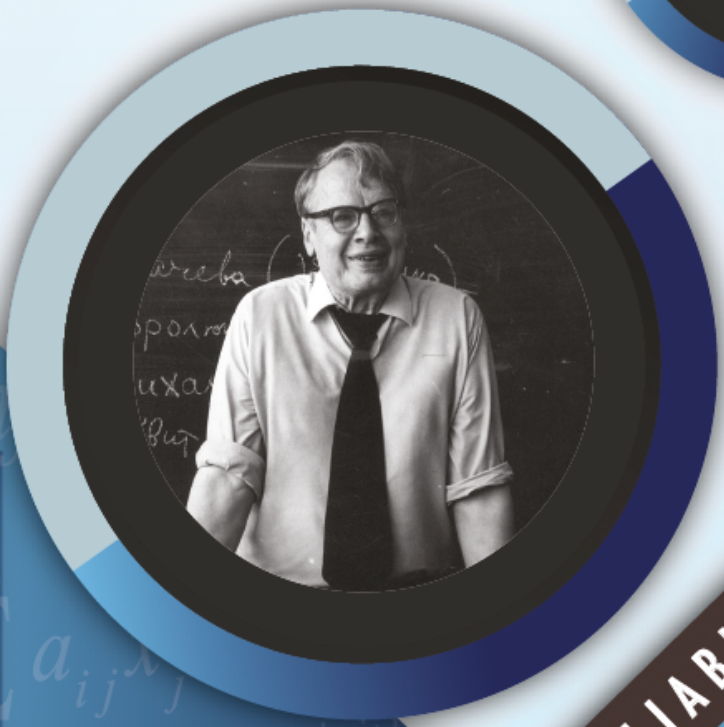
RTA

ISSN 1932-2321

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

RELIABILITY: THEORY & APPLICATIONS

INTERNATIONAL GROUP ON RELIABILITY



RELIABILITY
RISK ANALYSIS
MAINTENANCE
SAFETY

GNEDENKO FORUM PUBLICATIONS

(78) VOL.19

#2

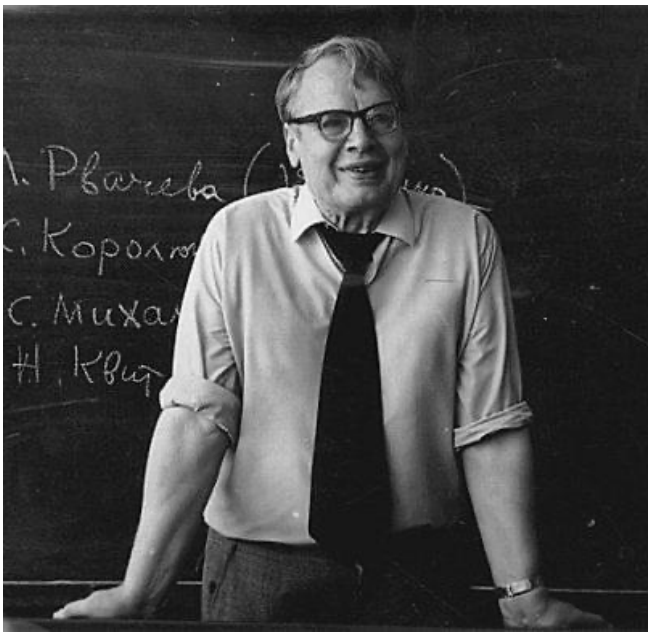
JUNE 2024

SAN DIEGO

ISSN 1932-2321

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





RELIABILITY:

THEORY & APPLICATIONS

Vol.19 No.2 (78),
June 2024

San Diego
2024

Editorial Board

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

The main aspects of the use of automatic monitoring and control systems in order to increase the reliability of power lines are considered. The article highlights modern technologies and techniques that make it possible to quickly identify and prevent possible emergency situations on power transmission lines. The advantages of automated systems compared to traditional monitoring and control methods are discussed, and examples of research and practical applications of such systems are presented. The results obtained can be useful for energy specialists and engineers involved in the design, operation and maintenance of power transmission lines, as well as for developers and manufacturers of automated monitoring and control systems.

This paper pictures the importance and the generalization of a new family of distribution developed on Triangle distribution. The new family provides some useful expansions, properties and a suitable alternative to some of existing models with same and higher number of parameters. Exponential distribution (one parameter) and Inverse Weibull distribution (Two parameter) play the role of sub-models. This new family distribution is used as a statistical model to estimate the parameters using the maximum likelihood estimation method. A complete study of Percentage points has been tabled. Two real-world data sets are investigated, demonstrating the suggested model's capacity to fit a variety of data sets along with some other models.

The work examines the main technical indicators of electromechanical converters callers with levitation elements, a generalized design method has been developed tions, as well as design diagrams and functional dependencies of the main varieties of electromechanical devices with elements of levitation. Analytical expressions for the levitation coefficient as a function of the dimensions of the magnetic core and coefficient factor of force multiplicity, technical characteristics of levitation material element, set superheat temperature. A mathematical model has been compiled for based on the parameters of the current mode and forces from the equations of electrical, magnetic, mechanical and thermal circuits of the magnetic system. As a result, the main dimensions of the magnetic system and dimensionless quantities. Analytic expressions for the main dimensions, the specified values of the winding overheating temperature are taken into account, input and output parameters, the condition of uniformity of the magnetic field in the working air nom gap. The optimal values of the dimensions of the magnetic circuit, active resistors have been determined winding voltages are minimal, resulting in minimization of losses active capacities.

This paper proposes a procedure for formulating the software reliability growth model using the non-homogeneous poisson process. We consider the software reliability growth model, which includes imperfect debugging, change points, and testing effort. Nevertheless, when formulating their software reliability models, the majority of scientists make the assumption of a constant detection rate per fault. When software is tested, they all suppose that each fault has an equal chance of being detected and that the rate is equal between generations. In practice, the fault detection rate varies depending on the test teams' abilities, program size, and software properties. Troubleshooting, even in the most realistic situations relevant to the error reintroduction rate due to incomplete debugging phenomena. In this case, changes in error detection and error introduction rates during software development program. Therefore, here we incorporate the generalized logic test workload function and change points. Parameters in software reliability modeling. Estimated using the least squares estimation method unknown parameters of the new model. Therefore, in our newly proposed model, we collect software testing data. use data from a practical application to illustrate the proposed model. Experimental results show that the proposed SRGM framework for imperfect debugging of integrated test jobs and change points has fairly accurate prediction capabilities.

HRV patterns. Rigorous training and validation reached to an 88% accuracy rate. These findings demonstrate the effectiveness of the proposed methodology. The LSTM model accurately forecasts stress responses, highlighting the potential of Apple Watch-acquired HRV data for stress assessment. Beyond prediction, the study enhances understanding of the complex interplay between HRV dynamics and unique stress reactions. This novel approach, leveraging Apple Watch features and intelligent computing, offers a personalized method to predict stress levels using K-Means Clustering Algorithm. Through integrating K-means clustering and person-specific HRV analysis, the research endeavours to advance our comprehension of the intricate interplay between physiological responses and stressors. The study offers a novel perspective on stress response variations by delving into the distinct autonomic patterns characterizing each cluster. It sets the stage for developing targeted interventions and personalized stress management strategies.

A shock model with two types of shocks functioning in the presence of an additional risk is proposed. Survival probability and mean residual life times of the proposed models are derived and assessed through the data of life testing experiment. Model validation and estimation of survival probability and mean residual life times is done through simulation studies. Comparison of survival probabilities and mean residual life times of models functioning without and with additional risk is made.

The condition evaluation for reinforced concrete framing requires comprehensive analysis of the factors influencing their performance such as strength, protective layer thickness, rebar diameter, thermal conductivity, humidity, adhesion of coatings, etc. Non-destructive methods are especially relevant when the characteristics of concrete and rebars are unknown and the scope of testing is considerable. Non-destructive testing allows to effectively monitor the conditions of technical devices, structures and buildings and enables to evaluate the timeliness and quality of repair and maintenance of a facility. Non-destructive testing provides the most reliable characteristics of the parameters defining the technical condition of the facilities under test. Non-destructive testing of the structural strength is applied in those areas, which have been exposed to loads due to natural and man-made contingencies.

This work applies the polar coordinates system of advanced calculus in the summation of the Gaussian distribution. In trying to achieve this aim, sub-concepts such as complex variables, gamma function of half, error function, and the relation between the error function and the standard normal distribution were defined and explained at various stages of the work. The embedded theorem which seems to be a new theorem also came up in the body of the work.

Aim of the present paper is to find suitable model for bimodal data. We have modelled mixture of two Weibull distributions in the presence of competing risks and also used Epanechnikov kernel to estimate hazard and survival functions. We considered prostate cancer data for application of the mixture model and kernel. We used maximum likelihood estimation (MLE) to estimate parameters of the mixture model, as the equations have no closed form, so we considered expectation–maximization (EM) algorithm. The mixture model and kernel gave good fit to the bimodal data. The prostate cancer data consists of three causes, we have estimated hazard function for these three causes using mixture model and kernel. The asymptotic confidence interval for the parameters of mixture model to all three causes were estimated. Also compared survival curve of mixture model with kernel and Kaplan-Meier survival curves for all the three causes.

In nowadays, the chief attention of the researcher is to study how the reliability analysis of manufacturing industrial systems by using the stochastic process. This topic tells us, how the manufacturing industries perform over time with the help of mathematical models which include randomness and uncertainty. Through stochastic processes we examine the reliability of these systems, technologists can recognize possible failure points and develop tactics to improve overall performance and effectiveness. The reliability of a manufacturing industrial system can be examined through a stochastic process, which permits for the estimate of failure rates and maintenance agendas. This analysis can lead to more well-organized and cost-effective procedure of the system. In this study, the researcher analysed the possibility plan for reliability through many distributions such as the normal distribution, gamma distribution, weibull distribution, and exponential distribution. The result of the study was prepared using Minitab software. The result of the study shows that the normal distribution of reliability fits best in comparison to the gamma, weibull, and exponential distributions.

In this paper, we consider a system with a finite number of components. It is assumed that the system architecture is a series format. The system fails when any one of the components fails. The case where the lifetimes of the components, are independently distributed and have pathway density is considered. Then the survival function, hazard function, the expected time to failure, general moments, etc. of the system lifetime are computed. It is shown that the hazard function can have many types of shapes, including bathtub shapes. The estimation of stress-strength reliability is considered based on the method of maximum likelihood estimation when both stress and strength variables follow the pathway model. Finally, to show the applicability of the proposed model in a real-life scenario, remission time data from cancer patients is analyzed.

better suit reliability data sets than alternative candidate distributions with an equivalent number of parameters. This distribution should be able to eliminate the Heavy-tail of the current distribution and model both monotonic and non-monotonic patterns of failure rates. Solar photovoltaic system reliability data was used to evaluate the performance of the proposed Extended Lomax distribution as well as the estimation methods.

Wararit Panichkitkosolkul

One of the lifetime distributions is the Iwueze distribution, which is constructed by combining the exponential and gamma distributions. In this paper, confidence intervals (CIs) are proposed for the parameter of the Iwueze distribution using the likelihood-based, Wald-type, bootstrap-t, and bias-corrected and accelerated (BCa) bootstrap methods. We evaluated the performance of the proposed CI methods through Monte Carlo simulation in terms of their coverage probability (CP) and average length (AL) in various scenarios. Furthermore, we had also derived the explicit formula for the Wald-type CI, which is straightforward for computation. The simulation results showed that the likelihood-based and Wald-type CIs returned satisfactory results according to coverage probabilities, even for the setting of small sample sizes. On the other hand, both the bootstrap-t and BCa bootstrap CIs yield CPs lower than the nominal confidence level when sample sizes are small. However, as the sample sizes increase, the CP of all CIs tend to approach the nominal confidence level. The parameter values also have a minor influence on the CP of all CIs when the sample size is fixed. Moreover, the AL of all CIs decreases as the sample size increases. The Wald-type and likelihood-based CIs have very similar ALs for all parameter values. In general, the bootstrap-t CI tends to yield the shortest interval. The effectiveness of all CIs was demonstrated by applying them to medical and engineering data, yielding results consistent with those of the simulation study.

Ngnassi Djami A.B., Samon J.B, Nzié W

The control of the maintenance of the industrial installations, in particular of the costs due to the implementation of the preventive policies is very interesting because of the growing importance of this service in the chains of production. The objective of this paper is to minimize the preventive maintenance costs of a production unit. For this, a state of the art on the maintenance cost models according to the policy used is first made, then a synthesis of the optimization methods is made in order to deploy the exact resolution methods and the genetic algorithms. The result of this paper is the proposal of a cost model corresponding to a periodic maintenance policy with minimal repair to the failure and the optimization of the periodicities of the partial revisions of the production unit.

Rajesh Singh, Kailash R. Kale and Pritee Singh

The Reliability is one of the key characteristics of software that operates flawlessly and in accordance with needs of users. The assessment of Reliability is very important but it is complicated. The oneparameter exponential class failure intensity function is used in this article to quantify the model and assess the software Reliability. The scale parameter and the number of existing total failures are the model's parameters. Using the Bayesian approach, the estimators of parameters are obtained under the assumption that gamma priors are suitable to provide prior information of the parameters. Using risk efficiencies computed under squared error loss, the performance of proposed estimators is studied with their corresponding maximum likelihood estimators. The suggested Bayes estimators are found to outperform over the equivalent maximum likelihood estimators.

