(n+r+2). \quad (27)$$

The *mean* of ETAD, that is, $E[x]$ is given as

$$\mu_r = -4G_n \theta^{2+n} [2\theta]^{-(n+r+2)} \Gamma(n+r+2)$$

$$E(x) = \mu_1 = -4G_n \theta^{2+n} [2\theta]^{-(n+1+2)} \Gamma(n+1+2) = - \left[\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \right] \quad (28)$$

The population variance, σ^2 of ETAD can be obtained as follows,

$$Var(x) = E[x^2] - [E(x)]^2 \tag{29}$$

$$\begin{aligned} E[x]^2 &= \mu_2 = -4G_n \theta^{2+n} [2\theta]^{-(n+4)} \\ &= -4G_n \theta^{2+n} 2^{-(n+4)} \theta^{-(n+4)} \Gamma(n+4) = -2^2 \cdot 2^{-(n+4)} G_n \theta^{2+n} \theta^{-(n+4)} \Gamma(n+4) \\ &= -\left[2^{-(n+2)} G_n \theta^{-2} \Gamma(n+4) \right] = -\left[\frac{G_n \Gamma(n+4)}{2^{(n+2)} \theta^2} \right] = -\frac{G_n \Gamma(n+4)}{2^{(n+2)} \theta^2} - \left[\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \right]^2 \\ Var(x) &= \frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} - \frac{G_n \Gamma(n+4)}{2^{(n+2)} \theta^2} = \frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \left[1 - \frac{(n+3)}{2\theta} \right] \end{aligned} \tag{30}$$

The Coefficient of Variance (CV) of ETAD can be obtained as follows

$$CV = \frac{\sigma}{\mu} 100\% = \frac{\sqrt{\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \left[1 - \frac{(n+3)}{2\theta} \right]}}{\frac{-G_n \Gamma(n+3)}{2^{(n+1)} \theta}} = \left[\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \left[1 - \frac{(n+3)}{2\theta} \right] \right]^{\frac{1}{2}} \frac{(-2^{(n+1)} \theta)}{G_n \Gamma(n+3)} \tag{31}$$

$$\begin{aligned} &= \left[\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \right]^{\frac{1}{2}} \left[\frac{-2^{(n+1)} \theta}{G_n \Gamma(n+3)} \right] \left[1 - \frac{(n+3)}{2\theta} \right]^{\frac{1}{2}} = -[G_n \Gamma(n+3)]^{\frac{1}{2}} [2^{(n+1)} \theta] [G_n \Gamma(n+3)]^{-1} \left[1 - \frac{(n+3)}{2\theta} \right]^{\frac{1}{2}} \\ &= -[G_n \Gamma(n+3)]^{\frac{1}{2}} [2^{(n+1)} \theta]^{\frac{1}{2}} \left[1 - \frac{(n+3)}{2\theta} \right]^{\frac{1}{2}} = -\left[\frac{2^{(n+1)} \theta}{G_n \Gamma(n+3)} \right]^{\frac{1}{2}} \left[1 - \frac{(n+3)}{2\theta} \right]^{\frac{1}{2}} \\ CV &= -\left[\frac{2^{(n+1)} \theta}{G_n \Gamma(n+3)} \left[1 - \frac{(n+3)}{2\theta} \right] \right]^{\frac{1}{2}} \times 100\% \end{aligned} \tag{32}$$

The Harmonic Mean (HM) of ETAD is obtained as follows

$$HM(X) = \int_0^\infty \frac{1}{x} f(x) dx \tag{33}$$

$$\begin{aligned} &= \int_0^\infty \frac{1}{x} (-4G_n \theta^{2+n} x^{1+n} e^{-2\theta x}) dx = -4G_n \theta^{2+n} \int_0^\infty x^{-n} x^{1+n} e^{-2\theta x} dx = -4G_n \theta^{2+n} \int_0^\infty x^n e^{-2\theta x} dx \\ &= -4G_n \theta^{2+n} \left[\frac{n!}{(2\theta)^{n+1}} \right] = -\frac{4G_n \theta^{2+n} n!}{2^{n+1} \theta^{n+1}} = -2^2 \cdot 2^{-(n+1)} G_n \theta^{2+n} \theta^{-(n+1)} n! = 2^{(1-n)} G_n \theta n! \end{aligned} \tag{34}$$

The Mode (M) of ETAD can be obtained as follows

$$M(x) = \frac{d}{dx} \ln(f(x, \theta)) \tag{35}$$

$$\begin{aligned} &= \frac{d}{dx} \ln[-4G_n \theta^{2+n} x^{1+n} e^{-2\theta x}] = \frac{1}{-4G_n \theta^{2+n} x^{1+n} e^{-2\theta x}} \frac{d}{dx} [-4G_n \theta^{2+n} x^{1+n} e^{-2\theta x}] \\ &= \frac{\left[-4G_n \theta^{2+n} \frac{d}{dx} [x^{1+n} e^{-2\theta x}] \right]}{-4G_n \theta^{2+n} x^{1+n} e^{-2\theta x}} = \frac{(n+1)x^n e^{-2\theta x} + x^{n+1} \cdot e^{-2\theta x} (-2\theta)}{x^{n+1} e^{-2\theta x}} \\ &= \frac{(n+1)x^n e^{-2\theta x} - 2x^{n+1} \cdot e^{-2\theta x} \theta}{x^{n+1} e^{-2\theta x}} = \frac{x^n e^{-2\theta x} [(n+1) - 2x\theta]}{x^n \cdot x \cdot e^{-2\theta x}} = \frac{(n+1) - 2\theta x}{x} = [(n+1) - 2\theta x] x^{-1} \\ M(x) &= (n+1)x^{-1} - 2\theta \end{aligned} \tag{36}$$

The *Median* can be derived by using equation (25) which is the derived quantile function of ETAD and fixing $u = 0.5$, the median of ETAD is obtained as follows:

$$x = \frac{-W_{-1} \left[-10W_0 \left(\frac{0.5}{10} \right) e^{-1} \right] - 1}{2\theta} = \frac{-W_{-1} \left[-10W_0 (0.05) e^{-1} \right] - 1}{2\theta} \quad (37)$$

Skewness [SK]: The skewness of a probability distribution is a measure of symmetry and the lack of symmetry of the probability distribution. The ETAD SK is derived from the 3rd moment of the mean.

$$SK = \frac{E(x^3) - 3\mu\sigma^2 - \mu^3}{\sigma^3} \quad (38)$$

$$SK = \frac{2^{\left[\frac{3n+7}{2} \right]} \left[\frac{G_n \Gamma n + 5}{2^{(n+2)}} + \frac{3[G_n \Gamma n + 3]^2 [2(n+3) + G_n \Gamma n + 3]}{2^{2n+3}} - \frac{[G_n \Gamma n + 3]^3}{2^{3n+2}} \right]}{\sqrt{i [G_n \Gamma n + 3 [2(n+3) + G_n \Gamma n + 3]]^{\frac{3}{2}}}} \quad (39)$$

Kurtosis [KT]: The kurtosis measure whether or not the probability distribution is heavy-tailed. For ETAD the KT is derived from the 4th moment and is obtained as

$$KT = \frac{E(x^4) - 4\mu E[x^3] + 6\sigma^2 E[x^2] - 3E[x^4]}{\sigma^4} \quad (40)$$

$$KT = \frac{3G_n \Gamma(n+4) [2(n+3) + G_n \Gamma(n+3)] - 4G_n \Gamma(n+5) + 2^{(n+2)} (n+5)(n+4)(n+3)}{G_n \Gamma(n+3) [2(n+3) + G_n \Gamma(n+3)]^2} \quad (41)$$

Dispersion Index [DI]: It is simply defined as the variance divided by the mean. It tells how dispersed the mean is from the variance. The DI for the ETAD is obtained as

$$DI = \frac{\sigma^2}{\mu} = \frac{\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta} \left[1 - \frac{(n+3)}{2\theta} \right]}{\frac{G_n \Gamma(n+3)}{2^{(n+1)} \theta}} = 1 - \frac{(n+3)}{2\theta} \quad (42)$$

Moment Generating Function: The moment generating function of X, say $M_x(t)$, is

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x, \theta) dx = \int_0^\infty e^{tx} \cdot -4\theta^{2+n} G_n x^{1+n} e^{-2\theta x} dx; e^{tx} = \sum_{j=0}^\infty \frac{(tx)^m}{m!} \quad (43)$$

By substitution equation (43) becomes

$$M_x(t) = \int_0^\infty \sum_{j=0}^\infty \frac{(tx)^m}{m!} \cdot -4\theta^{2+n} G_n x^{1+n} e^{-2\theta x} dx \quad (44)$$

$$= \int_0^\infty \sum_{j=0}^\infty \frac{t^m}{m!} x^m \cdot -4\theta^{2+n} G_n x^{1+n} e^{-2\theta x} dx = \phi \int_0^\infty x^{(1+n+m)} \theta^{2+n} e^{-2\theta x} dx \quad (45)$$

Let $u = 2\theta x \Rightarrow x = \frac{u}{2\theta}; \frac{d}{dx} = 2\theta \Rightarrow dx = \frac{1}{2\theta} du$. Where, $x = 0, u = 0, x \rightarrow \infty, u \rightarrow \infty$

Substituting into (45) to get

$$= \phi \int_0^\infty \left(\frac{u}{2\theta} \right)^{(1+n+m)} \theta^{2+n} e^{-u} \frac{du}{2\theta} = \phi \int_0^\infty u^{(1+n+m)} (2\theta)^{-(1+n+m)} \theta^{2+n} e^{-u} 2\theta^{-1} du$$

$$M_x(t) = \varphi(2\theta)^{-(2+n+m)} \theta^{2+n} \int_0^\infty u^{[1+n+m]-1} e^{-u} du = \phi[2\theta]^{-(2+n+m)} \theta^{2+n} \Gamma[2+n+m] \quad (46)$$

V. Entropy

Suppose a random variable X follows the ETA distribution, then the Renyi entropy (Renyi, 1961) which measures the uncertainty of information is expressed as

$$\begin{aligned} I_R(c) &= \frac{1}{1-c} \log \int_0^\infty f^c(x) dx = \frac{1}{1-c} \log \left[\int_0^\infty \left\{ -4\theta^2 x e^{-2\theta x} \left[\log(e^{-2\theta x} (2\theta x + 1)) \right] \right\}^c dx \right] \quad (47) \\ &= \frac{1}{1-c} \log \left[\int_0^\infty (-4\theta^2)^c x^c e^{-2\theta c x} \left[\log(e^{-2\theta c x} [2\theta x + 1]) \right]^c dx \right] = \frac{1}{1-c} \log \left[\int_0^\infty [-4\theta^{2+n} G_n x^{1+n} e^{-2\theta x}]^c dx \right] \\ &= \frac{1}{1-c} \log \left[\int_0^\infty [-4\theta^2 G_n]^c x^{c(1+n)} e^{-2\theta c x} dx \right] = \frac{1}{1-c} \log \left[[-4\theta^{2+n} G_n]^c \int_0^\infty x^{c(1+n)} e^{-2\theta c x} dx \right] \end{aligned}$$

Let $u = 2\theta c x$, $x = \frac{u}{2\theta c}$, $\frac{d}{dx} = 2\theta c$, $du = \frac{du}{2\theta c}$. Where $x = 0, u = 0, x \rightarrow \infty, u \rightarrow \infty$. Therefore,

$$\begin{aligned} I_R(c) &= \frac{1}{1-c} \log \left[(-4\theta^{2+n} G_n)^c \int_0^\infty \left(\frac{u}{2\theta c} \right)^{c(1+n)} e^{-u} \frac{du}{2\theta c} \right] \\ &= \frac{1}{1-c} \log \left[(-4\theta^{2+n} G_n)^c \int_0^\infty u^{c(1+n)} (2\theta c)^{-(n+1)} (2\theta c)^{-1} e^{-u} du \right] \\ &= \frac{1}{1-c} \log \left[(-4\theta^{2+n} G_n)^c (2\theta c)^{-(n+1)c-1} \int_0^\infty u^{c(1+n)} e^{-u} du \right] \\ &= \frac{1}{1-c} \log \left[(-4\theta^{2+n} G_n)^c (2\theta c)^{-(n+1)c-1} \int_0^\infty u^{c(1+n)+1-1} e^{-u} du \right] \\ &= \frac{1}{1-c} \log \left[(-4\theta^{2+n} G_n)^c (2\theta c)^{-c(n+1)-1} \Gamma(c[1+n]+1) \right] \\ &= \frac{1}{1-c} \log \left[(-4\theta^{2+n} G_n)^c \right] + \log \left[(2\theta c)^{-c(n+1)-1} \right] + \log \left[\Gamma(c(n+1)+1) \right] \quad (48) \end{aligned}$$

VI. ETAD Order statistics

Given an independence characteristic distribution of a random sample $X_1; X_2 \dots; X_n$, this sample can be in ordered form as $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ or in a better notation, $X_{(1:n)} \leq X_{(2:n)} \leq \dots \leq X_{(n:n)}$. These representations are called order statistics. The 1st order statistics, $X_{(1)}$ is the minimum, the 2nd order statistics, $X_{(2)}$ is the second while the n^{th} order statistics, $X_{(n)}$ is the maximum.

PDF of the k^{th} order statistics of ETAD: Suppose a random sample $X_1, X_2 \dots, X_n$ from the ETAD is ordered as $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, then the PDF $f_{(n,n)}(X)$ of the k^{th} order statistics can be define as:

$$f_{(k,n)}(x) = \frac{n!}{(k-1)!(n-k)!} f(x) \times F(x)^{k-1} \times [1 - F(x)]^{n-k} \quad (49)$$

where, $F(X)$ and $f(X)$ are the CDF and PDF respectively of the ETA distribution. Hence, for easier simplification, the binomial expansion of $[1 - F(x)]^{n-k}$ was used as

$$[1 - F(x)]^{n-k} = \sum_{c_5}^{\infty} \binom{n-k}{c_5} (-1)^{c_5} [F(x)]^{c_5} \quad (50)$$

Substituting equation (50) into (49) to obtain

$$f_{(k,n)}(x) = \frac{n!}{(k-1)!(n-k)!} f(x) \cdot F(x)^{k-1} \sum_{c_5}^{\infty} \binom{n-k}{c_5} (-1)^{c_5} [F(x)]^{c_5} \quad (51)$$

$$= \sum_{c_5}^{\infty} \frac{n!}{(k-1)!(n-k)!} f(x) \binom{n-k}{c_5} (-1)^{c_5} [F(x)]^{c_5} \quad (52)$$

Substituting the PDF and CDF of ETA distribution into (52) we have

$$= \sum_{c_5}^{\infty} \frac{n!(-1)^{c_5}}{(k-1)!(n-k-c_5)!c_5!} \left(-4\theta^2 x e^{-2\theta x} \ln[(2\theta x + 1)e^{-2\theta x}] \right) \binom{n-k}{c_5} (-1)^{c_5} \left[e^{-2\theta x} (2\theta x + 1) \{ \ln[(2\theta x + 1)e^{-2\theta x}] - 1 \} \right]^{c_5+k-1} \quad (53)$$

PDF of the Smallest and Largest Ordered Statistic

To obtain the 1st or minimum order statistic, let $k = 1$ be substituted into equation (53) to give

$$= \sum_{c_5}^{\infty} \frac{n!(-1)^{c_5}}{(n-1-c_5)!c_5!} \left\{ -4\theta^2 x e^{-2\theta x} \ln[(2\theta x + 1)e^{-2\theta x}] \right\} \left[e^{-2\theta x} (2\theta x + 1) \{ \ln[(2\theta x + 1)e^{-2\theta x}] - 1 \} \right]^{c_5} \quad (54)$$

Similarly, the n^{th} order statistic or maximum order statistic is obtained by substituting $k = n$, to give

$$= \sum_{c_5}^{\infty} \frac{n!(-1)^{c_5}}{(n-1)!c_5!} \left\{ -4\theta^2 x e^{-2\theta x} \ln[(2\theta x + 1)e^{-2\theta x}] \right\} \left[e^{-2\theta x} (2\theta x + 1) \{ \ln[(2\theta x + 1)e^{-2\theta x}] - 1 \} \right]^{c_5+n-1} \quad (55)$$

VII. Parameter Estimation using Maximum Likelihood Estimation

Suppose X_1, X_2, \dots, X_n is a random sample with size n drawn from ETA distribution. The likelihood function

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x; \theta) \quad (56)$$

$$= \prod_{i=1}^n \left[-4\theta^2 x e^{-2\theta x} \ln[(2\theta x + 1)e^{-2\theta x}] \right] = \left[-4\theta^2 \right]^n \prod_{i=1}^n x \prod_{i=1}^n e^{-2\theta x} \prod_{i=1}^n \left[\ln[(2\theta x + 1)e^{-2\theta x}] \right]$$

The log-likelihood for our PDF becomes

$$\begin{aligned} \log L(x_1, x_2, \dots, x_n; \theta) &= \log \left[\left(-4\theta^2 \right)^n \prod_{i=1}^n x \prod_{i=1}^n e^{-2\theta x} \prod_{i=1}^n \left[\ln[(2\theta x + 1)e^{-2\theta x}] \right] \right] = Z \\ &= n(-4\theta^2) + \sum_{i=1}^n \log(x_i) - 2\theta \sum_{i=1}^n x_i + \log \left[\prod_{i=1}^n \left[\ln[(2\theta x + 1)e^{-2\theta x}] \right] \right] \\ &= -4n\theta^2 + \sum_{i=1}^n \log(x_i) - 2\theta \sum_{i=1}^n x_i + \left[\sum_{i=1}^n \log \left[\ln[(2\theta x + 1)e^{-2\theta x}] \right] \right] \end{aligned} \quad (57)$$

The parameter estimate $\hat{\theta}$ of ETAD is obtained by differentiating equation (57) partially with respect to θ and equate to zero. This implies,

$$\frac{\partial Z}{\partial \theta} = -8n\theta - 2 \sum_{i=1}^n x_i - 4 \sum_{i=1}^n \left[\frac{\theta^2 x}{(2\theta x + 1) \ln \left[\left((2\theta x + 1) e^{-2\theta x} \right) \right]} \right] = 0 \quad (58)$$

$$\sum_{i=1}^n x_i = -4n\theta - 2 \sum_{i=1}^n \left[\frac{\theta^2 x}{(2\theta x + 1) \ln \left[\left((2\theta x + 1) e^{-2\theta x} \right) \right]} \right]$$

$$\hat{\theta} = - \left(4\bar{x}_i + \frac{1}{2n} \sum_{i=1}^n \left[\frac{\hat{\theta}^2 x_i}{(2\hat{\theta} x_i + 1) \ln \left[\left((2\hat{\theta} x_i + 1) e^{-2\hat{\theta} x_i} \right) \right]} \right] \right) \quad (59)$$

Equation (59) clearly reveal that there is no close form expression for θ and this implies it can only be estimated through numerical methods using software like SAS, R, Maple, etc., for this research the R open source software was used.

VIII. Goodness of Fit Test

The goodness of fit tests used are, Akaike IC (AIC), Bayesian IC (BIC), Corrected AIC (CAIC), and Hannan-Quinn IC (HQIC). Their respective formulas are presented as follows.

$$AIC = -2(l) + 2k \quad (60)$$

$$BIC = -2(l) + (k(\ln(n))) \quad (61)$$

$$CAIC = AIC + 2k(k+1) \div (n-k-1) \quad (62)$$

$$HQIC = -2(l) + (2 \cdot k \cdot \ln(\ln(n))) \quad (63)$$

where, l is the log-likelihood, n is the sample size, and k is the number of parameter to be estimated.

III. Results

I. Simulation Study

The usefulness of simulation is to determine the efficiency and consistency of the estimate of MLE method. Using different parameters values and sample sizes (20-1500), the estimation methods used for the comparing are based on absolute bias (AB), variance (S^2), standard error (SE), and root mean square error (RMSE).

$$AB = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\theta}_i - \theta_i) \quad (64)$$

$$RMSE = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (\hat{\theta}_i - \theta_i)^2} \quad (65)$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad (66)$$

$$SE = \frac{\sigma}{\sqrt{n}} \quad (67)$$

Table 1 presents the simulation results for up to 1500 simulations at different parameter values. The results demonstrate that the MLE method is consistent and efficient. This is due to the fact that

as the sample size increases from 20 to 1500 the variance and RMSE decreases to the minimum.

Table 1: Absolute Bias, Variance, Standard Error, and Root Mean Square Error

Sample Size	Parameter Value	Estimate	Bias	Variance	Standard Error	RMSE
20	0.1	0.1613	0.0613	0.0000	0.0000	0.0038
20	0.3	0.5012	0.2012	0.0003	0.0039	0.0408
20	0.5	0.8389	0.3389	0.0008	0.0063	0.1157
20	0.7	1.1764	0.4764	0.0015	0.0087	0.2285
40	0.1	0.1609	0.0609	0.0000	0.0000	0.0037
40	0.3	0.4911	0.1911	0.0001	0.0016	0.0366
40	0.5	0.8204	0.3204	0.0004	0.0032	0.1031
40	0.7	1.1497	0.4497	0.0008	0.0045	0.2030
80	0.1	0.1606	0.0606	0.0000	0.0000	0.0037
80	0.3	0.4859	0.1859	0.0001	0.0011	0.0347
80	0.5	0.8108	0.3108	0.0002	0.0016	0.0968
80	0.7	1.1357	0.4357	0.0004	0.0022	0.1902
100	0.1	0.1608	0.0608	0.0000	0.0000	0.0037
100	0.3	0.4856	0.1856	0.0001	0.0010	0.0345
100	0.5	0.8102	0.3102	0.0002	0.0014	0.0964
100	0.7	1.1347	0.4347	0.0003	0.0017	0.1893
150	0.1	0.1606	0.0606	0.0000	0.0000	0.0037
150	0.3	0.4839	0.1839	0.0000	0.0000	0.0338
150	0.5	0.8071	0.3071	0.0001	0.0008	0.0944
150	0.7	1.1303	0.4303	0.0002	0.0012	0.1854
250	0.1	0.1606	0.0606	0.0000	0.0000	0.0037
250	0.3	0.4831	0.1831	0.0000	0.0000	0.0335
250	0.5	0.8055	0.3055	0.0000	0.0000	0.0933
250	0.7	1.1279	0.4279	0.0001	0.0006	0.1832
500	0.1	0.1606	0.0606	0.0000	0.0000	0.0037
500	0.3	0.4824	0.1824	0.0000	0.0000	0.0333
500	0.5	0.8042	0.3042	0.0000	0.0000	0.0925
500	0.7	1.1260	0.4260	0.0000	0.0000	0.1815
750	0.1	0.1605	0.0605	0.0000	0.0000	0.0037
750	0.3	0.4820	0.1820	0.0000	0.0000	0.0331
750	0.5	0.8035	0.3035	0.0000	0.0000	0.0921
750	0.7	1.1250	0.4250	0.0000	0.0000	0.1806
1000	0.1	0.1605	0.0605	0.0000	0.0000	0.0037
1000	0.3	0.4819	0.1819	0.0000	0.0000	0.0331
1000	0.5	0.8033	0.3033	0.0000	0.0000	0.0920
1000	0.7	1.1246	0.4246	0.0000	0.0000	0.1803
1500	0.1	0.1605	0.0605	0.0000	0.0000	0.0037
1500	0.3	0.4818	0.1818	0.0000	0.0000	0.0331
1500	0.5	0.8030	0.3030	0.0000	0.0000	0.0918
1500	0.7	1.1243	0.4243	0.0000	0.0000	0.1800

II. Real Life Data Analysis

This research fitted six (6) real data sets to the ETAD and seven other competing distributions [Inverted Exponential Distribution (IED), Sine Exponential Distribution (SED), Ram Awadh Distribution (RAD), Prakaamy Distribution (PD), N-Sine Exponential Distribution (NSED), Exponential Distribution (ED), Ailamujia Distribution (AD)] for comparison. The data sets are nicotine measurement from different cigarettes brands [16], Carbon Fiber Tensile Strength of Length 20mm and 50mm [17], Survival Times of Growth Hormone Medication [18], Lung Cancer

Patients Tumours Size [19], and Glass Fiber Strength of 1.5cm [20]. The descriptive statistics, MLE, and GoF results are presented.

Table 2: Descriptive Statistics

Variable	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6
Sample size	346	69	66	35	76	63
Maximum value	2	2.585	4.9	13.7	11.18	2.24
Minimum Value	0.1	0.312	0.39	2.15	0.96	0.55
Mean	0.8525	1.451	2.751	5.298	3.529	1.507
Median	0.0041	1.49	2.035	4.51	2.7	1.54
Variance	0.1192	0.245	0.7948	8.509	6.595	0.1051

Table 3: The MLE Estimates for the Six Data Sets

Distribution	MLE 1	MLE 2	MLE 3	MLE 4	MLE 5	MLE 6
IED	0.839	1.81	1.064	1.993	0.345	1.999
SED	0.672	0.396	0.208	0.108	0.161	0.383
RAD	1.993	0.755	1.763	1.114	1.548	1.997
PD	1.927	0.401	1.879	1.117	1.589	2
NSED	1.174	0.217	0.441	0.778	1.62	1.205
ED	1.173	0.688	0.362	0.188	0.283	0.664
AD	1.985	0.689	0.362	1.665	0.283	0.664
ETAD	1.985	1.167	0.614	0.3204	0.32	1.122

Table 4: Log-likelihood and Information Criteria for Data Set I

	<i>ll</i>	AIC	BIC	CAIC	HQIC	Rank
IED	-467.823	937.646	941.492	937.657	939.177	5
SED	-273.346	548.692	552.539	548.704	550.224	3
RAD	-5076.81	10155.617	10159.464	10155.629	10157.149	6
PD	-693.204	1388.407	1392.254	1388.419	1389.939	7
NSED	-1267.3	2536.593	2540.440	2536.605	2538.125	8
ED	-290.826	583.652	587.498	583.663	585.183	4
AD	-193.354	388.708	392.554	388.720	390.240	2
ETAD	-144.56	291.119	294.965	291.130	292.650	1

Table 5: Log-likelihood and Information Criteria for Data Set II

	<i>ll</i>	AIC	BIC	CAIC	HQIC	Rank
IED	-89.1475	180.295	182.529	180.354	181.181	4
SED	-90.9535	183.907	186.141	183.967	184.793	5
RAD	-386.964	775.927	778.161	775.987	776.813	7
PD	-608.715	1219.429	1221.663	1219.488	1220.315	8
NSED	-60.315	122.630	124.864	122.690	123.517	2
ED	-94.7015	191.403	193.637	191.463	192.289	6
AD	-73.095	148.190	150.424	148.250	149.076	3
ETAD	-59.69	121.380	123.614	121.440	122.267	1

Table 6: Log-likelihood and Information Criteria for Data Set III

	<i>ll</i>	AIC	BIC	CAIC	HQIC	Rank
IED	-400.249	802.497	804.686	802.559	803.362	8
SED	-129.339	260.677	262.867	260.740	261.542	5
RAD	-104.968	211.936	214.126	211.999	212.801	4
PD	-100.044	202.088	204.278	202.151	202.953	2
NSED	-181.793	365.585	367.774	365.647	366.450	7

A MODIFIED AILAMUJIA DISTRIBUTION

ED	-132.995	267.989	270.178	268.051	268.854	6
AD	-112.004	226.008	228.197	226.070	226.873	3
ETAD	-98.697	199.394	201.584	199.457	200.259	1

Table 7: Log-likelihood and Information Criteria for Data Set IV

	<i>ll</i>	AIC	BIC	CAIC	HQIC	Rank
IED	-555.075	1112.150	1113.705	1112.271	1112.686	8
SED	-91.9045	185.809	187.364	185.930	186.346	5
RAD	-81.9755	165.951	167.506	166.072	166.487	4
PD	-81.9295	165.859	167.414	165.980	166.396	3
NSED	-307.259	616.518	618.073	616.639	617.055	2
ED	-93.4075	188.815	190.370	188.936	189.352	6
AD	-480	961.999	963.554	962.120	962.536	7
ETAD	-80.142	162.284	163.839	162.405	162.821	1

Table 8: Log-likelihood and Information Criteria for Data Set V

	<i>ll</i>	AIC	BIC	CAIC	HQIC	Rank
IED	-3948.32	7898.632	7900.963	7898.686	7899.563	8
SED	-169.826	341.651	343.982	341.705	342.582	3
RAD	-184.068	370.136	372.466	370.190	371.067	6
PD	-183.404	368.807	371.137	368.861	369.738	5
NSED	-910.383	1822.766	1825.097	1822.820	1823.698	7
ED	-171.901	345.802	348.133	345.856	346.734	4
AD	-159.121	320.241	322.572	320.295	321.173	2
ETAD	-80.142	162.284	163.839	162.405	162.821	1

Table 9: Log-likelihood and Information Criteria for Data Set VI

	<i>Ll</i>	AIC	BIC	CAIC	HQIC	Rank
IED	-70.438	142.876	145.019	142.942	143.719	3
SED	-84.9795	171.959	174.102	172.024	172.801	5
RAD	-110.009	222.017	224.160	222.082	222.860	7
PD	-107.312	216.623	218.767	216.689	217.466	6
NSED	-253.889	509.777	511.920	509.843	510.620	8
ED	-88.8305	179.661	181.804	179.726	180.504	4
AD	-66.3175	134.635	136.778	134.700	135.477	2
ETAD	-49.5355	101.071	103.214	101.136	101.914	1

From the results presented in Table 4 to Table 9 clearly shows that the ETAD is more flexible and a better fit distribution compared to the other one-parameter lifetime distributions because its *ll* values is larger than the other distributions *ll* values and it's AIC, BIC, CAIC, and HQIC values are smaller than that of the other seven competing distributions. To buttress these results, the fitted plots for the ETAD and the other competing distributions are presented below. Due to space, only very close competing distributions and our proposed distribution plots are presented in Figure 6 to Figure 8.

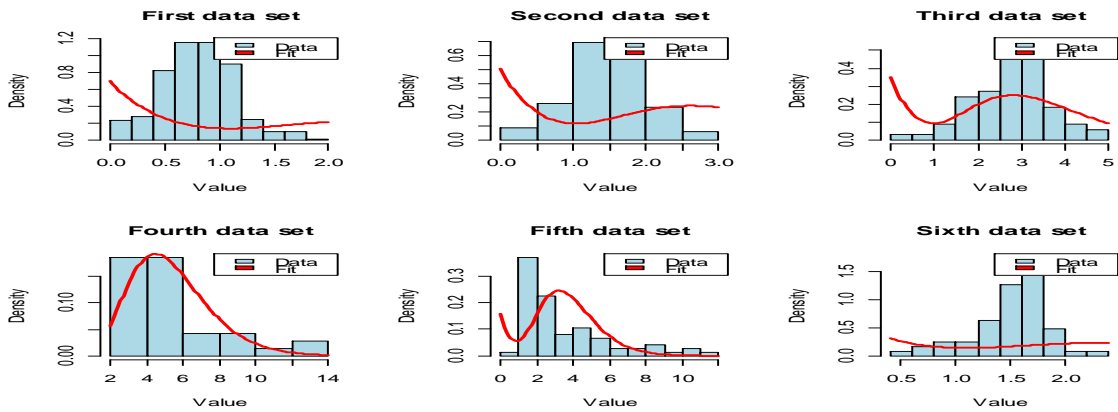


Figure 6: Density Plot for Sine Exponential Distribution

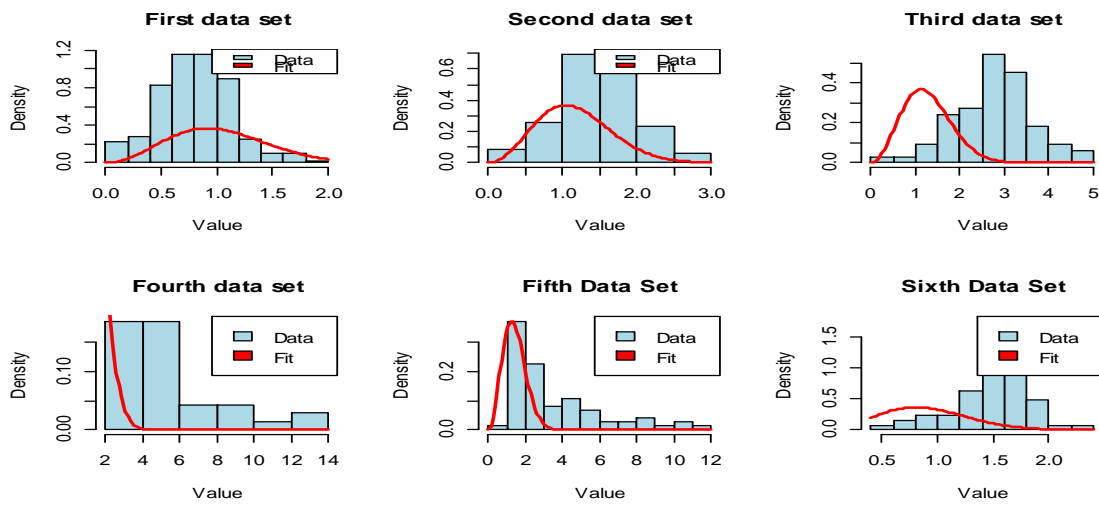


Figure 7: Density Plot for Ailamujia Distribution

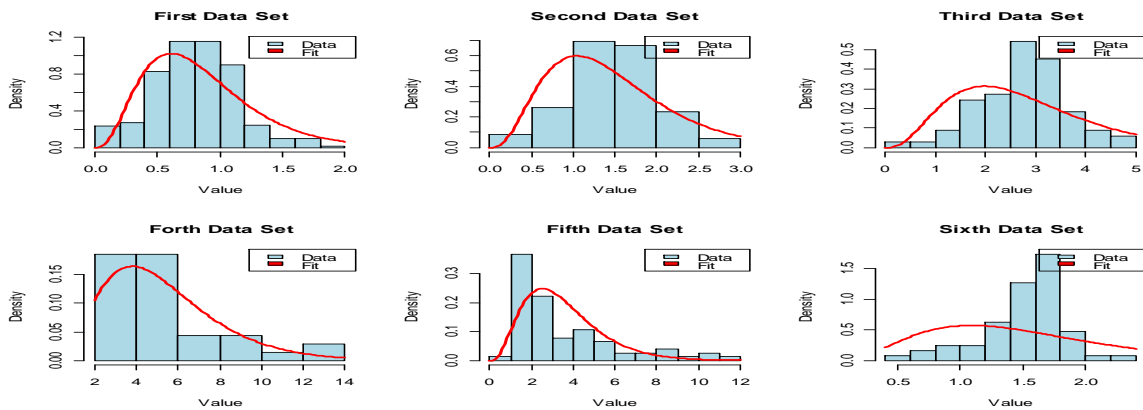


Figure 8: Density Plot for Entropy Transformed Ailamujia Distribution

From the plots presented, it clearly shows that the ETAD can flexibly fit right skewed and symmetric data sets well compared to the other seven competing distributions.

IV. Conclusion

This study has been able to utilize a novel approach known as the entropy transformation (ET) in

modifying a distribution without adding any extra parameter to the original distribution. Illustratively, the Ailamujia distribution was modified using the ET to form ETAD and it was found to perform exceedingly better than the Ailamujia distribution and other six one-parameter competing distributions. This was observed from the information criteria results and the fitted plots to six different data sets from literature. The MLE was used to estimate the parameter of the distribution which is consistent in estimating the parameter as observed from the simulation results.

References

- [1] David, I. J., Asiribo, O. E., Dikko, H. G., & Ikwuoche, P. O. (2023). Johnson-Schumacher Split-Plot Design Modelling of Rice Yield. *Biometrical Letters*, 60(1): 37-52.
- [2] David, I. J., Asiribo, O. E., & Dikko, H. G. (2023). Nonlinear Split-Plot Design Modeling and Analysis of Rice Varieties Yield. *Scientific African*, 19(9):e01444.
- [3] David, I. J., Asiribo, O. E., & Dikko, H. G. (2022). A Weibull Split-Plot Design Model and Analysis. *Thailand Statistician*, 20(2), 420-434.
- [4] David, I. J., Asiribo, O. E., & Dikko, H. G. (2022). A Bertalanffy-Richards Split-Plot Design Model and Analysis. *Journal of Statistical Modeling and Analysis*, 4(1), 56-71.
- [5] Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull G-family of probability distributions. *Journal of data science*, 12(1), 53-68.
- [6] Gupta, R. C., Gupta, P. L., & Gupta, R. D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4), 887-904.
- [7] Nofal, Z. M., Altun, E., Afify, A. Z., & Ahsanullah, M. (2019). The generalized kumaraswamy-G family of distributions. *Journal of statistical Theory and Applications*, 18(4), 329-342.
- [8] Tahir, M. H., Cordeiro, G. M., Alzaatreh, A., Mansoor, M., & Zubair, M. (2016). The logistic-X family of distributions and its applications. *Communications in statistics-Theory and methods*, 45(24), 7326-7349.
- [9] David, I. J., Mathew, S., & Eghwerido, J. T. (2023). Reliability Analysis with New Sine Inverse Rayleigh Distribution. *Journal of Reliability and Statistical Studies*, 16(2): 255-268.
- [10] David, I. J., Mathew, S., & Falgore, J. Y. (2024). Reliability Analysis with New Sine Inverted Exponential Distribution: Properties, Simulation and Application. *Journal of Reliability and Statistical Studies*, 16(2): 255-268.
- [11] Salih, A. N. (2013). Building a probabilistic distribution of the failure time spent using entropy transformation for Burr Type-XII distribution. M.Sc. thesis, Baghdad University.
- [12] Hassan, D. S., Ibrahim, N. A., & Abood, S. N. (2013). Comparing different estimators of reliability function for proposed probability distribution. *American Journal of Mathematics and Statistics*, 3(2), 84-94.
- [13] Ahmad, A., Ahmad, A., & Ozel, G. (2022). One Parameter discrete Ailamujia Distribution with statistical properties and Biodiversity and Abundance Data Applications. *Punjab University Journal of Mathematics*, 54(2), 111-125.
- [14] Lv, H. Q., Gao, L. H., & Chen, C. L., (2002). Ailamujia distribution and its application in supportability data analysis. *Journal of Academy of Armoured Force Engineering*, 16(3), 48-52.
- [15] Soleha, M. M. & Sewilam, I. A. (2007). Generalized Rayleigh Distribution revisited. *Interstat*. <http://interstat.statjournals.net/YEAR/2007/articles/0702006.pdf>
- [16] Awodutire, P. O. (2022). Statistical Properties and Applications of the Exponentiated Chen-G Family of Distributions: Exponential Distribution as Baseline Distribution. *Austrian Journal of Statistics*, 51, 57-59.
- [17] Surles, J. G. & Padgett, W. J. (1998). Inference for reliability and stress-strength for a scaled Burr Type X distribution. *Lifetime Data Analysis*, 7, 187-200.
- [18] Alizadeh, M., Bagheri, S., Bahrami, S. E., Ghobadi, S., & Nadarajah, S. (2018). Exponentiated Power Lindley Power Series Class of Distribution. *Theory and Application: Communications in Statistics, Simulation, and Computation*, 47, 2499-2531.
- [19] Tomy, L., Veena, G., & Chesneau, C. (2022). Applications of Sine Modified Lindley Distribution to Biomedical Data. *Mathematical Computation*, 27(3), 43. <https://doi.org/10.3390/mca27030043>
- [20] Shukla, K. K. (2019). Comparative study of one parameter lifetime distributions. *Biometrics & Biostatistics International Journal*, 8(4), 111-123.

MODIFIED GROUP RUNS CONTROL CHART FOR MONITORING PROCESS DISPERSION

Chandrakant G. Gardi and Vikas B. Ghute¹

•

Department of Statistics

Punyashlok Ahilyadevi Holkar Solapur University, Solapur (M.S.), India

¹Corresponding Author: Email- vbghute_stats@rediffmail.com

Abstract

Due to a rise in competitiveness, it has become an intense concern to the manufacturers to monitor process dispersion to avoid low quality production. To ensure quality production, the control chart that gives early detection of change in the dispersion is always encouraged. Researchers have suggested various control charts based on different estimators of process dispersion. Recently, many synthetic control charts based on such estimators are put forth by researchers to effectively monitor the dispersion in the process. Modified Group Runs (MGR) control chart is an extension of synthetic charts with further enhancement in the detection ability. In this paper, we propose a MGR control chart based on Downton's estimator (D). Comparison of MGR control chart with synthetic chart based on estimator D reveals the enhanced performance of MGR-D chart.

Keywords: Control chart, Process dispersion, Downton estimator, Modified group runs.

I. Introduction

For optimizing the cost and resources constraints for the manufacturing processes, it is vital to monitor the changes in the process location and/or dispersion. Control chart is a prominently used statistical process control tool to identify any change in the process parameters. Shewhart [1] proposed the idea of control charts for monitoring the process, in which certain thresholds are set for a test statistic and process is said to be in control as long as test statistic falls within these threshold values. Bourke [2] introduced the Conforming Run Length (CRL) chart for qualitative data. Wu and Spedding [3] introduced a synthetic control chart for monitoring process location as an enhancement to the Shewhart-type control chart, which includes \bar{x} chart and CRL chart. When the dispersion of the underlying process is of interest, dispersion control charts are utilized. Huang and Chen [4] extended the application of synthetic charts to monitor process dispersion by introducing the synthetic S chart, which combines S and CRL chart. Chen and Huang [5] constructed synthetic R chart, which comprises of R and CRL chart. Synthetic charts have been observed to surpass Shewhart-type control charts in performance. The development of synthetic control charts is well documented by Davis and Woodall [6], Ghute and Shirke [7], Ghute and Shirke [8], Ghute and Shirke [9]. A detailed overview of synthetic charts is given by Rakitzis et al. [10].

Klein [11] proposed runs rule control chart to enhance the effectiveness of Shewhart control chart. These two approaches, synthetic as well as runs rule charts, are superior to Shewhart-type charts. Combining these two approaches, Gadre and Rattihalli [12] constructed Group Runs (GR) control chart

for monitoring the process mean. GR control charts are demonstrated to be more effective than synthetic charts. Gadre and Kakade [13] proposed a nonparametric GR chart to detect shifts in the process median. Gadre and Rattihalli [14] introduced the Modified Group Runs scheme for detecting changes in the mean of a normally distributed process. The MGR chart is found to be more efficient than Shewhart, synthetic, and GR charts. The development of GR and MGR control charts is well documented by Gadre and Kakade [15], Rakitzis et al. [10], Khilare and Shirke [16], Ghadge and Ghute [17].

Researchers have also sought to enhance the effectiveness of control charts by considering alternative charting statistics. Abbasi and Miller [18] constructed a chart based on statistic D to monitor the process variability. The statistic D is proposed by Downton [19] as an unbiased estimator of standard deviation for normally distributed process. It has been demonstrated that the D chart is as effective as the Shewhart S chart in identifying changes in the standard deviation of a normally distributed process. Gardi and Ghute [20] constructed D chart using runs rule, whereas Rajmanya and Ghute [21] developed synthetic D chart as a combination of D chart and CRL chart, which enhanced the performance of D chart. This paper attempts to improve the performance of synthetic D chart by using the idea of Modified Group Runs (MGR-D chart).

The rest of the paper is structured as follows: The D chart and CRL chart are discussed in section 2. In section 3, MGR-D chart is proposed for monitoring the process dispersion. The performance evaluation and comparison of proposed chart with synthetic D chart is made in section 4. Conclusions are given in the section 5.

II. The D chart

For normally distributed process, Downton [19] proposed an unbiased estimator D of process standard deviation σ . Let X_1, X_2, \dots, X_n be a random sample of size n from $N(\mu, \sigma^2)$. Then Downton's statistic D is given as

$$D = \frac{2\sqrt{\pi}}{n(n-1)} \sum_{i=1}^n \left[i - \frac{1}{2}(n+1) \right] X_{(i)} \quad (1)$$

where $X_{(i)}$ is the i^{th} ($i = 1, 2, \dots, n$) order statistic for the given sample. Abbasi and Miler [18] developed Shewhart-type control chart based on the statistic D and revealed that the D-chart is as effective as Shewhart S chart in discovering a dispersion change. The probabiity limits for D-chart were obtained using the distribution of $Z = D/\sigma$. The Upper and Lower Control Limit (UCL and LCL) for the D chart is given as

$$\left. \begin{aligned} \text{UCL} &= Z_{1-\alpha} \bar{D} \text{ with } P(Z \geq Z_{1-\alpha}) = 1 - \alpha \\ \text{LCL} &= Z_{\alpha} \bar{D} \text{ with } P(Z < Z_{\alpha}) = \alpha \end{aligned} \right\} \quad (2)$$

Where Z_{α} is α^{th} quantile point of the distribution of Z and α is Type-I error probability, specified in advance. To monitor process dispersion, D-values are plotted on the chart. Decision regarding out-of-control signal is taken based on whether a D-value goes beyond UCL, for the case of detecting positive shift or whether D-value is smaller than LCL, for the case of detecting negative shift.

We assume that the process parameters μ and σ^2 are known, that is, $\mu = \mu_0$ and $\sigma^2 = \sigma_0^2$ are in-control mean and variance of the process respectively. Let $\sigma_1 = \delta\sigma_0$ ($0 < \delta \neq 1$) be the shifted value of standard deviation, with a shift of size δ . In order to detect a positive (negative) shift, that is, $\delta > 1$ ($\delta < 1$) in the process standard deviation, an upper control limit $k^+\sigma_0$ (a lower control limit $k^-\sigma_0$) of D-chart is required, and a signal is given if $D > k^+\sigma_0$ ($D < k^-\sigma_0$). The average run length (ARL) is the average number of D samples required to identify a shift in σ of the D-chart. It is calculated as

$$\text{ARL}_D(\delta) = \frac{1}{P(\delta)} \quad (3)$$

where $P(\delta)$ is the probability of detecting a shift δ in the process standard deviation.

For $\delta > 1$,

$$P(\delta) = P(D > k^+ \sigma_0 \mid \sigma = \delta \sigma_0) = P\left(Z > k^+ / \delta\right) = 1 - F(k^+ / \delta) \quad (4)$$

For $\delta < 1$,

$$P(\delta) = P(D < k^- \sigma_0 \mid \sigma = \delta \sigma_0) = P\left(Z < k^- / \delta\right) = F(k^- / \delta) \quad (5)$$

Where $F(\cdot)$ denotes the cumulative distribution function.

Conforming Run Length (CRL) chart

For monitoring attribute characteristic, the conforming run length (CRL) chart is used, which detects shift in the fraction nonconforming p . The random variable CRL is defined as the number of conforming units between two successive nonconforming units including ending nonconforming unit. The distribution of CRL is as follows.

$$F(\text{CRL}) = 1 - (1 - p)^{\text{CRL}}, \text{CRL} = 1, 2, 3, \dots \quad (6)$$

where p is the probability of the nonconforming unit. If practitioner is only concerned with detection of an increase in p , the lower control limit (L) of the CRL chart serves the purpose. It is given by

$$L = \frac{\ln(1 - \alpha_{\text{CRL}})}{\ln(1 - p_0)} \quad (7)$$

Where $\alpha_{\text{CRL}} = 1 - (1 - p_0)^L = F_{p_0}(L)$ is the type-I error probability of the CRL chart and p_0 is the in-control fraction nonconforming. If a sample CRL is not greater than L , it indicates the increase in fraction nonconforming and an out-of-control signal is issued.

The ARL of such CRL chart is given by

$$ARL_{\text{CRL}} = \frac{1}{F_{\text{CRL}}(L)} = \frac{1}{1 - (1 - p)^L} \quad (8)$$

III. Modified Group Runs Control Chart

In this section, we present the design of modified group runs control chart based on Downton estimator D. Modified group runs control chart is the integration of D-chart and extended version of CRL chart. The D chart has only upper control limit (UCL) and CRL chart has two limits, namely the warning limit L_1 and the lower limit L_2 . Let Y_r be the r^{th} group based CRL, then modified group runs control chart declares the process as out-of-control if $Y_1 \leq L_1$ or for some $r (> 1)$, $Y_r \leq L_1$ and $Y_{r+1} \leq L_2$.

The steps for implementing the MGR-D chart are as follows:

1. Fix the UCL for the D chart and L_1 and L_2 of the CRL chart.
2. Select a subgroup of n items at each inspection point j and compute the chart statistic, say D_j
3. If $D_j \leq \text{UCL}$, then the subgroup is considered 'conforming' and control flow returns to step 2. Else the subgroup is considered 'nonconforming' and control flow goes to the next step.
4. Check the number ($\text{CRL}_i, i = 1, 2, \dots$) of subgroups between the current and previous nonconforming groups.
5. If $\text{CRL}_1 \leq L_2$ or for some $i > 1$, $\text{CRL}_i \leq L_1$ and $\text{CRL}_{i+1} \leq L_2$, for $i = 2, 3, \dots$ for the first time, the process is thought to be out-of-control, and control flow moves to the next step. Else flow returns to step 2.
6. Signal the out-of-control state.
7. An assignable cause should be identified and corrective action should be taken to remove it.

Let the expected number of subgroups needed to identify a shift of magnitude δ in process dispersion be $ARL_{\text{MGR}}(\delta)$. Following Gadre and Rattihalli (2004), the performance measure $ARL_{\text{MGR}}(\delta)$ for the MGR-D chart can be expressed as follows:

For increase in the process dispersion, that is, when $\delta > 1$,

$$ARL_{\text{MGR}}(\delta) = \frac{1}{P(\delta)} \times \frac{[1 + Q(\delta)^{L_2} - Q(\delta)^{L_1}]}{[1 - Q(\delta)^{L_1}][1 - Q(\delta)^{L_2}]} \quad (9)$$

Where $Q(\delta) = 1 - P(\delta)$ and $P(\delta)$ is given in equation (4) when $\delta > 1$ and is given in equation (5) when $\delta < 1$.

For constructing MGR-D chart, the below ARL model is used.

$$\begin{aligned} & \text{Minimize } \text{ARL}_{\text{MGR}}(\delta) \\ & \text{Subject to } \text{ARL}_{\text{MGR}}(0) \geq \tau \end{aligned}$$

where τ is the minimum required value of $\text{ARL}_{\text{MGR}}(0)$.

Optimal Design Procedure

The optimal design procedure for MGR-D chart is given below:

1. Specify subgroup size n , shift size δ^* and in-control ARL as ARL_0 .
2. Initialize L_1 as 1.
3. Initialize L_2 as 1.
4. Obtain $P(0)$ by solving Equation (9) numerically. From the obtained value of $P(0)$, obtain k^+ (or k^-) using Equation (4) (or Equation (5)).
5. From the current values of L_1 , L_2 and k^+ (or k^-), obtain $P(0)$ from Equation (4) (or Equation (5)). Then compute $\text{ARL}_{\text{MGR}}(\delta)$ using Equation (9).
6. If $\text{ARL}_{\text{MGR}}(\delta)$ is reduced, then increase L_1 by 1 and go back to step 4. Else go to the next step.
7. Mark the minimum $\text{ARL}_{\text{MGR}}(\delta)$ for current combination of L_1 , L_2 . If currently marked $\text{ARL}_{\text{MGR}}(\delta)$ for (L_1, L_2) is less than previously marked $\text{ARL}_{\text{MGR}}(\delta)$, present pair of (L_1, L_2) is the optimum value. Else increase L_2 by 1, initialize L_1 as 1 and go back to step 4.

IV. Performance Evaluation

In this section the performance of proposed MGR-D chart is evaluated and is also compared with synthetic D chart. The comparison is made based on the Average Run Length (ARL). The underlying process is assumed to have normal distribution with mean 0 and variance 1. Since, the exact distribution of D statistic is not, we used simulation approach to obtain quantiles of distribution of D. We considered three different subgroup sizes as $n = 5, 8, 10$. For each subgroup size, 50000 subgroups were simulated. Based on D statistic of these 50000 subgroups, required quantile points of the distribution of D were determined. This process was repeated 500 times and average of these 500 quantiles was considered for obtaining control limits of the proposed chart. For fair comparison, all charts are fabricated such that in-control ARL remains the same as 200. The optimal design parameters L_1 , L_2 and k^+ (or k^-) of proposed MGR-D chart are obtained for a pre-determined shift of size $\delta = 1.2$ ($\delta = 0.8$) for positive (negative) shift in process deviation. The L_1 is initiated as 1, and L_2 is gradually incremented by 1, for each combination of (L_1, L_2) , ARL is determined. Once the minimum ARL is achieved, the combination (L_1, L_2) is noted and L_1 is incremented by 1 and same process is repeated. If the minimum ARL for present combination of (L_1, L_2) is greater than that for the earlier combination of (L_1, L_2) , the earlier combination of (L_1, L_2) is considered. That is, the combination of (L_1, L_2) at which ARL attains its minimum across L_1 as well as L_2 . Table 1 gives a demonstration for obtaining the optimal parameters for $n = 10$ for detecting positive shift. Here, the minimum ARL is 4.235941 and is attained at $(L_1 = 1, L_2 = 13)$.

Once the limits are set ensuring in-control ARL to be 200, out-of-control ARLs are determined for various shift sizes. For positive shift in the process dispersion, the shift sizes considered are $\delta = 1, 1.1, 1.2, 1.3, 1.4, 1.5, 2.0$ and that for negative shift in the process dispersion are $\delta = 0.9, 0.8, 0.7, 0.6, 0.5, 0.1$. The ARL values are determined using simulation of size 50000 for various sample sizes as well as for different subgroup sizes. Table 2 presents these ARL values as well as ARL values for synthetic D chart for positive shift in process dispersion, whereas Table 3 provides ARLs for negative shift. Since out-of-control ARL values of MGR-D chart are smaller than that of the synthetic-D chart for all considered shifts in the process standard deviation and considered subgroup sizes, the

proposed MGR-D chart is very superior to synthetic D chart for shifts in either direction as well as for different subgroup sizes.

Table 1: Optimal Parameters for MGR-D chart.

L_1	L_2	k^+	ARL	L_1	L_2	k^+	ARL
1	1	1.229106	11.51437	2	1	1.257766	11.71673
1	2	1.271325	8.132786	2	2	1.29923	8.422839
1	3	1.295223	6.644562	2	3	1.32252	7.002282
1	4	1.311976	5.79848	2	4	1.338762	6.204549
1	5	1.324845	5.264146	2	5	1.351146	5.696124
1	6	1.335386	4.913096	2	6	1.361216	5.359928
1	7	1.344217	4.66926	2	7	1.369726	5.130402
1	8	1.351881	4.502848	2	8	1.37713	4.97412
1	9	1.358633	4.38984	2	9	1.38362	4.866527
1	10	1.364715	4.316068	2	10	1.389399	4.794161
1	11	1.370216	4.269155	2	11	1.394621	4.74751
1	12	1.37528	4.244542	2	12	1.399379	4.721099
1	13	1.379935	4.235941	2	13	1.403793	4.710871
1	14	1.384273	4.239508	2	14	1.40785	4.711723

Table 2: ARL comparison for positive shift in process dispersion.

Shift (δ)	n = 5		n = 8		n = 10	
	Synthetic D chart	MGR D Chart	Synthetic D chart	MGR D Chart	Synthetic D chart	MGR D Chart
	$L = 17$ $k^+ = 1.843$	$L_1 = 1$ $L_2 = 24$ $k^+ = 1.647$	$L = 12$ $k^+ = 1.595$	$L_1 = 1$ $L_2 = 16$ $k^+ = 1.448$	$L = 12$ $k^+ = 1.519$	$L_1 = 1$ $L_2 = 13$ $k^+ = 1.380$
1.0	200 (2.371)	199.498 (1.99014)	200 (2.366)	199.864 (1.74324)	201 (2.388)	199.787 (1.63627)
1.1	43.92 (0.560)	22.1446 (0.27269)	32.80 (0.422)	16.271 (0.17797)	28.61 (0.366)	14.3427 (0.14795)
1.2	15.89 (0.202)	7.23788 (0.05486)	10.75 (0.132)	5.07486 (0.03393)	8.66 (0.103)	4.23606 (0.02686)
1.3	8.34 (0.093)	4.45104 (0.01932)	5.27 (0.057)	3.09928 (0.01272)	4.41 (0.045)	2.61958 (0.00995)
1.4	5.38 (0.055)	3.32406 (0.01248)	3.44 (0.032)	2.32074 (0.00785)	2.89 (0.024)	1.97842 (0.00617)
1.5	3.92 (0.036)	2.68268 (0.00952)	2.56 (0.028)	1.89742 (0.00585)	2.16 (0.016)	1.62993 (0.0045)
2.0	1.79 (0.012)	1.51418 (0.00395)	1.31 (0.007)	1.1828 (0.00208)	1.18 (0.005)	1.10386 (0.00151)

Table 3: ARL comparison for negative shift in process dispersion.

Shift (δ)	n = 5		n = 8		n = 10	
	Synthetic D chart	MGR-D Chart	Synthetic D chart	MGR-D Chart	Synthetic D chart	MGR-D Chart
	$L = 5$ $k^- = 0.396$	$L_1 = 1$ $L_2 = 36$ $k^+ = 0.423$	$L = 7$ $k^+ = 0.516$	$L_1 = 1$ $L_2 = 18$ $k^+ = 0.587$	$L = 6$ $k^+ = 0.5757$	$L_1 = 1$ $L_2 = 13$ $k^+ = 0.648$
1.0	200 (2.433)	200.11 (2.254)	200 (2.292)	201.67 (1.829)	200 (2.228)	199.9 (1.674)
0.9	68.14 (0.840)	54.32 (0.73)	67.11 (0.808)	32.74 (0.367)	55.55 (0.651)	26.31 (0.271)
0.8	30.20 (0.382)	16.33 (0.194)	22.28 (0.291)	7.57 (0.067)	15.47 (0.198)	5.47 (0.042)
0.7	15.98 (0.201)	7.56 (0.045)	8.00 (0.096)	3.46 (0.014)	5.19 (0.061)	2.54 (0.01)
0.6	10.07 (0.120)	4.51 (0.018)	3.32 (0.033)	2.05 (0.007)	2.28 (0.019)	1.55 (0.004)
0.5	6.99 (0.076)	2.78 (0.01)	1.78 (0.012)	1.34 (0.003)	1.35 (0.007)	1.13 (0.002)
0.1	2.78 (0.023)	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)

V. Conclusions

In this paper, MGR-D control chart is proposed for monitoring changes in the process dispersion of normally distributed process. The proposed chart is based on Downton's statistic D and is an integration of D chart and an extended version of CRL chart. The ARL comparison highlights that the proposed MGR-D chart performs better than the synthetic D chart.

References

- [1] Shewhart, W. A. (1931). Economic control of quality of manufactured product. Reprinted by the American Society for Quality Control in 1980.
- [2] Bourke, P. D. (1991). Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection. *Journal of Quality Technology*, 23(3), 225-238.
- [3] Wu, Z. and Spedding, T. A. (2000). A synthetic control chart for detecting small shifts in the process mean. *Journal of Quality Technology*, 32(1), 32-38.
- [4] Huang, H. J. and Chen, F. L. (2005). A synthetic control chart for monitoring process dispersion with sample standard deviation. *Computers and Industrial Engineering*, 49(2), 221-240.
- [5] Chen, F. L. and Huang, H. J. (2005). A synthetic control chart for monitoring process dispersion with sample range. *The International Journal of Advanced Manufacturing Technology*, 26, 842-851.
- [6] Davis, R. B., and Woodall, W. H. (2002). Evaluating and improving the synthetic control chart. *Journal of Quality Technology*, 34(2), 200-208.
- [7] Ghute, V. B. and Shirke, D. T. (2007). Joint monitoring of multivariate process using synthetic control charts, *International Journal of Statistics and Management System*, 2(1-2), 129-141.

- [8] Ghute, V. B., and Shirke, D. T. (2008a). A multivariate synthetic control chart for monitoring process mean vector. *Communications in Statistics-Theory and Methods*, 37(13), 2136-2148.
- [9] Ghute, V. B. and Shirke, D. T. (2008b). A multivariate synthetic control chart for process dispersion. *Quality Technology and Quantitative Management*, 5(3), 271-288.
- [10] Rakitzis, A. C., Chakraborti, S., Shongwe, S. C., Graham, M. A., and Khoo, M. B. C. (2019). An overview of synthetic-type control charts: Techniques and methodology. *Quality and Reliability Engineering International*, 35(7), 2081-2096.
- [11] Klein, M. (2000). Two alternatives to the Shewhart X control chart. *Journal of Quality Technology*, 32(4), 427-431.
- [12] Gadre, M.P. and Rattihalli, R.N. (2004). A group runs control chart for detecting shifts in the process mean. *Economic Quality Control*, 19, 29-43.
- [13] Gadre, M. P. and Kakade, V. C. (2014). A nonparametric group runs control chart to detect shift in the process median. *IAPQR Transactions*, 39(1), 29-53.
- [14] Gadre, M. P. and R. N. Rattihalli (2006) Modified Group Runs Control Charts to Detect Increases in Fraction Non Conforming and Shifts in the Process Mean, *Communications in Statistics - Simulation and Computation*, 35:1, 225-240, DOI: 10.1080/03610910500416256
- [15] Gadre, M. P. and Kakade, V. C. (2016). Some group runs based multivariate control charts for monitoring the process mean vector. *Open Journal of Statistics*, 6(6), 1098-1109.
- [16] Khilare, S. K., and Shirke, D. T. (2023). A nonparametric group runs control chart for location using sign statistic. *Thailand Statistician*, 21(1), 165-179.
- [17] Ghadge, O.D. and Ghute, V. B. (2023). Group runs and modified group runs control charts for monitoring linear regression profiles. *Reliability: Theory & Applications*, 18 (4 (76)), 513-524.
- [18] Abbasi, S. A. and Miller, A. (2011). D chart: An efficient alternative to monitor process dispersion. In *Proceedings of the World Congress on Engineering and Computer Science* (Vol. 2, pp. 19-21).
- [19] Downton, F. (1966). Linear estimates with polynomial coefficients. *Biometrika*, 53(1/2), 129-141.
- [20] Gardi, C. and Ghute, V. (2023). A new control chart for process dispersion based on ranked set sampling. *Reliability: Theory & Applications*, 18 (2 (73)), 154-166. doi: 10.24412/1932-2321-2023-273-154-166
- [21] Rajmanya, S. V. and Ghute, V. B. (2014). A synthetic control chart for monitoring process variability. *Quality and Reliability Engineering International*, 30(8), 1301-1309.

ON THE FLEXIBILITY OF TYPE I HALF LOGISTIC EXPONENTIATED FRECHET DISTRIBUTION

Olalekan Akanji Bello^{1*}

Sani Ibrahim Doguwa¹

Abukakar Yahaya¹

Haruna Mohammed Jibril²

Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria¹.

Department of Mathematics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria,
Nigeria².

olalekan4sure@gmail.com^{1*}

sidoguwa@gmail.com¹

ensiliyu2@yahoo.co.uk¹

alharun2004@yahoo.com²

Abstract

In this article, we delve into the modeling and analysis of lifetimes, which hold substantial importance across various scientific and industrial fields. Our focus is on introducing a novel distribution termed the Type I Half-Logistic Exponentiated Frechet (TIHLEtF) Distribution, which is an extension of the Frechet distribution. We have derived a crucial representation of the density function for this distribution. Furthermore, we explore several statistical properties associated with the TIHLEtF distribution. These properties encompass explicit expressions for the quantile function, probability-weighted moments, moments, moments generating function, reliability function, hazard function, and order statistics. To estimate the model parameters, we employ the maximum likelihood estimation technique and present the results of a simulation study. To emphasize the superiority of our newly introduced distribution, we apply it to two real datasets. The outcomes of our analysis reveal that the TIHLEtF distribution outperforms the other considered distributions in terms of fitting the data in these real-world cases.

Keywords Type I Half-Logistic Exponentiated-G, Frechet distribution, Quantile function, Hazard function, Maximum likelihood, Order Statistics

I. Introduction

The Frechet distribution, often referred to as the type II extreme value distribution, plays a crucial role in various fields, including engineering, actuarial science, environmental studies, medical sciences, economics, finance, and insurance. It serves as a fundamental statistical tool in these disciplines. It was introduced by [10] as a means of modeling extreme values in data. Despite the widespread use of traditional probability distributions in these areas, there is a growing need for more flexible forms of these distributions, and the Frechet distribution is one such example that has been proposed to meet this need. It is a flexible tool for modeling data and is widely used in extreme value theory. Various researchers have proposed several modifications to the Frechet

distribution in recent literature. The exponentiated Frechet was first introduced by [16], followed by the beta Frechet proposed by [17], and the transmuted Frechet proposed by [13]. The gamma extended Frechet defined by [7], while [12] introduced the Marshall-Olkin Frechet. The Kumaraswamy Frechet proposed by [14], while [9] studied the transmuted exponentiated Frechet. [1] investigated the transmuted Marshall-Olkin Frechet, and [3] proposed the Kumaraswamy Marshall-Olkin Frechet, [2] also studied the Weibull Frechet. A novel distribution family was introduced by [4] in their recent study, which they have named Type I Half-Logistic Exponentiated-G (TIHLEt-G). This distribution family is characterized by two positive shape parameters, denoted by λ and α , and can be applied to any arbitrary cumulative distribution function (cdf) $H(x, \mathcal{G})$. The cumulative distribution function (cdf) and the probability density function (pdf) for TIHLEt-G are given by

$$F_{TIHLEt-G}(x; \lambda, \alpha, \mathcal{G}) = \frac{1 - [1 - H^\alpha(x; \mathcal{G})]^\lambda}{1 + [1 - H^\alpha(x; \mathcal{G})]^\lambda}, \quad x > 0, \lambda, \alpha > 0 \quad (1)$$

and

$$f_{TIHLEt-G}(x; \lambda, \alpha, \mathcal{G}) = \frac{2\lambda\alpha h(x; \mathcal{G})H_{(x; \mathcal{G})}^{\alpha-1}[1 - H_{(x; \mathcal{G})}^\alpha]^{\lambda-1}}{[1 + [1 - H_{(x; \mathcal{G})}^\alpha]^\lambda]^2}, \quad x > 0, \lambda, \alpha > 0 \quad (2)$$

The TIHLEt-G family of distributions is noteworthy due to several factors, as explained in [4]. This family of distributions offers increased flexibility in terms of kurtosis compared to conventional models. It also enables the creation of skewed distributions even for symmetrical ones and can produce heavy-tailed distributions that better fit real data. The TIHLEt-G family can generate symmetric, left-skewed, right-skewed, and reversed-J shaped distributions and allows for special models with different types of hazard rate functions.

The Frechet distribution's cdf and pdf are provided as

$$H(x; \theta, \delta) = e^{-\left(\frac{\theta}{x}\right)^\delta}, \quad x > 0, \theta, \delta > 0 \quad (3)$$

$$h(x; \theta, \delta) = \delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta}, \quad x > 0, \theta, \delta > 0 \quad (4)$$

The rest of the paper is organized into several sections. The second section describes the materials and a method used to drive the pdf and cdf of Type I Half-Logistic Exponentiated Frechet (TIHLEtF) Distribution and presents the expansion of the density. In the third section, we explore the statistical properties of the distribution, including its moment, moment-generating function, probability-weighted moment, reliability function, hazard function, and quantile function. The fourth section derives order statistics. The fifth section describes how maximum likelihood estimation is used to estimate the unknown model parameters and presents the results of a simulation study. The sixth section demonstrates the flexibility of the TIHLEtF distribution using two real data sets. Finally, the paper concludes with a summary of the findings and some closing remarks in the seventh section.

II. The Type I Half-Logistic Exponentiated Frechet (TIHLEtF) Distribution

A new model called the TIHLEtF model is introduced, where a random variable X is said to follow the TIHLEtF model if its cumulative distribution function (cdf) is obtained by using equation (3) in equation (1), which is defined as follows:

$$F_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta) = \frac{1 - \left[1 - \left(e^{-\alpha \left(\frac{\theta}{x}\right)^\delta} \right) \right]^\lambda}{1 + \left[1 - \left(e^{-\alpha \left(\frac{\theta}{x}\right)^\delta} \right) \right]^\lambda}, \quad x > 0, \lambda, \alpha, \theta, \delta > 0 \quad (5)$$

and the pdf corresponding to equation (6) is

$$f_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta) = 2\lambda\alpha\delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta} e^{-(\alpha-1)\left(\frac{\theta}{x}\right)^\delta} \left[1 - e^{-\alpha\left(\frac{\theta}{x}\right)^\delta} \right]^{\lambda-1} \left[1 + \left[1 - e^{-\alpha\left(\frac{\theta}{x}\right)^\delta} \right]^\lambda \right]^{-2} \quad (6)$$

where θ is a scale parameter and λ, α, δ are shape parameters.

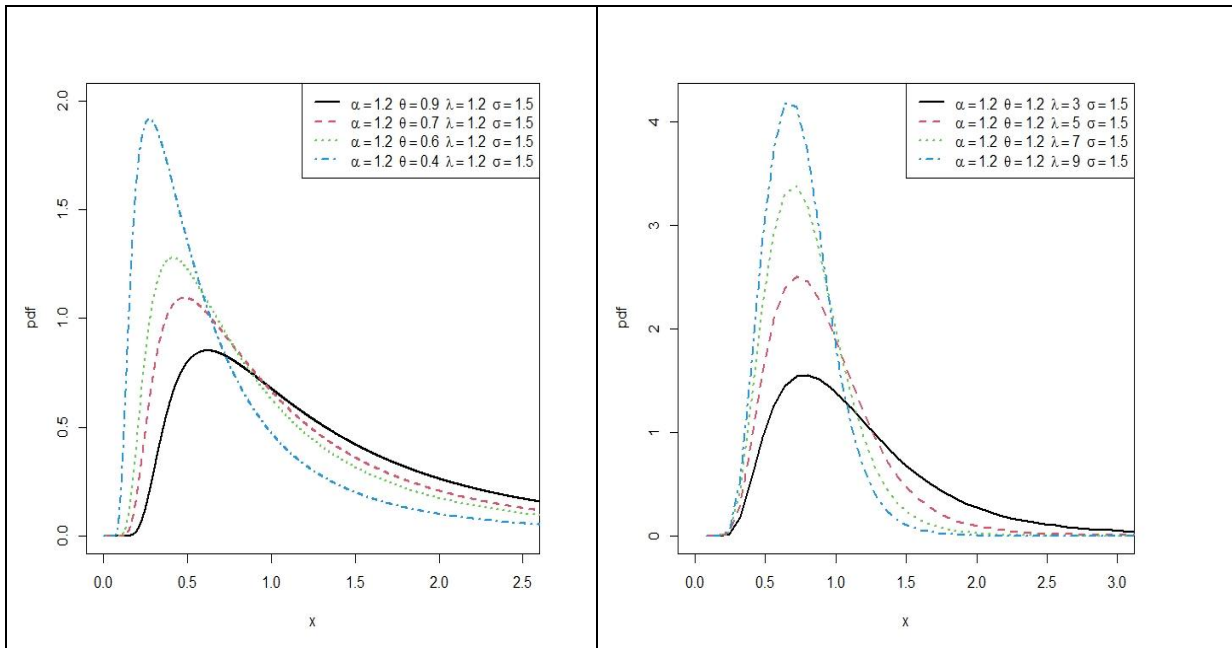


Figure 1: Plots of pdf of TIHLEtF distribution for different values of parameters

III. Expansion of Density for TIHLEtF Distribution

In this section, we present an advantageous expansion of the pdf and cdf for the TIHLEtF distribution. As a result of the generalized binomial series being

$$(1+z)^{-b} = \sum_{i=0}^{\infty} (-1)^i \binom{b+i-1}{i} z^i \quad (7)$$

Moreover, by applying the binomial theorem from equation (7) to equation (6)

$$f_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta) = 2\lambda\alpha\delta\theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta} e^{-(\alpha-1)\left(\frac{\theta}{x}\right)^\delta} \sum_{i=0}^{\infty} (-1)^i \binom{1+i}{i} \left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^\alpha \right]^{\lambda(i+1)-1}$$

also

$$\left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^\alpha \right]^{\lambda(i+1)-1} = \sum_{j=0}^{\infty} (-1)^j \binom{\lambda(i+1)-1}{j} \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha i}$$

Now, the pdf can be written as

$$f_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta) = \sum_{i,j=0}^{\infty} \eta_p \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha(j+1)} \tag{8}$$

where, $\eta_p = 2\lambda\alpha\delta\theta^\delta x^{-\delta-1} (-1)^{i+j} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j}$

Furthermore, an expansion for the $[F_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta)]^h$ is produced, with h being an integer, and the binomial expansion is worked out once more.

$$[F_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta)]^h = \sum_{p,m=0}^h (-1)^{p+m} \binom{h}{m} \binom{h+p-1}{p} \left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^\alpha \right]^{\lambda(p+m)}$$

Consider

$$\left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^\alpha \right]^{\lambda(p+m)} = \sum_{z=0}^{\infty} (-1)^z \binom{\lambda(p+m)}{z} \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha z}$$

The cdf can be written as:

$$[F_{TIHLEtF}(x; \lambda, \alpha, \theta, \delta)]^h = \sum_{p,m=0}^h \varphi_t \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha z} \tag{9}$$

where, $\varphi_t = \sum_{z=0}^{\infty} (-1)^{p+m+z} \binom{h+p-1}{p} \binom{h}{m} \binom{\lambda(m+p)}{z}$

IV. Statistical Properties

We derived some statistical properties of the new distribution.

I. Probability weighted moments

The probability-weighted moments (PWMs) were introduced by [11]. It is used to derive inverse form estimators for the parameters and quantiles of a distribution. The PWMs, is denoted by $K_{r,s}$ which can be derived for a random variable X using the following affiliations.

$$K_{r,s} = E[X^r F(x)^s] = \int_0^\infty x^r f(x) F(x)^s dx \tag{10}$$

The PWMs of TIHLEtF distribution is developed by substituting (8) and (9) into (10), and substituting h with s, as proceed

$$K_{r,s} = \sum_{i,j=0}^{\infty} \sum_{p,m=0}^s \eta_p \varphi_t \int_0^{\infty} x^r \left[e^{-\left(\frac{\theta}{x}\right)^{\delta}} \right]^{\alpha(j+1+z)} dx \quad (11)$$

Consider the integral

$$\int_0^{\infty} x^r \left[e^{-\left(\frac{\theta}{x}\right)^{\delta}} \right]^{\alpha(j+1+z)} dx$$

$$\text{Let } y = \alpha(j+1+z)\left(\frac{\theta}{x}\right)^{\delta} \Rightarrow x = \left[\frac{\alpha(j+1+z)\theta^{\delta}}{y} \right]^{\frac{1}{\delta}}; dx = \frac{dyx^{\delta-1}}{\delta\theta^{\delta}\alpha(j+1+z)}$$

Then

$$\int_0^{\infty} \left[\frac{\alpha(j+1+z)\theta^{\delta}}{y} \right]^{\frac{r}{\delta}} e^{-y} \frac{dyx^{\delta-1}}{\delta\theta^{\delta}\alpha(j+1+z)} = \int_0^{\infty} y^{-\frac{r}{\delta}} e^{-y} dy = \Gamma\left(1 - \frac{r}{\delta}\right)$$

The PWMs of TIHLEtF can be written as proceed

$$K_{r,s} = \sum_{i,j=0}^{\infty} \sum_{p,m=0}^s (\alpha)^{\frac{r}{\delta}} \theta^r (j+1+z)^{\frac{r}{\delta}-1} \eta_p \varphi_t \Gamma\left(1 - \frac{r}{\delta}\right) \quad (12)$$

now

$$\varphi_t = \sum_{z,q=0}^{\infty} (-1)^{p+m+z+q} \binom{s+p-1}{p} \binom{s}{m} \binom{\lambda(m+p)}{z}$$

and

$$\eta_p = 2\lambda(-1)^{i+j} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j}$$

II. Moments

Moments are fundamental to all statistical analysis, particularly in applications. So, for the new distribution, we determine the r^{th} moment.

$$E(X^r) = \int_0^{\infty} x^r f(x) dx \quad (13)$$

Using the important representation of the pdf as shown in equation (8), we have

$$E(X^r) = \sum_{i,j=0}^{\infty} \eta_p \int_0^{\infty} x^r \left[e^{-\left(\frac{\theta}{x}\right)^{\delta}} \right]^{\alpha(j+1)} dx \quad (14)$$

Consider the integral

$$\int_0^{\infty} x^r \left[e^{-\left(\frac{\theta}{x}\right)^{\delta}} \right]^{\alpha(j+1)} dx$$

$$\text{Let } w = \alpha(j+1)\left(\frac{\theta}{x}\right)^{\delta} \Rightarrow x = \left[\frac{\alpha(j+1)\theta^{\delta}}{w} \right]^{\frac{1}{\delta}}; dx = \frac{dw x^{\delta-1}}{\alpha(j+1)\theta^{\delta}}$$

Then

$$\int_0^{\infty} \left[\frac{\alpha(j+1)\theta^{\delta}}{w} \right]^{\frac{r}{\delta}} e^{-w} \frac{dw x^{\delta-1}}{\alpha(j+1)\theta^{\delta}} = \theta^r \alpha^{\frac{r}{\delta}} (j+1)^{\frac{r}{\delta}-1} \int_0^{\infty} w^{-\frac{r}{\delta}} e^{-w} dw = \int_0^{\infty} w^{-\frac{r}{\delta}} e^{-w} dw = \Gamma\left(1 - \frac{r}{\delta}\right)$$

The r^{th} moment of the TIHLEtF distribution can be expressed as follows

$$E(X^r) = \sum_{i,j=0}^{\infty} \eta_p \theta^r \alpha^{\frac{r}{\delta}} (j+1)^{\frac{r}{\delta}-1} \Gamma(1-\frac{r}{\delta}) \quad (15)$$

Now

$$\eta_p = 2\lambda\alpha(-1)^{i+j} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j}$$

The expected value and the spread of the TIHLEtF distribution can be described by the following

$$E(X) = \sum_{i,j=0}^{\infty} \eta_p \theta \alpha^{\frac{1}{\delta}} (j+1)^{\frac{1}{\delta}-1} \Gamma(1-\frac{1}{\delta})$$

and

$$\text{var}(x) = \sum_{i,j=0}^{\infty} \eta_p \theta \alpha^{\frac{1}{\delta}} (j+1)^{\frac{1}{\delta}-1} \Gamma(1-\frac{1}{\delta}) - \left[\sum_{i,j=0}^{\infty} \eta_p \theta \alpha^{\frac{1}{\delta}} (j+1)^{\frac{1}{\delta}-1} \Gamma(1-\frac{1}{\delta}) \right]^2$$

III. Moment-generating function (mgf)

The Moment-Generating Function of x is expressed as follows:

$$M_x(t) = \int_0^{\infty} e^{tx} f(x) dx \quad (16)$$

where the expansion of $e^{tx} = \sum_{m=0}^{\infty} \frac{(tx)^m}{m!}$

The moment-generating function of TIHLEtF distribution can be represented as follows

$$M_x(t) = \sum_{i,j=0}^{\infty} \sum_{m=0}^{\infty} \frac{t^m \eta_p \theta^m \alpha^{\frac{m}{\delta}} (j+1)^{\frac{m}{\delta}-1} \Gamma(1-\frac{m}{\delta})}{m!} \quad (17)$$

IV. Reliability function

The reliability function provides the likelihood that an individual will endure beyond a designated time frame. This function is defined as follows:

$$R(x; \lambda, \alpha, \theta, \delta) = \frac{2 \left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^{\delta}} \right]^{\alpha} \right]^{\lambda}}{1 + \left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^{\delta}} \right]^{\alpha} \right]^{\lambda}} \quad (18)$$

V. Hazard function

The hazard function represents the likelihood of the event of interest happening within a relatively brief time period. It can be defined as follows:

$$T(x; \lambda, \alpha, \theta, \delta) = \frac{\lambda \alpha \delta \theta^\delta x^{-\delta-1} e^{-\left(\frac{\theta}{x}\right)^\delta} \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha-1}}{\left[1 + \left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^\alpha \right]^\lambda \right] \left[1 - \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^\alpha \right]} \quad (19)$$

VI. Quantile Function

The quantile function, also known as the inverse CDF, of the TIHLEtF distribution is obtained by using the CDF in equation (5)

$$x = Q(u) = \frac{\theta}{\left[-\log \left[1 - \left[\frac{1-U}{U+1} \right]^{\frac{1}{\lambda}} \right]^{\frac{1}{\alpha}} \right]^{\frac{1}{\delta}}} \quad (20)$$

The median of the TIHLEtF distribution can be derived by substituting $U=0.5$ in equation (20) as follows:

$$\text{median} = Q(0.5) = \frac{\theta}{\left[-\log \left[1 - \left[\frac{1-0.5}{0.5+1} \right]^{\frac{1}{\lambda}} \right]^{\frac{1}{\alpha}} \right]^{\frac{1}{\delta}}} \quad (21)$$

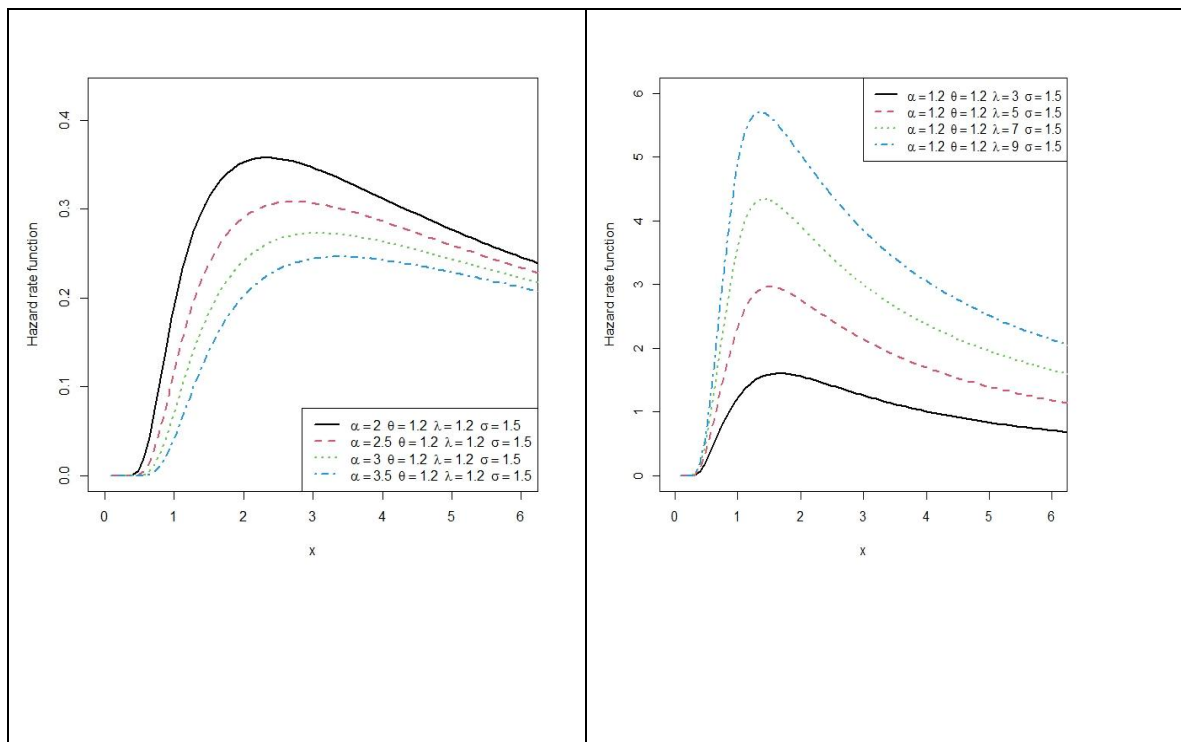


Figure 2: Plots of the hazard function of the TIHLEtF distribution for different parameter values.

VII. Order Statistics

Let X_1, X_2, \dots, X_n be independent and identically distributed (i.i.d) random variables with their corresponding continuous distribution function $F(x)$. Let $X_{1:n} < X_{2:n} < \dots < X_{n:n}$ the corresponding ordered random sample from the TIHLEtF distributions. Let $F_{r:n}(x)$ and $f_{r:n}(x)$, $r = 1, 2, 3, \dots, n$ denote the CDF and PDF of the r^{th} order statistics $X_{r:n}$ respectively. The PDF of the r^{th} order statistics of $X_{r:n}$ is given as:

$$f_{r:n}(x; \lambda, \alpha, \theta) = \frac{f(x)}{B(r, n-r+1)} \sum_{v=0}^{n-r} (-1)^v \binom{n-r}{v} [F(x)]^{v+r-1} \quad (22)$$

The PDF of r^{th} order statistic for TIHLEtF distribution is derived by substituting equation (8) and equation (9) into equation (22). Also replacing h with $v+r-1$ in equation (9), so we

$$f_{r:n}(x; \lambda, \alpha, \theta, \delta) = \frac{1}{B(r, n-r+1)} \sum_{v=0}^{n-r} \sum_{i,j=0}^{\infty} \sum_{p,m=0}^{r+v-1} (-1)^v \binom{n-r}{v} \eta_p \varphi_t \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha(j+i+z)} \quad (23)$$

The PDF of minimum order statistic of the TIHLEtF distribution is obtained by setting $r = 1$ in equation (23) as

$$f_{1:n}(x; \lambda, \alpha, \theta, \delta) = 2n\lambda\alpha\delta\theta^\delta x^{-\delta-1} \sum_{v=0}^{n-1} \sum_{i,j=0}^{\infty} \sum_{p,m=0}^v (-1)^{i+j} (-1)^{p+m+z} (-1)^v \binom{n-1}{v} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j} \binom{v+p-1}{p} \binom{v}{m} \binom{\lambda(m+p)}{z} \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha(j+i+z)} \quad (24)$$

Also, the PDF of maximum order statistic of the TIHLEtF distribution is obtained by setting $r = n$ in equation (23) as

$$f_{n:n}(x; \lambda, \alpha, \theta, \delta) = 2n\lambda\alpha\delta\theta^\delta x^{-\delta-1} \sum_{i,j=0}^{\infty} \sum_{p,m=0}^{n+v-1} (-1)^{i+j} (-1)^{p+m+z} (-1)^v \binom{i+1}{i} \binom{\lambda(i+1)-1}{j} \binom{n+v+p-2}{p} \binom{n+v-1}{m} \binom{\lambda(m+p)}{z} \left[e^{-\left(\frac{\theta}{x}\right)^\delta} \right]^{\alpha(j+i+z)} \quad (25)$$

V. Parameter Estimation

Given complete data, we investigate the maximum likelihood method to estimate the TIHLEtF distribution's unknown parameters. Maximum likelihood estimates (MLEs) are attractive because they can be used to produce confidence intervals and offer straightforward approximations that work well in finite samples. The resulting approximation for MLEs is simple to handle in distribution theory, both analytically and numerically. Let $x_1, x_2, x_3, \dots, x_n$ be a random sample of size n from the TIHLEtF distribution. Then, the likelihood function based on the observed sample for the vector of the parameter $(\lambda, \alpha, \theta, \delta)^T$ is given by

$$\log L = n \log(2) + n \log(\lambda) + n \log(\alpha) + n \log(\delta) + n\delta \log(\theta) - (\delta - 1) \sum_{i=1}^n \log(x_i) - \alpha \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta + (\lambda - 1) \sum_{i=1}^n \log \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right] - 2 \sum_{i=1}^n \log \left[1 + \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^\lambda \right] \quad (26)$$

The components of the score vector $\Delta L(\phi) = \left(\frac{\partial L(\phi)}{\partial \lambda}, \frac{\partial L(\phi)}{\partial \alpha}, \frac{\partial L(\phi)}{\partial \theta}, \frac{\partial L(\phi)}{\partial \delta} \right)^T$ are given as

$$\frac{\partial \log L}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=1}^n \log \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right] - 2 \sum_{i=1}^n \frac{\left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^\lambda \log \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]}{1 + \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^\lambda} = 0 \quad (27)$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta + (\lambda - 1) \sum_{i=1}^n \frac{\left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \log \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]}{1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha} \quad (28)$$

$$+ 2 \sum_{i=1}^n \frac{\lambda \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^{\lambda-1} \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \log \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]}{1 + \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^\lambda} = 0$$

$$\begin{aligned} \frac{\partial \log L}{\partial \delta} &= \frac{n}{\delta} + n \log(\theta) + \sum_{i=1}^n \log(x_i) - \alpha \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^\delta \log\left(\frac{\theta}{x_i}\right) \\ &\quad + (\lambda - 1) \sum_{i=1}^n \frac{\alpha \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^\delta} \left(\frac{\theta}{x_i}\right)^\delta \log\left(\frac{\theta}{x_i}\right)}{1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha} \\ &\quad - 2 \sum_{i=1}^n \frac{\lambda \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^{\lambda-1} \alpha \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^\delta} \left(\frac{\theta}{x_i}\right)^\delta \log\left(\frac{\theta}{x_i}\right)}{1 + \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^\lambda} = 0 \end{aligned} \quad (29)$$

$$\begin{aligned} \frac{\partial \log L}{\partial \theta} = & \frac{n\delta}{\theta} - \alpha\delta \sum_{i=1}^n \left(\frac{\theta}{x_i}\right)^{\delta-1} \frac{1}{x_i} + (\lambda-1) \sum_{i=1}^n \frac{\alpha \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^\delta} \delta \left(\frac{\theta}{x_i}\right)^{\delta-1} \frac{1}{x_i}}{1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha} \\ & - 2 \sum_{i=1}^n \frac{\lambda \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^{\lambda-1} \alpha \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^\delta} \delta \left(\frac{\theta}{x_i}\right)^{\delta-1} \frac{1}{x_i}}{1 + \left[1 - \left[e^{-\left(\frac{\theta}{x_i}\right)^\delta} \right]^\alpha \right]^\lambda} = 0 \end{aligned} \tag{30}$$

The MLEs are obtained by setting $\frac{\partial L(\phi)}{\partial \lambda}$, $\frac{\partial L(\phi)}{\partial \alpha}$, $\frac{\partial L(\phi)}{\partial \theta}$ and $\frac{\partial L(\phi)}{\partial \delta}$ to zero and solving these equations simultaneously. These equations cannot be solved analytically, so we have to appeal to numerical method.

I. Simulation Study

In this section, a numerical analysis will be conducted to evaluate the performance of MLE for TIHLEtF Distribution.

Table 1: MLEs, biases and RMSE for some values of parameters

n	Parameters	(1,5,1,2,1,5)			(2,1,3,2,5)		
		Estimated Values	Bias	RMSE	Estimated Values	Bias	RMSE
20	λ	1.5650	0.0650	0.7475	2.1548	0.1548	0.9708
	α	1.0471	0.0471	0.5045	1.0385	0.0385	0.4666
	θ	2.0348	0.0348	0.3533	3.0465	0.0465	0.3439
	δ	1.7878	0.2878	0.7086	2.8275	0.3275	0.9009
50	λ	1.5507	0.0507	0.5664	2.1074	0.1074	0.7485
	α	1.0106	0.0106	0.3523	1.0073	0.0073	0.3337
	θ	2.0310	0.0310	0.2845	3.0360	0.0360	0.2589
	δ	1.6328	0.1328	0.4472	2.6418	0.1418	0.5924
100	λ	1.5251	0.0251	0.4358	2.0546	0.0546	0.5630
	α	1.0024	0.0024	0.2587	1.0024	0.0024	0.2520
	θ	2.0271	0.0271	0.2265	3.0456	0.0456	0.1937
	δ	1.5729	0.0729	0.2982	2.5830	0.0830	0.4288
250	λ	1.5391	0.0391	0.3064	2.0524	0.0524	0.3855
	α	1.0018	0.0018	0.1615	1.0011	0.0011	0.1426
	θ	2.0280	0.0280	0.1695	3.0313	0.0313	0.1479
	δ	1.5153	0.0153	0.1928	2.5129	0.0129	0.2808
500	λ	1.5269	0.0269	0.2273	2.0285	0.0285	0.2806
	α	1.0004	0.0004	0.1024	1.0002	0.0002	0.0983
	θ	2.0284	0.0284	0.1306	3.0251	0.0251	0.1162
	δ	1.5034	0.0034	0.1372	2.5062	0.0062	0.2051
1000	λ	1.5215	0.0215	0.1693	2.0275	0.0275	0.2059

α	1.0001	0.0001	0.0729	1.0000	0.0000	0.0693
θ	2.0182	0.0182	0.1014	3.0206	0.0206	0.0868
δ	1.5012	0.0012	0.1005	2.5013	0.0013	0.1446

The presented table demonstrates that the biases and RMSE values converge towards zero. As the sample size increases, the estimations approach the original (true) values, indicating the efficiency and reliability of the estimates.

VI. Applications to Real Data

We fit the TIHLEtF distribution to two real data sets and give a comparative study with the fits to the Exponentiated Half-Logistic Frechet (EHLF) distribution by [6], Kumaraswamy Exponentiated Frechet (KExF) distribution by [8], Gompertz Frechet (GoFr) distribution by [18], Exponentiated Frechet (ExFr) distribution by [16], and Frechet distribution by [10] as comparator distributions for illustrative purposes.

The EHLF distribution by [6]

$$f(x; \alpha, \lambda, \theta, \beta) = 2\alpha\lambda\theta\beta^\theta x^{-(\theta+1)} e^{-\left(\frac{\beta}{x}\right)^\theta} \left[1 - e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^{\lambda-1} \left[1 - \left[1 - e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^\lambda \right]^{\alpha-1} \left[1 + \left[1 - e^{-\left(\frac{\beta}{x}\right)^\theta} \right]^\lambda \right]^{-(\alpha+1)} \quad (31)$$

The KExF distribution by [8]

$$f(x; \alpha, \sigma, \lambda, \theta, \beta) = \alpha\lambda\sigma\beta\theta^\lambda x^{-(1+\lambda)} e^{-\left(\frac{\theta}{x}\right)^\lambda} \left[1 - e^{-\left(\frac{\theta}{x}\right)^\lambda} \right]^{\alpha-1} \left[1 - \left[1 - e^{-\left(\frac{\theta}{x}\right)^\lambda} \right]^\alpha \right]^{\sigma-1} \left[1 - \left[1 - \left[1 - e^{-\left(\frac{\theta}{x}\right)^\lambda} \right]^\alpha \right]^\sigma \right]^{\beta-1} \quad (32)$$

The GoFr distribution by [18]

$$f(x; \alpha, \lambda, \theta, \beta) = \theta\beta\alpha^\beta x^{-\beta-1} e^{-\left(\frac{\alpha}{x}\right)^\beta} \left[e^{-\left(\frac{\alpha}{x}\right)^\beta} \right]^{\lambda-1} e^{\left[\frac{\theta}{\lambda} \left(1 - \left[1 - e^{-\left(\frac{\alpha}{x}\right)^\beta} \right] \right) \right]^{-\lambda}} \quad (33)$$

The ExFr distribution by [16]

$$f(x; \alpha, \lambda, \sigma) = \alpha\lambda\sigma^\lambda \left[1 - e^{-\left(\frac{\sigma}{x}\right)^\lambda} \right]^{\alpha-1} x^{-(1+\lambda)} e^{-\left(\frac{\sigma}{x}\right)^\lambda} \quad (34)$$

The Fr distribution by [10]

$$f(x; \theta, \sigma) = \delta\theta^\sigma x^{-\sigma-1} e^{-\left(\frac{\theta}{x}\right)^\sigma} \quad (35)$$

The two datasets used to illustrate the application offer practical proof of the adaptability and appropriateness of the new proposed distribution. This distribution is shown to be the optimal

selection for modeling the datasets, distinct from the comparator distributions mentioned earlier. All calculations were conducted using the R programming language.

Data set 1

The first data set shown below represents the strength of carbon fibers tested under tension at gauge lengths of 10mm, previously used by [5]:

1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

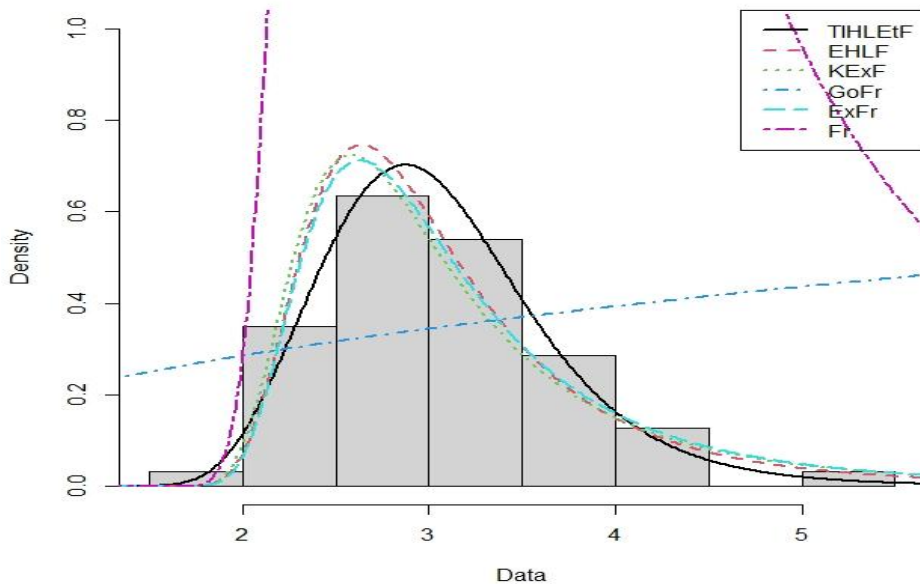


Figure 3: Fitted pdfs for the TIHLEtF, EHLF, KExF, GoFr, ExFr and Fr distributions to the data set 1

Table 2: The MLEs, log-likelihoods, and goodness of fit statistics of the models are based on the strength of carbon fibers tested under tension at gauge lengths of 10 mm (Data set 1).

Distributions	α	λ	θ	δ	β	LL	AIC
TIHLEtF	1.2456	3.7093	3.0756	3.1438		-56.4277	120.8555
EHLF	22.4675	0.2832	19.1580		1.3491	-58.9236	125.8472
KExF	2.4724	3.7429	2.0495	9.6236	0.4923	-60.2204	130.4407
GoFr	5.3750	2.7756	3.3750		6.3750	-67.2387	142.4774
ExFr	0.7937	5.8763		2.6529		-69.9499	145.8998
Fr			2.7215	5.4345		-79.5116	155.0232

Maximum likelihood was employed to estimate the parameters of both the newly developed distribution and five other comparable distributions. The outcomes are presented in Table 2. Extremely, the newly proposed distribution exhibited the lowest AIC value based on the goodness of fit measure, closely followed by EHLF. Furthermore, the superiority of the proposed distribution over its competitors is reinforced by visually assessing the fit, as illustrated in Figure 3. This underscores the fact that the newly recommended distribution is the most appropriate choice for accurately representing the carbon fiber dataset among the mentioned distributions.

Data set 2

The second data set shown below represents the survival times of one hundred and twenty-one (121) patients with breast cancer obtained from a large hospital in a period from 1929 to 1938, previously used by [19]:

0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 1.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0, 154.0.

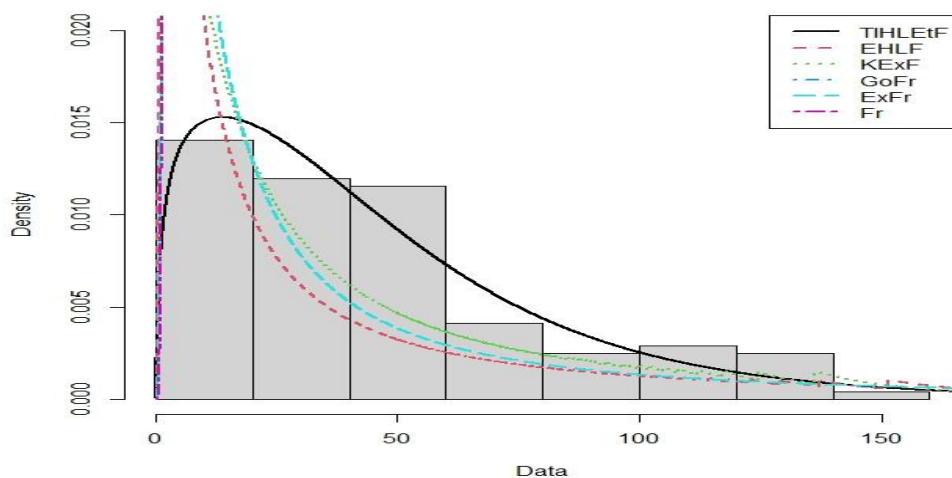


Figure 4: Fitted pdfs for the TIHLEtF, EHLF, KExF, GoFr, ExFr and Fr distributions to the data set 2

Table 3: The MLEs, log-likelihoods, and goodness of fit statistics of the models based on the survival time of patients with breast cancer (Data 2).

Distributions	α	λ	θ	δ	β	LL	AIC
TIHLEtF	4.6372	6.6628	3.7129	0.4224		- 599.0163	1206.033
EHLF	6.0551	0.1075	5.2159		0.1429	- 644.6529	1297.306
KExF	0.0343	5.5835	0.2036	5.9722	13.1474	- 619.2298	1248.46
GoFr	41.0479	59.0871	0.2511		0.5019	-1311.34	2630.68
ExFr	1.3538	0.6216		18.0606		- 632.1718	1270.344
Fr			16.8934	0.6524		-695.0586	1394.117

The parameters of the TIHLEtF distribution and five similar distributions were estimated using the maximum likelihood method, and the outcomes are detailed in Table 3. It's worth highlighting that the new distribution displayed the lowest AIC value when compared to the other options, which signifies that it is the most appropriate choice for modeling the survival time of patients with breast cancer based on the goodness of fit measure AIC. Additionally, a visual examination of the fit, as illustrated in Figure 4, reinforces the idea that the new distribution outperforms its rivals.

VII. CONCLUSION

In our research, we have introduced and explored a novel statistical distribution known as the Type I Half-Logistic Exponentiated Frechet Distribution, building upon the distribution family introduced by [4]. This study delved into various statistical aspects of this newly introduced distribution. These aspects encompassed explicit quantile functions, probability-weighted moments, moments, moments generating functions, reliability functions, hazard functions, and order statistics. For parameter estimation, we utilized the maximum likelihood method. Our analysis also included simulations to assess the effectiveness of this newly proposed distribution. To underscore the significance and versatility of this innovative distribution, we compared it with well-established models using two authentic datasets. The outcomes of our investigation underscore the superiority of the new distribution in comparison to the other models under consideration. This implies that the proposed distribution shows promise for effectively modeling data across various applications.

References

- [1] Afify, A. Z., Hamedani, G. G., Ghosh, I. & Mead, M. E. (2015). The transmuted Marshall-Olkin Frechet distribution: properties and applications. *International Journal of Statistics and Probability*, 4, 132-184.
- [2] Afify, A. Z., Yousof, H. M., Cordeiro, G. M., M. Ortega, E. M., & Nofal, Z. M. (2016). The Weibull Frechet distribution and its applications. *Journal of Applied Statistics*, 43(14), 2608-2626.
- [3] Afify, A. Z., Yousof, H. M., Cordeiro, G. M., Nofal, Z. M., & Ahmad, M. (2016). The Kumaraswamy Marshall Olkin Frechet Distribution with Applications. *Journal of ISOSS*, 2(2), 151-168.
- [4] Bello, O. A., Doguwa, S. I., Yahaya, A., & Jibril, H. M. (2021). A Type I Half Logistic Exponentiated-G Family of Distributions: Properties and Application. *Communication in Physical Sciences*, 7(3).
- [5] Bi, Q., & Gui, W. (2017). Bayesian and classical estimation of stress-strength reliability for inverse Weibull lifetime models. *Algorithms*, 10(2), 71.
- [6] Cordeiro, G. M., Alizadeh, M. & Ortega, E. M. M. (2014a). The exponentiated half logistic family of distributions: Properties and applications. *Journal of Probability and Statistics*, 1-21.
- [7] Da Silva, R. V., de Andrade, T. A., Maciel, D. B., Campos, R. P., & Cordeiro, G. M. (2013). A New Lifetime Model: The Gamma Extended Frechet Distribution. *Journal of Statistical Theory and Applications*, 12(1), 39-54.
- [8] Diab, L. S., & Elbatal, I. (2016). A new generalization of exponentiated Frechet distribution. *International Journal of Reliability and Applications*, 17(1), 65-84.
- [9] Elbatal, I., Asha, G., & Raja, A. V. (2014). Transmuted exponentiated Frechet distribution: properties and applications. *Journal of Statistics Applications & Probability*, 3(3), 379.
- [10] Frechet, M. (1924). Sur la Loi des Erreurs d'Observation. *Bulletin de la Soci et e Math ematique de Moscou*, 33(1), 5-8.
- [11] Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water resources research*, 15(5), 1049-1054.

- [12] Krishna, E., Jose, K. K., Alice, T., & Ristic, M. M. (2013). The Marshall-Olkin Frechet distribution. *Communications in Statistics-Theory and Methods*, 42(22), 4091-4107.
- [13] Mahmoud, M. R., & Mandouh, R. M. (2013). On the transmuted Fréchet distribution. *Journal of Applied Sciences Research*, 9(10), 5553-5561.
- [14] Mead, M. E. & Abd-Eltawab A. R. (2014). A note on Kumaraswamy Frechet distribution. *Australian Journal of Basic and Applied Sciences*, 8, 294-300.
- [15] Mead, M. E., Afify, A. Z., Hamedani, G. G., & Ghosh, I. (2017). The beta exponential Frechet distribution with applications. *Austrian Journal of Statistics*, 46(1), 41-63.
- [16] Nadaraja, S., & Kotz, S. (2003). The Exponentiated Frechet distribution, *Interstat Electronic Journal*, 1-7.
- [17] Nadarajah, S., & Gupta, A. K. (2004). The beta Frechet distribution. *Fareast journal of theoretical statistics*, 14(1), 15-24.
- [18] Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., & Okagbue, H. I. (2019). The Gompertz Frechet Distribution: Properties and Applications. *Cogent Mathematics & Statistics*, 6(1), 1568662.
- [19] Ramos, M.W. A., Marinho, P. R. D., Da Silva, R. V., & Cordeiro, G. M. (2013). The exponentiated Lomax Poisson distribution with an application to lifetime data. *Advances and Applications in Statistics*, 34(2), 107-135.

BAYESIAN ANALYSIS OF EXTENDED MAXWELL-BOLTZMANN DISTRIBUTION USING SIMULATED AND REAL-LIFE DATA SETS

¹NUZHAT AHAD; ²S.P.AHMAD; ^{3*}J.A.RESHI

•

^{1,2} Department of Statistics, University of Kashmir, Srinagar, India

^{3*} Department of Statistics, Govt. Degree College Pulwama, Srinagar, India

¹nuzhatahad01@gmail.com, ²sprvz@yahoo.com, ^{3*}reshijavaid19@gmail.com

Abstract

The objective of the study is to use Bayesian techniques to estimate the scale parameter of the 2Kth order weighted Maxwell-Boltzmann distribution (KWMBD). This involved using various prior assumptions such as extended Jeffrey's, Hartigan's, Inverse-gamma and Inverse-exponential, as well as different loss functions including squared error loss function (SELF), precautionary loss function (PLF), Al Bayyati's loss function (ALBF), and Stein's Loss Function (SLF). The maximum likelihood estimation (MLE) is also obtained. We compared the performances of MLE and Bayesian estimation under each prior and its associated loss functions. And demonstrated the effectiveness of Bayesian estimation through simulation studies and analyzing real-life datasets.

Keywords: 2Kth Order Weighted Maxwell-Boltzmann Distribution, Prior Distribution, Loss Function and Bayesian estimation.

1. INTRODUCTION

The Maxwell-Boltzmann distribution, characterizes the probability distribution of speeds for particles in a gas at various temperatures. It provides a statistical framework for understanding the distribution of kinetic energies among particles, which makes it vital for modeling physical systems and predicting their behavior. Because of its practical significance, scientists and engineers closely examine the Maxwell-Boltzmann distribution to attain a deeper understanding of various scientific phenomena and to create precise models of complex systems. Tyagi and Bhattacharya [15] were the first to explore the Maxwell distribution as a lifetime model, and introduced considerations of Bayesian and minimum variance unbiased estimation methods for determining its parameters and reliability function. Chaturvedi and Rani [6] derived classical and Bayesian estimators for the Maxwell distribution by extending it with an additional parameter. Various Statisticians and Mathematicians have carried out the Bayesian paradigm of Maxwell-Boltzmann distribution by using loss functions and prior distributions, See, Spiring and Yeung [14], Rasheed [11], Reshi [13], and Ahmad and Tripathi [1].

The 2Kth order weighted Maxwell-Boltzmann distribution (KWMBD) is a flexible, symmetric continuous univariate probability distribution suitable for modelling datasets of decreasing-increasing, increasing and constant behaviour. The probability density function (pdf) of KWMBD is given by:

$$f(x) = \frac{x^{2(k+1)}\alpha^{-(3+2k)}e^{-\frac{x^2}{2\alpha^2}}}{2^{k+\frac{1}{2}}\Gamma(k+\frac{3}{2})} \quad x > 0, \alpha > 0, k \in R. \quad (1)$$

And, the corresponding cumulative distribution function (cdf) of KWMBD is given by:

$$F(x) = 1 - \frac{\Gamma\left(k+\frac{3}{2}, \frac{x^2}{2\alpha^2}\right)}{\Gamma(k+\frac{3}{2})} \quad x > 0, \alpha > 0, k \in R. \quad (2)$$

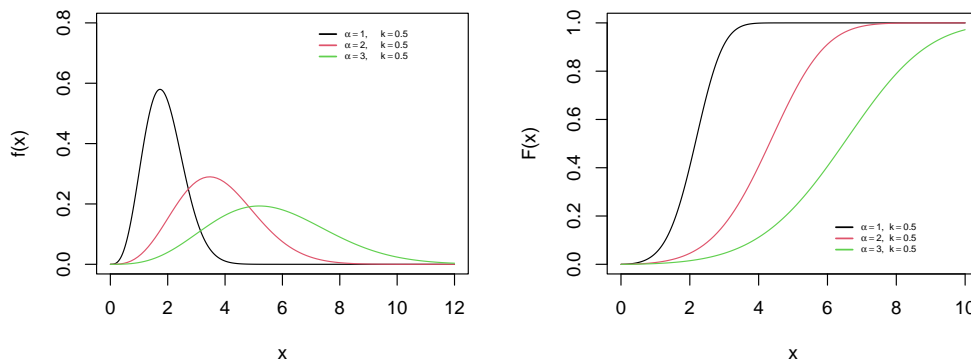


Figure 1: Probability density plot and cumulative distribution plot of KWMBD for different combinations of parameters.

2. METHODOLOGICAL PROCEDURE

Bayesian approach utilizes prior beliefs, observed data, and a loss function to make decision in a structured manner, and is considered more reliable for estimating distribution parameters. Compared to the classical approach, especially when the prior distribution accurately represents the parameter's random behavior. In Bayesian analysis, parameters are treated as uncertain variables, allowing prior knowledge to be incorporated into the analysis. This prior information is typically described using a probability distribution known as the prior distribution. Friesl and Hurt[7] noted that employing Bayesian theory is a viable approach for incorporating prior information into the model, potentially improving the inference process and reflects the parameter's behavior. However, there are no strict rules for choosing one prior over another, frequently, prior distributions are selected based on an individual's subjective knowledge and beliefs. When sufficient information about the parameter is available, informative priors are preferred; otherwise, non-informative priors, such as the uniform prior, are used. Aslam [4] demonstrated the application of prior predictive distribution for determining the prior density. In this study, we assume the parameter α follows an extension of Jeffrey's prior proposed by Al-Kutobi[3] and α^2 follows an inverse-gamma prior and are given by:

2.1. Extension of Jeffrey's prior

The prior, known as extension of Jeffrey's prior is given by:

$$g(\alpha) = [I(\alpha)]^{c_1}; \quad c_1 \in R^+$$

where $e, I(\alpha) = -nE\left\{\frac{d^2}{d\alpha^2} \log f(x)\right\}$ is fisher information matrix.

Thus, the resulting extension of Jeffrey's-prior for KWMBD will be:

$$g(\alpha) = \left[\frac{1}{\alpha^2}\right]^{c_1}; \quad c_1 \in R^+ \quad (3)$$

2.2. Inverse-gamma prior

The density of parameter α^2 on assuming it to follow Gamma (β, λ) distribution is given by:

$$g(\alpha^2) = \frac{\lambda^\beta}{\Gamma(\beta)} (\alpha^2)^{-\beta-1} e^{-\frac{\lambda}{\alpha^2}} \quad (4)$$

2.3. Loss functions

The idea of loss functions had been introduced first by Laplace, and later during the mid-20th century it was reintroduced by Weiss [16]. Loss function, serves as a measure of the discrepancy between observed data and the values predicted by a statistical model. Decisions in Bayesian inference, apart from relying on experimental data, are not entirely controlled by the loss function. Moreover, the relationship between the loss function and the posterior probability is significant. The choice of a loss function depends on the specific characteristics of the data and the goals of the analysis. Han [9] pointed out that, in Bayesian analysis choosing the right loss function and prior distribution is essential for making accurate statistical inferences. The Bayesian estimator is directly impacted by the choice of loss function, while the parameters of the prior density function may be affected by hyper parameters. Various symmetric and asymmetric loss functions have been demonstrated to be effective in research conducted by Zellner [17], Reshi [12], and Ahmad [2], among others. In this study, we have explored squared error, precautionary, Al-Bayyati's, and Stein's loss functions to enhance the comparison of Baye's estimators. And are given by:

2.3.1. Squared error loss function

The squared error loss function is given by:

$$l_{sq}(\hat{\alpha}, \alpha) = c(\hat{\alpha} - \alpha)^2; \quad c \in R^+ \quad (5)$$

2.3.2. Precautionary loss function

The Precautionary loss function is given by:

$$l_{pr}(\hat{\alpha}, \alpha) = \frac{c(\hat{\alpha} - \alpha)^2}{\hat{\alpha}} \quad (6)$$

2.3.3. Al-Bayyati's loss function

The Al-Bayyati's loss function is given by:

$$l_{AI}(\hat{\alpha}, \alpha) = \alpha^{c_2} (\hat{\alpha} - \alpha)^2; \quad c_2 \in R^+ \quad (7)$$

2.3.4. Stein's loss function

The Stein's loss function is given by:

$$l_{St}(\hat{\alpha}, \alpha) = \frac{\hat{\alpha}}{\alpha} - \log\left(\frac{\hat{\alpha}}{\alpha}\right) - 1 \quad (8)$$

3. PARAMETRIC ESTIMATION OF KWMBD

In this section, we discuss the various estimation methods for KWMB Distribution.

3.1. Maximum Likelihood Estimation

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample of size n from k th Order Weighted Maxwell-Boltzmann Distribution. Therefore the maximum likelihood estimator (MLE) of α is:

$$\hat{\alpha} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n(2k+3)}} \quad (9)$$

3.2. Baye's Estimator under Extension of Jeffrey's Prior

The Joint Probability Density Function of x and given α is given by:

$$L(\underline{x}|\alpha) = \frac{\prod_{i=1}^n x_i^{2(k+1)} \alpha^{-n(3+2k)} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}}}{\left(2^{k+\frac{1}{2}}\right)^n \left(\Gamma\left(k+\frac{3}{2}\right)\right)^n} \quad (10)$$

The posterior probability density function of α for given data x is given by:

$$\begin{aligned} \pi_1(\alpha|\underline{x}) &\propto L(\underline{x}|\alpha)g(\alpha) \\ \pi_1(\alpha|\underline{x}) &\propto \frac{\prod_{i=1}^n x_i^{2(k+1)} \alpha^{-n(3+2k)} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}}}{\left(2^{k+\frac{1}{2}}\right)^n \left(\Gamma\left(k+\frac{3}{2}\right)\right)^n} \frac{1}{\alpha^{2c_1}} \\ \pi_1(\alpha|\underline{x}) &= k\alpha^{-n(3+2k)-2c_1} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}} \end{aligned}$$

where k is normalising constant independent of α and is given by:

$$\begin{aligned} k^{-1} &= \int_0^\infty \alpha^{-n(3+2k)-2c_1} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}} d\alpha \\ k^{-1} &= \frac{\left(\sum_{i=1}^n x_i^2\right)^{\frac{-n(3+2k)-2c_1+1}{2}} \Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}{2^{\frac{-n(3+2k)-2c_1+3}{2}}} \end{aligned}$$

Therefore, the posterior probability density function is:

$$\pi_1(\alpha|\underline{x}) = \frac{2^{\frac{-n(3+2k)-2c_1+3}{2}} \alpha^{-n(3+2k)-2c_1} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}}}{\left(\sum_{i=1}^n x_i^2\right)^{\frac{-n(3+2k)-2c_1+1}{2}} \Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)} \quad (11)$$

3.2.1. Baye's Estimator under squared error loss function

The Risk Function Under SELF is given by:

$$R_{(sq,ej)}(\hat{\alpha}) = \int_0^{\infty} c(\hat{\alpha} - \alpha)^2 \pi_1(\alpha|\underline{x}) d\alpha$$

$$R_{(sq,ej)}(\hat{\alpha}) = c\hat{\alpha}^2 + \frac{\sum_{i=1}^n x_i^2}{(n(3+2k)+2c_1-3)} - 2\hat{\alpha}c \sqrt{\frac{\sum_{i=1}^n x_i^2}{2} \frac{\Gamma\left(\frac{n(3+2k)+2c_1-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}} \quad (12)$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(sq,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(ej,sq)} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{2} \frac{\Gamma\left(\frac{n(3+2k)+2c_1-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}} \quad (13)$$

3.2.2. Baye's Estimator under precautionary Loss function

The Risk Function Under PLF is given by:

$$R_{(pre,ej)}(\hat{\alpha}) = \int_0^{\infty} c \frac{(\hat{\alpha} - \alpha)^2}{\hat{\alpha}} \pi_1(\alpha|\underline{x}) d\alpha$$

$$R_{(pre,ej)}(\hat{\alpha}) = c\hat{\alpha} + c \frac{\sum_{i=1}^n x_i^2}{\hat{\alpha}(n(3+2k)+2c_1-3)} - 2c \sqrt{\frac{\sum_{i=1}^n x_i^2}{2} \frac{\Gamma\left(\frac{n(3+2k)+2c_1-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}} \quad (14)$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(pre,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(pre,ej)} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{(n(3+2k)+2c_1-3)}} \quad (15)$$

3.2.3. Baye's Estimator under Al-Bayyati's loss function

The Risk Function Under Al-Bayyati's loss function is given by:

$$R_{(alb,ej)}(\hat{\alpha}) = \int_0^{\infty} \alpha^{c_2} (\hat{\alpha} - \alpha)^2 \pi_1(\alpha|\underline{x}) d\alpha$$

$$R_{(alb,ej)}(\hat{\alpha}) = \hat{\alpha}^2 \left(\frac{\sum_{i=1}^n x_i^2}{2} \right)^{\frac{c_2}{2}} + \left(\frac{\sum_{i=1}^n x_i^2}{2} \right)^{\frac{c_2-2}{2}} \frac{1}{(n(3+2k)+2c_1-c_2-3)} - 2\hat{\alpha} \left(\frac{\sum_{i=1}^n x_i^2}{2} \right)^{\frac{c_2+1}{2}} \frac{\Gamma\left(\frac{n(3+2k)+2c_1-c_2-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-c_2-1}{2}\right)} \quad (16)$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(alb,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(alb,ej)} = \sqrt{\frac{\sum_{i=1}^n x_i^2 \Gamma\left(\frac{n(3+2k)+2c_1-c_2-2}{2}\right)}{2 \Gamma\left(\frac{n(3+2k)+2c_1-c_2-1}{2}\right)}} \quad (17)$$

3.2.4. Baye's Estimator under combination of Stein's loss function

The Risk Function Under SLF is given by:

$$R_{(ste,ej)}(\hat{\alpha}) = \int_0^{\infty} \left(\frac{\hat{\alpha}}{\alpha} - \log\left(\frac{\hat{\alpha}}{\alpha}\right) - 1 \right) \pi_1(\alpha|\underline{x}) d\alpha$$

$$R_{(ste,ej)}(\hat{\alpha}) = \hat{\alpha} \sqrt{\frac{2 \Gamma\left(\frac{n(3+2k)+2c_1}{2}\right)}{\sum_{i=1}^n x_i^2 \Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}} - \log(\hat{\alpha}) - m - 1 \quad (18)$$

where e, m is constant of integration.

Now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(ste,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(ste,ej)} = \sqrt{\frac{\sum_{i=1}^n x_i^2 \Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1}{2}\right)}} \quad (19)$$

3.3. Baye's Estimator under Inverse-Gamma Prior

The Joint Probability Density Function of x and given α^2 is given by:

$$L(\underline{x}|\alpha^2) = \frac{\prod_{i=1}^n x_i^{2(k+1)} (\alpha^2)^{-\frac{n(3+2k)}{2}} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}}}{\left(2^{k+\frac{1}{2}}\right)^n \left(\Gamma\left(k+\frac{3}{2}\right)\right)^n} \quad (20)$$

The posterior probability density function of α^2 for given data x is given by:

$$\pi_2(\alpha^2|\underline{x}) \propto L(\underline{x}|\alpha^2)g(\alpha^2)$$

$$\pi_2(\alpha^2|\underline{x}) \propto \frac{\prod_{i=1}^n x_i^{2(k+1)} (\alpha^2)^{-\frac{n(3+2k)}{2}} e^{-\frac{\sum_{i=1}^n x_i^2}{2\alpha^2}}}{\left(2^{k+\frac{1}{2}}\right)^n \left(\Gamma\left(k+\frac{3}{2}\right)\right)^n} \frac{\lambda^\beta}{\Gamma(\beta)} (\alpha^2)^{-\beta-1} e^{-\frac{\lambda}{\alpha^2}}$$

$$\pi_2(\alpha^2|\underline{x}) = k(\alpha^2)^{-\frac{n(3+2k)-2\beta-2}{2}} e^{-\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right) \frac{1}{\alpha^2}}$$

where k is normalising constant independent of α and is given by:

$$k^{-1} = \int_0^{\infty} (\alpha^2)^{-\frac{n(3+2k)-2\beta-2}{2}} e^{-\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right) \frac{1}{\alpha^2}} d\alpha^2$$

$$k^{-1} = \frac{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)}{\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)^{\frac{n(3+2k)+2\beta}{2}}}$$

Therefore, the posterior probability density function is:

$$\pi_2(\alpha^2 | \underline{x}) = \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)^{\frac{n(3+2k)+2\beta}{2}} (\alpha^2)^{-\frac{n(3+2k)-2\beta-2}{2}} e^{-\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right) \frac{1}{\alpha^2}}}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} \quad (21)$$

3.3.1. Baye's Estimator under squared error loss function

The Risk Function Under SELF is given by:

$$R_{(sq,igp)}(\hat{\alpha}^2) = \int_0^{\infty} c(\hat{\alpha}^2 - \alpha^2)^2 \pi_2(\alpha^2 | \underline{x}) d(\alpha^2)$$

$$R_{(sq,igp)}(\hat{\alpha}^2) = c(\hat{\alpha}^2)^2 + \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)^2}{\left(\frac{n(3+2k)+2\beta-2}{2}\right) \left(\frac{n(3+2k)+2\beta-4}{2}\right)} - \hat{\alpha}^2 c \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)}{\left(\frac{n(3+2k)+2\beta-2}{2}\right)} \quad (22)$$

now, the Baye's estimator is obtained by solving

$$\frac{R_{(sq,igp)}(\hat{\alpha}^2)}{d(\hat{\alpha}^2)} = 0$$

and, is given by:

$$\hat{\alpha}_{(sq,igp)} = \sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)}{(n(3+2k) + 2\beta - 2)}} \quad (23)$$

3.3.2. Baye's Estimator under precautionary Loss function

The Risk Function Under PLF is given by:

$$R_{(pre,igp)}(\hat{\alpha}^2) = \int_0^{\infty} c \frac{(\hat{\alpha}^2 - \alpha^2)^2}{\hat{\alpha}^2} \pi_2(\alpha^2 | \underline{x}) d\alpha^2$$

$$R_{(pre,igp)}(\hat{\alpha}^2) = c\hat{\alpha}^2 + c \frac{1}{\hat{\alpha}^2} \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)^2}{\left(\frac{(n(3+2k)+2\beta-2)(n(3+2k)+2\beta-4)}{4}\right)} - 2c \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda\right)}{\left(\frac{n(3+2k)+2\beta-2}{2}\right)} \quad (24)$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(pre,igp)}(\hat{\alpha}^2))}{d\hat{\alpha}^2} = 0$$

and, is given by:

$$\hat{\alpha}_{(pre,igp)} = \sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{\sqrt{(n(3+2k)+2\beta-2)(n(3+2k)+2\beta-4)}}} \quad (25)$$

3.3.3. Baye's Estimator under Al-Bayyati's loss function

The Risk Function Under Al-Bayyati's loss function is given by:

$$R_{(alb,igp)}(\hat{\alpha}^2) = \int_0^{\infty} (\alpha^2)^{c_2} (\hat{\alpha}^2 - \alpha^2)^2 \pi_2(\alpha^2 | \underline{x}) d\alpha^2$$

$$R_{(alb,igp)}(\hat{\alpha}^2) = (\hat{\alpha}^2)^4 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)^{c_2} \frac{\Gamma\left(\frac{n(3+2k)+2\beta-2c_2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} + \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)^{c_2+2} \frac{\Gamma\left(\frac{n(3+2k)+2\beta-2c_2-4}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} - 2\hat{\alpha}^2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)^{c_2+1} \frac{\Gamma\left(\frac{n(3+2k)+2\beta-2c_2-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} \quad (26)$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(alb,igp)}(\hat{\alpha}^2))}{d\hat{\alpha}^2} = 0$$

and, is given by:

$$\hat{\alpha}_{(alb,igp)} = \sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{(n(3+2k)+2\beta-2)}} \quad (27)$$

3.3.4. Baye's Estimator under combination of Stein's loss function

The Risk Function Under SLF is given by:

$$R_{(s,igp)}(\hat{\alpha}^2) = \int_0^{\infty} \left(\frac{\hat{\alpha}^2}{\alpha^2} - \log\left(\frac{\hat{\alpha}^2}{\alpha^2}\right) - 1 \right) \pi_2(\alpha^2 | \underline{x}) d\alpha^2$$

$$R_{(s,igp)}(\hat{\alpha}^2) = \hat{\alpha}^2 \frac{(n(3+2k)+2\beta)}{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)} - \log(\hat{\alpha}) - m - 1 \quad (28)$$

where, m is constant of integration.

Now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(s,igp)}(\hat{\alpha}^2))}{d\hat{\alpha}^2} = 0$$

and, is given by:

$$\hat{\alpha}_{(ste,igp)} = \sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{(n(3+2k)+2\beta)}} \quad (29)$$

Table 1: Baye's Estimation under Hartigan's Prior Distribution and Different Combinations of Loss Functions.

Prior	Loss Function	Baye's Estimator
Hartigan's (i.e. $c_1 = 3/2$)	Squared-error	$\sqrt{\frac{\sum_{i=1}^n x_i^2}{2} \frac{\Gamma\left(\frac{n(3+2k)+1}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2}{2}\right)}}$
	Precautionary	$\sqrt{\frac{\sum_{i=1}^n x_i^2}{n(3+2k)}}$
	Al-Bayyati's	$\sqrt{\frac{\sum_{i=1}^n x_i^2}{2} \frac{\Gamma\left(\frac{n(3+2k)-c_2+1}{2}\right)}{\Gamma\left(\frac{n(3+2k)-c_2+1}{2}\right)}}$
	Stein's	$\sqrt{\frac{\sum_{i=1}^n x_i^2}{2} \frac{\Gamma\left(\frac{n(3+2k)+2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+3}{2}\right)}}$

Table 2: Baye's Estimation under Inverse-Exponential Prior Distributions and Different Combinations of Loss Functions.

Prior	Loss Function	Baye's Estimator
Inverse-Exponential (i.e. $\beta = 1$)	Squared-error	$\sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{n(3+2k)}}$
	Precautionary	$\sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{\sqrt{(n(3+2k))(n(3+2k)-2)}}}$
	Al-Bayyati's	$\sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{n(3+2k)}}$
	Stein's	$\sqrt{\frac{2 \left(\frac{\sum_{i=1}^n x_i^2}{2} + \lambda \right)}{n(3+2k)+2}}$

3.4. Simulation Study

We conducted simulation studies using R software, generated samples of sizes $n=10, 50,$ and 100 to observe the effect of small, medium, and large samples on the estimators of scale parameter α of the $2k$ th order weighted Maxwell Boltzmann distribution. Each process is replicated 500 times to examine the performance of the MLEs and Bayesian estimators under different priors such as the extension of Jeffrey's prior, Hartigan's prior, inverse-Gamma prior, and inverse-exponential prior, across different loss functions in terms of average estimates, biases, variances, and mean squared errors by considering different parameter combinations. The results are presented in the tables below:

Table 3: Average estimate, Bias, Variance and Mean Squared Error under Extension of Jeffrey's prior.

n	α	k	c_1	c_2	Criterion	$\hat{\alpha}_{mle}$	$\hat{\alpha}_{sq}$	$\hat{\alpha}_{pre}$	$\hat{\alpha}_{alb}$	$\hat{\alpha}_{ste}$
10	3	-0.5	2	5	Estimate	2.97912	2.87293	2.90732	3.27915	2.80839
					Bias	-0.02088	-0.12707	-0.09268	0.27915	-0.19161
					Variance	0.23825	0.22157	0.22691	0.28866	0.21173
					MSE	0.23869	0.23772	0.23549	0.36658	0.24844
50	3	-0.5	2	5	Estimate	3.00890	2.98656	2.99396	3.06295	2.97196
					Bias	0.00890	-0.01344	-0.00604	0.06295	-0.02804
					Variance	0.04693	0.04624	0.04647	0.04864	0.04579
					MSE	0.04701	0.04642	0.04645	0.05260	0.04658
100	3	-0.5	2	5	Estimate	3.00411	2.99291	2.99663	3.03075	2.98551
					Bias	0.00411	-0.00709	-0.00337	0.03075	-0.01449
					Variance	0.02164	0.02148	0.02153	0.02203	0.02137
					MSE	0.02166	0.02153	0.02154	0.02297	0.02158
10	4	0.1	1.2	3	Estimate	3.96513	3.94644	3.97758	4.14350	3.88675
					Bias	-0.03487	-0.05356	-0.02242	0.14350	-0.11325
					Variance	0.25397	0.25158	0.25557	0.27733	0.24403
					MSE	0.25518	0.25445	0.25620	0.29792	0.25685
50	4	0.1	1.2	3	Estimate	4.00312	3.99937	4.00563	4.03732	3.98695
					Bias	0.00312	-0.00063	0.00563	0.03732	-0.01305
					Variance	0.05165	0.05155	0.05172	0.05254	0.05123
					MSE	0.05166	0.05155	0.05175	0.05393	0.05140
100	4	0.1	1.2	3	Estimate	3.99978	3.99790	4.00103	4.01676	3.99168
					Bias	-0.00022	-0.00210	0.00103	0.01676	-0.00832
					Variance	0.02381	0.02379	0.02382	0.02401	0.02371
					MSE	0.02381	0.02379	0.02382	0.02429	0.02378

Table 4: Average estimate, Bias, Variance and Mean Squared Error under Hartigan's prior.

n	α	k	c_1	c_2	Criterion	$\hat{\alpha}_{mle}$	$\hat{\alpha}_{sq}$	$\hat{\alpha}_{pre}$	$\hat{\alpha}_{alb}$	$\hat{\alpha}_{ste}$
10	3	-0.5	1.5	5	Estimate	2.98117	2.94416	2.98117	3.38551	2.87491
					Bias	-0.01883	-0.05584	-0.01883	0.38551	-0.12509
					Variance	0.20672	0.20162	0.20672	0.26660	0.19225
					MSE	0.20708	0.20474	0.20708	0.41521	0.20789
50	3	-0.5	1.5	5	Estimate	2.99573	2.98825	2.99573	3.06548	2.97350
					Bias	-0.00427	-0.01175	-0.00427	0.06548	-0.02650
					Variance	0.04357	0.04335	0.04357	0.04562	0.04292
					MSE	0.04359	0.04349	0.04359	0.04991	0.04363
100	3	-0.5	1.5	5	Estimate	2.99912	2.99537	2.99912	3.03344	2.98793
					Bias	-0.00088	-0.00463	-0.00088	0.03344	-0.01207
					Variance	0.02168	0.02163	0.02168	0.02218	0.02152
					MSE	0.02168	0.02165	0.02168	0.02330	0.02167
10	4	0.1	1.5	3	Estimate	3.96310	3.93226	3.96310	4.12731	3.87314
					Bias	-0.03690	-0.06774	-0.03690	0.12731	-0.12686
					Variance	0.25001	0.24614	0.25001	0.27116	0.23879
					MSE	0.25137	0.25073	0.25137	0.28737	0.25489
50	4	0.1	1.5	3	Estimate	3.99271	3.98647	3.99271	4.02426	3.97411
					Bias	-0.00729	-0.01353	-0.00729	0.02426	-0.02589
					Variance	0.04954	0.04939	0.04954	0.05033	0.04908
					MSE	0.04959	0.04957	0.04959	0.05092	0.04975
100	4	0.1	1.5	3	Estimate	4.00132	3.99819	4.00132	4.01704	3.99197
					Bias	0.00132	-0.00181	0.00132	0.01704	-0.00803
					Variance	0.02222	0.02219	0.02222	0.02240	0.02212
					MSE	0.02222	0.02219	0.02222	0.02269	0.02218

$\hat{\alpha}_{mle}$ = Estimate under maximum likelihood estimation, $\hat{\alpha}_{sq}$ = Bayes estimate under squared error loss function, $\hat{\alpha}_{pre}$ = Bayes estimate under precautionary loss function, $\hat{\alpha}_{alb}$ = Bayes estimate under Al-Bayyati's loss function, $\hat{\alpha}_{ste}$ = Bayes estimate under Stein's loss function.

Table 5: Average estimate, Bias, Variance and Mean Squared Error under Inverse-Gamma prior.

n	α	k	β	λ	c_2	Criterion	$\hat{\alpha}_{mle}$	$\hat{\alpha}_{sq}$	$\hat{\alpha}_{pre}$	$\hat{\alpha}_{alb}$	$\hat{\alpha}_{ste}$
10	3	-0.5	1.5	3.5	5	Estimate	2.96815	2.95500	3.02987	4.08292	2.82360
						Bias	-0.03185	-0.04500	0.02987	1.08292	-0.17640
						Variance	0.22183	0.20296	0.21337	0.38746	0.18531
						MSE	0.22284	0.20498	0.21426	1.56018	0.21642
50	3	-0.5	1.5	3.5	5	Estimate	3.01586	3.01248	3.02758	3.17368	2.98309
						Bias	0.01586	0.01248	0.02758	0.17368	-0.01691
						Variance	0.04434	0.04357	0.04400	0.04835	0.04272
						MSE	0.04459	0.04372	0.04476	0.07852	0.04301
100	3	-0.5	1.5	3.5	5	Estimate	3.00761	3.00593	3.01346	3.08362	2.99109
						Bias	0.00761	0.00593	0.01346	0.08362	-0.00891
						Variance	0.02039	0.02021	0.02031	0.02127	0.02001
						MSE	0.02045	0.02024	0.02049	0.02826	0.02009
10	4	0.1	1.2	3	3	Estimate	3.96991	3.96910	4.03283	4.39706	3.85199
						Bias	-0.03009	-0.03090	0.03283	0.39706	-0.14801
						Variance	0.24080	0.23493	0.24254	0.28833	0.22127
						MSE	0.24171	0.23589	0.24361	0.44598	0.24318
50	4	0.1	1.2	3	3	Estimate	3.97652	3.97628	3.98878	4.05281	3.95172
						Bias	-0.02348	-0.02372	-0.01122	0.05281	-0.04828
						Variance	0.05037	0.05013	0.05044	0.05207	0.04951
						MSE	0.05092	0.05069	0.05056	0.05486	0.05184
100	4	0.1	1.2	3	3	Estimate	4.00210	4.00195	4.00822	4.03995	3.98952
						Bias	0.00210	0.00195	0.00822	0.03995	-0.01048
						Variance	0.02531	0.02525	0.02533	0.02573	0.02509
						MSE	0.02531	0.02525	0.02534	0.02733	0.02520

Table 6: Average estimate, Bias, Variance and Mean Squared Error under Inverse-Exponential prior.

n	α	k	β	λ	c_2	Criterion	$\hat{\alpha}_{mle}$	$\hat{\alpha}_{sq}$	$\hat{\alpha}_{pre}$	$\hat{\alpha}_{alb}$	$\hat{\alpha}_{ste}$
10	3	-0.5	1	3.5	5	Estimate	2.93546	2.99603	3.07599	4.23702	2.85660
						Bias	-0.06454	-0.00397	0.07599	1.23702	-0.14340
						Variance	0.22744	0.21819	0.23000	0.43639	0.19836
						MSE	0.23161	0.21821	0.23577	1.96661	0.21892
50	3	-0.5	1	3.5	5	Estimate	2.98571	2.99747	3.01265	3.15961	2.96794
						Bias	-0.01429	-0.00253	0.01265	0.15961	-0.03206
						Variance	0.04084	0.04052	0.04093	0.04502	0.03973
						MSE	0.04105	0.04053	0.04109	0.07050	0.04076
100	3	-0.5	1	3.5	5	Estimate	2.99384	2.9997	3.00724	3.07762	2.98481
						Bias	-0.00616	-0.0003	0.00724	0.07762	-0.01519
						Variance	0.02369	0.0236	0.02372	0.02484	0.02336
						MSE	0.02373	0.0236	0.02377	0.03087	0.02360
10	4	0.1	1	3	3	Estimate	3.96977	3.99370	4.05866	4.43061	3.87445
						Bias	-0.03023	-0.00630	0.05866	0.43061	-0.12555
						Variance	0.25840	0.25532	0.26369	0.31424	0.24030
						MSE	0.25931	0.25536	0.26713	0.49966	0.25606
50	4	0.1	1	3	3	Estimate	3.99208	3.99679	4.00938	4.07391	3.97204
						Bias	-0.00792	-0.00321	0.00938	0.07391	-0.02796
						Variance	0.05112	0.05100	0.05132	0.05298	0.05037
						MSE	0.05118	0.05101	0.05140	0.05845	0.05115
100	4	0.1	1	3	3	Estimate	3.99707	3.99942	4.00569	4.03745	3.98698
						Bias	-0.00293	-0.00058	0.00569	0.03745	-0.01302
						Variance	0.02562	0.02559	0.02567	0.02608	0.02543
						MSE	0.02563	0.02559	0.02570	0.02749	0.02560

From the results of simulation tables 3,4,5, and 6, conclusions are drawn regarding the performance and behavior of the estimators under different priors, which are summarized below.

- The performances of the Bayesian and MLEs become better when the sample size increases.
- It has been observed that Bayesian estimation, outperforms MLE estimation.

- In terms of MSE, in most cases the Bayesian estimation under squared error loss function gives smaller MSEs as compared to other loss functions.

3.5. Fitting of real life data-set

For illustrative purposes, we analyze three different types of real datasets. The dataset I consists of tensile strength measurements (in GPA) from 69 carbon fibers tested under tension at gauge lengths of 20mm. These measurements were initially reported by Bader and Priest [5]. The dataset II consists of an accelerated life test conducted on 59 conductors, with failure times measured in hours. Reported first by Johnston [10]. The dataset III comprises times between arrivals of 25 customers at a facility and reported first Grubbs [8]. Our objective is to evaluate and contrast the performance of KWMBD estimates using mle and Bayesian estimation.

Table 7: Average estimate, Mean Squared Error, AIC, BIC for posterior distribution under different priors for dataset I.

critierion	MLE	Ex-Jeffreys Prior	Hartigan's Prior	I-Gamma Prior	I- Exponential Prior
Estimate	2.5001	2.4390	2.4911	2.4144	2.4819
MSE	0.2440	0.2418	0.24320	0.2430	0.2426
AIC	228.6145	197.1729	198.5274	196.5776	198.2806
BIC	230.8486	199.4070	200.7615	198.8117	200.5147

Table 8: Estimates and MSE for Extension of Jeffrey's and Inverse-Gamma Priors with different loss functions for dataset I.

$\hat{\alpha}_{mle}$		priors	$\hat{\alpha}_{sq}$		$\hat{\alpha}_{pre}$		$\hat{\alpha}_{alb}$		$\hat{\alpha}_{ste}$	
Estimate	MSE		Estimate	MSE	Estimate	MSE	Estimate	MSE	Estimate	MSE
2.5001	0.2440	EX-Jeffrey's Prior	2.4390	0.2418	2.4475	0.2417	2.4911	0.2432	2.4224	0.2424
		I-Gamma Prior	2.4391	0.2418	2.456	0.2416	2.5460	0.2506	2.4064	0.2436

Table 9: Average estimate, Mean Squared Error, AIC, BIC for posterior distribution under different priors for dataset II.

critierion	MLE	Ex-Jeffreys Prior	Hartegan's Prior	I-Gamma Prior	I- Exponential Prior
Estimate	7.16117	6.957312	7.129853	6.855565	7.077978
MSE	2.59377	2.561495	2.583413	2.576478	2.570564
AIC	319.9468	246.6048	247.7058	246.1869	247.3245
BIC	322.0243	248.6823	249.7833	248.2644	249.4021

Table 10: Estimates and MSE for Extension of Jeffrey's and Inverse-Gamma Priors with different loss functions for dataset II.

$\hat{\alpha}_{mle}$		priors	$\hat{\alpha}_{sq}$		$\hat{\alpha}_{pre}$		$\hat{\alpha}_{alb}$		$\hat{\alpha}_{ste}$	
Estimate	MSE		Estimate	MSE	Estimate	MSE	Estimate	MSE	Estimate	MSE
7.1612	2.5938	EX-Jeffrey's Prior	6.9577	2.5615	6.9858	2.5610	7.2538	2.6359	6.9027	2.5670
		I-Gamma Prior	6.9370	2.5628	6.9931	2.5612	7.5631	2.9010	6.8294	2.5837

Table 11: Average estimate, Mean Squared Error, AIC, BIC for posterior distribution under different priors for dataset III.

critierion	MLE	Ex-Jeffreys Prior	Hartegan's Prior	I-Gamma Prior	I- Exponential Prior
Estimate	4.0405	3.9242	4.0003	3.8025	3.9433
MSE	0.6053	0.6015	0.6010	0.6264	0.6003
AIC	108.1082	85.9577	86.3621	85.4566	86.0532
BIC	109.3270	87.1765	87.5810	86.6755	87.2721

Table 12: Estimates and MSE for Extension of Jeffrey's and Inverse-Gamma Priors with different loss functions for dataset III.

$\hat{\alpha}_{mle}$		priors	$\hat{\alpha}_{sq}$		$\hat{\alpha}_{pre}$		$\hat{\alpha}_{alb}$		$\hat{\alpha}_{ste}$	
Estimate	MSE		Estimate	MSE	Estimate	MSE	Estimate	MSE	Estimate	MSE
4.0405	0.6054	EX-Jeffrey's Prior	3.9242	0.6015	3.9621	0.5998	4.0811	0.6131	3.8522	0.6126
		I-Gamma Prior	3.9070	0.6032	3.9829	0.6001	4.2331	0.6713	3.7699	0.6381

The results of tables 7, 8, 9, 10, 11 and 12 demonstrate that the estimation of parameters for KWMBD under both priors (Extension of Jeffrey's and Inverse Gamma prior) and precautionary and square error loss function is better compared to the other three loss functions considered and mle estimation, owing to its lower Mean Squared Error (MSE).

4. CONCLUSION

We compared estimation methods for the scale parameter α of the 2kth order weighted Maxwell-Boltzmann distribution, utilizing both Maximum Likelihood Estimation (MLE) and Bayesian Estimation under various loss functions and prior distributions. This comparison is based on the simulated data and real-life datasets. Results of simulated data reveal that as the sample size increases, MSE decreases. and the Bayesian Estimation outperforms Maximum Likelihood Estimation (MLE). Furthermore, results from the real-life datasets demonstrate that the estimation of parameters of KWMBD under both prior distributions and precautionary loss function and square error loss function yields better performance, with smaller MSE compared to other estimators.

Conflict of interest: The authors confirm that they have no conflicts of interest to disclose regarding the publication of this paper.

REFERENCES

- [1] A. Ahmad and R. Tripathi. Bayesian estimation of weighted inverse maxwell distribution under different loss functions. *Earthline Journal of Mathematical Sciences*, 8(1):189–203, 2022.
- [2] A. Ahmed, S. Ahmad, and J. Reshi. Bayesian analysis of rayleigh distribution. *International Journal of Scientific and Research Publications*, 3(10):1–9, 2013.
- [3] H. Al-Kutobi. *On comparison estimation procedures for parameter and survival function exponential distribution using simulation*. PhD thesis, Ph. D. Thesis, Baghdad University, College of Education (Ibn-Al-Haitham), 2005.
- [4] M. Aslam. An application of prior predictive distribution to elicit the prior density. *Journal of Statistical Theory and applications*, 2(1):70–83, 2003.
- [5] M. Bader and A. Priest. Statistical aspects of fibre and bundle strength in hybrid composites. *Progress in science and engineering of composites*, pages 1129–1136, 1982.
- [6] A. Chaturvedi and U. Rani. Classical and bayesian reliability estimation of the generalized maxwell failure distribution. *Journal of Statistical Research*, 32(1):113–120, 1998.

- [7] M. Friesl and J. Hurt. On bayesian estimation in an exponential distribution under random censorship. *Kybernetika*, 43(1):45–60, 2007.
- [8] F. E. Grubbs. Approximate fiducial bounds on reliability for the two parameter negative exponential distribution. *Technometrics*, 13(4):873–876, 1971.
- [9] M. Han. E-bayesian estimation and its e-mse under the scaled squared error loss function, for exponential distribution as example. *Communications in Statistics-Simulation and Computation*, 48(6):1880–1890, 2019.
- [10] G. Johnston. Statistical models and methods for lifetime data, 2003.
- [11] H. Rasheed. Minimax estimation of the parameter of the maxwell distribution under quadratic loss function. *Journal of Al-Rafidain University College For Sciences (Print ISSN: 1681-6870, Online ISSN: 2790-2293)*, (1):43–56, 2013.
- [12] J. Reshi, A. Ahmed, and K. Mir. Some important statistical properties, information measures and estimations of size biased generalized gamma distribution. *Journal of Reliability and Statistical Studies*, pages 161–179, 2014.
- [13] J. A. Reshi, B. A. Para, and S. A. Bhat. Parameter estimation of weighted maxwell-boltzmann distribution using simulated and real life data sets. In *Adaptive Filtering-Recent Advances and Practical Implementation*. IntechOpen, 2021.
- [14] F. A. Spiring and A. S. Yeung. A general class of loss functions with industrial applications. *Journal of Quality Technology*, 30(2):152–162, 1998.
- [15] R. Tyagi and S. Bhattacharya. Bayes estimation of the maxwell's velocity distribution function. *Statistica*, 29(4):563–567, 1989.
- [16] L. Weiss. Introduction to wald (1949) statistical decision functions. In *Breakthroughs in Statistics: Foundations and Basic Theory*, pages 335–341. Springer, 1992.
- [17] A. Zellner. Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, 81(394):446–451, 1986.

BAYESIAN NON-INFERIORITY TEST BETWEEN TWO BINOMIAL PROPORTIONS

W. B. Yahya¹, C. P. Ezenweke², O. R. Olaniran³, I. A. Adeniyi⁴, K. Jimoh⁵,
R. B. Afolayan⁶, M. K. Garba⁷, I. Ahmed⁸

^{1,3,6,7}Department of Statistics, University of Ilorin, Ilorin, Nigeria

¹wbyahya@unilorin.edu.ng; ³rid4stat@yahoo.com;

⁶afolayan.rb@unilorin.edu.ng ; ⁷garba.mk@unilorin.edu.ng

^{2,4}Department of Mathematics, Federal University Lokoja, Nigeria

²chinenye.ezenweke@fulokoja.edu.ng; ⁴isaac.adeniyi@fulokoja.edu.ng

⁵Department of Physical Sciences, Al-Hikmah University, Ilorin, Nigeria

⁵jimkaminsha@alhikmah.edu.ng

⁸Department of Statistics, Nasarawa State University, Keffi, Nigeria

⁸ibrahimloko@nsuk.edu.ng

Abstract

The paper aimed to propose a new Bayesian test method for establishing a non-inferiority measure between an active treatment (drug) and a new (cheaper) treatment using two independent binomial samples. A Bayesian test statistic was developed for testing non-inferiority between two independent binomial proportions. Conjugate Beta prior was assumed for the binomial proportions to elicit posterior from the same Beta family of distributions. The efficiency of this test method was established via power analysis and its ability to yield the nominal Type I error rate (alpha) in a detailed Monte-Carlo study. Results from this study showed that the proposed test method yielded higher powers and good estimates of the Type I error rate at the chosen sample sizes and varying non-inferiority margins (effect sizes). Thus, the new Bayesian test method is very efficient at detecting the significance of the non-inferiority margin between two independent binomial proportions when such is not negligible at all sample sizes. Further results showed that the size of the two population proportions being tested influences the power and the estimated nominal Type I error rate with an increase in power and a good estimate of Type I error rate achieved when both population proportions being tested are less than 0.5. It is therefore concluded that the new Bayesian test method can be employed whenever it is desirable to establish the existence of non-inferiority or otherwise between a pair of (clinical) treatments (drugs). All the simulations and analyses were performed with the R statistical package.

Keywords Non-inferiority test, Test of proportion, Bayesian Inference, Conjugate prior, Power

1. Introduction

In pharmacological and drug discovery studies, the need for the development of new drugs to compete with some of the existing ones for a particular ill-health condition is often desirable for many reasons. Clinicians and end-users may prefer a new drug (or treatment) if significant advantages can be derived from its use such as having fewer side effects, being less toxic, being relatively cheaper, being relatively more convenient in its formulation and administration (e.g. tablets instead of infusion or infusion/drug instead of surgery) and so on (Yahya et al. [1]). Therefore, it is always a welcome development that improved, better, and (possibly) cheaper drugs or treatments are discovered and introduced for the treatment of various forms of ailments.

In clinical trials, a new drug (or treatment) for a particular medical condition can only be accepted as being good if its performance is at least as good as an existing one (the active control). In other words, a new drug may be preferred if its performance is equivalent to or better than (not inferior to) the active control (Chen and Peace [2], Ng [3], Norleans [4]). This is the scenario that brought about the concept of non-inferiority between two drugs or treatments in clinical trials.

There are different methods for determining if a new drug (or treatment), often referred to as the test product, is as good as an existing one, most of which involve the use of statistical tools. Some of the procedures for determining if the performance of a new drug is not inferior to an existing one include *superiority trials* and *non-inferiority trials* (Lesaffre [5], Yahya and Jolayemi [6]).

A superiority trial is used to prove that a new drug is better than an existing one, while a non-inferiority trial is used to show that the new drug is not (much) worse than the existing one (Kawasaki et al. [7]). Many clinical trials that compare a test product with an active comparator are designed as *non-inferiority trials* (Ng [3], EMA [8]). This is attributed to the fact that a placebo (controlled trial) is considered unethical or impractical (Temple [9]) among other reasons. In this study, our focus is much on *non-inferiority trials* based on nominal outcomes from two independent proportions, and a new test method to handle such a situation is proposed.

The term '*non-inferiority trial*' is commonly used to refer to a randomized clinical trial in which a new treatment is compared with a standard active treatment rather than a placebo or untreated control group (D'Agostino et al. [10], Pocock [11]). The non-inferiority trial is a trial to show how an experimental treatment is statistically and clinically not inferior to the active control, see D'Agostino et al. [10]. In clinical trials, non-inferiority tests are frequently used to demonstrate that the response for study drugs is not much worse than the response for reference drugs, see Kawasaki et al. [7], Kawasaki and Miyaoka [12].

1.1 Non-Inferiority Tests

A *non-inferiority test* is used to indicate whether the responses from new drugs are clinically not much worse than the response to active control. It is often conducted in clinical trials. Many clinical trials that compare a test product with an active comparator or control are designed as non-inferiority trials. A non-inferiority trial is sometimes stated as being to demonstrate that the test product is not inferior to the active control. However, only a superiority trial can demonstrate this. A non-inferiority trial aims to demonstrate that the test product is not worse than the active control by more than a pre-specified threshold, often called the *non-inferiority margin*, Δ , see Ng [3], EMA [8], D'Agostino et al. [10].

Non-inferiority studies are typically confirmatory trials that employ a randomized parallel-group design with an active control group. Some trials also include a placebo control group. The placebo is used to validate the study and to demonstrate the superiority of the test treatment to the

placebo (Garrett [13]). Non-inferiority methods are frequently used in more serious, acute, and sometimes life-threatening situations such as oncology and infectious diseases, see Garrett [13].

Without loss of generality, a non-inferiority test is used to indicate whether the responses from a new drug are clinically not much worse than the responses from the active control (an existing drug) as it is often conducted in clinical trials. Thus, a prior judgment is made that, for the new treatment to be of merit, it only needs to be as good as the active control regarding appropriate outcome measure(s) of response, see Pocock [11].

Non-inferiority tests are examples of directional (one-sided) tests. There are many forms of this test depending on the form of the response. The *non-inferiority* tests for two-sample designs in which the outcome is a continuous normal random variable were carried out using the two-sample t-test procedure and the analysis of variance (ANOVA) k -sample F-test procedure when there are more than two samples (Yahya and Jolayemi [6]). When the response or outcome is binary, the non-inferiority test procedure, which is the main focus of this study, involves determining if the difference between two binomial proportions is beyond a specified non-inferiority margin Δ .

1.2 Non-Inferiority Test Between Two Independent Proportions

The responses from non-inferiority studies can be binary, for example, cure or no cure, alive or dead, cancerous or non-cancerous, and so on. When this is the case, the test product group and active control group are referred to as two independent binominal samples. The non-inferiority tests for this case are carried out using the tests for the differences between two proportions. For instance, suppose that the current drug (treatment) for a given disease works 80% of the time. Suppose also that this treatment is either too expensive or occasionally exhibits serious side effects or that its administration is considered quite tedious (e.g. surgery) and a new promising treatment or drug has been developed to the point where it can be tested. One of the questions that must be answered is whether the new treatment is as good as the current treatment. In other words, do at least 80% of treated subjects respond positively to the new treatment?