The article considers multi-component and multi-stage mathematical models of queuing systems (QS) with the distribution of the incoming flow simultaneously between the system components, which consist of a certain number of service channels and waiting places in the queue. The maintenance of requirements with a lack of time to stay in the service channel and waiting is considered, while the service process in the QS of each component consists of several stages with the corresponding duration, and the full-service period is equal to the sum of such time intervals. The number of components and their parameters correspond to the similar characteristics of the production divisions of the repair enterprise. The study of the effectiveness of the operation of the repair enterprise as a multi-component and multi-stage QS consists in determining the values of the initial parameters of the QS components, taking into account the restrictions imposed on them, in order to obtain the largest values of the probabilities of servicing the requirements of the QS components and the system as a whole. The model is implemented using Any Logic University Researcher, which allows you to combine the principles of system dynamics with the paradigms of agent and discrete-event modelling. The proposed approach to the modelling of maintenance and repair processes by production divisions of the enterprise as a multi-component and multi-phase QS allows to determine the effectiveness of the functioning of such a QS and to obtain arguments for increasing the efficiency of its operation.

This paper relates to the reliability measures analysis of two identical unit system with reboot facility. Initially, one unit of the system is in operative mode and another unit is kept in cold standby mode. A technician is always available with the system to perform repairing and rebooting activities. Here, the system operative unit failed in safe mode and unsafe mode. During unsafe failure, repair activity cannot be done immediately but first rebooting is done to transform unsafe failure into safe failure, and then repair activity is performed as usual. Sometimes, the technician needs refreshments due to continuous work and provides better services after taking refreshments. The unit works like a new one after repair. The failure time of the unit in safe mode, unsafe mode and technician refreshment request time are assumed to be general while the repair time of the unit, rebooting delay time and technician refreshment time are taken as exponential. Reliability measures such as mean time to system failure, availability of the system, busy period of the repairman, the expected number of visits by the technician and profit values are calculated using tables.

The escalating connectivity of devices in the Internet of Things (IoT) era necessitates robust security measures while accommodating resource constraints. Lightweight cryptography addresses this need, focusing on algorithm development for devices with limited resources. This research proposes the Lightweight Crypto System (LCS) as a hybrid cryptosystem, integrating the Lightweight Symmetric Algorithm (LSA) and the Lightweight Hash Algorithm (LHA). LSA is a modified AES-128 variant, enhancing data confidentiality, while LHA, derived from SHA-256, verifies data integrity. The study evaluates the proposed LCS on criteria such as execution time, memory usage, avalanche effect, collision resistance, and entropy, emphasizing the optimal balance between performance and security achieved by LSA and LHA. The findings position LCS as a compelling solution for securing IoT devices without compromising on stringent security requirements.

This paper describes the role of queuing theory in developing queuing networks in companies. Queuing networks can be considered as a collection of nodes where each node stands for a service facility. It is a powerful and versatile tool for modeling facilities in manufacturing products. In the realm of service industries like scooter manufacturing, the queuing theory and simulation play a vital role. These concepts help in predicting queue lengths and waiting durations when multiple scooters are manufactured and distributed using first come first serve discipline. Tables are used to explore the availability of furnished scooters in the companies and their comparative study analyzes the waiting scooters and space availability in the companies.

A REVIEW ON BULK QUEUE WITH SERVER BREAKDOWN MODELS

Rani R[1] and Indhira K*

•

[1] Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore - 632 014, Tamil Nadu, India.
*Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore - 632 014, Tamil Nadu, India.
Email: rani.r2020@vitstudent.ac.in
Correspondence Email:kindhira@vit.ac.in

Abstract

This review article presents an overview of bulk arrival and bulk service with breakdown QM's. The concept of bulk arrivals and bulk service has a new significance in the world of reality. To prevent the problem of traffic congestion, researchers must concentrate their efforts on developing models and processes to address the issue. Numerical methods of QM's are critical in many industries, notably in production lines, to alleviate traffic congestion. This study seeks to give analysts, researchers, and industry professionals enough information to model congestion problems and extract various performance indicators to improve the QS's.

Keywords: Bulk arrival, Bulk service, Vacations, Breakdown.

1. Background and preliminaries

In order to improve the total service of the customers, queueing theory has been widely used as an operations management strategy to assess and simplify workforce needs, scheduling, and inventory. A.k.Erlang, a Danish mathematician, statistician, and architect, is credited with inventing not only queuing theory, but also the field of telephone traffic engineering as a whole.

Chaudhry and Templeton [73] provided a thorough examination of bulk queuing. A great place to start with customised modelling is with bulk arrivals analysis, which is a simplified version of our normal named customer analysis. To simulate a hospital outpatient department with a weekly clinic, a fixed-capacity transportation link, and an elevator, batch arrivals were used to represent bulk supplies and batch services. Under some conditions, networks of such queues are known to exhibit a product form of fixed distribution. Users enter in groups, and each group is served continuously in batch-arrival batch-service QM. In networking and telephony devices, such as multiple processors computer networks, where each programme requires the loading of memory units from a primary memory store, such queueing strategies begin to provide an example for performance measurement. A circuit-switched telecoms system that accommodates a number of traffic types, such as voice, video, and data, all of which have varying broadband requirements and holding durations.

Bailey [74] gets credit for developing bulk service QM's. He created the method, which he called "fixed-batch service." The server always serves a particular lot of customers in each group in fixed-batch service QS's. Kendall pioneered the embedded Markov-chain approach. This is accomplished through the use of regeneration points. Using Kendall's terminology for single

queueing nodes, $MX/MY/1$ indicates a $M/M/1$ queue with entries in groups defined by the random variable X and services in bulk given by the random variable Y. $GI^X/G^Y/1$ is extended in the same way as the $GI/G/1$ queue.

The following is a summary of the paper: The models of bulk arrival queues with breakdown are discussed in Section 2. The objective of Section 3 is on bulk service queues with breakdown. Finally, Section 4 presents the conclusion and summary.

2. Bulk arrival queues with server breakdown

A. M. Sultan et al. [34] explored a multi server, bulk arrival $(M^{[x]}/M/C; C-1/FCFS)$ QM using an extended Monte Carlo simulation. Because of the system breakage, the system can only serve with C or (C-1) servers. Average queue size, Average waiting time, and blocking probability were all established as measures of system efficiency. Simulation of the complete system yields numerical results. The performance evaluation of the finite buffer bulk and bunch service $GeoX/GY/1/K+B$ queue with several vacations, which could be employed in large wireless connections as well as other systems analysed by Seok Ho Chang and Dae Won Choi[35]. The Steady State(SS) probabilities and periods of the quantity of items in the system were provided at three different eras: departure, random, and arrival. Various helpful performance measurements were also been offered, such as the loss probability, the mean delay in the packet queue. The operation of a $M[x]/G/1$ QS under vacation criteria with startup/closedown timings was examined by Jau-Chuan Ke [36]. When all of the clients in the system have been processed, the server stops working at the closedown time. The server uses one of two vacation policies after shutdown: (1) multi vacation policies or (2) SV policies. In particular, the system features for the vacation models were examined. In a SS situation, Ahmed M.M Sultan [37] developed a workable solution for batch arrival queuing difficulties caused by the breakage of one of the heterogeneous servers. Due to the limitless number of possible variations in bulk queues and the breakage of one of the servers, massive tables with exact results were generated with different queueing variables. An $M^{[x]}/G/1$ QM's with an unreliable server and a SV policy was presented by M. Haridass and R. Arumuganathan [38].

While the server was running, it was susceptible to breakage, and the arrival time was determined by the server's up and down states. The time taken for anything to fail is exponentially distributed, and repair times are also distributed in a basic way. With the help of a numerical illustration, a cost model for QS's was also explored. The Two stage N-policy $M^X/M/1$ QS with startup times and server breakdowns was explored by V. Vasanta Kumar et al. [39]. The ideal value of N was determined using a cost function. Moreover,in addition to the N policy, the study of a QS with alternative vacation rules was also looked forward. In an MX/G(a, b)/1 queuing model with periodic vacations and closedown periods, S.Jeyakumar and B. Senthilnathan [40] tested the effect of the server breakage without interruption. It's also been noticed that when the rate of refurbishment rises, the projected line length lowers. G.Ayyappan and S.Shyamala [41] extended a single server with a BV and a random breakage. Transient solutions SS and probability generating function (PGF) were both assessed explicitly. N.A. Hassan and S.A. Hoda Ibrahim [42] demonstrated a recursive method for solving problems involving multi-level QM's in SS. When modifying the characteristics of a system, this method has been used to provide many of the system efficiency metrics. When modifying system parameters, the breakage of servers has an impact on system efficiency metrics. Charan Jeet Singh et al. [43] developed a single server QM with vacation, where goods are sent in bulk. The PGF of the number of components in the system was calculated using the Supplimentary Variable Technique (SVT), which can then be used to produce evaluation metrics like the average number of parts of the system, average waiting time, and so on.

In their examination of a bulk arrival two-stage retrial queueing models with balking, reneging, orbital search, and server breakdown, J.Ebenesar Anna Bagyam et al.[44] discussed the problem of delay time and reserved time. SVT provides an analytical solution to this model. In their study, Sushil Ghimire et al.[45] looked at a bulk QS with a fixed batch size of 'b' with users entering

the system in a Poisson manner and being segregated exponentially with the rate. Using PGF method, we may derive equations for W_q, W_s, L_q and L_s after constructing the mathematical model. S. Suganya [46] investigated into a $M[X]/G/1$ that has SOS, Multi Vacation, breakdown, and Repairs. The SVT is used to calculate the PGF of the number of consumers in the wait. This concept is applicable to large-scale production and communication networks. Madhu Jain and Amita Bhagat [47] generalized at how to deal with impatient consumers in the bulk arrival $M/G/1$ retry wait and changed the vacation policy. Zaiming Liu and Yang Song [48] examined a batch arrival $M^X/M/1$ queue model with functioning breakdown. WV is not the same as taking a break from work. Communication systems, transportation systems, production systems, and so on are all examples of this.

S. Maragathasundari [49] investigated a three stage heterogeneous service bulk arrival QM with various vacation policies. All arriving consumers must go through all three steps of service. The server can take a long vacation if he or she desires. The QS SS results are derived. In an $M^X/G(a,b)/1$ queuing model, M. Haridass and R.P. Nithya [50] generalised the server breakdown with interrupted vacation. In addition, a cost model has been built. A feedback QM with Bernoulli server vacation, multiple phases of unit service, and random server failure was reviewed by Sundar Rajan et al.[51]. The expected number of units and SS PGF had been calculated. Gautam Choudhury and Mitali Deka [52] analysed the queue size distribution due to busy time commencement era, and the waiting time distribution at random eras. In addition, several reliability indices and the system reliability function's LT were calculated. The concept of bulk arrivals was researched by S.P. Niranjan and K. Indhira [53]. To avoid a congestion problem, researchers must concentrate their efforts on developing models and processes to address the issues. It can assist researchers, engineers, and statisticians in the application of these models. G.Ayyappan et al. [54] investigated a bulk arrival with two different types of general bulk service QM with server breakage and modified M-vacation. The stationary queue size distribution at a random era, the busy period distribution, and the waiting time distribution were all generalised by Gautam Choudhury and Mitali Deka[55]. The LT of the system reliability function were also calculated.

R. P. Nithya and M. Haridass [56] developed a bulk QM that included a breakdown, batch control, and several vacations. The findings can also be utilised to make managerial decisions about how to reduce overall costs and they found the optimum operating policy in a QM. In practise, Charan Jeet Singh et al.[57] proved that a wide range of queuing system models for various capacity issues may be investigated in a general environment by taking into account the broad dispersion of service operations, bulk arrival, and unreliable servers.In this model, the stochastic principle was utilised to analyse a huge, unreliable arrival queue with vital services under the hypothesis of a Bernoulli feedback schedule. Many industries, such as telephone, wireless mobile networks, industrial production systems, and others, have real-world queueing challenges. G. Ayyappan and R. Supraja [58] explored an $M^X/G(a,b)/1$ QM with two stages of service subjected to system breakages and BV. The SVT was used to find performance indicators such as system state probabilities, average queue size, and queue length in the queue. S. Jeyakumar and B. Senthilnathan [59] investigated a variable bulk service estimation approach with numerous WV's and server breakdowns. The queue size was calculated for various arrival rates, service rates during WV's, service rates during regular periods, and WV duration. G. Ayyappan and P.Thamizhselvi [60] presented two categories of batch arrivals: high-priority and non-priority (retrial) clients, both of whom were treated with non-preemptive priority. Bernoulli Feedback, low-priority client collisions, orbiting search, and a revised BV for an unpredictable server, featured a breakdown and a time delay before repairs can begin. Customers arrive at this location via the compound Poisson process.

G. Ayyappan and S. Karpagam [61] looked at a generic bulk service with a backup server, an unpredictable arrival rate, multiple vacations, a server failure, and a second optional repair. Using a multiple vacation policy, G. Ayyappan and M. Nirmala [62] investigated the transient and SS behaviour of the $MX/G(a,b)/1$ queue with breakdown and two stages of repair with delay. In addition, the PGF of the queue size at any arbitrary and departing era was obtained.

According to the numeric values, the average queue length and waiting time grows as the arrival rate, breakge rate, and average queue size and waiting time of the batch of consumers increases. It's also been noticed that as service and repair rate rises, the projected line size and waiting time for a batch of clients decreases. $M^{[X]}/G(a,b)/1$, a QS with multiple vacation, closedown, essential, and optional repairs was generalised by G. Ayyappan and T. Deepa [63]. When the queue length drops below, the server close down and goes into multiple vacations. G. . The general bulk service queueing strategy with broken and repaired backup servers, numerous vacations, and a reservice request control policy was studied by G. Ayyappan and S. Karpagam [64]. It has also been observed that the predicted wait length grows when the primary server's rate of vacation does.

An innovative recoverable server QM with bulk input and state-dependent levels was examined by Charan Jeet Singh et al. [65], taking into consideration generic repair possibility, time to repair, and service processes. The server can offer two levels of service, the first of which is required and the second of which is optional. Jitendra Kumar and Vikas Shinde [66] discovered explicit mathematical equations for real-world challenges like customer dispatching methods for bulk arrivals and bulk services with multi-servers that include a mix of customers with systems holding and cancellation methods. MATLAB-9 was used to calculate the numerical results. This model can be analysed with time dependent arrival and service rates, which gives our model a more realistic feel. G. Ayyappan and R. Supraja [67] explored a bulk arrival non-Markovian queueing system with balking under BV, breakage, and repair under BV, breakage, and repair. A single server Markovian WV queue with customer balked and breakage was developed by R. Kalyanaraman and A. Sundaramoorthy [68]. Furthermore, the arrival rates has been affected by the state of the server. For a system that provides three stage of heterogeneous services for consumers that renege during server vacation and system outage periods, Samuel Ugochukwu Enogwe and Sidney Iheanyi Onyeagu [69] proposed a single server batch arrival QM. Queuing performance metrics including the probability of the system being idle, the utilisation factor Lq, Ls, Wq, and Ws were also calculated. G. Ayyapan and J. Udayageetha [70] examined a general retrial queueing system with priority services using M[X1], M[X2], /G1,G2/1. In this research, the system is entered by two distinct client types from separate classes using several independent compound Poisson processes. And also, they examined the server adheres to the pre-emptive priority principle how when it comes to working breakdown, startup/closedown times, and Bernoulli vacations with generic vacation periods. According to the non-preemptive priority service rule, G. Ayyappan et al.[71] suggested a single server serve two groups of consumers. In this paper, defined breakdowns are explored along with admission control, balking, and Bernoulli vacation. The server slows down service for the current client when the system experiences a breakdown, and then the repair work starts. Additionally, it is explained that a policy of admission control is in place to prevent the server from allowing all users to access the system. Two individualistic batch arrival queues with rapid feedback, a modified Bernoulli vacation, and server breakdown are incorporated in this study's steady state analysis by G. Ayyappan at al.[72] Priority and ordinary clients go into two different categories that need to be taken into account. And they are also talking about the non-preemptive priority discipline suggested by this approach.

3. Bulk service queues with server

K. C. Madan [1] explored a single channel QM in groups of fixed size $b(\geq 1)$ with Poisson arrivals and exponential service. The service channel, on the other hand, is liable to breakagess that happen at times. For the SS, the PGF of the queue size was obtained. I. P. Singh et al. [2] investigated a system that serves a fixed-size batch and is prone to random breakage. After repairs, the system enters an idle state before returning to a working state. The LT techniques were used to obtain the various transition probabilities as well as the SS solution. R. Nadarajan and D. Jayaraman [3] studied a Markovian tandem queue with two units and included general bulk service in unit II and server vacation in unit II, random breakage in both units, and a

finite interval waiting room. The SS probabilities and conditions were derived using the Matrix Geometric concept. Madhu Jain [4] used the PGF method to produce an analytical and explicit queue size distribution for distinct states. The mean queue length in a SS for various states were also determined. Madhu Jain and Poonam Singh [5] generalised a broad bulk service QM with repeated delayed vacations and a service time that's also state-dependent. In these environments, several real-life transportation systems, such as shuttle bus routes, cabs, fast lifts, and tour operators, can be shown.

Madhu Jain and Praveen Kumar Agrawal [6] examined a state-dependent $M/E_k/1$ QM's with server breakage and vacation time. The length of a server's vacation and its duration has been exponentially distributed. In both a WV and a busy period, service times has been considered to be Erlangian distributed. Investigation can be improved by adding bulk input/service. Lotfi Tadj and Gautam Choudhury [7] analysed a bulk service QS that had an ineffective server, Poisson input, and regular maintenance and processing times. A condition of stability has been established, along with SS system size distributions. The best management strategy was explored, with examples given to illustrate the point. Mathu Jain and Anamika Jain[8] studied a QM that includes WV's and server breakages, both of which require a series of phases of repair before operation can be restored. Some performance criteria were established. The influence of different variables was investigated using a sensitivity analysis. By applying the concept of multi optional repair, Madhu Jain et al.[9] investigated the unstable $M/E_k/1/WV$ queue. This approach is more stable and flexible since it combines the concepts of WV and service disruption due to server breakage. The SS equations that defines the model has also been built. $M/G/1$ queue Gautam Choudhury and Mitali Deka[10] presented that each customer requires two stages of service, the server is unprdictable and may fail at any time during the service, and the server is on a BV schedule.

Vikas Shinde and Deepali Patankar [11] studied with the departure of anxious clients and server vacations. They created the equations for SS probability as well as various system performance metrics, and also developed a cost model to establish the best service charge. Jau-Chuan Ke et al. [12] examined a multi-server QM with infinite capacity and a second optional service (SOS) channel. The optimization problem was solved using the quasi-Newton method and the Particle Swarm Optimization (PSO) method. The multistage batch arrival queue, including reneging on vacation and breakage times, was thoroughly analyzed. Sivagnanasundararam et al. [13] used SVT to produce SS solutions, and the average waiting time and average delay length. This concept can be applied to communication networks as well as large-scale manufacturing businesses. Under SS conditions, Sanjeet Singh and Naveen Kapil[14] examined the optimal operation of a single replaceable and server in a Markovian queuing system. The server's breakage and repair times are expected to be exponentially distributed. R P Nithya and M Haridass [15] analysed queue length distribution and also discussed the implications of various parameters on the performance of the system. S. Sasikala and K. Indhira [16] analysed a QM where customers were being served in batches in a bulk service, which can be fixed or variable in size. They explained why the service rate may be affected by the number of people in line for service. S. Jeyakumar and B. Senthilnathan [17] explained a model for variable bulk service queueing with various WV's and server breakages During WV's, the length of the queue was also obtained for different arrival and service rates.

M.Thangaraj and P.Rajendran [18] investigated batch arrival QM with two types of service patterns and a SV. S. Sasikala et al. [19] studied the SS behaviour of the M X / G(a,b) /1 queue in the presence of server downtime, numerous vacations, setup time, and N-policy. Using the PGF technique, the performance of the proposed QM may be measured. S.Bharathidass et al.[20] discussed single server Markovian arrival and Erlangian bulk service queues with state reliable rates. The system's state probabilities and expected number of units were explicitly calculated. G.Ayyappan and S.Karpagam [21] examined a batch arrival general bulk service single server QS with server breakage and optional second repair, stand-by server, balking, variable arrival rate, and many vacation. It has also been found that as the service rate of the main server increases, the projected queue size and waiting time decreases. Jitendra Kumar and Vikas Shinde [22]

developed a methodology for dealing with bulk arrivals and bulk service queues. L_q, L_s, W_q and W_s response times, as well as the efficiency of the server corresponding to consumers, have all been measured. MATLAB-9 was used to calculate the numerical results. Using the additional variable technique, Madhu Jain et al.[23] generalised the permanance modelling and analysis of a single server general service QM's with service interruption. M.Thangaraj and P.Rajendran [24] created and investigated batch arrival QM's with two types of service patterns and two types of vacations. Messaoud Bounkhel et al.[25] analysed a Markovian QS and used a numerical method based on operators to calculate the SS system size probabilities, as well as used a scientific method to calculate the PGF of these probabilities. The performance of a non-Markovian bulk service queuing models with an server, a stand-by server, loss and feedback, N-policy, and varied vacation Bernoulli schedules was investigated by G. Ayyappan and S. Karpagam [26] .The stand-by server has only been used while the primary server is being repaired. The queue size PGF was calculated, along with certain key performance indicators.

Srinivas R. Chakravarthy et al. [27] reported their findings in the setting of a single server queue with batch Markovian arrivals and a general service time distribution, regardless of batch size. For the general model, the SS probability vector has been determined almost directly, and we showed how explicit solutions look in a number of particular instances. Using reasoning from renewal theory, Niek Baer et al. [28] proposed a novel decomposition-based solution strategy for such queues, as well as they found the range of the wait period metric for multi-type server systems. The performance analysis of a non-preemptive priority $M/M/1$ queuing model with system breakage and repair time was conducted by K. Ruth Evangelin and V. Vidhya [29]. Investigation was done on the SS changes that occur in the equations of the queuing system using the complementary variable approach. Finally, the waiting time can also be computed using Little's formula, and a graphical depiction.Shanthi et al.[30] generalised a new numerical technique and utilized it to evaluate the transient stability of a bulk service queueing system with a server maintenance and breakage model using an infinite generator matrix and basic matrices. N.A. Hassan [31] demonstrated a simulation technique for solving SS problems involving multi-level QM's. When modifying the characteristics of a system, this method has been used to provide many of the metrics of system efficiency. When modifying system parameters, the breakage of servers had an impact on system efficiency metrics. In this study, G.Ayyappan and M.Nirmala [32] investigate the frustration of the client with an unreliable bulk queueing system with two forms of vacation. The server is allowed to take either a type I vacation or a type II vacation, subject to the size of the queue. And they also say that the server may experience malfunctions, and it requires a lot of setup time before repairs can commence. In this research, Rani Rajendiran and Indhira Kandaiyan [33] analysed the transient scrutiny of a batch arrival feedback queueing system with balking and two phases of variable service with differing levels of service subjected to Bernoulli vacation. They say that if the server is unable to accommodate the customer's request when they arrive, they also have the choice to deny services and exit the service area.

4. Conclusion

In this paper, a comprehensive review has been done on bulk queue with server Breakage models. One of the main reasons for conducting this survey is to get insight into the bulk arrival and bulk service models with breakages. These models are widely used and it plays a prominent part in the sectors of telephones, wireless mobile networks, and industrial production systems. The topics connected to bulk models with Breakage, which has been discussed in various fields, have been synthesized. A wide spectrum of literature has been examined, with suitable citations.

References

[1] K. C. Madan, A single channel queue with bulk seryice subject to interruptions, Microelectron Retlab Vol. 29, No. 5, pp. 813-818, (1989).

[2] I.P. Singh, Chottu Ram and Dinesh Kumar, The service channel subject to breakdown and idleness with bulk service, Microelectron. Reliab., Vol. 30, No. 4, pp. 667-671, (1990).
[3] R. Nadarajan and D. Jayaraman, Series Queue with general Bulk service, random breakdown and vacation,Microelectron. Reliab., Vol. 31, No. 5, pp. 861-863, (1991).
[4] Madhu Jain, Single Server Queue with Server Breakdown Including Priority and Varying Rates, JKAU: Eng.Sci.,vol.11 No.2,pp. 51-60 (1999).
[5] Madhu jain and Poonam singh, State Dependent Bulk Service Queue with Delayed Vacations, JKAU: Eng. Sci., vol. 16 no. 1, pp. 3-15 (2005 A.D./1426 A.H.).
[6] Madhu Jain and Praveen Kumar Agrawal,$M/E_k/1$ queueing system with working vacation, Quality Technology & Quantitative Management vol.4, No.4, pp. 455-470,(2007).
[7] Lotfi Tadj and Gautam Choudhury,A quorum queueing system with an unreliable server,Applied Mathematics Letters, 22, 1710–1714, (2009)
[8] Mathu jain and Anamika Jain, working vacations queueing model with multiple types of server breakdowns, Applied Mathematical Modelling, 34, 1–13, (2010).
[9] Madhu Jain, G.C.Sharma and Richa Sharma, Working vacation Queue with Service Interruption and Multi Optional Repair, International Journal of Information and Management Sciences, 22, 157-175, (2011).
[10] Gautam Choudhury, Mitali Deka, A single server queueing system with two phases of service subject to server breakdown and Bernoulli vacation, Applied mathematical modelling, 36, 6050-6060, (2012).
[11] Vikas Shinde and Deepali Patankar, Performance Analysis of State Dependent Bulk Service Queue with Balking, Reneging and Server Vacation, Shinde and Patankar / IJORN 1 , 61 – 69, (2012).
[12] Jau-Chuan Ke , Chia-Huang Wu , Wen Lea Pearn, Analysis of an infinite multi-server queue with an optional service, Computers & Industrial Engineering, 65, 216–225, (2013).
[13] Sivagnanasundararam, Maragathasundari and Santhanagopalan Srinivasan, A Non-Markovian Multistage Batch Arrival Queue with Breakdown and Reneging, Hindawi Publishing Corporation Mathematical Problems in Engineering , Article ID 519579, 16 pages, (2014).
[14] Sanjeet Singh, Naveen Kapil, Analysis of a Bulk Arrival Bulk Service queueing model for Non Reliable Server, IJCEM International Journal of Computational Engineering & Management, Vol. 17 Issue 6, November(2014).
[15] R P Nithya, M Haridass ,Analysis of a queueing system with two phases of bulk service, closedown and interrupted vacation, International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 11 No.1 (2016).
[16] S.Sasikala and K.Indhira, Bulk Service queueing models – A Survey, International Journal of Pure and Applied Mathematics, Volume 106 No. 6, 43-56, (2016).
[17] S.Jeyakumar and B.Senthilnathan, Modelling and analysis of a bulk service queueing model with multiple working vacations and server breakdown, RAIRO-Oper. Res. 51, 485–508, (2017).
[18] M.Thangaraj and P.Rajendran, Analysis of bulk queueing system with single service and SV, IOP Conf. Series: Materials Science and Engineering 263 (2017).
[19] S Sasikala, K Indhira and V M Chandrasekaran, General bulk service queueing system with N-policy, multiplevacations, setup time and server breakdown without interruption, IOP Conf. Series: Materials Science and Engineering 263, (2017).
[20] Bharathidass.S, Arivukkarasu.V and Ganesan.V, Bulk service queue with server breakdown and repairs, International Journal of Statistics and Applied Mathematics, 3(1): 136-142, (2018).
[21] G.Ayyappan and S.Karpagam, An $M^{[X]}/G(a,b)/1$ queueing system with Breakdown and Second Optional Repair, Stand-by Server, Balking, Variant Arrival Rate and Multiple Vacation,Int.J. Math. And Appl., 6(2–A), 145–156, (2018).