Due to the many benefits (being cheaper, having relatively fewer side effects, and being very easy to administer) of the new treatment, clinicians would be willing to adopt the new treatment even if it is slightly less effective than the current treatment by some small tolerable margin. In other words, it is of interest to determine how much less effective the new treatment can be relative to the existing one (the active control) for which the new treatment can still be confidently adopted. This is called the *non-inferiority margin* of the new drug compared to the existing one (Ng [3], Yahya and Jolayemi [12]). The non-inferior margin represents how much worse the new treatment can be compared with the standard treatment, yet still, be considered 'similar' or 'not worse' than the standard treatment, Leung et al. [14]. In essence, there is a performance level of the new treatment as compared to the existing active control treatment below or above which is no longer considered ignorable.

In the above example, suppose it was found that the performance rate of the new drug is 75% among the population of users and this was considered acceptable by the clinicians and other users due to its numerous attendant benefits over the existing treatment. This simply shows that the *non-inferiority margin* of the new drug is 5%, i.e. $\Delta = 5\%$ and the new treatment is thus, considered non-inferior to the existing treatment. Thus, the drug developers need to design an experiment to test the hypothesis that the response rate of the new drug (treatment) is at least 0.75. In other words, the test hypothesis problem here is to establish whether the non-inferiority margin of 5% is exceeded by the new drug or not.

Different test methods have been proposed for establishing non-inferiority between a pair of two independent population proportions if such exists. Most of these methods were developed through the frequentist techniques while a few had appeared within the Bayesian framework. In

this work, a new Bayesian test method for testing the difference between two population proportions within the context of the non-inferiority formulation is proposed. In the development of this new test method, the emphasis is more on the effect sizes rather than the p-values of the statistical tests as opined by Leung et al. [14].

2. Methods

Let X_j be an independent Bernoulli random variable that indicates whether a clinician or patient prefers treatment/drug j with the associated probability of success π_j , while $j = 1$ for the existing active treatment and $j = 2$ for the new treatment. Therefore, in a total of n_j end-users of treatment/drug j , the random variable $\sum_{i=1}^{n_j} X_{ij}$ that represents the number of end-users that preferred treatment/drug j is distributed Binomial with parameters n_j and π_j . That is, $\sum_{i=1}^{n_j} X_{ij} \sim \text{Bin}(n_j, \pi_j)$. The sample estimate of π_j is obtained as $\hat{\pi}_j = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j}$, $j = 1, 2$. For simplicity, the terms 'treatment' and 'drug' shall be used interchangeably in the subsequent discussions to represent the existing or the new clinical treatment adopted or preferred by the end-users.

To establish the efficiency of the new drug relative to the existing drug within the context of non-inferiority, the hypothesis of non-inferiority between the two drugs in terms of their relative preferences by end-users is constructed around the two independent population proportions π_1 (for existing drug) and π_2 (for new drug), and the non-inferiority margin Δ_0 as follows:

$$H_0: \pi_1 \geq \pi_2 + \Delta_0 \text{ against } H_a: \pi_1 < \pi_2 + \Delta_0 \quad (1)$$

The inequality statement $\pi_1 \geq \pi_2 + \Delta_0$ in the null hypothesis H_0 simply indicates that the proportion (π_2) of users that preferred the new drug is only up to the proportion (π_1) of those that preferred the existing drug by a margin of the non-inferiority parameter value Δ_0 .

Therefore, if the null hypothesis H_0 is not rejected by the test, it simply shows that the new drug is not inferior to the existing drug an indication that the non-inferiority margin Δ_0 is negligible. However, a rejection of the null hypothesis H_0 in favour of the alternative hypothesis H_a is a strong indication that the new drug is superior (more preferred by users) to the existing drug. Hence, the non-inferiority margin Δ_0 , which has contributed significantly to the rejection of H_0 is not negligible.

Without loss of generality, the hypothesis set in (1) can be re-expressed as follows (with the null hypothesis H_0 indicating no difference between the treatment and the active control groups):

$$H_0: \pi_1 - \pi_2 = \Delta_0 \text{ against } H_a: \pi_1 - \pi_2 < \Delta_0 \quad (2)$$

Some of the existing methods for testing the non-inferiority hypothesis set in (2) include the z-test with a *pooled variance* given by

$$Z_p = \frac{(\hat{\pi}_1 - \hat{\pi}_2) - \Delta_0}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \hat{\pi}_c (1 - \hat{\pi}_c)}} \sim N(0,1) \quad (3)$$

where, $\hat{\pi}_c$ is computed as $\hat{\pi}_c = \frac{\sum_{i=1}^{n_1} x_{i1} + \sum_{i=1}^{n_2} x_{i2}}{n_1 + n_2}$.

Another test statistic for hypothesis in (2) is the Wald statistic, Ward and Ahlquist [15] given by:

$$Z_w = \frac{(\hat{\pi}_1 - \hat{\pi}_2) - \Delta_0}{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}} \sim N(0,1) \quad (4)$$

It has been indicated in some studies that the performance of the Wald statistic suffers when the sample size is small, Di-Caterina and Kosmidis [16]. However, Munzel and Hsuschke [17] have shown the framework of the non-inferiurity test for ordered categorical data. When the number of categories is two, it can be regarded as a problem of the difference between two proportions. Hence, the Wald test statistic is derived by extending the method proposed by Munzel and Hsuschke [17] to the non-inferiurity test for deriving the difference between two proportions.

In this paper, an attempt is made to extend the Wald-test procedures from a Bayesian perspective.

2.1. The Posterior Distribution of Bayesian Binomial Proportions

The posterior distribution $p(\pi|X)$ is defined as:

$$p(\pi|X) = \frac{L(\pi|X) \times p(\pi)}{\int_{\pi} L(\pi|X) \times p(\pi) d\pi} \quad (5)$$

Given that $f(x) = \pi^x (1 - \pi)^{1-x}$; then, the likelihood distribution function $L(\pi|X)$ is defined by

$$\begin{aligned} L(\pi|X) &= \prod_{i=1}^n \pi^{x_i} (1 - \pi)^{1-x_i} \\ \rightarrow L(\pi|X) &= \pi^{\sum x_i} (1 - \pi)^{n - \sum x_i} \end{aligned} \quad (6)$$

Therefore, the Maximum Likelihood Estimator (MLE) of π is given by

$$\text{MLE}(\pi) = \hat{\pi} = \frac{\sum x_i}{n} \quad (7)$$

Assuming a conjugate beta prior for the likelihood in order to ascertain the same class of the posterior, the prior distribution, $p(\pi)$ is defined as;

$$p(\pi) = \frac{\pi^{a_0-1} (1-\pi)^{b_0-1}}{\Gamma(a_0) \times \Gamma(b_0)} \Gamma(a_0 + b_0), \quad 0 < \pi < 1 \quad (8)$$

Therefore, the posterior distribution $p(\pi|X)$ is determined using (5) as

$$\begin{aligned} p(\pi|X) &= \frac{\pi^{\sum x_i} (1 - \pi)^{n - \sum x_i} \times \frac{\pi^{a_0-1} (1 - \pi)^{b_0-1}}{\Gamma(a_0) \times \Gamma(b_0)} \Gamma(a_0 + b_0)}{\int_0^1 \pi^{\sum x_i} (1 - \pi)^{n - \sum x_i} \times \frac{\pi^{a_0-1} (1 - \pi)^{b_0-1}}{\Gamma(a_0) \times \Gamma(b_0)} \Gamma(a_0 + b_0) d\pi} \\ p(\pi|X) &= \frac{\pi^{\sum x_i} (1 - \pi)^{n - \sum x_i} \times \pi^{a_0-1} (1 - \pi)^{b_0-1}}{\int_0^1 \pi^{\sum x_i} (1 - \pi)^{n - \sum x_i} \times \pi^{a_0-1} (1 - \pi)^{b_0-1} d\pi} \\ p(\pi|X) &= \frac{\pi^{a_0 + \sum x_i - 1} (1 - \pi)^{b_0 + n - \sum x_i - 1}}{\int_0^1 \pi^{a_0 + \sum x_i - 1} (1 - \pi)^{b_0 + n - \sum x_i - 1} d\pi} \end{aligned} \quad (9)$$

Recall from the Beta function that,

$$\int_0^1 \pi^{a_0 + \sum x_i - 1} (1 - \pi)^{b_0 + n - \sum x_i - 1} d\pi = \frac{\Gamma(a_0 + \sum x_i - 1) \times \Gamma(b_0 + n - \sum x_i - 1)}{\Gamma(a_0 + b_0 + n - 2)} \quad (10)$$

Therefore, the prior distribution in (10) becomes;

$$p(\pi|X) = \pi^{a_0 + \sum x_i - 1} (1 - \pi)^{b_0 + n - \sum x_i - 1} \times \frac{\Gamma(a_0 + \sum x_i - 1) \Gamma(b_0 + n - \sum x_i - 1)}{\Gamma(a_0 + b_0 + n - 2)} \quad (11)$$

Thus,

$$p(\pi|X) \sim \text{Beta}(a_0 + \sum x_i, b_0 + n - \sum x_i), 0 < \pi < 1 \quad (12)$$

If we let $a = a_0 + \sum x_i$; $b = b_0 + n - \sum x_i$ in (11) and (12), the posterior mean, $E(\pi|x)$ and posterior variance, $\text{Var}(\pi|x)$ can be obtained as;

$$E(\pi|x) = \frac{a}{a+b} \quad (13)$$

$$\text{Var}(\pi|x) = \frac{ab}{(a+b)^2(a+b+1)} \quad (14)$$

2.2 Prior Elicitation

2.2.1 Choosing a Conjugate Prior by Matching Location and Scale Parameters

Given that the $\text{Beta}(ab)$ family of distributions is the conjugate family for $\text{Binomial}(n, \pi)$ distribution, the posterior will be a member of the same family, with the parameters updated by simple rules. Then, we can find the posterior without integration. Note that, the beta distribution can have many shapes. Therefore, the prior chosen should correspond to one's belief. We suggest choosing a $\text{Beta}(a_0, b_0)$ that matches one's prior belief about the mean (location) and standard deviation (scale).

Let π_0 be the prior mean for the proportion and let σ_0^2 be the prior variance for the proportion. But, we know that the mean of $\text{Beta}(ab)$ distribution is $\frac{a}{a+b}$. If this is set to equal to what the prior belief about the mean of the proportion is, we have;

$$\pi_0 = \frac{a_0}{a_0 + b_0} \quad (16)$$

Also, the variance of the $\text{Beta}(ab)$ distribution is $\frac{ab}{(a+b)^2(a+b+1)}$. This shall be set to equal to the prior belief about the variance for the proportion to have:

$$\sigma_0^2 = \frac{a_0 b_0}{(a_0 + b_0)^2 (a_0 + b_0 + 1)} \quad (17)$$

Note that $\pi_0 = \frac{a_0}{a_0 + b_0}$ from (16), this implies that $1 - \pi_0 = \frac{b_0}{a_0 + b_0}$. Substituting these in (17) to have;

$$\sigma_0^2 = \frac{\pi_0(1-\pi_0)}{a_0 + b_0 + 1} \quad (18)$$

Solving equations (16) and (18) for a_0 and b_0 gives the $\text{Beta}(a_0, b_0)$ prior parameters.

Proof: Given that $\pi \sim \text{Beta}(ab)$ then,

$$E(\pi) = \frac{a}{a+b}; \quad \text{Var}(\pi) = \frac{ab}{(a+b)^2(a+b+1)} \quad (19)$$

Solving for a_0 in terms of b_0 and π_0 from (16) gives

$$a_0 = \frac{b_0\pi_0}{1-\pi_0} \quad (20)$$

Substituting (20) in (18) and solving for b_0 gives

$$\begin{aligned} \sigma_0^2 &= \frac{\pi_0(1-\pi_0)}{\frac{b_0\pi_0}{1-\pi_0} + b_0 + 1} \\ \sigma_0^2 &= \frac{\pi_0(1-\pi_0)}{\frac{b_0\pi_0 + b_0 - b_0\pi_0 + 1 - \pi_0}{1-\pi_0}} \\ \sigma_0^2 &= \frac{\pi_0(1-\pi_0)}{b_0 + 1 - \pi_0} \\ \sigma_0^2 &= \pi_0(1-\pi_0) \times \frac{1-\pi_0}{b_0 + 1 - \pi_0} \\ \sigma_0^2 &= \frac{\pi_0(1-\pi_0)^2}{b_0 + 1 - \pi_0} \\ \sigma_0^2(b_0 + 1 - \pi_0) &= \pi_0(1-\pi_0)^2 \\ b_0\sigma_0^2 + \sigma_0^2 - \sigma_0^2\pi_0 &= \pi_0(1-\pi_0)^2 \\ b_0\sigma_0^2 + \sigma_0^2(1-\pi_0) &= \pi_0(1-\pi_0)^2 \\ b_0\sigma_0^2 &= \pi_0(1-\pi_0)^2 - \sigma_0^2(1-\pi_0) \\ b_0 &= \frac{\pi_0(1-\pi_0)^2 - \sigma_0^2(1-\pi_0)}{\sigma_0^2} \\ b_0 &= \frac{(1-\pi_0)[\pi_0(1-\pi_0) - \sigma_0^2]}{\sigma_0^2} \end{aligned} \quad (21)$$

Substituting b_0 in (21) into (20) to have

$$\begin{aligned} a_0 &= \frac{\frac{(1-\pi_0)[\pi_0(1-\pi_0) - \sigma_0^2]}{\sigma_0^2} \pi_0}{1-\pi_0} \\ a_0 &= \frac{\pi_0(1-\pi_0)[\pi_0(1-\pi_0) - \sigma_0^2]}{(1-\pi_0)\sigma_0^2} \\ a_0 &= \frac{\pi_0[\pi_0(1-\pi_0) - \sigma_0^2]}{\sigma_0^2} \end{aligned} \quad (22)$$

Therefore, the posterior mean, $\hat{\pi}$ and posterior variance, $\hat{\sigma}_\pi^2$ are determined as follows. We recall from (13) and (14) that

$$E(\pi|x) = \frac{a}{a+b} = \hat{\pi}$$

and

$$V(\pi|x) = \frac{ab}{(a+b)^2(a+b+1)}$$

respectively.

$$\rightarrow V(\pi|x) = \frac{a}{a+b} \times \frac{b}{a+b} \times \frac{1}{a+b+1} = \hat{\sigma}_\pi^2$$

Also, recall that $a = a_0 + \sum x_i$; $b = b_0 + n - \sum x_i$ then, by proper substitution, we have that

$$\hat{\pi} = \frac{a_0 + \sum x_i}{a_0 + \sum x_i + b_0 + n - \sum x_i}$$

$$\rightarrow \hat{\pi} = \frac{a_0 + \sum x_i}{a_0 + b_0 + n} \quad (23)$$

$$\hat{\sigma}_{\pi}^2 = \frac{a_0 + \sum x_i}{a_0 + \sum x_i + b_0 + n - \sum x_i} \times \frac{b_0 + n - \sum x_i}{a_0 + \sum x_i + b_0 + n - \sum x_i} \times \frac{1}{a_0 + \sum x_i + b_0 + n - \sum x_i + 1}$$

$$\rightarrow \hat{\sigma}_{\pi}^2 = \frac{a_0 + \sum x_i}{a_0 + b_0 + n} \times \frac{b_0 + n - \sum x_i}{a_0 + b_0 + n} \times \frac{1}{a_0 + b_0 + n + 1} \quad (24)$$

By substituting (21) and (22) in (23) for a_0 and b_0 the Bayesian posterior mean of binomial proportion π is determined as follows;

$$\hat{\pi} = \frac{\frac{\pi_0[\pi_0(1 - \pi_0) - \sigma_0^2]}{\sigma_0^2} + \sum x_i}{\frac{\pi_0[\pi_0(1 - \pi_0) - \sigma_0^2]}{\sigma_0^2} + \frac{(1 - \pi_0)[\pi_0(1 - \pi_0) - \sigma_0^2]}{\sigma_0^2} + n}$$

$$\hat{\pi} = \frac{\pi_0[\pi_0(1 - \pi_0) - \sigma_0^2] + \alpha_0^2 \sum x_i}{\pi_0[\pi_0(1 - \pi_0) - \sigma_0^2] + \pi_0(1 - \pi_0)^2 - \alpha_0^2(1 - \pi_0) + \sigma_0^2 n}$$

$$\hat{\pi} = \frac{\pi_0[\pi_0(1 - \pi_0) - \sigma_0^2] + \alpha_0^2 \sum x_i}{[\pi_0(1 - \pi_0) - \alpha_0^2][\pi_0 + (1 - \pi_0)] + \alpha_0^2 n}$$

$$\rightarrow \hat{\pi} = \frac{\pi_0[\pi_0(1 - \pi_0) - \sigma_0^2] + \alpha_0^2 \sum x_i}{\pi_0(1 - \pi_0) + \alpha_0^2(n - 1)} \quad (25)$$

Also, the Bayesian posterior variance $\hat{\sigma}_{\pi}^2$ of binomial proportion π is derived by substituting (21) and (22) in (24) as

$$\hat{\sigma}_{\pi}^2 = \frac{\hat{\pi}(1 - \hat{\pi})\sigma_0^2}{[\pi_0(1 - \pi_0) - \alpha_0^2] + \alpha_0^2 n + \alpha_0^2}$$

$$\rightarrow \hat{\sigma}_{\pi}^2 = \frac{\hat{\pi}(1 - \hat{\pi})\sigma_0^2}{\pi_0(1 - \pi_0) + n\alpha_0^2} \quad (26)$$

where $\hat{\pi} = \frac{\sum x_i}{n}$ as given in (7).

2.3 Proposed Bayesian Non-inferiority Test of Two Independent Binomial population proportions

Following the above Bayesian estimates of mean and variance of a binomial proportion, the proposed Bayesian non-inferiority test statistic for testing the hypothesis set

$$H_0: \pi_1 - \pi_2 = \Delta_0 \text{ against } H_1: \pi_1 - \pi_2 < \Delta_0$$

as earlier stated in (2) is;

$$Z_B = \frac{\hat{\pi}_1 - \hat{\pi}_2 - \Delta_0}{\sqrt{\sigma_{\hat{\pi}_1}^2 + \sigma_{\hat{\pi}_2}^2}} \sim N(0,1) \quad (27)$$

where

$$\hat{\pi}_1 = \frac{\pi_{01}[\pi_{01}(1 - \pi_{01}) - \sigma_{01}^2] + \alpha_{01}^2 \sum x_{1i}}{\pi_{01}(1 - \pi_{01}) + \alpha_{01}^2(n_1 - 1)} \quad (28)$$

$$\hat{\pi}_2 = \frac{\pi_{02}[\pi_{02}(1 - \pi_{02}) - \sigma_{02}^2] + \alpha_{02}^2 \sum x_{2i}}{\pi_{02}(1 - \pi_{02}) + \alpha_{02}^2(n_2 - 1)} \quad (29)$$

$$\hat{\sigma}_{\hat{\pi}_1}^2 = \frac{\hat{\pi}_1(1 - \hat{\pi}_1) \times \sigma_{01}^2}{\pi_{01}(1 - \pi_{01}) + \sigma_{01}^2 n_1} \quad (30)$$

$$\hat{\sigma}_{\hat{\pi}_2}^2 = \frac{\hat{\pi}_2(1 - \hat{\pi}_2) \times \sigma_{02}^2}{\pi_{02}(1 - \pi_{02}) + \sigma_{02}^2 n_2} \quad (31)$$

and $\hat{\pi}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \quad j = 1, 2$.

2.4 Decision Rule

The decision rule for the test function is to reject H_0 if $\hat{Z}_B < -Z_{1-\alpha}$ where $Z_{1-\alpha}$ is the quantile of the standard normal distribution at Type I error rate α .

If the null hypothesis H_0 is not rejected by the test function (27), it simply shows that the new drug is not inferior to the existing drug an indication that the non-inferiority margin Δ_0 is negligible. However, a rejection of the null hypothesis H_0 in favour of the alternative hypothesis H_a by the test is a strong indication that the new drug is superior (more preferred by users) to the existing drug. Hence, the non-inferiority margin Δ_0 , which has contributed significantly to the rejection of H_0 is not negligible.

2.5 The Power and Type I Error Rate (α) of the Proposed Bayesian Non-inferiority Test

In computing the power and nominal Type I error rate of the proposed Bayesian Non-inferiority test method, we consider the hypothesis of the non-inferiority test stated in (2), its test statistic as well as its decision rule. The hypothesis set to test is

$$H_0: \pi_1 - \pi_2 = \Delta_0 \text{ against } H_a: \pi_1 - \pi_2 < \Delta_0$$

with the proposed test statistic for the above test according to (27) is of the form

$$Z_B = \frac{\hat{\Delta}_0 - \Delta_0}{SE(\hat{\Delta}_0)} \sim N(0,1) \quad (32)$$

where $\hat{\Delta}_0 = \hat{\pi}_1 - \hat{\pi}_2$ and $SE(\hat{\Delta}_0)$ is the standard error of $\hat{\Delta}_0$. The decision rule for the test function as earlier stated is to reject H_0 if $\hat{Z}_B < -Z_{1-\alpha}$ where $Z_{1-\alpha}$ is the quantile of the standard normal distribution at Type I error rate α . Note that, $SE(\hat{\Delta}_0) = \sqrt{\sigma_{\hat{\pi}_1}^2 + \sigma_{\hat{\pi}_2}^2}$.

2.5.1 Power Computation

The power of the statistical test of size α for testing the null hypothesis H_0 against the alternative set H_a is the probability that the test rejects H_0 when H_a is true. That is;

$$\text{Power} = P[\text{Test function Rejects } H_0 | H_a \text{ is true}]$$

Therefore, given the above decision rule for testing the hypothesis (2) above at some specified Type I error rate α , the power of the test is defined by

$$\text{Power} = P[Z_{cal} < -Z_{1-\alpha} | H_a \text{ is true}]$$

$$\rightarrow \text{Power} = P \left[\frac{\hat{\Delta}_0 - \Delta_0}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} \mid \pi_1 - \pi_2 < \Delta_0 \right] \quad (33)$$

Suppose Δ_1 is the true difference between the two proportions π_1 and π_2 for the alternative hypothesis H_a to be true, hence, the test hypothesis (2) can be re-expressed as

$$H_0: \pi_1 - \pi_2 = \Delta_0 \text{ against } H_a: \pi_1 - \pi_2 = \Delta_1, \text{ with } \Delta_1 < \Delta_0.$$

The expression for power of the test in (33) then becomes;

$$\text{Power} = P \left[\frac{\hat{\Delta}_0 - \Delta_0}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} \mid \Delta_1 \right] \quad (34)$$

Thus, we have;

$$\begin{aligned} \text{Power} &= P \left[\frac{\hat{\Delta}_0 - \Delta_1}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} \right] \\ \text{Power} &= P \left[\frac{\hat{\Delta}_0}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} + \frac{\Delta_1}{SE(\hat{\Delta}_0)} \right] \\ \rightarrow \text{Power} &= P \left[Z < -Z_{1-\alpha} + \frac{\Delta_1}{SE(\hat{\Delta}_0)} \right] \\ \therefore \text{Power} &= \Phi \left[-Z_{1-\alpha} + \frac{\Delta_1}{SE(\hat{\Delta}_0)} \right] \end{aligned} \quad (35)$$

where Φ is the cumulative distribution function of the normal density and $SE(\hat{\Delta}_0)$ is the standard error of $\hat{\Delta}_0 = \hat{\pi}_1 - \hat{\pi}_2$. From (30) and (31), $SE(\hat{\Delta}_0)$ is computed by

$$SE(\hat{\Delta}_0) = \sqrt{\frac{\hat{\pi}_1(1 - \hat{\pi}_1) \times \sigma_{01}^2}{\pi_{01}(1 - \pi_{01}) + \sigma_{01}^2 n_1} + \frac{\hat{\pi}_2(1 - \hat{\pi}_2) \times \sigma_{02}^2}{\pi_{02}(1 - \pi_{02}) + \sigma_{02}^2 n_2}}$$

Therefore, the power of the proposed Bayesian non-inferiority test is given by (35).

2.5.2 Nominal Type I Error Rate Computation

The Type I error rate of a statistical test α is the probability that the test function rejects the null hypothesis H_0 when H_0 is actually true. Let the actual size alpha level of the test be denoted by $\hat{\alpha}$. This can be computed by;

$$\hat{\alpha} = P \left[\frac{\hat{\Delta}_0 - \Delta_0}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} \mid \Delta_0 \right] \quad (36)$$

where Δ_0 is the null difference. Thus, we have that;

$$\begin{aligned} \hat{\alpha} &= P \left[\frac{\hat{\Delta}_0 - \Delta_0}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} \right] \\ \hat{\alpha} &= P \left[\frac{\hat{\Delta}_0}{SE(\hat{\Delta}_0)} < -Z_{1-\alpha} + \frac{\Delta_0}{SE(\hat{\Delta}_0)} \right] \\ \rightarrow \hat{\alpha} &= P \left[Z < -Z_{1-\alpha} + \frac{\Delta_0}{SE(\hat{\Delta}_0)} \right] \\ \therefore \hat{\alpha} &= \Phi \left[-Z_{1-\alpha} + \frac{\Delta_0}{SE(\hat{\Delta}_0)} \right] \end{aligned} \quad (37)$$

Thus, the power function and the estimator of the significance level α for the proposed Bayesian test statistic considered in this paper are presented in Table 1.

Table 1: The Power function (Z_B -Power) and estimator of the Nominal Type I error rate α (Z_B -Alpha (α)) of the proposed Bayesian test for two independent Binomial Proportions.

Description	Estimator
Z_B -Alpha (α)	$\Phi \left[-Z_{1-\alpha} + \Delta_0 \div \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1) \times \sigma_{01}^2}{\pi_{01}(1-\pi_{01}) + \sigma_{01}^2 n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2) \times \sigma_{02}^2}{\pi_{02}(1-\pi_{02}) + \sigma_{02}^2 n_2}} \right]$
Z_B -Power	$\Phi \left[-Z_{1-\alpha} + \Delta_1 \div \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1) \times \sigma_{01}^2}{\pi_{01}(1-\pi_{01}) + \sigma_{01}^2 n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2) \times \sigma_{02}^2}{\pi_{02}(1-\pi_{02}) + \sigma_{02}^2 n_2}} \right]$

3. Simulation Study

3.1 Simulation Scheme

The data utilized for this work were simulated using the R statistical package (R Core Team [18]). Two independent binomial random variables X_1 and X_2 were generated at different sample sizes $n = 20, 40, 100$ using varying probabilities of success π_1 and π_2 with varying effect size Δ_0 where $\pi_1 = \pi_2 + \Delta_0$. The Power (%) and the nominal size α (%) values of the proposed test statistic were computed using 10,000 iterations of the basic experiment. The implementation of the Bayesian methodology is in two phases:

1. Obtaining the Bayesian estimate of $\hat{\pi}_1$ and $\hat{\pi}_2$ by updating the data using conjugate beta-prior.
2. Estimating the corresponding Power (%) and the nominal size α (%) values of the proposed Bayesian test statistic.

Recall that the non-inferiority hypothesis test is given as $H_0: \pi_1 - \pi_2 = \Delta_0$ vs. $H_a: \pi_1 - \pi_2 < \Delta_0$.

The section that follows presents the results for different values of π_1 and π_2 at various sample sizes.

4. Results

The results of the proposed non-inferiority test between two binomial proportions are presented in phases based on the different parameters combinations as provided under the simulation scheme.

Table 2 presents the estimated power of the proposed Bayesian test first at the effect size $\Delta_0 = 0.05$ but under two different sizes of proportion pairs ($\pi_1 = 0.2, \pi_2 = 0.25$) and ($\pi_1 = 0.6, \pi_2 = 0.65$) with the associated Bayesian test statistics Z_{B1} and Z_{B2} respectively at varying sample sizes. Table 2 equally presents the results of the estimated powers of the proposed test at the effect size $\Delta_0 = 0.1$ and under two different sizes of proportion pairs ($\pi_1 = 0.2, \pi_2 = 0.3$) and ($\pi_1 = 0.6, \pi_2 = 0.7$) with the associated Bayesian test statistics Z_{B3} and Z_{B4} respectively also at varying sample sizes. Expectedly, it can be observed that the power of the proposed Bayesian test method increases as the sample size increases.

Table 2: The Power (in %) of the proposed Binomial Non-inferiority Test (ZB) at various sample sizes (n), with the following test parameters; $\pi_1 = 0.2, \pi_2 = 0.25$, effect size $\Delta_0 = 0.05$ for test Z_{B1} ; $\pi_1 = 0.6, \pi_2 = 0.65$ effect size $\Delta_0 = 0.05$ for test Z_{B2} ; $\pi_1 = 0.2, \pi_2 = 0.3$, effect size $\Delta_0 = 0.1$ for test Z_{B3} ; $\pi_1 = 0.6, \pi_2 = 0.7$ effect size $\Delta_0 = 0.1$ for test Z_{B4} .

Power of Bayesian Non-Inferiority Tests				
Sample Size (n)	Effect Size $\Delta_0 = 0.05$		Effect Size $\Delta_0 = 0.1$	
	Z_{B1} $(\pi_1, \pi_2) = (0.2, 0.25)$	Z_{B2} $(\pi_1, \pi_2) = (0.6, 0.65)$	Z_{B3} $(\pi_1, \pi_2) = (0.2, 0.3)$	Z_{B4} $(\pi_1, \pi_2) = (0.6, 0.7)$
20	21.91	16.35	70.42	65.69
40	34.92	28.16	85.01	79.77
60	44.29	37.23	91.78	89.25
80	50.86	43.48	96.16	93.90
100	56.80	49.50	97.96	96.51
120	64.02	55.29	99.03	98.17
140	68.96	59.67	99.62	98.92
160	73.81	65.24	99.77	99.28
180	78.11	69.20	99.88	99.67
200	81.25	71.57	99.96	99.86

It can also be observed from the results in Table 2 that, the power of the test increases remarkably as the effect size Δ_0 increases. All these results are presented in the line graphs in Fig 1 (left) in which the powers of the test were plotted against the various sample sizes under the two effect sizes chosen.

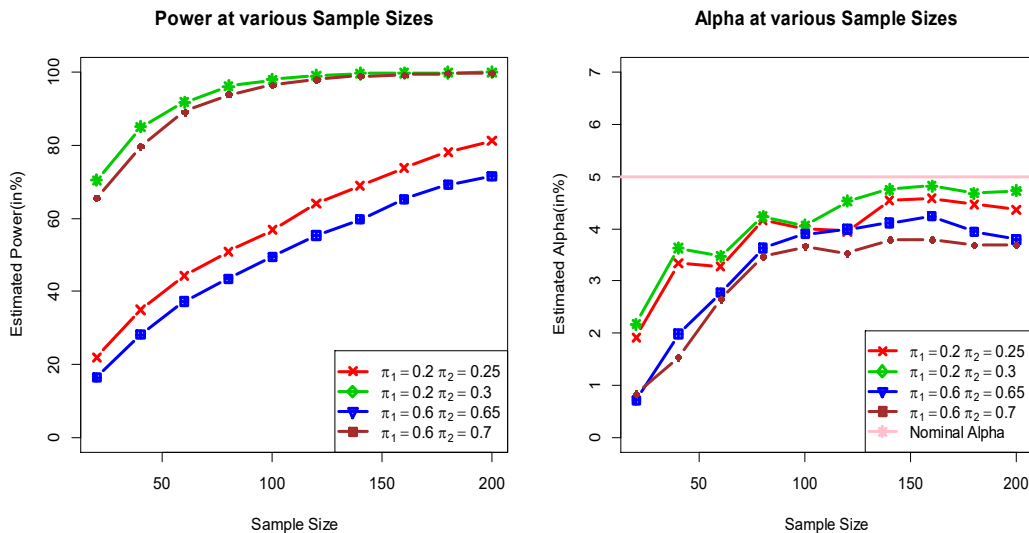


Figure 1: The graphs of the estimated Powers (left) and nominal Type I error rate α (right) of the proposed Bayesian Non-inferiority test for two Independent Population Proportions at different sample sizes and varying effect sizes.

Regarding the ability of the proposed test at returning the 5% nominal Type I error rate α set for it, it can be observed from Table 3 that the test returns nominal alpha values that are quite close to the 5% alpha level set for it at all the effect sizes most especially at higher sample sizes. However, the test under-estimated this nominal 5% α level of the test as all the estimated nominal values are below the 5% line as shown in Fig. 1 (right).

Table 3: The estimated nominal Type I error rate α (in %) of the Proposed Binomial Non-inferiority Test (ZB) at various sample sizes (n), with the following test parameters; $\pi_1 = 0.2, \pi_2 = 0.25$, effect size $\Delta_0 = 0.05$ for test Z_{B1} ; $\pi_1 = 0.6, \pi_2 = 0.65$ effect size $\Delta_0 = 0.05$ for test Z_{B2} ; $\pi_1 = 0.2, \pi_2 = 0.3$, effect size $\Delta_0 = 0.1$ for test Z_{B3} ; $\pi_1 = 0.6, \pi_2 = 0.7$ effect size $\Delta_0 = 0.1$ for test Z_{B4} .

Estimated Nominal Type I Error Rate α (in %) of Bayesian Non-Inferiority Tests				
Sample Size (n)	Effect Size $\Delta_0 = 0.05$		Effect Size $\Delta_0 = 0.1$	
	Z_{B1} $(\pi_1, \pi_2) = (0.2, 0.25)$	Z_{B2} $(\pi_1, \pi_2) = (0.6, 0.65)$	Z_{B3} $(\pi_1, \pi_2) = (0.2, 0.3)$	Z_{B4} $(\pi_1, \pi_2) = (0.6, 0.7)$
20	1.91	0.71	2.17	0.83
40	3.33	1.98	3.63	1.54
60	3.28	2.77	3.47	2.65
80	4.17	3.63	4.23	3.46
100	4.00	3.90	4.06	3.66
120	3.95	3.99	4.53	3.53
140	4.54	4.11	4.75	3.78
160	4.58	4.24	4.82	3.79
180	4.47	3.94	4.68	3.68
200	4.36	3.79	4.73	3.69

5. Discussion of Results

In this paper, the Bayesian test statistic for testing the difference of two binomial proportions under the non-inferiority condition was presented. The estimators of the power and significance level for the proposed Bayesian test were summarized in Table 1. Table 2 presents the empirical percentage power of the proposed test at varying sample sizes under the four scenarios considered while Table 3 presents the nominal Type I error rate of the test at various sample sizes under different effect sizes.

The Bayesian test statistic of non-inferiority Z_{B1} examines a test condition with a low effect size ($\Delta_0 = 0.05$) and relatively small sizes of the proportion pair ($\pi_1 = 0.2, \pi_2 = 0.25$). Under this scenario, the proposed Bayesian test method requires an average sample size of about 150 observations to achieve an approximate 70% power as can be observed in Table 2 and Fig 1 (left).

Unlike the test statistic Z_{B1} , the Bayesian test statistic Z_{B2} examines a test condition with a low effect size ($\Delta_0 = 0.05$) but relatively large sizes of the proportion pair ($\pi_1 = 0.6, \pi_2 = 0.65$) in the hypothesis set to be tested. Here, the Bayesian test Z_{B2} requires more sample units in the neighborhood of 180 samples before it could achieve about 70% power. It can therefore be concluded that the size of the binomial proportion pair (π_1, π_2) influences the size of the power of the test. Thus, the smaller the sizes of the two population proportions π_1 and π_2 being tested, the higher the power of the test.

The Bayesian test statistic Z_{B3} unlike Z_{B1} , examines a test condition with a relatively large effect size ($\Delta_0 = 0.1$) and smaller sizes of the proportion pair ($\pi_1 = 0.2, \pi_2 = 0.3$). Here, the proposed Bayesian test method achieved a reasonable power of about 70% even at a very small sample size as low as 20 samples as can be observed from the results in Table 1 and clearly shown by the power plot in Fig 1 (left).

Finally, the Bayesian test statistic Z_{B4} in Table 2 tested a pair of relatively large binomial proportions ($\pi_1 = 0.6, \pi_2 = 0.7$) with the same effect size ($\Delta_0 = 0.1$) as used in Z_{B3} test. Given these parameters settings, the Bayesian test Z_{B4} requires more samples to attain the same feat of about

70% power achieved by Z_{B3} at a relatively smaller sample size simply due to an increase in the sizes of the two proportions π_1 and π_2 . In all cases considered, the power of the proposed Bayesian test increases as the sample size increases.

To further examine the goodness of the proposed test method, its ability to retain its size α of 5% set for it was evaluated. Table 3 presents the empirical percentage significance levels α returned by the test at varying sample sizes under the four parameters combinations considered. It could be observed from Table 3 that the empirical percentage significance levels α provided by the test are closer to the 5% nominal level set for it, especially at sample sizes 100 and above. Finally, all these results in Table 3 showed that the ability of the proposed Bayesian test method to commit the Type I error is lower than the 5% nominal level set for it. At all the sample sizes considered, the test under-estimates the 5% nominal level set for it.

6. Conclusion

In this work, an efficient Bayesian test method for testing non-inferiority between two independent binomial proportions is proposed. The goodness of the proposed test method was assessed based on the power and the empirical Type I error rates provided by the test across the various sample sizes considered. Results in Table 2 clearly showed that the proposed test is quite efficient at detecting the significance of the non-inferiority parameter value when such is not negligible in all sample sizes.

Without loss of generality, the various results from power analysis and analysis of nominal Type I error rates reported by the test are a clear indication that the proposed Bayesian test method is quite efficient and good for testing and establishing non-inferiority between two binomial population proportions. Besides the high power reported by the proposed test, the empirical levels of significance $\hat{\alpha}$ estimated by the test method that was closer to the 5% nominal level set for the test in all the cases considered also confirmed the goodness of the proposed Bayesian test method. It is therefore recommended that the proposed Bayesian test method be employed whenever it is desirable to establish the existence of non-inferiority or otherwise between a pair of treatments in which the preference of users is of the essence.

References

- [1] Yahya, W. B., Olaniran, O. R., Garba, M. K., Banjoko, A. W., Dauda, K. A., and Oloredo, K. O. (2016). A Test Procedure for Ordered Hypothesis of Population Proportions against a Control. *Türkiye Klinikleri Journal of Biostatistics*, 8(1):1-12. doi:10.5336/biostatic.2016-50196.
- [2] Chen, D. G. and Peace, K. E. (2013). *Clinical Trial Data Analysis using R*. Chapman and Hall/CRC Press, UK. <https://doi.org/10.1177/0962280211425588>
- [3] Ng T-H. (2014). *Noninferiority Testing in Clinical Trials: Issues and Challenges*. 1st Edition, Chapman and Hall/CRC, UK.
- [4] Norleans, M. X. (2019). *Statistical Methods for Clinical Trials*. 1st Edition. Chapman and Hall/CRC Press, UK.
- [5] Lesaffre, E. (2008). Superiority, Equivalence and Non-Inferiority Trials. *Bulletin of the NYU Hospital for Joint Diseases*, 66(2):150-154.
- [6] Yahya, W. B. and Jolayemi, E. T. (2008). A note on the Test of Hypothesis with Ordered Alternatives. *Scientific Annals of "Alexandru Ioan Cuza"*, 54(1), 197-208.
- [7] Kawasaki, Y., Zhang, F. and Miyaoka, E. (2010). Comparisons of Test Statistics for Non-inferiority Test for the Difference between Two Independent Binominal Proportions. *American Journal of Biostatistics*, 1(1): 23-31.

- [8] European Medicines Agency (EMA) (2005). Pre-authorization Evaluation of Medicines for Human Use. Canary Wharf, London. Doc. Ref. EMEA/CPMP/EWP/2158/99.
- [9] Temple, R. J. (2022). FDA Guidance on Non-Inferiority Trials General Issues, 2022. URL: www.fda.gov (Accessed on 30th March 2022).
- [10] D'Agostino, R. B., Massaro, J. M. and Sullivan, L. M. (2003). Non-Inferiority trials. Design concepts and issues – the encounters of academic consultants in statistics. *Statistics in Medicine*, **22**(2): 169 – 186. DOI:[10.1002/sim.1425](https://doi.org/10.1002/sim.1425)
- [11] Pocock, J. S. (2003). The pros and cons of non-inferiority trials. *Blackwell Publishing Fundamental & Clinical Pharmacology*, **17**:483 – 490.
- [12] Kawasaki, Y. and Miyaoka, E. (2013). A Bayesian non-inferiority test for two independent binomial proportions. *Pharmaceutical Statistics*, **12**(4): 201-206. <https://doi.org/10.1002/pst.1571>
- [13] Garrett, A. (2003). Therapeutic equivalence; Fallacies and falsification. *Statistics in Medicine*, **22**: 741 – 762.
- [14] Leung, J. T., Barnes, S. L., Lo, S. T. and Leung, D. Y. (2020). Non-inferiority trials in cardiology: what clinicians need to know. *Heart*, **106**: 99–104. doi:10.1136/heartjnl-2019-315772
- [15] Ward, M. D. and Ahlquist, J. S. (2018). *Maximum Likelihood for Social Science: Strategies for Analysis*. Cambridge University Press. P. 36.
- [16] Di-Caterina, C., Kosmidis, I. (2018). Location-adjusted Wald statistics for scalar parameters. *Statistics (Methodology)*, 1-21. <https://arxiv.org/pdf/1710.11217.pdf>
- [17] Munzel, U. and Hsuschke, D. (2003). A nonparametric test for proving non-inferiority in clinical trials with ordered categorical data. *Pharma. Stat.*, **2**:31 – 37. DOI: 10.1002/pst.17.
- [18] R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

BAYESIAN AND E-BAYESIAN ESTIMATION OF EXPONENTIATED INVERSE RAYLEIGH DISTRIBUTION USING CONJUGATE PRIOR

RAMESH KUMAR¹, HEMANI SHARMA*², RAHUL GUPTA², ABLEEN KAUR³

Department of statistics, University of Jammu,
Jammu, J&K, India 180006

rk1825308@gmail.com¹, hemanisharma124@gmail.com^{*2}, rahulgupta68@gmail.com²,
ableenkaur23@gmail.com³

Abstract

This study explores the application of Bayesian and E-Bayesian techniques to estimate the scale parameter of the Exponentiated Inverse Rayleigh distribution. Bayesian estimates for the parameter are derived using an informative Gamma prior and evaluated under three distinct loss functions: De-Groot, Squared Error, and Al-Bayyati loss functions. Various Properties of the E-Bayesian estimators under different loss functions have also been studied. To compare the effectiveness of E-Bayesian estimates against the Bayesian counterpart, a simulation study is conducted using MatLab. The various derived estimators were compared in terms of their Mean Squared Error. The results of a simulation study reveal that E-Bayesian estimates exhibit a smaller Mean Squared Error in comparison to Bayesian estimates, thereby demonstrating their enhanced efficiency. Among the E-Bayesian estimates, the third one stands out as the most effective. Moreover, the analysis highlights that the Squared Error loss function outperforms the Al-Bayyati and De-Groot loss functions, exhibiting a smaller MSE. Furthermore, the efficacy of these estimators is demonstrated through an analysis of a real-life dataset.

Keywords: Al-Bayyati loss function, De-Groot loss function, Exponentiated inverse Rayleigh distribution, Gamma prior, Squared error loss function.

1. Introduction

The Exponentiated Inverse Rayleigh distribution (EIRD) finds extensive utility in life testing and reliability studies, playing a crucial role in domains like electronic component longevity and wind speed analysis. Its significance also extends to physics and signal processing, facilitating investigations into radiations, sounds, and light phenomena. This versatility prompts statisticians to frequently employ the EIRD across diverse datasets.

Rehman and Dar [1] conducted a comprehensive examination of the Exponentiated Inverse Rayleigh distribution, delving into its mathematical properties and harnessing Bayesian estimation techniques for parameter estimation. The probability density function (PDF) and cumulative distribution function (CDF) of the EIRD, characterized by scale parameter θ and shape parameter α , are as follows:

$$f(x, \theta, \alpha) = \frac{2\alpha\theta e^{-\frac{\alpha\theta}{x^2}}}{x^3} \quad ; x > 0, \alpha, \theta > 0 \quad (1)$$

$$F(x, \theta, \alpha) = e^{-\frac{\alpha\theta}{x^2}} \quad ; x > 0, \alpha, \theta > 0 \quad (2)$$

Numerous authors have explored the Inverse Rayleigh distribution (IRD) from various angles. Voda [2] delved into essential properties such as Maximum Likelihood Estimation (MLE), confidence intervals, and hypothesis tests. Siddiqui [3] focused on the diverse practical applications of the Inverse Rayleigh Distribution. Soliman et al. [4] utilized squared error and zero-one loss functions to devise Bayesian estimators for IRD, centered around lower record values. Reshi et al. [5] tackled parameter estimation for the Generalized Inverse Rayleigh distribution.

Dey [6] derived Bayes estimators for IRD parameters using distinct loss functions and a non-informative prior. Sindhua et al. [7] explored Bayesian estimators and associated risks for IRD parameters, emphasizing left-censored data and showcasing the efficacy of the gamma prior under Quasi-Quadratic loss functions. Okasha [8] explored E-Bayesian estimation for the Lomax distribution with type-II censored data.

This paper's objective is to conduct a statistical comparison between Bayesian estimators and Expected Bayesian estimators for the Exponentiated Inverse Rayleigh distribution's scale parameter. The analysis involves the utilization of gamma priors and different loss functions. The ensuing layout of the paper is outlined as follows: Section 2 outlines the derivation of the likelihood function, prior distribution, and posterior distribution. Section 3 presents Bayesian estimators for the EIRD scale parameter using Al-Bayyati, Squared Error, and De-Groot loss functions. In Section 4, E-Bayesian estimates are derived and their properties are examined. Section 5 is dedicated to a simulation study comparing Bayes and E-Bayes estimates. Real data analysis is tackled in Section 6, while Section 7 concludes by summarizing the findings.

2. Likelihood function, Prior and Posterior Distribution

2.1 Likelihood function

Let $\underline{x} = x_1, x_2, \dots, x_n$ be a random sample of size n drawn from EIRD. Then the likelihood function is given by

$$L(\underline{x}, \theta) = 2^n \alpha^n \theta^n \prod_{i=1}^n \frac{1}{x_i^3} e^{-\alpha \theta \sum_{i=1}^n x_i^{-2}} \quad (3)$$

In the context of Bayesian estimation, the selection of an appropriate prior holds paramount importance in parameter estimation. When a substantial understanding of the parameter(s) is available, the inclination is towards informative priors; however, when such knowledge is lacking, non-informative priors may be more appropriate. In this study, we opt for an informative prior, specifically the Gamma Prior, to derive the corresponding posterior distribution.

2.2 Prior distribution

The gamma distribution is employed as a conjugate prior distribution for the parameter θ . The subsequent Probability Density Function (PDF) is formulated using the shape parameter 'c' and the scale parameter 'r'.

$$h(\theta|c, r) = \frac{r^c \theta^{c-1} e^{-\theta r}}{\Gamma c} \quad ; c, r > 0, \theta > 0 \quad (4)$$

2.3 Posterior distribution

The posterior distribution for the parameter θ using (3) and (4), is given as

$$g(\underline{x}, \theta) = \frac{L(\underline{x}, \theta) * h(\theta)}{\int_0^\infty L(\underline{x}, \theta) * h(\theta) d\theta}$$

$$= \frac{2^n \alpha^n \theta^n \prod_{i=1}^n \frac{1}{x_i^3} e^{-\alpha \theta \sum_{i=1}^n x_i^{-2}} * \frac{r^c \theta^{c-1} e^{-\theta r}}{\Gamma c}}{\int_0^\infty 2^n \alpha^n \theta^n \prod_{i=1}^n \frac{1}{x_i^3} e^{-\alpha \theta \sum_{i=1}^n x_i^{-2}} * \frac{r^c \theta^{c-1} e^{-\theta r}}{\Gamma c} d\theta}$$

As a result, the posterior distribution of is equal to

$$g(\underline{x}, \theta) = \frac{\theta^{(n+c)-1} e^{-\theta(\alpha \sum_{i=1}^n x_i^{-2} + r)}}{\int_0^\infty \theta^{(n+c)-1} e^{-\theta(\alpha \sum_{i=1}^n x_i^{-2} + r)} d\theta} = \frac{S^{n+c} \theta^{(n+c)-1} e^{-S\theta}}{\Gamma(n+c)}, \quad \theta > 0 \quad \text{where } S = \left\{ \alpha \sum_{i=1}^n x_i^{-2} + r \right\} \quad (5)$$

3. Bayesian Estimation

In this section, we find the Bayes estimate of scale parameter of EIRD under three different loss functions as:

3.1 Under the Al-Bayyati loss function

Al-Bayyati [9] proposed a loss function, defined as

$L(\hat{\theta}, \theta) = \theta^d (\hat{\theta} - \theta)^2$; where $\hat{\theta}$ is the estimate of θ .

By using the Al-Bayyati loss function, the bayes estimator is given as

$$\begin{aligned} E\{L(\hat{\theta}, \theta)\} &= \int_0^\infty L(\hat{\theta}, \theta) * g(\underline{x}, \theta) d\theta \\ &= \int_0^\infty \theta^d (\hat{\theta} - \theta)^2 * \frac{S^{n+c} \theta^{(n+c)-1} e^{-S\theta}}{\Gamma(n+c)} d\theta \\ &= \int_0^\infty \theta^d (\hat{\theta}^2 + \theta^2 - 2\hat{\theta}\theta)^2 * \frac{S^{n+c} \theta^{(n+c)-1} e^{-S\theta}}{\Gamma(n+c)} d\theta \\ &= \hat{\theta}^2 \int_0^\infty \frac{S^{n+c} \theta^{(n+c+d)-1} e^{-S\theta}}{\Gamma(n+c)} d\theta + \int_0^\infty \frac{S^{n+c} \theta^{(n+c+d+2)-1} e^{-S\theta}}{\Gamma(n+c)} d\theta - 2\hat{\theta} \int_0^\infty \frac{S^{n+c} \hat{\theta}^{(n+c+d+1)-1} e^{-S\theta}}{\Gamma(n+c)} d\theta \end{aligned}$$

By solving the above integral, finally we get

$$E\{L(\hat{\theta}, \theta)\} = \hat{\theta}^2 * \frac{\Gamma(n+c+d)}{S^d \Gamma(n+c)} + \frac{\Gamma(n+c+d+2)}{S^{d+2} \Gamma(n+c)} - 2\hat{\theta} * \frac{\Gamma(n+c+d+1)}{S^{d+1} \Gamma(n+c)}$$

And consequently, the Bayes estimator is

$$\hat{\theta}_{BA} = \frac{(n+c+d)}{S} \quad (6)$$

3.2 Under the Squared error loss function

The Squared Error Loss Function [10] is defined as follows:

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$$

By using the Squared error loss function, the bayes estimator is given as

$$\begin{aligned} E\{L(\hat{\theta}, \theta)\} &= \int_0^\infty L(\hat{\theta}, \theta) * g(\underline{x}, \theta) d\theta \\ &= \int_0^\infty (\hat{\theta} - \theta)^2 * \frac{S^{n+c} \theta^{(n+c)-1} e^{-S\theta}}{\Gamma(n+c)} d\theta \end{aligned}$$

$$= \int_0^{\infty} (\hat{\theta}^2 + \theta^2 - 2\hat{\theta}\theta)^2 * \frac{S^{n+c}\theta^{(n+c)-1}e^{-S\theta}}{\Gamma(n+c)} d\theta$$

By solving the above integral, finally we get

$$E\{L(\hat{\theta}, \theta)\} = \hat{\theta}^2 + \frac{\Gamma(n+c+2)}{S^2\Gamma(n+c)} - 2\hat{\theta} \frac{(n+c)}{S}$$

And consequently, the Bayes estimator as

$$\hat{\theta}_{BS} = \frac{(n+c)}{S} \tag{7}$$

3.3 Under the De-Groot loss function

The De-Groot loss function [11] is defined as follows:

$$L(\hat{\theta}, \theta) = \frac{(\hat{\theta} - \theta)^2}{\hat{\theta}^2}$$

By using the De-Groot loss function, the bayes estimator is given as

$$\begin{aligned} E\{L(\hat{\theta}, \theta)\} &= \int_0^{\infty} L(\hat{\theta}, \theta) * g(x, \theta) d\theta \\ &= \int_0^{\infty} \frac{(\hat{\theta} - \theta)^2}{\hat{\theta}^2} * \frac{S^{n+c}\theta^{(n+c)-1}e^{-S\theta}}{\Gamma(n+c)} d\theta \\ &= \frac{1}{\hat{\theta}^2} \int_0^{\infty} (\hat{\theta}^2 + \theta^2 - 2\hat{\theta}\theta)^2 * \frac{S^{n+c}\theta^{(n+c)-1}e^{-S\theta}}{\Gamma(n+c)} d\theta \end{aligned}$$

By solving the above integral, finally we get

$$E\{L(\hat{\theta}, \theta)\} = 1 + \frac{1}{\hat{\theta}^2} \frac{\Gamma(n+c+2)}{S^2\Gamma(n+c)} - \frac{2}{\hat{\theta}} \frac{\Gamma(n+c+1)}{S\Gamma(n+c)}$$

And consequently, the Bayes estimator as

$$\hat{\theta}_{BD} = \frac{(n+c+1)}{S}. \tag{8}$$

4. E-Bayesian Estimation

According to Han [12], the prior parameters 'c' and 'r' should be chosen so that the prior given in (4) is a decreasing function of θ .

$$\frac{d}{d\theta} h(\theta|c, r) = \frac{r^c}{\Gamma c} \theta^{c-2} e^{-\theta r} \{(c-1) - r\theta\},$$

As a result, our prior distribution (4) becomes a decreasing function of θ , for $0 < c < 1$ and $r > 0$.

The E-Bayesian estimate of is calculated as follows:

$$\hat{\theta}_{EB} = \int_0^1 \int_0^t \hat{\theta}_{BE} * \pi(\theta, c, r) * drdc$$

The intervals of integration for the first and second integrals correspond to the domains of the hyperparameters 'c' and 'r,' respectively, ensuring that our prior density function exhibits a decreasing trend with respect to θ . The Bayesian estimate of θ , denoted as $\hat{\theta}_{EB}$, is calculated utilizing three distinct loss functions.

Subsequently, for the E-Bayesian estimates of θ , we deliberate on the choice of prior distributions for the hyperparameters 'c' and 'r.' These distributions serve primarily to explore the influence of different prior choices on the E-Bayesian estimations of θ . The hyperparameters 'c' and 'r' are governed by the following distributions

$$\pi_1(\theta, c, r) = \frac{2(t-r)}{t^2} ; 0 < c < 1; 0 < r < t \tag{9}$$

$$\pi_2(\theta, c, r) = \frac{1}{t} \quad ; \quad 0 < c < 1; \quad 0 < r < t \quad (10)$$

$$\pi_3(\theta, c, r) = \frac{2r}{t^2} \quad ; \quad 0 < c < 1; \quad 0 < r < t \quad (11)$$

Now follows the E-Bayesian estimates of the scale parameter of EIRD under proposed loss functions

4.1 E-Bayesian estimation of θ under the Al-Bayyati loss function

E-Bayesian estimate of the parameter θ based on $\pi_1(\theta, c, r)$, is provided by

$$\begin{aligned} \hat{\theta}_{EBA_1} &= \int_0^1 \int_0^t \hat{\theta}_{BA} * \pi_1(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c+d}{S} \right\} * \left\{ \frac{2(t-r)}{t^2} \right\} * dr dc \\ &= \frac{2}{t^2} \left\{ \int_0^1 (n+c+d) dc * \int_0^t \left(\frac{t-r}{S} \right) * dr \right\} \end{aligned}$$

On solving the above equation, we get

$$\hat{\theta}_{EBA_1} = \frac{2n+2d+1}{t^2} * \left\{ (t+P) * \log\left(\frac{t+P}{P}\right) - t \right\} \quad (12)$$

E-Bayesian estimate of the parameter θ based on $\pi_2(\theta, c, r)$, is provided by

$$\begin{aligned} \hat{\theta}_{EBA_2} &= \int_0^1 \int_0^t \hat{\theta}_{BA} * \pi_2(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c+d}{S} \right\} * \frac{1}{t} * dr dc \\ \text{where } P &= (\alpha \sum_{i=1}^n x_i^{-2}) \text{ and } \int_0^1 (n+c+d) dc = \frac{2n+2d+1}{2} \end{aligned}$$

Hence, on solving we get

$$\hat{\theta}_{EBA_2} = \frac{2n+2d+1}{2t} * \left\{ \log\left(\frac{t+P}{P}\right) \right\} \quad (13)$$

E-Bayesian estimate of the parameter θ based on $\pi_3(\theta, c, r)$, and is given by

$$\begin{aligned} \hat{\theta}_{EBA_3} &= \int_0^1 \int_0^t \hat{\theta}_{BA} * \pi_3(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c+d}{S} \right\} * \frac{2r}{t^2} * dr dc \end{aligned}$$

Hence on solving, we get

$$\hat{\theta}_{EBA_3} = \frac{(2n+2d+1)}{t^2} * \left\{ P * \log\left(\frac{P}{t+P}\right) + t \right\} \quad (14)$$

4.2 E-Bayesian estimation of θ under the Squared error loss function

E-Bayesian estimate of the parameter θ under based on $\pi_1(\theta, c, r)$, and is provided by

$$\begin{aligned} \theta_{EBS_1} &= \int_0^1 \int_0^t \hat{\theta}_{BS} * \pi_1(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left(\frac{n+c}{S} \right) * \left\{ \frac{2(t-r)}{t^2} \right\} * dr dc \\ &= \frac{2}{t^2} \left\{ \int_0^1 (n+c) dc * \int_0^t \left(\frac{t-r}{P+r} \right) * dr \right\} \end{aligned}$$

Where $S = \left\{ \alpha \sum_{i=1}^n x_i^{-2} + r \right\}$ and $P = \alpha \sum_{i=1}^n x_i^{-2}$

On solving the above integrals, finally we get

$$\hat{\theta}_{EBS_1} = \frac{2n+1}{t^2} * \left\{ (t+P) * \log\left(\frac{t+P}{P}\right) - t \right\} \quad (15)$$

E-Bayesian estimate of the parameter θ based on $\pi_2(\theta, c, r)$, and is provided by

$$\begin{aligned} \hat{\theta}_{EBS_2} &= \int_0^1 \int_0^t \hat{\theta}_{BS} * \pi_2(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c}{S} \right\} * \frac{1}{t} * dr dc \\ &= \frac{1}{t} \left\{ \int_0^1 (n+c) dc * \int_0^t \left(\frac{1}{S} \right) * dr \right\} \end{aligned}$$

where $S = \left\{ \alpha \sum_{i=1}^n x_i^{-2} + r \right\}$ and $P = \alpha \sum_{i=1}^n x_i^{-2}$

On solving the above integrals, finally we get

$$\hat{\theta}_{EBS_2} = \frac{2n+1}{2t} * \left\{ \log\left(\frac{t+P}{P}\right) \right\} \quad (16)$$

E-Bayesian estimate of the parameter θ based on $\pi_3(\theta, c, r)$, and is provided by

$$\begin{aligned} \hat{\theta}_{EBS_3} &= \int_0^1 \int_0^t \hat{\theta}_{BS} * \pi_3(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c}{S} \right\} * \frac{2r}{t^2} * dr dc \end{aligned}$$

on solving the above intervals, finally we get

$$\hat{\theta}_{EBS_3} = \left\{ \left(\frac{2n+1}{t^2} \right) * \left\{ P * \log\left(\frac{P}{t+P}\right) + t \right\} \right\} \quad (17)$$

4.3 E-Bayesian estimation of θ under the De-Groot loss function

E-Bayesian estimate of the parameter θ based on $\pi_1(\theta, c, r)$, and is provided by

$$\begin{aligned} \hat{\theta}_{EBD_1} &= \int_0^1 \int_0^t \hat{\theta}_{BD} * \pi_1(\theta, c, r) dr dc \\ \hat{\theta}_{EBD_1} &= \int_0^1 \int_0^t \left(\frac{n+c+1}{S} \right) * \left\{ \frac{2(t-r)}{t^2} \right\} * dr dc \\ &= \frac{2}{t^2} \left\{ \int_0^1 (n+c+1) dc * \int_0^t \left(\frac{t-r}{P+r} \right) * dr \right\} \end{aligned}$$

Where $S = \left\{ \alpha \sum_{i=1}^n x_i^{-2} + r \right\}$ and $P = \alpha \sum_{i=1}^n x_i^{-2}$

On solving the above integrals, we get

$$\hat{\theta}_{EBD_1} = \frac{(2n+3)}{t^2} * \left\{ (t+P) * \log\left(\frac{t+P}{P}\right) - t \right\} \quad (18)$$

E-Bayesian estimate of the parameter θ based on $\pi_2(\theta, c, r)$, and is provided by

$$\begin{aligned} \hat{\theta}_{EBD_2} &= \int_0^1 \int_0^t \hat{\theta}_{BD} * \pi_2(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c+1}{S} \right\} * \frac{1}{t} * dr dc \\ &= \frac{1}{t} \left\{ \int_0^1 (n+c) dc * \int_0^t \left(\frac{1}{S} \right) * dr \right\} \\ &= \frac{1}{t} \left\{ \int_0^1 (n+c+1) dc * \int_0^t \left(\frac{1}{P+r} \right) * dr \right\} \end{aligned}$$

Where $S = \left\{ \alpha \sum_{i=1}^n x_i^{-2} + r \right\}$ and $P = \alpha \sum_{i=1}^n x_i^{-2}$

$$\hat{\theta}_{EBD_2} = \frac{(2n+3)}{2t} * \left\{ \log \left(\frac{t+P}{P} \right) \right\} \tag{19}$$

E-Bayesian estimate of the parameter θ based on $\pi_3(\theta, c, r)$, and is provided by

$$\begin{aligned} \hat{\theta}_{EBD_3} &= \int_0^1 \int_0^t \hat{\theta}_{BD} * \pi_3(\theta, c, r) dr dc \\ &= \int_0^1 \int_0^t \left\{ \frac{n+c+1}{S} \right\} * \frac{2r}{t^2} * dr dc \\ &= \frac{2}{t^2} \left\{ \int_0^1 (n+c+1) dc * \int_0^t \left(\frac{r}{P+r} \right) * dr \right\} \\ &= \frac{2}{t^2} \left\{ \frac{2n+3}{2} * \left\{ t - P * \log \left(\frac{t+P}{P} \right) \right\} \right\} \end{aligned}$$

On solving the above intervals, finally we get

$$\hat{\theta}_{EBD_3} = \left\{ \frac{(2n+3)}{t^2} * \left\{ P * \log \left(\frac{P}{t+P} \right) + t \right\} \right\} \tag{20}$$

4.4 Properties of E-Bayesian estimates under Different Loss Functions

In this section, we will discuss the relationship amongst the different E-Bayesian estimators obtained under the Al-Bayyati loss function i.e, $\hat{\theta}_{EBA_1}, \hat{\theta}_{EBA_2}, \hat{\theta}_{EBA_3}$ ($i = 1,2,3$)

Theorem 4.1 E-Bayesian estimators obtained under the Al-Bayyati loss function will follow the following results:

- (i) $\hat{\theta}_{EBA_3} < \hat{\theta}_{EBA_2} < \hat{\theta}_{EBA_1}$
- (ii) $\lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_1}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_3})$

Proof (i) From (12) and (13), we get

$$\begin{aligned} \hat{\theta}_{EBA_1} - \hat{\theta}_{EBA_2} &= \frac{(2n+2d+1)}{t^2} * \left\{ (t+P) * \log \left(\frac{t+P}{P} \right) - t \right\} - \frac{2n+2d+1}{2t} * \left\{ \log \left(\frac{t+P}{P} \right) \right\} \\ &= \frac{(2n+2d+1)}{t} \left\{ \log \left(1 + \frac{t}{P} \right) \left(\frac{P}{t} + \frac{1}{2} \right) - 1 \right\} \end{aligned} \tag{21}$$

From (13) and (14), we get

$$\begin{aligned} \hat{\theta}_{EBA_2} - \hat{\theta}_{EBA_3} &= \frac{(2n+2d+1)}{2t} * \left\{ \log \left(\frac{t+P}{P} \right) \right\} - \frac{(2n+2d+1)}{t^2} * \left\{ t - P * \log \left(\frac{t+P}{P} \right) \right\} \\ &= \frac{(2n+2d+1)}{t} \left\{ \log \left(1 + \frac{t}{P} \right) \left(\frac{P}{t} + \frac{1}{2} \right) - 1 \right\} \end{aligned} \tag{22}$$

Since $\log \left(1 + \frac{t}{P} \right) = \left\{ \frac{t}{P} - \frac{t^2}{2P^2} + \frac{t^3}{3P^3} - \dots \dots \right\}$

$$\hat{\theta}_{EBA_1} - \hat{\theta}_{EBA_2} = \frac{(2n+2d+1)}{t} \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} - \dots \dots \right\}$$

$$\hat{\theta}_{EBA_1} - \hat{\theta}_{EBA_2} > 0$$

hence

$$\hat{\theta}_{EBA_1} > \hat{\theta}_{EBA_2} \tag{23}$$

Similarly,

$$\hat{\theta}_{EBA_2} - \hat{\theta}_{EBA_3} = \frac{(2n+2d+1)}{t} \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} - \dots \dots \right\}$$

$$\hat{\theta}_{EBA_2} - \hat{\theta}_{EBA_3} > 0, \text{ and}$$

hence

$$\hat{\theta}_{EBA_2} > \hat{\theta}_{EBA_3} \tag{24}$$

Combining (23) and (24), we get

$$\hat{\theta}_{EBA_3} < \hat{\theta}_{EBA_2} < \hat{\theta}_{EBA_1}$$

Proof (ii): From (21) and (22), we get

$$\hat{\theta}_{EBA_1} - \hat{\theta}_{EBA_2} = \hat{\theta}_{EBA_2} - \hat{\theta}_{EBA_3} \frac{(2n+2d+1)}{t} \left\{ \log \left(1 + \frac{t}{P} \right) \left(\frac{P}{t} + \frac{1}{2} \right) - 1 \right\}$$

After taking the limit, we get

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_1} - \hat{\theta}_{EBA_2}) = \lim_{P \rightarrow \infty} \frac{(2n + 2d + 1)}{t} \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} \dots \dots \right\}$$

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_2} - \hat{\theta}_{EBA_3}) = \lim_{P \rightarrow \infty} \frac{(2n + 2d + 1)}{t} \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} \dots \dots \right\}$$

On solving the above, we have

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_1} - \hat{\theta}_{EBA_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_2} - \hat{\theta}_{EBA_3}) = 0$$

Hence

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_1}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBA_3})$$

Now we will discuss the relationship amongst the different E-Bayesian estimators obtained under the Square Error loss function i.e, $\hat{\theta}_{EBS_1}, \hat{\theta}_{EBS_2}, \hat{\theta}_{EBS_3}$ ($i = 1,2,3$)

Theorem 4.2 E-Bayesian estimators obtained under the Squared Error loss function will follow the following results:

- (i) $\hat{\theta}_{EBS_3} < \hat{\theta}_{EBS_2} < \hat{\theta}_{EBS_1}$
- (ii) $\lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_1}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_3})$

Proof (i): From (15) and (16), we get

$$\begin{aligned} \hat{\theta}_{EBS_1} - \hat{\theta}_{EBS_2} &= \frac{2n + 1}{t^2} * \left\{ (t + P) * \log \left(\frac{t + P}{P} \right) - t \right\} - \frac{2n + 1}{2t} * \left\{ \log \left(\frac{t + P}{P} \right) \right\} \\ \hat{\theta}_{EBS_1} - \hat{\theta}_{EBS_2} &= \frac{(2n + 1)}{t} * \left\{ \log \left(\frac{t + P}{P} \right) \left(\frac{P}{t} + \frac{1}{2} \right) - 1 \right\} \end{aligned} \tag{25}$$

$$\begin{aligned} \hat{\theta}_{EBS_1} - \hat{\theta}_{EBS_2} &> 0 \\ \hat{\theta}_{EBS_1} &> \hat{\theta}_{EBS_2} \end{aligned} \tag{26}$$

Similarly, from (16) and (17), we get

$$\begin{aligned} \hat{\theta}_{EBS_2} - \hat{\theta}_{EBS_3} &= \frac{(2n + 1)}{2t} * \left\{ \log \left(\frac{t + P}{P} \right) \right\} - \frac{(2n + 1)}{t^2} * \left\{ t - P * \log \left(\frac{t + P}{P} \right) \right\} \\ \hat{\theta}_{EBS_2} - \hat{\theta}_{EBS_3} &= \frac{(2n + 1)}{t} * \left\{ \log \left(1 + \frac{t}{P} \right) \left(\frac{P}{t} + \frac{1}{2} \right) - 1 \right\} \end{aligned} \tag{27}$$

$$\begin{aligned} \hat{\theta}_{EBS_2} - \hat{\theta}_{EBS_3} &> 0 \\ \hat{\theta}_{EBS_2} &> \hat{\theta}_{EBS_3} \end{aligned} \tag{28}$$

Combining (26) and (28), we get

$$\hat{\theta}_{EBS_3} < \hat{\theta}_{EBS_2} < \hat{\theta}_{EBS_1}$$

Proof (ii): From (25) and (27), we get

$$\hat{\theta}_{EBS_1} - \hat{\theta}_{EBS_2} = \frac{(2n + 1)}{t} * \left\{ \log \left(1 + \frac{t}{P} \right) \left(\frac{P}{t} + \frac{1}{2} \right) - 1 \right\} = \hat{\theta}_{EBS_2} - \hat{\theta}_{EBS_3}$$

After taking the limit, we get

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_1} - \hat{\theta}_{EBS_2}) = \lim_{P \rightarrow \infty} \frac{(2n + 1)}{t} \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} - \dots \dots \right\} = 0$$

and

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_2} - \hat{\theta}_{EBS_3}) = \lim_{P \rightarrow \infty} \frac{(2n+1)}{t} \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} - \dots \dots \right\} = 0$$

On solving, we have

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_1} - \hat{\theta}_{EBS_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_2} - \hat{\theta}_{EBS_3}) = 0$$

hence

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_1}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBS_3})$$

Now we will discuss the relationship amongst the different E-Bayesian estimators obtained under the De-Groot loss function i.e, $\hat{\theta}_{EBD_1}, \hat{\theta}_{EBD_2}, \hat{\theta}_{EBD_3}$ ($i = 1,2,3$)

Theorem 4.3 E-Bayesian estimators obtained under the Square Error loss function will follow the following results:

- (i) $\hat{\theta}_{EBD_3} < \hat{\theta}_{EBD_2} < \hat{\theta}_{EBD_1}$
- (ii) $\lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_1}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_3})$

Proof (i): From (18) and (19), we get

$$\begin{aligned} \hat{\theta}_{EBD_1} - \hat{\theta}_{EBD_2} &= \frac{2n+3}{t^2} * \left\{ (t+P) * \log\left(\frac{t+P}{P}\right) - t \right\} - \frac{2n+3}{2t} * \left\{ \log\left(\frac{t+P}{P}\right) \right\} \\ \hat{\theta}_{EBD_1} - \hat{\theta}_{EBD_2} &= \frac{(2n+3)}{t} * \left\{ \log\left(\frac{t+P}{P}\right) \left(\frac{P}{t} + \frac{1}{2}\right) - 1 \right\} \end{aligned} \quad (29)$$

$$\begin{aligned} \hat{\theta}_{EBD_1} - \hat{\theta}_{EBD_2} &> 0 \\ \hat{\theta}_{EBD_1} &> \hat{\theta}_{EBD_2} \end{aligned} \quad (30)$$

Similarly, from eq. (19) and (20), we get

$$\begin{aligned} \hat{\theta}_{EBD_2} - \hat{\theta}_{EBD_3} &= \frac{(2n+3)}{2t} * \left\{ \log\left(\frac{t+P}{P}\right) \right\} - \frac{(2n+3)}{t^2} * \left\{ t - P * \log\left(\frac{t+P}{P}\right) \right\} \\ \hat{\theta}_{EBD_2} - \hat{\theta}_{EBD_3} &= \frac{(2n+3)}{t} * \left\{ \log\left(1 + \frac{t}{P}\right) \left(\frac{P}{t} + \frac{1}{2}\right) - 1 \right\} \end{aligned} \quad (31)$$

$$\begin{aligned} \hat{\theta}_{EBD_2} - \hat{\theta}_{EBD_3} &> 0 \\ \hat{\theta}_{EBD_2} &> \hat{\theta}_{EBD_3} \end{aligned} \quad (32)$$

Combining (30) and (32), we get

$$\hat{\theta}_{EBD_3} < \hat{\theta}_{EBD_2} < \hat{\theta}_{EBD_1}$$

Proof (ii): From (29) and (31), we get

$$\hat{\theta}_{EBD_1} - \hat{\theta}_{EBD_2} = \frac{(2n+3)}{t} * \left\{ \log\left(1 + \frac{t}{P}\right) \left(\frac{P}{t} + \frac{1}{2}\right) - 1 \right\} = \hat{\theta}_{EBD_2} - \hat{\theta}_{EBD_3}$$

After taking the limit, we get

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_1} - \hat{\theta}_{EBD_2}) = \lim_{P \rightarrow \infty} \frac{(2n+3)}{t} * \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} - \dots \dots \right\} = 0$$

And

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_2} - \hat{\theta}_{EBD_3}) = \lim_{P \rightarrow \infty} \frac{(2n+3)}{t} * \left\{ \frac{t^2}{12P^2} + \frac{t^3}{6P^3} - \dots \dots \right\} = 0$$

On solving, we have

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_1} - \hat{\theta}_{EBD_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_2} - \hat{\theta}_{EBD_3}) = 0$$

$$\lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_1}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_2}) = \lim_{P \rightarrow \infty} (\hat{\theta}_{EBD_3})$$

Hence the proof is complete.

The part (a) of theorem 4.(i), 4.(ii) and 4.(iii) shows that with different priors (9)-(11) of the parameters c and r , the associated E-Bayesian estimate $\hat{\theta}_{EBA_i}$, $\hat{\theta}_{EBD_i}$, and $\hat{\theta}_{EBS_i}$; ($i=1,2,3$) are different. The property (b) of the theorems shows that $\hat{\theta}_{EBA_i}$, c ; ($i=1,2,3$) are asymptotically equivalent to each other as $\sum_{i=1}^n x_i^{-2} \rightarrow \infty$, that means $\hat{\theta}_{EBA_i}$; ($i=1,2,3$) are all close to each other when $\sum_{i=1}^n x_i^{-2}$ is sufficiently large and $\hat{\theta}_{EBD_i}$, and $\hat{\theta}_{EBS_i}$; ($i = 1,2,3$) are also close to each other.

5. Simulation Study

In order to compare the performance of Bayesian and E-Bayesian techniques of estimation, a simulation study was conducted using MatLab. We chose a sample of size of $n=20, 50, 70, 100, 120$ to represent small, medium and large data set. The following steps were conducted:

1. The shape (α) and scale (θ) parameters has been fixed at 0.5 and 0.25 respectively.
2. For given value of t , we generate c and r from uniform and gamma distribution respectively.
3. For given value of α and θ , we generate a random sample of different sizes from Exponentiated inverse Rayleigh distribution (EIRD) using the quantile function.
4. The above steps are iterated 1000 times to find the MSE of Bayesian and E-Bayesian estimates of scale parameter using different loss functions.

5. The MSE for Bayesian and E-Bayesian estimates under different loss function are shown in table 1.
6. The MSE of θ for Bayesian and E-Bayesian estimation under different loss functions are also illustrated in Figure 1, 2 and 3.