[22] Jitendra Kumar and Vikas Shinde, Performance Evaluation Bulk Arrival and Bulk Service with Multi Server using Queue Model, International Journal of Research in Advent Technology, Vol.6, No.11, November (2018).
[23] Madhu Jain, Sandeep Kaur, Parminder Singh, Supplementary variable technique (SVT) for non?Markovian single server queue with service interruption (QSI), operational Research, https://doi.org/10.1007/s12351-019-00519-8.
[24] M.Thangaraj and P.Rajendran, Analysis of Batch Arrival Bulk Service queueing system with Breakdown, Different Vacation Policies, and Multiphase Repair, Advances in Algebra and Analysis, Trends in Mathematics.
[25] Messaoud Bounkhel, Lotfi Tadj, Ramdane Hedjar SS Analysis of a Flexible Markovian Queue with Server Breakdowns,Entropy, 21, 259, (2019). doi:10.3390/e21030259.
[26] G. Ayyappan and S. Karpagam Analysis of a bulk queue with unreliable server, immediate feedback, N-policy, Bernoulli schedule multiple vacation and stand-by server, Ain Shams Engineering Journal, (2019).
[27] Srinivas R.Chakravarthy, Shruti, Alexander Rumyantsev, Analysis of a queueing model with Batch Markovian Arrival Process and General Distribution for Group Clearance, Methodology and Computing in Applied Probability, https://doi.org/10.1007/s11009-020-09828-4.
[28] Niek Baer, Nishant Mishra and Debjit Roy, Batch service systems with heterogeneous servers, queueing systems, https://doi.org/10.1007/s11134-020-09654-y.
[29] K. Ruth evangelin and V. Vidhya, $M/M/1$ non-preemptive priority model with system breakdown and repair times, Advances in Mathematics: Scientific Journal 9, no.10, 8197–8205, (2020).
[30] Shanthi, Muthu Ganapathi Subramanian and Gopal Sekar, Computational approach for transient behaviour of $M/M(a,b)/1$ bulk service queueing system with server breakdown and repair, Malaya Journal of Matematik, Vol. S, No. 1, 443-447,(2021).
[31] N. A.Hassan, Analysis of Multi-Level queueing systems with Servers Breakdown by Using Simulation Technique.
[32] G. Ayyappan, and M. Nirmala. "Analysis of customer's impatience on bulk service queueing system with unreliable server, setup time and two types of multiple vacations." International Journal of Industrial and Systems Engineering 38.2, 198-222, (2021).
[33] Rajendiran, Rani, and Indhira Kandaiyan. "Transient scrutiny of $M^X/G(a,b)/1$ queueing system with feedback, balking and two phase of service subject to server failure under Bernoulli vacation." AIMS Mathematics 8.3, 5391-5412, (2023).
[34] A. M. Sultan, N. A. Hassan and N. M. Elhamy, Computational analysis of a multi-server bulk arrival with two modes server breakdown, Mathematical and Computational Applications, Vol. 10, No. 2, pp. 249-259, (2005).
[35] Seok Ho Chang and Dae Won Choi, Performance analysis of a finite-buffer discrete-time queue with bulk arrival, bulk service and vacations, Computers & Operations Research 32, 2213 – 2234, (2005).
[36] Jau-Chuan Ke, Batch arrival queues under vacation policies with server breakdowns and startup/closedown times, Applied Mathematical Modelling 31, 1282–1292, (2007).
[37] Ahmed M. M. Sultan, Multi-channel bi-level heterogeneous servers bulk arrival queueing system with erlangian service time, Mathematical and Computational Applications, Vol. 12, No. 2, pp. 97-105, (2007).
[38] M. Haridass and R. Arumuganathan, Analysis of a Bulk Queue with Unreliable Server and SV, Int. J. Open Problems Compt. Math., Vol. 1, No. 2, September (2008).
[39] V. Vasanta Kumar , B.V.S.N. Hari Prasad and K. Chandan, Optimal Strategy Analysis of an N-policy Twophase $M^X/M/1$ Gated queueing system with Server Startup and Breakdowns, Int. J. Open Problems Compt. Math., Vol. 3, No. 4, December (2010).
[40] S. Jeyakumar and B. Senthilnathan, A study on the behaviour of the server breakdown without interruption in a $M^x/G(a,b)/1$ queueing system with multiple vacations and closedown time, Applied Mathematics and Computation 219, 2618–2633, (2012).

[41] G. Ayyappan and S. Shyamala, $M[X]/G/1$ with Bernoulli Schedule Server Vacation Random Break Down and second optional Repair, Journal of Computations & Modelling, vol.3, no.3, 159-175, (2013).

[42] N.A. Hassan and S.A. Hoda Ibrahim, Analysis of multi-level queueing systems with servers breakdown by using recursive solution technique, Applied Mathematical Modelling 37, 3714–3723, (2013).

[43] Charan Jeet Singh , Madhu Jain and Binay Kumar, Analysis of unreliable bulk queue with state dependent arrivals, Journal of Industrial Engineering International, 9:21, (2013).

[44] J.Ebenesar Anna Bagyam, K.Udaya Chandrika and K.Prakash Rani, Bulk Arrival Two Phase Retrial queueing system with Impatient Customers, Orbital Search, Active Breakdowns and Delayed Repair, International Journal of Computer Applications (0975 – 8887), Volume 73– No.11, July (2013).

[45] Sushil Ghimire , R. P. Ghimire and Gyan Bahadur Thapa, Mathematical Models of $M^b/M/1$ Bulk Arrival queueing system, Journal of the Institute of Engineering, Vol. 10, No. 1, pp. 184–191.

[46] S. Suganya, $M^{[X]}/G/1$ with Second Optional Service, Multiple Vacation, Breakdown and Repair, International Journal of Research in Engineering and Science (IJRES), Volume 2 Issue 11, November, PP.70-77, (2014).

[47] Madhu Jain and Amita Bhagat, Unreliable bulk retrial queues with delayed repairs and modified vacation policy, International Journal of Research in Engineering and Science (IJRES), J Ind Eng Int, 10:63, (2014).

[48] Zaiming Liu and Yang Song, The $M^x/M/1$ queue with working breakdown, RAIRO-Oper. Res. 48, 399–413, (2014).

[49] S. Maragathasundari, A Bulk arrival queueing model of three stages of service with different vacation policies, service interruption and delay time, American International Journal of Research in Science, Technology, Engineering & Mathematics,11(1), pp. 52-56, (2015).

[50] M. Haridass and R.P. Nithya, Analysis of a bulk queueing system with server breakdown and vacation interruption, International Journal of Operations Research Vol. 12, No. 3, (2015).

[51] Sundar Rajan , Ganesan. V and Rita .S, *Feedback Queue With Multi-Stage Heterogeneous Services And Random Breakdown*, Global Journal of Pure and Applied Mathematics. ISSN 0973-1768 Volume 11, Number 2, pp.1135-1145, (2015).

[52] Gautam Choudhury and Mitali Deka, A batch arrival unreliable server delaying repair queue with two phases of service and Bernoulli vacation under multiple vacation policy, Quality Technology & Quantitative Management, (2016).

[53] S.P. Niranjan and K. Indhira, *A review on classical bulk arrival and batch service queueing model*, International Journal of Pure and Applied Mathematics Volume 106 No. 8, 45-51, (2016).

[54] G.Ayyappan, R.Vimala Devi and S.Suganya, A Batch Arrival Two Types of Bulk Service Queue with Server Breakdown and Modified M-Vacation, International Journal of Mathematics and its Applications Volume 4, Issue 4, 389–396, (2016).

[55] Gautam Choudhury and Mitali Deka, A batch arrival unreliable server delaying repair queue with two phases of service and Bernoulli vacation under multiple vacation policy, Quality Technology & Quantitative Management, (2016).

[56] R.P. Nithya and M. Haridass, Optimum cost analysis of a bulk queueing system with breakdown, controlled arrival and multiple vacations, International Journal of Pure and Applied Mathematics Volume 109 No. 9, 81– 89, (2016).

[57] Charan Jeet Singh, Sandeep Kaur and Madhu Jain, Waiting Time of Bulk Arrival Unreliable Queue with Balking and Bernoulli Feedback using Maximum Entropy Principle, Journal of Statistical Theory and Practice, (2016).

[58] G. Ayyappan and R. Supraja, Analysis of $M^X/G(a,b)/1$ queueing system with Two Phases of Service Subject to Server Breakdown and Extended Bernoulli vacations, (IJSIMR) Volume 5, Issue 11, PP 32-51, (2017).

[59] S. Jeyakumar and B. Senthilnathan, Modelling and analysis of a bulk service queueing model with multiple working vacations and server breakdown, RAIRO-Oper. Res. 51, 485–508, (2017).

[60] G. Ayyappan and P. Thamizhselvi, Transient analysis of bulk arrival general service retrial queueing system with priority, Bernoulli feedback, collisions, orbital search, modified Bernoulli vacation, random breakdown and delayed repair, (IJSS). ISSN 0973-2675 Volume 12, Number 1, pp. 57–70, (2017).

[61] G. Ayyappan and S. Karpagam, An $M^{[X]}/G(a,b)/1$ queueing system with Breakdown and Second Optional Repair, Stand-by Server, Balking, Variant Arrival Rate and Multiple Vacation, Int. J. Math. And Appl., 6(2–A), 145–156, (2018).

[62] G. Ayyappan and M. Nirmala, An $M^X/G(a,b)/1$ queue with breakdown and delay time to two phase repair under multiple vacation, Appl. Math. ISSN: 1932-9466 Vol. 13, Issue 2, pp. 639 – 663, (December 2018).

[63] G. Ayyappan and T. Deepa , Analysis of batch arrival bulk service queue with multiple vacation closedown essential and optional repair, Appl. Math. ISSN: 1932-9466 Vol. 13, Issue 2 , pp. 578 – 598, (December 2018)

[64] G. Ayyappan and S. Karpagam , An $M^{[X]}/G(a,b)/1$ queueing system with Breakdown and Repair, Stand-By Server, Multiple Vacation and Control Policy on Request for Re-Service, Mathematics, 6, 101, (2018).

[65] Charan Jeet Singh , Madhu Jain and Sandeep Kaur, Performance analysis of bulk arrival queue with balking, optional service, delayed repair and multi-phase repair, Ain Shams Engineering Journal ,9 , 2067–2077, (2018).

[66] Jitendra Kumar and Vikas Shinde, Performance Evaluation Bulk Arrival and Bulk Service with Multi Server using Queue Model, International Journal of Research in Advent Technology, Vol.6, No.11, November (2018).

[67] G. Ayyappan and R. Supraja, Transient Analysis of $M^X/G(a,b)/1$ queueing system with Balking under Bernoulli Schedule Vacation and Random Breakdown, Journal of Computer and Mathematical Sciences, Vol.9(5), 455-473, May (2018).

[68] R. Kalyanaraman and A. Sundaramoorthy, A Markovian single server working vacation queue with server state dependent arrival rate, balking and with breakdown, Malaya Journal of Matematik, Vol. S, No. 1, 160-166, (2019).

[69] Samuel Ugochukwu Enogwe and Sidney Iheanyi Onyeagu, Single Channel Batch Arrival queueing model for Systems that Provides Three-Stage Service for Customers that Renege During Server Vacation and Breakdown Periods, Journal of Xidian University , October (2021).

[70] G. Ayyappan and J. Udayageetha, Transient analysis of M [X1], M [X2]/G1, G2/1 retrial queueing system with priority services, working breakdown, start up/close down time, Bernoulli vacation, reneging and balking, Pakistan Journal of Statistics and Operation Research, 203-216, (2020).

[71] G. Ayyappan, P.Thamizhselvi, B.Somasundaram, and J. Udayageetha, Analysis of an retrial queueing system with priority services, working breakdown, Bernoulli vacation, admission control and balking. Journal of Statistics and Management Systems, 24(4), 685-702, (2021).

[72] G. Ayyappan, S. Nithya, and B.Somasundaram, Analysis of M [X1], M [X2]/G1, G2(a,b)/1 Queue with Priority Services, Server Breakdown, Repair, Modified Bernoulli Vacation, Immediate Feedback, Applications and Applied Mathematics, 17.2, (2022).

[73] M. Chaudhry and J. Templeton, A first course in bulk queues, John Wiley & Sons, (1983).

[74] N. T. Bailey, On queueing processes with bulk service, Journal of the Royal Statistical Society, Series B (Methodological), 80-87, (1954).

COMPARATIVE ANALYSIS OF METHODS FOR ESTIMATING ELECTRICITY LOSSES IN PROBLEMS OF OPERATIONAL OPTIMIZATION OF POWER SYSTEM MODES

V.Kh. Nasibov R.R. Alizade E.J. Isgenderov

•

Azerbaijan Scientific-Research and Design-Prospecting Power Engineering Institute
nvaleh@mail.ru , rena_alizade@mail.ru , iskenderelvin@gmail.com

Abstract

The distribution of unscheduled capacity corresponding to the difference between current and forecast values should be carried out according to the criterion of minimum costs for the units involved to cover this capacity. In operational management, the optimization of power distribution is the process of adjusting the regime for active power, obtained during its short-term planning and optimization. Some of the optimization parameters, such as the relative increase in energy consumption and a measure of the efficiency of the use of water resources, can be determined in the optimization of short-term regimes and used in the optimization of operational regimes. Other parameters included in the operational optimization equation are either calculated during operational control using telemetered parameters, or are set. During the operational optimization of the regime, unscheduled power between stations must be distributed in such a way as to ensure the same relative increases in energy consumption at power plant units, taking into account the relative increases in power losses in the network from the power of these stations. The article considers a comparison of two methods for the operational assessment of the relative increments of power losses for the tasks of operational optimization of the mode by active power.