Table 1: Mean Squared Error (MSE) of θ under different loss functions for $\alpha = 0.25, \theta = 0.5, t = 1.5, d = 3$.

n	$\hat{\theta}_{BA}$	$\hat{\theta}_{EBA_1}$	$\hat{\theta}_{EBA_2}$	$\hat{\theta}_{EBA_3}$	$\hat{\theta}_{BS}$	$\hat{\theta}_{EBS_1}$	$\hat{\theta}_{EBS_2}$	$\hat{\theta}_{EBS_3}$	$\hat{\theta}_{BD}$	$\hat{\theta}_{EBD_1}$	$\hat{\theta}_{EBD_2}$	$\hat{\theta}_{EBD_3}$
20	0.1357	0.1315	0.1304	0.1294	0.1029	0.0994	0.0987	0.0979	0.1134	0.1096	0.1087	0.1079
50	0.0958	0.0948	0.0945	0.0942	0.0854	0.0844	0.0842	0.0839	0.0888	0.0878	0.0876	0.0873
70	0.0735	0.0727	0.0725	0.0724	0.0677	0.0668	0.0667	0.0666	0.0696	0.0688	0.0686	0.0685
100	0.0604	0.0599	0.0599	0.0598	0.0569	0.0565	0.0564	0.0563	0.0581	0.0576	0.0576	0.0575
120	0.0546	0.0543	0.0543	0.0542	0.0519	0.0517	0.0517	0.0516	0.0528	0.0526	0.0525	0.0525

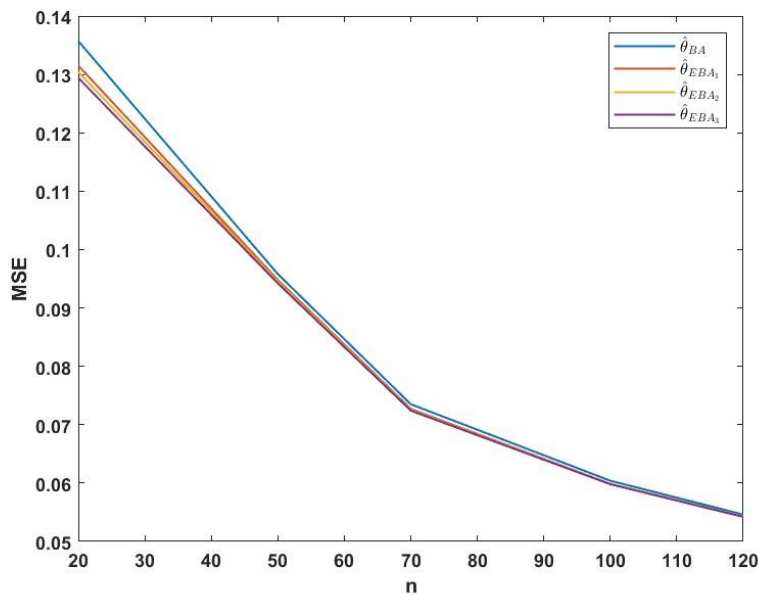


Figure 1: MSE of Bayesian and E-Bayesian estimates of θ under Al-Bayyati loss function.

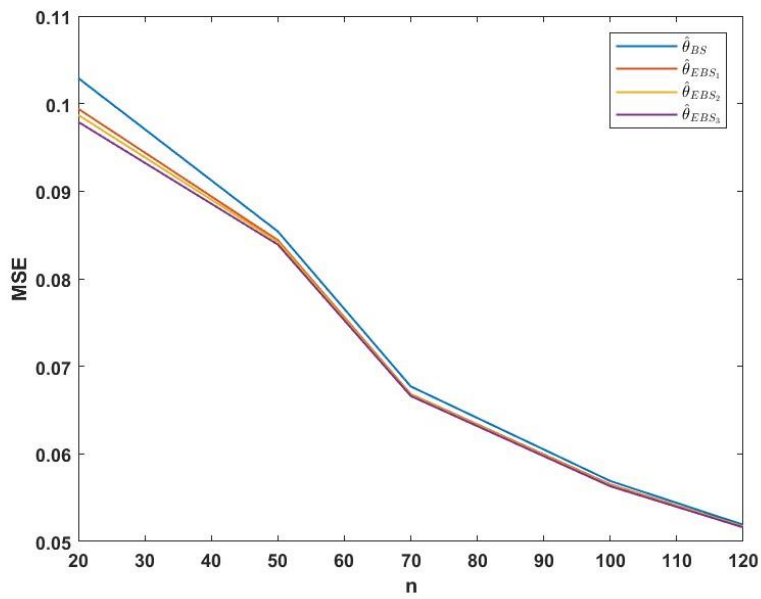


Figure2: MSE of Bayesian and E-Bayesian estimates of θ under Squared error loss function.

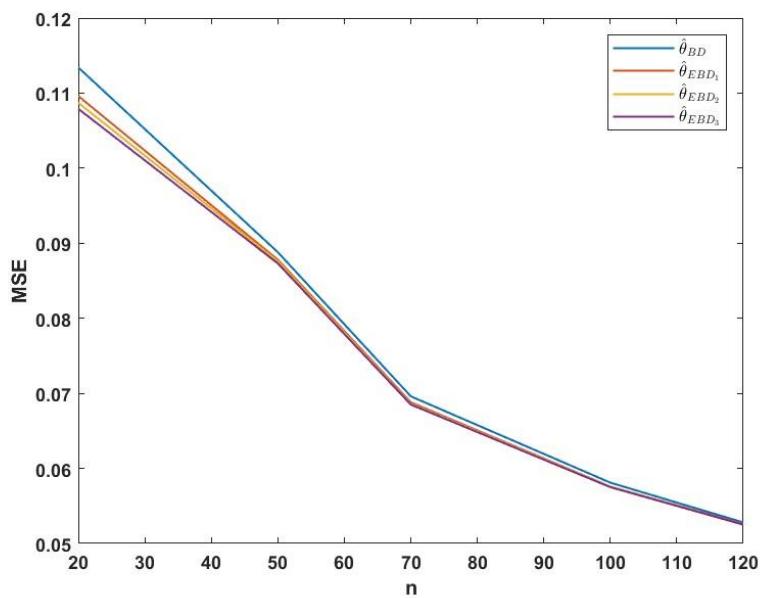


Figure 3: MSE of Bayesian and E-Bayesian estimates of θ under De-Groot loss function.

6. Real Data Analysis

The dataset was sourced from [13], comprising monthly actual tax revenues in Egypt spanning from January 2006 to November 2010. The data, expressed in 1000 million Egyptian pounds, are as follows: 5.9, 20.4, 14.9, 16.2, 17.2, 7.8, 6.1, 9.2, 10.2, 9.6, 13.3, 8.5, 21.6, 18.5, 5.1, 6.7, 17, 8.6, 9.7, 39.2, 35.7, 15.7, 9.7, 10, 4.1, 36, 8.5, 8, 9.2, 26.2, 21.9, 16.7, 21.3, 35.4, 14.3, 8.5, 10.6, 19.1, 20.5, 7.1, 7.7, 18.1, 16.5, 11.9, 7, 8.6, 12.5, 10.3, 11.2, 6.1, 8.4, 11, 11.6, 11.9, 5.2, 6.8, 8.9, 7.1, 10.8. The Kolmogorov-Smirnov (K-S) statistic's value is 0.082194, with an associated p-value of 0.8203. This suggests that the Exponentiated Inverse Rayleigh Distribution (EIRD) is the best fit for this dataset. Based on this data, the Maximum Likelihood estimates yield $\hat{\theta} = 8.9362$ and $\hat{\alpha} = 9.8028$. Figure 4 and 5 displays the histogram and the estimated Cumulative Distribution Function (CDF) of the EIRD for the dataset, while the estimated Bayesian and E-Bayesian values are presented in Table 2.

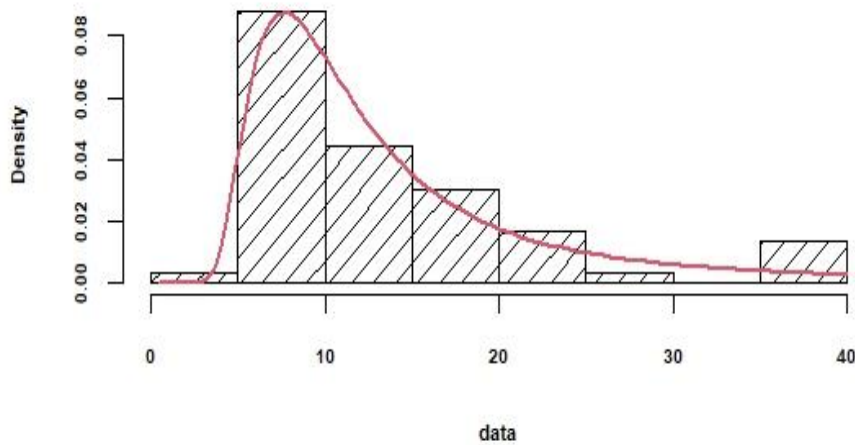


Figure 4: Histogram and the fitted density for the monthly actual taxes revenue.

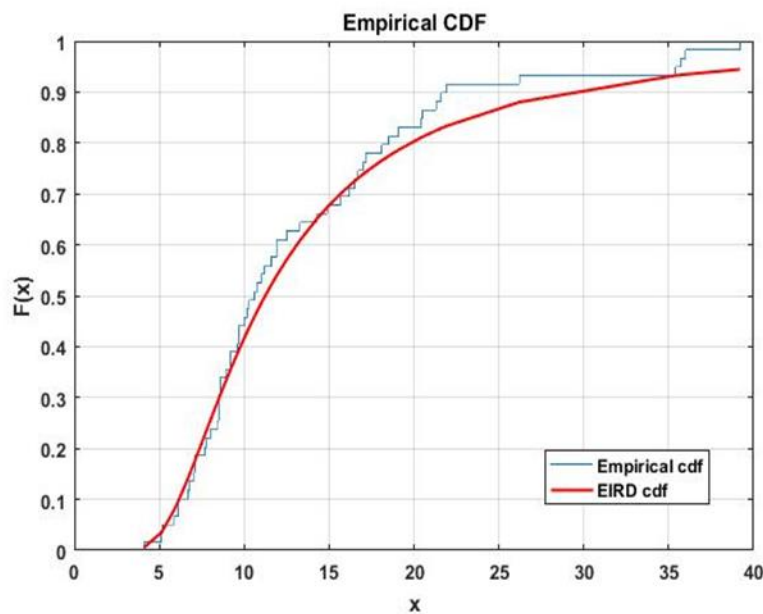


Figure 5: Plot for the ECDF of the EIRD model.

Table 2: Bayesian and E-Bayesian estimates of θ based on real dataset.

$\hat{\theta}_{BA}$	$\hat{\theta}_{EBA_1}$	$\hat{\theta}_{EBA_2}$	$\hat{\theta}_{EBA_3}$	$\hat{\theta}_{BS}$	$\hat{\theta}_{EBS_1}$	$\hat{\theta}_{EBS_2}$	$\hat{\theta}_{EBS_3}$	$\hat{\theta}_{BD}$	$\hat{\theta}_{EBD_1}$	$\hat{\theta}_{EBD_2}$	$\hat{\theta}_{EBD_3}$
94.257	129.236	95.558	71.446	89.769	123.032	90.971	68.017	91.265	125.100	92.500	69.160

7. Conclusion

This paper focuses on employing Bayesian and E-Bayesian methods to estimate the scale parameter of the Exponentiated Inverse Rayleigh distribution (EIRD) through the use of diverse loss functions. Additionally, certain properties of the E-Bayesian estimates are explored. Notably, the results of a simulation study reveal that E-Bayesian estimates exhibit a smaller Mean Squared Error (MSE) in comparison to Bayesian estimates, thereby demonstrating their enhanced efficiency. Among the E-Bayesian estimates, the third one stands out as the most effective.

Moreover, the analysis highlights that the Squared Error loss function outperforms the Al-Bayyati and De-Groot loss functions, exhibiting a smaller MSE. The conclusions drawn from the simulation study are further substantiated by validating the findings through a real-life dataset.

References

- [1] Rehman, S. and I, Dar, Sajjad. (2015). Bayesian analysis of Exponentiated Inverse Rayleigh distribution under different priors. *M. Phil. Thesis*.
- [2] Voda, V. Gh. (1972). On the Inverse Rayleigh Distributed Random Variable. *Journal of applied and statistical research*. JUSE. 19(4): 13-21.
- [3] Siddiqui, M.M. (1962). Some problems connected with Rayleigh distributions. *Journal of Research of the National Institute of Standards and Technology* 60D, 167-174.
- [4] Soliman, A., Essam A. Amin and Alaa A. Abd-EI Aziz. (2010). Estimation and Prediction from Inverse Rayleigh Distribution Based on Lower Record Values. *Applied Mathematical Sciences*. 4(62): 3057-3066.
- [5] Reshi, J.A., Ahmed, A., Ahmad, S.P. (2014). Bayesian Analysis of Scale Parameter of the Generalized Inverse Rayleigh Model Using Different Loss Functions. *International Journal of Modern Mathematical Sciences*,10(2):151- 162.
- [6] Dey, S. (2012). Bayesian estimation of the parameter and reliability function of an inverse Rayleigh distribution. *Malaysian Journal of Mathematical Sciences*. 6(1):113-124.
- [7] Sindhua, T.N., Aslama M. and Ferozeb, N. (2013). Bayes estimation of the parameters of the inverse Rayleigh distribution for left censored data. *Probability Statistical Forum*, 6:42-59.
- [8] Okasha, H. M., (2014). E -Bayesian estimation for the Lomax distribution based on type-II censored data. *Journal of the Egyptian Mathematical Society*, 22:489-495.
- [9] Al-Bayyati, H.N. (2002). Comparing methods of estimating Weibull failure models using simulation, Unpublished PhD thesis, College of Administration and Economics, Baghdad University, Iraq.
- [10] Mood, A., Graybill, F. A., Boes, D. (1974). Introduction to the Theory of Statistics. *McGraw-Hill Series in Probability and Statistics*.
- [11] Degroot, M. h. (1970). Optimal Statistical Decision, McGraw-Hill Inc.
- [12] Han, M. (2007). E-Bayesian estimation of failure probability and its application. *Mathematical and Computer Modelling*, 45:1272-1279.
- [13] Mead, M.E. (2016). On five- parameter Lomax distribution: Properties and applications. *Pak. J. Stat. Oper. Res.* 1:185-199.

AVAILABILITY ANALYSIS FOR IDENTIFICATION OF CRITICAL FACTOR OF A THERMAL POWER PLANT

Pardeep Kumar¹, Vipin Kumar Sharma², Dinesh Kumar³

¹Maharishi Markandeshwar (Deemed to be University), Ambala, India,

²IIMT University, Meerut, U.P., India

³NIT Kurukshetra, Haryana, India

¹pardeepkamboj@yahoo.com

²vipin2871985@gmail.com

³dinesh_61900120@nitkkr.ac.in

Abstract

In the present stimulated business environment, power sector is playing a major role in the economic growth of India. During the last 20 years, the country had been facing a poor supply of energy and this supply-demand gap is increasing continuously. So, it is important for power plants to improve its power generation capacity drastically by reducing the failure rate. In the present paper, to analyze the causes of poor availability, thermal power plant has divided into six different systems and a system comprising of waste gases heating system has been considered. With the help of transition diagram, mathematical equations have been used to find out the availability. After analyzing, it was found that the value of availability is very low and boiler tube failure is one of the most critical factors for this low availability of system. Economizer zone has identified having long existence time of failures and frequency of occurrence is very high. So, minimizing the failure rate with the help of a proper maintenance schedule will result in decreasing the shutdown period of the plant and increasing the system availability.

Keywords: Thermal Power Plant; Performance Evaluation; Availability, Boiler; Tube Failure; Economizer

I. Introduction

In today's competitive world, it becomes necessary that thermal power plant will be available for long run without any failure. In India, total installed capacity of electricity generation is 3,30,354 MW while total thermal installed capacity is 2,20,456 MW i.e. 66.8% of the total installed capacity (refer table 1). The major contribution almost 59% in thermal installed capacity is coal fired thermal power plant. For continuous power production, boiler becomes the backbone of a thermal power plant. Boiler tube failure is one of the critical problems which are facing the thermal power plant and influence the rate of power generation. This loss of generation increases the operating cost of plant and significant amount of water is being waste. Availability analysis gives the necessary information about various parameters of the system.

Table 1: Installed Capacity for Different Source of Fuel

Fuel Used	Installed Capacity (MW)	% of Total
Total Thermal	2,20,456	66.8 %
Coal	1,94,433	58.9 %
Gas	25,185	7.6 %
Oil	838	0.3 %
Hydro	44,614	13.5 %
Nuclear	6,780	2.1 %
Renewable Energy Sources	58,303	17.7 %
Total	3,30,354	

II. Literature Review

From last decades, Complexity in the industries is increasing day by day, so many researchers using Markov Method for the performance evaluation of complex system in process industry. Tsarouhas [1] computed the parameters on which the reliability of the machine of the ice cream plant is dependent. Dai et al. [2] presented as a model for a system which is centralized heterogeneously widely used in distributed system design. With the help of this model the reliability of the distributed service is find out to provide a best service in a distributed environment. Gupta et al [3] studied the critical components on which the reliability of the plant is mainly dependent and on the basis of these components, a Decision Support system for for minimizing the failure rate of the industry has been decided. Gupta [4] discussed the DSS and performance modelling for a subsystem of feed water in a system of thermal power plant by using performance Modelling and analysis. With the help of a transition diagram, differential equations are generated and then study state probabilities are applied to find out the performance level for the combination of different failure and repair rate of all the sub-system. Using Markov process and probability theory, Gupta [5, 8] found the availability and reliability of a system of a thermal power plant. The author found that the reliability or availability decreases with increasing failure rate while the availability improves with increasing the repair rate for different sub-systems. On the basis of this, author made a maintenance policy for all the subsystems of a thermal power plant. To evaluate the reliability parameter, Gupta [6] expressed the mathematical formulation and expression for mission reliability and availability of a complex polymer powder production system with more realistic and practical assumption. Swiderski et al. [7] discussed the models of semi-Markov and Markov as best conventional tools to evaluate the availability and reliability of each subsystem of the full system. Kumar [9] developed a mathematical model based on the Markov birth-death process and developed differential equations based on probabilistic approach and solved these equations recursively. Khanduja et al. [10-11] developed a system for performance evaluation with the help of mathematical formulation and this performance evaluation deals with quantitative analysis of all critical factors which affects the maintenance decision support system of a paper plant. Lai et al. [12] designed a Markov based model and derived the equations to obtain the steady-state availability by considering both software and hardware failures. Kumar et al. [13-15] studied the behavior of a thermal power plant to improve the performance by minimizing the failure rate so that all the systems in thermal power plant can be function effectively. Sabouhi [16] proposed reliability oriented analysis to drive the

mathematical expression for the analysis of availability of the critical component of the plant so that an effective maintenance plan can be scheduled. Hassan et al. [17] purposed a stochastic model for liquefied natural gas plant using Markov analysis. Parkash and Tiwari [18-19] suggested a approach which can helps both engineers and managers to enhance the performability of the system by utilizing the best combination of failure and repair rates.

III. Availability Analysis of Waste Gases System of a Thermal Power Plant

The flow diagram of thermal power plant consisting of waste gases system (refer figure 1) shows that flue or waste gases from furnace flow upward and this waste heat is utilized in superheater, economizer and air preheater to raise the temperature of some extent of steam, feed water and air.

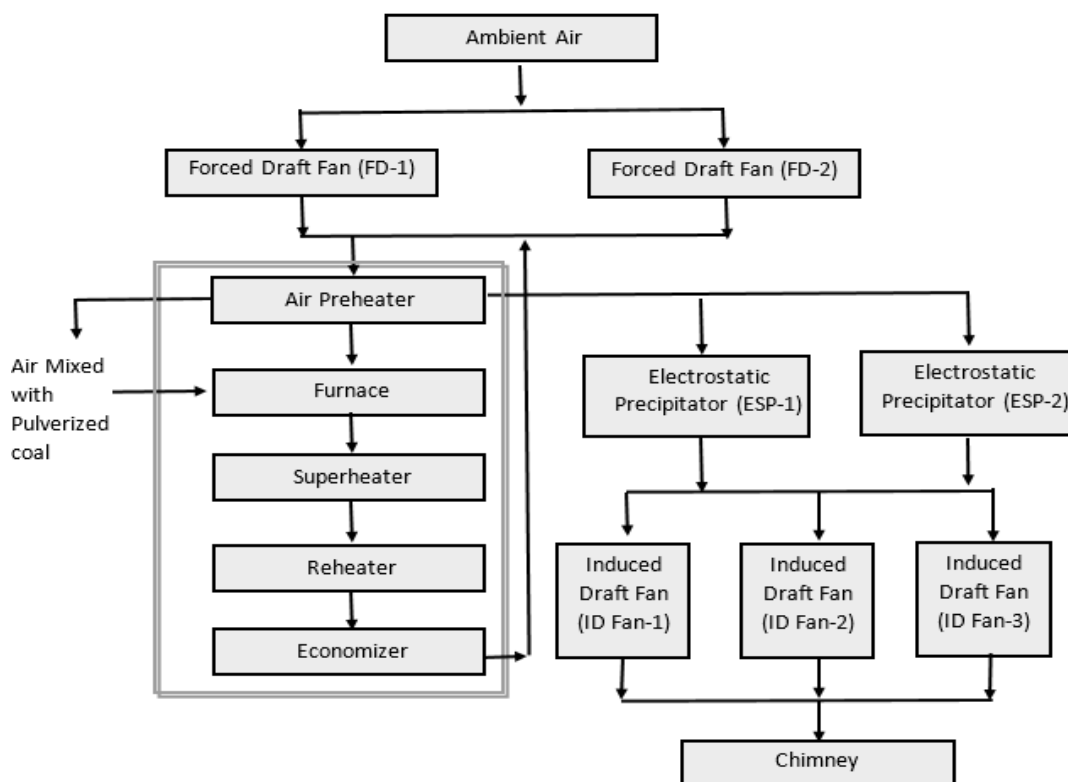


Figure 1: Flow Diagram of Waste Gases Heating System

To find the availability, this system is divided further in four different subsystems.

Subsystem A: it consists of furnace, superheater, economizer and air preheater and arranged in series to establish a single subsystem.




Subsystem B: It consists of two electrostatic precipitator (ESP) which makes a single subsystem.

Subsystem C: Two forced draught fans working in parallel consists a subsystem.

Subsystem D: Three induced draught fans (ID Fan) arranged in parallel, creating one subsystem.

The table 2 shows some notation which are used to construct the transition diagram as shown in figure 2.

Table 2: Notation Used

Full capacity States (without standby)	A_2, A_4, A_5, A_6
One unit of subsystem A_1 and A_3 are in failed state and the system is working in full capacity with stand by unit.	A_1^* and A_3^*
Failed states	$a_1, a_2, a_3, a_4, a_5, a_6$
Reduced capacity states	A_4^f
Failure rates	$\alpha_i, i=1$ to 6
Repair rates	$\beta_i, i=1$ to 6
System Working at Full Capacity	
System Working at Reduced Capacity	
System in Failed State	

I. Performance Modelling of Waste Gases System

The mathematical equations are derived using Chapman–Kolmogorov equation with the help of transition diagram.

$$P_0'(t) + \sum_{i=1}^4 \lambda_i P_0(t) = \beta_1 P_{12}(t) + \beta_2 P_6(t) + \beta_3 P_3(t) + \beta_4 P_1(t) \quad (1)$$

$$P_1'(t) + \sum_{i=1}^4 (\lambda_i + \beta_4) P_1(t) = \beta_1 P_{11}(t) + \beta_2 P_7(t) + \beta_3 P_4(t) + \beta_4 P_2(t) + \beta_4 P_0(t) \quad (2)$$

$$P_2'(t) + \sum_{i=1}^4 (\lambda_i + \beta_4) P_2(t) = \beta_1 P_{10}(t) + \beta_2 P_8(t) + \beta_3 P_5(t) + \beta_4 P_9(t) + \beta_4 P_1(t) \quad (3)$$

$$P_3'(t) + \sum_{i=1}^4 (\lambda_i + \beta_3) P_3(t) = \sum_{i=1}^3 \beta_i P_{i-18}(t) + \lambda_3 P_0(t) + \beta_4 P_4(t) \quad (4)$$

$$P_4'(t) + \sum_{i=1}^4 (\lambda_i + \beta_3 + \beta_4) P_4(t) = \sum_{i=1}^3 \beta_i P_{i-15}(t) + \lambda_{33} P_1(t) + \lambda_4 P_3(t) + \beta_4 P_5(t) \quad (5)$$

$$P_5'(t) + \sum_{i=1}^4 (\lambda_i + \beta_4 + \beta_3) P_5(t) = \sum_{i=1}^4 \lambda_i P_{i-12}(t) + \lambda_4 P_4(t) + \lambda_3 P_2(t) \quad (6)$$

$$P_i'(t) + \beta_m P_i(t) = \lambda_m P_j(t) \quad (7)$$

When

$$m = 1, \text{ then } i = 12, j = 0; i = 11, j = 1; i = 10, j = 2; i = 13, j = 3; i = 16, j = 4; i = 19, j = 5;$$

$$m = 2, \text{ then } i = 6, j = 0; i = 7, j = 1; i = 8, j = 2; i = 14, j = 3; i = 17, j = 4; i = 20, j = 5$$

$$m = 3, \text{ then } i = 15, j = 3; i = 18, j = 4; i = 21, j = 5$$

$$m = 4, \text{ then } i = 9, j = 2; i = 22, j = 5$$

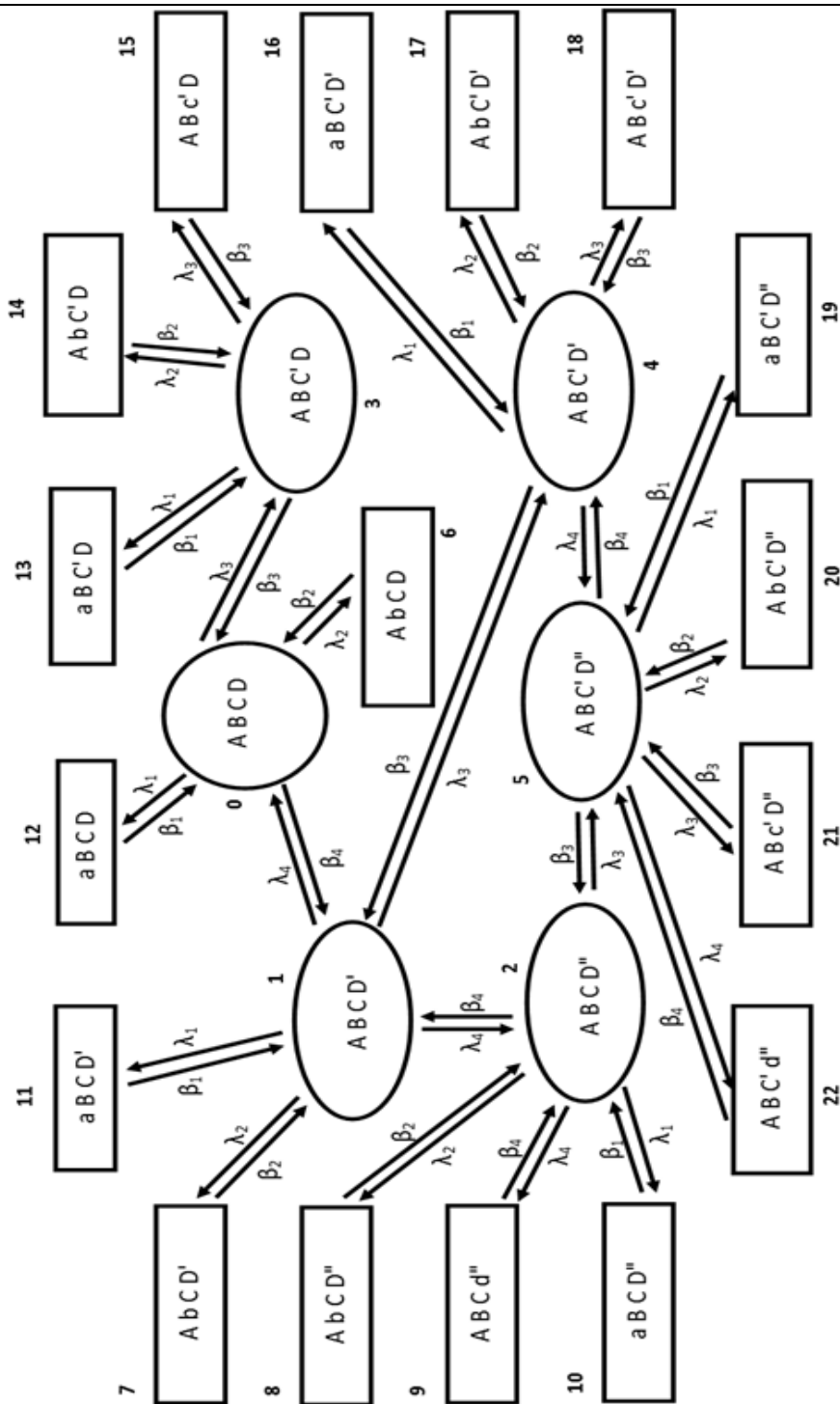


Figure 2: Transition Diagram of Waste Gases Heating System

IV. Steady State Availability of Waste Gases System

By putting derivatives = 0 as $t \rightarrow \infty$ in equations 1 to 7 and solved by recursive method, the following values obtained of all 23 states probabilities (P_0 to P_{22}) in terms of full working state probability i.e. P_0 .

$$P_1 = C_{12}P_0$$

$$P_2 = C_{13}P_0$$

$$P_3 = C_{14}P_0$$

$$P_4 = C_{11}P_0$$

$$\begin{array}{llll}
 P_5 = C_{15}P_0 & P_6 = \frac{\phi_2}{\lambda_2} P_0 & P_7 = \frac{\phi_2}{\lambda_2} C_{12}P_0 & P_8 = \frac{\phi_2}{\lambda_2} C_{13}P_0 \\
 P_9 = \frac{\phi_4}{\lambda_4} C_{13}P_0 & P_{10} = \frac{\phi_1}{\lambda_1} C_{13}P_0 & P_{11} = \frac{\phi_1}{\lambda_1} C_{12}P_0 & P_{12} = \frac{\phi_1}{\lambda_1} P_0 \\
 P_{13} = \frac{\phi_1}{\lambda_1} C_{14}P_0 & P_{14} = \frac{\phi_2}{\lambda_2} C_{14}P_0 & P_{15} = \frac{\phi_3}{\lambda_3} C_{14}P_0 & P_{16} = \frac{\phi_1}{\lambda_1} C_{11}P_0 \\
 P_{17} = \frac{\phi_2}{\lambda_2} C_{11}P_0 & P_{18} = \frac{\phi_3}{\lambda_3} C_{11}P_0 & P_{19} = \frac{\phi_1}{\lambda_1} C_{15}P_0 & P_{20} = \frac{\phi_2}{\lambda_2} C_{15}P_0 \\
 P_{21} = \frac{\phi_3}{\lambda_3} C_{15}P_0 & P_{22} = \frac{\phi_4}{\lambda_4} C_{15}P_0 & &
 \end{array}$$

The probability of full working capacity, namely, P_0 determined by using normalizing condition: i.e (sum of the probabilities of all working states, reduced capacity and failed states is equal to 1)

$$\sum_{i=0}^{22} P_i = 1, \text{ therefore}$$

$$P_0 = \frac{1}{\left[(1 + C_{11} + C_{12} + C_{13} + C_{14} + C_{15}) \left(1 + \left(\frac{\lambda_1}{\beta_1} + \frac{\lambda_2}{\beta_2} \right) + \frac{\lambda_3}{\beta_3} (C_{11} + C_{14} + C_{15}) + \frac{\lambda_4}{\beta_4} (C_{14} + C_{15}) \right) \right]}$$

Where

$$\begin{array}{lll}
 C_1 = \lambda_3 + \beta_4 & C_2 = \lambda_3 + \lambda_4 + \beta_4 & C_3 = \lambda_3 + \beta_4 \\
 C_4 = \lambda_4 + \beta_3 & C_5 = \lambda_4 + \beta_3 + \beta_4 & C_6 = \beta_3 + \beta_4 \\
 C_7 = \frac{\lambda_4 \beta_3}{C_3 C_6 - \lambda_3 \beta_3} & C_8 = \frac{\lambda_4 C_6}{C_3 C_6 - \lambda_3 \beta_3} & C_9 = \frac{\lambda_3 + C_7 \beta_4}{C_3 - \beta_4 C_8} \\
 C_{10} = \frac{\lambda_4}{C_2 - \beta_4 C_8} & C_{11} = \frac{C_3 C_4 - \lambda_3 \beta_3 - C_{10} C_4 \beta_4}{\beta_3 \beta_4 - C_9 C_4 \beta_4} & \\
 C_{12} = C_9 C_{11} + C_{10} & C_{13} = C_7 C_{11} + C_8 C_{12} & C_{14} = \frac{C_{11} \beta_4 + \lambda_3}{C_4} \\
 C_{15} = \frac{C_{13} \lambda_3 + C_{11} \lambda_4}{C_6} & &
 \end{array}$$

Hence

$$Av = P_0 + P_1 + P_2 + P_3 + P_4 + P_5$$

$$Av = P_0 (1 + C_{11} + C_{12} + C_{13} + C_{14} + C_{15})$$

Table 3 shows the variation of system availability with different possible combination of failure and repair rates for waste gases heating system. System availability decreases (0.9402 - 0.7985) appreciably by 14.2 % with increasing the failure rate from 0.005 (once in 200 hrs) to 0.040 (once in 25 hrs). Similarly other values show the decreasing trend of availability. Correspondingly, repair rate also effect the value of availability, as repair rate increases from 0.10 (once in 10 hrs) to 0.50

(once in 2 hrs), the system availability increases (0.9784-0.9402) drastically by 3.82 %.

Table 3: Availability Matrix for Waste Gases Heating System

λ_1 ↓	β_1 →	0.1	0.2	0.3	0.4	0.5	Constant values
0.005		0.9402	0.9596	0.9676	0.9712	0.9784	$\lambda_2 = 0.010,$ $\lambda_3 = 0.0033,$ $\lambda_4 = 0.005,$ $\beta_2 = 0.20,$ $\beta_3 = 0.25,$ $\beta_4 = 0.20$
0.0066		0.9298	0.9424	0.9639	0.9701	0.9731	
0.100		0.9192	0.9322	0.9412	0.9565	0.9698	
0.020		0.9085	0.9408	0.9533	0.9599	0.9653	
0.040		0.7985	0.9164	0.9398	0.9492	0.9619	

This table also shows that the failure rate influences the availability of the system. System availability can be improved by decreasing the failure rate. Maintenance data shows that boiler tube failure is one of the most critical reasons for low availability of waste gases heating system.

V. Results

The performance evaluation of waste gases heating system has been done with the help of simulation modeling. Table 3 showed the variation in the system performance with the variation in failure and repair rates of its different components. The various availability levels (A_v) for different combinations of failure and repair rates were also calculated and found that availability of system decreases appreciably as failure rate increases. To improve the availability of the system it becomes essential that reduces the significant causes of failure.

References

- [1] Tsarouhas P., "Reliability, Availability, and Maintainability (RAM) Study of an Ice Cream Industry", *Appl. Sci.* 2020, Vol.10, pp.1-20.
- [2] Dai, Yuan-Shun, Min Xie, Kim-Leng Poh, and G. Q. Liu. "A study of service reliability and availability for distributed systems," *Reliability Engineering & System Safety*, vol. 79, no. 1, pp. 103-112, 2003.
- [3] Gupta, P., J. Singh, and I. P. Singh. "Availability Analysis of Soap Cakes production System–A Case Study," In *Proc. National Conference on Emerging Trends in Manufacturing System*, SLIET, Longowal (Punjab) Pb, 2004, pp. 283-295..
- [4] Gupta, S., P. C. Tewari, and A. K. Sharma. "A performance modeling and decision support system for a feed water unit of a thermal power plant," *South African Journal of Industrial Engineering*, vol. 19, no. 2, pp. 125-134, 2008.
- [5] Gupta, S., P. C. Tewari, and A. K. Sharma. "Reliability and availability analysis of the ash handling unit of steam thermal power plant," *South African Journal of Industrial Engineering*, vol. 20, no. 1, pp. 147-158, 2009.
- [6] Guha, P., J. Singh, and I. P. Singh. "Mission reliability and availability prediction of flexible polymer powder production system," *OPSEARCH-New Delhi*, vol. 42, no. 2, pp. 152-167, 2005.
- [7] Swiderski A., Borucka A., Grzelak M. and Gil L., "Evaluation of Machinery Readiness Using Semi-Markov Processes", *Appl. Sci.* 2020, Vol.10, pp.1-15.
- [8] Gupta, S., P. C. Tewari, and A. K. Sharma. "Reliability and availability analysis of the ash handling unit of steam thermal power plant." *South African Journal of Industrial Engineering* vol. 20, no. 1, pp. 147-158, 2009.

- [9] Kumar, Ravinder. "Availability analysis of thermal power plant boiler air circulation system using Markov approach." *Decision Science Letters*, vol. 3, no. 1, 65-72, 2014.
- [10] Khanduja, Rajiv, P. C. Tewari, and Dinesh Kumar. "Availability analysis of bleaching system of paper plant," *Journal of Industrial Engineering*, Udyog Pragati, NITIE Mumbai (India) 32, no. 1, pp. 24-29, 2008.
- [11] Khanduja, Rajiv, and P. C. Tewari. "Development of performance evaluation system for screening unit of a paper plant," *International Journal of Applied Engineering Research*, vol. 3, no. 3, pp. 451-460, 2008.
- [12] Lai, C. D., Min Xie, Kim-Leng Poh, Yuan-Shun Dai, and P. Yang. "A model for availability analysis of distributed software/hardware systems," *Information and software technology*, vol. 44, no. 6, pp. 343-350, 2002.
- [13] P. Kumar, P.C. Tewari and D. Khanduja, "Six Sigma application in a process industry for capacity waste reduction: A case study" *Management Science Letters*, vol. 7, pp. 423-430, 2017.
- [14] P. Kumar, P.C. Tewari and D. Khanduja, "Maintenance Priorities for a Repairable System of a Thermal Power Plant Subject To Availability Constraint" *International Journal of Performability Engineering*, vol. 12, no.6, pp. 551-572, 2016.
- [15] Kumar, Pardeep, Ansar Ali, and Sandeep Kumar. "Coal-fired thermal power plant performance optimization using Markov and CFD analysis." *SN Applied Sciences* 2 (2020): 1-10.
- [16] Sabouhi, Hamed, Ali Abbaspour, Mahmud Fotuhi-Firuzabad, and Payman Dehghanian. "Reliability modeling and availability analysis of combined cycle power plants", *International Journal of Electrical Power & Energy Systems*, vol. 79, pp. 108-119, 2016.
- [17] Hassan, J., Thodi, P. and Khan, F., "Availability analysis of a LNG processing plant using the Markov process", *Journal of Quality in Maintenance Engineering*, Volume 22 , 2016, pp. 302-320.
- [18] Parkash, Shanti, and P. C. Tewari. "Performance modeling and dss for assembly line system of leaf spring manufacturing plant." *Reliability: Theory & Applications* 17, no. 2 (68) (2022): 403-412.
- [19] Parkash, Shanti, and P. C. Tewari. "Critical Review Of Rams Tools And Techniques For The Analysis Of Multi Component Complex Systems." *Reliability: Theory & Applications* 16.3 (63) (2021): 220-227.

WEIGHTED R-NORM ENTROPY FOR LIFETIME DISTRIBUTIONS: PROPERTIES AND APPLICATION

Bilal Ahmad Bhat^{1*} and M.A.K Baig²

Department of Statistics, University of Kashmir, Srinagar, J&K-190006, India
¹bhatbilal3819@gmail.com, ²baigmak@gmail.com

*Corresponding Author

Abstract

In the field of information theory, different uncertainty measures have been introduced by various researchers. These measures are widely used in reliability and survival studies. In this article, we introduce two new weighted uncertainty measures which are known as weighted R-Norm entropy (WRNE) and weighted R-Norm residual entropy (WRNRE). WRNE and WRNRE are "length-biased" shift-dependent uncertainty measures in which higher weight is assigned to large values of the observed random variable. Several important properties of these measures are studied. Some significant characterization results and the relationships of WRNRE with other reliability measures are presented. We also show that the survival function is uniquely determined by the WRNRE. Finally, based on a real life data set of bladder cancer patients, we illustrate the importance of WRNE and WRNRE.

Keywords: Weighted entropy, weighted R-Norm entropy, hazard rate function, mean residual life function and characterization results.

1. Introduction

A very important concept that has attracted the attention of researchers in the field of information theory is the measurement of uncertainty of probability distributions. The fundamental uncertainty measure (UM) which is well known by means of applications not only in the field of information theory but also in different other research fields is the Shannon's entropy [1]. Let Y be an absolutely continuous non-negative r.v with p.d.f $f(y)$, then the Shannon's entropy (SE) is defined as

$$H_Y(f) = - \int_0^{\infty} f(y) \log f(y) dy = -E[\log f(y)]. \quad (1)$$

Throughout this article, the notations r.v and p.d.f represent an absolutely continuous non-negative random variable and a probability density function respectively.

For a lifetime component that has survived up to an age t_0 , the SE is not a useful technique for measuring the uncertainty about its residual life. So, the concept of residual entropy was proposed by Ebrahimi [2] and is defined as

$$H_Y(f; t_0) = - \int_{t_0}^{\infty} \frac{f(y)}{\bar{F}(t_0)} \log \frac{f(y)}{\bar{F}(t_0)} dy, \quad (2)$$

where, $\bar{F}(t_0) = 1 - F(t_0)$ is the survival function (s.f) of the r.v Y .

The above UM's have been widely used in different research fields, but these UM's consider that a lifetime system or component serves the same in its whole life from the aspects of some given qualitative characteristic set by the experimenter. Due to this drawback, these UM's provide the same importance to the occurrence of every event of a probabilistic experiment and therefore these measures take the designation of shift-independent UM's. But in our real life, there exist several situations where the shift-dependent UM's are desirable. So, with the contribution of Belis and Guiasu [3], the first shift-dependent UM, simply known as weighted entropy was introduced and is defined as

$$\begin{aligned} H_Y^w(f) &= - \int_0^\infty w(y) f(y) \log f(y) dy \\ &= - \int_0^\infty y f(y) \log f(y) dy, \end{aligned} \quad (3)$$

where, the factor y in the integrand of (3) represents the weight which linearly emphasizes the occurrence of the event $\{Y = y\}$ and therefore yields a shift-dependent UM.

Similarly, Di Crescenzo and Longobardi [4] have extended the UM (3) to its dynamic (residual) version and therefore proposed the concept of weighted residual entropy as follows

$$H_Y^w(f; t_0) = - \int_{t_0}^\infty y \frac{f(y)}{\bar{F}(t_0)} \log \frac{f(y)}{\bar{F}(t_0)} dy. \quad (4)$$

It is clear from the available literature that the classical SE has been generalized in different ways by introducing some additional parameters to it. A well-known generalization that plays a very important role in different sciences is the concept of R-Norm entropy introduced by Boeke and Lube [5]. For more work and applications of R-Norm entropy, one can see Kumar and Choudhary [6] and Kumar et al. [7]. The continuous version of R-Norm entropy (RNE) was given by Nanda and Das [8] and is given by

$$H_{(Y,R)}(f) = \frac{R}{R-1} \left[1 - \left\{ \int_0^\infty f^R(y) dy \right\}^{\frac{1}{R}} \right], \quad R > 0 (\neq 1). \quad (5)$$

Similarly, analogous to (2), Nanda and Das [8] have extended the R-Norm entropy to its dynamic (residual) version, known as R-Norm residual entropy for the residual lifetime $Y - t_0 | Y > t_0$ and is defined as

$$H_{(Y,R)}(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ \int_{t_0}^\infty \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy \right\}^{\frac{1}{R}} \right], \quad R > 0 (\neq 1). \quad (6)$$

From the recent literature, it is seen that the measurement of uncertainty (entropy) of probability distributions is widely being used in the research work of various researchers with respect to different sciences. After the existence of fundamental UM's, the various researchers have introduced their weighted versions (i.e. weighted entropies) for measuring the uncertainty of such real life problems which are best fitted by weighted probability distributions. The researchers who have been attracted in the recent past by the concept of weighted entropy and therefore introduced some new flexible weighted UM's are: Bhat et al. [9], Bhat and Baig [10], Bhat et al. [11], Khammar and Jahanshahi [12], Kayal [13], Mirali and Baratpour [14], Nair et al. [15], Rajesh et al. [16], Nourbakhsh and Yari [17], Misagh et al. [18], Misagh and Yari [19] etc. Motivated with this research literature and the usefulness of R-Norm entropy and R-Norm residual entropy, here in this article, we introduce the concept of weighted R-Norm entropy and weighted R-Norm residual entropy. The article is continued as follows: In section 2, we consider the weighted R-Norm entropy (WRNE) in the form of its definition and several important properties. The section 3 studies the dynamic (residual) version of WRNE, known as weighted R-Norm residual entropy

(WRNRE) and also presents various significant characterization results of this UM. The various important properties of WRNRE and also its relationship with other well-known reliability measures are focused in section 4. The section 5 presents an application of the WRNE and WRNRE by using a real life data. Finally, in the last section, some concluding remarks are illustrated.

2. Weighted R-Norm Entropy (WRNE)

Analogous to (3), here in this section, we generate a new weighted UM which is actually the weighted version of R-Norm entropy (5) and is known as weighted R-Norm entropy (WRNE).

Definition 2.1 The WRNE for a r.v Y having p.d.f $f(y)$ denoted by $H_{(Y,R)}^W(f)$ is defined as

$$H_{(Y,R)}^W(f) = \frac{R}{R-1} \left[1 - \left(\int_0^\infty (yf(y))^R dy \right)^{\frac{1}{R}} \right], \quad R > 0 (\neq 1), \tag{7}$$

where, the factor y in the integrand is defined in (3).

The following example makes it clear that two different probability distributions can have the same RNE's, but unequal WRNE's.

Example 2.1. Let Y and Z be two r.v's with pdf's

$$f_Y(t) = \begin{cases} \frac{t}{2}, & 0 < t < 2 \\ 0, & \text{otherwise} \end{cases} \quad g_Z(t) = \begin{cases} 1 - \frac{t}{2}, & 0 < t < 2 \\ 0 & \text{otherwise} \end{cases}$$

By using (5), we obtain that

$$H_{(Y,R)}(f) = H_{(Z,R)}(g) = \frac{R}{R-1} \left[1 - \left(\frac{2}{R+1} \right)^{\frac{1}{R}} \right].$$

But, the WRNE's of Y and Z are not identical as follow

$$H_{(Y,R)}^W(f) = \frac{R}{R-1} \left[1 - \left(\frac{2^R}{2R+1} \right)^{\frac{1}{R}} \right] \text{ and } H_{(Z,R)}^W(g) = \frac{R}{R-1} \left[1 - \{2^{R+1}B(R+1, R+1)\}^{\frac{1}{R}} \right],$$

where,

$$B(\alpha, \beta) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \alpha, \beta > 0.$$

Hence, even though $H_{(Y,R)}(f) = H_{(Z,R)}(g)$, but $H_{(Y,R)}^W(f) \neq H_{(Z,R)}^W(g), \forall R > 0 (\neq 1)$.

Lemma 2.1. If $U = cY$, with $c > 0$, then

$$H_{(U,R)}^W(f) = \frac{R}{R-1} \left(1 - c^{\frac{1}{R}} \right) + c^{\frac{1}{R}} H_{(Y,R)}^W(f).$$

Example 2.2. Let $f(y)$ be the p.d.f of a r.v Y distributed as:

(a) Uniformly over $[m, n]$ with $f(y) = \frac{1}{n-m}, m < y < n$, then

$$H_{(Y,R)}(f) = \frac{R}{R-1} \left[1 - (n-m)^{\frac{1-R}{R}} \right] \text{ and } H_{(Y,R)}^W(f) = \frac{R}{R-1} \left[1 - \left(\frac{n^{R+1} - m^{R+1}}{(n-m)^R} \right)^{\frac{1}{R}} \right].$$

(b) Exponentially with $f(y) = \eta e^{-\eta y}, y > 0, \eta > 0$, then

$$H_{(Y,R)}(f) = \frac{R}{R-1} \left[1 - \left(\frac{\eta^{R-1}}{R} \right)^{\frac{1}{R}} \right] \text{ and } H_{(Y,R)}^w(f) = \frac{R}{R-1} \left[1 - \left(\frac{\Gamma(R+1)}{\eta R^{R+1}} \right)^{\frac{1}{R}} \right].$$

(c) Gamma with $f(y) = \frac{1}{\Gamma(\eta)} e^{-y} y^{\eta-1}$, $0 < y < \infty, \eta > 0$, then

$$H_{(Y,R)}(f) = \frac{R}{R-1} \left[1 - \frac{1}{(\Gamma(\eta))^R} \left(\frac{\Gamma((\eta-1)R+1)}{R(\eta-1)^{R+1}} \right)^{\frac{1}{R}} \right] \text{ and } H_{(Y,R)}^w(f) = \frac{R}{R-1} \left[1 - \frac{1}{(\Gamma(\eta))^R} \left(\frac{\Gamma(\eta R+1)}{R \eta^{R+1}} \right)^{\frac{1}{R}} \right].$$

(d) Weibull with $f(y) = \eta y^{\eta-1} e^{-y^\eta}$, $y > 0, \eta > 0$, then

$$H_{(Y,R)}(f) = \frac{R}{R-1} \left[1 - \left\{ \frac{\eta^{R-1} \Gamma\left(\left(1-\frac{1}{\eta}\right)(R-1)+1\right)}{R^{(R-1)\left(1-\frac{1}{\eta}\right)+1}} \right\}^{\frac{1}{R}} \right] \text{ and } H_{(Y,R)}^w(f) = \frac{R}{R-1} \left[1 - \left(\frac{\eta^{R-1} \Gamma\left(R+\frac{1}{\eta}\right)}{R^{R+\frac{1}{\eta}}} \right)^{\frac{1}{R}} \right].$$

(e) Rayleigh with $f(y) = \eta y e^{-\frac{\eta}{2}y^2}$, $y \geq 0, \eta > 0$, then

$$H_{(Y,R)}(f) = \frac{R}{R-1} \left[1 - \left(\frac{(2\eta)^{\frac{R-1}{2}} \Gamma\left(\frac{R+1}{2}\right)}{(\sqrt{R})^{R+1}} \right)^{\frac{1}{R}} \right] \text{ and } H_{(Y,R)}^w(f) = \frac{R}{R-1} \left[1 - \left(\frac{\Gamma\left(R+\frac{1}{2}\right) \left(\frac{2^{2R-1}}{\eta}\right)^{\frac{1}{2}}}{R^{R+\frac{1}{2}}} \right)^{\frac{1}{R}} \right].$$

Theorem 2.1. Let Y be a r.v having SE $H_Y(f)$, then

$$H_{(Y,R)}^w(f) \leq \frac{R}{R-1} \left[1 - \exp\left(\frac{(1-R)}{R} H_Y(f) + E(\log Y)\right) \right].$$

Proof. By applying the log-sum inequality, we have

$$\begin{aligned} \int_0^\infty f(y) \log \frac{f(y)}{(yf(y))^R} dy &\geq \int_0^\infty f(y) dy \log \frac{\int_0^\infty f(y) dy}{\int_0^\infty (yf(y))^R dy} \\ &= -\log \int_0^\infty (yf(y))^R dy. \end{aligned}$$

Due to (7) and after simple simplification, we obtain the desired result.

3. Weighted R-Norm Residual Entropy (WRNRE)

This section presents the weighted R-Norm entropy for residual lifetimes by utilizing the equation (7) which is the weighted version of (6). Some important characterization results of this UM are also discussed.

Definition 3.1 For a r.v Y having p.d.f $f(y)$ and s.f $\bar{F}(t_0)$, the WRNRE of order R at time $t_0 > 0$ is defined as

$$H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ \int_{t_0}^\infty \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy \right\}^{\frac{1}{R}} \right], \quad R > 0 (\neq 1). \tag{8}$$

Here, we study the expressions of WRNRE of some well-known lifetime distributions.

Example 3.1. If a r.v Y has the p.d.f $f(y)$ and s.f $\bar{F}(t_0)$ as:

(a) $f(y) = \frac{1}{d-c}$, $c < y < d$ and $\bar{F}(t_0) = \frac{d-t_0}{d-c}$, then

$$H_{(Y,R)}(y; t_0) = \frac{R}{R-1} \left[1 - (d - t_0)^{\frac{1-R}{R}} \right] \quad \text{and} \quad H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ \frac{d^{R+1} - t_0^{R+1}}{(d-t_0)^{R(R+1)}} \right\}^{\frac{1}{R}} \right].$$

(b) $f(y) = \eta e^{-\eta y}$, $y > 0, \eta > 0$ and $\bar{F}(t_0) = e^{-\eta t_0}$, then

$$H_{(Y,R)}(f; t_0) = \frac{R}{R-1} \left(1 - \frac{\eta^{R-1}}{R} \right) \quad \text{and} \quad H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - \left(\frac{\Gamma(R+1, \eta R t_0)}{\eta e^{-\eta R t_0} R^{R+1}} \right)^{\frac{1}{R}} \right].$$

(c) $f(y) = \frac{\eta^\mu}{\Gamma(\mu)} e^{-\eta y} y^{\mu-1}$, $0 < y < \infty, \eta, \mu > 0$ and $\bar{F}(t_0) = \frac{\Gamma(\mu, \eta t_0)}{\Gamma(\mu)}$, then

$$H_{(Y,R)}(f; t_0) = \frac{R}{R-1} \left[1 - \frac{\eta^\mu}{\Gamma(\mu, \eta t_0)} \left\{ \frac{\Gamma(R(\mu-1)+1, \eta R t_0)}{(\eta R)^{R(\mu-1)+1}} \right\}^{\frac{1}{R}} \right]$$

and

$$H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - \frac{\eta}{\Gamma(\mu, \eta t_0)} \left\{ \frac{\Gamma(R\mu+1, \eta R t_0)}{(\eta R)^{R\mu+1}} \right\}^{\frac{1}{R}} \right].$$

(d) $f(y) = \eta y^{\eta-1} e^{-y^\eta}$, $y > 0, \eta > 0$ and $\bar{F}(t_0) = e^{-t_0^\eta}$, then

$$H_{(Y,R)}(f; t_0) = \frac{R}{R-1} \left[1 - \frac{\eta^{\frac{R-1}{R}} e^{t_0^\eta}}{\frac{\eta(R+1)-R}{R} \eta^R} \left\{ \Gamma \left(\frac{R(\eta-1)+1}{\eta}, R t_0^\eta \right) \right\}^{\frac{1}{R}} \right]$$

and

$$H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ \frac{\eta^{R-1} \Gamma \left(R + \frac{1}{\eta}, R t_0^\eta \right)}{\frac{\eta^{R+1}}{R} e^{-R t_0^\eta}} \right\}^{\frac{1}{R}} \right].$$

(e) $f(y) = \eta y e^{-\frac{\eta}{2} y^2}$, $y \geq 0, \eta > 0$ and $\bar{F}(t_0) = e^{-\frac{\eta}{2} t_0^2}$, then

$$H_{(Y,R)}(f; t_0) = \frac{R}{R-1} \left[1 - e^{\frac{\eta}{2} t_0^2} \left\{ \left(\frac{(2\eta)^{R-1}}{R^{R+1}} \right)^{\frac{1}{2}} \Gamma \left(\frac{R+1}{2}, \frac{\eta R}{2} t_0^2 \right) \right\}^{\frac{1}{R}} \right]$$

and

$$H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - e^{\frac{\eta}{2} t_0^2} \left\{ \Gamma \left(R + \frac{1}{2}, \frac{\eta R}{2} t_0^2 \right) \left(\frac{2^{2R-1}}{\eta^{R^2 R+1}} \right)^{\frac{1}{2R}} \right\}^{\frac{1}{R}} \right].$$

where, $\Gamma(\beta, \alpha z) = \alpha^\beta \int_z^\infty e^{-\alpha u} u^{\beta-1} du$, $\alpha, \beta > 0$ is an upper incomplete gamma function.

Theorem 3.1 Let Y be a r.v having WRNRE and RNRE $H_{(Y,R)}^w(f; t_0)$ and $H_{(Y,R)}(f; t_0)$ respectively. Then for all $t_0 > 0$, we have

$$H_{(Y,R)}^w(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ t_0^R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}(f; t_0) \right)^R + \int_{z=t_0}^\infty z^{R-1} \left(\frac{\bar{F}(z)}{\bar{F}(t_0)} \right)^R (R - (R-1) H_{(Y,R)}(f; z)) dz \right\}^{\frac{1}{R}} \right].$$

Proof.

$$\begin{aligned} \int_{t_0}^{\infty} \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy &= \int_{t_0}^{\infty} \left(\int_0^y R z^{R-1} dz \right) \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy \\ &= R \int_{t_0}^{\infty} \left[\int_0^{t_0} z^{R-1} dz + \int_{t_0}^y z^{R-1} dz \right] \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy \\ &= t_0^R \int_{t_0}^{\infty} \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy + R \int_{z=t_0}^{\infty} z^{R-1} \left(\int_{y=z}^{\infty} \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy \right) dz. \end{aligned} \tag{9}$$

From (6), we have

$$\int_{t_0}^{\infty} \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy = \left(1 - \frac{(R-1)}{R} H_{(Y,R)}(f; t_0) \right)^R. \tag{10}$$

and

$$\int_{t_0}^{\infty} f^R(y) dy = \bar{F}^R(t_0) \left(1 - \frac{(R-1)}{R} H_{(Y,R)}(f; t_0) \right)^R. \tag{11}$$

Using (9), (10) and (11) in (8), the required result will be obtained.

The following theorem shows that $H_{(Y,R)}^W(f; t_0)$ determines the s.f $\bar{F}(t_0)$ uniquely.

Theorem 3.2. Let Y be a r.v having p.d.f $f(y)$, s.f $\bar{F}(t_0)$ and WRNRE $H_{(Y,R)}^W(f; t_0) < \infty, \forall R > 0 (\neq 1)$ respectively. If $H_{(Y,R)}^W(f; t_0)$ is increasing in t_0 , then $H_{(Y,R)}^W(f; t_0)$ uniquely determines the corresponding s.f $\bar{F}(t_0)$.

Proof. Rewriting (8) as

$$1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) = \left(\int_{t_0}^{\infty} \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy \right)^{\frac{1}{R}}. \tag{12}$$

Differentiating (12) both sides w.r.t t_0 , we have

$$\frac{(1-R)}{R} \frac{\partial}{\partial t_0} H_{(Y,R)}^W(f; t_0) = \frac{1}{R} \left(\int_{t_0}^{\infty} \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy \right)^{\frac{1-R}{R}} \left[R h_F(t_0) \int_{t_0}^{\infty} \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy - t_0^R h_F^R(t_0) \right], \tag{13}$$

where, $h_F(t_0) = \frac{f(t_0)}{\bar{F}(t_0)}$ represents the hazard rate of Y . Using (12), we can rewrite (13) as

$$t_0^R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right)^{1-R} h_F^R(t_0) - \{ R - (R-1) H_{(Y,R)}^W(f; t_0) \} h_F(t_0) - (R-1) \frac{\partial}{\partial t_0} H_{(Y,R)}^W(f; t_0) = 0. \tag{14}$$

For fixed $t_0 > 0$, $h_F(t_0)$ is a solution of $\psi(x) = 0$, where

$$\psi(x) = t_0^R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right)^{1-R} x^R - R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right) x - (R-1) \frac{\partial}{\partial t_0} H_{(Y,R)}^W(f; t_0) = 0.$$

Differentiating $\psi(x)$ w.r.t x , we have

$$\frac{\partial}{\partial x} \psi(x) = R t_0^R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right)^{1-R} x^{R-1} - R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right).$$

Also,

$$\frac{\partial^2}{\partial x^2} \psi(x) = R(R-1) t_0^R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right)^{1-R} x^{R-2}.$$

Now, $\frac{\partial}{\partial x} \psi(x) = 0$ gives

$$x = \left(\frac{R - (R-1)H_{(Y,R)}^W(f; t_0)}{Rt_0} \right)^{\frac{R}{R-1}} = x_0 \text{ (say).}$$

Case I. Let $R > 1$, then $\frac{\partial^2}{\partial x^2} \psi(x_0) > 0$. Thus, $\psi(x)$ attains minimum at x_0 . Also, $\psi(0) < 0$ and $\psi(\infty) = \infty$. Further, we can also observe it that $\psi(x)$ first decreases for $0 < x < x_0$ and then increases for $x > x_0$. So, $x = h_F(t_0)$ is the unique solution to $\psi(x) = 0$.

Case II. Let $R < 1$, then $\frac{\partial^2}{\partial x^2} \psi(x_0) < 0$. Thus, $\psi(x)$ attains maximum at x_0 . Also, $\psi(0) > 0$ and $\psi(\infty) = -\infty$. Further, we can easily see it that $\psi(x)$ first increases for $0 < x < x_0$ and then decreases for $x > x_0$. So, $x = h_F(t_0)$ is the unique solution to $\psi(x) = 0$. By combining both the cases, it is concluded that $H_{(Y,R)}^W(f; t_0)$ uniquely determines $h_F(t_0)$, which in turns determines $\bar{F}(t_0)$.

4. Properties and Inequalities of $H_{(Y,R)}^W(f; t_0)$

In this section, we study some interesting properties and inequalities of WRNRE.

Definition 4.1. A r.v Y is said to be smaller than in WRNRE of order R (denoted by $Y \stackrel{WRNRE}{\leq} Z$), if $H_{(Y,R)}^W(f; t_0) \leq H_{(Z,R)}^W(f; t_0), t_0 > 0$.

Definition 4.2. A r.v Y or a s.f \bar{F} has increasing (decreasing) R-Norm entropy for residual life IWRNERL (DWRNERL), if $H_{(Y,R)}^W(f; t_0)$ is increasing (decreasing) in $t_0, t_0 > 0$.

Lemma 4.1. If $Z = \lambda Y$, with $\lambda > 0$ is a constant, then

$$H_{(Z,R)}^W(f; t_0) = \frac{R}{R-1} \left(1 - \lambda^{\frac{1}{R}} \right) + \lambda^{\frac{1}{R}} H_{(Y,R)}^W \left(f; \frac{t_0}{\lambda} \right).$$

Proof.

$$H_{(Z,R)}^W(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ \int_{t_0}^{\infty} \left(z \frac{f(z)}{Pr(Z > t_0)} \right)^R dz \right\}^{\frac{1}{R}} \right],$$

where, $f(z)$ is the p.d.f of Z .

Setting $Z = \lambda Y$, we obtain

$$H_{(Z,R)}^W(f; t_0) = \frac{R}{R-1} \left[1 - \left\{ \int_{\frac{t_0}{\lambda}}^{\infty} \lambda \left(y \frac{f(y)}{\bar{F}(\frac{t_0}{\lambda})} \right)^R dy \right\}^{\frac{1}{R}} \right].$$

By using (8), we obtain the required result.

Theorem 4.1. For two r.v's Y and Z , let us define $X_1 = \alpha_1 Y$ and $X_2 = \alpha_2 Z$, with $\alpha_1, \alpha_2 > 0$. Let $Y \stackrel{WRNRE}{\leq} Z$ and $\alpha_1 \leq \alpha_2$. Then $X_1 \stackrel{WRNRE}{\leq} X_2$, if $H_{(Y,R)}^W(f; t_0)$ or $H_{(Z,R)}^W(f; t_0)$ is decreasing in $t_0 > 0$.

Poof. Suppose $H_{(Y,R)}^W(f; t_0)$ is decreasing in t_0 .

Now, $Y \stackrel{WRNRE}{\leq} Z$ implies

$$H_{(Y,R)}^W \left(f; \frac{t_0}{\alpha_2} \right) \leq H_{(Z,R)}^W \left(f; \frac{t_0}{\alpha_2} \right). \tag{15}$$

Further, since $\frac{t_0}{\alpha_1} \geq \frac{t_0}{\alpha_2}$, we have

$$H_{(Y,R)}^w \left(f; \frac{t_0}{\alpha_1} \right) \leq H_{(Y,R)}^w \left(f; \frac{t_0}{\alpha_2} \right). \tag{16}$$

From (15) and (16), we get

$$H_{(Y,R)}^w \left(f; \frac{t_0}{\alpha_1} \right) \leq H_{(Z,R)}^w \left(f; \frac{t_0}{\alpha_2} \right). \tag{17}$$

Using Lemma 4.1 in (17), we obtain $X_1 \stackrel{WRNRE}{\leq} X_2$.

Theorem 4.2. Let Y be a r.v with support $(0,m]$, $m > 0$, p.d.f $f(y)$ and s.f $\bar{F}(t_0)$, $t_0 > 0$, then for $R > 0 (\neq 1)$, the following inequality holds

$$H_{(Y,R)}^w(f; t_0) \geq \frac{R}{R-1} \left[1 - \exp \left\{ \frac{\int_{t_0}^m \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R \log \left(y \frac{f(y)}{\bar{F}(t_0)} \right) dy}{R \int_{t_0}^m \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy} + \log(m - t_0) \right\} \right].$$

Proof. Using log-sum inequality and (8), we have

$$\begin{aligned} \int_{t_0}^m \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R \log \left(y \frac{f(y)}{\bar{F}(t_0)} \right) dy &\geq \int_{t_0}^m \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy \log \frac{\int_{t_0}^m (y f(y))^R dy}{\int_{t_0}^m (\bar{F}(t_0))^R dy} \\ &= \int_{t_0}^m \left(y \frac{f(y)}{\bar{F}(t_0)} \right)^R dy \left[\log \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right\}^R - \log(m - t_0) \right]. \end{aligned}$$

After simple calculations, we can easily obtain the required result.

Theorem 4.3. If Y is IWRNERL (DWRNERL) and $R > 0 (\neq 1)$, then

$$h_F(t_0) \leq (\geq) \left[\frac{R}{t_0^R} \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right\}^R \right]^{\frac{1}{R-1}}.$$

Proof. From (14), we have

$$(R - 1) \frac{\partial}{\partial t_0} H_{(Y,R)}^w(f; t_0) = t_0^R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right)^{1-R} h_F^R(t_0) - R \left(1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right) h_F(t_0).$$

Since Y is IWRNERL (DWRNERL), therefore

$$h_F(t_0) \left[t_0^R \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right\}^{1-R} h_F^{R-1}(t_0) - R \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right\} \right] \geq (\leq) 0.$$

which leads to

$$h_F(t_0) \leq (\geq) \left[\frac{R}{t_0^R} \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^w(f; t_0) \right\}^R \right]^{\frac{1}{R-1}}.$$

Theorem 4.4. If \bar{F} is IWRNERL (DWRNERL), then

$$H_{(Y,R)}^w(f; t_0) \geq (\leq) \frac{R}{R-1} \left[1 - t_0 \left(\frac{\left(1 + \frac{\partial}{\partial t_0} \delta_F(t_0) \right)^{R-1}}{\delta_F(t_0)} \right)^{\frac{1}{R}} \right],$$

where $\delta_F(t_0)$ is the mean residual life function of Y .

Proof. From (14), we have

$$\frac{\partial}{\partial t_0} H_{(Y,R)}^W(f; t_0) = \frac{1}{R-1} \left[t_0^R \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right\}^{1-R} h_F^R(t_0) - R \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right\} h_F^R(t_0) \right].$$

Since, \bar{F} is IWRNERL and $R > 0 (\neq 1)$, therefore, we have

$$t_0^R \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right\}^{1-R} h_F^R(t_0) - R \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right\} h_F^R(t_0).$$

which gives

$$H_{(Y,R)}^W(f; t_0) \geq \frac{R}{(R-1)} \left[1 - t_0 \left(\frac{h_F^{R-1}(t_0)}{R} \right)^{\frac{1}{R}} \right].$$

Using $h_F(t_0) = \frac{1 + \frac{\partial}{\partial t_0} \delta_F(t_0)}{\delta_F(t_0)}$, we get

$$H_{(Y,R)}^W(f; t_0) \geq \frac{R}{R-1} \left[1 - t_0 \left\{ \frac{1}{R} \left(\frac{1 + \frac{\partial}{\partial t_0} \delta_F(t_0)}{\frac{\partial}{\partial t_0} \delta_F(t_0)} \right)^{R-1} \right\}^{\frac{1}{R}} \right].$$

The proof of DWRNERL is similar.

Theorem 4.5. Let Y be the lifetime of a system with p.d.f $f(y)$ and s.f $\bar{F}(t_0)$, $t_0 > 0$, then for $R > 0 (\neq 1)$, we have

$$H_{(Y,R)}^W(f; t_0) \leq \frac{R}{R-1} \left[1 - \exp \left\{ R \int_{t_0}^{\infty} \frac{f(y)}{\bar{F}(t_0)} \log y dy + (1 - R) H_Y(f; t_0) \right\} \right]. \tag{18}$$

Proof. From log-sum inequality, we have

$$\begin{aligned} \int_{t_0}^{\infty} f(y) \log \frac{f(y)}{\left(\frac{f(y)}{\bar{F}(t_0)} \right)^R} dy &\geq \int_{t_0}^{\infty} f(y) dy \log \frac{\int_{t_0}^{\infty} f(y) dy}{\int_{t_0}^{\infty} \left(\frac{f(y)}{\bar{F}(t_0)} \right)^R dy} \\ &= \bar{F}(t_0) \left[\log \bar{F}(t_0) - R \log \left\{ 1 - \frac{(R-1)}{R} H_{(Y,R)}^W(f; t_0) \right\} \right]. \end{aligned} \tag{19}$$

where (19) is obtained from (8).

The L.H.S of (19) leads to

$$(1 - R) \int_{t_0}^{\infty} f(y) \log f(y) dy - R \int_{t_0}^{\infty} f(y) \log y dy + R \bar{F}(t_0) \log \bar{F}(t_0). \tag{20}$$

Using (20) in (19), we obtain (18).

5. Application

In this section, we demonstrate a real life data set to analyze the performance of WRNE and WRNRE in practice. The data set represents the remission times (in months) of a random sample of 128 bladder cancer patients given in Lee and Wang [20] and is given as follows:

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28,

9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

According to Afaq et al. [21] this data set is best fitted by length biased Lomax distribution (LBLD). So, for computing the uncertainty of this data set, the simple entropy techniques are not appropriate. Therefore, it is necessary to apply the weighted entropy techniques rather than the simple entropy. For the weighted entropy, here we must consider the parameters of the Lomax distribution (LD) not of the LBLD. The MLE's of the parameters of LD having p.d.f $f(y) = \frac{\alpha}{\beta} \left(1 + \frac{y}{\beta}\right)^{-(\alpha+1)}$, $y > 0, \alpha, \beta > 0$ from this data set are obtained as: $\alpha = 8.43$ (shape parameter) and $\beta = 70.29$ (scale parameter) respectively. Now, for $\alpha = 8.43, \beta = 70.29, R = 2$ and $t_0 = 5$, the values of WRNE and WRNRE are obtained as: $H_{(Y,R)}^W(f) = 1.028$ and $H_{(Y,R)}^W(f; t_0) = 0.585$. Similarly, for the same values of α and β , if we take $R = 4$ and $t_0 = 10$, we can obtain $H_{(Y,R)}^W(f) = 1.111$ and $H_{(Y,R)}^W(f; t_0) = 1.015$ respectively.

6. Conclusion

In this article, we considered weighted R-Norm entropy of order R and also its dynamic (residual) version. These are shift-dependent uncertainty measures which assign the higher weight to the larger values of the observed random variable. We have also studied the various significant properties of these measures. Some of the important relationships of the proposed dynamic measure with hazard rate and mean residual life functions have been discussed. Finally, we have illustrated the importance of the proposed measures with the help of a real life data set.

References

- [1] Shannon, C. E. (1948). A mathematical theory of communications, *Bell System Technical Journal*, 27, 379-423.
- [2] Ebrahimi, N. (1996). How to measure uncertainty in the residual lifetime distribution. *Sankhya Series A*, 58, 48-56.
- [3] Belis, M. and Guiasu, S. (1968). A quantitative-qualitative measure of information in cybernetic systems. *IEEE Transactions on Information Theory*, IT., 4, 593-594.
- [4] Di, Crescenzo, A. and Longobardi, M. (2006). On weighted residual and past entropies. *Scientiae Mathematicae Japonicae*, 64, 255-266.
- [5] Boekee, D. E. and Van der Lubbe, J. C. (1980). The R-norm information measure. *Information and control*, 45(2), 136-155.
- [6] Kumar, S. and Choudhary, A. (2011). R-Norm Shannon-Gibbs Type Inequality. *Journal of Applied Sciences*, 11(15), 2866-2869.
- [7] Kumar, S., Ram, G. and Gupta, V. (2012). On 'useful' R-norm relative information and J-divergence measures. *International journal of pure and applied mathematics*, 77(3), 349-358.
- [8] Nanda, A. K. and Das, S. (2006). Study on R Norm Residual Entropy. *Calcutta Statistical Association Bulletin*, 58(3-4), 197-210.
- [9] Bhat, B.A., Sultan, M. and Baig, M. A. K. (2023). Weighted generalized entropy: properties and application. *Reliability: Theory and Applications*, 3(74), 113-120.
- [10] Bhat, B.A. and Baig, M. A. K. (2019). A New Two Parametric Weighted Generalized Entropy for Lifetime Distributions. *Journal of Modern Applied Statistical Methods*, 18(2).

- [11] Bhat, B. A., Mudasir, S., and Baig, M. A. K. (2019). Some Characterization Results on Length-Biased Generalized Interval Entropy for Lifetime Distributions. *Pakistan Journal of Statistics*, 35(2), 155-170.
- [12] Khammar, A., and Jahanshahi, S. (2018). On weighted cumulative residual Tsallis entropy and its dynamic version. *Physica A: Statistical Mechanics and Its Applications*, 491, 678–692.
- [13] Kayal, S. (2017). On weighted generalized cumulative residual entropy. *Springer Science+Business Media New York*, 1-17.
- [14] Mirali, M. and Baratpour, S. (2017). Dynamic version of weighted cumulative residual entropy. *Communications in Statistics-Theory and Methods*, 46(22), 11047-11059.
- [15] Nair, R. S., Sathar, E. I. A. and Rajesh, G. (2017). A study on dynamic weighted failure entropy of order α . *American Journal of Mathematical and Management Sciences*, 36(2), 137-149.
- [16] Rajesh, G., Abdul-Sathar, E., and Rohini, S. N. (2017). On dynamic weighted survival entropy of order α . *Communications in Statistics-Theory and Methods*, 46(5), 2139–2150.
- [17] Nourbakhsh, M. and Yari, G. (2016). Weighted Renyi's entropy for lifetime distributions. *Communications in Statistics-Theory and Methods*, doi: 10.1080 /03610926.2016.1148729.
- [18] Misagh, F., Panahi, Y., Yari, G. H. and Shahi, R. (2011). Weighted cumulative entropy and its estimation. In *Quality and Reliability (ICQR)*, IEEE International conference, 477-480.
- [19] Misagh, F. and Yari, G. H. (2011). On weighted interval entropy. *Statistics and Probability Letters*, 81, 188–194.
- [20] Lee, E. T., and Wang, J. (2003). *Statistical methods for survival data analysis*. John Wiley and Sons, Vol. 476.
- [21] Ahmad, A., Ahmad, S. P. and Ahmed, A. (2016). Length-biased Weighted Lomax Distribution: Statistical Properties and Application. *Pakistan Journal of Statistics and Operation Research*, 245-255.

AN IMPROVED ESTIMATOR OF FINITE POPULATION MEAN UNDER RANKED SET SAMPLING

FRANCIS DELALI BAETA ^{1,2*}, DIOGGBAN JAKPERIK ², MICHAEL JACKSON ADJABUI ²

•
Ho Technical University, Ho, Ghana ¹

C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana ²

baetadelali@gmail.com

jdioggban@cktutas.edu.gh

madjabui@cktutas.edu.gh

Abstract

To obtain reliable estimates of population parameters, data that is sampled for estimation must accurately represent the underlying population. Sampled data that is representative of the underlying population depends also on the sampling technique that was used in obtaining them. This is very important since sampling bias could lead to over or under estimation of parameters. Ranked Set Sampling is considered to be a better alternative to the classical sampling designs in obtaining such data. Ranked Set Sampling is designed to minimize the number of measured observations required to achieve a desired precision in making inferences, and thus it is more economical to use for the purposes of estimation, compared to the classical sampling designs. This is also an added advantage in cases where it is difficult to obtain data. Many estimators have been developed recently for the estimation of finite population mean under ranked set sampling. This paper aims to improve estimation by modifying an existing estimator using a simple linear combination of the known population mean, square root of the known coefficient of variation, and the known median of an auxiliary variable. The theoretical properties of the proposed estimator, such as the bias and mean squared error were derived up to the first order of approximation, using Taylor's expansion. The bias, mean squared error, absolute relative bias, and the relative efficiency were used as means of evaluation and comparison between the proposed modified estimator and its competitors. The R software was used to aid computations. Empirical applications to real data showed that the proposed modified estimator is superior to the competing estimators that were compared since it has least bias, the least mean squared error, the least absolute relative bias, and the highest relative efficiency in all sample sizes that were considered. The bias and mean squared error of the modified estimator under Ranked Set Sampling was found to be smaller than those of the existing estimators that were compared. Hence it is more efficient and capable of providing reliable estimates than the existing estimators that were compared and so we recommend that it should be used in survey estimations.

Keywords: Ranked Set Sampling, Ratio Estimator, Bias, Mean Squared Error, Auxiliary Variable

1. INTRODUCTION

Over the years, researchers have been preoccupied with the development of new estimators for finite heterogeneous population mean with the aim of reducing the associated bias and MSE of existing estimators to the barest minimum [1, 5, 9, 16]. For reliable estimates, data that is employed for estimation must be representative of the underlying population. Sampled data that is representative of the underlying population depends also on the sampling method [18]. Sampling bias could lead to over or under estimation of population parameters. Consequently,

one field of interest currently has been in the area of identifying designs that generate representative samples for super populations. Estimation of finite population mean has been based disproportionately on the classical sampling designs, especially simple random sampling (SRS). However, the SRS procedure as noted by [3] is incapable of generating representative samples for certain populations. The consequence, as noted by [10] is that, a specific sample which is not truly representative of the underlying population can possibly be included for estimation, and that can lead to unreliable estimates. Therefore, to improve accuracy and precision in the estimation of finite population mean, sampling procedures which do not suffer such weaknesses as the SRS must be considered. Among the sampling methods, the Ranked Set Sampling (RSS) technique is a good alternative to SRS for obtaining data that are truly representative of the population under study [2]. The goal of RSS is to collect observations that are more likely to span the full range of values in the population and therefore produces more representative samples than SRS [10]. RSS was first introduced by [12] and was used to estimate pasture yield. RSS was introduced for circumstances where difficulty exist in taking actual measurements for sample units. [17] established the statistical methodology for RSS.