Keywords: Knowledge Assessment, Training, Fuzzy Knowledge Base

I. Introduction

Operational optimization of load distribution in a mixed power system has a number of features, both in terms of the sequence of algorithmic constructions, and in terms of software implementation of the developed algorithms. The possible participation of HPPs in covering unplanned capacity, corresponding to the difference between current and forecast values, makes it necessary to conduct operational optimization of the regime, taking into account the efficient use of water resources in the energy system, and should be carried out according to the criterion of minimum energy consumption at power plants involved in covering this capacity. The implementation of this principle is algorithmically fraught with difficulties associated with the use of the current telemetered mode parameters. At the same time, it is especially difficult to estimate the relative increases in power losses, since the relative increase in power losses from the power of power plants changes both with a change in the network layout and operating parameters [1].

II. Methods for assessing relative growth of power loss

In operational management, the optimization of active power distribution is the process of adjusting the regime for active power obtained during its short-term planning [2-4]. As with short-term planning, in a hydrothermal power system with operational control, the equation for the optimal distribution of active capacities between power plants is the equation:

$$\frac{b_1}{1-\sigma_1} = \lambda_a \frac{q_a}{1-\sigma_a} = \lambda_b \frac{q_b}{1-\sigma_b} = \cdots = \lambda_n \frac{q_n}{1-\sigma_n} \quad (1)$$

The operational optimization algorithm consists in the implementation of the mode re-optimization equation (1) for active power.

It should be noted that of the parameters included in the operational optimization equation, only the coefficient of efficiency in the use of water resources λ and the coefficients of the characteristics of relative increases in energy consumption at power plant units are determined during short-term forecasting and from preliminary calculations, and the rest are either calculated during operational control using remotely measured parameters, or are set. At the same time, λ, found in the course of short-term forecasting, participate as constant coefficients in characterizing the relative increase in water consumption at HPPs.

In contrast to short-term planning, with operational optimization of the regime for active power, the optimization equation is characterized as follows: when the regime changes, unplanned power between stations must be distributed in such a way as to ensure the same relative increases in energy consumption at power plant units, taking into account the relative increases in power losses in the network from power these stations. At the same time, taking into account the relative increases in power losses from the power of power plants is a correction of the relative increases in energy consumption at the corresponding power plants. To implement the principle of optimality, it is necessary to quickly estimate the power losses and the relative increases in active power losses in the network.

An operational assessment of the relative increments of active power losses in the network is carried out either on the basis of the current values of the power of power plants or voltages in controlled nodes, and are found either by regression equations or by the method of average voltages.

As shown above, the distribution of unscheduled power, corresponding to the difference between current and forecast values, should be carried out according to the criterion of minimum fuel consumption on the units involved in covering this power and correspond to formula (1).

Included in the denominator of equation (1), the variable σi - is the relative increase in active power losses in the network from the power corresponding to the power plant. The use of σi values found from short-term calculations is impossible, since with a change in both the network scheme and the mode of operation of the power system, the values of σi also change, so it is necessary to use methods for determining σi at the pace of the process using the current telemetered mode parameters. Below are two methods for determining the relative increases in active power losses in the network.

III. Construction of analytical characteristics of relative increments of active power losses in electrical networks by node voltage vectors

In operational management, to determine the relative growth of power losses in the power system, telemetry of source voltage vectors can be used without introducing data on the

parameters of the electrical network, loads of power plants and consumers into the calculations [5-6].

As is known, in the general case for heterogeneous networks.

$$\sigma_i^P + \vartheta_i^Q = \frac{2\,(U_1 - U_0)}{U} \quad (2)$$

$$\sigma_i^Q - \vartheta_i^P = 2\sin\delta_i \quad (3)$$

Where,

σ_i^P, σ_i^Q, ϑ_i^P, ϑ_i^Q – relative increments of active power losses by source active power, relative gains of active power by source reactive power, relative increments of reactive power losses by source active power and relative increments of reactive power losses by source reactive power, respectively, U_i- source voltage, U_0 - balancing node voltage, U is the average network voltage, σi is the angle between the voltage vectors of the sources and the balancing nodes.

In the general case, these equations are not enough to determine the relative power losses, since the number of unknowns is greater than the number of equations. For an approximate solution of these equations, we first neglect the inhomogeneity of the network, when the quality factor of all branches is assumed to be the same, i.e.

$$\psi_s = \text{arctg}\frac{x_s}{r_s} = \text{idem} \quad (4)$$

Then, taking into account

$$\vartheta_i^p = -\,\sigma_i^P \text{tg}\psi \quad (5)$$

$$\vartheta_i^Q = -\,\sigma_i^Q \text{tg}\psi \quad (6)$$

Equations (2) and (3) will take the form:

$$\sigma_i^p - \sigma_i^Q \text{tg}\psi = \frac{2(U_i - U_0)}{U} \quad (7)$$

$$\sigma_i^Q + \sigma_i^P tg\psi = 2\sin\delta_i \quad (8)$$

By solving equations (7), (8) we obtain the following formulas for the relative gains in active power losses:

$$\sigma_i^p = \frac{2(U_i - U_0)}{U}\cos 2\psi + \sin\delta 1 \sin 2\psi \quad (9)$$

$$\sigma_i^Q = \frac{-(U_i - U_0)}{U}\sin 2\psi + 2\sin\delta 1 \cos 2\psi \quad (10)$$

And if we neglect the difference in the phase angles of the power of the nodes, then $\text{tg}\varphi_i = \frac{Q_i}{P_i}$, will be the same for all nodes, and then there will be the following relations between the relative gains in losses: $\sigma_i^Q = \sigma_i^p tg\varphi$ and $\vartheta_i^Q = \vartheta_i^P \text{tg}\varphi$

In this case, equations (2) and (3) will take the following form:

$$\sigma_i^p + \vartheta_i^P \text{tg}\varphi = \frac{2(U_i - U_0)}{U} \quad (11)$$

$$\sigma_i^p tg\varphi - \vartheta_i^P = 2\sin\delta i \quad (12)$$

By solving equations (11) and (12), we obtain the following equations for relative loss increments:

$$\sigma_i^P = \frac{2(U_i - U_0)}{U} \cos 2\varphi + \sin\delta 1 \sin 2\varphi \quad (13)$$

$$\sigma_i^Q = \frac{(U_i - U_0)}{U} \sin 2\varphi + 2\sin\delta 1 \sin 2\psi \quad (14)$$

As can be seen from Figure 1, the sum of the right and left parts of equations (9) and (10) for different values of the quality factor of all branches is a band of possible solutions limited between $tg\psi_{max}$ and $tg\psi_{min}$ on the voltage plane and an increase in active power losses[7-9].

And the sum of the right and left parts of the equation (13) and (14) at different values of the phase angles of power is a band of possible solutions limited between $tg\varphi_{max}$ and $tg\varphi_{min}$ on the voltage plane and increase in active power losses.

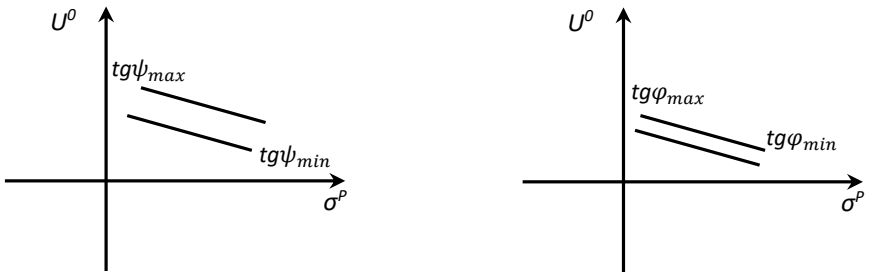


Figure 1.*a) The dependence of σP on voltage at the same quality factors of the branches*
b) The dependence of σP on voltage at the same phase angles of power in the nodes

Adding the left and right parts of equations (9), (10), (13), (14) and dividing by 2, we obtain the equations for the relative increments of active power losses in the network for two averaged parameters.

$$\sigma_\iota = \frac{(U_i - U_0)}{U} (2\cos 2\psi + 2\cos 2\varphi - \sin 2\psi + \sin 2\varphi)$$
$$\sin\delta 1 \, (\sin 2\psi + \sin 2\varphi + 2\cos 2\psi + 2\sin 2\varphi) \quad (15)$$

Equation (15) is a combination of two graphs shown in fig. 1 and therefore the boundary of possible changes in the relative increments of losses becomes even narrower and therefore the accuracy of the calculations increases. Let us denote the coefficients of two variables $\frac{(U_i - U_0)}{U}$, sinδ1,, inside the brackets, respectively, A and B, then equation (15) will take the form:

$$\sigma_\iota = A \frac{(U_i - U_0)}{U} + B\sin\delta 1 \quad (16)$$

The relative increase in active power losses is obtained as a function of two variables $\frac{(U_i - U_0)}{U}$ and sinδ1, Ui and U0 are the voltages of the source and the balancing node, they are usually maintained at the nominal value, telemetered and can be used at the pace of the control process.

It should be noted that the resulting equation can be used for operational control if the coefficients for two variables $\frac{(U_i - U_0)}{U}$ and sinδ1 are determined from preliminary calculations, since it is not possible to determine the components of the equation inside the brackets at the pace of the process. To obtain a working formula for a quick assessment of the relative increase in power losses, it is necessary to calculate a number of characteristic modes of the power system with the determination of the coefficients A and B in equation 16, and also determine U - the average voltage of the network. The fact is that the use of the U-average voltage of the network is difficult because usually the voltage telemetry contains significant measurement errors. Therefore, to determine the average network voltage, one can use the regression dependences of the average

voltage as a function of the voltage of some nodes. To determine these nodes, it is possible to carry out steady state calculations for normal and post-accident modes with the determination of the nodes on the voltage of which the mains voltage mode depends to the greatest extent, the so-called sensor nodes. The voltage of these nodes can be taken as factors for determining the average network voltage.

To determine the coefficients A and B, calculations were made for some characteristic modes of the Azerenerji energy system. The results of calculations with the determination of the coefficients A and B are given below:

№	A	B
1	0.0021	0.172
2	0.0025	0.165
3	0.0022	0.167
4	0.0024	0.169
5	0.0023	0.168

It can be seen from the table that coefficient A changes in ten-thousand digits, and B - in thousandth digits, so the arithmetic average value will correspond quite accurately to the desired coefficients:

Avr.= 0.0023, Vav.= 0.168 and equation (16) will take the form:

$$\sigma_i = 0.0023 \frac{(U_i - U_0)}{U} + 0.168\sin\delta 1 \qquad (17)$$

To determine U - the average voltage of the network as a function of the voltages of some nodes, calculations of normal modes and post-emergency modes were carried out in the Azerenergy system, obtained by typical outages of some power lines, a total of 64 calculations. As a result, 7 nodes were identified, on the voltage of which the mains voltage mode depends to the greatest extent, the so-called sensor nodes. The voltages of these nodes were determined as factors for determining the average network voltage. Below, without details, the found regression coefficients for the mean stress are given.

$$U = 121 + 0.7722\ UHOV + 0.4930\ UMUSH + 0.5923\ UKHUR + 0.2215\ UMAS + 0.4622\ UAKSU + 0.2123\ UAGC + 0.3311\ UAGS \qquad (18)$$

In this case, the standard deviation (SD) at the experimental points is 0.7%, and in the basic modes, when all factors are taken at an average level, the SD is 1.5%, which is at the level of the error in measuring telemetered node voltages. Thus, according to the telemetered values of the voltages of the sources, the balancing node and the 7 nodes listed above, it is possible to determine the relative increases in power losses in the network at the pace of the process.

Numerous calculations of the relative increases in power losses by the voltage vectors of the nodes show that the error for the minimum and maximum modes is greater (8%) than for the typical average modes (5%), when operational optimization of the modes by active power is possible.

IV. Construction of analytical characteristics of the relative increase in losses in the network by the active capacities of power plants

Here, it is required to build the dependence ΔP(Pi) on the basis of the experiment planning matrix,

where Pi is the load of power plants participating in the operational optimization of the regime. As is known, on the basis of the experiment planning matrix, equations are obtained with normalized values of the factors, i.e. varying from +1 to -1. For management purposes, not normalized, but natural values of factors are needed. In this regard, the construction of equations by natural values and their analysis are considered [10].

To obtain a regression equation with natural values of the factors, regression equations were constructed to determine the power losses in the power system, where the following power plants were the factors with the ranges of change in their total load for characteristic modes.

1. Shimal PP (760-560 MW)
2. Sumgayit PP (500-300 MW)
3. Canub PP (700-500 MW)
4. Azerbaijan TPP (1780-1580 MW)
5. Gobu PP (380-180 MW)

Regression models were obtained in a fractional-factorial experiment of the type N=25-1=24=16, in which the results of ΔP analyzes show a fairly high accuracy of the models. The planning matrix, indicating the interaction of factors used as additional factors and the interaction used in the regression equation, are given below. The losses ΔP in the network, obtained as a result of calculating the steady state using the corresponding program, are also indicated there.

Tab. 1

Table 1: *Planning matrix*

№	Shimal PP X1	Sumgayit PP X2	Canub PP X3	Azerbaijan TPPX4	Gobu PP X5	ΔP
1	760+j400	500+j300	700+j400	1780+j900	380+j200	110,7
2	530+j280	500+j300	700+j400	1780+j900	180+ j100	99,1
3	760+j400	300+ j160	700+j400	1780+j900	180+ j100	98,4
4	530+j280	300+ j160	700+j400	1780+j900	380+j200	97,1
5	760+j400	500+j300	500+ j280	1780+j900	180+ j100	101,8
6	530+j280	500+j300	500+ j280	1780+j900	380+j200	99,9
7	760+j400	300+ j160	500+ j280	1780+j900	380+j200	99,2
8	530+j280	300+ j160	500+ j280	1780+j900	180+ j100	104,9
9	760+j400	500+j300	700+j400	1580+ j800	180+ j100	100,9
10	530+j280	500+j300	700+j400	1580+ j800	380+j200	99,1
11	760+j400	300+ j160	700+j400	1580+ j800	380+j200	98,4
12	530+j280	300+ j160	700+j400	1580+ j800	180+ j100	104,1
13	760+j400	500+j300	500+ j280	1580+ j800	380+j200	102
14	530+j280	500+j300	500+ j280	1580+ j800	180+ j100	107,4
15	760+j400	300+ j160	500+ j280	1580+ j800	180+ j100	106,5
16	530+j280	300+ j160	500+ j280	1580+ j800	380+j200	105,3

Below, without details, the regression equation ΔP and the relative increase in power losses for characteristic modes are carried out.