The procedure for obtaining a ranked set sample is briefly outlined by the following steps:

1. Randomly select a sample of size m^2 from the targeted population.
2. Distribute the m^2 selected units in m sets, each of size m .
3. Rank the units within each set with respect to the attribute of interest, using the judgement of an expert or by the aid of an auxiliary variable that is correlated with the study variable.
4. Select the i^{th} ranked unit from the i^{th} set for actual measurement of the attribute of interest, in the order $i = 1, 2, 3, \dots, m$.
5. Repeat steps i to iv for r cycles if it is desired to obtain a sample of size, $n = mr$.

RSS is preferred when mechanisms are readily available for ranking a set of sample units, whether by the use of an auxiliary variable, or by the use of the judgement of an expert. [7] proved that the ranked set sample mean is an unbiased estimator for population mean, even in cases of imperfect ranking. [10] adduced that an auxiliary variable, X could be used to rank any variable under study, Y in cases where judgement ranking of Y is difficult. Consequently, a lot of estimators have been developed under RSS, employing a variety of auxiliary variables for ranking.

[16] introduced the classical ratio estimator under RSS. Several other authors have since extended the work of [16], employing a variety of auxiliary variables. [11] suggested a modified ratio estimator for population mean under RSS utilizing the quartile deviations and the known mean of an auxiliary variable. [4] proposed a generalized ratio estimator for population mean under RSS using the known population mean of an auxiliary variable and some pre-assigned constants. [13] suggested a modified ratio-cum-product estimator for finite population mean under RSS using the known population information on the mean, the coefficients of variation and of kurtosis of an auxiliary variable under RSS. [15] proposed a ratio-type estimator for population mean under RSS, using the known population mean and quartiles of an auxiliary variable. [8] proposed a ratio-type estimator under RSS based on known population mean and population deciles of an auxiliary variable. [9] proposed a generalised ratio-type estimator based on RSS, employing known parameters of the population such as the coefficients of variation, kurtosis and skewness as well as the mean of the auxiliary variable. [14] suggested a ratio type estimator of population mean based on RSS employing the known coefficient of variation, known median, as well as the known population mean of the auxiliary variable. These estimators were more efficient and superior to their competitors. Notwithstanding, the existing estimators wield significant biases and are fraught with large mean squared errors. Therefore, this study sought to improve estimation by modifying an existing estimator of finite population mean that was based on RSS.

2. REVIEW OF EXISTING ESTIMATORS

Suppose the study variable Y and the auxiliary variable, X are positively correlated. Then [16] expressed the classical ratio estimator of population mean under RSS as

$$\bar{y}_{R,RSS} = \bar{y}_{[n]} \left(\frac{\bar{X}}{\bar{x}_{(n)}} \right) \quad (2.1)$$

where

$$Bias(\bar{y}_{R,RSS}) = \bar{Y} \left[\theta \left(C_x^2 - \rho C_x C_y \right) - \left(W_{x(i)}^2 - W_{yx(i)} \right) \right] \quad (2.2)$$

$$MSE(\bar{y}_{R,RSS}) = \bar{Y}^2 \left[\theta \left(C_x^2 - 2\rho C_x C_y + C_y^2 \right) - \left(W_{x(i)}^2 - 2W_{yx(i)} + W_{y[i]}^2 \right) \right] \quad (2.3)$$

Through out of this study,

$$W_{x(i)}^2 = \frac{1}{r} \sum_{i=1}^m \left(\frac{\mu_{x(i)} - \bar{X}}{m\bar{X}} \right)^2, W_{y[i]}^2 = \frac{1}{r} \sum_{i=1}^m \left(\frac{\mu_{y[i]} - \bar{Y}}{m\bar{Y}} \right)^2, W_{yx(i)} = \sum_{i=1}^m \frac{(\mu_{y[i]} - \bar{Y})(\mu_{x(i)} - \bar{X})}{m^2 r \bar{Y} \bar{X}}, \theta = \frac{1}{m r}, C_x^2 = \frac{s_x^2}{\bar{X}^2}$$

and $C_y^2 = \frac{s_y^2}{\bar{Y}^2}$.

[13] modified the classical ratio estimator of population mean under RSS respectively using C_x and $\beta_2(x)$ as

$$\bar{y}_{M1,RSS} = \bar{y}_{[n]} \left[\frac{\bar{X} + C_x}{\bar{x}_{(n)} + C_x} \right] \quad (2.4)$$

and

$$\bar{y}_{M2,RSS} = \bar{y}_{[n]} \left[\frac{\bar{X} C_x + \beta_2(x)}{\bar{x}_{(n)} C_x + \beta_2(x)} \right]. \quad (2.5)$$

with the respective biases

$$B(\bar{y}_{M1,RSS}) = \bar{Y} \left[\theta \left(\varphi^2 C_x^2 - \varphi \rho C_x C_y \right) - \left(\varphi^2 W_{x(i)}^2 - \varphi W_{yx(i)} \right) \right] \quad (2.6)$$

$$B(\bar{y}_{M2,RSS}) = \bar{Y} \left[\theta \left(v^2 C_x^2 - v \rho C_x C_y \right) - \left(v^2 W_{x(i)}^2 - v W_{yx(i)} \right) \right] \quad (2.7)$$

and the respective MSEs

$$MSE(\bar{y}_{M1,RSS}) = \bar{Y}^2 \left[\theta \left(\varphi^2 C_x^2 + C_y^2 - 2\varphi \rho C_x C_y \right) - \left(\varphi^2 W_{x(i)}^2 + W_{y[i]}^2 - 2\varphi W_{yx(i)} \right) \right] \quad (2.8)$$

$$MSE(\bar{y}_{M2,RSS}) = \bar{Y}^2 \left[\theta \left(v^2 C_x^2 + C_y^2 - 2v \rho C_x C_y \right) - \left(v^2 W_{x(i)}^2 + W_{y[i]}^2 - 2v W_{yx(i)} \right) \right] \quad (2.9)$$

where

$$\varphi = \frac{\bar{X}}{\bar{X} + C_x}$$

$$v = \frac{\bar{X} C_x}{\bar{X} C_x + \beta_2(x)}$$

[14] used the coefficient of variation (C_x) and the median of an auxiliary variable (M_d) to propose a ratio-type estimator of population mean as

$$\bar{y}_{P,RSS} = \bar{y}_{[n]} \left[\frac{\bar{X} + C_x M_d}{\bar{x}_{(n)} + C_x M_d} \right]. \quad (2.10)$$

with the respective bias and MSE as

$$B(\bar{y}_{P,RSS}) = \bar{Y} \left[\theta \left(\lambda^2 C_x^2 - \lambda \rho C_x C_y \right) - \left(\lambda^2 W_{x(i)}^2 - \lambda W_{yx(i)} \right) \right] \quad (2.11)$$

and

$$MSE(\bar{y}_{P,RSS}) = \bar{Y}^2 \left[\theta \left(\lambda^2 C_x^2 + C_y^2 - 2\lambda \rho C_x C_y \right) - \left(\lambda^2 W_{x(i)}^2 + W_{y[i]}^2 + 2\lambda W_{yx(i)} \right) \right] \quad (2.12)$$

where

$$\lambda = \frac{\bar{X}}{\bar{X} + C_x M_d}$$

3. THE PROPOSED ESTIMATOR

Motivated by [14], this study proposes modified ratio-type estimator for finite population mean under Ranked Set Sampling, which utilizes the coefficient of variation (C_x) and the median (M_d) of the employed auxiliary variable as

$$\bar{y}_{B,RSS} = \bar{y}_{[n]} \left[\frac{\bar{X} + M_d \sqrt{C_x}}{\bar{x}_{(n)} + M_d \sqrt{C_x}} \right]. \quad (3.1)$$

Using large sample properties, the following assumptions are made: $\bar{y}_{[n]} = \bar{Y}(1 + e_0)$ and $\bar{x}_{(n)} = \bar{X}(1 + e_1)$, where $E(e_0) = E(e_1) = 0$.

Therefore, equation (3.1) evolves as

$$\begin{aligned} \bar{y}_{B,RSS} &= \bar{y}_{[n]} \left[\frac{\bar{X} + M_d \sqrt{C_x}}{\bar{X}(1 + e_1) + M_d \sqrt{C_x}} \right] \\ &= \bar{y}_{[n]} \left[\frac{\bar{X} + M_d \sqrt{C_x}}{\bar{X} + M_d \sqrt{C_x} + \bar{X}e_1} \right] \\ &= \bar{y}_{[n]} \left[\frac{1}{1 + \left(\frac{\bar{X}}{\bar{X} + M_d \sqrt{C_x}} \right) e_1} \right] \\ &= \bar{y}_{[n]} \left(\frac{1}{1 + \omega e_1} \right) \end{aligned}$$

where

$$\omega = \frac{\bar{X}}{\bar{X} + M_d \sqrt{C_x}}.$$

Now,

$$\begin{aligned} \bar{y}_{B,RSS} &= \bar{y}_{[n]} \left(\frac{1}{1 + \omega e_1} \right) \\ &= \bar{Y}(1 + e_0)(1 + \omega e_1)^{-1}. \end{aligned}$$

Assuming $|\omega e_1| < 1$ and using Taylor's expansion to the second order,

$$\begin{aligned} \bar{y}_{B,RSS} &= \bar{Y}(1 + e_0) \left(1 - \omega e_1 + \omega^2 e_1^2 + \dots \right) \\ &= \bar{Y} \left(1 - \omega e_1 + \omega^2 e_1^2 + e_0 - \omega e_0 e_1 + \dots \right). \end{aligned}$$

Therefore the bias of the proposed estimator is obtained as

$$B(\bar{y}_{B,RSS}) = \bar{Y} \left[\omega^2 E(e_1^2) - \omega E(e_0 e_1) \right]$$

$$\begin{aligned} \Rightarrow B(\bar{y}_{B,RSS}) &= \bar{Y} \left[\omega^2 \left(\theta C_x^2 - W_{x(i)}^2 \right) - \omega \left(\theta \rho C_x C_y - W_{yx(i)} \right) \right] \\ &= \bar{Y} \left[\theta \left(\omega^2 C_x^2 - \omega \rho C_x C_y \right) - \left(\omega^2 W_{x(i)}^2 - \omega W_{yx(i)} \right) \right] \end{aligned} \quad (3.2)$$

The mean squared error of the proposed estimator is obtained as

$$\begin{aligned} MSE(\bar{y}_{B,RSS}) &= \bar{Y}^2 \left[\omega^2 E(e_1^2) - 2\omega E(e_0 e_1) + E(e_0^2) \right] \\ \Rightarrow MSE(\bar{y}_{B,RSS}) &= \bar{Y}^2 \left[\omega^2 \left(\theta C_x^2 - W_{x(i)}^2 \right) + \left(\theta C_y^2 - W_{y[i]}^2 \right) - 2\omega \left(\theta \rho C_x C_y - W_{yx(i)} \right) \right] \\ &= \bar{Y}^2 \left[\theta \left(\omega^2 C_x^2 + C_y^2 - 2\omega \rho C_x C_y \right) - \left(\omega^2 W_{x(i)}^2 + W_{y[i]}^2 - 2\omega W_{yx(i)} \right) \right] \end{aligned} \quad (3.3)$$

4. EFFICIENCY COMPARISON

The proposed estimator, $\bar{y}_{B,RSS}$ was compared to the RSS estimators of [13] and that of [14]. The proposed estimator, $\bar{y}_{B,RSS}$ is more efficient than the estimator of [14] if

$$\begin{aligned} MSE(\bar{y}_{B,RSS}) &< MSE(\bar{y}_{P,RSS}) \\ \Rightarrow \omega^2 \left(\theta C_x^2 - W_{x(i)}^2 \right) - 2\omega \left(\theta \rho C_x C_y - W_{yx(i)} \right) &< \lambda^2 \left(\theta C_x^2 - W_{x(i)}^2 \right) - 2\lambda \left(\theta \rho C_x C_y - W_{yx(i)} \right) \\ \Rightarrow \left(\omega^2 - \lambda^2 \right) \left(\theta C_x^2 - W_{x(i)}^2 \right) &< 2(\omega - \lambda) \left(\theta \rho C_x C_y - W_{yx(i)} \right) \end{aligned}$$

Hence, provided $\omega < \lambda$, the proposed estimator is more efficient than the estimator of [14] if

$$\rho < \frac{(\omega + \lambda) \left(\theta C_x^2 - W_{x(i)}^2 \right) + 2W_{yx(i)}}{2\theta C_x C_y} \quad (4.1)$$

where ρ is the correlation coefficient between the auxiliary variable X and the study variable Y . Let $P = \theta C_x^2 - W_{x(i)}^2$, $Q = 2W_{yx(i)}$ and $R = 2\theta C_x C_y$. Then the proposed estimator, $\bar{y}_{B,RSS}$ is respectively more efficient than $\bar{y}_{M1,RSS}$ and $\bar{y}_{M2,RSS}$ if

$$\rho < [P(\omega + \varphi) + Q] / R \text{ and } \rho < [P(\omega + v) + Q] / R,$$

where

$$\omega = \frac{\bar{X}}{\bar{X} + M_d \sqrt{C_x}}, \quad \varphi = \frac{\bar{X}}{\bar{X} + C_x}, \quad v = \frac{\bar{X} C_x}{\bar{X} C_x + \beta_2(x)}.$$

5. EMPIRICAL APPLICATION

The dataset that was used for evaluating the estimators is taken from page 34 of [6] and a general description is given below.

X : Weekly family income.

Y : Weekly family expenditure.

Objective: To estimate mean weekly family expenditure.

$N = 33$, $\bar{X} = 72.5454$, $\bar{Y} = 27.4909$, $\rho = 0.2521$, $M_d = 69$, $\beta_2(x) = 2.1429$, $C_x = 0.1436$, $C_y = 0.3629$

The ARB of the various proposed estimators were obtained by the formula

$$ARB = \left| \frac{Bias(\bar{y}_i)}{Bias(\bar{y}_{R,RSS})} \right|, \quad (5.1)$$

where $i = (M1,RSS), (M2,RSS), (P,RSS), (B,RSS)$.

The Percent Relative Efficiency (PRE) of an estimator \bar{y}_i compared to the classical ratio estimator \bar{y}_R of [5], was obtained by

$$PRE = \frac{MSE(\bar{y}_R)}{MSE(\bar{y}_i)} \times 100, \tag{5.2}$$

where $i = (M1,RSS), (M2,RSS), (P,RSS), (B,RSS)$.

Six ranked set sample sizes were considered with the data for different set sizes m and the corresponding number of cycles r and the results displayed in Tables 1 to 6. For each case, corresponding values of $W_{x(i)}^2, W_{y[i]}^2$ and $W_{yx(i)}$ were determined for the sample size $n = m \times r$.

For sample size $n = 9$, where $m = 3$ and $r = 3$, then $W_{x(i)}^2 = 0.0012, W_{y[i]}^2 = 0.0064, W_{yx(i)} = 0.0028$ and the corresponding performance of the various estimators is displayed in Table 1.

Table 1: $m=3, r=3$

Estimator	Bias	MSE	ARB	PRE
$\bar{y}_{R, RSS}$	0.0668	9.0725	1.0000	100.0
$\bar{y}_{M1, RSS}$	0.0667	9.0653	0.9985	100.1
$\bar{y}_{M2, RSS}$	0.0562	8.6695	0.8413	104.6
$\bar{y}_{P, RSS}$	0.0556	8.6428	0.8323	105.0
$\bar{y}_{B, RSS}$	0.0433	8.1567	0.6482	111.2

If $n = 12$ where $m = 3, r = 4$, then $W_{x(i)}^2 = 0.0009, W_{y[i]}^2 = 0.0050, W_{yx(i)} = 0.0021$ and the performance of the various estimators is displayed in Table 2.

Table 2: $m=3, r=4$

Estimator	Bias	MSE	ARB	PRE
$\bar{y}_{R, RSS}$	0.0501	6.6533	1.0000	100.0
$\bar{y}_{M1, RSS}$	0.0450	6.6478	0.8982	100.1
$\bar{y}_{M2, RSS}$	0.0418	6.3319	0.8343	105.1
$\bar{y}_{P, RSS}$	0.0417	6.3310	0.8323	105.1
$\bar{y}_{B, RSS}$	0.0325	5.9664	0.6487	111.5

If $n = 15$ where $m = 3, r = 5$, then $W_{x(i)}^2 = 0.0007, W_{y[i]}^2 = 0.0040, W_{yx(i)} = 0.0017$ and the performance of the various estimators is displayed in Table 3.

Table 3: $m=3, r=5$

Estimator	Bias	MSE	ARB	PRE
$\bar{y}_{R, RSS}$	0.0412	5.3680	1.0000	100.0
$\bar{y}_{M1, RSS}$	0.0411	5.3635	0.9976	100.1
$\bar{y}_{M2, RSS}$	0.0346	5.1103	0.8398	105.0
$\bar{y}_{P, RSS}$	0.0343	5.1031	0.8325	105.2
$\bar{y}_{B, RSS}$	0.0267	4.8055	0.6481	111.7

If $n = 16$ where $m = 4, r = 4$, then $W_{x(i)}^2 = 0.0008, W_{y[i]}^2 = 0.0044, W_{yx(i)} = 0.0019$ and the performance of the various estimators is displayed in Table 4.

Table 4: $m=4, r=4$

Estimator	Bias	MSE	ARB	PRE
$\bar{y}_{R, RSS}$	0.0431	4.8955	1.0000	100.0
$\bar{y}_{M1, RSS}$	0.0429	4.8908	0.9954	100.1
$\bar{y}_{M2, RSS}$	0.0367	4.5020	0.8515	108.7
$\bar{y}_{P, RSS}$	0.0364	4.5020	0.8445	108.7
$\bar{y}_{B, RSS}$	0.0291	4.2936	0.6752	114.1

For $n = 20$ where $m = 4, r = 5$, then $W_{x(i)}^2 = 0.0006, W_{y[i]}^2 = 0.0036, W_{yx(i)} = 0.0015$ and the performance of the various estimators is displayed in Table 5.

Table 5: $m=4, r=5$

Estimator	Bias	MSE	ARB	PRE
$\bar{y}_{R, RSS}$	0.0350	3.8559	1.0000	100.0
$\bar{y}_{M1, RSS}$	0.0349	3.8521	0.9971	100.1
$\bar{y}_{M2, RSS}$	0.0310	3.6368	0.8857	106.0
$\bar{y}_{P, RSS}$	0.0296	3.6292	0.8457	106.3
$\bar{y}_{B, RSS}$	0.0234	3.3685	0.6686	114.5

If $n = 25$ where $m = 5, r = 5$, then $W_{x(i)}^2 = 0.0006, W_{y[i]}^2 = 0.0033, W_{yx(i)} = 0.0013$ and the performance of the various estimators is displayed in Table 6.

Table 6: $m=5, r=5$

Estimator	Bias	MSE	ARB	PRE
$\bar{y}_{R, RSS}$	0.0275	2.8278	1.0000	100.0
$\bar{y}_{M1, RSS}$	0.0274	2.8248	0.9963	100.1
$\bar{y}_{M2, RSS}$	0.0239	2.6750	0.8691	105.7
$\bar{y}_{P, RSS}$	0.0235	2.6487	0.8545	106.8
$\bar{y}_{B, RSS}$	0.0190	2.4400	0.6909	115.9

6. CONCLUSION

The study modified the ratio estimator of [14] and derived the theoretical properties of the modified estimator up to order $O(n^{-1})$. The modified estimator was compared to the all the RSS estimators that were considered by [14] using the classical RSS ratio estimator of [16] as the basis of comparison. Ranked Set Samples various sizes were considered to test the performance of the various estimators and all sizes, the proposed modified estimator had the least bias and MSE. Compared to the classical RSS ratio estimator of [16], the efficiency of the proposed modified estimator ranged from 11% to 16% whilst the efficiency of the estimator of [14] ranged from 5% to 8%. The bias of the proposed modified estimator was also the least in all the sample combinations that were considered. This study therefore recommend the use of improved estimator for estimation since it can provide more efficient and more accurate estimates, compared to its competitors.

Funding: This work received no funding.

Conflict of Interest: The authors declare that they have no conflicts of interest.

Data Availability: The data that was used in this study is available within the paper.

REFERENCES

- [1] Afzal, I. and Masood, N. (2022). An exponentially ratio type estimator under ranked set sampling (rss) and its efficiency. *Journal of Statistics*, 26:44–53.
- [2] Ahmed, S. and Shabbir, J. (2019). Model based estimation of population total in presence of non-ignorable non-response. *PLoS ONE*, 14(10):e0222701.
- [3] Al-Saleh, M. F. and Samawi, H. M. (2007). A note on inclusion probability in ranked set sampling and some of its variations. *Test*, 16:198–209.
- [4] Brar, S. S. and Malik, S. (2014). Generalized ratio type estimator of population mean under ranked set sampling. *International Journal of Statistics and Reliability Engineering*, 1(2):178–192.
- [5] Cochran, W. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*, 30(2):262–275.
- [6] Cochran, W. G. (1977). *Sampling Techniques*. John Wiley & Sons, New York, NY, USA.
- [7] Dell, T. and Clutter, J. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, 28:545–555.
- [8] Jeelan, M. I., Bouza, C. N., and Sharma, M. (2017). Modified ratio estimator under rank set sampling. *Investigación Operacional*, 38(1):103–106.
- [9] Khan, Z. and Ismail, M. (2019). Modified ratio estimators of population mean based on ranked set sampling. *Pakistan Journal of Statistics and Operation Research*, 2:445–449.
- [10] Lynne Stokes, S. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics-Theory and Methods*, 6(12):1207–1211.
- [11] Maqbool, S. and Javaid, S. (2014). Modified ratio estimator using quartiles of auxiliary variable in ranked set sampling. *Int. J. Agricult. Stat. Sci*, 10(2):333–334.
- [12] McIntyre, G. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, 3(4):385–390.
- [13] Mehta, N. and Mandowara, V. (2016). A modified ratio-cum-product estimator of finite population mean using ranked set sampling. *Communications in Statistics-Theory and Methods*, 45(2):267–276.
- [14] Riyaz, S., Rather, K. U. I., Maqbool, S., and Jan, T. (2023). Ratio estimator of population mean using a new linear combination under ranked set sampling. *Reliability: Theory & Applications*, 18(2 (73)):479–485.
- [15] Saini, M. and Kumar, A. (2017). Ratio estimators for the finite population mean under simple random sampling and rank set sampling. *International Journal of System Assurance Engineering and Management*, 8:488–492.
- [16] Samawi, H. M. and Muttalak, H. A. (1996). Estimation of ratio using rank set sampling. *Biometrical Journal*, 38(6):753–764.
- [17] Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the institute of statistical mathematics*, 20(1):1–31.
- [18] Wolfe, D. A. (2010). Ranked set sampling. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):460–466.

BAYESIAN ESTIMATION OF PARAMETERS AND RELIABILITY CHARACTERISTICS IN THE INVERSE GOMPERTZ DISTRIBUTION

*¹TAIWO. M. ADEGOKE, ²LATIFAT A. ABIMBOLA, ³OLADAPO M. OLADOJA,
⁴OYINDAMOLA. R. OYEBANJO & ⁵K.O. OBISESAN

*^{1,2,3,4,5} Department of Mathematics and Statistics, First Technical University, Ibadan, Nigeria.

*¹taiw.o.adegoke@tech-u.edu.ng, ²latifat.abimbola@tech-u.edu.ng,

³oladapo.oladoja@tech-u.edu.ng, ⁴oyindamola.o.yebanjo@tech-u.edu.ng

⁵obisesan.olalekan@tech-u.edu.ng

Abstract

In this study, we derive Bayes' estimators for the unknown parameters of the Inverse Gompertz Distribution (IGD) using three alternative loss functions: the Squared Error Loss Function (SELF), the Entropy Loss Function (ELF), and the Linex Loss Function. Closed-form formulas for Bayes estimators are not possible when both parameters are unknown, hence Lindley's approximation (L-Approximation) is used for computation. We examine the performance of these estimators using their simulated hazards and assess their effectiveness in parameter estimation. It was discovered that as the sample size increases, parameter estimations became more precise and accurate across all functions. However, ELF consistently has lower MSE values than SELF and LINEX, indicating better parameter estimation. This pattern was also seen in the estimation of the hazard function, where ELF regularly beat SELF and LINEX, implying more efficient parameter estimation overall.

Keywords: Likelihood Function, Prior Distribution, Posterior Distribution, Bayes Estimates, Lindley Approximation

1. INTRODUCTION

Gompertz [1] proposed a probability distribution with two parameters, which is widely used in survival analysis to represent human mortality and behavioral sciences data. This distribution, a generalization of the exponential distribution, has many practical uses, particularly in medical and actuarial studies. It has considerable similarities to well-known distributions such as the Gumbel, Weibull, generalized logistic, exponential, and double exponential distributions [2].

However, the Gompertz distribution (GD) only shows an increasing failure rate, restricting its potential to represent occurrences across several fields. As a result, many authors have contributed to methodological studies and characterizations of this distribution to address real-world challenges in a variety of fields, including medical sciences, economics, behavioral sciences, engineering, biological studies, actuarial science, environmental studies, and lifetime analysis.

The Gompertz distribution and its variants have been the subject of extensive research. Read [3] offers a fundamental overview of the Gompertz distribution, including its features and applications in statistical fields. Makany [4] explores the theoretical foundations of Gompertz's curve and provides insights into its mathematical representation. Franses [5] discusses practical issues of fitting Gompertz curves to actual data. Wu and Lee [6] investigate combinations of Gompertz distributions, offering a framework for defining complicated systems. El-Gohary et al.

[7] introduce the generalized Gompertz distribution, which improves modeling flexibility. The beta-Gompertz distribution, proposed by Jafari et al. [8], enhances the flexibility of data capture. Khan et al. [9] introduce the transmuted Gompertz distribution, which can accommodate a wider range of data patterns. El-Bassiouny et al. [10, 11] study mixture models that combine the Gompertz distribution with other distributions to improve applicability in reliability and survival analysis. Rasool et al. [12] introduced the McDonald Gompertz distribution, which improves its ability to capture complicated data patterns. [13] introduced Topp-Leone Inverse Gompertz Distribution with different estimation procedures and application. Sanku et al. [14] assess and compare various estimating methodologies for the Gompertz distribution, assisting researchers and practitioners in selecting relevant methods.

2. INVERSE GOMPERTZ DISTRIBUTION

The random variable X is said to have an Inverse Gaussian Distribution (IGD) with shape parameter λ and scale parameter γ , if its cumulative distribution function (CDF) is given by

$$F(x) = e^{-\frac{\lambda}{\gamma} \left(e^{\frac{\gamma}{x}} - 1 \right)}, \quad x > 0, \quad \lambda, \gamma > 0 \quad (1)$$

The probability density function (PDF) of the Inverse Gaussian Distribution (IGD) is expressed as

$$f(x) = \frac{\lambda}{x^2} e^{-\frac{\lambda}{\gamma} \left(e^{\frac{\gamma}{x}} - 1 \right) + \frac{\gamma}{x}} \quad (2)$$

Furthermore, the reliability function is provided as follows:

$$R(x) = 1 - e^{-\frac{\lambda}{\gamma} \left(e^{\frac{\gamma}{x}} - 1 \right)} \quad (3)$$

The quantile function for the IGD distribution can be expressed as

$$q = \frac{\gamma}{\ln \left(1 - \frac{\gamma}{\lambda} \ln q \right)}, \quad 0 < q < 1. \quad (4)$$

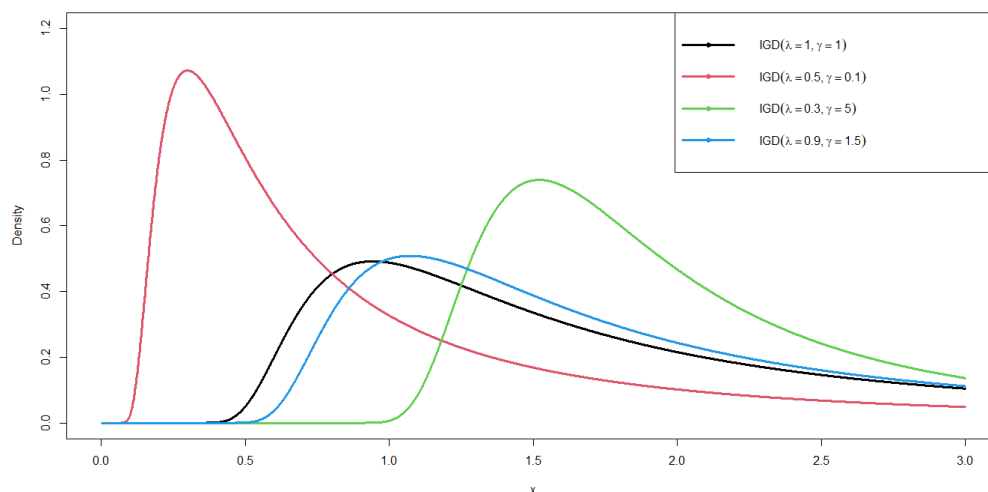


Figure 1: PDF of the IGD

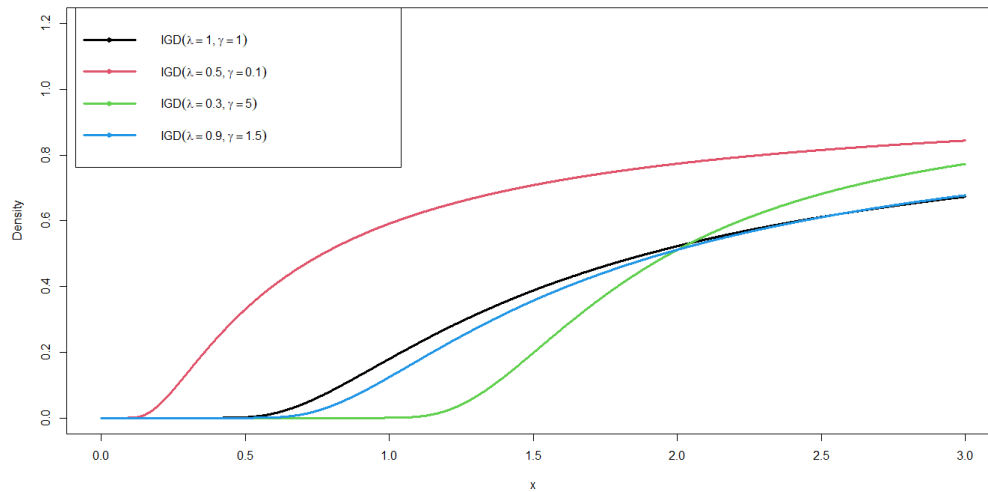


Figure 2: CDF of the IGD

3. BAYESIAN ESTIMATION TECHNIQUES

Let $x = (x_1, x_2, \dots, x_n)$ be a random variable with parameters λ and γ having a size n . From the bayes' the posterior probability density function of the parameters λ and γ given x can be expressed as

$$\Pr(\lambda, \gamma, \eta|x) = \frac{\pi(\lambda, \gamma)l(\lambda, \gamma)}{\int \int \int \pi(\lambda, \gamma)l(\lambda, \gamma)d(\lambda, \gamma)} \quad (5)$$

where $l(\lambda, \gamma)$ is the likelihood and (λ, γ) is the prior probability distribution.

3.1. Likelihood Function

Given a series of observations $x = (x_1, x_2, \dots, x_n)$ with parameters λ and γ having a size n for IG distribution (2), the likelihood function can be expressed as

$$l = \frac{\lambda^n}{\sum x^2} e^{-\frac{\lambda}{\gamma} \sum (e^{\frac{\gamma}{x}} - 1) + \sum (\frac{\gamma}{x})} \quad (6)$$

The log likelihood of IG distribution can be expressed as

$$L = \log l = n \log \lambda + \sum \left(\frac{\gamma}{x} \right) - 2 \sum \log(x) - \frac{\lambda}{\gamma} \sum \left(e^{\frac{\gamma}{x}} - 1 \right) \quad (7)$$

The maximum likelihood estimator of the shape and scale parameters for the parameters λ and γ is obtained by differentiating the (7) on parameters λ and γ . The maximum likelihood differential equations are:

$$\frac{dL}{d\lambda} = \frac{1}{\lambda} - \sum_{i=1}^n \left(e^{\frac{\gamma}{x_i}} - 1 \right) \frac{1}{\gamma} \quad (8)$$

$$\frac{dL}{d\gamma} = \sum_{i=1}^n \frac{1}{x_i} + \frac{\lambda \sum_{i=1}^n \left(\exp \left(\frac{\gamma}{x_i} \right) - 1 \right)}{\gamma^2} - \frac{\lambda \sum_{i=1}^n \left(\frac{\exp \left(\frac{\gamma}{x_i} \right)}{x_i} \right)}{\gamma} \quad (9)$$

Analytical solutions to equations (8) and (9) are not viable. The estimated values for the parameters λ and γ can be derived numerically using an iterative approach known as the Newton-Raphson method [15, 17, 16]. The Fisher information matrix elements for parameters λ and γ can

be represented as follows:

$$J_k = \begin{bmatrix} \frac{\partial^2 l(\lambda, \gamma)}{\partial \lambda^2} & \frac{\partial^2 l(\lambda, \gamma)}{\partial \lambda \partial \gamma} \\ \frac{\partial^2 l(\lambda, \gamma)}{\partial \lambda \partial \gamma} & \frac{\partial^2 l(\lambda, \gamma)}{\partial \gamma^2} \end{bmatrix} \quad (10)$$

The Jacobian matrix must be a non-singular symmetric matrix so its inverse must exist. So, using the Newton Raphson method we have

$$\begin{bmatrix} \lambda_{k+1} \\ \gamma_{k+1} \end{bmatrix} = \begin{bmatrix} \lambda_k \\ \gamma_k \end{bmatrix} - J_k^{-1} \begin{bmatrix} \frac{\partial l(\lambda, \gamma)}{\partial \lambda} \\ \frac{\partial l(\lambda, \gamma)}{\partial \gamma} \end{bmatrix} \quad (11)$$

with error term ϵ being the absolute differences between the new and the previous value of λ and γ in the iterative algorithm. That is

$$\epsilon \begin{bmatrix} \epsilon_{k+1}(\lambda) \\ \epsilon_{k+1}(\gamma) \end{bmatrix} = \left[\begin{bmatrix} \lambda_{k+1} \\ \gamma_{k+1} \end{bmatrix} - \begin{bmatrix} \lambda_k \\ \gamma_k \end{bmatrix} \right] \quad (12)$$

where λ_k and γ_k are the initial values of λ and γ respectively.
 where

$$L_{\lambda\lambda} = \frac{d^2 L}{d\lambda^2} = -\frac{1}{\lambda^2} \quad (13)$$

$$L_{\gamma\gamma} = \frac{d^2 L}{d\gamma^2} = -2 \cdot \frac{\lambda \sum_{i=1}^n \left(\exp\left(\frac{\gamma}{x_i}\right) - 1 \right)}{\gamma^3} + \frac{2 \cdot \lambda \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i} \right)}{\gamma^2} - \frac{\lambda \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i^2} \right)}{\gamma} \quad (14)$$

$$L_{\lambda\gamma} = \frac{d^2 L}{d\lambda d\gamma} = \frac{d^2 L}{d\gamma d\lambda} = \frac{-n + \sum_{i=1}^n \exp\left(\frac{\gamma}{x_i}\right)}{\gamma^2} - \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i} \right) \cdot \frac{1}{\gamma} \quad (15)$$

3.2. Prior Distribution

From (6), it can be observed that there is no proper conjugate distribution for the parameters λ and γ . Therefore, we will consider the use of independent gamma prior distribution for the scale with parameters a_1 and b_1 and shape parameters a_2 and b_2 . That is $\lambda \sim \text{Gamma}(a_1, b_1)$ and $\gamma \sim \text{Gamma}(a_2, b_2)$. The joint prior distribution can be expressed as

$$\pi(\lambda, \gamma) \propto \lambda^{a_1-1} \gamma^{a_2-1} e^{-b_1\lambda} e^{-b_2\gamma} \quad (16)$$

where a_1, a_2, b_1 and b_2 are hyper parameters.

3.3. Posterior Distribution

To obtain the posterior distribution for the IG distribution, we combine (6) and (16) and can be expressed as

$$P(\lambda, \gamma | X) = k^{-1} \lambda^{a_1+n-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \sum (e^{\frac{\gamma}{x}} - 1) + \sum (\frac{\gamma}{x}) - b_1\lambda - b_2\gamma} \quad (17)$$

where

$$k = \int_0^\infty \int_0^\infty \lambda^{a_1+n-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \sum (e^{\frac{\gamma}{x}} - 1) + \sum (\frac{\gamma}{x}) - b_1\lambda - b_2\gamma} d\lambda d\gamma$$

Analytical solutions for λ and γ from the posterior equation (17) are not viable due to its complicated nature, necessitating the use of numerical approaches such as Gibbs sampling, Metropolis-Hastings, EM algorithm, Lindley approximation, among others. In this study, we will use the Lindley approximation approach to obtain Bayesian estimates of λ and γ .

3.4. Loss Functions

The squared error is commonly employed as a loss function, however, its symmetric nature may not be acceptable in estimating issues with asymmetric losses. This disparity is especially pronounced in disciplines such as life testing and reliability estimation. In response, asymmetric loss functions, such as Varian's LINEX loss function [18], have gained popularity. [19] investigated the features of the LINEX loss function and discovered that the squared error loss is a specific instance of it. Another useful option is the entropy loss function.

In recent years, many authors have used Bayesian estimation for estimating the parameters of distributions. Examples include the works of Ahmed et al. [20], Basu & Ebrahimi [21], Nassar & Eissa [22], Pandey [23], Roio [24], Soliman et al. [31, 32, 33], Singh et al. [30, 25, 26], Adegoke et al [27], Ogunsanya et al. [28], Nzei et al. [29], and others.

We achieve the appropriate Bayesian estimates by using predefined loss functions such as squared error, LINEX, and entropy, which are defined as follows:

$$\begin{aligned} L_S(\hat{d}(\theta), d(\theta)) &= (\hat{d}(\theta) - d(\theta))^2, \\ L_L(\hat{d}(\theta), d(\theta)) &= e^{h(\hat{d}(\theta) - d(\theta))} - h(\hat{d}(\theta) - d(\theta)) - 1, \quad h \neq 0, \\ L_E(\hat{d}(\theta), d(\theta)) &\propto \left(\frac{\hat{d}(\theta)}{d(\theta)}\right)^w - w \log\left(\frac{\hat{d}(\theta)}{d(\theta)}\right) - 1, \quad w \neq 0, \end{aligned}$$

We get the desired Bayesian estimates. Here, $\hat{d}(\theta)$ is an estimate of $d(\theta)$. In the Bayesian paradigm, an optimal estimate for a certain loss function can be obtained by minimizing the average risk of $\hat{d}(\theta)$ relative to a weight function, also known as the prior distribution of θ . The Bayesian estimate, \hat{d}_{BS} , under the loss L_S , corresponds to the posterior mean of $d(\theta)$. by applying specified loss functions: squared error, LINEX, and entropy, which are described as follows. The Bayesian estimate of $d(\theta)$ for the loss function LL is provided as:

$$\hat{d}_{BL} = -\frac{1}{h} \log\left(\mathbb{E}_\theta\left[e^{-h\theta}|x\right]\right)$$

the equivalent estimate for the loss function LE is as follows:

$$\hat{d}_{BE} = \left(\mathbb{E}_\theta(\theta^{-w}|x)\right)^{-\frac{1}{w}}$$

given that the corresponding expectations $\mathbb{E}_\theta(\cdot)$ exist. We use loss functions L_S , L_L , and L_E to get Bayesian estimates of λ , γ , θ , the reliability function $R(t)$, and the hazard function $h(t)$.

Initially, we compute the Bayesian estimate for λ under the loss function L_S using the posterior distribution $P(\lambda, \gamma|x)$. This estimate is calculated as:

$$\hat{\lambda}_{BS} = k^{-1} \int_0^\infty \int_0^\infty \lambda^{a_1+n} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \Sigma(e^{\frac{\gamma}{x}} - 1) + \Sigma(\frac{\gamma}{x}) - b_1 \lambda - b_2 \gamma} \partial \lambda \partial \gamma \quad (18)$$

For the L_L loss function, the Bayesian estimate for λ is as follows:

$$\hat{\lambda}_{BL} = -\frac{1}{h} \log\left(\mathbb{E}\left[e^{-h\lambda}|x\right]\right) \quad h \neq 0$$

where

$$\mathbb{E}_\lambda\left[e^{-h\lambda}|x\right] = k^{-1} \int_0^\infty \int_0^\infty \lambda^{a_1+n-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \Sigma(e^{\frac{\gamma}{x}} - 1) + \Sigma(\frac{\gamma}{x}) - b_1 \lambda - b_2 \gamma - h\lambda} \partial \lambda \partial \gamma \quad (19)$$

Finally, when considering the loss function LE , we determine that

$$\hat{\lambda}_{BE} = \left(\mathbb{E}(\lambda^{-w}|x)\right)^{-\frac{1}{w}}$$

where

$$\mathbb{E}_\lambda(\lambda^{-w}|x) = k^{-1} \int_0^\infty \int_0^\infty \lambda^{a_1+n-w-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \Sigma(e^{\frac{\gamma}{x}}-1) + \Sigma(\frac{\gamma}{x}) - b_1\lambda - b_2\gamma} \partial\lambda\partial\gamma \quad (20)$$

Similarly, we proceed to derive Bayesian estimates for γ under the specified loss functions.

Assuming that λ and γ are unknown, we obtain equations for Bayesian estimates of the reliability function $R(t)$ in a similar manner. For the loss function L_S it is given as

$$\hat{R}(t) = k^{-1} \int_0^\infty \int_0^\infty \lambda^{a_1+n-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \Sigma(e^{\frac{\gamma}{x}}-1) + \Sigma(\frac{\gamma}{x}) - b_1\lambda - b_2\gamma} \left(1 - e^{-\frac{\lambda}{\gamma} \left(e^{\frac{\gamma}{t}} - 1 \right)} \right) \partial\lambda\partial\gamma \quad (21)$$

For the L_L loss function, we have

$$\hat{R}(t)_{BL} = -\frac{1}{h} \log \left(\mathbb{E} \left[e^{-hR(t)} | x \right] \right) \quad h \neq 0$$

where

$$\mathbb{E}_\lambda \left[e^{-hR(t)} | x \right] = k^{-1} \int_0^\infty \int_0^\infty \lambda^{a_1+n-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \Sigma(e^{\frac{\gamma}{x}}-1) + \Sigma(\frac{\gamma}{x}) - b_1\lambda - b_2\gamma} e^{-h \left(1 - e^{-\frac{\lambda}{\gamma} \left(e^{\frac{\gamma}{t}} - 1 \right)} \right)} \partial\lambda\partial\gamma \quad (22)$$

Finally, for the loss function L_E , it is found that

$$\hat{\lambda}_{BE} = (\mathbb{E}(R(t)^{-w}|x))^{-\frac{1}{w}}$$

$$\hat{R}(t)_{BE} = k^{-1} \int_0^\infty \int_0^\infty \lambda^{a_1+n-1} \gamma^{a_2-1} \sum x^{-2} e^{-\frac{\lambda}{\gamma} \Sigma(e^{\frac{\gamma}{x}}-1) + \Sigma(\frac{\gamma}{x}) - b_1\lambda - b_2\gamma} \left(1 - e^{-\frac{\lambda}{\gamma} \left(e^{\frac{\gamma}{t}} - 1 \right)} \right)^{-w} \partial\lambda\partial\gamma \quad (23)$$

3.5. Lindley Approximation

In the preceding section, we derived Bayes estimators for λ , γ , and θ using various loss functions, such as squared error, linex, and entropy. It is worth noting that these estimators are expressed as ratios of two integrals, which resist simplification into closed forms. Nonetheless, using the methods developed by Lindley [34], these Bayes estimators can be estimated to a form devoid of integrals. In practice, this strategy produces simple Bayes estimators that are easy to implement. Consider the ratio of the integral $I(X)$,

$$I(x) = E[u(\lambda, \gamma)|x] = \frac{\int \int u(\lambda, \gamma) e^{L(\lambda, \gamma) + G(\lambda, \gamma)} d\lambda d\gamma}{\int \int e^{L(\lambda, \gamma) + \rho(\lambda, \gamma)} d\lambda d\gamma}, \quad (24)$$

where:

- $u(\lambda, \gamma)$ is a function of λ and γ only;
- $L(\lambda, \gamma)$ is the log of likelihood;
- $\rho(\lambda, \gamma)$ is the log of joint prior of λ and γ .

This can be evaluated as

$$\begin{aligned}
 I(x) = & u(\hat{\lambda}, \hat{\gamma}) + \frac{1}{2} [(u_{\gamma\gamma} + 2\hat{u}_{\gamma}\hat{p}_{\gamma})\hat{\sigma}_{\gamma\gamma} + (u_{\lambda\gamma} + 2\hat{u}_{\lambda}\hat{p}_{\gamma})\hat{\sigma}_{\lambda\gamma} \\
 & + (u_{\gamma\lambda} + 2\hat{u}_{\gamma}\hat{p}_{\lambda})\hat{\sigma}_{\gamma\lambda} + (u_{\lambda\lambda} + 2\hat{u}_{\lambda}\hat{p}_{\lambda})\hat{\sigma}_{\lambda\lambda}] \\
 & + \frac{1}{2} [(u_{\gamma}\hat{\sigma}_{\gamma\gamma} + \hat{u}_{\lambda}\hat{\sigma}_{\gamma\lambda})(L_{\gamma\gamma\gamma}\hat{\sigma}_{\gamma\gamma} + L_{\gamma\lambda\gamma}\hat{\sigma}_{\gamma\lambda} \\
 & + L_{\lambda\gamma\gamma}\hat{\sigma}_{\lambda\gamma} + L_{\lambda\lambda\gamma}\hat{\sigma}_{\lambda\lambda}) + (u_{\gamma}\hat{\sigma}_{\lambda\gamma} + \hat{u}_{\lambda}\hat{\sigma}_{\lambda\lambda})(L_{\lambda\gamma\gamma}\hat{\sigma}_{\gamma\gamma} \\
 & + L_{\gamma\lambda\lambda}\hat{\sigma}_{\gamma\lambda} + L_{\lambda\gamma\lambda}\hat{\sigma}_{\lambda\gamma} + L_{\lambda\lambda\lambda}\hat{\sigma}_{\lambda\lambda})]
 \end{aligned} \tag{25}$$

where e:

- $\hat{\lambda}$ = MLE of λ ;
- $\hat{\gamma}$ = MLE of γ ;
- $\hat{u}_{\gamma} = \frac{\partial u(\hat{\lambda}, \hat{\gamma})}{\partial \gamma}$, $\hat{u}_{\lambda} = \frac{\partial u(\hat{\lambda}, \hat{\gamma})}{\partial \lambda}$, $\hat{u}_{\gamma\lambda} = \frac{\partial^2 u(\hat{\lambda}, \hat{\gamma})}{\partial \gamma \partial \lambda}$, $\hat{u}_{\lambda\gamma} = \frac{\partial^2 u(\hat{\lambda}, \hat{\gamma})}{\partial \lambda \partial \gamma}$;
- $\hat{u}_{\gamma\gamma} = \frac{\partial^2 u(\hat{\lambda}, \hat{\gamma})}{\partial \gamma^2}$, $\hat{u}_{\lambda\lambda} = \frac{\partial^2 u(\hat{\lambda}, \hat{\gamma})}{\partial \lambda^2}$;
- $\hat{L}_{\lambda\gamma\gamma} = \hat{L}_{\gamma\lambda\gamma} = \hat{L}_{\gamma\gamma\lambda} = \frac{\partial^3 L(\hat{\lambda}, \hat{\gamma})}{\partial \gamma \partial \gamma \partial \lambda}$, $\hat{L}_{\gamma\gamma\gamma} = \frac{\partial^3 L(\hat{\lambda}, \hat{\gamma})}{\partial \gamma \partial \gamma \partial \gamma}$, $\hat{L}_{\lambda\lambda\lambda} = \frac{\partial^3 L(\hat{\lambda}, \hat{\gamma})}{\partial \lambda \partial \lambda \partial \lambda}$;
- $\hat{L}_{\gamma\lambda\lambda} = \hat{L}_{\lambda\lambda\gamma} = \hat{L}_{\lambda\gamma\lambda} = \frac{\partial^3 L(\hat{\lambda}, \hat{\gamma})}{\partial \gamma \partial \lambda \partial \lambda}$;
- $\hat{p}_{\lambda} = \frac{\partial \pi(\hat{\lambda}, \hat{\gamma})}{\partial \lambda}$, $\hat{p}_{\gamma} = \frac{\partial \pi(\hat{\lambda}, \hat{\gamma})}{\partial \gamma}$.

where e

$$L_{\lambda\lambda\gamma} = 0; \quad L_{\lambda\lambda\lambda} = \frac{2}{\lambda^3} \tag{26}$$

$$L_{\gamma\gamma\lambda} = -2 \cdot \frac{-n + \sum_{i=1}^n \exp\left(\frac{\gamma}{x_i}\right)}{\gamma^3} + \frac{2 \cdot \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i}\right)}{\gamma^2} - \frac{\sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i^2}\right)}{\gamma} \tag{27}$$

$$L_{\gamma\gamma\gamma} = 6 \cdot \frac{\lambda \sum_{i=1}^n \left(\exp\left(\frac{\gamma}{x_i}\right) - 1\right)}{\gamma^4} - 6 \cdot \frac{\lambda \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i}\right)}{\gamma^3} + 3 \cdot \frac{\lambda \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i^2}\right)}{\gamma^2} - \frac{\lambda \sum_{i=1}^n \left(\frac{\exp\left(\frac{\gamma}{x_i}\right)}{x_i^3}\right)}{\gamma} \tag{28}$$

$$\log \pi(\lambda, \gamma) = (a_1 - 1) * \log(\lambda) + (a_2 - 1) * \log(\gamma) - b_1\lambda - b_2\gamma$$

$$\rho_{\lambda} = \frac{a_1 - 1}{\lambda} - b_1; \quad \rho_{\gamma} = \frac{a_2 - 1}{\gamma} - b_2$$

3.5.1 Bayes estimates of the parameters of IGD and its reliability

To obtain the bayes estimate under SELF for $\hat{\lambda}$, $u(\hat{\lambda}, \hat{\gamma}) = \hat{\lambda}$, $u_{\lambda\lambda} = u_{\lambda\gamma} = u_{\gamma\gamma} = u_{\gamma\lambda} = u_{\gamma} = 0$ and $u_{\lambda} = 1$. Substituting these values into (25), we have

$$\hat{\lambda}_{BS} = \hat{\lambda} + \hat{p}_{\lambda}\hat{\sigma}_{\lambda\lambda} + \frac{1}{2}L_{\lambda\lambda\lambda}\hat{\sigma}_{\lambda\lambda} \tag{29}$$

also to obtain the bayes estimate under SELF for $\hat{\gamma}$, $u(\hat{\lambda}, \hat{\gamma}) = \hat{\gamma}$, $u_{\lambda\lambda} = u_{\lambda\gamma} = u_{\gamma\gamma} = u_{\gamma\lambda} = u_{\lambda} = 0$ and $u_{\gamma} = 1$. Substituting these values into (25), we have

$$\hat{\gamma}_{BS} = \hat{\gamma} + \hat{p}_{\gamma}\hat{\sigma}_{\gamma\gamma} + \frac{1}{2}\hat{\sigma}_{\gamma\gamma}L_{\gamma\gamma\gamma}\hat{\sigma}_{\gamma\gamma} \tag{30}$$

To obtain the bayes estimate of $\hat{\lambda}$ under the ELF, $u(\hat{\lambda}, \hat{\gamma}) = \lambda^{-w}$, then $u_{\lambda} = -w\lambda^{-w-1}$, $u_{\lambda\lambda} = w(w+1)\lambda^{-w-2}$ and $u_{\lambda\gamma} = u_{\gamma\lambda} = u_{\gamma\lambda} = u_{\gamma} = 0$. Substituting these values into (25), we have

$$\hat{\lambda}_{BE} = \hat{\lambda}^{-w} + \frac{1}{2} [\hat{\sigma}_{\lambda\lambda} (\hat{u}_{\lambda\lambda} + 2\hat{u}_{\lambda}\rho_{\lambda})] + \frac{1}{2} [(\hat{u}_{\lambda}\hat{\sigma}_{\lambda\lambda} (L_{\lambda\lambda\lambda}\hat{\sigma}_{\lambda\lambda}))] \quad (31)$$

also to obtain the bayes estimate of $\hat{\gamma}$ under the ELF, $u(\hat{\lambda}, \hat{\gamma}) = \gamma^{-w}$, then $u_{\gamma} = -w\lambda^{-w-1}$, $u_{\gamma\gamma} = w(w+1)\lambda^{-w-2}$ and $u_{\lambda\gamma} = u_{\lambda\lambda} = u_{\gamma\lambda} = u_{\gamma} = 0$. Substituting these values into (25), we have

$$\hat{\gamma}_{BE} = \hat{\gamma}^{-w} + \frac{1}{2} [\hat{\sigma}_{\gamma\gamma} (\hat{u}_{\gamma\gamma} + 2\hat{u}_{\gamma}\rho_{\gamma})] + \frac{1}{2} [(\hat{u}_{\gamma}\hat{\sigma}_{\gamma\gamma} (L_{\gamma\gamma\gamma}\hat{\sigma}_{\gamma\gamma}))] \quad (32)$$

To obtain the bayes estimate of λ under the LLF, $u(\hat{\lambda}, \hat{\gamma}) = e^{-h\lambda}$, then $u_{\lambda} = -he^{-h\lambda}$, $u_{\lambda\lambda} = h^2e^{-h\lambda}$ and $u_{\lambda\gamma} = u_{\gamma\lambda} = u_{\gamma\lambda} = u_{\gamma} = 0$. Substituting these values into (25), we have

$$\hat{\lambda}_{BL} = e^{-h\lambda} + \frac{1}{2} [\hat{\sigma}_{\lambda\lambda} (\hat{u}_{\lambda\lambda} + 2\hat{u}_{\lambda}\rho_{\lambda})] + \frac{1}{2} [(\hat{u}_{\lambda}\hat{\sigma}_{\lambda\lambda} (L_{\lambda\lambda\lambda}\hat{\sigma}_{\lambda\lambda}))] \quad (33)$$

also to obtain the bayes estimate of γ under the LLF, $u(\hat{\lambda}, \hat{\gamma}) = e^{-h\gamma}$, then $u_{\gamma} = -he^{-h\lambda}$, $u_{\gamma\gamma} = h^2e^{-h\lambda}$. and $u_{\lambda\gamma} = u_{\lambda\lambda} = u_{\gamma\lambda} = u_{\gamma} = 0$. Substituting these values into (25), we have

$$\hat{\gamma}_{BL} = e^{-h\gamma} + \frac{1}{2} [\hat{\sigma}_{\gamma\gamma} (\hat{u}_{\gamma\gamma} + 2\hat{u}_{\gamma}\rho_{\gamma})] + \frac{1}{2} [(\hat{u}_{\gamma}\hat{\sigma}_{\gamma\gamma} (L_{\gamma\gamma\gamma}\hat{\sigma}_{\gamma\gamma}))] \quad (34)$$

Under the SELF the bayes estimates for the reliability of IGD can be obtained by equating

$$\begin{aligned} u &= 1 - e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}; & u_{\lambda} &= \frac{(e^{\frac{\gamma}{t}} - 1) \cdot e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}}{\gamma} \\ u_{\gamma} &= -\left(\frac{\lambda(e^{\frac{\gamma}{t}} - 1)}{\gamma^2} - \frac{\lambda e^{\frac{\gamma}{t}}}{\gamma t}\right) \cdot e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}; & u_{\lambda\lambda} &= -\frac{(e^{\frac{\gamma}{t}} - 1)^2 \cdot e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}}{\gamma^2} \\ u_{\gamma\gamma} &= -\left(-2\frac{\lambda(e^{\frac{\gamma}{t}} - 1)}{\gamma^3} + 2\frac{\lambda e^{\frac{\gamma}{t}}}{\gamma^2 t} - \frac{\lambda e^{\frac{\gamma}{t}}}{\gamma t^2}\right) e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)} - \left(\frac{\lambda(e^{\frac{\gamma}{t}} - 1)}{\gamma^2} - \frac{\lambda e^{\frac{\gamma}{t}}}{\gamma t}\right)^2 e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)} \\ u_{\lambda\gamma} &= u_{\gamma\lambda} = -\frac{(e^{\frac{\gamma}{t}} - 1) \cdot e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}}{\gamma^2} + \frac{e^{\frac{\gamma}{t}} \cdot e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}}{\gamma t} + \frac{(e^{\frac{\gamma}{t}} - 1) \cdot \left(\frac{\lambda(e^{\frac{\gamma}{t}} - 1)}{\gamma^2} - \frac{\lambda e^{\frac{\gamma}{t}}}{\gamma t}\right) \cdot e^{-\frac{\lambda}{\gamma}(e^{\frac{\gamma}{t}} - 1)}}{\gamma} \end{aligned}$$

and substituting the values into (25). We have

$$\begin{aligned} I(x) &= u(\hat{\lambda}, \hat{\gamma}) + \frac{1}{2} [(\hat{u}_{\gamma\gamma} + 2\hat{u}_{\gamma}\hat{\rho}_{\gamma})\hat{\sigma}_{\gamma\gamma} + (\hat{u}_{\lambda\lambda} + 2\hat{u}_{\lambda}\hat{\rho}_{\lambda})\hat{\sigma}_{\lambda\lambda}] \\ &+ \frac{1}{2} [(\hat{u}_{\gamma}\hat{\sigma}_{\gamma\gamma})(L_{\gamma\gamma\gamma}\hat{\sigma}_{\gamma\gamma} + L_{\lambda\lambda\gamma}\hat{\sigma}_{\lambda\lambda}) + (\hat{u}_{\lambda}\hat{\sigma}_{\lambda\lambda})(L_{\lambda\gamma\gamma}\hat{\sigma}_{\gamma\gamma} \\ &+ L_{\lambda\lambda\lambda}\hat{\sigma}_{\lambda\lambda})]. \end{aligned} \quad (35)$$

Similarly, we can evaluate the Bayes estimators for the reliability function using the ELF and LLF.

3.6. Simulation Study

In this part, we undertake a simulation research to estimate the parameters and reliability of the Inverse Gamma (IG) distribution across several λ and γ combinations: (0.9, 0.6), (1.0, 1.0), (1.0, 0.7), and (1.2, 0.8). The population parameter is created with R programming version 4.3.1. Sampling distributions are calculated for various sample sizes $n = [30, 50, 100, 500]$ using $R = 1000$ replications. Tables 1 and 2 show the calculated estimates and mean square errors (MSE) in brackets.

Table 1: Bayes estimates for different parameter values under the SELF, ELF and LINEX

		SELF		ELF		LINEX	
		$\hat{\lambda}_{BS}$	$\hat{\gamma}_{BS}$	$\hat{\lambda}_{BE}$	$\hat{\gamma}_{BE}$	$\hat{\lambda}_{BL}$	$\hat{\gamma}_{BL}$
n =30	$\lambda = 0.9$	0.9300 (0.0077)	0.7191 (0.1106)	0.8560 (0.0027)	0.8274 (0.0884)	0.6464 (0.0646)	0.6600 (0.0182)
	$\gamma = 0.6$	0.9524 (0.0061)	0.6090 (0.0430)	0.8626 (0.0018)	0.7691 (0.0475)	0.6465 (0.0647)	0.6988 (0.0172)
n =50	$h = 0.6$	0.967 (0.0059)	0.6103 (0.0273)	0.8689 (0.0011)	0.7744 (0.0417)	0.6414 (0.0670)	0.6964 (0.0139)
	$w = -0.5$	0.9859 (0.0076)	0.6026 (0.0050)	0.8759 (0.0005)	0.775 (0.0327)	0.6373 (0.0690)	0.6971 (0.0103)
n = 100	$a_1 = 1$	1.0589 (0.2014)	1.2097 (0.3130)	1.2692 (0.0623)	0.9871 (0.0775)	1.1225 (0.1425)	1.1300 (0.0203)
	$a_2 = 1$	1.0731 (0.1880)	1.0216 (0.1199)	1.2397 (0.0728)	1.0376 (0.0482)	1.1253 (0.1404)	1.1081 (0.0131)
n = 500	$b_1 = 1$	1.0956 (0.1654)	1.0206 (0.0760)	1.2359 (0.0723)	1.0171 (0.0226)	1.1277 (0.1385)	1.1078 (0.0125)
	$b_2 = 0.5$	1.1195 (0.1448)	1.0051 (0.01408)	1.2265 (0.0752)	1.0021 (0.0033)	1.1307 (0.1363)	1.1058 (0.0113)
n =30	$\lambda = 1.0$	1.4021 (0.2226)	0.8344 (0.1419)	1.2489 (0.0903)	0.8529 (0.1103)	0.8726 (0.0166)	0.9204 (0.04964)
	$\gamma = 0.7$	1.4760 (0.2700)	0.7099 (0.0553)	1.2998 (0.1089)	0.7538 (0.0441)	0.8666 (0.0180)	0.9317 (0.0541)
n =50	$h = 0.1$	1.4773 (0.2530)	0.7113 (0.0350)	1.3025 (0.1025)	0.7575 (0.0288)	0.8662 (0.018)	0.9315 (0.0538)
	$w = -0.8$	1.4968 (0.2519)	0.7030 (0.0065)	1.3174 (0.1029)	0.7536 (0.0076)	0.8644 (0.0184)	0.9321 (0.0539)
n = 100	$a_1 = 1$	0.9083 (0.1015)	0.9659 (0.2014)	0.8734 (0.1193)	0.9612 (0.1689)	1.2290 (0.0016)	1.217 (0.1844)
	$a_2 = 1$	0.8943 (0.1192)	0.8157 (0.0772)	0.8592 (0.1357)	0.8283 (0.0662)	1.2290 (0.0019)	1.1788 (0.1478)
n = 500	$b_1 = 1$	0.9272 (0.0853)	0.8157 (0.0488)	0.8867 (0.1067)	0.8299 (0.0418)	1.2373 (0.0018)	1.1782 (0.1458)
	$b_2 = 1$	0.9533 (0.0618)	0.8039 (.0090)	0.9077 (0.0863)	0.8211 (0.0080)	1.245 (0.0020)	1.1746 (0.1408)

Table 2: Bayes estimates for the hazard function under the SELF, ELF and LINEX

		$\hat{R}(t)_{BS}$	$\hat{R}(t)_{BE}$	$\hat{R}(t)_{BL}$
n = 30	$\lambda = 0.9$	0.5728 (0.1114)	0.7343 (0.0288)	0.7212 (0.0326)
	$\gamma = 0.6$			
n = 50	$h = 0.6$	0.5499 (0.1256)	0.7208 (0.0328)	0.7304 (0.0291)
	$w = -0.5$			
n = 100	$a_1 = 1$	0.5517 (0.1230)	0.7214 (0.0323)	0.7301 (0.0290)
	$a_2 = 1$			
n = 500	$b_1 = 1$	0.5501 (0.1227)	0.7206 (0.0322)	0.7309 (0.0286)
	$b_2 = 0.5$			
t = 1				
n = 30	$\lambda = 1.0$	0.3971 (1.218)	1.6789 (0.0375)	1.0409 (0.2107)
	$\gamma = 1.0$			
n = 50	$h = -0.1$	0.36002 (0.2917)	1.9986 (1.2129)	1.0371 (0.0188)
	$w = 0.5$			
n = 100	$a_1 = 0.5$	0.36199 (0.2895)	1.9976 (1.2085)	1.0373 (0.0188)
	$a_2 = 0.5$			
n = 500	$b_1 = 0.5$	0.3650 (0.2862)	1.9876 (1.1837)	1.0376 (0.0189)
	$b_2 = 0.5$			
t = 2				
n = 30	$\lambda = 1$	0.4004 (0.3605)	0.4639 (0.2881)	0.9609 (0.0015)
	$\gamma = 0.7$			
n = 50	$h = 0.1$	0.4051 (0.3542)	0.4674 (0.2838)	0.9605 (0.0015)
	$w = -0.8$			
n = 100	$a_1 = 1$	0.4077 (0.3511)	0.4696 (0.2813)	0.9603 (0.0015)
	$a_2 = 1$			
n = 500	$b_1 = 0.5$	0.4124 (0.3452)	0.4738 (0.2768)	0.9598 (0.0016)
	$b_2 = 0.5$			
t = 3				
n = 30	$\lambda = 1.2$	0.1776 (1.046)	0.2037 (0.9935)	0.9656 (0.0549)
	$\gamma = 0.8$			
n = 50	$h = -0.2$	0.1695 (1.063)	0.1950 (1.011)	0.9672 (0.0542)
	$w = -0.9$			
n = 100	$a_1 = 1$	0.1741 (1.0529)	0.1995 (1.0015)	0.9664 (0.0545)
	$a_2 = 1$			
n = 500	$b_1 = 1$	0.1771 (1.0464)	0.2021 (0.9957)	0.9659 (0.0548)
	$b_2 = 1$			
t = 5				

Table 1 shows Bayesian estimates for various parameter values using three loss functions: SELF, ELF, and LINEX, with varied sample sizes. Each cell includes the estimated value of parameters ($\hat{\lambda}$ and $\hat{\gamma}$) with their standard errors in parentheses. Generally, as the sample size grows, the estimates get more precise, as evidenced by decreasing standard errors. The three loss functions act differently depending on the parameter values. However, it is clear that the ELF loss function consistently produces estimates with fewer standard errors than SELF and LINEX, implying greater performance in parameter estimation. This trend persists across a wide range of sample sizes and parameter values, demonstrating the efficiency of the ELF loss function in Bayesian estimation.

Table 2 shows Bayesian estimates of the hazard function for three different loss functions: SELF, ELF, and LINEX, across a range of sample sizes and parameter values. Increasing sample sizes often results in lower mean squared error (MSE) across all three functions, indicating better parameter estimate accuracy. However, performance differences exist amongst the loss algorithms

at different parameter settings. with example, with $\lambda = 0.9, \gamma = 0.6, h = 0.6, w = -0.5, a_1 = 1, a_2 = 1, b_1 = 1,$ and $b_2 = 0.5,$ the ELF loss function consistently produces the lowest MSE compared to SELF and LINEX. This pattern holds true across other parameter settings, implying that the ELF function outperforms the MSE.

3.7. Real life Application

In this section, we look at the dataset published by Balakrishnan et al. [35], which includes 134 entries representing scores on the General Rating of Affective Symptoms for Preschoolers (GRASP) scale. Using Bayesian approaches, we obtain the parameter estimates and reliability ratings for the Inverse Gamma (IG) distribution over a variety of loss functions.

Table 3: Bayes estimate for the parameter of IGD under different loss functions when $a_1 = 1, a_2 = 1, b_1 = 0.5$ and $b_2 = 0.5$

	SELF	ELF w = -0.7	ELF w = 1.2	LINEX h = -0.5	LINEX h = 0.5
$\hat{\lambda}$	0.2959	0.2962	0.29226	0.3520	0.3812
$\hat{\gamma}$	153.1028	152.3344	161.8959	156.055	156.0564

Table 4: Bayes estimate for the reliability function under different loss functions for different parameter values

		$a_1 = a_2 = 1, b_1 = b_2 = 0.5$	$a_1 = a_2 = 1, b_1 = b_2 = 1$
SELF	t = 1	0.3820	0.3804
	t = 5	0.3805	0.3789
ELF	w = -1.5 t = 1	0.3821	0.3811
	w = 1.5 t = 5	0.3830	0.3799
LINEX	h = 1 t = 1	0.3678	0.3651
	h = -1 t = 5	0.3679	0.3645

Table 3 shows the the Bayes estimates for the parameters of IG distribution under different loss functions. Also, Table 4 display the reliability estimates under different loss functions and parameter values.

4. CONCLUSION

Table 1 compares Bayesian parameter estimation for three different loss functions: SELF, ELF, and LINEX. Overall, as sample size grows, parameter estimates become more precise and accurate across all loss functions. However, the ELF loss function consistently produces lower mean squared error (MSE) values than SELF and LINEX, indicating more effective parameter estimation. This shows that the ELF loss function may perform better in terms of balancing precision and accuracy, making it an attractive option for Bayesian parameter estimation applications. Table 2 shows Bayesian estimates for the hazard function using three alternative loss functions: SELF, ELF, and LINEX. It demonstrates how the performance of these estimators fluctuates with sample size and parameter values. In general, as sample size increases, mean squared error (MSE) decreases across all three loss functions, indicating that parameter estimations are more accurate and precise. The ELF loss function regularly produces lower MSE values than SELF and LINEX, indicating more efficient parameter estimation.

REFERENCES

- [1] Gompertz, B. (1819). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 109, 513–585.
- [2] Wilkenskens, F. (1999). Continuous-time demographic models for the population of the United States: United States working paper. *United Nations Population Division*.
- [3] Read, C. B. (1975). Gompertz distribution. *Encyclopedia of Statistical Sciences*. Wiley, New York.
- [4] Makany, R. A. (2012). Theoretical basis of Gompertz's curve. *Biometrical Journal*, 33, 121–128.
- [5] Franses, P. H. (1998). Fitting a Gompertz curve. *Journal of the Operational Research Society*, 45, 109–113.
- [6] Wu, J. W. and Lee, W. C. (2003). Characterization of the mixtures of Gompertz distributions by conditional expectation of order statistics. *Biometrical Journal*, 41, 371–381.
- [7] El-Gohary, A., Ahmad, A., and Adel Naif, A. (2010). The generalized Gompertz distribution. *Applied Mathematics and Modeling*, 37, 13–24.
- [8] Jafari, A., Saeid, T., and Morad, A. (2016). The beta-Gompertz distribution. *Revista Colombiana de Estadística*, 37(1), 139–156.
- [9] Khan, M. S., Robert, K., and Irene, L. H. (2018). Transmuted Gompertz distribution: properties and estimation. *Pakistan Journal of Statistics*, 32(3), 161–182.
- [10] El-Bassiouny, A. H., Medhat, E. L.-D., Abdelfattah, M., and Eliwa, M. S. (2018). Mixture of exponentiated generalized Weibull–Gompertz distribution and its applications in reliability. *Journal of Statistical Applications in Probability*, 5(3), 1–14.
- [11] El-Bassiouny, A. H., Medhat, E. L.-D., Abdelfattah, M., and Eliwa, M. S. (2019). Exponentiated generalized Weibull–Gompertz distribution with application in survival analysis. *Journal of Statistical Applications in Probability*, 6(1), 7–16.
- [12] Rasool, R., Saeid, T., and Jafari, A. (2017). The McDonald Gompertz distribution: properties and applications. *Communications in Statistics - Simulation and Computation*, 47(5), 3341–3355.
- [13] Adegoke, T. M., Oladoja, O. M., Bashir, S. O., Mustapha, A. A., Aderupatan, D. E., & Nzei, L. C. (2023). Topp-Leone inverse Gompertz distribution: Properties and different estimation techniques and applications. *Pakistan Journal of Statistics*, 39(4), 433–456. <https://www.pakjst.com/wp-content/uploads/2023/08/39402.pdf>
- [14] Sanku, D., Tanmay, K., and Yogesh, M. (2018). Evaluation and comparison of estimators in the Gompertz distribution. *Annals of Data Science*, 5(2), 235–258.
- [15] Obisesan, K.O., Adegoke, T.M., Adekanmbi, D.B., and Lawal, M. (2015). Numerical approximation to intractable likelihood functions. *Perspectives and Developments in Mathematics*, 301–324.
- [16] Bakari, H.R., Adegoke, T.M., and Yahya, A.M. (2016). Application of Newton Raphson method to non-linear models. *International Journal of Mathematics and Statistics Studies*, 4(4), 21–31.
- [17] Yahya, A.M., Dibal, N.P., Bakari, H.R., and Adegoke, T.M. (2016). Obtaining parameter estimate from the truncated Poisson probability distribution. *North Asian International Research Journal of Sciences, Engineering & I.T.*, 2(9), 3–10.
- [18] Varian, H. R. (1975). A Bayesian approach to real estate assessment. In *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage* (pp. 195–208).
- [19] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, 81.
- [20] Ahmed, A., Al-Kutubi, H., & Ibrahim, N. (2010). Comparison of the Bayesian and maximum likelihood estimation for Weibull distribution. *Journal of Mathematics and Statistics*, 6, 100–104.
- [21] Basu, A., & Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of Statistical Planning and Inference*, 29, 21–31.
- [22] Nassar, M., & Eissa, F. H. (2005). Bayesian estimation for the exponentiated Weibull model. *Communications in Statistics-Theory and Methods*, 33, 2343–2362.

- [23] Pande y, B. (1997). Testimator of the scale parameter of the exponential distribution using linex loss function. *Communications in Statistics-Theory and Methods*, 26, 2191–2202.
- [24] Roio, J. (1987). On the admissibility of $c[\bar{x}] + d$ with respect to the linex loss function. *Communications in Statistics-Theory and Methods*, 16, 3745–3748.
- [25] Singh, S. K., Singh, U., & Shar ma, V. K. (2013a). Bayesian estimation and prediction for flexible Weibull model under type-II censoring scheme. *Journal of Probability and Statistics*, 2013.
- [26] Singh, S. K., Singh, U., & Shar ma, V. K. (2013b). Bayesian estimation and prediction for the generalized Lindley distribution under asymmetric loss function. *Haceteppe Journal of Mathematics and Statistics*.
- [27] Adegoke, T. M., Obisesan, K. O., Oladoja, O. M., & Adegoke, G. K. (2023). Bayesian and classical estimations of transmuted inverse Gompertz distribution. *Reliability: Theory & Applications*, 2(73), 24–38. DOI: 10.24412/1932-2321-2023-273-207-222
- [28] Ogunsanya, A. S., Yahya, W. B., Adegoke, T. M., Iluno, C., Ader ele, O. R., & Ekum, M. I. (2021). A new three-parameter Weibull inverse Rayleigh distribution: Theoretical development and applications. *Mathematics and Statistics*, 9(3), 249–272. DOI: 10.13189/ms.2021.090306
- [29] Nzei, C. Lawrence, Adegoke, M. Taiwo, Ekhosuehi, N., & Mbegbu, I. Julian. (2024). Bayesian estimation of Topp-Leone Lindley (TLL) distribution parameters under different loss functions using Lindley approximation. *Reliability: Theory & Applications*, 1(77), 50–64. DOI: 10.24412/1932-2321-2024-177-50-64
- [30] Singh, U., Gupta, P. K., & Upadhy ay, S. (2002). Estimation of exponentiated Weibull shape parameters under linex loss function. *Communications in Statistics-Simulation and Computation*, 31, 523–537.
- [31] Soliman, A. et al. (2002). Reliability estimation in a generalized life-model with application to the Burr-XII. *IEEE Transactions on Reliability*, 51, 337–343.
- [32] Soliman, A. et al. (2005). Estimation of parameters of life from progressively censored data using Burr-XII model. *IEEE Transactions on Reliability*, 54, 34–42.
- [33] Soliman, A. A., Ellah, A. A., & Sultan, K. (2006). Comparison of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches. *Computational Statistics & Data Analysis*, 51, 2065–2077.
- [34] Lindley, D. V. (1980). Approximate Bayesian methods. *Trabajos de Estadística y de Investigación Operativa*, 31, 223–245.
- [35] Balakrishnan, N., Victor, L., & Antonio, S. (2010). A mixture model based on Birnbaum-Saunders distributions: A study conducted by authors regarding the scores of the GRASP (general rating of affective symptoms for preschoolers), in a city located at the south part of Chile.

A TWO-PARAMETER ARADHANA DISTRIBUTION WITH APPLICATIONS TO RELIABILITY ENGINEERING

Ravi Shanker and Nitesh Kumar Soni

•

Department of Mathematics, G.L.A. College, Nilamber-Pitamber University, Daltonganj,
Jharkhand, India

E-mail: ravi.shanker74@gmail.com ; nitinpearl26@gmail.com

Rama Shanker, Mousumi Ray, and Hosenuur Rahman Prodhani

•

Department of Statistics, Assam University, Silchar, India

E-mail: shankerrama2009@gmail.com; mousumiray616@gmail.com; hosenuur72@gmail.com

Abstract

The search for a statistical distribution for modelling the reliability data from reliability engineering is challenging and the main cause is the stochastic nature of the data and the presence of skewness, kurtosis and over-dispersion. During recent decades several one and two-parameter statistical distributions have been proposed in statistics literature but all these distributions were unable to capture the nature of data due to the presence of skewness, kurtosis and over-dispersion in the data. In the present paper, two-parameter Aradhana distribution, which includes one parameter Aradhana distribution as a particular case, has been proposed. Using convex combination approach of deriving a new statistical distribution, a two-parameter Aradhana distribution has been proposed. Various interesting and useful statistical properties including survival function, hazard function, reverse hazard function, mean residual life function, stochastic ordering, deviation from mean and median, stress-strength reliability, Bonferroni and Lorenz curve and their indices have been discussed. The raw moments, central moments and descriptive measures based on moments of the proposed distribution have been obtained. The estimation of parameters using the maximum likelihood method has been explained. The simulation study has been presented to know the performance in terms of consistency of maximum likelihood estimators as the sample size increases and. The goodness of test of the proposed distributions has been tested using the values of Akaike Information criterion and Kolmogorov-Smirnov statistics. Finally, two examples of real lifetime datasets from reliability engineering have been presented to demonstrate its applications and the goodness of fit, and it shows a better fit over two-parameter generalized Aradhana distribution, quasi Aradhana distribution, new quasi Aradhana distribution, Power Aradhana distribution, weighted Aradhana distribution, gamma distribution and Weibull distribution. The flexibility, tractability and usefulness of the proposed distribution show that it is very much useful for modelling reliability data from reliability engineering. As this is a new distribution and it has wide applications, it will draw the attention of researchers in reliability engineering and biomedical sciences to search many more applications in the future.

Keywords: Aradhana distribution, reliability properties, maximum likelihood estimation, applications.

I. INTRODUCTION

Shanker [1] proposed the Aradhana distribution, a one parameter lifetime distribution designed to characterize lifetime data originating from the fields of biomedical sciences and engineering. This distribution is characterized by its probability density function (pdf) and cumulative distribution function (cdf) as

$$f(x; \theta) = \frac{\theta^3}{\theta^2 + 2\theta + 2} (1+x)^2 e^{-\theta x}; x > 0, \theta > 0$$

$$F(x; \theta) = 1 - \left[1 + \frac{\theta x(\theta x + 2\theta + 2)}{\theta^2 + 2\theta + 2} \right] e^{-\theta x}; x > 0, \theta > 0$$

It has been shown by Shanker [1] that in most of the real lifetime datasets it exhibited superior fit in comparison to exponential, Lindley [2], Shanker and Akash distributions introduced by Shanker [3,4]. The Aradhana distribution is a convex combination of exponential (θ), gamma ($2, \theta$) and gamma ($3, \theta$) distributions with their proportions $\frac{\theta^2}{\theta^2 + 2\theta + 2}$, $\frac{2\theta}{\theta^2 + 2\theta + 2}$ and $\frac{2}{\theta^2 + 2\theta + 2}$ respectively.

The mean and variance of Aradhana distribution are

$$E(X) = \frac{\theta^2 + 4\theta + 6}{\theta(\theta^2 + 2\theta + 2)} \text{ and } Var(X) = \frac{\theta^4 + 8\theta^3 + 24\theta^2 + 24\theta + 12}{\theta^2(\theta^2 + 2\theta + 2)^2}.$$

Important statistical properties of Aradhana distribution including shapes for varying values of parameter, moments related measures, hazard function, mean residual life function, stochastic ordering, mean deviations, distribution of order Statistics, Bonferroni and Lorenz curves, Renyi entropy measure and stress-strength reliability have been discussed and also studied estimation of parameter and applications of Aradhana distribution for modelling lifetime data by Shanker [1].

Shanker et al [5] have introduced a quasi Aradhana distribution (QAD) with its pdf and cdf as

$$f(x, \theta, \alpha) = \frac{\theta}{\alpha^2 + 2\alpha + 2} (\alpha + \theta x)^2 e^{-\theta x}; x > 0, \theta > 0, \alpha > 0$$

$$F(x, \theta, \alpha) = 1 - \left[1 + \frac{\theta x(\theta x + 2\alpha + 2)}{\alpha^2 + 2\alpha + 2} \right] e^{-\theta x}; x > 0, \theta > 0, \alpha > 0.$$

The detailed studies on various statistical properties, estimation of parameters and applications of QAD are available in Shanker et al [5]. Anthony and Elangovan [6] discussed the length-biased version of QAD and study its properties and applications. Further, Anthony and Elangovan [7] proposed a new generalization of QAD by introducing an additional parameter in the QAD and study its statistical properties and applications.

Shanker et al [8] proposed a new quasi Aradhana distribution defined by its pdf and cdf

$$f(x; \theta, \alpha) = \frac{\theta^3}{\theta^4 + 2\theta^2\alpha + 2\alpha^2} (\theta + \alpha x)^2 e^{-\theta x}; x > 0, \theta > 0, \alpha > 0$$

$$F(x; \theta, \alpha) = 1 - \left[1 + \frac{\theta \alpha x(\theta \alpha x + 2\theta^2 + 2\alpha)}{\theta^4 + 2\theta^2\alpha + 2\alpha^2} \right] e^{-\theta x}; x > 0, \theta > 0, \alpha > 0.$$

The detailed studies on various statistical properties, estimation of parameters and applications of NQAD are available in Shanker et al [8].

In this paper an attempt has been made to suggest a two-parameter Aradhana distribution and study its statistical properties, estimation of parameters and applications. The whole paper is divided into eleven sections. Section one is introductory in nature. Section 2 deals with the derivation of pdf and cdf of a two-parameter Aradhana distribution and the behaviors of its pdf

and cdf. Descriptive measures based on moments have been discussed in section three. Reliability properties of the distribution have been studied in section four. Deviation from mean and median, Bonferroni and Lorenz curves, stress-strength reliability have been discussed in sections five, six and seven, respectively. Maximum likelihood estimation and simulation study of the proposed distribution are given in sections eight and nine respectively. Finally, the applications and concluding remarks are presented in section ten and eleven, respectively.

II. A TWO-PARAMETER ARADHANA DISTRIBUTION

A two-parameter Aradhana distribution (ATPAD) can be defined by its pdf and cdf

$$f(x; \theta, \alpha) = \frac{\theta^3}{\theta^2 \alpha^2 + 2\theta\alpha + 2} (\alpha + x)^2 e^{-\theta x} ; x > 0, \theta > 0, \alpha > 0$$

$$F(x; \theta, \alpha) = 1 - \left[1 + \frac{\theta x (\theta x + 2\theta\alpha + 2)}{\theta^2 \alpha^2 + 2\theta\alpha + 2} \right] e^{-\theta x} ; x > 0, \theta > 0, \alpha > 0$$

Aradhana distribution with parameter θ , gamma $(3, \theta)$ distribution and exponential distribution are special cases of ATPAD for $(\alpha = 1)$, $(\alpha = 0)$ and $\alpha \rightarrow \infty$. The behavior of the pdf and the cdf of ATPAD for different values of parameters are shown in figures 1 and 2 respectively.

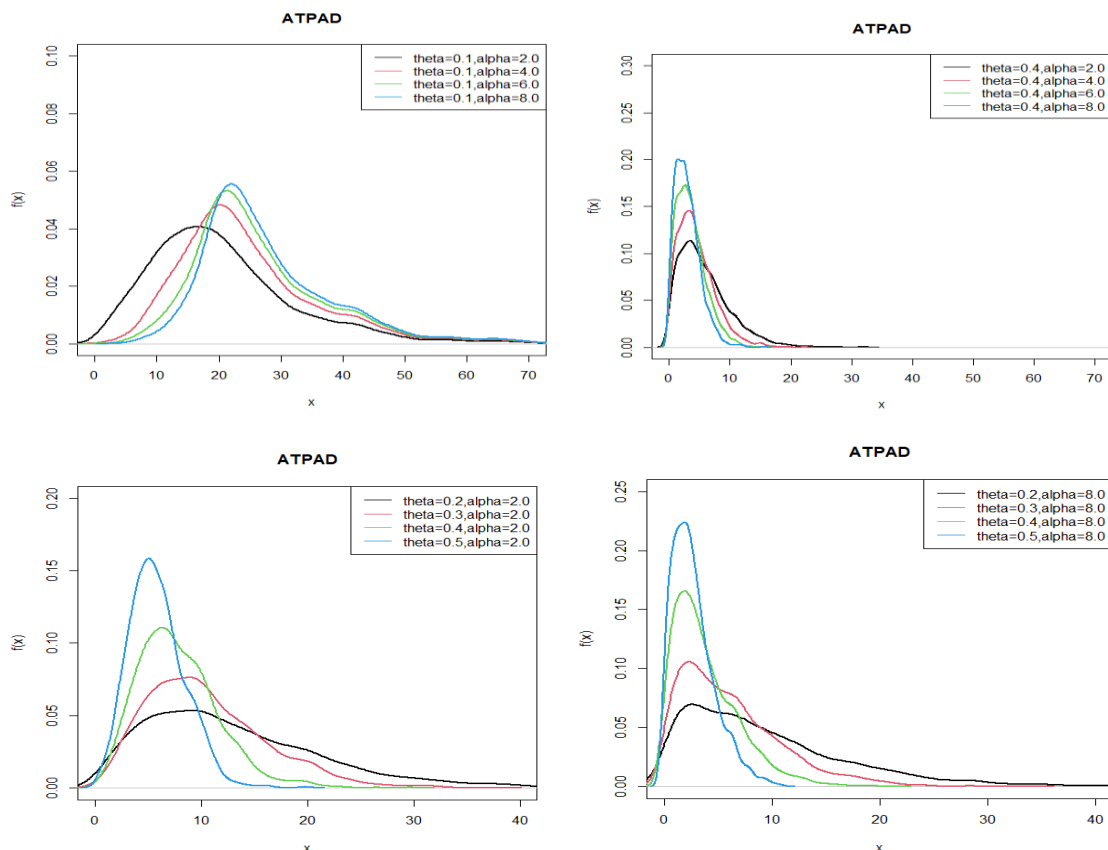


Figure 1: pdf of ATPAD for different values of parameters

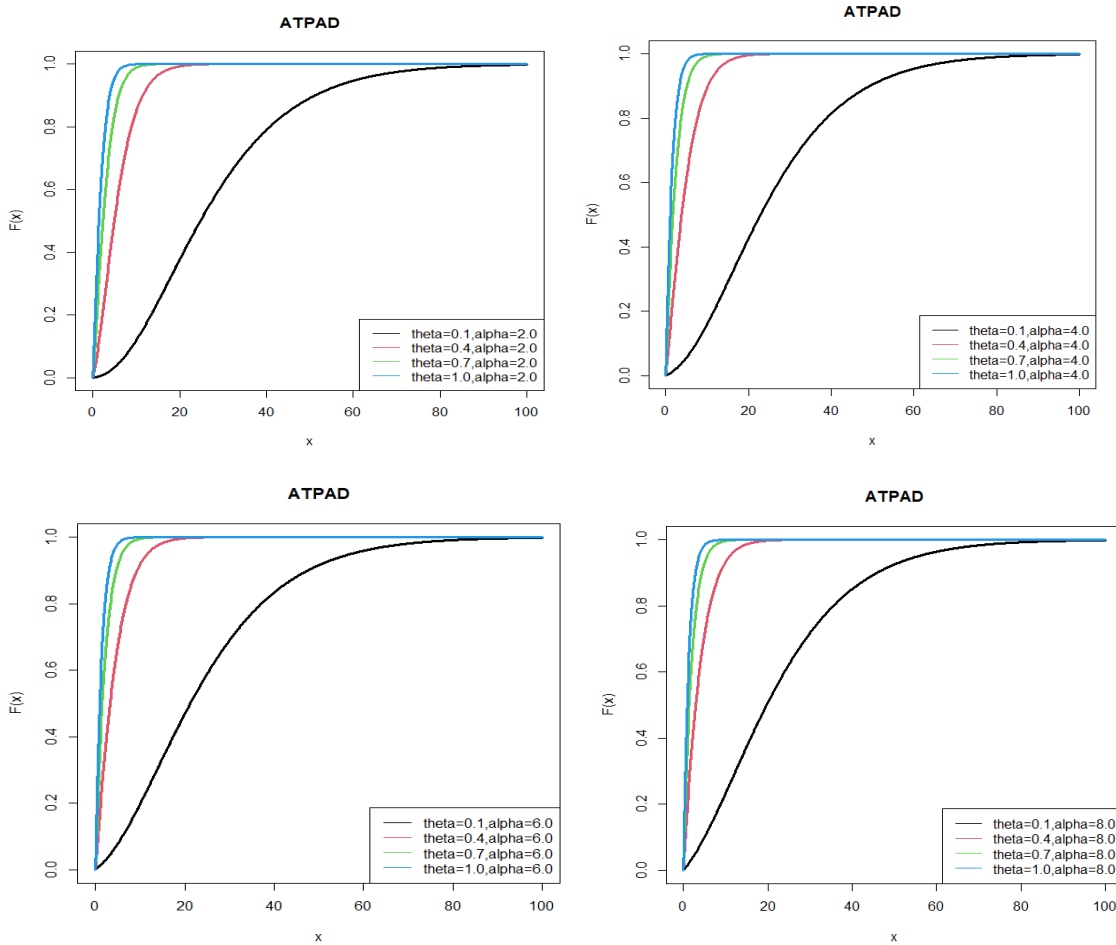


Figure 2: cdf of ATPAD for different values of parameters

III. DESCRIPTIVE MEASURES

The r th moment about origin of ATPAD can be obtained as

$$\begin{aligned} \mu_r' = E(X^r) &= \frac{\theta^3}{\theta^2\alpha^2 + 2\theta\alpha + 2} \int_0^\infty x^r (\alpha^2 + 2\alpha x + x^2) e^{-\theta x} dx \\ &= \frac{r! \{ \theta^2\alpha^2 + 2(r+1)\theta\alpha + (r+1)(r+2) \}}{\theta^r (\theta^2\alpha^2 + 2\theta\alpha + 2)}; r = 1, 2, 3, \dots \end{aligned}$$

Substituting $r = 1, 2, 3, 4$ in the above expression, the first four moments about origin (raw moments) of ATPAD can be obtained as

$$\begin{aligned} \mu_1' &= \frac{\theta^2\alpha^2 + 4\theta\alpha + 6}{\theta(\theta^2\alpha^2 + 2\theta\alpha + 2)}, & \mu_2' &= \frac{2(\theta^2\alpha^2 + 6\theta\alpha + 12)}{\theta^2(\theta^2\alpha^2 + 2\theta\alpha + 2)}, \\ \mu_3' &= \frac{6(\theta^2\alpha^2 + 8\theta\alpha + 20)}{\theta^3(\theta^2\alpha^2 + 2\theta\alpha + 2)}, & \mu_4' &= \frac{24(\theta^2\alpha^2 + 10\theta\alpha + 30)}{\theta^4(\theta^2\alpha^2 + 2\theta\alpha + 2)} \end{aligned}$$

The central moments, using relationship between central moments and raw moments, can thus be obtained as

$$\mu_2 = \frac{\theta^4\alpha^4 + 8\theta^3\alpha^3 + 24\theta^2\alpha^2 + 24\theta\alpha + 12}{\theta^2(\theta^2\alpha^2 + 2\theta\alpha + 2)^2}$$

$$\mu_3 = \frac{2(\theta^6\alpha^6 + 12\theta^5\alpha^5 + 54\theta^4\alpha^4 + 100\theta^3\alpha^3 + 108\theta^2\alpha^2 + 72\theta\alpha + 24)}{\theta^3(\theta^2\alpha^2 + 2\theta\alpha + 2)^3}$$

$$\mu_4 = \frac{3(3\theta^8\alpha^8 + 48\theta^7\alpha^7 + 304\theta^6\alpha^6 + 944\theta^5\alpha^5 + 1816\theta^4\alpha^4 + 2304\theta^3\alpha^3 + 1920\theta^2\alpha^2 + 960\theta\alpha + 240)}{\theta^4(\theta^2\alpha^2 + 2\theta\alpha + 2)^4}$$

Thus, the coefficient of variation (C.V), coefficient of skewness ($\sqrt{\beta_1}$), coefficient of kurtosis (β_2), and index of dispersion (γ) of ATPAD are obtained as

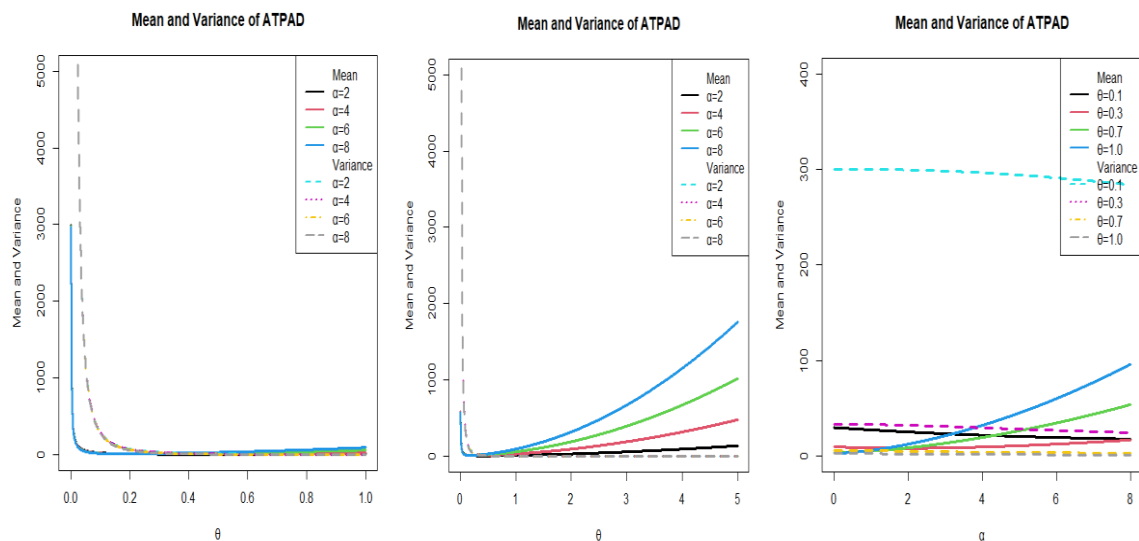
$$C.V. = \frac{\sigma}{\mu_1'} = \frac{\sqrt{\theta^4\alpha^4 + 8\theta^3\alpha^3 + 24\theta^2\alpha^2 + 24\theta\alpha + 12}}{\theta^2\alpha^2 + 4\theta\alpha + 6}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2(\theta^6\alpha^6 + 12\theta^5\alpha^5 + 54\theta^4\alpha^4 + 100\theta^3\alpha^3 + 108\theta^2\alpha^2 + 72\theta\alpha + 24)}{(\theta^4\alpha^4 + 8\theta^3\alpha^3 + 24\theta^2\alpha^2 + 24\theta\alpha + 12)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{33(3\theta^8\alpha^8 + 48\theta^7\alpha^7 + 304\theta^6\alpha^6 + 944\theta^5\alpha^5 + 1816\theta^4\alpha^4 + 2304\theta^3\alpha^3 + 1920\theta^2\alpha^2 + 960\theta\alpha + 240)}{(\theta^4\alpha^4 + 8\theta^3\alpha^3 + 24\theta^2\alpha^2 + 24\theta\alpha + 12)^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{\theta^4\alpha^4 + 8\theta^3\alpha^3 + 24\theta^2\alpha^2 + 24\theta\alpha + 12}{\theta(\theta^2\alpha^2 + 2\theta\alpha + 2)(\theta^2\alpha^2 + 4\theta\alpha + 6)}$$

The graphical relationship between mean and variance of ATPAD to see the over-dispersion, equi-dispersion and under-dispersion are shown in the following figure 3.



3:

Figure 3: Mean and variance of ATPAD

Behavior of coefficient of variation, skewness, kurtosis and index of dispersion of ATPAD shown in figure 4.

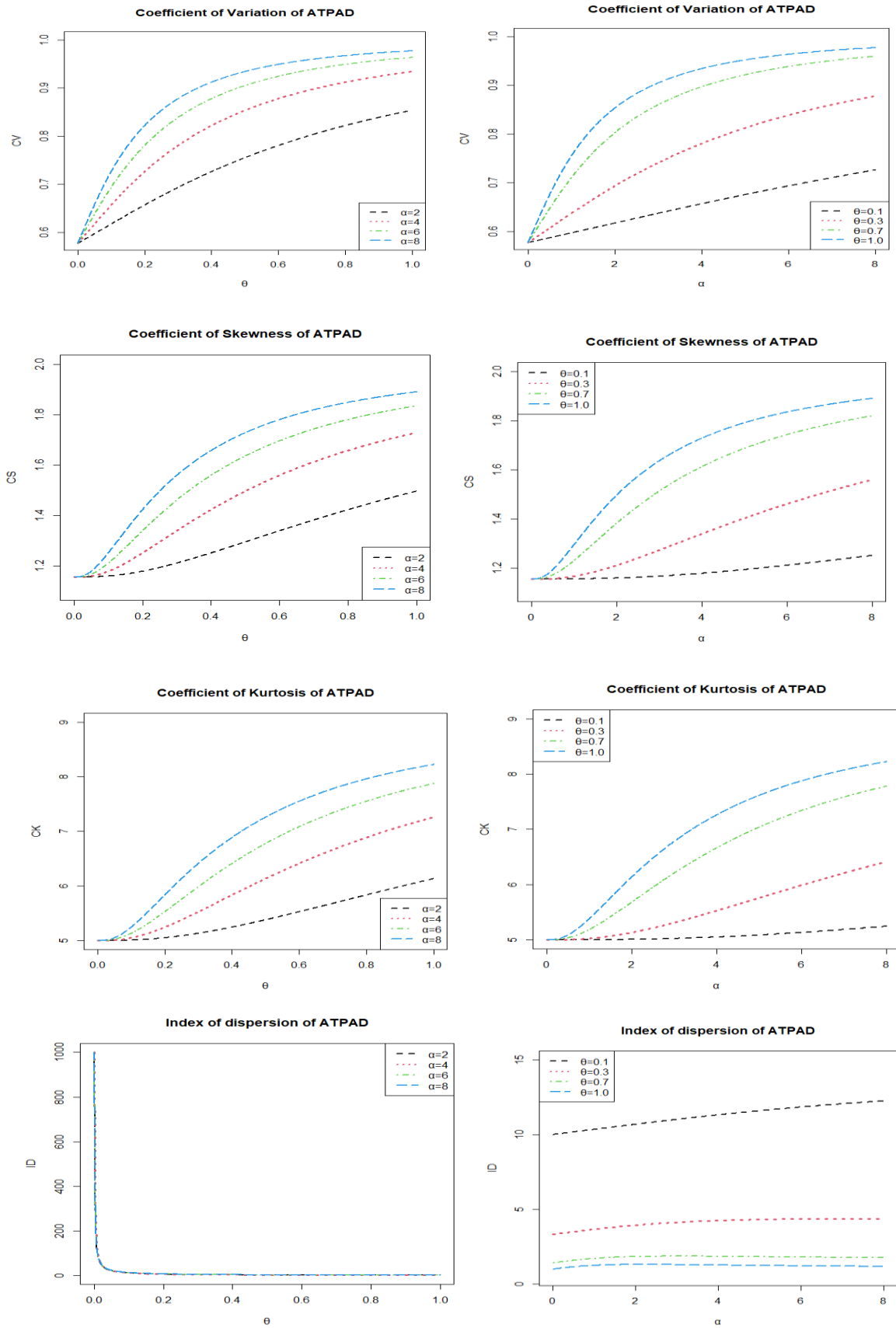


Figure 4: Behaviors of coefficient of variation, skewness, kurtosis and index of dispersion of ATPAD

IV. SOME RELIABILITY PROPERTIES

I. Survival Function

The survival function of ATPAD can be obtained as

$$S(x; \theta, \alpha) = 1 - F(x; \theta, \alpha) = \frac{[\theta x(\theta x + 2\theta\alpha + 2) + (\theta^2\alpha^2 + 2\theta\alpha + 2)]e^{-\theta x}}{\theta^2\alpha^2 + 2\theta\alpha + 2}; x > 0, \theta > 0, \alpha > 0$$

II. Hazard Function and Mean Residual Life Function

The hazard function of ATPAD can be obtained as

$$h(x; \theta, \alpha) = \frac{f(x; \theta, \alpha)}{S(x; \theta, \alpha)} = \frac{\theta^3(\alpha + x)^2}{\theta x(\theta x + 2\theta\alpha + 2) + (\theta^2\alpha^2 + 2\theta\alpha + 2)}$$

The mean residual life function of ATPAD can be obtained as

$$\begin{aligned} m(x, \theta, \alpha) &= \frac{1}{1 - F(x; \theta, \alpha)} \int_0^x [1 - F(t; \theta, \alpha)] dx \\ &= \frac{\theta^2 x^2 + 2(\theta\alpha + 2)\theta x + (\theta^2\alpha^2 + 4\theta\alpha + 6)}{\theta[\theta x(\theta x + 2\theta\alpha + 2) + (\theta^2\alpha^2 + 2\theta\alpha + 2)]} \end{aligned}$$

The graphical representation of hazard function and mean residual life function are presented in the figure 5 and 6 respectively. From the figure 5 it is cleared that all values of the parameters θ and α hazard function is monotonically increasing. From the figure 6 it is cleared that for all values of the parameters θ and α mean residual life function is monotonically decreasing.

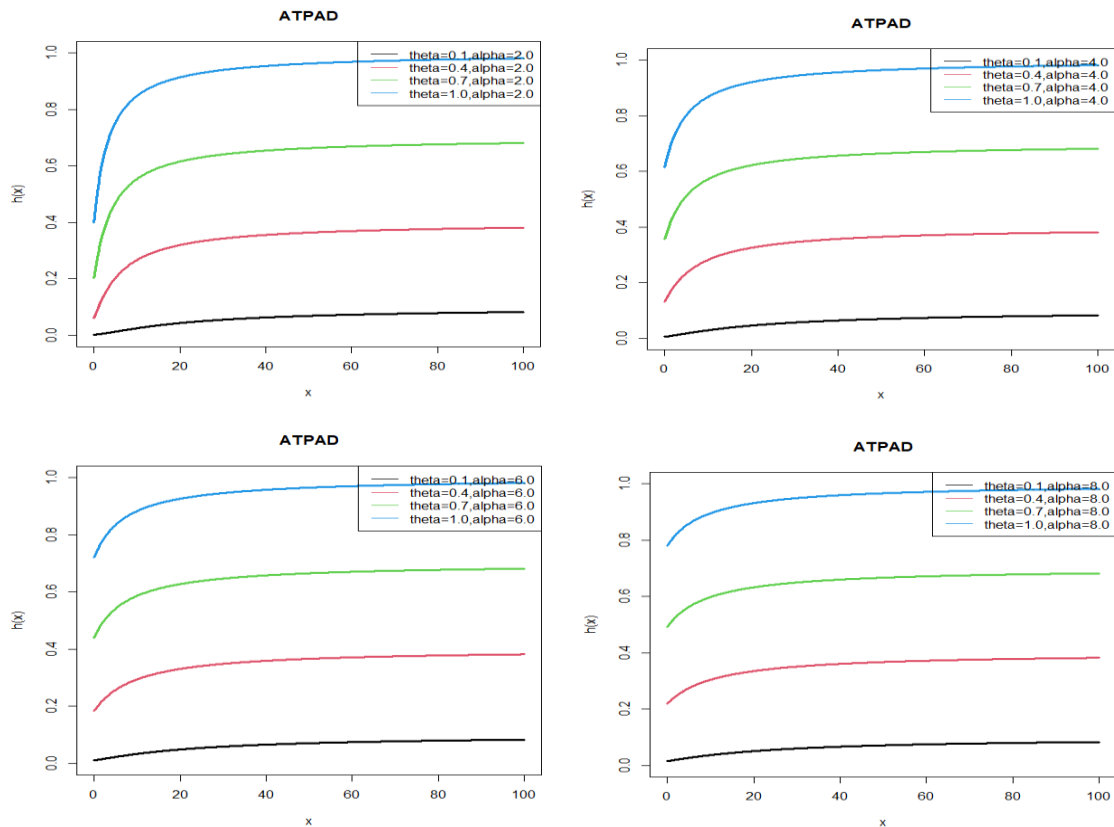


Figure 5: Hazard function of ATPAD for various values of the parameters

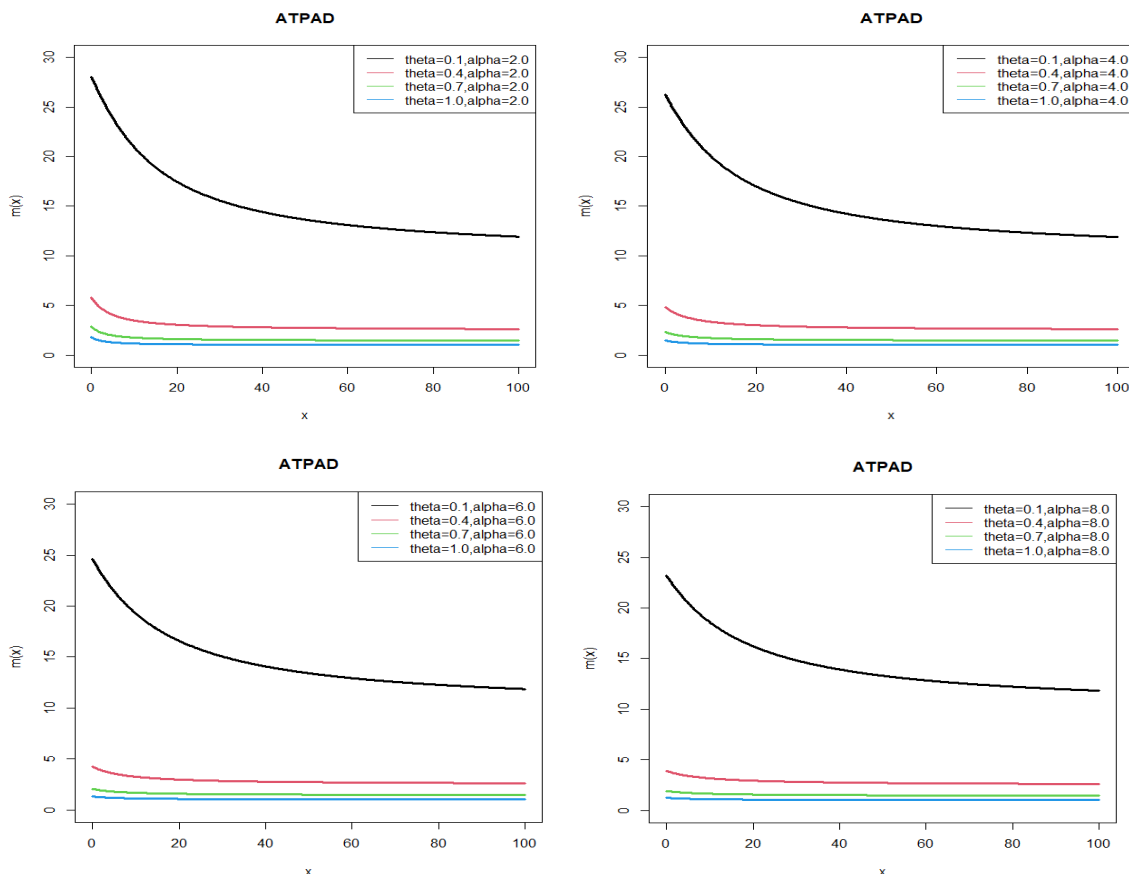


Figure 6: Mean residual life function of ATPAD for various values of the parameters

III. Reverse Hazard Function

Reverse hazard function of ATPAD can be obtained as

$$r_x(x; \theta, \alpha) = \frac{f(x; \theta, \alpha)}{F(x; \theta, \alpha)} = \frac{\theta^3 (\alpha + x)^2 e^{-\theta x}}{(\theta^2 \alpha^2 + 2\theta\alpha + 2) - [\theta x(\theta x + 2\theta\alpha + 2) + (\theta^2 \alpha^2 + 2\theta\alpha + 2)] e^{-\theta x}}$$

IV. Stochastic Ordering

In probability theory and statistics, a stochastic order quantifies the concept of one random variable being bigger than another. A random variable X is said to be smaller than a random variable Y in the:

- i. Stochastic order ($X \leq_{st} Y$) if $F_X(x) \geq F_Y(x)$ for all x
- ii. Hazard rate order ($X \leq_{hr} Y$) if $h_X(x) \geq h_Y(x)$ for all x
- iii. Mean residual life order ($X \leq_{mrl} Y$) if $m_X(x) \geq m_Y(x)$ for all x
- iv. Likelihood ratio order ($X \leq_{lr} Y$) if $\frac{f_X(x)}{f_Y(x)}$ decrease in x

The following results due to Shaked and Shantikumar [9] are well known for establishing stochastic ordering of distributions

$$X <_{lr} Y \Rightarrow X <_{hr} Y \Rightarrow X <_{mrl} Y$$

$$\Downarrow$$

$$X <_{st} Y$$

Theorem: Let $X \sim \text{ATPAD}(\theta_1, \alpha_1)$ and $Y \sim \text{ATPAD}(\theta_2, \alpha_2)$. , If $\alpha_1 \geq \alpha_2$ and $\theta_1 > \theta_2$ or $\theta_1 = \theta_2$ and $\alpha_1 > \alpha_2$ then $X <_{lr} Y$ hence $X <_{hr} Y$, $X <_{mrl} Y$ and $X <_{st} Y$.

Proof: We have

$$\frac{f_X(x)}{f_Y(x)} = \frac{\theta_1^3 (\alpha_2^2 \theta_2^2 + 2\alpha_2 \theta_2 + 2) (\alpha_1 + x)^2}{\theta_2^3 (\alpha_1^2 \theta_1^2 + 2\alpha_1 \theta_1 + 2) (\alpha_2 + x)^2} e^{-(\theta_1 - \theta_2)x}$$

Now

$$\log \frac{f_X(x)}{f_Y(x)} = \log \left[\frac{\theta_1^3 (\alpha_2^2 \theta_2^2 + 2\alpha_2 \theta_2 + 2)}{\theta_2^3 (\alpha_1^2 \theta_1^2 + 2\alpha_1 \theta_1 + 2)} \right] + 2 \log \left(\frac{\alpha_1 + x}{\alpha_2 + x} \right) - (\theta_1 - \theta_2)x$$

Therefore

$$\frac{d}{dx} \left(\log \frac{f_X(x)}{f_Y(x)} \right) = \frac{2(\alpha_2 - \alpha_1)}{(\alpha_1 + x)(\alpha_2 + x)} - (\theta_1 - \theta_2)$$

Thus, if $\alpha_1 \geq \alpha_2$ and $\theta_1 > \theta_2$ or $\theta_1 = \theta_2$ and $\alpha_1 > \alpha_2$, $\frac{d}{dx} \left(\log \frac{f_X(x)}{f_Y(x)} \right) < 0$. This means $X <_{lr} Y$

hence $X <_{hr} Y$, $X <_{mrl} Y$ and $X <_{st} Y$.

V. DEVIATION FROM MEAN AND MEDIAN

The amount of scatter in a population is an evidently measured to some extent by the totality of deviations from the mean and median. These are known as the mean deviation about the mean and mean deviation about median and are defined by

$$\delta_1(x) = 2\mu F(\mu) - 2 \int_0^{\mu} x f(x; \theta, \alpha) dx \quad \text{and} \quad \delta_2(x) = -\mu + 2 \int_M^{\infty} x f(x; \theta, \alpha) dx$$

Thus $\delta_1(x)$ and $\delta_2(x)$ of ATPAD are obtained as

$$\delta_1(x) = \frac{2 \left[\theta^2 \mu^2 + 2\theta^2 \alpha \mu + 4\theta \mu + (\theta^2 \alpha^2 + 4\theta \alpha + 6) \right] e^{-\theta \mu}}{\theta (\theta^2 \alpha^2 + 2\theta \alpha + 2)}$$

$$\delta_2(x) = \frac{2 \left\{ \theta^3 M^3 + (2\theta \alpha + 3)\theta^2 M^2 + (\theta^2 \alpha^2 + 4\theta \alpha + 6)\theta M + (\theta^2 \alpha^2 + 4\theta \alpha + 6) \right\} e^{-\theta M}}{\theta (\theta^2 \alpha^2 + 2\theta \alpha + 2)} - \mu$$

VI. BONFERRONI AND LORENZ CURVES

The Bonferroni and Lorenz curves [10] and Bonferroni and Gini indices have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medicine. The Bonferroni and Lorenz curves are defined as

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x) dx = \frac{1}{p\mu} \left[\int_0^{\infty} x f(x) dx - \int_q^{\infty} x f(x) dx \right] = \frac{1}{p\mu} \left[\mu - \int_q^{\infty} x f(x) dx \right]$$

and

$$L(p) = \frac{1}{\mu} \int_0^q xf(x)dx = \frac{1}{\mu} \left[\int_0^{\infty} xf(x)dx - \int_q^{\infty} xf(x)dx \right] = \frac{1}{\mu} \left[\mu - \int_q^{\infty} xf(x)dx \right]$$

respectively or equivalently.

The Bonferroni and Gini indices are obtained as

$$B = 1 - \int_0^1 B(p)dp \quad \text{and} \quad G = 1 - 2 \int_0^1 L(p)dp, \text{ respectively.}$$

Using the pdf of ATPAD, we get

$$B(p) = \frac{1}{p} \left[1 - \frac{\left\{ \theta^3 q^3 + (2\theta\alpha + 3)\theta^2 q^2 + (\theta^2\alpha^2 + 4\theta\alpha + 6)\theta q + (\theta^2\alpha^2 + 4\theta\alpha + 6) \right\} e^{-\theta q}}{(\theta^2\alpha^2 + 4\theta\alpha + 6)} \right]$$

$$L(p) = 1 - \frac{\left\{ \theta^3 q^3 + (2\theta\alpha + 3)\theta^2 q^2 + (\theta^2\alpha^2 + 4\theta\alpha + 6)\theta q + (\theta^2\alpha^2 + 4\theta\alpha + 6) \right\} e^{-\theta q}}{(\theta^2\alpha^2 + 4\theta\alpha + 6)}$$

Finally, after little algebraic simplification, the Bonferroni and Gini indices of ATPAD are obtained as

$$B = \frac{\left\{ \theta^3 q^3 + (2\theta\alpha + 3)\theta^2 q^2 + (\theta^2\alpha^2 + 4\theta\alpha + 6)\theta q + (\theta^2\alpha^2 + 4\theta\alpha + 6) \right\} e^{-\theta q}}{(\theta^2\alpha^2 + 4\theta\alpha + 6)}$$

$$G = \frac{2 \left\{ \theta^3 q^3 + (2\theta\alpha + 3)\theta^2 q^2 + (\theta^2\alpha^2 + 4\theta\alpha + 6)\theta q + (\theta^2\alpha^2 + 4\theta\alpha + 6) \right\} e^{-\theta q}}{(\theta^2\alpha^2 + 4\theta\alpha + 6)} - 1$$

VII. STRESS-STRENGTH RELIABILITY

The stress-strength reliability of a component illustrates the life of the component which has random strength X that is subjected to a random stress Y . When the stress of the component Y applied to it exceeds the strength of the component X , the component fails instantly, and the component will function satisfactorily until $X > Y$. Therefore, $R = P(Y < X)$ is a measure of the component reliability and is known as stress-strength reliability in statistical literature. It has extensive applications in almost all areas of knowledge especially in engineering such as structures, deterioration of rocket motors, static fatigue of ceramic components, aging of concrete pressure vessels, etc.

$$R = P(Y < X) = \int_0^{\infty} P(Y < X | X = x) f_X(x) dx$$

$$= 1 - \frac{\theta_1^3 \left[\begin{array}{l} 24\theta_2^2 + 48\{\theta_2^2\alpha_1 + \theta_2(\theta_2\alpha_2 + 1)\}(\theta_1 + \theta_2) \\ + 2\{\theta_2^2\alpha_1 + 4\theta_2(\theta_2\alpha_2 + 1)\alpha_1 + (\theta_2^2\alpha_2^2 + 2\theta_2\alpha_2 + 2)\}(\theta_1 + \theta_2)^2 \\ + 2\{\theta_2(\theta_2\alpha_2 + 1)\alpha_1^2 + \alpha_1(\theta_2^2\alpha_2^2 + 2\theta_2\alpha_2 + 2)\}(\theta_1 + \theta_2)^{3/2} \\ + \alpha_1^2(\theta_2^2\alpha_2^2 + 2\theta_2\alpha_2 + 2)(\theta_1 + \theta_2)^4 \end{array} \right]}{(\theta_1^2\alpha_1^2 + 2\theta_1\alpha_1 + 2)(\theta_2^2\alpha_2^2 + 2\theta_2\alpha_2 + 2)(\theta_1 + \theta_2)^5}$$

VIII. ESTIMATION AND INFERENCE

Let (x_1, x_2, \dots, x_n) be a random sample from ATPAD (θ, α) , the likelihood function L and the log-likelihood function, $\log L$ are given by

$$L = \left(\frac{\theta^3}{\theta^2 \alpha^2 + 2\theta\alpha + 2} \right)^2 \prod_{i=1}^n (\alpha + x_i)^2 e^{-n\theta\bar{x}}$$

$$\log L = 3n \log \theta - n \log(\theta^2 \alpha^2 + 2\theta\alpha + 2) + 2 \sum_{i=1}^n \log(\alpha + x_i) - n\theta\bar{x}$$

The maximum likelihood estimates (MLEs) $\hat{\theta}$ and $\hat{\alpha}$ of θ and α are then the solutions of the following non-linear equations

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{\theta} - \frac{2n\alpha(\theta\alpha + 1)}{\theta^2 \alpha^2 + 2\theta\alpha + 2} - n\bar{x} = 0$$

$$\frac{\partial \log L}{\partial \alpha} = -\frac{2n\theta(\theta\alpha + 1)}{\theta^2 \alpha^2 + 2\theta\alpha + 2} + 2 \sum_{i=1}^n \frac{1}{\alpha + x_i} = 0$$

These two natural log likelihood equations do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. We have

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{3n}{\theta^2} + \frac{2n\alpha^3 \theta(\theta\alpha + 2)}{(\theta^2 \alpha^2 + 2\theta\alpha + 2)^2}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{2n\alpha\theta^3(\theta\alpha + 2)}{(\theta^2 \alpha^2 + 2\theta\alpha + 2)^2}$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = \frac{-2n(\theta^2 \alpha^2 + 4\theta\alpha + 2)}{(\theta^2 \alpha^2 + 2\theta\alpha + 2)^2} = \frac{\partial^2 \log L}{\partial \alpha \partial \theta}$$

The following equations can be solved for MLEs $\hat{\theta}$ and $\hat{\alpha}$ of θ and α of ATPAD

$$\begin{pmatrix} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \log L}{\partial \alpha \partial \theta} & \frac{\partial^2 \log L}{\partial \alpha^2} \end{pmatrix}_{\substack{\hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0}} \begin{pmatrix} \hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0 \end{pmatrix} = \begin{pmatrix} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \alpha} \end{pmatrix}_{\substack{\hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0}}$$

IX. THE SIMULATION STUDY

In this section, we carried out simulation study to examine the performance of maximum likelihood estimators of the ATPAD. We examined the mean estimates, biases (B), mean square errors (MSEs) and variances of the MLEs. The mean, bias, MSE and variance are computed using the formulae

$$Mean = \frac{1}{n} \sum_{i=1}^n \hat{H}_i, B = \frac{1}{n} \sum_{i=1}^n (\hat{H}_i - H), MSE = \frac{1}{n} \sum_{i=1}^n (\hat{H}_i - H)^2, Variance = MSE - B^2$$

where $H = \theta, \alpha$ and $\hat{H} = \hat{\theta}_i, \hat{\alpha}_i$.

The simulation results for different parameter values of ATPAD are presented in tables 1 and 2 respectively. The steps for simulation study are as follows:

- a. Data is generated using the acceptance-rejection method of simulation. The acceptance-rejection method is a commonly used approach in simulation studies to generate random samples from a target distribution when inverse transform method of simulation is not feasible or efficient. Acceptance rejection method for generating random samples from the

ATPAD consists of following steps.

- i. Generate a random variable Y distributed as exponential (θ)
- ii. Generate U distributed as Uniform (0,1)
- iii. If $U \leq \frac{f(y)}{M g(y)}$, then set $x = y$ ("accept the sample"); otherwise ("reject the sample")

and if reject then repeat the process: step (i-iii) until getting the required samples.

Where M is a constant.

- b. The sample sizes are taken as $n = 50, 100, 150, 200$
- c. The parameter values are set as values $\theta = 0.2, \alpha = 1.8$ and $\theta = 0.2, \alpha = 4.0$
- d. Each sample size is replicated 10000 times

The results obtained in Tables 1 and 2 show that as the sample size increases, biases, MSEs and variances of the MLEs of the parameters become smaller respectively. This result is in line with the first-order asymptotic theory.

Table-1: The Mean values, Biases, MSEs and Variances of ATPAD for parameter $\theta = 0.2, \alpha = 1.8$

Parameters	Sample Size	Mean	Bias	MSE	Variance
$\hat{\theta}$	20	0.20148	0.00148	0.000004212	0.000002011
	40	0.20094	0.00094	0.000002244	0.000001354
	50	0.20083	0.00083	0.000002477	0.000002009
	100	0.20081	0.00081	0.000002422	0.000001785
	150	0.20079	0.00079	0.000002254	0.000001591
	200	0.20068	0.00068	0.000002251	0.000001557
$\hat{\alpha}$	20	1.72273	-0.07726	0.127852500	0.121882500
	40	1.76513	-0.03486	0.062430860	0.061215480
	50	1.77232	-0.02767	0.051223530	0.050457760
	100	1.78933	-0.01066	0.026087550	0.025973800
	150	1.79298	-0.00701	0.017392750	0.017343480
	200	1.78781	-0.00121	0.015850450	0.025702030

Table-2: The Mean values, Biases, MSEs and Variances of ATPAD for parameter $\theta = 0.2, \alpha = 4.0$

Parameters	Sample Size	Mean	Bias	MSE	Variance
$\hat{\theta}$	20	0.20138	0.001380	0.00000342	0.000001517
	40	0.20110	0.001107	0.00000242	0.000001195
	50	0.20107	0.001077	0.00000224	0.000001082
	100	0.20107	0.001074	0.00000223	0.000001084
	150	0.20096	0.000963	0.00000206	0.000001139
	200	0.20074	0.000742	0.00000173	0.000001799
$\hat{\alpha}$	20	4.01832	0.018320	0.00659461	0.00625895
	40	4.00915	0.009156	0.00329730	0.00321346
	50	4.00726	0.007266	0.00263803	0.00258523
	100	4.00484	0.004849	0.00175868	0.00173517
	150	4.00364	0.003640	0.00141901	0.00140576
	200	4.00363	0.003630	0.00131901	0.00130576

X. APPLICATIONS

The following real lifetime datasets have been considered for testing the goodness of fit of ATPAD over the other two-parameter lifetime distributions. The goodness of fit based on K-S statistic, fitted plots of considered distributions for the datasets, p-p plots of considered distributions for the datasets and total time in test (TTT) plots for the datasets and the ATPAD confirm that among all considered distributions, ATPAD provides much closure fit.

Data set-1: This censored tri-modal data contains 30 items that is tested when test is stopped after 20-th failure. The following data discussed by Murthy et al [11] and the values are:

0.0014, 0.0623, 1.3826, 2.0130, 2.5274, 2.8221, 3.1544, 4.9835, 5.5462, 5.8196, 5.8714, 7.4710, 7.5080, 7.6667, 8.6122, 9.0442, 9.1153, 9.6477, 10.1547, 10.7582.

Description of the data set-1

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.0014	2.7484	5.8455	5.7081	8.7202	10.7582

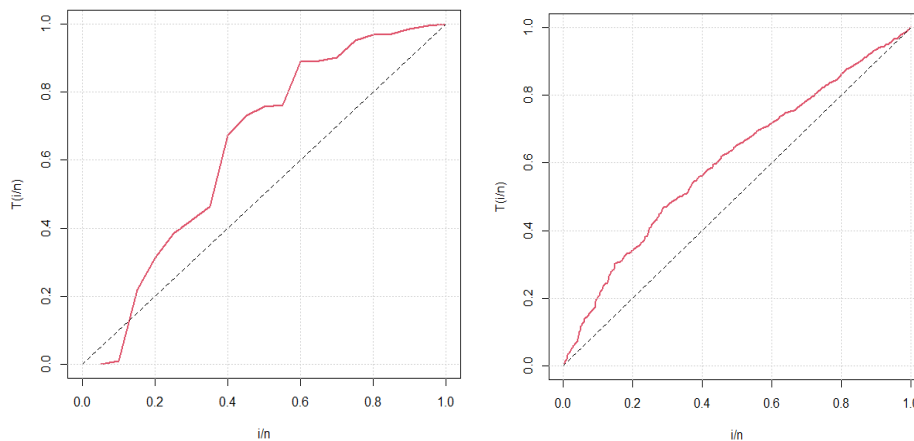


Figure 7: TTT- plot of the observed dataset 1 and simulated data of ATPAD respectively

Data set-2: The following skewed to right, a complete set of data, discussed by Murthy et al [11], and reports the failure time of 20 electric bulbs and the observations are:

1.32, 12.37, 6.56, 5.05, 11.58, 10.56, 21.82, 3.60, 1.33, 12.62, 5.36, 7.71, 3.53, 19.61, 36.63, 0.39, 21.35, 7.22, 12.42, 8.92.

Description of the data set-2

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.390	4.688	8.315	10.498	12.470	36.630

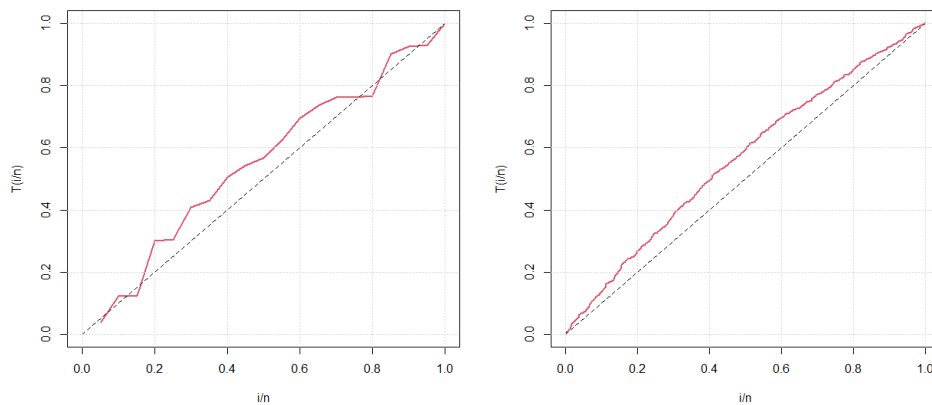


Figure 8: TTT- plot of the observed dataset 2 and simulated data of ATPAD respectively

In order to compare lifetime distributions, values of $-2\log L$, Akaike Information Criterion (AIC), Kolmogorov-Smirnov Statistics (K-S) and the corresponding probability value (p-value) for the above data set has been computed. The formulae for computing AIC and K-S are as follows:

$$AIC = -2\log L + 2k, \quad D = \sup_x |F_n(x) - F_0(x)|$$

where, k = number of parameters, n = sample size, $F_n(x)$ = empirical cdf of considered distribution and $F_0(x)$ = cdf of considered distribution

The distribution corresponding to the lower values of $-2\log L$, AIC and K-S Statistics is the best fit distribution. The MLEs of parameters of the considered distributions along with their standard error, $-2\log L$, AIC, K-S and p-value of the considered distributions for datasets 1 and 2 are presented in tables 3 and 4 respectively.

Table 3: ML estimates, Standard errors, $-2\log L$, AIC, K-S and p-value of the considered distributions for the dataset-1

Distribution	ML estimates	$-2\log L$	AIC	K-S	p-value
	$\hat{\theta}$ $SE(\hat{\theta})$ $\hat{\alpha}$ $SE(\hat{\alpha})$				
ATPAD	0.3896 (0.0806) 2.4576 (1.9436)	105.8912	109.8912	0.1837	0.5105
GAD	0.3896 (0.0806) 0.4068 (0.3217)	105.8912	109.8912	0.1949	0.4341
QAD	0.3896 (0.0806) 0.9577 (0.6199)	105.8912	109.8912	0.1971	0.3908
NQAD	0.3896 (0.0806) 0.1585 (0.1518)	105.8912	109.8912	0.1930	0.4696
PAD	0.5935 (0.1558) 0.8366 (0.1394)	106.0269	110.0269	0.1915	0.4687
WAD	0.4535 (0.1142) 0.0100 (0.5170)	107.4734	111.4734	0.1902	0.4346
GD	0.1513 (0.0552) 0.8637 (0.2373)	109.3792	113.3792	0.2530	0.1624
WD	0.1469 (0.0721) 1.0892 (0.2209)	109.5036	113.5036	0.9000	0.0000

Table 4: ML estimates, Standard errors, $-2\log L$, AIC, K-S and p-value of the considered distributions for the dataset-2

Distributions	ML estimates $\hat{\theta}$ $SE(\hat{\theta})$ $\hat{\alpha}$ $SE(\hat{\alpha})$	$-2\log L$	AIC	K-S	p-value
ATPAD	0.1866 (0.0565) 8.1128(9.4833)	132.9421	136.9421	0.1220	0.9276
GAD	0.1867 (0.0565) 0.1232 (0.1440)	132.9421	136.9421	0.1355	0.8570
QAD	0.1866 (0.0565) 1.5147 (1.3803)	132.9421	136.9421	0.1307	0.8972
NQAD	0.1869 (0.0558) 0.0231 (0.0329)	132.9421	136.9421	0.1459	0.8073
PAD	0.4421 (0.1225) 0.7755 (0.1097)	133.1672	137.1672	0.1343	0.8711
WAD	0.2625 (0.0710) 0.0100 (0.6242)	137.0825	141.0825	0.1425	0.8289
GD	0.1272(0.0438) 1.3361 (0.3811)	133.0916	137.0916	0.1340	0.8729
WD	0.1000 (0.0776) 1.0150 (0.2550)	134.0518	138.0518	0.9830	0.000

From Table-3 and 4 we observed that the ATPAD has the same $-2\log L$, AIC values but least K-S values as compared to GAD (Generalized Aradhana Distribution) of Daniel and Shanker [12] and , QAD (Quasi Aradhana Distribution), NQAD (New Quasi Aradhana distribution) and has the least $-2\log L$, AIC, K-S values as compared to PAD (Power Aradhana distribution) of Shanker and Shukla [13], WAD (Weighted Aradhana distribution) by Ganaie et al [14] and subsequently critical study done by Shanker et al [15], GD (gamma Distribution) and WD (Weibull distribution) by Weibull [16].

Hence, we may conclude that ATPAD provides the better fit than GAD, QAD, NQAD, PAD, WAD, GD and WD. Further, it is also clear from the fitted plot and P-P plot of two dataset of considered distributions in figure 9, 10 and 11, that ATPAD provides a much better fit over GAD, QAD, NQAD, PAD, WAD, GD and WD.

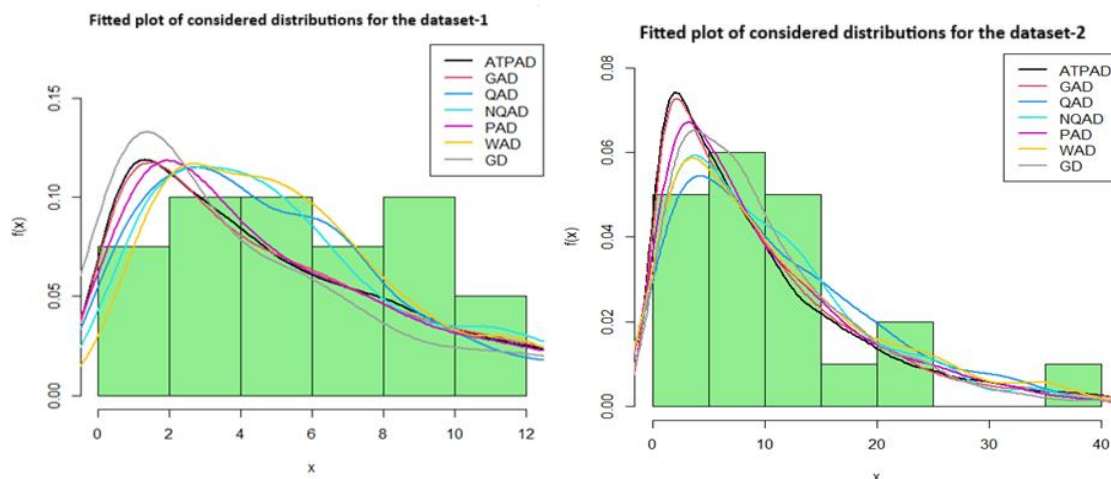


Figure 9: Fitted plot of the considered distribution for the data set-1 and data set-2

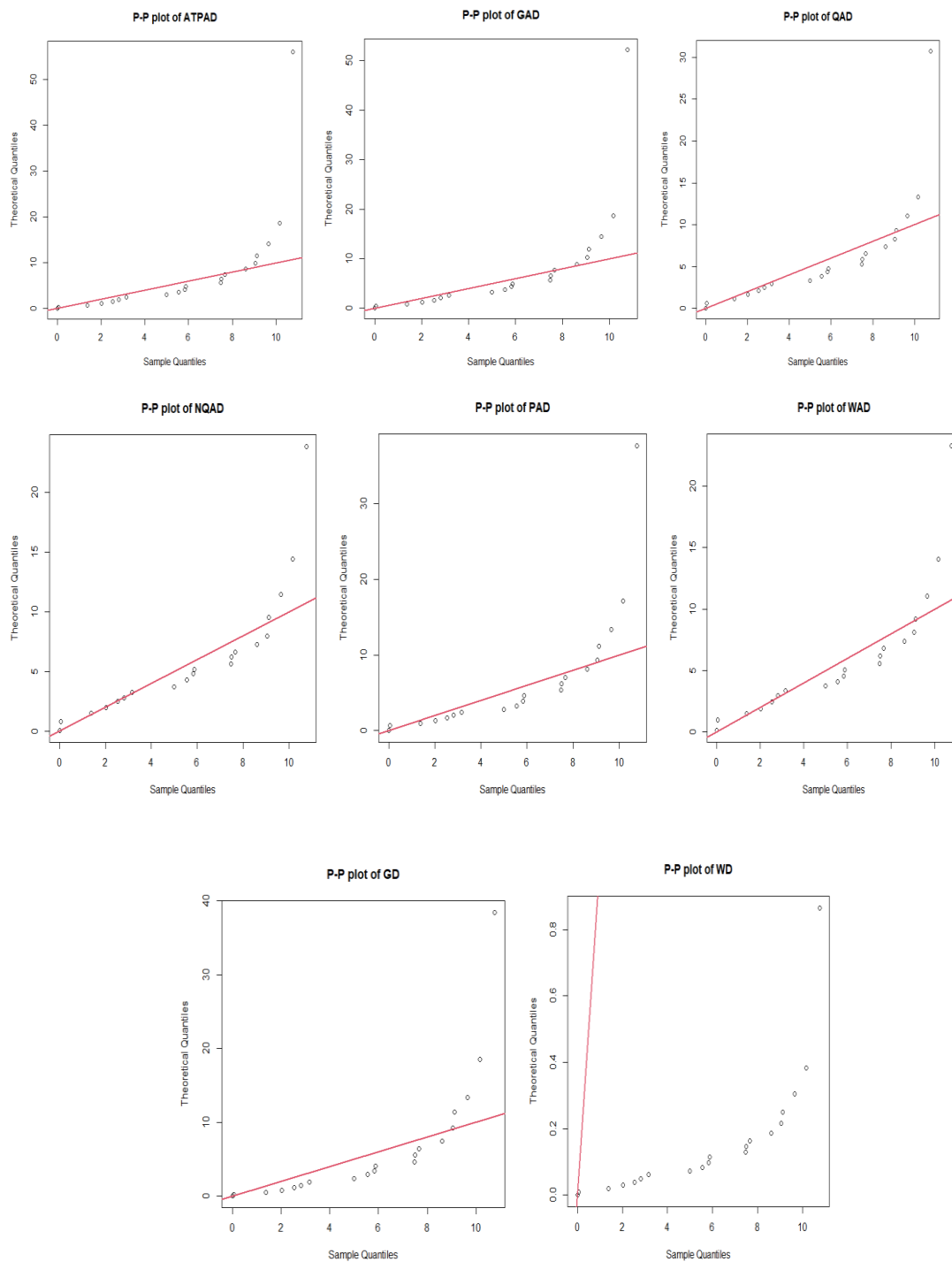


Figure 10: P-P plot for considered distributions of the data set-1

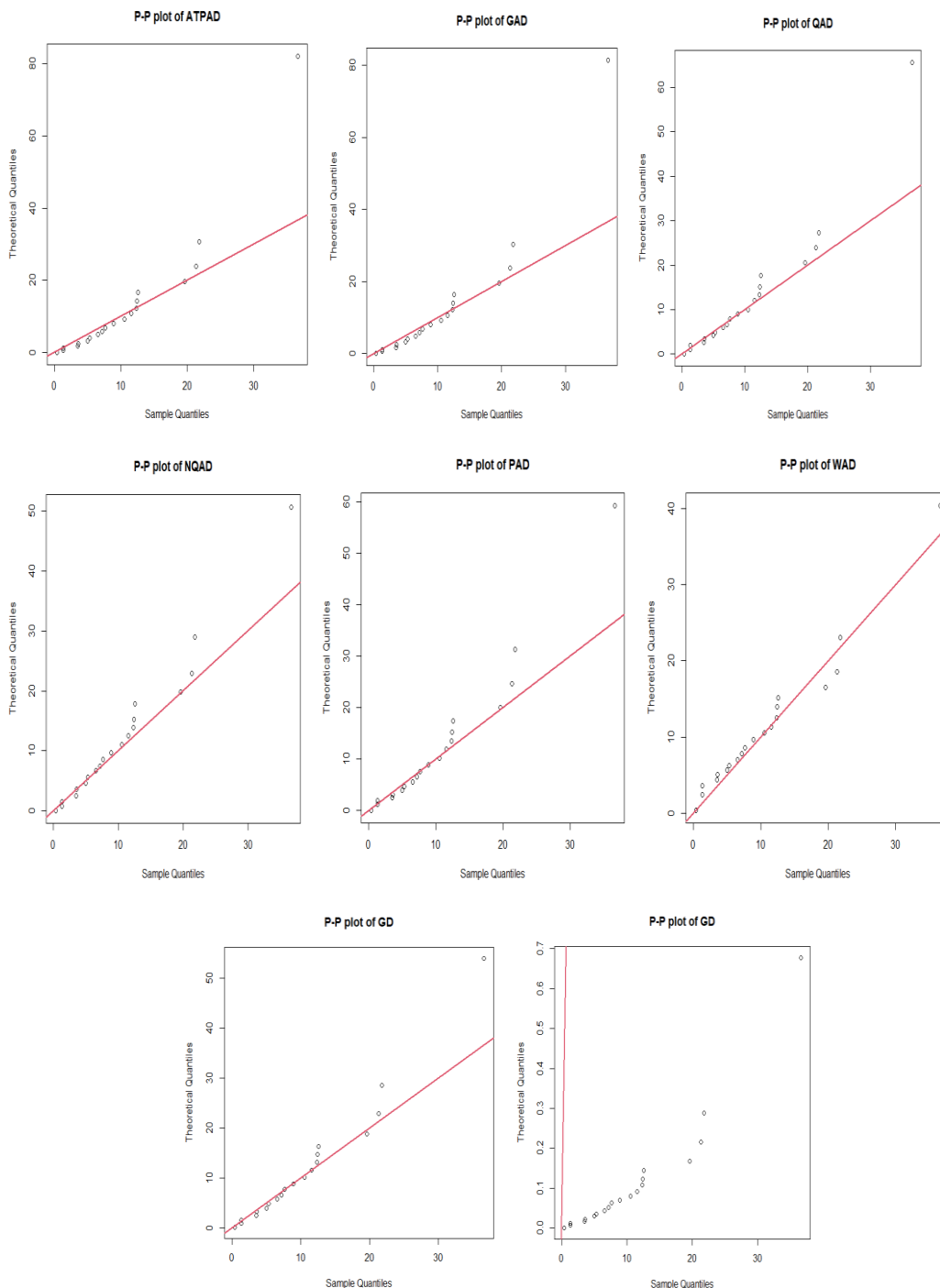


Figure 11: P-P plot for considered distributions of the data set-2

XI. CONCLUSION AND FUTURE WORKS

In this paper, a two-parameter Aradhana distribution which includes Aradhana distribution, gamma distribution and exponential distribution are proposed. Its moments and statistical properties including survival function, hazard function, mean residual life function, reverse hazard function, stochastic ordering have been discussed. Deviations from mean and median, Bonferroni and Lorenz curve and their indices, stress strength reliability have also been discussed. The parameters of this distribution have been estimated using maximum likelihood estimation. To know the performance of maximum likelihood estimates of parameters, a simulation study has been presented. Finally, two examples of real lifetime datasets have been considered for

applications and compared with GAD, QAD, NQAD, PAD, WAD, GD and WD. It has been found that ATPAD provide the best fit than the GAD, QAD, NQAD, PAD, WAD, GD and WD. As this is a new two-parameter lifetime distribution, it has the possibility of extension by adding more parameter in the distribution to see its performance over other lifetime distributions of same parameter. Further, Bayesian method of estimation and ranked set sampling method of estimation can also be considered in future to see the efficiency of these two methods of estimation over the classical maximum likelihood estimation.

References

- [1] Shanker, R. (2016). Aradhana distribution and Its applications. *International Journal of Statistics and Applications*, 6(1): 23-34.
- [2] Lindley, D.V. (1958). Fiducial distributions and Bayes' theorem, *Journal of the Royal Statistical Society, Series B*, 20(1), 102-107.
- [3] Shanker, R. (2015). Shanker distribution and its applications. *International Journal of Statistics and Application*, 5(6): 338-348.
- [4] Shanker, R. (2015). Akash Distribution and its application. *International Journal of Probability and Statistics*, 4(3): 65-75.
- [5] Shanker, R. Shukla, K.K., and Shanker, R. (2018). A Quasi Aradhana Distribution with properties and Application. *International Journal of Statistics and Systems*, 13(1): 61 – 80.
- [6] Anthony, M. and Elangovan, R. (2020). Length biased quasi Aradhana distribution and its applications to survival data, *Journal of XIDIAN University*, 14(10): 893 – 908.
- [7] Anthony, M. and Elangovan, R. (2023). A new generalization of quasi Aradhana distribution with properties and applications, conference proceedings on, *New Trends on Stochastic Processes, SQC & Reliability Using R Programming*, 266 – 278.
- [8] Shanker, R., Soni, N.K., Shanker, R., and Prodhani, H.R. (2023). A New Quasi Aradhana Distribution with Properties and Applications. *Journal of XIDIAN University*, 17(11): 472 – 493.
- [9] Shaked, M. and Shanthikumar, J. Stochastic Orders and Their Applications. Academic Press, New York, 1994.
- [10] Bonferroni, C.E. (1930). Elementi di Statistica generale, Seeber, Firenze
- [11] Murthy, D.N.P., Xie M. and Jiang R. Weibull models, John Wiley & Sons Inc., Hoboken 2004.
- [12] Daniel, W., and Shanker, R (2018). A generalized Aradhana distribution with Properties and applications, *Biometrics & Biostatistics International Journal*, 7(4): 374 – 385.
- [13] Shanker, R. and Shukla, K. K. (2018). A two-parameter Power Aradhana distribution with properties and Application, *Indian Journal of Industrial and Applied Mathematics*, 9(2):210 – 220.
- [14] Ganaie R.A., Rajgopalan V. and Rather A. A. (2019). Weighted Aradhana distribution-properties and applications. *Journal of Information and Computational Science*, 9(8):392-406.
- [15] Shanker, R., Shukla, K.K. and Shanker, R. (2022). A note on weighted Aradhana distribution with an application. *Biometrics & and Biostatistics International Journal*, 11(1): 22-26.
- [16] Weibull, W. (1951). A statistical distribution function of wide applicability, *ASME Journal of Applied Mechanics*, 293-297.

STATISTICAL DESIGN OF CONDITIONAL REPETITIVE GROUP SAMPLING PLAN BASED ON TRUNCATED LIFE TEST FOR PERCENTILE LIFETIME USING EXPONENTIATED GENERALIZED FRECHET DISTRIBUTION

S. JAYALAKSHMI¹, S. VIJILAMERY²

¹ Assistant Professor, Department of Statistics, Bharathiar University,
Coimbatore, India- 641046

² Research Scholar, Department of Statistics, Bharathiar University,
Coimbatore, India- 641046

¹statjayalakshmi16@gmail.com,² vijilamerysrsj@gmail.com

Abstract

Reliability Acceptance sampling plan is used to assess whether to accept or reject a product depending on its lifetime. An inspection carried out for the purpose of determining if lifetime inspections are performing properly can be tested by submitting a truncated lifetime test. In this paper describes a new approach on Conditional Repetitive Group Sampling Plan based on Truncated life test is proposed and the lifetime follows an Exponentiated Generalized Frechet Distribution. For each consumer risk, it is determined whether minimum sample sizes are required to assert a percentile life. It is calculated that the operating characteristic function values of the sampling plans as well as the producer's risk ratio corresponding to the sampling plans. The results are illustrated with numerical examples and a real-world data set is used to demonstrate the impact and performance of the suggested acceptance sampling plans.

Keywords: Conditional Repetitive Group Sampling Plan, Exponentiated Generalized Frechet Distribution, Producer's risk, percentile life.

1. INTRODUCTION

Quality items always need greater attention and must maintain the manufacturers' standards in the highly competitive worldwide market. In order to monitor the product quality, which is among the most important operations in industries, outgoing or incoming products are checked thoroughly. A method of quality control which uses statistical methods is known as statistical quality control. It may be categorized into process control and product control. Acceptance sampling is one of the statistical methods of statistical quality control. It is used to determine whether to accept or reject a decision based on the inspection of a sample of items from the lot. Reliability Acceptance sampling plan is used to determine to accept or reject based on lifetime of product. Truncated life test is adopted at which the test will terminated at a certain point of time in the sense that observing the lifetime of the products until it fails is not possible. It helps to minimize the inspection of time and cost. In acceptance sampling based on truncated life tests have the following assumptions: (i) the units are destructible or are degraded after the life test, and (ii) there are several distributions that model the product life reasonably well. The

purpose of truncated life test used to save the inspection time and cost. A truncated life test may be conducted to determine the smallest sample size to ensure a certain percentile life of products when the life test is terminated at a preassigned time t_0 , and the number of failures observed does not exceed a given acceptance number. For example, some production companies that routinely inspect and sample their products based on their production stages before they are released to market. Accepted lots are sent on for further processing, while rejected lots are either reworked or scrapped. The sampling plan process are carried out throughout the production process to ensure that the product meets the desired specifications. Regular audits are also conducted to make sure all quality standards are consistently met. Finally, customer feedback is used to improve the production process.

An Exponentiated Generalized Frechet Distribution is used in this paper to calculate a Conditional Repetitive Group Sampling plan based on a truncated life test. As part of acceptance sampling plans, it is important to find the minimum sample size, operating characteristics function, and producer's risk that exceeds the specified life of the product. Using the reliable life criterion as a basis for selecting the parameters of the plan, a methodological procedure is proposed for selecting the parameters of the plan with the desired discrimination protecting the interests of the producer and the consumer in terms of the acceptable reliable life and the unacceptable reliable life. As a measure of discrimination, the operating ratio is used to design the proposed reliability sampling plan.

In Reliability Acceptance Sampling Plan based on truncated life test using various distribution of review of literature given below: Epstein [5] has developed truncate life test in the Exponential cases. Goode and kao [6] have developed Sampling Plans Based on the Weibull Distribution. Gupta [7] has designed the life test sampling plans for Normal and Lognormal Distributions. Rosaiah and kantam [11] have progressed an acceptance sampling based on the Inverse Rayleigh Distribution. Rosaiah et.al [12] have described the source of reliability of test plans for Exponentiated Log-Logistic Distribution. Balakrishnan et.al [4] have designed an acceptance sampling plan from truncated life tests based on the Generalized Birnbaum-Saunders Distribution. Aslam et.al [3] have designed acceptance sampling plan using generalized exponential Distribution. Pradeepa Veerakumari et.al [10] have developed for Exponentiated Rayleigh Distribution. Kaviyarasu et.al [9] have developed for Weibull-Poisson Distribution.

Robert Sherman [14] has introduced an inspection procedure and developed Repetitive Group Sampling (RGS) plans. Shankar and Mohapatra [13] have introduced the GERT Analysis of Conditional Repetitive Group Sampling Plan. Anburajan and Ramaswamy [2] have developed a Conditional Repetitive Group Sampling plan based on Truncated life test using Various distributions. Jayalakshmi and Kavyamani [8] have designed the Quick Switching conditional Repetitive Group Sampling Plan through quality decision Region.

Abd-Elfattah et.al [1] have developed a statistical distribution of Exponentiated Generalized Frechet Distribution. The real time application of Exponentiated Generalized Frechet Distributions are reliability studies, hydrology, finance and so on.

2. EXPONENTIATED GENERALIZED FRECHET DISTRIBUTION

The Cumulative Distribution Function of Exponentiated Generalized Frechet distribution is

$$F(x) = \left[1 - \left(1 - \exp - \left\{ \left(\frac{\sigma}{x} \right)^\lambda \right\} \right)^\alpha \right]^\beta \quad (1)$$

The Probability density function of Exponentiated Generalized Frechet distribution is

$$f(x) = \alpha\beta\lambda\sigma^\lambda x^{-(\lambda-1)} \exp \left[- \left(\frac{\sigma}{x} \right)^\lambda \right] \left\{ 1 - \exp \left[- \left(\frac{\sigma}{x} \right)^\lambda \right] \right\}^{\alpha-1} \left\{ 1 - \left[1 - \exp \left(- \left(\frac{\sigma}{x} \right)^\lambda \right) \right]^\alpha \right\}^{\beta-1} \quad (2)$$

Where, α, β, λ are Shape parameters and σ be a Scale parameter.

$$Pr(T \leq t_q) = q$$

$$t_q = \sigma \left[-\ln \left(1 - \left(1 - (1-u)^{\frac{1}{\beta}} \right)^{\frac{1}{\alpha}} \right) \right]^{-\frac{1}{\lambda}} \tag{3}$$

$$\varphi_q = \left[-\ln \left(1 - \left(1 - (1-u)^{\frac{1}{\beta}} \right)^{\frac{1}{\alpha}} \right) \right]^{-\frac{1}{\lambda}}$$

Replacing the scale parameter σ in equation (1) then we get Cumulative Distribution function for Exponentiated Generalized Frechet distribution is

$$F(x) = \left[1 - \left(1 - \exp - \left\{ \left(\frac{\sigma}{\varphi_q \delta_q} \right)^\lambda \right\} \right)^\alpha \right]^\beta \tag{4}$$

Taking a first derivative of partial differentiation with respect to δ_q , we get

$$\frac{\partial F(t, \delta_q)}{\partial \delta_q} = \frac{e^{-\frac{1}{\varphi_q \delta_q}}}{\varphi_q \delta_q^2} \beta \left[1 - \left(1 - \exp \left\{ - \left(\frac{1}{\varphi_q \delta_q} \right)^\lambda \right\} \right)^\alpha \right]^{\beta-1}$$

$$- \alpha \left[1 - \exp \left\{ - \left(\frac{1}{\varphi_q \delta_q} \right)^\lambda \right\} \right]^{\alpha-1} \cdot -\lambda \exp \left\{ - \left(\frac{1}{\varphi_q \delta_q} \right)^{\lambda-1} \right\} \tag{5}$$

$\frac{\partial F(t, \delta_q)}{\partial \delta_q} > 0$. The cumulative distribution function $F(t, \delta_q)$ is a non- decreasing function of δ_q .

3. CONDITIONAL REPETITIVE GROUP SAMPLING PLAN (CRGS) BASED ON TRUNCATED LIFE TEST

Conditional Repetitive Group sampling plan is the extension of Repetitive Group Sampling plan. The following notations similar to those of Sherman (1965), the proposed conditional RGS plan is carried out through the following steps:

3.1. Conditions for the application of CRGS

- The production is steady, so the results of previous, present, and future lots can be used as broad indicators of a process that will continue into the future.
- It is possible to submit isolated lots or a series of lots.
- Inspection is by attributes, when lot quality is defined as the proportion of defective.
- Variation in lot quality may exist.
- Lot has at least one defective unit.
- Lots submitted for inspection may be of low quality.

3.2. Operating procedure of Conditional Repetitive Group Sampling plan

- Step 1: Draw a random sample of size n and determine the number of defectives d found therein.
- Step 2: Accept the lot, if $d \leq c_1$. Reject the lot, if $d > c_2$.

- Step 3: If $c_1 < d \leq c_2$, repeat the steps (1), (2) and (3) provided previous 'i' lots are accepted (i.e. in each of the previous 'i' lots $d \leq c_1$); otherwise reject the lot.