For the operational optimization of the active power mode, the controlled parameters are the capacities of the power plants Shimal ES, Sumgayit ES, Janub ES and Azerbaijan TPP, and therefore the relative increases in power losses are determined as partial derivatives of the equation of power losses with respect to the capacities of the corresponding stations.

$$\Delta P = 622,97- 0,3201*X1 - 0,3405X2 - 0,3445X3 - 0,2185X4 - 0,357X5+0,000118X1X2 +0,000106X1X3+0,000107X1X4+0,000105X2X3+0,000105X2X4+0,000102X3X4$$

$$\sigma1=-0,3201+0,000118X2+0,000107X4$$
$$\sigma2=-0,3405+0,000118X1+0,000105X3$$
$$\sigma3=-0,3445+0,000106X1+0,000105X2$$
$$\sigma4=-0,2185+0,000107X1+0,000105X2$$

The resulting models at all experimental points have a very high accuracy. Numerous calculations show that the standard deviation of the calculated data from the experimental data is on average 4% for the maximum and minimum modes, and 7% for the middle modes, when operational optimization is possible. If we take into account that the proposed regression equations will be used in the additional optimization of the regime with the operational management of the distribution of approximately 200 MW of unscheduled active power, then the above deviations fall within the accuracy of the initial data and they can be used for the purposes of operational optimization.

V. Discussion

[1] N. Yusifbayli, V. Nasibov, R. Alizade Power consumption management and equalization of the load schedules of Azerbaijan power system. Rudenko International Conference "Methodological Problems in Reliability Study of Large Energy Systems" (RSES 2022) Volume 384, 2023

[2] Christoph Graf, Federico Quaglia&Frank A. Wolak Simplified Electricity Market Models with Significant Intermittent Renewable Capacity: Evidence from Italy
https://www.nber.org/papers/w27262

[3] Mariano Ventosa, Alvaro Baillo, Andres Ramos, Michel Rivier Electricity market modeling trends. Energy Policy, Volume 33, Issue 7, May 2005

[4] The European Electricity Market Model EMMA Model documentation, Latest version
http://neon-energie.de/emma

[5] M. Y. Hassan; M. P. Abdullah; A. S. Arifin; F. Hussin; M. S. Majid Electricity market models in restructured electricity supply industry, https://ieeexplore.ieee.org/document/4762618

[6] L.S. Belyaev Electricity market problems, 2009
https://isem.irk.ru/upload/iblock/228/228f43e79d0cd2692c29ae64199ad2e2.pdf

[7] Hawker G., Bell K., Gill S. Electricity security in the European Union. The conflict between national Capacity Mechanisms and the Single Market. Energy Research & Social Science, №24, 2017

[8] L.S. Belyaev, O.V. Marchenko, S.V. Podkovalnikov Growth of electrical energy prices necessary for energy system development at transition to competitive market
https://www.researchgate.net/publication/292808323_Growth_of_electrical_energy_prices_necessary_for_energy_system_development_at_transition_to_competitive_market

[9] Site of State Statistical Committee of the Republic of Azerbaijan, https://www.stat.gov.az/index.php

[10] L.A. Barroso; T.H. Cavalcanti; P. Giesbertz; K. Purchala Classification of electricity market models worldwide. International Symposium CIGRE/IEEE PES, New Orleans, LA, USA, 2005

TRANSIENT AND METAHEURISTIC COST SCRUTINY OF $M^X/G(A,B)/1$ RETRIAL QUEUE WITH RANDOM FAILURE UNDER EXTENDED BERNOULLI VACATION WITH IMPATIENT CUSTOMERS

Rani R[1] and Indhira K*

•

[1,*]Vellore Institute of Technology, Vellore - 632 014, Tamil Nadu, India.
kindhira@vit.ac.in.

Abstract

The transient and metaheuristic cost analysis of a $M^X/G(a,b)/1$ retrial queue with random failure during an extended Bernoulli vacation with impatient clients is covered in this study. Any batch that arrives and discovers the server is busy, down, or on vacation joins an orbit. In the alternative, only one new customer from the group joins the service right away, while the others join the orbit. After providing each service, the server either waits to serve the following customer with probability $(1-\theta)$ or goes on vacation with probability θ. It has been found that these systems express steady-state solutions and are dependent on time probability generating functions in consideration of their Laplace transforms. We also discuss a few exceptional and particular instances. After that, the impact of different parameters on the system's effectiveness is evaluated. We are also talking about ANFIS. Additional approaches employed in this study to swiftly determine the system's optimum cost include genetic algorithms (GA), artificial bee colonies (ABC), and particle swarm optimization (PSO). We also examined the graph-based convergence of several optimization algorithms.

Keywords: Batch arrival, Retrial queues, Feedback, Extended Bernoulli Vacation, ANFIS, Cost Optimization.

1. Introduction

For the development, capacity planning, performance assessment, and optimization of numerous real-world systems, queueing theory offers a potent tool. Chaudhry and Templeton[1] provided a comprehensive analysis of bulk queuing. Bulk arrival analysis, a condensed form of customer examination, is a great place to start with customised models. Bulk service queuing models were created by Bailey [2]. He invented the process known as "fixed-batch service". The server continuously offers a specific batch of services to each set of users in fixed-batch service queueing systems (QS).

The "retrial queueing"system, which is used when a customer enters and the server is occupied, requires the customer to leave the appropriate area and repeat his request after a certain period of time. This property is essential for network technologies, cognitive networks, online computing systems, manufacturing systems, and other systems.

Sumitha and Udaya Chandrika [3] investigated a retrial queuing system with starting failure, single vacation, and orbital search. In batch arrival retrial queues, Radha et al. [4] studied some system performance measures are evaluated using the supplementary variable technique (SVT) and the steady-state (SS) probability generating function (PGF) for system size.

Gomez-Corral has talked a lot about a retrial QS with FCFS discipline and typical retrial

periods. The M/G/1 retrial queue with feedback and starting failures was described by Krishna Kumar et al. [5]. Yang, Tao, and Hui Li[6] investigated an M/G/1 retrial queue with a starting failure-prone server. An analysis of a feedback retrial queuing system with starting failures and a single vacation was studied by Mokaddis et al. [7].

In a Vacation, queueing system the server could be temporarily unavailable for a number of reasons, including maintenance monitoring, tending to other queues, or simply taking a break. When the server is unavailable to users, that time period is referred to as a "vacation ". A single server batch arrival Bernoulli feedback QS with a waiting server, K-variant vacations, and anxious clients was examined by Bouchentouf et al. [8]. The transient behaviour of a batch arrival feedback retrial queue with starting failure and Bernoulli vacation (BV) was investigated by Ayyappan and Sathiya [9]. Assuming that repair, service, and vacation times are randomly distributed, the time-dependent PGF are also computed in relation to their Laplace transforms(LT).

Numerous academics who have studied queueing techniques with interruptions have as their primary tenet that, in the event of a failure, the service channel will be promptly repaired. A transient analysis of the $M^{[X_1]}, M^{[X_2]}/G1, G2/1$ retrial QS's with priority services, working breakdown, start up/close down time, BV, reneging, and balking was studied by Ayyappan et al. [10]. Kulkarni et al. [11] established a retrial queue with a server prone to failures and maintenance. Ayyappan and Shyamala [12] created an $M^{[X]}/G/1$ with Bernoulli schedule, server vacation, random break down and second optional repair. And also calculate the typical length of the line and the typical wait period in closed form. When the repair is finished, a number of consumers who had previously used the services wait for the remainder to be provided. Jau-Chuan Ke et al. [13] demonstrated a waiting line with customers complaining and providing feedback the servers malfunctioned. Furthermore, if all servers are already in use when a customer arrives, he will either join a retrial orbit or decline. When a service is finished, the client can exit the system or rejoin the retrial group to receive more services. They can also design a cost function to determine the system's ideal parameter settings under the stability condition. Computer telecommunication systems is a example of application for these types.

A consumer may try again until they are happy if they are not satisfied with the service they received. Takacs [14] investigates this at first, allowing the consumer who has finished the service to provide feedback to the rear of the line. An M/(G1, G2)/1 feedback retrial queue with two phase service, variant vacation policy under delaying repair for impatient Customers was analysed by Rajadurai et al. [15].

Many real-world systems have impatient customers as a built-in feature, particularly when the customer is a human, a perishable product, or some moving object that can depart the service area and their waiting period in the queue reaches certain pre-defined threshold values. This clearly explains why queueing literature frequently discusses the impatience phenomenon. Accounting impatience is crucial in the setting of lines for group service because a client could spend a large amount of time in the system while waiting for the accumulation of a sufficient number of customers.

More focus has been placed on the numerous retrial lineups with non-persistent (impatient) consumers. A discussion about the study of a retrial queue with group service of impatient clients involved D'rienzo et al. [16]. A batch arrival retrial queuing model with starting failures and customer impatience was addressed by Nila and Sumitha [17]. Customers arrive in batches in line with the Poisson process. In certain situations, the clients refuse and break their promises. The analysis of a retrial QS with priority services, working breakdown, BV, admission control, and balking was explained by Ayyappan et al. [18]. Ayyappan and Nirmala [19] explored an analysis of customer's impatience on bulk service QS's with an unreliable server, setup time and two types of multiple vacations. Sethi.R et al. [20] investigated the cost optimization and ANFIS computing of an unreliable M/M/1 queueing system with customers' impatience under n-policy. The ideal Cost Analysis for Discrete-Time Recurrent Queue with Bernoulli Feedback and Emergency Vacation was described by M. Vaishnawi [21]. In order to calculate costs, PSO, ABC, and GA are also used. To ensure the best deal, these methods compare and contrast the outputs.

The paper's structure is as follows: Section 2 provides a detailed explanation of the mathematical model. Section 3 discusses the ideas and formulae governing our system as well as how to obtain the time-dependent solution of our model. The PGF for the queue length at each given epoch and the SS performance of the system are explicitly determined in Section 4. In Section 5, the pertinent stability condition has been uncovered. In Section 6, we precisely estimate the mean queue size, mean queue waiting time, and efficiency features for each state of the system. In Section 7, we present a practical illustration. We offer a numerical study and associated graphs in Section 8. Furthermore, an ANFIS was provided in Section 9. The Cost optimization is offered by Section 10. The conclusion is presented in Section 11.

2. Model Description and Analysis

We suppose that the underlying queueing model is as follows:
Arrival process: Customers enter a poisson stream, and bulk service is offered on an FCFS basis. Considering that a batch of "i"customers enters the system, $\Lambda > 0$ represents the average batch arrival rate, and $\Lambda c_i dt (i \geq 1)$ represents the first order probability during the short interval of time $(\varpi, \varpi + d\varpi]$. We define a batch arrival and a bulk service as having a smallest batch size of "a"and a highest batch size of "b".
Retrial process: When a customer arrives and discovers that the server is busy, unavailable, or broken, the customer has two options: (1) leave the service area with a probability of d and join a pool of blocked customers known as an orbit; or (2) balk the system with a probability of $\bar{d}$ in accordance with FCFS, which implies that only the customer at the head of the orbit queue is permitted access to the server.
When the server is idle, the customer at the head of the retrial queue engages with potential primary customers to see who can cancel their service request and, with prob., g, either move up in the retrial queue or leave the system with prob., $(1-g)$.
A general (arbitray) distribution with the distribution function $A(u)$ and the density function $a(u)$ determines the retrial interval.
Let $g(\varsigma)d\varsigma$ be the conditional prob., density of completing the retrial within the range $(\varsigma, \varsigma + d\varsigma]$, where ς is the elapsed retrial time.

$$g(\varsigma) = \frac{a(\varsigma)}{1 - A(\varsigma)}$$

and therefore,

$$a(u) = g(u)e^{-\int_0^u g(\varsigma)d\varsigma}$$

Inter-retrial times have an arbitrary dist., $A(\varsigma)$ with correponding Laplace-Stieltijes transforms (LST) $A^*(u)$.
Service process: The server enters an idle state wherever a fresh or returning user comes before quickly resuming regular operations for the newcomers. A generic (arbitrary) distance with the distance function $B(\varsigma)$ and the density function $b(\varsigma)$ follows the service time.
Given the elapsed retrial time ς, define $\phi(\varsigma)d\varsigma$ as the conditional probability of service completion within the range $(\varsigma, \varsigma + d\varsigma]$.

$$\phi(\varsigma) = \frac{b(\varsigma)}{1 - B(\varsigma)}$$

and therefore,

$$b(\varpi) = \phi(\varpi)e^{-\int_0^{\varpi} \phi(\varsigma)d\varsigma}$$

The random variable B with the dist., function $B(\varsigma)$ and LST $B^*(\varpi)$ denotes the service time.
Random failure: Failures are anticipated to occur sporadically throughout the system and ought to follow a poisson stream with an average failure rate of $\tau > 0$. The repair times follow a general dist.,, which is represented by the random variable D and the dist., function $D(\varsigma)$, with the LST $D^*(\varpi)$.
The length of repairs is determined by a general (arbitrary) dist., with a dist., function $D(\varsigma)$ and a density function $d(\varsigma)$. Given an elapsed repair time of ς, define $\alpha(\varsigma)d\varsigma$ as the conditional probability of completing repairs within the range $(\varsigma, \varsigma + d\varsigma]$.