The proposed plan parameters are $(n, c_1, c_2, i, \frac{t}{t_q})$ Where, i=acceptance criterion, Figure 1 repre-

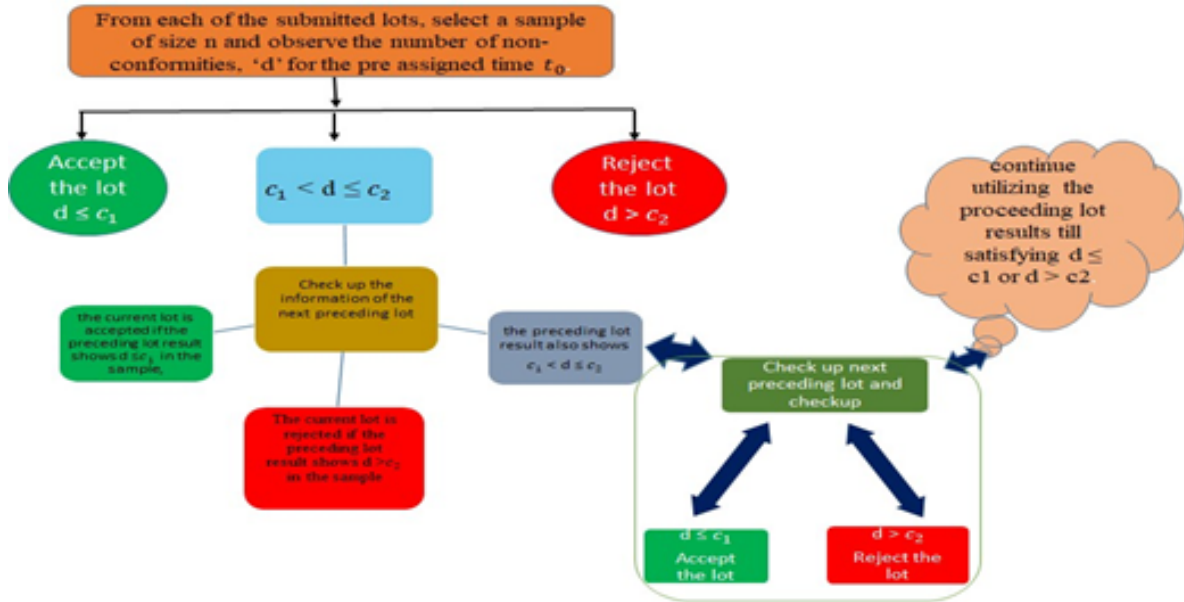


Figure 1: Flow Chart of operating procedure of Conditional Repetitive Group Sampling Plan based on Truncated Life Test

sents the Operating Procedure of Conditional Repetitive Group Sampling Plan based on Truncated Life test. We have used binomial models to determine the number of samples. The Operating Characteristics function of Conditional Repetitive Group Sampling Plan is,

$$L(P) = \frac{P_1}{1 - P_3 P_1^i} \tag{6}$$

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^i - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right]^i} \tag{7}$$

Where,

$$P_1 = \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i$$

$$P_2 = 1 - \sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^i$$

$$P_3 = 1 - P_1 - P_2 \tag{8}$$

The failure probability is expressed as 'p'. As a result of the cumulative distribution function of the lifetime distributions, these failure probabilities can be determined. These probabilities can then be used to estimate the percentile lifetime of a product. They can also be used to assess the reliability of a product, as well as the risk of failure.

3.3. Minimum Sample Size

Our sampling plan parameters are $(n, c_1, c_2, i, \frac{t}{t_q})$ for a given Probability of acceptance P^* . We determine that acceptable exhaust sized lots are possible, as well as that the binomial distribution may be used. The study is designed to determine the minimum sample size 'n' necessary to ensure that $t_q < t_q^0$.

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^{n-i}}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^{n-i} - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^{n-i} \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^{n-i} \right]^i} \leq 1 - P^* \tag{9}$$

Where, P^* is the Probability of acceptance.

The minimum sample size n is determined by calculating the binomial distribution, where n is the number of samples, c_1 is the number of successes in n trials, c_2 is the number of failures in n trials, i is the acceptance criterion and $\frac{t}{t_q}$ is the number of successes that are expected in n trials.

As shown in Table 1, the Exponentiated Generalized Frechet Distribution requires a minimum sample size n to exceed the actual 50th percentile value.

Table 1: The minimum sample size n for the specified 50th percentile value of the Exponentiated Generalized Frechet Distribution exceeding the actual 50th percentile value , with probability p^* and acceptance number c using binomial approximation.

P^*	i	c_1	c_2	$\frac{t}{t_q}$				
				0.9	1	1.1	1.2	1.3
0.75	1	0	1	19	3	1	1	1
	1	0	2	20	3	2	2	2
	1	1	1	36	5	3	2	2
	1	1	2	37	5	3	2	2
	1	1	3	38	5	3	3	3
	3	0	1	20	2	1	1	1
	3	0	2	21	3	2	2	2
	3	1	1	37	5	3	2	2
	3	1	2	38	6	3	2	2
0.90	1	0	1	30	4	2	1	1
	1	0	2	32	6	4	2	2
	1	1	1	51	7	4	2	2
	1	1	2	53	7	5	2	2
	1	1	3	54	7	3	3	3
	3	0	1	30	4	2	1	1
	3	0	2	32	5	2	2	2
	3	1	1	52	7	3	2	2
	3	1	2	53	8	6	2	2
0.95	1	0	1	39	5	3	1	1
	1	0	2	40	6	4	2	2
	1	1	1	63	8	4	3	3
	1	1	2	64	9	5	4	3
	1	1	3	65	10	7	4	3
	3	0	1	41	6	3	2	2
	3	0	2	42	7	3	2	2
	3	1	1	64	9	4	3	2
	3	1	2	65	10	6	4	3
0.99	1	0	1	60	7	3	2	2
	1	0	2	62	8	4	2	2
	1	1	1	88	11	5	3	3
	1	1	2	89	12	6	4	3
	1	1	3	90	13	7	5	4
	3	0	1	62	8	3	2	2
	3	0	2	65	10	6	2	2
	3	1	1	89	12	5	3	3
	3	1	2	90	13	6	4	3

3.4. Operating Characteristic Function

The operating characteristic curve plots the relationship between the probability of acceptance of a product and the specified lifetime of that product. It is required that the operating characteristic function of the Conditional Repetitive Group sampling plan satisfy the following equation,

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^i - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right]^i} \tag{10}$$

Where, $p = F(t, \delta)$ and δ can be indicated as a function of $\delta = \frac{t}{t_q}$

$$\therefore p = F\left(\frac{t}{t_q} \cdot \frac{1}{d_q}\right) \tag{11}$$

Where $d_q = \frac{t_q}{t_q^0}$.

The operating characteristic curve helps to identify the optimal sample size for a sampling plan. It can also be used to compare different sampling plans and determine which one is more efficient. Table 2 represents the values of operating Characteristic function for using the acceptance numbers $c_1 = 0, c_2 = 2$ and $i=1,3$.

3.5. Producer’s Risk Ratio

A producer’s risk is explained as the probability of rejecting a lot when a lot has a rejection rate $t_q > t_q^0$. Consider the case of a given producer’s risk, say α . Table 3 represents the minimum ratio of true lifetime to specified lifetime for the acceptability of a lot with producer’s risk of 0.05. Producer’s risk ratio, which is based on the Conditional Repetitive Group Sampling plan in the ratio of the specified lifetime to the actual lifetime. It must meet the following conditions:

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^i - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right]^i} > 1 - \alpha \tag{12}$$

Where, $p = F\left(\frac{t}{t_q} \cdot \frac{1}{d_q}\right)$

4. APPLICATIONS

4.1. Numerical Illustration

We consider the inspector wants to conduct the inspection for lifetime of battery in Laptop and he suggests to use the Exponentiated Generalized Frechet distribution with pre-determined known shape parameters are $\alpha = 3, \beta = 6$ and $\lambda = 6$. Figure 2 represents the images of battery for Laptop. The investigator wants to run the experiment for 2700 hours, but the laboratory has testers to true percentile life time $t_{0.50} = 3000$ hours, $c_1 = 0, c_2 = 2, i=3, \alpha = 0.05, \beta = 0.10$, then, $\phi = 1.7045$ is computed from the equation get under the percentile estimator and the minimum ratio and the minimum sample size is $n=3$ get the information from the Table 1.

The probability of acceptance is given by $(n, c_1, c_2, i, t/t_{0.50}) = (3, 0, 2, 3, 1.1)$ with $p^* = 0.95$ under the Exponentiated Generalized Frechet distribution.

Table 4 represents the Operating Characteristic values with $p^* = 0.95$ under Exponentiated Generalized Frechet Distribution. Figure 3 represents the Operating Characteristic curve for

Table 2: Operating Characteristic values for Conditional Repetitive Group sampling Plan ($n, c_1 = 0, c_2 = 2, t/t_{0.50}$) for a given p^* under Exponentiated Generalized Frechet distribution.

P^*	i	n	$\frac{t}{t_q}$	t_q/t_q^0					
				1	1.1	1.2	1.25	1.3	1.4
0.75	1	20	0.9	0.2498	0.9997	1.0000	1.0000	1.0000	1.0000
	1	3	1	0.1399	0.9112	1.0000	1.0000	1.0000	1.0000
	1	2	1.1	0.0271	0.3108	0.9950	0.9998	1.0000	1.0000
	1	2	1.2	0.0019	0.0342	0.6472	0.9113	0.9994	1.0000
	1	2	1.3	0.0001	0.0031	0.1227	0.3108	0.8845	0.9985
	1	21	0.9	0.2020	0.9991	0.1000	0.1000	0.1000	0.1000
	3	3	1	0.1268	0.8199	0.9999	1.0000	1.0000	1.0000
	3	2	1.1	0.0264	0.2552	0.9862	0.9995	1.0000	1.0000
	3	2	1.2	0.0019	0.0342	0.6472	0.9113	0.9994	1.0000
0.90	1	32	0.9	0.0906	0.9993	1.0000	1.0000	1.0000	1.0000
	1	6	1	0.0161	0.7092	0.9999	1.0000	1.0000	1.0000
	1	4	1.1	0.0006	0.0662	0.9802	0.9994	1.0000	1.0000
	1	2	1.2	0.0019	0.0342	0.6472	0.9113	0.9994	1.0000
	1	2	1.3	0.0001	0.0031	0.1227	0.3108	0.8845	0.9985
	3	34	0.9	0.0868	0.9981	0.1000	0.1000	0.1000	0.1000
	3	5	1	0.0319	0.6520	0.9999	1.0000	1.0000	1.0000
	3	2	1.1	0.0264	0.2552	0.9862	0.9995	1.0000	1.0000
	3	2	1.2	0.0019	0.0331	0.5159	0.8191	0.9984	1.0000
0.95	1	40	0.9	0.0479	0.9989	1.0000	1.0000	1.0000	1.0000
	1	6	1	0.0161	0.7092	0.9999	1.0000	1.0000	1.0000
	1	4	1.1	0.0006	0.0662	0.9802	0.9994	1.0000	1.0000
	1	2	1.2	0.0019	0.0342	0.6472	0.9113	0.9994	1.0000
	1	2	1.3	0.0001	0.0031	0.1227	0.3108	0.8845	0.9985
	3	42	0.9	0.0404	0.9967	0.1000	0.1000	0.1000	0.1000
	3	7	1	0.0080	0.5148	0.9999	1.0000	1.0000	1.0000
	3	3	1.1	0.0043	0.1268	0.8704	0.9990	1.0000	1.0000
	3	2	1.2	0.0019	0.0331	0.5159	0.8191	0.9984	1.0000
0.99	1	62	0.9	0.0087	0.9975	1.0000	1.0000	1.0000	1.0000
	1	8	1	0.0040	0.5695	0.9999	1.0000	1.0000	1.0000
	1	4	1.1	0.0006	0.0662	0.9802	0.9994	1.0000	1.0000
	1	2	1.2	0.0019	0.0342	0.6472	0.9113	0.9994	1.0000
	1	2	1.3	0.0001	0.0031	0.1227	0.3108	0.8845	0.9985
	3	89	0.9	0.9985	0.9999	0.1000	0.1000	0.1000	0.1000
	3	12	1	0.0090	0.7167	1.0000	1.0000	1.0000	1.0000
	3	6	1.1	0.0033	0.1114	0.9984	0.9999	1.0000	1.0000
	3	4	1.2	0.0005	0.0208	0.6963	0.9618	0.9999	1.0000
	3	3	1.3	0.0003	0.0091	0.2604	0.5284	0.9808	0.9999

Table 3: Minimum ratio of true lifetime to specified lifetime for the acceptability of a lot with producer's risk of 0.05

P*	i	c ₁	c ₂	$\frac{t}{t_q}$				
				0.9	1	1.1	1.2	1.3
0.75	1	0	1	1.0307	1.0792	1.1147	1.2160	1.3175
	1	0	2	1.0290	1.0747	1.1600	1.2655	1.3710
	1	1	1	1.0297	1.0684	1.1401	1.1995	1.2992
	1	1	2	1.0232	1.0531	1.1161	1.1563	1.2528
	1	1	3	1.0211	1.0477	1.1098	1.2105	1.3115
	3	0	1	1.0388	1.0741	1.1406	1.2443	1.3478
	3	0	2	1.0390	1.0905	1.1801	1.2874	1.3946
	3	1	1	1.0304	1.0686	1.1399	1.1993	1.2994
	3	1	2	1.0275	1.0711	1.1290	1.1811	1.2796
	3	1	3	1.0278	1.0772	1.1652	1.2295	1.3320
0.90	1	0	1	1.0546	1.1130	1.1929	1.2560	1.3606
	1	0	2	1.0529	1.1224	1.2192	1.2966	1.4047
	1	1	1	1.0521	1.1085	1.1922	1.2817	1.3885
	1	1	2	1.0421	1.0888	1.1788	1.1913	1.2907
	1	1	3	1.0381	1.1071	1.1343	1.2375	1.3406
	3	0	1	1.0618	1.1251	1.2101	1.2832	1.3901
	3	0	2	1.0620	1.1309	1.2087	1.3187	1.4286
	3	1	1	1.0526	1.1086	1.1747	1.2465	1.3503
	3	1	2	1.0457	1.1015	1.1976	1.2156	1.3166
	3	1	3	1.0446	1.1038	1.2026	1.2562	1.3611
0.95	1	0	1	1.0678	1.1330	1.2271	1.2788	1.3854
	1	0	2	1.0651	1.1343	1.2332	1.3146	1.4244
	1	1	1	1.0650	1.1271	1.2101	1.3031	1.4117
	1	1	2	1.0524	1.1092	1.1913	1.2846	1.3664
	1	1	3	1.0486	1.1071	1.2014	1.2759	1.3575
	3	0	1	1.0760	1.1496	1.2416	1.3379	1.4493
	3	0	2	1.0751	1.1520	1.2403	1.3364	1.4477
	3	1	1	1.0654	1.1311	1.2104	1.3031	1.3790
	3	1	2	1.0560	1.1192	1.2088	1.2951	1.3807
	3	1	3	1.0544	1.1232	1.2127	1.2916	1.3773
0.99	1	0	1	1.0912	1.1652	1.2550	1.3542	1.4675
	1	0	2	1.0890	1.1642	1.2603	1.3513	1.4624
	1	1	1	1.0883	1.1624	1.2525	1.3416	1.4538
	1	1	2	1.0710	1.1377	1.2232	1.3130	1.3994
	1	1	3	1.0653	1.1328	1.2222	1.3036	1.3910
	3	0	1	1.0978	1.1757	1.2679	1.3695	1.4850
	3	0	2	1.0974	1.1768	1.2885	1.3690	1.4838
	3	1	1	1.0876	1.1636	1.2506	1.3412	1.4534
	3	1	2	1.0729	1.1439	1.2307	1.3196	1.4107
	3	3	3	1.0714	1.1455	1.2327	1.3162	1.4069



Figure 2: Image of laptop battery

Table 4: $(n, c_1, c_2, i, \frac{t}{t_{0.50}}) = (3, 0, 2, 3, 1.1)$ with $p^* = 0.95$ under Exponentiated Generalized Frechet Distribution

$\frac{t_{0.50}}{t_{0.50}^0}$	1	1.1	1.2	1.25	1.3	1.4
L(p)	0.0043	0.1268	0.8904	0.9790	1.0000	1.0000

truncated lifetest using Exponentiated Generalized Frechet distribution.

It reveals that if the true 50th percentile is almost equal to the true 50th percentile ($t_{0.50}/t_{0.50}^0 = 1.2$) the producer's risk nearly 0.9957 (1-0.0043). The producer's risk is an almost nearly equal to Zero whenever the actual 50th percentile is greater than or equal to 1.2 times the specified 50th percentile.

4.2. Real Data Study

According to Lawless (2012), consider the real data study for lifetime for brake pads to pre-determine the minimum of thickness of product. In this study follows lifetime distribution of Exponentiated Generalized Frechet Distribution with known shape Parameters. The values of data as follows:

First, we check the goodness of fit for given data. The Kolmogorov-Smirnov test, Anderson-Darling test, Shapiro-Wilk's normality test, Histogram, box-plot and the Q-Q plot are all considered for goodness of fit.

Figure 7 graphically represents the Histogram satisfies the Normality of the given data. We get the result of Kolmogorov Smirnov test statistic is 0.999, Anderson Darling test is 0.94590 and the Shapiro-Wilk test is 0.95462. Figure 6 represents the graphically satisfies the Q-Q plot and

38.7,	49.2,	42.4,	73.8,	46.7	44.1,	61.9,	39.3,	49.8,	46.3,	56.2,
50.5,	54.9,	54,	49.2,	44.8,	72.2,	107.8,	81.6,	45.2,	124.6,	64,
83,	143.6,	43.4,	69.6,	74.8,	32.9,	51.5,	31.8,	77.6,	63.7,	83,
24.8,	68.8,	68.8,	89.1,	65,	65.1,	59.3,	53.9,	79.4,	47.4,	61.4,
72.8,	61.4,	72.8,	54,	37.2,	44.2,	50.8,	65.5,	86.7,	43.8,	100.6,
67.6,	89.5,	60.3,	103.6,	82.6,	88,	42.4,	68.9,	95.7,	78.1,	83.6,
18.6,	92.6,	42.4,	34.3,	105.6,	68.9,	78.7,	165.5,	79.5,	55,	46.8,
124.5,	92.5,	110,	101.2,	59.4,	27.8,	33.6,	69,	75.2,	58.4,	105.6,
56.2,	55.9,	83.8,	123.5,	69,	101.9,	87.6,	38.8,	74.7		

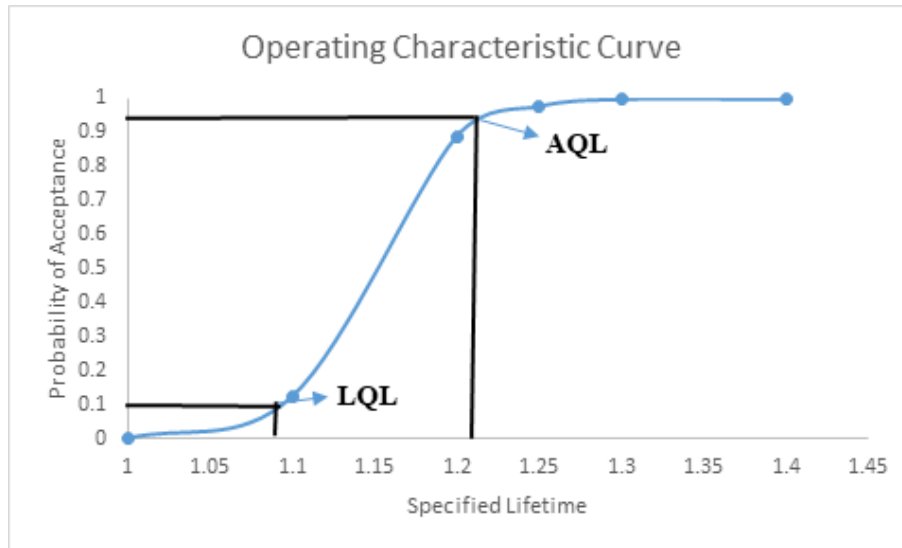


Figure 3: Operating Characteristic Curve for Exponentiated Generalized Frechet Distribution

Figure 5 represents the box plot of the given data. Hence, the data provides reasonable goodness of fits for data set.

Assume that the inspector wants to inspect the lifetime of brake pad and he interested to using Double Sampling plan follows lifetime as Exponentiated Generalized Frechet Distribution. He wants to conduct the runtime of experiment is 2600hrs but the laboratory has the testers to true percentile life time $t_{0.10} = 2450$ hrs. $c_1 = 0, c_2 = 2$, Producer’s Risk (α) = 0.05, Consumer’s Risk (β) = 0.10, $i=1$ then, $\phi = 1.704$ is computed from the equation get under the percentile estimator and the minimum ratio, $t/t_q = 1$ and minimum sample size is $n = 7$ get the information from the Table 1. The probability of acceptance is characterized by $((n, c_1, c_2, i, t/t_0.50) = (7, 0, 2, 3, 1)$ with $P^* = 0.95$ under Exponentiated Generalized Frechet the values of tables are given. Since there were no items with a failure time less than or equal to 2600 hours in the given sample of $n=7$ observations, the experimenter would accept the lot, assuming the 50th percentile lifetime $t_{0.50}$ of at least 2450 hours with a confidence level of $P^- = 0.95$.

5. CONSTRUCTION OF TABLES

Step 1: Find the value of ϕ for fixing the values of parameters are λ, β, α and $q=0.50$. Set the evaluated $\phi=1.0745, c_1 = 0, 1, c_2 = 1, 2, 3$ and $\frac{t}{t_q} = 0.9, 1.1, 1.2, 1.3$ and 1.4.

Step 2: Find the minimum value of n satisfying

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^i - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right]^i} \leq 1 - P^*$$

Step 3: Find the operating characteristic function of the Conditional Repetitive Group sampling plan must satisfy,

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^i - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^i \right]^i}$$

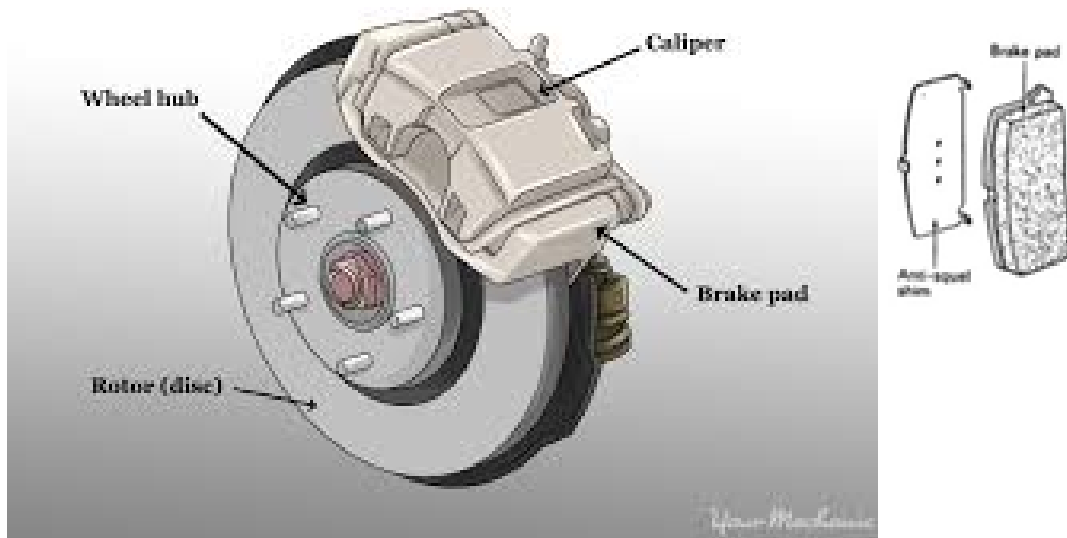


Figure 4: Image of brakepads

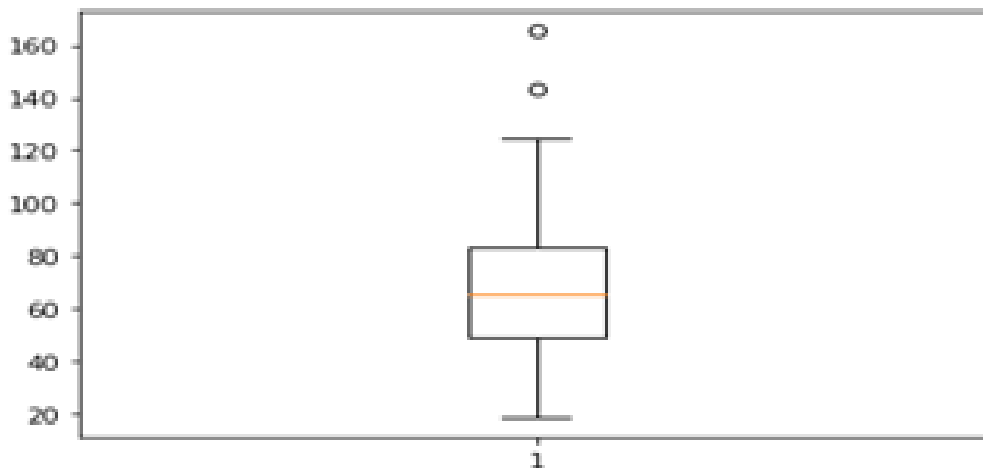


Figure 5: Box Plot

Where $d_q = \frac{t_q}{t_0^q}$. Set the evaluated $\frac{t_q}{t_0^q} = 1, 1.1, 1.2, 1.25, 1.3$ and 1.4 .

Step 4: Find the minimum ratio for the acceptability of a lot with producer's risk must satisfy the condition is,

$$L(P) = \frac{\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^{n-i}}{1 - \left[\sum_{i=0}^{c_2} \binom{n}{i} p^i (1-p)^{n-i} - \sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^{n-i} \right] \left[\sum_{i=0}^{c_1} \binom{n}{i} p^i (1-p)^{n-i} \right]^i} > 1 - \alpha$$

Where, $p = F\left(\frac{t}{t_0} \cdot \frac{1}{d_q}\right)$

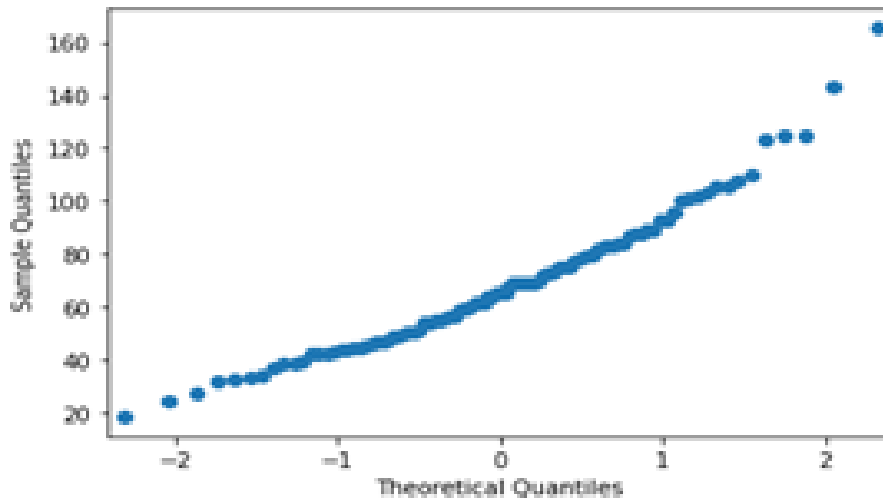


Figure 6: Q-Q plot

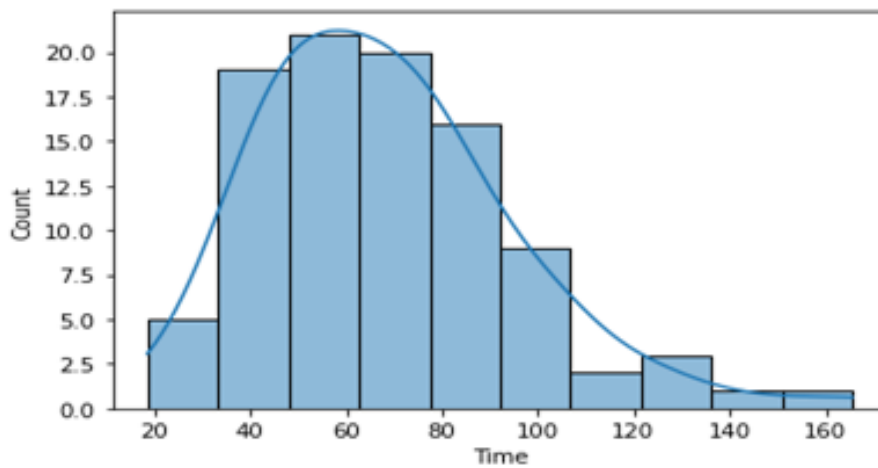


Figure 7: Histogram

6. CONCLUSION

In this article deals with Conditional Repetitive Group Sampling plan are designed when the life-times of the items follow Exponentiated Generalized Frechet Distribution. Conditional Repetitive group sampling plan is quite flexible and reliable. This sampling plan can also be used to increase the accuracy of the life tests. It can also help to reduce the sample size, which in turn can help reduce the costs associated with the life tests. According to the results of this study, as the time termination ratio increase, the sample size decreases. Further, it has been shown that there is an increase in the operating characteristic values when the quality is improved. So, it is strongly suggested that industrial practitioners test electrical components using the proposed plan. This plan is also cost-effective and efficient, making it an ideal choice for industrial practitioners. It can be easily adjusted to fit different production requirements. The study will also concentrate the effectiveness of the sampling plan in terms of accuracy and precision. Finally, the study will provide recommendations for future research.

REFERENCES

- [1] Abd-Elfattah, A. M., Assar, S. M., Abd-Elghaffar, H. I. (2016). Exponentiated Generalized Frechet Distribution. *International Journal of Mathematical Analysis and Applications*,3(5):39-48.
- [2] Anburajan, P., Ramaswamy, A. R. S.(2015). A Conditional Repetitive Group Sampling Plan for Truncated Life Tests Using Log - Logistic, Exponentiated Log - Logistic, Rayleigh and Inverse Rayleigh Distributions. *Journal of Progressive Research in Mathematics* ,2(2):118-130.
- [3] Aslam, M., Kundu, D. and Ahmad, M. (2010). Time Truncated Acceptance Sampling Plans for Generalized Exponential Distribution. *Journal of Applied Statistics*,37(4):555-566.
- [4] Balakrishnan, N., Leiva, L. and Lopez, J. (2007). Acceptance Sampling Plans from Truncated Life Tests Based on the Generalized Birnbaum-Saunders Distribution. *Communications in Statistics - Simulation and Computation*,36(3):643-656.
- [5] Epstein, B. (1954). Truncated Life Tests in the Exponential Case. *Annals of Mathematical Statistics*,25(3):555-564.
- [6] Goode, H. P. and Kao, J. H. K.(1961). Sampling Plans Based on the Weibull Distribution. *Proceeding of the Seventh National Symposium on Reliability and Quality Control, Philadelphia*, 24-40.
- [7] Gupta, S. S. (1962). Life Test Sampling Plans for Normal and Lognormal Distributions. *Technometrics*,4(2):151-175.
- [8] Jayalakshmi, S. and Kaviyamani, K. (2020). Designing quick switching conditional repetitive group sampling system indexed through quality decision region using MATLAB program. *AIP Conference Proceedings*,2261(030099) .
- [9] Kaviyarasu, V., Fawaz, P. (2017). Design of acceptance sampling plan for life tests based on percentiles using Weibull-Poisson distribution. *Int J Stat Appl Math* ,2(5):51-57 .
- [10] Pradeepa Veerakumari, K. and Ponneswari, P. (2016). Designing of acceptance sampling plan for life tests based on percentiles of exponentiated Rayleigh distribution. *International Journal of Current Engineering and Technology* ,6(4):1148-1153.
- [11] Rosaiah, K. and Kantam, R.R.L. (2005). Acceptance Sampling Based on the Inverse Rayleigh Distribution. *Economic Quality Control*,20(2):277-286.
- [12] Rosaiah, K., Kantam, R. R. L. and Kumar, S. (2006). Reliability of Test Plans for Exponentiated Log-Logistic Distribution. *Economic Quality Control*, 21(2):165-175.
- [13] Shankar, G. and Mohapatra, B.N. (1993). GERT Analysis of Conditional Repetitive Group Sampling Plan. *International Journal of Quality Reliability Management*,10(2).
- [14] Sherman, E. (1965). Design and Evaluation of Repetitive Group Sampling Plans. *Technometrics*,7(1):11-21.

NEW EXTENSION OF INVERTED MODIFIED LINDLEY DISTRIBUTION WITH APPLICATIONS

DEVENDRA KUMAR^a, ANJU GOYAL^b, P. PAREEK^c AND M. SAHA^a



^aDepartment of Statistics, Faculty of Mathematical Sciences, University of Delhi, India

^bDepartment of statistics, Panjab University Chandigarh, India

^cDepartment of Statistics, Central University of Rajasthan, Kishangarh, India

corresponding author: anju.statistics@gmail.com

Abstract

In this article we, proposed a new two parameter distribution called inverted power modified Lindley distribution. The main objective is to introduce an extension to inverted modified Lindley distribution as an alternative to the inverted exponential, inverted gamma and inverted modified Lindley distributions, respectively. The proposed distribution is more flexible than the above mentioned distributions in terms of its hazard rate function. In the part of estimation of the proposed model, we first utilize the maximum likelihood (ML) estimator and parametric bootstrap confidence intervals, viz., standard bootstrap, percentile bootstrap, bias-corrected percentile (BCPB), bias-corrected accelerated bootstrap (BCAB) from the classical point of view as well the Bayesian estimation under different loss functions, squared error loss function, modified squared error loss function, and Bayes credible interval as to obtain the model parameter based on order statistics. A simulation study is carried out to check the efficiency of the classical and the Bayes estimators in terms of mean squared errors and posterior risks, respectively. Two real life data sets, have been analyzed for order statistics to demonstrate how the proposed methods may work in practice.

Keywords: Inverted modified Lindley distribution, moments, maximum likelihood estimator, order statistics, bootstrap confidence intervals, Bayes estimators.

1. INTRODUCTION

The inverted modified Lindley (IML) distribution is one of the most famous one-parameter distributions used for modeling count data, which was introduced by [5] as a mixture of inverted exponential and inverted gamma distributions with mixing proportion $\theta/(1 + \theta)$, to illustrate difference between fiducial distribution and posterior distribution. [5] pointed out IML distribution outperforms the classical inverse Lindley distribution for some real data sets. They studied many properties of this distribution such as moments and inverse moments and also, noted down that the first four moments of this distribution. Furthermore, the IML distribution does not provide a reasonable parametric fit for modeling phenomenon with non-monotone failure rates, such as the upside-down bathtub failure rates, which are common in reliability and biological studies. For example, such failure rates curves can be observed in the course of a disease whose mortality reaches a peak after some finite period and then declines gradually.

Several generalizations of Lindley distribution have been attempted by many researchers in the existing literature such as [18] studied the generalized Lindley, [3] proposed an extended Lindley, [10] proposed the power Lindley distribution, [2] introduced the exponentiated power Lindley distribution, [4] proposed exponential Poisson Lindley distribution, [1] proposed a

new weighted Lindley distribution, [12] proposed Wrapped Lindley distribution, [7] proposed alpha power transformed inverse Lindley distribution, [7] proposed alpha-power transformed Lindley distribution, [6] proposed a new modified Lindley distribution without considering any special function or additional parameters. Recently, [13] introduced power modified Lindly (PML) distribution. They showed that PML distribution provides better fit than Lindley, Weibull, gamma, generalized exponential (GE) and power Lindley (PL) distributions and it was suitable for modeling constant, increasing, decreasing and unimodal shaped hazard rate function.

Many researchers considered the inverted modified Lindley (IML) distribution in their studies. For example, [14] studied the moments of order statistics and also estimation of the parameters by using maximum likelihood methods, [15] have established relations for moments of generalized order statistics and also proposed the estimation procedures under complete and censored data. This study presents a one parameter extension of the IML distribution by [5]. The presented distribution shows the flexible shapes of the density and hazard functions and gives better fits than some well-known lifetime distributions, such as inverted modified Lindley, Modified Lindley and Lindley distributions. In this article, we propose a three-parameter distribution, referred to as inverted power modified Lindley (IPML) distribution using a similar idea [18], which is the linear combination of inverted power exponential and inverted power gamma distribution. We are motivated to introduce the IPML distribution because (i) it contain lots of aforementioned of known lifetime models; (ii) it is capable of modelling monotonically increasing, decreasing, hazard rates; (iii) it can be viewed as a suitable model for fitting the skewed data which may not be properly fitted by other common distributions and can also be used in a variety of problems in various areas such as public health, biomedical studies, environmental studies and industrial reliability and survival analysis; and, (iv) Three real life data applications show that it compares well with other competing lifetime distributions in modelling lifetime data.

The objective of this paper is three fold: First, we obtain the estimates of model parameters based on maximum likelihood method of estimation. The performance of the MLE is demonstrated in terms of their mean squared errors (MSEs) based on simulated samples and for different sample sizes through a simulation study. The second objective is to obtain four bootstrap confidence intervals (BCIs) of model parameters based on MLE. The performances of the BCIs are demonstrated in terms of their estimated coverage probabilities (CPs) and average widths (AWs). The third objective is to obtain Bayes estimates (BEs) of the model parameters under four loss functions (symmetric as well as asymmetric loss functions).

The rest of the paper is organized as follows: In Section 2, we described proposed model PIML. In Section 3, dealt with some statistical and mathematical properties of PIML distribution. Section 4 described the MLE and BCIs, namely, standard bootstrap (SB), percentile bootstrap (PB), bias-corrected percentile bootstrap (BCPB) and bias-corrected accelerated bootstrap (BCAB) based on MLE have been discussed. Also, we derive the Bayes estimators of the model parameters under four loss functions. In Section 5, a Monte Carlo simulation study has been carried out to assess the performances of the above cited classical and Bayes estimators in terms of their MSEs. Also, we assess the performances of different BCIs and Bayes credible intervals in terms of coverage probabilities (CPs) and average widths (AWs). For illustrative purposes, two real data sets are analyzed in Section 6. Finally, concluding remarks are given in Section 7.

2. MODEL DESCRIPTION

The one parameter inverted modified Lindley (IML) distribution proposed by [5] with cumulative distribution function (CDF)

$$F(y) = \left(1 + \frac{\eta}{1 + \eta} \frac{1}{y} e^{-\eta/y}\right) e^{-\eta/y}, \quad y > 0, \quad \eta > 0.$$

Now, we introduce a skewness parameter to the inverted modified Lindley distribution using a similar idea to [9], [10], [16] and [13] i.e., $X = Y^{1/\tau}$, $\tau > 0$ and to obtain a power inverted

modified Lindley (PIML) distribution. The CDF of the two parameter PIML distribution is given by

$$F(x) = \left(1 + \frac{\eta}{1 + \eta} \frac{1}{x^\tau} e^{-\eta/x^\tau}\right) e^{-\eta/x^\tau}, \quad x > 0, \eta > 0, \tau > 0, \tag{1}$$

and the corresponding probability density function (PDF) given by

$$f(x) = \frac{\eta}{1 + \eta} \frac{\tau e^{-2\eta/x^\tau}}{x^{\tau+1}} \left((1 + \eta) e^{\eta/x^\tau} + \frac{2\eta}{x^\tau} - 1 \right), \quad x > 0, \eta > 0, \tau > 0, \tag{2}$$

The corresponding survival function for a specified value $X = x$ is obtained as

$$S(x) = 1 - F(x) = 1 - \left(1 + \frac{\eta}{1 + \eta} \frac{1}{x^\tau} e^{-\eta/x^\tau}\right) e^{-\eta/x^\tau}, \quad x > 0, \eta > 0, \tau > 0, \tag{3}$$

Thus, we can also express the corresponding hazard rate function (HRF) for specified $X = x$ as

$$h(x) = \frac{\frac{\eta}{1 + \eta} \frac{\tau e^{-2\eta/x^\tau}}{x^{\tau+1}} \left[(1 + \eta) e^{\eta/x^\tau} + \frac{2\eta}{x^\tau} - 1 \right]}{1 - \left(1 + \frac{\eta}{1 + \eta} \frac{1}{x^\tau} e^{-\eta/x^\tau}\right) e^{-\eta/x^\tau}}, \quad x > 0, \eta > 0, \tau > 0, \tag{4}$$

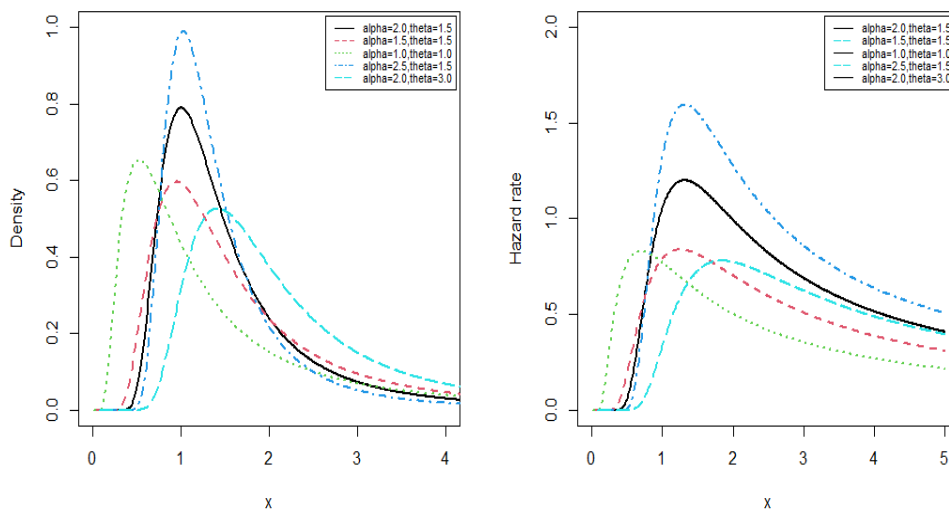


Figure 1: PDF and HRF of the PIML distribution.

From the Figure 1, it is clear that the PDF and HRF of the PIML distribution is right skewed distribution and initially increasing and then decreasing behaviour for the considered parameters values and for specified time. The corresponding cumulative hazard rate function is defined by

$$C(x) = -\log S(x) = -\log \left\{ 1 - \left(1 + \frac{\eta}{1 + \eta} \frac{1}{x^\tau} e^{-\eta/x^\tau}\right) e^{-\eta/x^\tau} \right\}, \quad x > 0, \eta > 0, \tau > 0. \tag{5}$$

When $\tau = 1$, the PIML distribution reduces to IML distribution. An advantage of the definition of $f(x)$ is that we can write it as a linear combination of well established PDFs as

$$f(x) = f_1(x) + \frac{1}{2(1 + \eta)} (f_2(x) - f_3(x)), \tag{6}$$

where, $f_1(x)$ is inverted exponential with parameter (η, τ) , $f_2(x)$ is inverted gamma with parameter $(2\eta, 2\tau)$ and $f_3(x)$ is inverted exponential with parameter $(2\eta, \tau)$

$$f_1(x) = \frac{\tau \eta e^{-\eta/x^\tau}}{x^{\tau+1}}, \quad f_2(x) = \frac{(2\eta)^{2\tau}}{x^{2\tau+1}} e^{-2\eta/x^\tau} \quad \text{and} \quad f_3(x) = \frac{2\eta \tau}{x^{\tau+1}} e^{-2\eta/x^\tau}$$

3. STATISTICAL AND MATHEMATICAL PROPERTIES OF PIML DISTRIBUTION

Here, we have discussed and derived several mathematical and statistical properties, which are given in the following subsections.

3.1. Moments and moment generating function

Let X be a random variable from PIML distribution with PDF given in (2), then its moments is given by the following

$$\begin{aligned} \mu'_r &= \int_0^\infty x^r f(x) dx = \int_0^\infty x^r \frac{\eta}{1+\eta} \frac{\tau e^{-\frac{2\eta}{x^\tau}}}{x^{\tau+1}} \left((1+\eta)e^{\frac{\eta}{x^\tau}} + \frac{2\eta}{x^\tau} - 1 \right) dx \\ &= \int_0^\infty x^{r-\tau-1} \tau \eta e^{-\frac{\eta}{x^\tau}} dx + \frac{1}{2(1+\eta)} \left(\int_0^\infty \tau (2\eta)^2 x^{r-2\tau-1} e^{-\frac{2\eta}{x^\tau}} - \int_0^\infty \tau (2\eta) x^{r-\tau-1} e^{-\frac{2\eta}{x^\tau}} \right) dx \\ &= \eta^{r/\tau} \Gamma\left(1 - \frac{r}{\tau}\right) \left(1 - \frac{2^{r/\tau-1}}{1+\eta} \left(\frac{r}{\tau}\right)\right). \end{aligned} \tag{7}$$

Also, the first four inverse moments are given by

$$\begin{aligned} E(Y^{-1}) &= \frac{1}{\eta^{1/\tau}} \Gamma\left(1 + \frac{1}{\tau}\right) \left(1 + \frac{1}{2^{\frac{1}{\tau}+1}(1+\eta)} \left(\frac{1}{\tau}\right)\right) \\ E(Y^{-2}) &= \frac{1}{\eta^{2/\tau}} \Gamma\left(1 + \frac{2}{\tau}\right) \left(1 + \frac{1}{2^{\frac{2}{\tau}+1}(1+\eta)} \left(\frac{2}{\tau}\right)\right) \\ E(Y^{-3}) &= \frac{1}{\eta^{3/\tau}} \Gamma\left(1 + \frac{3}{\tau}\right) \left(1 + \frac{1}{2^{\frac{3}{\tau}+1}(1+\eta)} \left(\frac{3}{\tau}\right)\right) \\ E(Y^{-4}) &= \frac{1}{\eta^{4/\tau}} \Gamma\left(1 + \frac{4}{\tau}\right) \left(1 + \frac{1}{2^{\frac{4}{\tau}+1}(1+\eta)} \left(\frac{4}{\tau}\right)\right). \end{aligned}$$

Table 1 presents the numerical values of these inverse moments for various values of t . For any $t < \eta$, the moment generating function of PIML distribution can be computed as

$$M_x(t) = \int_0^\infty e^{tx} f(x) dx = \sum_{p=0}^\infty \frac{t^p}{p!} \eta^{p/\tau} \Gamma\left(1 - \frac{p}{\tau}\right) \left(1 - \frac{2^{p/\tau-1}}{1+\eta} \left(\frac{p}{\tau}\right)\right).$$

The characteristic function of PML distribution, $\phi(t) = E(e^{itx})$, and the cumulant generating function of X , $K(t) = \log \phi(t)$, are given by

$$\phi_x(t) = \sum_{p=0}^\infty \frac{(it)^p}{p!} \eta^{p/\tau} \Gamma\left(1 - \frac{p}{\tau}\right) \left(1 - \frac{2^{p/\tau-1}}{1+\eta} \left(\frac{p}{\tau}\right)\right),$$

and

$$K(t) = \log \left(\sum_{p=0}^\infty \frac{(it)^p}{p!} (\eta)^{p/\tau} \right) + \log \left(\Gamma\left(1 - \frac{p}{\tau}\right) \left(1 - \frac{2^{p/\tau-1}}{1+\eta} \left(\frac{p}{\tau}\right)\right) \right).$$

Table 1: Numerical values related to the moments of the PIML distribution for different values of parameters τ and η .

τ	η	$E(Y^{-1})$		$E(Y^{-2})$		$E(Y^{-3})$		$E(Y^{-4})$	
		Sim.	Exact	Sim.	Exact	Sim.	Exact	Sim.	Exact
2	0.1	3.2638	3.2529	12.3493	12.2727	52.6211	52.1709	248.0657	245.4545
	1	0.9622	0.9646	1.1202	1.1250	0.5104	0.5115	2.2327	2.2500
	2	0.6637	0.6636	0.5413	0.5417	1.4967	1.5056	0.5392	0.5417
	3	0.5332	0.5343	0.3526	0.3542	0.2712	0.2728	0.2347	0.2361
	4	0.4585	0.4588	0.2621	0.2625	0.1746	0.1750	0.1309	0.1313
	5	0.4074	0.4080	0.2076	0.2083	0.1235	0.1242	0.0829	0.0833
	10	0.2843	0.2848	0.1020	0.1023	0.0429	0.0431	0.0203	0.0205
	15	0.2312	0.2314	0.0677	0.0677	0.0232	0.0233	0.0090	0.0090
	30	0.1631	0.1627	0.0338	0.0336	0.0082	0.0082	0.0023	0.0022
3	0.1	2.1535	2.1552	4.9804	4.9901	12.2332	12.2727	31.6766	31.8211
	1	0.9525	0.9520	0.9985	0.9975	1.1270	1.1250	1.3522	1.3481
	2	0.7387	0.7400	0.6071	0.6085	0.5407	0.5417	0.5138	0.5142
	3	0.6388	0.6396	0.4553	0.4568	0.3523	0.3542	0.2911	0.2934
	4	0.5778	0.5774	0.3738	0.3733	0.2630	0.2625	0.1979	0.1974
	5	0.5332	0.5337	0.3190	0.3195	0.2079	0.2083	0.1451	0.1454
	10	0.4196	0.4195	0.1983	0.1982	0.1024	0.1023	0.0567	0.0566
	15	0.3648	0.3651	0.1501	0.1504	0.0676	0.0677	0.0326	0.0327
	30	0.2890	0.2886	0.0944	0.0941	0.0337	0.0336	0.0130	0.0129

3.2. Conditional moment, mean deviation, mean residual life and Bonferroni and Lorenz curves

For the PML distribution, it can be easily seen that the conditional moments $E[X^n|X > t]$, can be written as $E[X^n|X > t] = \frac{1}{S(x)}\mu'_n(t)$, where

$$\begin{aligned}
 \mu'_n(t) &= E(X^n) = \int_t^\infty x^n f(x) dx = \int_t^\infty x^n \frac{\eta}{1+\eta} \frac{\tau e^{-\frac{2\eta}{x^\tau}}}{x^{\tau+1}} \left((1+\eta)e^{\frac{\eta}{x^\tau}} + \frac{2\eta}{x^\tau} - 1 \right) dx \\
 &= \tau\eta \int_t^\infty x^{n-\tau-1} e^{-\frac{\eta}{x^\tau}} dx + \frac{\tau\eta}{(1+\eta)} \left(2\eta \int_t^\infty x^{n-2\tau-1} e^{-\frac{2\eta}{x^\tau}} dx - \int_t^\infty x^{n-\tau-1} e^{-\frac{2\eta}{x^\tau}} dx \right) \\
 &= \eta^{n/\tau} \gamma\left(\frac{\eta}{t^\tau}, 1 - \frac{n}{\tau}\right) + \frac{\eta^{n/\tau} 2^{n/\tau-1}}{1+\eta} \left(\gamma\left(\frac{2\eta}{t^\tau}, 2 - \frac{n}{\tau}\right) - \gamma\left(\frac{2\eta}{t^\tau}, 1 - \frac{n}{\tau}\right) \right). \tag{8}
 \end{aligned}$$

The MRL function in terms of the first conditional moment as

$$\eta_1(t) = E[X|x > t] = \frac{\mu'_1(t)}{S(x)},$$

where $\mu'_1(t)$ can be obtained from (8) where $n = 1$.

If we denote the median by M , then the mean deviations from the mean and the median can be calculated as

$$\begin{aligned}
 \delta_{\mu'_1} &= 2\mu'_1 F(\mu'_1) - 2\mu'_1 + 2 \int_{\mu'_1}^\infty x f(x) dx = 2\mu'_1 F(\mu'_1) - 2\mu'_1 \\
 &+ \tau\eta^2(1+\eta)^{i+1} \sum_{(k,l) \in J} \sum_{r=0}^\infty \sum_{i=0}^r \sum_{z=0}^{z+1} \sum_{y=0}^y \binom{k+l}{r+1} \binom{i}{i} \binom{z+1}{z} \binom{z+1}{y} \\
 &\times (r+1) W_{k,l} \frac{(-1)^{r+i} \eta^z \Gamma\left(\frac{1}{\tau} + y + 1, \mu'_1\right)}{(1+\eta)^{i+1} [\eta i + \eta]^{\frac{1}{\tau} + y + 1}}.
 \end{aligned}$$

Similarly, the mean deviation of median (δ_M) is obtained as follows

$$\delta_M = 2MF(M) - M - \mu'_1 + 2 \int_M^\infty xf(x)dx$$

and by using the steps used to solve the integral , we get

$$\delta_M = 2MF(M) - M - \mu'_1 + 2\tau\eta^2 \sum_{(k,l) \in J} \sum_{r=0}^\infty \sum_{i=0}^r \sum_{z=0}^i \sum_{y=0}^{z+1} \binom{k+l}{r+1} \binom{r}{i} \binom{i}{z} \binom{z+1}{y} \\ \times (r+1)W_{k,l} \frac{(-1)^{r+i}\eta^z\Gamma(\frac{1}{\tau} + y + 1, \mu'_1)}{(1 + \delta)^{i+1}[\eta i + \eta]^{\frac{1}{\tau} + y + 1}}.$$

respectively. Where $\mu'_1(\mu)$ and $\mu'_1(M)$ can obtained from (8). Also, $F(\mu)$ and $F(M)$ are easily calculated from (1).

The Bonferroni and Lorenz curves are defined as

$$B(P) = \frac{1}{P\mu} \int_0^Q xf(x)dx \text{ and } L(P) = \frac{1}{\mu} \int_0^Q xf(x)dx,$$

respectively, where $Q = F^{-1}(P)$. The Bonferroni and Gini indices are defined by

$$B = 1 - \int_0^1 B(P)dP \text{ and } G = 1 - 2 \int_0^1 L(P)dP,$$

respectively. If X has the pdf in (2), then one can obtain Bonferroni curve of the MPL distribution as By replacing $n=1$ and $t=q$ in (8) we get-

$$B(P) = \frac{\eta^{1/\tau}}{P\mu} \left(\Gamma\left(\frac{\eta}{q^\tau}, 1 - \frac{1}{\tau}\right) + \frac{\eta^{1/\tau}2^{1/\tau-1}}{1 + \eta} \left(\Gamma\left(\frac{2\eta}{q^\tau}, 2 - \frac{1}{\tau}\right) - \Gamma\left(\frac{2\eta}{q^\tau}, 1 - \frac{1}{\tau}\right) \right) \right) \tag{9}$$

and the Lorenz curves $L(p) = pB(p)$.

3.3. Entropy

If X is a continuous random variable having probability density function $f(\cdot)$, then Renyi entropy is defined as

$$R_r = \frac{1}{1-r} \log \left(\int_0^\infty f^r(x)dx \right), \quad r \neq 1, r > 0 \\ = \frac{1}{1-r} \log \left(\tau^{r-1} \eta^{\frac{1-r}{\tau}} \sum_{i=0}^r \sum_{j=0}^i (-1)^j \binom{r}{i} \binom{i}{j} \frac{2^{i-j} \Gamma(i-j+r+\frac{r-1}{\tau})}{(1+\eta)^i (r+i)^{i-j+r-\frac{r-1}{\tau}}} \right). \tag{10}$$

The r -entropy, say $I_r(x)$, is defined by

$$I_r(x) = \frac{1}{1-r} \log \left(1 - \int_0^\infty f^r(x)dx \right), \quad r \neq 1, r > 0$$

and then it follows from equation (10).

3.4. Stress-strength Reliability

The stress-strength reliability for PIML random variables $X \sim PIML(\tau_1, \eta_1)$ and $Y \sim PIML(\tau_2, \eta_2)$ is given by

$$\begin{aligned}
 R = P(X_2 < X_1) &= \int_0^\infty F_2(x)f_1(x)dx = 1 - \int_0^\infty F_2(x)f_1(x)dx \\
 &= 1 - \left(\sum_{i=0}^\infty \frac{(-1)^i}{i!} \left(\frac{\eta_2}{(\eta_1)^{\tau_2/\tau_1}} \right)^i \Gamma\left(\frac{i\tau_2}{\tau_1} + 1\right) + \frac{1}{2(1+\eta_1)} \sum_{i=0}^\infty \frac{(-1)^i}{i!} \left(\frac{\eta_2}{(2\eta_1)^{\frac{\tau_2}{\tau_1}}} \right)^i \right. \\
 &\times \Gamma\left(\frac{i\tau_2}{\tau_1} + 2\right) - \frac{1}{2(1+\eta_1)} \sum_{i=0}^\infty \frac{(-1)^i}{i!} \left(\frac{\eta_2}{(2\eta_1)^{\tau_2/\tau_1}} \right)^i \Gamma\left(\frac{i\tau_2}{\tau_1} + 1\right) \\
 &+ \left(\frac{\eta_2}{1+\eta_2} \right) \left(\frac{1}{\eta_1^{\tau_2/\tau_1}} \right) \sum_{i=0}^\infty \frac{(-1)^i}{i!} \left(\frac{2\eta_2}{(\eta_1)^{\tau_2/\tau_1}} \right)^i \Gamma\left(\frac{(i+1)\tau_2}{\tau_1} + 1\right) \\
 &+ \left(\frac{\eta_1\eta_2}{(1+\eta_1)(1+\eta_2)} \right) \frac{1}{(2\eta_1)^{\frac{\tau_2}{\tau_1}+1}} \sum_{i=0}^\infty \frac{(-1)^i}{i!} \left(\frac{2\eta_2}{(2\eta_1)^{\tau_2/\tau_1}} \right)^i \Gamma\left(\frac{(i+1)\tau_2}{\tau_1} + 2\right) \\
 &\left. - \left(\frac{\eta_2}{2(1+\eta_1)(1+\eta_2)} \right) \frac{1}{(2\eta_1)^{\tau_2/\tau_1}} \sum_{i=0}^\infty \frac{(-1)^i}{i!} \left(\frac{2\eta_2}{(2\eta_1)^{\tau_2/\tau_1}} \right)^i \Gamma\left(\frac{(i+1)\tau_2}{\tau_1} + 1\right) \right).
 \end{aligned}$$

3.5. Order statistics

Let X_1, X_2, \dots, X_n be a random sample of size n from the PIML distribution and $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the corresponding order statistics. The probability density function of the r th order statistics is obtained as follow:

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} [F(x)]^{r-1} [1-F(x)]^{n-r} f(x).$$

For the PIML distribution, the pdf of r^{th} order statistic is obtained as

$$\begin{aligned}
 f_{r:n}(x) &= \frac{n!}{(r-1)!(n-r)!} \sum_{i=0}^{n-r} \sum_{j=0}^{r+i-1} \binom{n-r}{i} \binom{r+i-1}{j} (-1)^i \left(\frac{\eta}{1+\eta} \right)^{j+1} \frac{1}{x^{\tau(j+1)}} \\
 &\times e^{\frac{-2\eta(j+1)}{x^\tau}} e^{\frac{-\eta(r+i-1)}{x^\tau}} \left((1+\eta)e^{\frac{\eta}{x^\tau}} + \frac{2\eta}{x^\tau} - 1 \right).
 \end{aligned}$$

The r^{th} ordered moment is obtained as

$$\begin{aligned}
 \mu_{r:n}(x) &= \int_0^\infty x f_{r:n}(x) dx = \frac{n!}{(r-1)!(n-r)!} \sum_{i=0}^{n-r} \sum_{j=0}^{r+i-1} \binom{n-r}{i} \binom{r+i-1}{j} (-1)^i \\
 &\times \left(\frac{\eta}{1+\eta} \right)^{j+1} \left(\frac{(1+\eta)}{\eta^{\frac{\tau j-1}{\tau}+1}} \frac{\Gamma\left(\frac{\tau j-1}{\tau} + 1\right)}{(2j+r+i)^{\frac{\tau j-1}{\tau}+1}} + \frac{2}{\eta^{\frac{\tau j-1}{\tau}+1}} \frac{\Gamma\left(\frac{\tau j-1}{\tau} + 1\right)}{(2j+r+i+1)^{\frac{\tau j-1}{\tau}+2}} \right. \\
 &\left. - \frac{1}{\eta^{\frac{\tau j-1}{\tau}+1}} \frac{1}{(2j+r+i+1)^{\frac{\tau j-1}{\tau}+1}} \Gamma\left(\frac{\tau j-1}{\tau} + 1\right) \right).
 \end{aligned}$$

4. PARAMETRIC ESTIMATION OF THE PARAMETERS OF PIML DISTRIBUTION

Here, in this Section, we have derived the classical and the Bayesian point and interval estimation of the model parameters, respectively.

4.1. Classical estimation

Let x_1, x_2, \dots, x_n be a random sample of size n from the PIML distribution. Then, the likelihood function is given by

$$\begin{aligned} L = \prod_{i=1}^n f(x_i) &= \prod_{i=1}^n \frac{\tau\eta}{1+\eta} \frac{e^{-2\eta/x_i^\tau}}{x_i^{\tau+1}} \left((1+\eta)e^{\eta/x_i^\tau} + \frac{2\eta}{x_i^\tau} - 1 \right) \\ &= \frac{\tau^n \eta^n}{(1+\eta)^n} e^{-2\eta \sum_{i=1}^n \frac{1}{x_i^\tau}} \prod_{i=1}^n \left((1+\eta)e^{\eta/x_i^\tau} + \frac{2\eta}{x_i^\tau} - 1 \right) \prod_{i=1}^n \frac{1}{x_i^{\tau+1}} \end{aligned}$$

The corresponding log-likelihood function is

$$\begin{aligned} \ln L &= n \ln(\tau) + n \ln(\eta) - n \ln(1+\eta) - 2\eta \sum_{i=1}^n \frac{1}{x_i^\tau} + \\ &\quad \sum_{i=1}^n \ln \left((1+\eta)e^{\eta/x_i^\tau} + \frac{2\eta}{x_i^\tau} - 1 \right) - \sum_{i=1}^n \ln(x_i^{\tau+1}) \end{aligned}$$

The maximum likelihood estimates of η and τ can be obtained by solving the following non-linear equations:

$$\begin{aligned} \frac{\partial \ln L}{\partial \eta} &= \frac{n}{\eta(1+\eta)} - 2 \sum_{i=1}^n \frac{1}{x_i^\tau} + \sum_{i=1}^n \frac{x_i^\tau e^{\eta/x_i^\tau} + (1+\eta)e^{\eta/x_i^\tau} + 2}{x_i^\tau (1+\eta)e^{\eta/x_i^\tau} + 2\eta - x_i^\tau} = 0, \\ \frac{\partial \ln L}{\partial \tau} &= \frac{n}{\tau} + 2\eta \sum_{i=1}^n x_i^\tau \ln(x_i) - \sum_{i=1}^n \frac{x_i^{2\tau} \eta (1+\eta) e^{\eta/x_i^\tau} \ln(x_i) + 2\eta x_i^{2\tau} \ln(x_i)}{x_i^\tau (1+\eta)e^{\eta/x_i^\tau} + 2\eta - x_i^\tau} \\ &\quad - \sum_{i=1}^n \ln(x_i) = 0. \end{aligned}$$

To solve the above equations, non-linear optimization methods such as the quasi-Newton algorithm can be used to obtain the MLEs of τ and η and are denoted by $\hat{\tau}_{mle}$ and $\hat{\eta}_{mle}$. To estimate δ and γ , we use two methods of estimation, namely maximum likelihood method and Bayesian method. Bayesian estimation method will be discussed in the subsequent Section.

Bootstrap confidence interval

Here, we provide a detailed method for constructing the CIs based on bootstrap method. Here, we consider four CIs based on bootstrap methods: (i) standard bootstrap (SB), (ii) percentile bootstrap (PB), (iii) bias-corrected percentile bootstrap (BCPB), and (iv) bias-corrected accelerated bootstrap (BCAB). Below, we provide the algorithm for construction of the bootstrap CIs based on method of maximum likelihood.

1. Let (X_1, X_2, \dots, X_n) be a random sample of size n drawn from PIML(η, τ). $(\hat{\eta}_{mle}, \hat{\tau}_{mle})$ of (η, τ) . A bootstrap sample of size n is obtained from the original sample by multiplying $1/n$ as mass at each point, denoted by $(X_1^*, X_2^*, \dots, X_n^*)$.
2. Compute the MLEs $(\hat{\eta}_{mle}^*, \hat{\tau}_{mle}^*)$ of (η, τ) . The M -th bootstrap estimator of (η, τ) are computed as

$$\begin{aligned} \hat{\eta}_{mle}^{*(M)} &= \hat{\eta}_{mle} \left(X_1^{*(M)}, X_2^{*(M)}, \dots, X_n^{*(M)} \right) \\ \hat{\tau}_{mle}^{*(M)} &= \hat{\tau}_{mle} \left(X_1^{*(M)}, X_2^{*(M)}, \dots, X_n^{*(M)} \right) \end{aligned}$$

3. There are total number of n^n re-samples. From these re-samples, the entire collection of R values of $\hat{\eta}_{mle}^*, \hat{\tau}_{mle}^*$ from smallest to largest would constitute an empirical bootstrap distribution as:

$$\begin{aligned} &\left\{ \hat{\eta}_{mle}^{*(I)}; I = 1(1)R \right\} \\ &\left\{ \hat{\tau}_{mle}^{*(I)}; I = 1(1)R \right\} \end{aligned}$$

SB

Let

$$\bar{\eta}_{mle}^* = \frac{1}{R} \sum_{I=1}^R \hat{\eta}_{mle}^{*(I)}, \quad s(\hat{\eta}_{mle}^*) = \sqrt{\frac{1}{(R-1)} \sum_{I=1}^R \left(\hat{\eta}_{mle}^{*(I)} - \bar{\eta}_{mle}^* \right)^2},$$

$$\bar{\tau}_{mle}^* = \frac{1}{R} \sum_{I=1}^R \hat{\tau}_{mle}^{*(I)}, \quad s(\hat{\tau}_{mle}^*) = \sqrt{\frac{1}{(R-1)} \sum_{I=1}^R \left(\hat{\tau}_{mle}^{*(I)} - \bar{\tau}_{mle}^* \right)^2}$$

be the sample means and standard deviations of $\left\{ \hat{\eta}_{mle}^{*(I)}; I = 1(1)R \right\}$, $\left\{ \hat{\tau}_{mle}^{*(I)}; I = 1(1)R \right\}$, respectively. Then, $100(1 - \gamma)\%$ SB confidence interval of (η, τ) are given as:

$$\left\{ \bar{\eta}_{mle}^* - Z_{(\gamma/2)} \times s(\hat{\eta}_{mle}^*), \bar{\eta}_{mle}^* + Z_{(\gamma/2)} \times s(\hat{\eta}_{mle}^*) \right\},$$

$$\left\{ \bar{\tau}_{mle}^* - Z_{(\gamma/2)} \times s(\hat{\tau}_{mle}^*), \bar{\tau}_{mle}^* + Z_{(\gamma/2)} \times s(\hat{\tau}_{mle}^*) \right\},$$

where $Z_{(\gamma/2)}$ is obtained by using upper $(\gamma/2)$ -th point of the standard normal deviate.

PB

Let $\hat{\eta}_{mle}^{*(\xi)}$, $\hat{\tau}_{mle}^{*(\xi)}$ are the ξ percentile of $\left\{ \hat{\eta}_{mle}^{*(I)}; I = 1(1)R \right\}$, $\left\{ \hat{\tau}_{mle}^{*(I)}; I = 1(1)R \right\}$, respectively. Then, a $100(1 - \gamma)\%$ PB confidence interval of (τ, η) are given as:

$$\left\{ \hat{\tau}_{mle}^{*(R \times (\gamma/2))}, \hat{\tau}_{mle}^{*(R \times (1-\gamma/2))} \right\},$$

$$\left\{ \hat{\eta}_{mle}^{*(R \times (\gamma/2))}, \hat{\eta}_{mle}^{*(R \times (1-\gamma/2))} \right\},$$

respectively.

To study the different confidence intervals, we consider their estimated average widths (AWs) and coverage probabilities (CPs) for each of the considered methods and are given as

$$AW(\tau) = \frac{\sum_{i=1}^K (U_{cli} - L_{cli})}{K} \quad \text{and} \quad CP(\tau) = \frac{\text{number}(L_{cl} \leq \tau \leq U_{cl})}{K},$$

$$AW(\eta) = \frac{\sum_{i=1}^K (U_{cli} - L_{cli})}{K} \quad \text{and} \quad CP(\eta) = \frac{\text{number}(L_{cl} \leq \eta \leq U_{cl})}{K}.$$

4.2. Bayesian estimation

As a powerful and valid alternative to classical estimation, the Bayesian approach suggests a procedure to combine the observed information with the prior knowledge. Here, for the purpose of framing the Bayesian analysis, we set assumptions as:

$$\tau \sim \text{Gamma}(\tau_0, \tau_1), \quad \eta \sim \text{Gamma}(\eta_0, \eta_1).$$

We now consider several (symmetric and asymmetric) loss functions (LS), namely, SELF, WSELF, MSELF, and PLF. These loss functions with corresponding Bayesian estimators (BS) and posterior risks (PR) are provided in Table 2.

Table 2: Five loss functions with corresponding BS and PR.

LS: $L(\psi, \delta)$	BS of parameter ψ_B	PR of parameter ρ_ψ
$SELF = (\psi - d)^2$	$E(\psi x)$	$Var(\psi x)$
$WSELF = \frac{(\psi-d)^2}{\psi}$	$(E(\psi^{-1} x))^{-1}$	$E(\psi x) - (E(\psi^{-1} x))^{-1}$
$MSELF = \left(1 - \frac{d}{\psi}\right)^2$	$\frac{E(\psi^{-1} x)}{E(\psi^{-2} x)}$	$1 - \frac{E(\psi^{-1} x)^2}{E(\psi^{-2} x)}$
$PLF = \frac{(\psi-d)^2}{d}$	$\sqrt{E(\psi^2 x)}$	$2 \left(\sqrt{E(\psi^2 x)} - E(\psi x)\right)$

Posterior distributions

The joint prior distribution of parameters τ and η under the independent prior distributions

$$\tau \sim Gamma(\tau_0, \tau_1), \eta \sim Gamma(\eta_0, \eta_1),$$

is given as

$$\pi(\tau, \eta) = \frac{\tau_1^{\tau_0} \eta_1^{\eta_0}}{\Gamma(\tau_0)\Gamma(\eta_0)} \tau^{\tau_0-1} \eta^{\eta_0-1} e^{-(\tau_1\tau + \eta_1\eta)}, \tag{11}$$

where all the hyper-parameters τ_0, τ_1, η_0 and η_1 are positive. Now, let ζ be

$$\zeta(\tau, \eta) = \tau^{\tau_0-1} \eta^{\eta_0-1} e^{-(\tau_1\tau + \eta_1\eta)}, \tau > 0, \eta > 0,$$

then, the joint posterior distribution is proportional to the joint prior distribution $\pi(\tau, \eta)$ and a given likelihood function $L(data)$ as

$$\pi^*(\tau, \eta|data) \propto \pi(\tau, \eta)L(data). \tag{12}$$

In the case of PML distribution, the exact joint posterior PDF of parameters τ and η , is given by

$$\pi^*(\tau, \eta|x) = CL(x, Y)\zeta(\tau, \eta) \tag{13}$$

where

$$L(x; Y) = \frac{\tau^n \eta^n}{(1 + \eta)^n} e^{-2\eta \sum_{i=1}^n \frac{1}{x_i^\tau}} \prod_{i=1}^n \left[(1 + \eta)e^{\eta/x_i^\tau} + \frac{2\eta}{x_i^\tau} - 1 \right] \prod_{i=1}^n \frac{1}{x_i^{\tau+1}}, \tag{14}$$

$Y = (\tau, \eta)$ and K is normalizing constant and is given by

$$C^{-1} = \int_0^\infty \int_0^\infty L(x, Y)\zeta(\tau, \eta)\partial\eta\partial\tau.$$

Consequently, the marginal posterior PDF for the elements of vector Y with $Y = (Y_1, Y_2) = (\tau, \eta)$, is given by

$$\pi(Y_i|x) = \int_0^\infty \pi^*(Y|x)\partial Y_j, \tag{15}$$

where $i, j = 1, 2, i \neq j$ and Y_i is the i th element of vector parameter Y .

Generating posterior samples

Let $f(x|v)$ be a general PDF that is labeled with parameter vector $v = (v_1, v_2, \dots, v_p)$. Based on a given sample x and initial parameter vector $v_0 = (v_1^{(0)}, v_2^{(0)}, \dots, v_p^{(0)})$, the Gibbs sampler gives the values for each iteration with p steps by extracting a new value for each parameter from its full conditional PDF. In symbols, the steps for each iteration (iteration l), are as follows:

- Set an initial parameter vector $(v_1^{(0)}, v_2^{(0)}, \dots, v_p^{(0)})$
- Extract v_1^l from $\pi(v_1|v_2^{l-1}, v_3^{l-1}, \dots, v_p^{l-1}, \underline{x})$
- Extract v_2^l from $\pi(v_2|v_1^l, v_3^{l-1}, \dots, v_p^{l-1}, \underline{x})$; and so on down to
- Extract v_p^l from $\pi(v_p|v_1^l, v_2^l, \dots, v_{p-1}^l, \underline{x})$.

Making use the above GS algorithm, the posterior samples of the parameters τ and η of PML distribution are generated from the full conditional posterior PDFs

$$\pi(\tau|\eta^{k-1}, \underline{x}) \propto \tau^{\tau_0+n-1} e^{-\tau_1\tau} \prod_{i=1}^n \left((1+\eta)x^{\tau-1} e^{\eta x_i^\tau} + 2\eta x_i^{2\tau-1} - x_i^{\tau-1} \right) e^{-2\eta x_i^\tau}$$

and

$$\pi(\eta|\tau^{k-1}, \underline{x}) \propto \frac{\eta^{\eta_0+n+1} e^{-\eta_1\eta}}{(1+\eta)^n} \prod_{i=1}^n \left((1+\eta)x^{\tau-1} e^{\eta x_i^\tau} + 2\eta x_i^{2\tau-1} - x_i^{\tau-1} \right) e^{-2\eta x_i^\tau},$$

respectively.

5. COMPARISON VIA MONTE-CARLO SIMULATION

Here, we have carried out a Monte Carlo simulation study to compare the performances of the classical and the Bayesian methods of estimation of the parameters (τ, η) of PIML distribution. The performance of the estimates (classical as well as Bayes) are compared in terms of their MSEs and posterior risks, respectively. Also, we have obtained four BCIs, namely, SB, PB, BCPB and BCAB and high posterior density (HPD) credible intervals, respectively. The performance of the CIs are compared in terms of their AWs and CPs. Here, for the simulation study, we have considered the sample sizes $n = 20, 30, 50, 100$ and $(\tau, \eta) = (0.5, 2.0), (1.0, 2.0), (0.5, 3.0), (1.0, 3.0), (2.0, 2.0)$, respectively. For each of the designs, $\mathcal{R} = 1,000$ bootstrap samples each of size n are drawn from the original sample and replicated $\mathcal{K} = 1,000$ times.

This section presents Monte Carlo simulation results to assess the performance of MLE mentioned in the previous section. First, we generate different samples with size n from (1) based upon the inversion method. We compute the mean square errors (MSEs) and biases of the MLEs of the parameters based on $N = 10,000$ iterations. The results are summed up in Table 2 for some selected parameter values and several sample sizes, n . The results in Table 2 indicate that the MSEs and biases of the MLEs decrease when the sample size n increases. So, the MLEs of the parameters are consistent.

5.1. Simulation results using mean squared errors, Bayes risks and nominal coverage probability as the criterion.

This section is devoted to calculate posterior risk values of Bayes estimators under different loss functions based on Monte Carlo simulation. We generated samples of different sizes $n = \{30, 50, 75, 100\}$ from the PIML distribution for true value of parameters (i) $(\tau, \eta) = (2, 0.5)$ and (ii) $(\tau, \eta) = (1, 2)$. Table 3 reports the posterior risk values of Bayes estimators under prior distributions defined in (11) and the aforementioned five loss functions as shown in Table 1. These results provided by considering hyper parameters values as $(\tau_0, \tau_1) = (2, 1), (\eta_0, \eta_1) = (4, 2)$ for case (i) and $(\tau_0, \tau_1) = (10, 1), (\eta_0, \eta_1) = (1, 2)$ and for case (ii) based on 10000 replicates with 1000 burn-in of MCMC procedure in Open BUGS software. It is evident from Table 4 that with increasing sample size n , the posterior risk decreases and this confirms the consistency property. We also observe that as n increases, Bayes estimate of τ based on KL loss function provide superior performance than other Bayes estimates whereas Bayes estimate of η based on PL loss function perform better than other loss functions as η decreases.

Table 3: AE, MSE, AW and CP of BCI of the model parameters τ and η by using MLE.

Sample size	"	$\hat{\tau}_{mle}$				BCI($\hat{\tau}_{mle}$)				$\hat{\eta}_{mle}$				BCI($\hat{\eta}_{mle}$)			
		AE	MSE	AW	SB	CP	AW	PB	CP	AW	MSE	AW	SB	CP	AW	PB	CP
20	0.5	0.64620	0.0397	0.590470	0.443	0.59545	0.64700	2	2.62260	0.62260	3.10983	0.11200	3.38438	1.00000			
30	0.5	0.63180	0.0285	0.447900	0.335	0.45046	0.50600	2	2.57140	0.44800	2.14232	0.01400	2.20525	0.51400			
50	0.5	0.61780	0.0197	0.327200	0.199	0.32802	0.28500	2	2.55280	0.36780	1.47214	0.00100	1.49041	0.00000			
100	0.5	0.61200	0.0153	0.225480	0.033	0.22584	0.04300	2	2.57160	0.35310	0.96979	0.00000	0.97314	0.00000			
20	1	1.29500	0.1555	1.170150	0.446	1.17866	0.67000	2	2.60850	0.59370	3.15528	0.11100	3.45266	1.00000			
30	1	1.26660	0.1148	0.881140	0.364	0.88604	0.52000	2	2.57540	0.46550	2.14217	0.00600	2.21428	0.51100			
50	1	1.23550	0.0803	0.659850	0.183	0.66214	0.25000	2	2.54390	0.36330	1.47791	0.00000	1.48917	0.00500			
100	1	1.22060	0.0605	0.449170	0.026	0.45009	0.03800	2	2.55780	0.33290	0.97038	0.00000	0.97416	0.00000			
20	0.5	0.56860	0.0163	0.481160	0.808	0.48511	0.92200	3	3.67660	1.48850	5.83612	0.81000	6.66360	1.00000			
30	0.5	0.55520	0.0105	0.363930	0.775	0.36614	0.88700	3	3.52840	0.85260	3.63643	0.76200	3.80360	1.00000			
50	0.5	0.54990	0.0068	0.266610	0.717	0.26766	0.80200	3	3.43990	0.45950	2.38948	0.68000	2.43261	0.96200			
100	0.5	0.53920	0.0034	0.180740	0.581	0.18113	0.64600	3	3.33840	0.21570	1.48682	0.46700	1.49615	0.71100			
20	1	1.13460	0.0661	0.967300	0.787	0.97562	0.92300	3	3.66990	1.62120	5.79667	0.80300	6.73772	1.00000			
30	1	1.11470	0.0405	0.724170	0.791	0.72811	0.89100	3	3.51640	0.78410	3.65657	0.76300	3.82785	1.00000			
50	1	1.09950	0.0258	0.530040	0.742	0.53161	0.82200	3	3.41530	0.41190	2.36034	0.69100	2.39949	0.96600			
100	1	1.07630	0.0137	0.360450	0.604	0.36163	0.66900	3	3.33670	0.21360	1.48351	0.46600	1.49077	0.71700			
20	2	2.57430	0.6227	2.326790	0.482	2.34762	0.67400	2	2.63590	0.65470	3.13810	0.10500	3.43835	1.00000			
30	2	2.52970	0.4622	1.795760	0.337	1.80633	0.47500	2	2.58900	0.47190	2.13065	0.01400	2.19423	0.51300			
50	2	2.47730	0.3309	1.317320	0.176	1.32084	0.26200	2	2.56400	0.38580	1.47180	0.00000	3.74158	0.00400			
100	2	2.45010	0.2468	0.894830	0.033	0.89621	0.05100	2	2.56410	0.34630	0.96020	0.00000	0.96310	0.00000			

Table 4: Posterior risk values of Bayesian estimators under different loss functions based on simulation data set for different sample sizes.

n	Loss function	$(\tau, \eta) = (2, 0.5)$		$(\tau, \eta) = (1, 2)$	
		$r_{\hat{\tau}}$	$r_{\hat{\eta}}$	$r_{\hat{\tau}}$	$r_{\hat{\eta}}$
20	SELF	0.183806	0.009075	0.042515	0.154721
	WSELF	0.071691	0.026926	0.030225	0.077895
	MSELF	0.030278	0.098003	0.022891	0.043705
	PLF	0.068412	0.025424	0.029432	0.075577
30	SELF	0.065705	0.007869	0.026909	0.085396
	WSELF	0.033654	0.017964	0.021658	0.046136
	MSELF	0.018129	0.046110	0.018359	0.026846
	PLF	0.032622	0.017514	0.021233	0.045574
50	SELF	0.035856	0.007735	0.014231	0.051113
	WSELF	0.020319	0.011983	0.011751	0.027568
	MSELF	0.011905	0.019580	0.009982	0.015532
	PLF	0.020002	0.011839	0.011603	0.027314
100	SELF	0.024259	0.003219	0.010833	0.029405
	WSELF	0.011909	0.006025	0.008129	0.015245
	MSELF	0.005946	0.011653	0.006210	0.008089
	PLF	0.011784	0.005991	0.008063	0.015200

6. APPLICATIONS

In this section, we examine the versatility of the PIML model in comparison with the inverted modified Lindley (IML), modified Lindley (ML) and inverse Lindley (IL) distributions by usage of three real data sets presented below, which are available in [5]. The box plot of the considered data set are displayed in Figure 2. To check the validity of the considered data sets with the proposed model, the goodness-of-fit statistics is considered. Here, we have used built-in package *fitdistrplus* of the R open source software (see, Ihaka and Gentleman (1996)) for goodness-of-fit test. And we derived the unknown parameters by the maximum likelihood estimation (MLE) method, log likelihood function evaluated at the MLEs (\hat{I}), the values of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), the values of the Kolmogorov-Smirnov (K-S) statistic, the corresponding p values and the values of the Anderson-Darling (AD) and Cramér von Mises (CM) are compared with IML, IL and also are reported in Table 5.

Data set I: This first data set has been analyzed by [19]. The Open University (1993), which relates to the prices of the 31 various children’s wooden toys on sale in a Suffolk craft shop in April 1991, is the source of the first data set. Originally, the data set is: 4.2, 1.12, 1.39, 2, 3.99, 2.15, 1.74, 5.81, 1.7, 0.5, 0.99, 11.5, 5.12, 0.9, 1.99, 6.24, 2.6, 3, 12.2, 7.36, 4.75, 11.59, 8.69, 9.8, 1.85, 1.99, 1.35, 10, 0.65, 1.45.

Data set II: The second data set, which was obtained from [17], includes the intervals between failures for repairable items and the data set is: 1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24, 1.97, 1.86, 1.17.

Data set III: The third actual data set includes 30 iterations of [11] reported March precipitation figures for Minneapolis/St. Paul (in inches). The set of data is: 0.77, 1.74, 0.81, 1.2, 1.95, 1.2, 0.47, 1.43, 3.37, 2.2, 3, 3.09, 1.51, 2.1, 0.52, 1.62, 1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.9, 2.05.

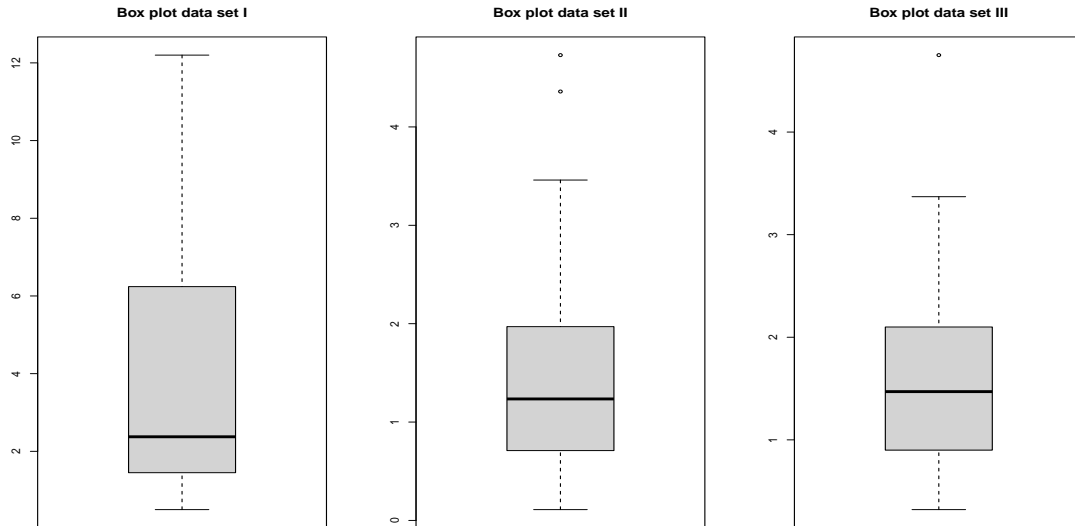


Figure 2: Box plot of the considered data sets I, II and III [[5]].

Table 5: The model fitting summary of the considered data sets I, II and III.

Distribution	n	$(\hat{\tau}, \hat{\eta})$	$-\hat{l}$	AIC	BIC	KS Statistic	p-value	AD	CM
Data Set I									
PIML	30	(1.093,2.233)	73.011	150.023	152.825	0.1017	0.9154	0.4138	0.0546
IML	30	(2.1537)	73.187	148.375	149.776	0.1225	0.7589	0.4082	0.0487
ML	30	(0.2825)	73.00	148.000	149.4016	0.18521	0.2548	0.9004	0.1556
L	30	(0.3999)	73.232	148.464	149.865	0.1832	0.2661	0.8631	0.1478
Data Set II									
PIML	30	(0.955,0.941)	45.227	94.454	97.257	0.12767	0.7124	0.9657	0.1387
IML	30	(0.9201)	45.301	92.603	94.004	0.1404	0.5951	0.9454	0.1405
ML	30	(0.7302)	40.749	83.499	84.901	0.0979	0.9355	0.4283	0.0629
L	30	(0.9767)	41.537	85.0740	86.4752	0.1278	0.7108	0.7125	0.1111
Data Set III									
PIML	30	(1.362,1.222)	41.608	87.216	90.018	0.1392	0.6058	0.6605	0.0985
IML	30	(1.2473)	43.868	89.736	91.137	0.1974	0.1925	1.391	0.217
ML	30	(0.6644)	41.945	85.889	87.291	0.1566	0.4532	1.1278	0.1723
L	30	(0.9096)	43.1437	88.2874	89.6886	0.1882	0.2383	1.5908	0.2618

Table 6: Widths of BCIs of τ and η for the considered data sets I, II and III.

Data sets	τ						η					
	PB			SB			PB			SB		
	L	U	W	L	U	W	L	U	W	L	U	W
I	0.97156	1.74531	0.77375	0.91424	1.71277	0.79853	2.14218	3.78504	1.64285	1.94844	3.58389	1.63544
II	1.23733	3.70488	2.46755	0.80434	3.48136	2.67702	1.42575	2.00554	0.57978	1.39314	1.99245	0.59931
III	1.54637	3.66720	2.12083	1.26633	3.39761	2.13127	1.59218	2.38249	0.79030	1.52205	2.32971	0.80765

Table 7: Bayes estimate of τ and η for the considered data sets I, II and III..

Data sets	τ						η					
	Bayes estimate			Bayes estimate			Bayes estimate			Bayes estimate		
	SELF	WSELF	MSELF	SELF	WSELF	MSELF	SELF	WSELF	MSELF	SELF	WSELF	MSELF
I	0.024264	0.022644	0.022422	0.022129	0.104306	0.049351	0.024988	0.048547	0.024988	0.024988	0.024988	0.048547
II	0.012078	0.013285	0.015249	0.012792	0.021774	0.0218064	0.023251	0.021587	0.023251	0.023251	0.023251	0.021587
III	0.029582	0.022593	0.018129	0.021859	0.033645	0.049351	0.022685	0.026307	0.022685	0.022685	0.022685	0.026307

The MLEs of the parameters given in Table 5. The widths of the BCIs and the Bayes estimates as well as Bayes credible intervals of the model parameters are given in Tables 6 and 5, respectively.

7. CONCLUDING REMARKS

In this article, we have proposed a new probability distribution, namely, PIML distribution by considering the IML distribution. Different statistical characteristics have been deliberated. Maximum likelihood estimates of the models parameters as well bootstrap confidence intervals from classical point of view and the Bayes estimates have been obtained. The consistency of the point and interval estimates have been shown through the simulation study in terms of mean squared errors, average widths and corresponding coverage probabilities. With the lowest values of AIC, BIC, AD, CM, KS and highest values of KS p values among all the competitive models, viz., L, ML and IML, the PIML distribution has been chosen the best fitted model to fit the considered three data sets.

REFERENCES

- [1] Asgharzadeh, A., Bakouch, H. S., Nadarajah, S. and Shara, F. (2016). A New Weighted Lindley Distribution with Application. *Brazilian Journal of Probability and Statistics*, 30, 1–27.
- [2] Ashour SK, Eltehiwy MA (2015) Exponentiated power Lindley distribution. *J Adv Res* 6, 895–905.
- [3] Bakouch, H., Al-Zahrani, B., Al-Shomrani, A., Marchi, V. and Louzada F (2012) An extended lindley distribution. *J Korean Stat Soc* 41(1), 75–85.
- [4] Barreto-SouzaW, Bakouch HS (2013). A new lifetime model with decreasing failure rate. *Statistics*, 47, 465–476.
- [5] Chesneau, C., Tomy, L., Gillariose, J. and Jamal, F. (2020). The inverted modified Lindley distribution, *J Stat Theor Pract*, 14(3), 1–17.
- [6] Chesneau, C., Tomy, L., and Gillariose, J. (2021). A New Modified Lindley Distribution with Properties and Applications." *Journal of Statistics Management and System*, 24, 1383–1403.
- [7] Dey, S., Nassar, M. and Kumar, D. (2019a). Alpha power transformed inverse Lindley distribution: A distribution with an upside-down bathtub-shaped hazard function, *Journal of Computational and Applied Mathematics*, 348 (2019) 130–145.
- [8] Dey, S., Gosh, I., and Kumar, D. (2019b). Alpha-power transformed Lindley distribution: Properties and associated inference with application to earthquake data, *Ann. Data. Sci.*, 6(4), 623–650
- [9] Gupta, R. D. and Kundu, D. (2009). Introduction of shape/skewness parameter(s) in a probability distribution, *Comput. Stat.* 7, 153–171.
- [10] Ghitany ME, Al-Mutairi DK, Balakrishnan N, Al-Enezi LJ (2013) Power Lindley distribution and associated inference. *Comput Stat Data Anal* 6, 20–33.
- [11] Hinkley D (1977) On quick choice of power transformations. *Appl Stat* 26, 67–69
- [12] Joshi S, Jose KK (2018). Wrapped Lindley Distribution." *Communications in Statistics-Theory and Methods*, 47, 1013–1021.
- [13] Kharazmi, O. Kumar, D. and Dey, S. (2023). Power modified Lindley distribution: Properties, classical and Bayesian estimation and regression model with applications, *Austrian Journal of Statistics*, 52, 71–95.
- [14] Kumar, D., Yadav, P. and Kumar, J. (2022a). Classical inferences of order statistics for inverted modified Lindley distribution with applications, *Strength of Materials*, 55, 441–455.
- [15] Kumar, D., Nassar, M., Dey, S. and Diyali, B. (2022b). Analysis of an inverted modified Lindley distribution using dual generalized order statistics, *Strength of Materials*, 54, 889–904.
- [16] Kumar, D. and Kumar, V. (2022c). An extension of exponentiated gamma distribution: A new regression model with application, *Lobachevskii Journal of Mathematics*, 43, 2525–2543.

- [17] Murthy DNP, Xie M, Jiang R (2004). Weibull models. Wiley series in probability and statistics, Wiley, Hoboken.
- [18] Nadarajah, S., Bakouch, H. S. and Tahmasbi, R. (2011). A Generalized Lindley Distribution." *Sankhya B*, 73, 331–359.
- [19] Shafei S, Darijani S, Saboori H (2016) Inverse Weibull power series distributions: properties and applications. *J Stat Comput Simul* 86(6):1069–1094.

A $MAP/PH_1, PH_2/2$ INVENTORY QUEUEING SYSTEM WITH TWO COMMODITY, MULTIPLE VACATION, SERVER FEEDBACK, WORKING BREAKDOWN, REPAIR AND EMERGENCY REPLENISHMENT

G. AYYAPPAN, N. ARULMOZHI



Department of Mathematics,
Puducherry Technological University,
Puducherry, India.

ayyappan@ptuniv.edu.in, arulmozhisathya@gmail.com,

Abstract

We investigate a continuous review inventory queuing system in the present study that has two heterogeneous servers: Server-2, which is reliable, and Server-1, which is unreliable. An exponentially distributed random time is used to describe the repair process when server-1 has an interruption. On the other hand, server-2 is completely dependable, but it goes on vacation when the system is empty. These two goods can be reordered under ordering regulations. To ensure customer satisfaction, an emergency replenishment of one item with no lead time occurs when the on-hand inventory level falls to zero. We use the matrix analytic approach for the QBD process under a steady-state probability vector. We also take into account the overall cost and the busy time. Furthermore, numerical data shows the benefits of the suggested approach in a range of random circumstances.

Keywords: Markovian arrival process, PH-distribution, multiple vacation, two commodity, working breakdown, emergency replenishment.

AMS Subject Classification (2010): 60K25, 68M30, 90B22.

1. INTRODUCTION

Substitution methods are essential for reducing client losses in an inventory-based organization. While the demand-driven item's stock-out time, emergency replenishment may be utilized. Stockouts may be managed and frequently avoided by employing excellent inventory management methods such as accurate demand forecasts, establishing reorder points, utilizing buffers and safety stock, and discovering stockout trends and emergency orders using inventory management software. In circumstances where stockouts are beyond a retailer's control, it's necessary to take efforts to save expenses and prevent dissatisfied consumers, such as proposing or inventing a product substitute or finding an alternate supplier.

The study [8] examined a $(0, S)$ in a service facility with multiple server vacations and impatient clients. The combined probability distribution of the inventory level and the number of customers in the waiting area is calculated under steady-state scenarios. A few system performance metrics are obtained, and the total estimated cost rate is computed. The two-commodity (TC) inventory system, in which commodities are split into primary and supplemental products, was examined by Jeganathan et al. [13]. They examined an individual reordering strategy for a single commodity inventory with no replenishment period and another commodity utilised for the (s, Q) policy. An inventory queuing system with Poisson arrival, randomly

distributed service times, zero lead times, requiring one item per client, and limited waiting room capacity was studied by Berman and Sapna [3].

Kalyanaraman and Senthilkumar [7] studied two heterogeneous server Markovian queues, with the first server's service mode altering at a threshold. If both servers are idle, the consumer will be served by the faster server. Using MAP arrival and group services, Chakravarthy et al. [4] investigated the multi-server finite capacity model. They proved that the invariant distribution of the sojourn time is phase-type if the inter-arrival durations are as well. They also developed mathematical approaches for determining the invariant density of waiting queue size. Yang et al. [15] examined a Markovian queue with two servers, malfunctions, and leaves of absence. They used the usual particle swarm optimization technique for numerical analysis, developed a heuristic cost model, and investigated their system using the matrix approach.

In a recent article, Laxmi and Soujanya [11] discussed QIS with multiple server working vacations. They consider both working vacation and multiple vacation, and numerically analyze the system to determine the best strategy. An (s, Q) replenishment technique was used by Manikandan and Nair [12] to assess a single server QIS while taking breaks and interruptions into consideration. Performance data are examined, the steady-state probability vector is calculated, and the stability condition of the system is established. A busy period analysis and a derived stationary waiting time distribution for the queue are included in the study.

Yadavalli et al. [16] used a finite source QIS with interruption to conduct research on two heterogeneity servers. Jeganathan et al. [6] studied a Markovian inventory system with two queues and customers jockeying. Yadavalli and Jeganathan [17] investigate the impact of two servers and partial vacations on inventory systems. Suganya et al. [14] proposed an inventory system with numerous server vacations, as opposed to partial ones. They did, however, take into account the Markovian Arrival Process (MAP) of clients. A queueing model subject to a single server that can offer two types of heterogeneous services, closedown, vacation, setup, immediate feedback, breakdown, and repair was determined by Ayyappan and Gowthami [2]. Ayyappan and Karpagam [1] have discussed a classical model with unreliable server, vacation and immediate feedback. They have assumed that the unsatisfied customers will be given re-service immediately without any delay. They have found the PGF of the size of the waiting line and some interesting performance measures.

2. MODEL DESCRIPTION

We investigate the two server inventory queueing model, where customers arrive at the system based on MAP. The parameter matrices D_0 and D_1 indicate that no customers have arrived at the system and that customers have arrived, respectively. Dimension m is used by the matrices D_0 and D_1 . We examine two different kinds of heterogeneous servers here: One among them does not go on vacation and the other one goes for a vacation after service completion. That is, the server-1 will always be there in the system but the server-2 will move forward to vacation after service completion to the customers. The service rendering by the server-1 during normal mode with representations (γ_1, U_1) of dimension n_1 such that $U_1^0 + U_1 e = 0$. Likewise, The service rendering by the server-2 during normal mode is follows the PH-type with representation (γ_2, U_2) and of dimension n_2 such that $U_2^0 + U_2 e = 0$. During the server-1 providing service to the customers who may struck with breakdown and the system has an option: continue delivering slow service to current customer. During the working breakdown period, the server provide slow service to current customer and service times are phase type distributed with notation $(\gamma_1, \theta_1 U_1)$; $0 < \theta_1 < 1$ of order n_1 and rate is $\mu_{bd} = [\gamma_1 (-\theta_1 U_1)^{-1} e]^{-1}$.

After service completion during this working breakdown period, the system automatically enters a repair phase. The server will begin a rejuvenating service for the customer, after the repair process. After service completion, no customer in the system server-2 go for multiple vacation. At the end of a vacation, when the customers are staying in the system then the server-2 is provided normal service. Otherwise, server-2 takes another vacation immediately. After received service

from the service station, the satisfied customer would left out of the system with probability p_i , for $i = 1,2$ and if the customer is not satisfied with probability q_i , for $i=1,2$ then they will get feedback immediatly that is, the instantaneous feedback will offer by the same server. The breakdown and repair times of server-1 follows an exponentially distributed with parameter ψ and τ respectively. We analyse a stochastic inventory system with two distinct goods in stock: server-1 (I-commodity) and server-2 (II-commodity). The maximal storage capacity for the commodity i^{th} is S_i ($i = 1,2$). The initial commodity with zero lead time follows a $(0, S_1)$ ordering strategy. When the on-hand inventory level of the second commodity falls below a predetermined level, s_2 , an order for S_2 units is made. The lead time for this order is exponentially distributed with parameter β . Furthermore, suppose an emergency replenishment of one item with zero lead time occurs when the on-hand inventory falls to zero. Emergency replenishment is included into the system to ensure client happiness. The server-2 does not offer feedback service during emergency replenishment. The schematic picture of this model is provided in Figure 1.

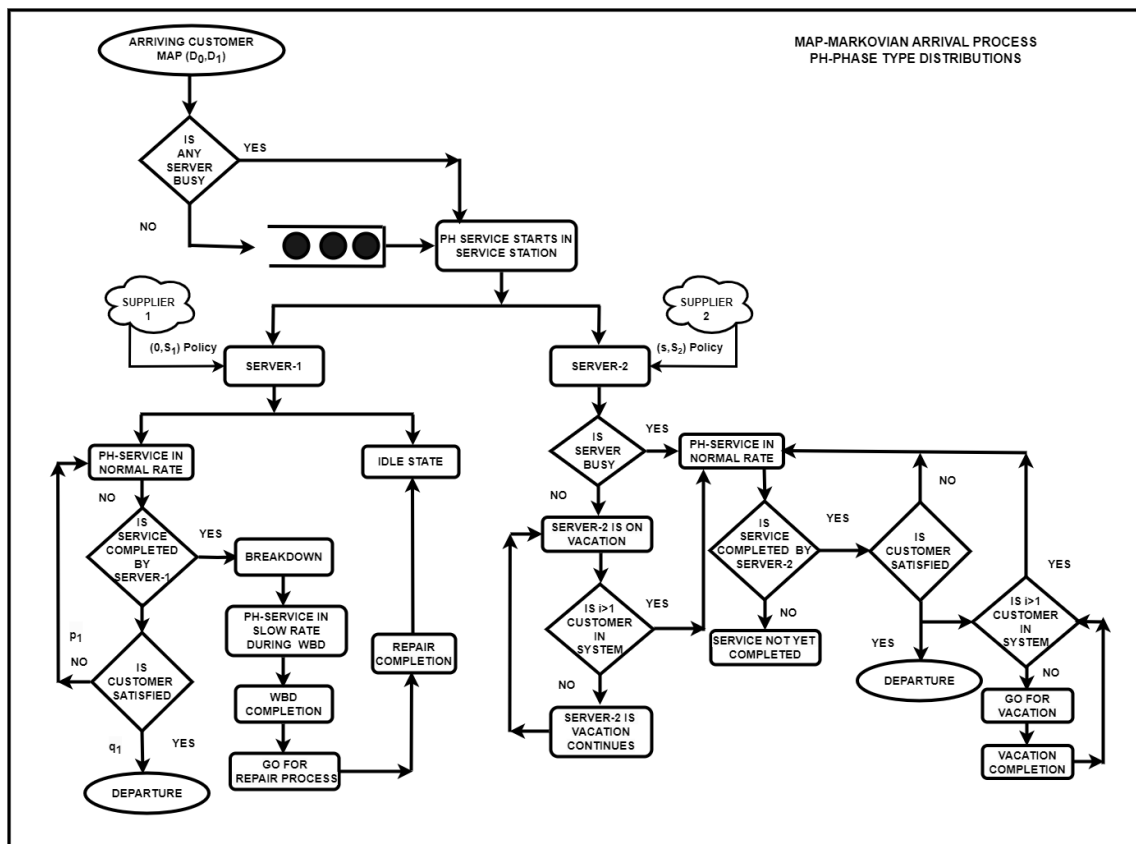


Figure 1: Schematic representation

3. ANALYSIS

In the following section, we establish the queueing-inventory system's transition rate matrix. Assume that $N(t), J_1(t), J_2(t), I_1(t), I_2(t), S_1(t), S_2(t), M(t)$ described total customers in the system, status of server-1, status of server-2, stock level for commodity I, stock level for commodity II, service phase for server-1, service phase for server-2, arrival phases, respectively.

$$J_1(t) = \begin{cases} 0, & \text{server-1 is idle ,} \\ 1, & \text{server-1 is busy ,} \\ 2, & \text{server-1 is busy in WBD mode,} \\ 3, & \text{server-1 is repair,} \end{cases}$$

$$J_2(t) = \begin{cases} 0, & \text{server-2 is vacation,} \\ 1, & \text{server-2 is busy in normal mode,} \end{cases}$$

Consider $X(t) = \{N(t), J_1(t), J_2(t), I_1(t), I_2(t), S_1(t), S_2(t), M(t)\}$ is a CTMC with state space

$$\Phi = \phi(0) \cup \phi(1) \cup_{i=1}^{\infty} \phi(i). \tag{1}$$

where

$$\begin{aligned} \phi(0) = & \{(0, 0, 0, u_1, u_2, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_5 \leq m\} \\ & \cup \{(0, 3, 0, u_1, u_2, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_5 \leq m\} \end{aligned}$$

$$\begin{aligned} \phi(1) = & \{(1, 0, 1, u_1, u_2, u_4, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_4 \leq n_2, 1 \leq u_5 \leq m\} \\ & \cup \{(1, 1, 0, u_1, u_2, u_3, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_3 \leq n_1, 1 \leq u_5 \leq m\} \\ & \cup \{(1, 2, 0, u_1, u_2, u_3, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_3 \leq n_1, 1 \leq u_5 \leq m\} \\ & \cup \{(1, 3, 0, u_1, u_2, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_5 \leq m\} \\ & \cup \{(1, 3, 1, u_1, u_2, u_4, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_4 \leq n_2, 1 \leq u_5 \leq m\} \end{aligned}$$

and for $i \geq 2$,

$$\begin{aligned} \phi(i) = & \{(i, 1, 0, u_1, u_2, u_3, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_3 \leq n_1, 1 \leq u_5 \leq m\} \\ & \cup \{(i, 1, 1, u_1, u_2, u_3, u_4, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_3 \leq n_1, 1 \leq u_4 \leq n_2, 1 \leq u_5 \leq m\} \\ & \cup \{(i, 2, 0, u_1, u_2, u_3, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_3 \leq n_1, 1 \leq u_5 \leq m\} \\ & \cup \{(i, 2, 1, u_1, u_2, u_3, u_4, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_3 \leq n_1, 1 \leq u_4 \leq n_2, 1 \leq u_5 \leq m\} \\ & \cup \{(i, 3, 0, u_1, u_2, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_5 \leq m\} \\ & \cup \{(i, 3, 1, u_1, u_2, u_4, u_5) : 1 \leq u_1 \leq S_1, 1 \leq u_2 \leq S_2, 1 \leq u_4 \leq n_2, 1 \leq u_5 \leq m\} \end{aligned}$$

Notations:

- \otimes - Kronecker product of two matrices of different dimensions.
- \oplus - Kronecker sum of two matrices of different dimensions.
- e - Column vector has a suitable size with each of its entries as 1.
- $e_0 - e_{2S_1 S_2 m}$.
- $e_1 - e_{2S_1 S_2 n_1 m + 2S_1 S_2 n_2 m + S_1 S_2 m}$.
- $e_2 - e_{2S_1 S_2 n_1 m + 2S_1 S_2 n_1 n_2 m + S_1 S_2 n_2 m + S_1 S_2 m}$.
- I_j - Square matrix with $j \times j$ size with diagonal entries as 1.
- 0 - It denotes zero matrices in the suitable order.

3.1. Construction of the QBD process for our Model

The generator matrix of the Markov chain under (s, S) policy is given by:

$$Q = \begin{bmatrix} A_{00} & A_{01} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots \\ A_{10} & A_{11} & A_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots \\ \mathbf{0} & A_{21} & F_1 & F_0 & \mathbf{0} & \mathbf{0} & \dots & \dots \\ \mathbf{0} & \mathbf{0} & F_2 & F_1 & F_0 & \mathbf{0} & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & F_2 & F_1 & F_0 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots & \dots \end{bmatrix}.$$

The entries in the block matrices of Q are defined as follows:

$$A_{00} = \begin{bmatrix} A_{00}^{11} & \mathbf{0} \\ A_{00}^{21} & A_{00}^{22} \end{bmatrix}, A_{00}^{11} = I_{S_1} \otimes P_1, A_{00}^{22} = I_{S_1} \otimes P_2,$$

$$P_1 = \begin{bmatrix} C_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & C_1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & \mathbf{0} & C_1 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_1 & \mathbf{0} & \dots & \mathbf{0} & C_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_3 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_3 \end{bmatrix}, P_2 = \begin{bmatrix} C_5 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & C_5 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & \mathbf{0} & C_5 & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_5 & \mathbf{0} & \dots & \mathbf{0} & C_6 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_7 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_7 \end{bmatrix},$$

$$C_1 = D_0 - \beta I_m, C_2 = \beta I_m, C_3 = D_0, C_4 = I_{S_2} \otimes \tau I_m, C_5 = D_0 - (\tau + \beta) I_m, C_6 = \beta I_m, \\ C_7 = D_0 - \tau I_m,$$

$$A_{00}^{21} = \begin{bmatrix} C_4 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_4 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_4 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_4 \end{bmatrix}, A_{01} = \begin{bmatrix} \mathbf{0} & A_{01}^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{01}^{24} & \mathbf{0} \end{bmatrix},$$

$$A_{01}^{12} = \begin{bmatrix} C_8 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_8 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_8 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_8 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_8 \end{bmatrix}, A_{01}^{24} = \begin{bmatrix} C_9 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_9 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_9 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_9 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_9 \end{bmatrix},$$

$$C_8 = I_{S_2} \otimes \gamma_1 \otimes D_1, C_9 = I_{S_2} \otimes D_1,$$

$$A_{10} = \begin{bmatrix} A_{10}^{11} & \mathbf{0} \\ A_{10}^{21} & A_{10}^{22} \end{bmatrix}, A_{10}^{11} = I_{S_1} \otimes C_{10}, A_{10}^{21} = \begin{bmatrix} \mathbf{0} & C_{11} \\ I_{S_1-1} \otimes C_{11} & \mathbf{0} \end{bmatrix}, A_{10}^{22} = \begin{bmatrix} \mathbf{0} & C_{12} \\ I_{S_1-1} \otimes C_{12} & \mathbf{0} \end{bmatrix},$$

$$A_{10}^{52} = I_{S_1} \otimes C_{13}, C_{10} = \begin{bmatrix} U_2^0 \otimes I_m & \mathbf{0} \\ I_{S_2-1} \otimes q_2 U_2^0 \otimes I_m & \mathbf{0} \end{bmatrix}, C_{11} = I_{S_2} \otimes q_1 U_1^0 \otimes I_m,$$

$$C_{12} = I_{S_2} \otimes \theta_1 U_1^0 \otimes I_m, C_{13} = \begin{bmatrix} U_2^0 \otimes I_m & \mathbf{0} \\ I_{S_2-1} \otimes q_2 U_2^0 \otimes I_m & \mathbf{0} \end{bmatrix},$$

$$A_{11} = \begin{bmatrix} A_{11}^{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{11}^{22} & A_{11}^{23} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_{11}^{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{11}^{42} & \mathbf{0} & A_{11}^{44} & A_{11}^{45} \\ A_{11}^{51} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{11}^{55} \end{bmatrix}, A_{11}^{11} = I_{S_1} \otimes P_3, A_{11}^{22} = \begin{bmatrix} P_4 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & P_5 \\ P_5 & P_4 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & P_5 & P_4 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_5 & P_4 \end{bmatrix},$$

$$P_3 = \begin{bmatrix} C_{14} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{15} \\ C_{16} & C_{14} & 0 & \dots & 0 & 0 & \dots & 0 & C_{15} \\ 0 & C_{16} & C_{14} & \dots & 0 & 0 & \dots & 0 & C_{15} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{14} & 0 & \dots & 0 & C_{15} \\ 0 & 0 & 0 & \dots & C_{16} & C_{17} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{17} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{16} & C_{17} \end{bmatrix}, C_{14} = U_2 \oplus D_0 - \beta I_{n_2 m}, C_{15} = \beta I_{n_2 m},$$

$$C_{16} = p_2 U_2^0 \gamma_2 \otimes I_m, C_{17} = U_2 \oplus D_0, P_4 = \begin{bmatrix} C_{18} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{19} \\ 0 & C_{18} & 0 & \dots & 0 & 0 & \dots & 0 & C_{19} \\ 0 & 0 & C_{18} & \dots & 0 & 0 & \dots & 0 & C_{19} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{18} & 0 & \dots & 0 & C_{19} \\ 0 & 0 & 0 & \dots & 0 & C_{20} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{20} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{20} \end{bmatrix},$$

$$C_{18} = U_1 \oplus D_0 - (\beta + \psi) I_{n_1 m}, C_{19} = \beta I_{n_1 m}, C_{20} = U_1 \oplus D_0, P_5 = I_{S_2} \otimes p_1 U_1^0 \gamma_1 \otimes I_m,$$

$$A_{11}^{23} = \begin{bmatrix} C_{21} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{21} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{21} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{21} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{21} \end{bmatrix}, C_{21} = I_{S_2} \otimes e_{n_1} \otimes \gamma_1 \otimes \psi I_m,$$

$$A_{11}^{33} = I_{S_1} \otimes P_6, A_{11}^{42} = \begin{bmatrix} C_{25} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{25} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{25} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{25} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{25} \end{bmatrix}, A_{11}^{44} = I_{S_1} \otimes P_7,$$

$$P_6 = \begin{bmatrix} C_{22} & 0 & 0 & \dots & 0 & 0 & \dots & C_{23} \\ 0 & C_{22} & 0 & \dots & 0 & 0 & \dots & C_{23} \\ 0 & 0 & C_{22} & \dots & 0 & 0 & \dots & C_{23} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{22} & 0 & \dots & C_{23} \\ 0 & 0 & 0 & \dots & 0 & C_{24} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{24} \end{bmatrix}, P_7 = \begin{bmatrix} C_{26} & 0 & 0 & \dots & 0 & 0 & \dots & C_{27} \\ 0 & C_{26} & 0 & \dots & 0 & 0 & \dots & C_{27} \\ 0 & 0 & C_{26} & \dots & 0 & 0 & \dots & C_{27} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{26} & 0 & \dots & C_{27} \\ 0 & 0 & 0 & \dots & 0 & C_{28} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{28} \end{bmatrix},$$

$$C_{22} = \theta U_1 \oplus D_0 - \beta I_{n_1 m}, C_{23} = \beta I_{n_1 m}, C_{24} = \theta U_1 \oplus D_0, C_{25} = I_{S_2} \otimes \gamma_1 \otimes \tau I_m, \\ C_{26} = D_0 - (\tau + \eta + \beta) I_m, C_{27} = \beta I_m, C_{28} = D_0 - (\tau + \eta) I_m,$$

$$A_{11}^{45} = \begin{bmatrix} C_{29} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{29} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{29} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{29} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{29} \end{bmatrix}, A_{11}^{51} = \begin{bmatrix} C_{30} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{30} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{30} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{30} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{30} \end{bmatrix},$$

$$C_{29} = I_{S_2} \otimes \gamma_2 \otimes \eta I_m, C_{30} = I_{S_2} \otimes I_{n_2} \otimes \tau I_m, A_{11}^{55} = I_{S_1} \otimes P_7,$$

$$P_7 = \begin{bmatrix} C_{31} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{32} \\ C_{33} & C_{31} & 0 & \dots & 0 & 0 & \dots & 0 & C_{32} \\ 0 & C_{33} & C_{31} & \dots & 0 & 0 & \dots & 0 & C_{32} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{31} & 0 & \dots & 0 & C_{32} \\ 0 & 0 & 0 & \dots & C_{33} & C_{34} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{34} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{33} & C_{34} \end{bmatrix},$$

where

$$C_{31} = U_2 \oplus D_0 - (\tau + \beta)I_{n_2m}, C_{32} = \beta I_{n_2m}, C_{33} = p_2 U_2^0 \gamma_2 \otimes I_m, C_{34} = U_2 \oplus D_0 - \tau I_{n_2m},$$

$$A_{12} = \begin{bmatrix} 0 & A_{12}^{12} & 0 & 0 & 0 & 0 \\ A_{12}^{21} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{12}^{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_{12}^{45} & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{12}^{55} \end{bmatrix}, A_{12}^{12} = \begin{bmatrix} C_{35} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{35} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{35} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{35} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{35} \end{bmatrix},$$

$$A_{12}^{21} = \begin{bmatrix} C_{36} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{36} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{36} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{36} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{36} \end{bmatrix}, C_{35} = I_{S_2} \otimes \gamma_1 \otimes I_{n_2} \otimes D_1, C_{36} = I_{S_2} \otimes I_{n_1} \otimes D_1,$$

$$A_{12}^{33} = \begin{bmatrix} C_{37} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{37} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{37} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{37} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{37} \end{bmatrix}, A_{12}^{45} = \begin{bmatrix} C_{38} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{38} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{38} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{38} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{38} \end{bmatrix},$$

$$C_{37} = I_{S_2} \otimes I_{n_1} \otimes D_1, C_{38} = I_{S_2} \otimes D_1, A_{12}^{45} = \begin{bmatrix} C_{39} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{39} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{39} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{39} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{39} \end{bmatrix},$$

$$C_{39} = I_{S_2} \otimes I_{n_2} \otimes D_1, A_{21} = \begin{bmatrix} \mathbf{0} & A_{21}^{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_{21}^{21} & A_{21}^{22} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{21}^{34} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_{21}^{43} & \mathbf{0} & A_{21}^{45} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & A_{21}^{65} \end{bmatrix}, A_{21}^{12} = \begin{bmatrix} \mathbf{0} & C_{40} \\ I_{S_{1-1}} \otimes C_{40} & \mathbf{0} \end{bmatrix},$$

$$A_{21}^{21} = I_{S_1} \otimes C_{41}, A_{21}^{22} = \begin{bmatrix} \mathbf{0} & C_{42} \\ I_{S_{1-1}} \otimes C_{42} & \mathbf{0} \end{bmatrix}, C_{40} = I_{S_2} \otimes q_1 U_1^0 \gamma_1 \otimes I_m,$$

$$C_{41} = \begin{bmatrix} e_{n_1} \otimes U_2^0 \gamma_1 \otimes I_m & \mathbf{0} \\ I_{S_{2-1}} \otimes e_{n_1} \otimes q_2 U_2^0 \gamma_1 \otimes I_m & \mathbf{0} \end{bmatrix}, C_{42} = I_{S_2} \otimes e_{n_2} \otimes q_1 U_1^0 \gamma_2 \otimes I_m, A_{21}^{34} = \begin{bmatrix} \mathbf{0} & C_{43} \\ I_{S_{1-1}} \otimes C_{43} & \mathbf{0} \end{bmatrix},$$

$$A_{21}^{43} = I_{S_1} \otimes C_{44}, A_{21}^{45} = \begin{bmatrix} \mathbf{0} & C_{45} \\ I_{S_{1-1}} \otimes C_{45} & \mathbf{0} \end{bmatrix}, A_{21}^{65} = I_{S_1} \otimes C_{46}, C_{43} = I_{S_2} \otimes \theta U_1^0 \otimes I_m,$$

$$C_{44} = \begin{bmatrix} I_{n_1} \otimes U_2^0 \otimes I_m & \mathbf{0} \\ I_{S_{2-1}} \otimes I_{n_1} \otimes q_2 U_2^0 \otimes I_m & \mathbf{0} \end{bmatrix}, C_{45} = I_{S_2} \otimes \theta U_1^0 \otimes I_{n_2 m}, C_{46} = \begin{bmatrix} U_2^0 \gamma_2 \otimes I_m & \mathbf{0} \\ I_{S_{2-1}} \otimes q_2 U_2^0 \gamma_2 \otimes I_m & \mathbf{0} \end{bmatrix},$$

$$F_1 = \begin{bmatrix} F_1^{11} & F_1^{12} & F_1^{13} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F_1^{22} & F_1^{23} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & F_1^{33} & F_1^{34} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & F_1^{44} & \mathbf{0} & \mathbf{0} \\ F_1^{51} & \mathbf{0} & \mathbf{0} & \mathbf{0} & F_1^{55} & F_1^{56} \\ \mathbf{0} & F_1^{62} & \mathbf{0} & \mathbf{0} & \mathbf{0} & F_1^{66} \end{bmatrix}, F_1^{11} = \begin{bmatrix} P_8 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & P_9 \\ P_9 & P_8 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & P_9 & P_8 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_8 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_9 & P_8 \end{bmatrix},$$

$$P_8 = \begin{bmatrix} C_{47} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{48} \\ \mathbf{0} & C_{47} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{48} \\ \mathbf{0} & \mathbf{0} & C_{47} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{48} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{47} & \mathbf{0} & \dots & \mathbf{0} & C_{48} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{49} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & C_{49} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{49} \end{bmatrix}, P_9 = I_{S_2} \otimes p_1 U_1^0 \gamma_1 \otimes I_m,$$

$$C_{47} = U_1 \oplus D_0 - (\eta + \beta + \psi) I_{n_1 m}, C_{48} = \beta I_{n_1 m}, C_{49} = U_1 \oplus D_0 - (\eta + \psi) I_{n_1 m},$$

$$F_1^{12} = \begin{bmatrix} C_{50} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{50} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_{50} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{50} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{50} \end{bmatrix}, F_1^{13} = \begin{bmatrix} C_{51} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{51} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_{51} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{51} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{51} \end{bmatrix},$$

$$C_{50} = I_{S_2} \otimes I_{n_1} \gamma_2 \otimes \eta I_m, C_{51} = I_{S_2} \otimes e_{n_1} \otimes \gamma_1 \otimes \psi I_m,$$

$$F_1^{22} = \begin{bmatrix} P_{10} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & P_{11} \\ P_{11} & P_{10} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & P_{11} & P_{10} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_{10} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_{11} & P_{10} \end{bmatrix}, F_1^{24} = \begin{bmatrix} C_{57} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{57} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_{57} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & C_{57} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & C_{57} \end{bmatrix},$$

$$P_{10} = \begin{bmatrix} C_{52} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{53} \\ C_{54} & C_{52} & 0 & \dots & 0 & 0 & \dots & 0 & C_{53} \\ 0 & C_{54} & C_{52} & \dots & 0 & 0 & \dots & 0 & C_{53} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{52} & 0 & \dots & 0 & C_{52} \\ 0 & 0 & 0 & \dots & C_{54} & C_{55} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{55} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{54} & C_{55} \end{bmatrix}, P_{11} = \begin{bmatrix} C_{56} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{56} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{56} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{56} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{56} \end{bmatrix},$$

where

$$C_{52} = U_1 \oplus U_2 \oplus D_0 - (\psi + \beta)I_{n_1 n_2 m}, C_{53} = \beta I_{n_1 n_2 m}, C_{54} = I_{n_1} \otimes p_2 U_2^0 \gamma_2 \otimes I_m, \\ C_{55} = U_1 \oplus U_2 \oplus D_0 - \psi I_{n_1 n_2 m}, C_{56} = I_{S_2} \otimes p_1 U_1^0 \gamma_1 \otimes I_{n_2 m}, C_{57} = I_{S_2} \otimes e_{n_1} \otimes \gamma_1 \otimes \psi I_{n_2 m}, \\ F_1^{33} = I_{S_1} \otimes P_{12},$$

$$P_{12} = \begin{bmatrix} C_{58} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{59} \\ 0 & C_{58} & 0 & \dots & 0 & 0 & \dots & 0 & C_{59} \\ 0 & 0 & C_{58} & \dots & 0 & 0 & \dots & 0 & C_{59} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{58} & 0 & \dots & 0 & C_{59} \\ 0 & 0 & 0 & \dots & 0 & C_{60} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{60} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{60} \end{bmatrix},$$

$$C_{58} = \theta U_1 \oplus D_0 - (\eta + \beta)I_{n_1 m}, C_{59} = \beta I_{n_1 m}, C_{60} = \theta U_1 \oplus D_0 - \eta I_{n_1 m},$$

$$F_1^{34} = \begin{bmatrix} C_{61} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{61} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{61} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{61} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{61} \end{bmatrix}, F_1^{44} = \begin{bmatrix} C_{62} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{63} \\ C_{64} & C_{62} & 0 & \dots & 0 & 0 & \dots & 0 & C_{63} \\ 0 & C_{64} & C_{62} & \dots & 0 & 0 & \dots & 0 & C_{63} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{62} & 0 & \dots & 0 & C_{63} \\ 0 & 0 & 0 & \dots & C_{64} & C_{65} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{65} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{64} & C_{65} \end{bmatrix},$$

where

$$C_{61} = I_{S_2} \otimes I_{n_1} \otimes \gamma_2 \otimes \eta I_m, C_{62} = \theta U_1 \oplus U_2 \oplus D_0 - \beta I_{n_1 n_2 m}, C_{63} = \beta I_{n_1 n_2 m}, \\ C_{64} = I_{n_1} \otimes p_2 U_2^0 \gamma_2 \otimes I_m, C_{65} = \theta U_1 \oplus U_2 \oplus D_0,$$

$$F_1^{51} = \begin{bmatrix} C_{66} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{66} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{66} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{66} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{66} \end{bmatrix}, F_1^{55} = I_{S_1} \otimes P_{13},$$

$$F_1^{56} = \begin{bmatrix} C_{70} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{70} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{70} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{70} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{70} \end{bmatrix}, P_{13} = \begin{bmatrix} C_{67} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{68} \\ 0 & C_{67} & 0 & \dots & 0 & 0 & \dots & 0 & C_{68} \\ 0 & 0 & C_{67} & \dots & 0 & 0 & \dots & 0 & C_{68} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{67} & 0 & \dots & 0 & C_{68} \\ 0 & 0 & 0 & \dots & 0 & C_{69} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{69} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{69} \end{bmatrix},$$

where

$$C_{66} = I_{S_2} \otimes \gamma_1 \otimes \tau I_m, C_{67} = D_0 - (\tau + \eta + \beta)I_m, C_{68} = \beta I_m, C_{69} = D_0 - (\eta + \tau)I_m,$$

$$C_{70} = I_{S_2} \otimes \gamma_2 \otimes \eta I_m, F_1^{62} = \begin{bmatrix} C_{71} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{71} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{71} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{71} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{71} \end{bmatrix}, C_{71} = I_{S_2} \otimes \gamma_1 \otimes \tau I_{n_2 m},$$

$$F_1^{66} = \begin{bmatrix} C_{72} & 0 & 0 & \dots & 0 & 0 & \dots & 0 & C_{73} \\ C_{74} & C_{72} & 0 & \dots & 0 & 0 & \dots & 0 & C_{73} \\ 0 & C_{74} & C_{72} & \dots & 0 & 0 & \dots & 0 & C_{73} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{72} & 0 & \dots & 0 & C_{73} \\ 0 & 0 & 0 & \dots & C_{74} & C_{75} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{75} & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & C_{74} & C_{75} \end{bmatrix},$$

where

$$C_{72} = U_2 \oplus D_0 - (\tau + \beta) I_{n_2 m}, C_{73} = \beta I_{n_2 m}, C_{74} = p_2 U_2^0 \gamma_2 \otimes I_m, C_{75} = U_2 \oplus D_0 - \tau I_{n_2 m},$$

$$F_0 = \begin{bmatrix} F_0^{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & F_0^{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & F_0^{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & F_0^{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & F_0^{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & F_0^{66} \end{bmatrix}, F_0^{11} = \begin{bmatrix} C_{76} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{76} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{76} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{76} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{76} \end{bmatrix},$$

$$F_0^{22} = \begin{bmatrix} C_{77} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{77} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{77} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{77} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{77} \end{bmatrix}, C_{76} = I_{S_2} \otimes I_{n_1} \otimes D_1, C_{77} = I_{S_2} \otimes I_{n_1 n_2} \otimes D_1,$$

$$F_0^{33} = \begin{bmatrix} C_{78} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{78} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{78} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{78} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{78} \end{bmatrix}, F_0^{44} = \begin{bmatrix} C_{79} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{79} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{79} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{79} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{79} \end{bmatrix},$$

where

$$C_{78} = I_{S_2} \otimes I_{n_1} \otimes D_1, C_{79} = I_{S_2} \otimes I_{n_1 n_2} \otimes D_1$$

$$F_0^{55} = \begin{bmatrix} C_{80} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{80} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{80} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{80} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{80} \end{bmatrix}, F_0^{66} = \begin{bmatrix} C_{81} & 0 & 0 & \dots & 0 & 0 \\ 0 & C_{81} & 0 & \dots & 0 & 0 \\ 0 & 0 & C_{81} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{81} & 0 \\ 0 & 0 & 0 & \dots & 0 & C_{81} \end{bmatrix},$$

$$C_{80} = I_{S_2} \otimes D_1, C_{81} = I_{S_2} \otimes I_{n_2} \otimes D_1,$$

$$F_2 = \begin{bmatrix} F_2^{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & F_2^{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & F_2^{35} & 0 \\ 0 & 0 & 0 & F_2^{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & F_2^{66} \end{bmatrix}, F_2^{11} = \begin{bmatrix} 0 & C_{82} \\ I_{S_1-1} \otimes C_{82} & 0 \end{bmatrix},$$

$$F_2^{22} = \begin{bmatrix} P_{14} & 0 & 0 & \dots & 0 & P_{15} \\ P_{15} & P_{14} & 0 & \dots & 0 & 0 \\ 0 & P_{15} & P_{14} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & P_{14} & 0 \\ 0 & 0 & 0 & \dots & P_{15} & P_{14} \end{bmatrix},$$

where

$$C_{82} = I_{S_2} \otimes q_1 U_1^0 \gamma_1 \otimes I_m, P_{14} = \begin{bmatrix} I_{n_1} \otimes U_2^0 \gamma_2 \otimes I_m & 0 \\ I_{S_2-1} \otimes I_{n_1} \otimes q_2 U_2^0 \gamma_2 \otimes I_m & 0 \end{bmatrix}, P_{15} = I_{S_2} \otimes q_1 U_1^0 \gamma_1 \otimes I_{n_2 m},$$

$$F_2^{35} = \begin{bmatrix} 0 & C_{83} \\ I_{S_1-1} \otimes C_{83} & 0 \end{bmatrix}, F_2^{44} = \begin{bmatrix} 0 & C_{84} \\ I_{S_1-1} \otimes C_{84} & 0 \end{bmatrix}, F_2^{66} = I_{S_2} \otimes C_{85},$$

where

$$C_{83} = I_{S_2} \otimes \theta U_1^0 \otimes I_m, C_{84} = I_{S_2} \otimes \theta U_1^0 \otimes I_{n_2 m}, C_{85} = \begin{bmatrix} U_2^0 \gamma_2 \otimes I_m & 0 \\ I_{S_2-1} \otimes q_2 U_2^0 \gamma_2 \otimes I_m & 0 \end{bmatrix}.$$

3.2. Stability condition

To discuss the stability condition, we first consider the generator matrix $F = F_0 + F_1 + F_2$. The vector χ is the invariant vector of the matrix F . Then, relations $\chi F = 0$ and $\chi e = 1$ and The LIQBD fashion with infinitesimal generator Q is stable if and only if

$$\chi F_0 e < \chi F_2 e.$$

The stability obtained after some mathematical rearranging is shown below:

$$\begin{aligned} & \chi_0 [e_{S_1 S_2 n_1} \otimes D_1 e_m] + \chi_1 [e_{S_1 S_2 n_1 n_2} \otimes D_1 e_m] + \chi_2 [e_{S_1 S_2 n_1} \otimes D_1 e_m] + \chi_3 [e_{S_1 S_2 n_1 n_2} \otimes D_1 e_m] \\ & + \chi_4 [e_{S_1 S_2} \otimes D_1 e_m] + \chi_5 [e_{S_1 S_2 n_2} \otimes D_1 e_m] < \chi_0 [e_{S_2} \otimes q_1 U_1^0 \otimes e_m] \\ & + \chi_1 \left(e_{S_1} \otimes [(e_{n_1} \otimes U_2^0 \otimes e_m + q_1 U_1^0 \otimes e_{n_2 m}) + e_{S_2-1} \otimes (e_{n_1} q_2 U_2^0 \otimes e_m + q_1 U_1^0 \otimes e_{n_2 m})] \right) \\ & + \chi_2 [e_{S_1 S_2} \otimes \theta U_1^0 \otimes e_{n_1 m}] + \chi_3 [e_{S_1 S_2} \otimes \theta U_1^0 \otimes e_{n_2 m}] \\ & + \chi_5 \left(e_{S_1} \otimes [U_2^0 \otimes e_m + e_{S_2-1} \otimes q_2 U_2^0 \otimes e_m] \right). \end{aligned}$$

3.3. The steady state probability vector

Let X be the steady state probability vector of the infinitesimal generator Q of the process $\{X(t): t \geq 0\}$. The subdivision of $X = (x_0, x_1, x_2, \dots)$, where x_0 is of dimension $2(S_1 S_2 m)$, x_1 is of dimension $2(S_1 S_2 n_2 m) + 2(S_1 S_2 n_1 m) + S_1 S_2 m$ and x_2, x_3, \dots are of dimension $2(S_1 S_2 n_1 m) + 2(S_1 S_2 n_1 n_2 m) + S_1 S_2 m + S_1 S_2 n_2 m$. As X is a vector satisfies the relation

$$XQ = 0 \text{ and } Xe = 1.$$

The probability vector X follows a matrix geometric structure under the steady state is

$$x_j = x_2 R^{j-1}, j \geq 3 \tag{2}$$

where R is the quadratic equation's lowest non-negative solution

$$R^2F_2 + RF_1 + F_0 = 0$$

and the vector x_0, x_1 and x_2 are obtained with the help of succeeding equations:

$$x_0A_{00} + x_1A_{10} = 0, \tag{3}$$

$$x_0A_{01} + x_1A_{11} + x_2A_{21} = 0, \tag{4}$$

$$x_1A_{12} + x_2[F_1 + RF_2] = 0, \tag{5}$$

subject to a condition normalization

$$x_0e_0 + x_1e_1 + x_2[I - R]^{-1}e_2 = 1. \tag{6}$$

Computing the rate matrix R is necessary before attempting to solve the set of equations mentioned above. However, [9] used Logarithmic reduction approach, an algorithm that makes it simple to produce R.

4. SYSTEM CHARACTERISTICS

- Probability of the system is empty:

$$P_{empty} = x_0e_0.$$

- The probability of the server-1 is idle:

$$P_{idle} = \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{000u_1u_2u_5} + \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{101u_1u_2u_4u_5}.$$

- The probability of the server-2 is on vacation:

$$\begin{aligned} P_{vac} = & \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{000u_1u_2u_5} + \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{030u_1u_2u_5} \\ & + \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{110u_1u_2u_3u_5} + \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{120u_1u_2u_3u_5} \\ & + \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{130u_1u_2u_5} + \sum_{i=2}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{i10u_1u_2u_3u_5} \\ & + \sum_{i=2}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{i20u_1u_2u_3u_5} + \sum_{i=2}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{i30u_1u_2u_5}. \end{aligned}$$

- The probability of the server-1 is offering service in normal mode:

$$P_{S_1B} = \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{i10u_1u_2u_3u_5} + \sum_{i=2}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i11u_1u_2u_3u_4u_5}.$$

- The probability of the server-1 is offering service in working breakdown:

$$P_{S_1WBD} = \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{i20u_1u_2u_3u_5} + \sum_{i=2}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m x_{i21u_1u_2u_3u_4u_5}.$$

- The probability of the server-2 is offering service in normal mode:

$$P_{S_2B} = \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{101u_1u_2u_4u_5} + \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i31u_1u_2u_4u_5} \\ + \sum_{i=2}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i21u_1u_2u_3u_4u_5}.$$

- The probability of the server-1 is on Repair:

$$P_{S_1R} = \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{030u_1u_2u_5} + \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_5=1}^m x_{i30u_1u_2u_5} \\ + \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i31u_1u_2u_4u_5}.$$

- Expected number of customers in the system:

$$E_{system} = \sum_{i=1}^{\infty} ix_i e = x_1 e_1 + x_2 [2(I - R)^{-1} + R(I - R)^{-2}] e_2.$$

- Expected first inventory level:

$$E_{IL_1} = \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} u_1 x_i(u_1, u_2)$$

- Expected first inventory level:

$$E_{IL_2} = \sum_{i=1}^{\infty} \sum_{u_1=1}^{S_1} \sum_{u_2=1}^{S_2} u_2 x_i(u_1, u_2)$$

- Expected reorder rate with first commodity

$$E_{RR_1} = \sum_{i=1}^{\infty} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_5=1}^m [x_{i101u_2u_3u_5} (U_1^0 \gamma_1 \otimes I_m) e + x_{i201u_2u_3u_5} (\theta U_1^0 \otimes I_m) e] \\ + \sum_{i=2}^{\infty} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i111u_2u_3u_4u_5} (U_1^0 \gamma_1 \otimes I_{n_2 m}) e \\ + \sum_{i=2}^{\infty} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i211u_2u_3u_4u_5} (U_1^0 \gamma_1 \otimes I_{n_2 m}) e.$$

- Expected reorder rate with second commodity

$$E_{RR_2} = \sum_{u_1=1}^{S_1} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{101u_1(s_2+1)u_5} (U_2^0 \otimes I_m) e \\ + \sum_{i=2}^{\infty} \sum_{u_2=1}^{S_2} \sum_{u_3=1}^{n_1} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m [x_{i11u_1(s_2+1)u_3u_4u_5} + x_{i21u_1(s_2+1)u_3u_4u_5}] (I_{n_1} \otimes U_2^0 \gamma_2 \otimes I_m) e \\ + \sum_{i=1}^{\infty} \sum_{u_2=1}^{S_2} \sum_{u_4=1}^{n_2} \sum_{u_5=1}^m x_{i31u_1(s_2+1)u_4u_5} (U_2^0 \gamma_2 \otimes I_{n_2 m}) e.$$

5. COST ANALYSIS

The cost function for our model was created with the premise that each cost element (per unit of time) correlates to a distinct system measure.

- C_{I_1} - the first item in inventory with a cost per unit
- C_{I_2} - the second item in the inventory with a cost per unit
- C_H - storing a customer's cost in the system for each unit of time.
- C_{R_1} -setup costs for each order of the primary item
- C_{R_2} - Setup costs for each order of complimentary items

$$TC = C_{I_1}E_{IL_1} + C_{I_2}E_{IL_2} + C_H E_{system} + C_{R_1}E_{RR_1} + C_{R_2}E_{RR_2}$$

6. NUMERICAL IMPLEMENTATION

To compute numerical outcomes, we have employed diverse MAP demonstrations for the incoming arrival in a manner that ensures their mean values are 1, as recommended by [5].

- **Erlang arrival (ERA):**

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix} D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

- **Exponential arrival (EXA):**

$$D_0 = [-1] D_1 = [1]$$

- **Hyper exponential arrival (HEXA):**

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix} D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}$$

- **MAP-Negative Correlation arrival (MNCA):**

$$D_0 = \begin{bmatrix} -1.00243 & 1.00243 & 0 \\ 0 & -1.00243 & 0 \\ 0 & 0 & -225.797 \end{bmatrix} D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.01002 & 0 & 0.99241 \\ 223.539 & 0 & 2.258 \end{bmatrix}$$

- **MAP-Positive Correlation arrival (MPCA):**

$$D_0 = \begin{bmatrix} -1.00243 & 1.00243 & 0 \\ 0 & -1.00243 & 0 \\ 0 & 0 & -225.797 \end{bmatrix} D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.99241 & 0 & 0.01002 \\ 2.258 & 0 & 223.539 \end{bmatrix}$$

Consider the following PH-distributions for the service and repair progression:

- **Erlang service (ERS):**

$$\gamma = [1, 0] \quad U = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

- **Exponential service (EXS):**

$$\gamma = [1] \quad U = [-1]$$

- **Hyper exponential service (HEXS):**

$$\gamma = [0.8, 0.2] \quad U = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

6.1. Illustrative 1

We examine the consequence of the service rate of server-2 (μ_2) versus the expected system size (E_{system}) in the Table 1-3. We fix $\lambda = 1, \mu_1 = 12, \eta = 1, \beta = 1, \psi = 1, \tau = 2, S_1 = 4, S_2 = 6, s_2 = 3, p_1 = 0.5, q_1 = 1 - p_1, p_2 = 0.5, q_2 = 1 - p_2, \theta = 0.6$, such that the system remains stable. The observation from Table 1-3 as follows:

- While we maximize the service rate of the server-2 then the corresponding E_{system} decreases with the combination of arrival and service times.
- From the point of view of arrival times, E_{system} decreases highly for *HYP A* and decreases slowly for *ERLA* while an increase the server-2's service rate. However, consider the service times, E_{system} decreases highly in *ERLS* and decreases slowly in *HYP S* with the combination of *EXPA, MNCA* and *MPCA*, but in the case of *ERLA* and *HYP A*, E_{system} decreases slowly in *EXPS* and decreases fastly in *HYP S*.

6.2. Illustration 2

With the support of the Tables 4-6, we visualize the influence of the service rate (μ_2) of the server-2 upon the probability of server-2 is undergoing vacation (P_{vac}). Fix $\lambda = 1, \mu_1 = 12, \eta = 1, \beta = 1, \psi = 1, \tau = 2, S_1 = 4, S_2 = 6, s_2 = 3, p_1 = 0.5, q_1 = 1 - p_1, p_2 = 0.5, q_2 = 1 - p_2, \theta = 0.6$, so that the stability condition is satisfied.

- Observation of Tables 4-6 discloses the fact that P_{vac} maximizes while maximizing the service rate of the server for the distinct feasible ordering of service and arrival times.
- This is because an increase in service rate leads to a decrease in the duration of service time respectively. As a result, the server will be getting more chances to go on vacation. Besides, the increment rate is higher in the case of *HYP A* and lower in the case of *ERLA*. In the same way, from the service time point of view, the increment rate is rapid for *HYP S* and gradual for *ERLS*.

6.3. Illustrative 3

With the support of Tables 7-9, we visualize the impact of the repair rate of the server-1(τ) on the Probability of server-1 being busy(P_{S_1B}). Fix $\lambda = 1, \mu_1 = 12, \mu_2 = 10, \eta = 1, \beta = 1, \psi = 1, S_1 = 4, S_2 = 6, s_2 = 3, p_1 = 0.5, q_1 = 1 - p_1, p_2 = 0.5, q_2 = 1 - p_2, \theta = 0.6$, so that the stability condition is satisfied.

- From Table 7-9, we may view that as the repair rate of server-1(τ) increases, (P_{S_1B}) increases for all possible groupings of service and arrival times.
- An increase in repair rate implies that the server takes minimum time to complete the repair process.
- As a result, the server will get more time to serve the customer and so (P_{S_1B}) increases. Hence, the system can focus on decreasing the time taken to repair the failed server for its optimal utilization. Moreover, the speed of increment of (P_{S_1B}) is high for *ERLA* and low for *HYP A*. In the same way, it is high for *ERLS* and low for *HYP S*.

Table 1: Server-2 service rate (μ_2) vs E_{system} - ERLS

μ_2	ERLA	EXPA	HYP A	MNCA	MPCA
8.0	0.197384249	0.224815927	0.2713795337	0.290089474	4.308746051
8.5	0.196918323	0.224009689	0.269865612	0.288694105	4.167282560
9.0	0.196498144	0.223275819	0.268486466	0.287407096	4.034807543
9.5	0.196117354	0.222604963	0.267225347	0.286215133	3.910580674
10.0	0.195770689	0.221989289	0.266068055	0.285107103	3.793926042
10.5	0.195453765	0.221422194	0.265002498	0.284073649	3.684228530
11.0	0.195162900	0.220898080	0.264018333	0.283106828	3.580929431
11.5	0.194894989	0.220412169	0.263106672	0.282199845	3.483521781
12.0	0.194647394	0.219960366	0.262259844	0.281346850	3.391545721
12.5	0.194417864	0.219539139	0.261471205	0.280542773	3.304584052

Table 2: Server-2 service rate (μ_2) vs E_{system} - EXPS

μ_2	ERLA	EXPA	HYP A	MNCA	MPCA
8.0	0.202436650	0.229325675	0.276190307	0.288189790	4.284049580
8.5	0.201923153	0.228463897	0.274610523	0.286720757	4.143969323
9.0	0.201458764	0.227678126	0.273168421	0.285367819	4.012769799
9.5	0.201036863	0.226958737	0.27184722	0.284116698	3.889718360
10.0	0.200651935	0.226297636	0.270632665	0.282955451	3.774147068
10.5	0.200299357	0.225687981	0.269512585	0.281874005	3.665448646
11.0	0.199975227	0.225123948	0.268476541	0.280863797	3.563071842
11.5	0.199676236	0.224600554	0.267515540	0.279917499	3.466516647
12.0	0.199399561	0.224113511	0.266621796	0.279028795	3.375329599
12.5	0.199142780	0.223659113	0.265788546	0.278192214	3.289099329

Table 3: Server-2 service rate (μ_2) vs E_{system} - HYP S

μ_2	ERLA	EXPA	HYP A	MNCA	MPCA
8.0	0.227728762	0.247910753	0.291707349	0.281256741	4.144213513
8.5	0.226937445	0.246779024	0.289865528	0.279466679	4.010955831
9.0	0.226211855	0.245738665	0.288170867	0.277819490	3.886146349
9.5	0.225544345	0.244779141	0.286606470	0.276298631	3.769067923
10.0	0.224928388	0.243891462	0.285157942	0.274889987	3.659073727
10.5	0.224358381	0.243067914	0.283812936	0.273581450	3.555580430
11.0	0.223829490	0.242301839	0.282560799	0.272362579	3.458061772
11.5	0.223337523	0.241587463	0.281392290	0.271224328	3.366042632
12.0	0.222878828	0.240919754	0.280299352	0.270158822	3.279093637
12.5	0.222450210	0.240294308	0.279274929	0.269159176	3.196826301

Table 4: Server-2 service rate (μ_2) vs P_{vac} - ERLS

μ_2	ERLA	EXPA	HYP A	MNCA	MPCA
8.0	0.992696356	0.988326139	0.981073741	0.978176264	0.960252482
8.5	0.993059887	0.988853514	0.981898129	0.979032770	0.961666866
9.0	0.993389158	0.989334618	0.982654586	0.979817153	0.962989215
9.5	0.993688677	0.989775249	0.983350966	0.980538758	0.964227717
10.0	0.993962221	0.990180294	0.983993994	0.981205329	0.965389685
10.5	0.994212971	0.990553889	0.984589456	0.981823341	0.966481656
11.0	0.994443622	0.990899561	0.985142349	0.982398253	0.967509488
11.5	0.994656468	0.991220336	0.985657014	0.982934708	0.968478435
12.0	0.994853476	0.991518816	0.986137231	0.983436680	0.969393223
12.5	0.995036336	0.991797261	0.986586309	0.983907592	0.970258104

Table 5: Server-2 service rate (μ_2) vs P_{vac} - EXPS

μ_2	ERLA	EXPA	HYP A	MNCA	MPCA
8.0	0.991933009	0.987434189	0.980074405	0.976954919	0.960264988
8.5	0.992324808	0.987994543	0.980928958	0.977881661	0.961679429
9.0	0.992680715	0.988506494	0.981713965	0.978729839	0.963001556
9.5	0.993005319	0.988976019	0.982437430	0.979509541	0.964239645
10.0	0.993302482	0.989408162	0.983106196	0.980229156	0.965401072
10.5	0.993575474	0.989807206	0.983726147	0.980895720	0.966492420
11.0	0.993827078	0.990176811	0.984302365	0.981515186	0.967519581
11.5	0.994059676	0.990520121	0.984839265	0.982092622	0.968487832
12.0	0.994275317	0.990839849	0.985340695	0.982632376	0.969401917
12.5	0.994475769	0.991138351	0.985810023	0.983138200	0.970266099

Table 6: Server-2 service rate (μ_2) vs P_{vac} - HYP S

μ_2	ERLA	EXPA	HYP A	MNCA	MPCA
8.0	0.987301352	0.982924032	0.975917875	0.971500408	0.960007628
8.5	0.987838969	0.983624068	0.976870150	0.972668334	0.961435754
9.0	0.988334016	0.984269283	0.977749342	0.973743753	0.962769344
9.5	0.988791273	0.984865824	0.978563672	0.974737072	0.964017128
10.0	0.989214835	0.985418958	0.979320150	0.975657230	0.965186832
10.5	0.989608227	0.985933220	0.980024790	0.976511937	0.966285315
11.0	0.989974506	0.986412541	0.980682791	0.977307876	0.967318682
11.5	0.990316333	0.986860343	0.981298668	0.978050865	0.968292380
12.0	0.990636036	0.987279615	0.981876366	0.97874599	0.969211284
12.5	0.990935659	0.987672986	0.982419347	0.979397714	0.970079769

Table 7: Server-1 repair rate (τ) vs P_{S_1B} - ERLS

τ	ERLA	EXPA	HYP A	MNCA	MPCA
2.1	0.141008388	0.13790536	0.13246213	0.130940643	0.075845503
2.2	0.141253797	0.138188589	0.132887094	0.131195807	0.075919075
2.3	0.141473036	0.138443996	0.133273312	0.131426853	0.07598541
2.4	0.141669716	0.138675233	0.133625511	0.131636941	0.076045499
2.5	0.141846844	0.138885360	0.133947705	0.131828717	0.076100165
2.6	0.142006944	0.139076961	0.134243323	0.132004409	0.076150095
2.7	0.142152144	0.139252229	0.134515310	0.132165904	0.076195868
2.8	0.142284249	0.139413038	0.134766207	0.132314815	0.076237974
2.9	0.142404799	0.139560993	0.134998218	0.132452523	0.076276831
3.0	0.142515115	0.139697482	0.135213260	0.132580217	0.076312796

Table 8: Server-1 repair rate (τ) vs P_{S_1B} - EXPS

τ	ERLA	EXPA	HYP A	MNCA	MPCA
2.1	0.138021661	0.134843332	0.129303375	0.127808856	0.074550912
2.2	0.138263883	0.135118502	0.129710507	0.128054649	0.074624807
2.3	0.138480662	0.135366864	0.130080690	0.128277310	0.074691540
2.4	0.138675482	0.135591923	0.130418429	0.128479865	0.074752084
2.5	0.138851247	0.135796620	0.130727549	0.128664843	0.074807246
2.6	0.139010396	0.135983435	0.131011317	0.128834378	0.074857702
2.7	0.139154988	0.136154477	0.131272540	0.128990278	0.074904020
2.8	0.139286770	0.136311545	0.131513638	0.129134086	0.074946683
2.9	0.139407235	0.136456186	0.131736709	0.129267126	0.074986104
3.0	0.139517664	0.136589733	0.131943580	0.129390538	0.075022635

Table 9: Server-1 repair rate (τ) vs P_{S_1B} - HYPS

τ	ERLA	EXPA	HYP A	MNCA	MPCA
2.1	0.121762987	0.118860529	0.113611432	0.112325780	0.067357513
2.2	0.121967364	0.119083143	0.113924156	0.112523046	0.067428853
2.3	0.122151232	0.119284568	0.114208672	0.112702075	0.067493742
2.4	0.122317344	0.119467554	0.114468443	0.112865232	0.067553018
2.5	0.122468001	0.119634405	0.114706401	0.113014495	0.067607384
2.6	0.122605135	0.119787070	0.114925047	0.113151535	0.067657429
2.7	0.122730382	0.119927203	0.115126525	0.113277768	0.067703651
2.8	0.122845133	0.120056218	0.115312681	0.113394403	0.067746476
2.9	0.122950579	0.120175328	0.115485115	0.113502478	0.067786270
3.0	0.123047742	0.120285582	0.115645216	0.113602890	0.067823346

7. CONCLUSION

In this study, we explain inventory management at service facilities using two types of servers: reliable servers and unreliable servers. Under steady-state conditions, matrix analytic techniques are used to determine the number of customers in the system, the server status, and the inventory level. Measures of important system features are derived in the steady state. We determined the optimality of this model by numerical analysis. As a result, this approach is appropriate for situations involving working vacation allocation when the server is reliable and for service disruption, or emergency vacation where the other server is unreliable.

REFERENCES

- [1] Ayyappan, G. and Karpagam, S. (2019). Analysis of a bulk queue with unreliable server, immediate feedback, N-policy, Bernoulli schedule multiple vacation and stand-by server. *Ain Shams Engineering Journal*, 10(4):873–880.
- [2] Ayyappan, G. and Gowthami, R. (2022) 'A MAP/PH₁, PH₂/1 system with two types of heterogeneous service, setup, closedown, vacation, immediate feedback, breakdown and repair', *Int. J. Mathematics in Operational Research*, Vol. 22, No. 3, pp.313–348.
- [3] Berman, O. and Sapna, K. P. (2000). Inventory management at service facilities for systems with arbitrarily distributed service times, *Stochastic Models*, 16, 343-360.
- [4] Chakravarthy, S. R. and Alfa, A. S. (1994). A finite capacity queue with Markovian arrivals and two servers with group services. *Journal of Applied Mathematics and Stochastic Analysis*, 7(2):161–178.
- [5] Chakravarthy, S.R. Markovian arrival processes. In *Wiley Encyclopedia of Operations Research and Management Science*; Wiley: Hoboken, NJ, USA, 2010.
- [6] Jeganathan, K., Sumathi, J. and Mahalakshmi, G. (2016). Markovian inventory model with two parallel queues, jockeying and impatient customers, *Yugoslav Journal of Operations Research*, 26(4), 467506.
- [7] Kalyanaraman, R. and Senthilkumar, R. (2019). Analysis of Heterogeneous Two Server Markovian Queue with Switching for Service Mode of First server. *AIP Conference Proceedings* 2177, 020031:1-5.
- [8] Jeganathan, K. 2014. Perishable inventory system at service facilities with multiple server vacations and impatient customers. *Journal of Statistics Applications and Probability Letters*, 2014, Vol. 3, Issue 3, P. 63-73.
- [9] Latouche, G. and Ramaswami, V. (1993) 'A Logarithmic reduction algorithm for quasi-birth-death processes', *Journal of Applied Probability*, Vol. 30, No. 3, pp.650-674.
- [10] Latouche, G. and Ramaswami, V. (1999) 'Introduction to Matrix Analytic Methods in Stochastic Modeling', Society for Industrial and Applied Mathematics, Philadelphia, PA.
- [11] Laxmi, P.V. and Soujanya, M.L. (2017). Retrial Inventory Model with Negative Customers and Multiple Working Vacations, *International Journal of Management Science and Engineering Management* 12(4):1-8.
- [12] Manikandan, R., & Nair, S. S. (2020). An M/M/1 Queueing-Inventory System with Working Vacations, Vacation Interruptions and Lost Sales. *Automation and Remote Control*, 81(4), 746–759. doi:10.1134/s0005117920040141.
- [13] Nithya, M., Sugapriya, C., Selvakumar, S., Jeganathan, K. and Harikrishnan, T. (2022). A Markovian two commodity queueing-inventory system with compliment item and classical retrial facility, *URAL MATHEMATICAL JOURNAL*, Vol. 8, No. 1, pp. 90–116.
- [14] Suganya, C., Sivakumar, B. and Arivarignan, G. (2017). Numerical investigation on MAP/PH(1), PH(2)/2 inventory system with multiple server vacations, *International Journal of Operational Research*, 29(1), 1-33.
- [15] Yang, D.-Y., Chen, Y.-H., and Wu, C.-H. (2019). Modelling and optimisation of a two-server queue with multiple vacations and working breakdowns. *International Journal of Production Research*, 58(10):3036–3048.
- [16] Yadavalli, V.S.S., Anbazhagan, N. and Jeganathan, K. (2015). A two heterogeneous servers perishable inventory system of a finite population with one unreliable server and repeated attempts, *Pakistan Journal of Statistics*, 31(1), 135-158.
- [17] Yadavalli, V.S.S. and Jeganathan, K. (2015). Perishable inventory model with two heterogeneous servers including one with multiple vacation and retrial customers, *Journal of Control and System Engineering*, 3(1), 10-34.