$$\alpha(\varsigma) = \frac{d(\varsigma)}{1 - D(\varsigma)}$$

and therefore,

$$d(\varpi) = \alpha(\varpi)e^{-\int_0^{\varpi} \alpha(\varsigma)d\varsigma}$$

Extended Bernoulli vacation: If there are any unfinished parts of the service, the server has two options: either accept the BV with a probability of θ or keep serving them with a probability of $(1 - \theta)$. After the vacation is over, the server either undertakes the second type of optional extended Bernoulli vacation with a prob., of μ or continues to serve the remaining batches with a prob., of $(1 - \mu)$.
The random variable F with the distance function $F(\varsigma)$ and LST $F^*(\varpi)$ is employed to represent the server's leisure time. This arbitrary variable F follows a general distribution.
The server's vacation time follows a general(arbitrary) dist., function $F(\varpi)$ and density function $f(\varpi)$. Let $\beta(\varsigma)d\varsigma$ be the conditional prob., of a completion of a vacation during the interval $(\varsigma, \varsigma + d\varsigma]$, given that the elapsed repair time is ς, so that

$$\beta(\varsigma) = \frac{f(\varsigma)}{1 - F(\varsigma)}$$

and therefore,

$$f(\varpi) = \beta(\varpi)e^{-\int_0^{\varpi} \beta(\varsigma)d\varsigma}$$

The system's stochastic processes are all considered to be independent of one another.
Feedback Rule: Clients who are unhappy with their offerings can re-join the line once they've been completed, give feedback to receive another service with minimal difficulty, or both p $(0 \leq p \leq 1)$, otherwise the system must be terminated with complement prob. $q = (1 - p)$

3. DEFINITIONS:

We define
1. $P_n(\varsigma, \varpi)$= Prob., that the server will be idle at time ϖ with $n(n \geq 0)$ customers in the orbit and ς for the customer's elapsed retrial time.
2. $Q_n(\varsigma, \varpi)$= Prob., that the server will be busy at time ϖ with $n(n \geq 0)$ customers in the orbit and η for the customer's elapsed retrial time.
3. $R_n(\varsigma, \varpi)$= Prob., that at time ϖ, there are $n(n \geq 0)$ customers in the orbit and the server is offline due to system repair and waiting for repairs to start with elapsed repair time ς.

4. $V_n(\varsigma,\varpi)$= Prob., that there are $n(n \geq 0)$ consumers in orbit at time ϖ and the server is on vacation with elapsed vacation time ς.
5. There are no customers in the orbit at time ϖ, and the server is inactive but still available in the system, according to the probability $P_0(\varpi)$.
The following differential-difference equations regulate the model:

$$\frac{d}{d\varpi}P_0(\varpi) = -\Lambda P_0(\varpi) + (1-\theta)\bar{d}\int_0^\infty Q_0(\varsigma,\varpi)\phi(\varsigma)d\varsigma + (1-\mu)\int_0^\infty V_0(\varsigma,\varpi)\beta(\varsigma)d\varsigma \tag{1}$$

$$\frac{\partial}{\partial\varsigma}P_n(\varsigma,\varpi) + \frac{\partial}{\partial\varpi}P_n(\varsigma,\varpi) = -[\Lambda + g(\varsigma)]P_n(\varsigma,\varpi), n \geq 1 \tag{2}$$

$$\frac{\partial}{\partial\varsigma}Q_0(\varsigma,\varpi) + \frac{\partial}{\partial\varpi}Q_0(\varsigma,\varpi) = -[\Lambda + \tau + \phi(\varsigma)]Q_0(\varsigma,\varpi) \tag{3}$$

$$\frac{\partial}{\partial\varsigma}Q_n(\varsigma,\varpi) + \frac{\partial}{\partial\varpi}Q_n(\varsigma,\varpi) = -[\Lambda + \tau + \phi(\varsigma)]Q_n(\varsigma,\varpi) + \Lambda\sum_{k=1}^{n} C_k Q_{n-k}(\varsigma,\varpi), n \geq 1 \tag{4}$$

$$\frac{\partial}{\partial\varsigma}R_0(\varsigma,\iota,\varpi) + \frac{\partial}{\partial\varpi}R_0(\varsigma,\varpi) = -[\Lambda + \alpha(\varsigma)]R_0(\varsigma,\varpi), n = 0 \tag{5}$$

$$\frac{\partial}{\partial\varsigma}R_n(\varsigma,\iota,\varpi) + \frac{\partial}{\partial\varpi}R_n(\varsigma,\varpi) = -[\Lambda + \alpha(\varsigma)]R_n(\varsigma,\varpi) + \Lambda\sum_{k=1}^{n} C_k R_{n-k}(\varsigma,\varpi), n \geq 1 \tag{6}$$

$$\frac{\partial}{\partial\varsigma}V_0(\varsigma,\varpi) + \frac{\partial}{\partial\varpi}V_0(\varsigma,\varpi) = -[\Lambda + \beta(\varsigma)]V_0(\varsigma,\varpi), n = 0 \tag{7}$$

$$\frac{\partial}{\partial\varsigma}V_n(\varsigma,\varpi) + \frac{\partial}{\partial\varpi}V_n(\varsigma,\varpi) = -[\Lambda + \beta(\varsigma)]V_n(\varsigma,\varpi) + \Lambda\sum_{k=1}^{n} C_k V_{n-k}(\varsigma,\varpi), n \geq 1 \tag{8}$$

The following boundary conditions must be met in order to answer the given equation:

$$\begin{aligned} P_n(0,\varpi) =&(1-\theta)\bar{d}\int_0^\infty Q_n(\varsigma,\varpi)\phi(\varsigma)d\varsigma + (1-\theta)d\int_0^\infty Q_{n-1}(\varsigma,\varpi)\phi(\varsigma)d\varsigma \\ &+ \int_0^\infty R_n(\varsigma,\varpi)\alpha(\varsigma)d\varsigma + (1-\mu)\int_0^\infty V_n(\varsigma,\varpi)\beta(\varsigma)d\varsigma, n \geq 1 \end{aligned} \tag{9}$$

$$\begin{aligned} Q_0(0,\varpi) =&\Lambda p(1-g)\sum_{r=a}^{b}\sum_{k=0}^{a-1} C_k\int_0^\infty P_{n-k+b}(\varsigma,\varpi)d\varsigma \\ &+ (1-\theta)p\sum_{r=a}^{b}\int_0^\infty P_r(\varsigma,\varpi)g(\varsigma)d\varsigma + \sum_{r=a}^{b}\int_0^\infty V_r(\varsigma,\varpi)\beta(\varsigma)d\varsigma \end{aligned} \tag{10}$$

$$\begin{aligned} Q_n(0,\varpi) =&\Lambda p(1-g)\sum_{k=0}^{a-1} C_k\int_0^\infty P_{n-k+b}(\varsigma,\varpi)d\varsigma + p\int_0^\infty P_{n+b}(\varsigma,\varpi)g(\varsigma)d\varsigma \\ &+ \Lambda g\int_0^\infty P_{n+b}(\varsigma,\varpi)d\varsigma + \int_0^\infty V_{n+b}(\varsigma,\varpi)\beta(\varsigma)d\varsigma \end{aligned} \tag{11}$$

$$R_0(\varsigma,0,\varpi) = \tau Q_0(\varsigma,\varpi), n = 0 \tag{12}$$

$$R_n(\varsigma,0,\varpi) = \tau Q_n(\varsigma,\varpi), n \geq 1 \tag{13}$$

$$V_n(0,\varpi) = \theta\int_0^\infty Q_n(\varsigma,\varpi)\phi(\varsigma)d\varsigma, n \geq 1 \tag{14}$$

We presume that the system is initially empty of users and that the server is idle. Thus, the initial conditions are

$$\begin{aligned} V_n(0) = R_n(0) = Q_n(0) = 0, n \geq 0 \\ P_0(0) = 1, P_n^i(0) = 0, n \geq 1 \end{aligned} \tag{15}$$

Generating functions of the queue length (The time-dependent solution):

$$\begin{aligned}
P(\varsigma,\Psi,\varpi) &= \sum_{n=0}^{\infty}\Psi^n P_n^i(\varsigma,\varpi); P(\Psi,\varpi) = \sum_{n=0}^{\infty}\Psi^n P_n(\varpi)\\
Q(\varsigma,\Psi,\varpi) &= \sum_{n=0}^{\infty}\Psi^n Q_n(\varsigma,\varpi);\ Q(\Psi,\varpi) = \sum_{n=0}^{\infty}\Psi^n Q_n(\varpi)\\
R(\varsigma,\iota,\Psi,\varpi) &= \sum_{n=0}^{\infty}\Psi^n R_n(\varsigma,\iota,\varpi);\ R(\varsigma,\Psi,\varpi) = \sum_{n=0}^{\infty}\Psi^n R_n(\varsigma,\varpi)\\
V(\varsigma,\Psi,\varpi) &= \sum_{n=0}^{\infty}\Psi^n V_n(\varsigma,\varpi);\ V(\Psi,\varpi) = \sum_{n=0}^{\infty}\Psi^n V_n(\varpi)\\
C(\Psi) &= \sum_{n=1}^{\infty} C_n\Psi^n;\ Q(\Psi) = \sum_{r=0}^{a-1} Q_r\Psi^r
\end{aligned} \tag{16}$$

which define the LT of a function $f(\varpi)$ as it converges within the circle defined by $z \leq 1$.

$$\bar{f}(s) = \int_0^{\infty} e^{-s\varpi} f(\varpi)d\varpi, \mathcal{R}(s) \geq 0 \tag{17}$$

Using (15) and the LT from equations (1) through (14), we arrive at

$$(s+\Lambda)\bar{p}_0(s) = 1 + (1-\theta)\bar{d}\int_0^{\infty}\bar{Q}_0(\varsigma,s)\phi(\varsigma)d\varsigma + (1-\mu)\int_0^{\infty}\bar{V}_0(\varsigma,s)\beta(\varsigma)d\varsigma \tag{18}$$

$$\frac{\partial}{\partial\varsigma}\bar{P}_n(\varsigma,s) + [s+\Lambda+g(\varsigma)]\bar{P}_n(\varsigma,s) = 0, n \geq 1 \tag{19}$$

$$\frac{\partial}{\partial\varsigma}\bar{Q}_0(\varsigma,s) + [s+\Lambda+\phi(\varsigma)]\bar{Q}_0(\varsigma,s) = 0 \tag{20}$$

$$\frac{\partial}{\partial\varsigma}\bar{Q}_n(\varsigma,s) + [s+\Lambda+\phi(\varsigma)]\bar{Q}_n(\varsigma,s) = \Lambda\sum_{k=1}^{n} C_k\bar{Q}_{n-k}(\varsigma,s), n \geq 1 \tag{21}$$

$$\frac{\partial}{\partial\varsigma}\bar{R}_0(\varsigma,\iota,s) + [s+\Lambda+\alpha(\varsigma)]\bar{R}_0(\varsigma,s) = 0 \tag{22}$$

$$\frac{\partial}{\partial\varsigma}\bar{R}_n(\varsigma,\iota,s) + [s+\Lambda+\alpha(\varsigma)]\bar{R}_n(\varsigma,s) = \Lambda\sum_{k=1}^{n} C_k\bar{R}_{n-k}(\varsigma,s), n \geq 1 \tag{23}$$

$$\frac{\partial}{\partial\varsigma}\bar{V}_0(\varsigma,s) + [s+\Lambda+\beta(\varsigma)]\bar{V}_0(\varsigma,s) = 0 \tag{24}$$

$$\frac{\partial}{\partial\varsigma}\bar{V}_n(\varsigma,s) + [s+\Lambda+\beta(\varsigma)]\bar{V}_n(\varsigma,s) = \Lambda\sum_{k=1}^{n} C_k\bar{V}_{n-k}(\varsigma,s), n \geq 1 \tag{25}$$

$$\bar{P}_n(0,s) = (1-\theta)\bar{d}\int_0^\infty \bar{Q}_n(\varsigma,s)\phi(\varsigma)d\varsigma + (1-\theta)d\int_0^\infty \bar{Q}_{n-1}(\varsigma,s)\phi(\varsigma)d\varsigma + \int_0^\infty \bar{R}_n(\varsigma,s)\alpha(\varsigma)d\varsigma + (1-\mu)\int_0^\infty \bar{V}_n(\varsigma,s)\beta(\varsigma)d\varsigma, n \geq 1 \tag{26}$$

$$\bar{Q}_0(0,s) = \Lambda p(1-g)\sum_{r=a}^{b}\sum_{k=0}^{a-1} C_k \int_0^\infty \bar{P}_{n-k+b}(\varsigma,s)d\varsigma + (1-\theta)p\sum_{r=a}^{b}\int_0^\infty \bar{P}_r(\varsigma,s)g(\varsigma)d\varsigma + \sum_{r=a}^{b}\int_0^\infty \bar{V}_r(\varsigma,s)\beta(\varsigma)d\varsigma \tag{27}$$

$$\bar{Q}_n(0,s) = \Lambda p(1-g)\sum_{k=0}^{a-1} C_k \int_0^\infty \bar{P}_{n-k+b}(\varsigma,s)d\varsigma + p\int_0^\infty \bar{P}_{n+b}(\varsigma,s)g(\varsigma)d\varsigma + \Lambda g\int_0^\infty \bar{P}_{n+b}(\varsigma,s)d\varsigma + \int_0^\infty \bar{V}_{n+b}(\varsigma,s)\beta(\varsigma)d\varsigma \tag{28}$$

$$\bar{R}_0(\varsigma,0,s) = \tau\bar{Q}_0(\varsigma,s), n = 0 \tag{29}$$

$$\bar{R}_n(\varsigma,0,s) = \tau\bar{Q}_n(\varsigma,s), n \geq 1 \tag{30}$$

$$\bar{V}_n(0,s) = \theta\int_0^\infty \bar{Q}_n(\varsigma,s)\phi(\varsigma)d\varsigma, n \geq 1 \tag{31}$$

By multiplying equations (19) through (31) by Ψ^n and adding the results over n, we can obtain using the generating function mentioned in equation (16).

$$\frac{\partial}{\partial\varsigma}\bar{P}(\varsigma,\Psi,s) + [s+\Lambda+g(\varsigma)]\bar{P}(\varsigma,\Psi,s) = 0 \tag{32}$$

$$\frac{\partial}{\partial\varsigma}\bar{Q}(\varsigma,\Psi,s) + [s+\Lambda(1-C(\Psi))+\phi(\varsigma)]\bar{Q}(\varsigma,\Psi,s) = 0 \tag{33}$$

$$\frac{\partial}{\partial\varsigma}\bar{R}(\varsigma,\iota,\Psi,s) + [s+\Lambda(1-C(\Psi))+\alpha(\varsigma)]\bar{R}(\varsigma,\Psi,s) = 0 \tag{34}$$

$$\frac{\partial}{\partial\varsigma}\bar{V}(\varsigma,\Psi,s) + [s+\Lambda(1-C(\Psi))+\beta(\varsigma)]\bar{V}(\varsigma,\Psi,s) = 0 \tag{35}$$

$$\bar{P}(0,\Psi,s) = (1-\theta)(\bar{d}+d\Psi)\int_0^\infty \bar{Q}(\varsigma,\Psi,s)\phi(\varsigma)d\varsigma + \int_0^\infty \bar{R}(\varsigma,\Psi,s)\alpha(\varsigma)d\varsigma + (1-\mu)\int_0^\infty \bar{V}(\varsigma,\Psi,s)\beta\varsigma d\varsigma - \bar{d}(1-\theta)\int_0^\infty \bar{Q}_0(\varsigma,s)\phi(\varsigma)d\varsigma - (1-\mu)\int_0^\infty \bar{V}_0(\varsigma,s)\beta(\varsigma)d\varsigma, n \geq 1 \tag{36}$$

$$\Psi^b\bar{Q}(0,\Psi,s) = \Lambda(1-g)pC(\Psi)\int_0^\infty \bar{P}(\varsigma,\Psi,s)d\varsigma + p\int_0^\infty \bar{P}(\varsigma,\Psi,s)g(\varsigma)d\varsigma + \Lambda g\int_0^\infty \bar{P}(\varsigma,\Psi,s)d\varsigma + \int_0^\infty \bar{V}(\varsigma,\Psi,s)\beta(\varsigma)d\varsigma \tag{37}$$

$$\bar{R}(\varsigma,0,\Psi,s) = \tau\bar{Q}(\varsigma,\Psi,s), n \geq 1 \tag{38}$$

$$\bar{V}(0,\Psi,s) = \theta\int_0^\infty \bar{Q}(\varsigma,\Psi,s)\phi(\varsigma)d\varsigma, n \geq 1 \tag{39}$$

Equation (18) in (36) gives us

$$\bar{P}(0,\Psi,s) = [1-(s+\Lambda)\bar{P}_0(s)] + (1-\theta)(\bar{d}+d\Psi)\int_0^\infty \bar{Q}(\varsigma,\Psi,s)\phi(\varsigma)d\varsigma + \int_0^\infty \bar{R}(\varsigma,\Psi,s)\alpha(\varsigma)d\varsigma + (1-\mu)\int_0^\infty \bar{V}(\varsigma,\Psi,s)\beta(\varsigma)d\varsigma \tag{40}$$

Equation (32), when integrated between 0 and ς, yields

$$\bar{P}(\varsigma,\Psi,s) = \bar{P}(0,\Psi,s)e^{-(s+\Lambda)\varsigma-\int_0^\varsigma g(\omega)d\omega} \tag{41}$$

Once more, integrating equation (41) by parts with respect to ς yiedls,

$$\bar{P}(\Psi,s) = \bar{P}(0,\Psi,s)\left[\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right] \tag{42}$$

where,

$$\bar{A}(s+\Lambda) = \int_0^\infty e^{-(s+\Lambda)\varsigma}dA(\varsigma)$$

When integrating equations (33) to (35) from 0 to ς, similar outcomes are found.

$$\bar{Q}(\varsigma,\Psi,s) = \bar{Q}(0,\Psi,s)e^{-\zeta(\Psi,s)\varsigma-\int_0^\varsigma \phi(\omega)d\omega} \tag{43}$$

$$\bar{R}(\varsigma,\iota,\Psi,s) = \bar{R}(\varsigma,0,\Psi,s)e^{-\zeta(\Psi,s)\varsigma-\int_0^\varsigma \alpha(\omega)d\omega}$$

$$\bar{R}(\varsigma,\Psi,s) = \bar{R}(\varsigma,0,\Psi,s)\left[\frac{1-\bar{D}(\zeta(\Psi,s))}{\zeta(\Psi,s)}\right] \tag{44}$$

$$\bar{V}(\varsigma,\Psi,s) = \bar{V}(0,\Psi,s)e^{-\zeta(\Psi,s)\varsigma-\int_0^\varsigma \beta(\omega)d\omega} \tag{45}$$

where the values of $\bar{P}(0,\Psi,s)$,$\bar{Q}(0,\Psi,s)$,$\bar{R}(0,\Psi,s)$ and $\bar{V}(0,\Psi,s)$ are given by (37) to (40).
Taking into account ς yiedls, integrate equations (43) to (45) by parts once more.

$$\bar{Q}(\Psi,s) = \bar{Q}(0,\Psi,s)\left[\frac{1-\bar{B}(\zeta(\Psi,s))}{\zeta(\Psi,s)}\right] \tag{46}$$

$$\bar{R}(\Psi,s) = \tau\bar{Q}(0,\Psi,s)\left[\frac{1-\bar{B}(\zeta(\Psi,s))}{\zeta(\Psi,s)}\right]\left[\frac{1-\bar{D}(\zeta(\Psi,s))}{\zeta(\Psi,s)}\right] \tag{47}$$

$$\bar{V}(\Psi,s) = \bar{V}(0,\Psi,s)\left[\frac{1-\bar{F}(\zeta(\Psi,s))}{\zeta(\Psi,s)}\right] \tag{48}$$

Where,

$$\bar{B}(\zeta(\Psi,s)) = \int_0^\infty e^{-\zeta(\Psi,s)\varsigma}dB(\varsigma)$$

$$\bar{D}(\zeta(\Psi,s)) = \int_0^\infty e^{-\zeta(\Psi,s)\varsigma}dD(\varsigma)$$

$$\bar{F}(\zeta(\Psi,s)) = \int_0^\infty e^{-\zeta(\Psi,s)\varsigma}dF(\varsigma)$$

are, in order, the LST of the following values: retrial time $A(\varsigma)$, service time $B(\varsigma)$, repair time $D(\varsigma)$, and vacation time $F(\varsigma)$.
Now, multiplying both side of equations (41),(43) to (45) by $g(\varsigma)$,$\phi(\varsigma)$,$\alpha(\varsigma)$ and $\beta(\varsigma)$ and integrating over ς, we obtain

$$\int_0^\infty \bar{P}(\varsigma,\Psi,s)g(\varsigma)d\varsigma = \bar{P}(0,\Psi,s)\bar{A}(s+\Lambda) \tag{49}$$

$$\int_0^\infty \bar{Q}(\varsigma,\Psi,s)\phi(\varsigma)d\varsigma = \bar{Q}(0,\Psi,s)\bar{B}(\zeta(\Psi,s)) \tag{50}$$

$$\int_0^\infty \bar{R}(\varsigma,\iota,\Psi,s)\alpha(\varsigma)d\varsigma = \bar{R}(\varsigma,0,\Psi,s)\bar{D}(\zeta(\Psi,s)) \tag{51}$$

$$\int_0^\infty \bar{V}(\varsigma,\Psi,s)\beta(\varsigma)d\varsigma = \bar{V}(0,\Psi,s)\bar{F}(\zeta(\Psi,s)) \tag{52}$$

Using equations (50) in (39)

$$\bar{V}(0,\Psi,s) = \theta\bar{Q}(0,\Psi,s)\bar{B}(\zeta(\Psi,s)) \tag{53}$$

Using equations (49) in (37) and (38), we get

$$\bar{Q}(0,\Psi,s) = \frac{\bar{P}(0,\Psi,s)}{\Psi^b - \theta\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s))}$$
$$\left[\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right) + p\bar{A}(s+\Lambda) + \Lambda g\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right)\right] \tag{54}$$

$$\bar{R}(\varsigma,0,\Psi,s) = \tau\bar{Q}(0,\Psi,s)\left(\frac{1-\bar{B}(\zeta(\Psi,s))}{(\zeta(\Psi,s))}\right) \tag{55}$$

Using equation (50) to (52) in (40) we get

$$\bar{P}(0,\Psi,s) = \frac{Nr(\Psi)}{Dr(\Psi)} \tag{56}$$

$$\begin{aligned}
Nr(\Psi) =& [1-(s+\Lambda)\bar{P}_0(s)][\Psi^b - \theta\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s))] \\
Dr(\Psi) =& \Psi^b - \theta\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s)) \\
&- \left[\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right) + p\bar{A}(s+\Lambda) + \Lambda g\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right)\right] \\
&\left[(1-\theta)(\bar{d}+d\Psi)\bar{B}(\zeta(\Psi,s)) + \tau\bar{D}(\zeta(\Psi,s))\left(\frac{1-\bar{B}(\zeta(\Psi,s))}{(\zeta(\Psi,s))}\right)\right. \\
&\left. + \theta(1-\mu)\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s))\right]
\end{aligned}$$

where,

$$\zeta(\Psi,s) = s + \Lambda(1-C(\Psi))$$

Subs/- $\bar{P}(0,\Psi,s)$ from equation (56) into equation (53) to (55)

$$\bar{Q}(0,\Psi,s) = \left[\frac{\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right) + p\bar{A}(s+\Lambda) + \Lambda g\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right)}{\Psi^b - \theta\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s))}\right]$$
$$\left[\frac{Nr(\Psi)}{Dr(\Psi)}\right] \tag{57}$$

$$\begin{aligned}
\bar{R}(\varsigma,0,\Psi,s) =& \tau\left(\frac{1-\bar{B}(\zeta(\Psi,s))}{(\zeta(\Psi,s))}\right)\left[\frac{Nr(\Psi)}{Dr(\Psi)}\right] \\
&\left[\frac{\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right) + p\bar{A}(s+\Lambda) + \Lambda g\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right)}{\Psi^b - \theta\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s))}\right]
\end{aligned} \tag{58}$$

$$\begin{aligned}
\bar{V}(0,\Psi,s) =& \theta\bar{B}(\zeta(\Psi,s))\left[\frac{Nr(\Psi)}{Dr(\Psi)}\right] \\
&\left[\frac{\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right) + p\bar{A}(s+\Lambda) + \Lambda g\left(\frac{1-\bar{A}(s+\Lambda)}{s+\Lambda}\right)}{\Psi^b - \theta\bar{F}(\zeta(\Psi,s))\bar{B}(\zeta(\Psi,s))}\right]
\end{aligned} \tag{59}$$

Updating equations (56) to (59) in (42), (46) to (48) We determine the PGF of various conditions in the system under a transient condition.

4. The Steady state's findings:

To define the SS prob., we disregard the argument ϖ wherever it appears in the time-dependent analysis.

$$\lim_{s\to 0} s\bar{f}(s) = \lim_{\varpi\to\infty} f(\varpi)$$

$$P(\Psi) = P(0,\Psi)\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right) \tag{60}$$

$$Q(\Psi) = \left(\frac{1-\bar{B}(\zeta(\Psi))}{\zeta(\Psi)}\right)P(0,\Psi) \tag{61}$$

$$\left[\frac{\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)+p\bar{A}(\Lambda)+\Lambda g\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)}{\Psi^b-\theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))}\right] \tag{62}$$

$$R(\Psi) = \tau\left(\frac{1-\bar{B}(\zeta(\Psi))}{\zeta(\Psi)}\right)\left(\frac{1-\bar{D}(\zeta(\Psi))}{\zeta(\Psi)}\right)$$
$$P(0,\Psi)\left[\frac{\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)+p\bar{A}(\Lambda)+\Lambda g\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)}{\Psi^b-\theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))}\right] \tag{63}$$

$$V(\Psi) = \theta\bar{B}(\zeta(\Psi))\left(\frac{1-\bar{F}(\zeta(\Psi))}{\zeta(\Psi)}\right)$$
$$P(0,\Psi)\left[\frac{\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)+p\bar{A}(\Lambda)+\Lambda g\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)}{\Psi^b-\theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))}\right] \tag{64}$$

where,

$$P(0,\Psi) = \frac{Nr(\Psi)}{Dr(\Psi)}$$
$$Nr(\Psi) = [1-\Lambda\bar{P}_0][\Psi^b-\theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))]$$
$$Dr(\Psi) = \Psi^b-\theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))$$
$$-\left[\Lambda(1-g)pC(\Psi)\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)+p\bar{A}(\Lambda)+\Lambda g\left(\frac{1-\bar{A}(\Lambda)}{\Lambda}\right)\right]$$
$$\left[(1-\theta)(\bar{d}+d\Psi)\bar{B}(\zeta(\Psi))+\tau\bar{D}(\zeta(\Psi))\left(\frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))}\right)\right.$$
$$\left.+\theta(1-\mu)\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))\right] \tag{65}$$

4.1. Queue sizes distribution at a certain epoch:

The PGF is a of the queue size dist., at a random interval, is obtained by adding (60) to (63) with the idle term.

$$K(\Psi) = \frac{Nr(\Psi)}{Dr(\Psi)} \tag{66}$$

$$\begin{aligned}
Nr(\Psi) &= \Lambda P_0 \zeta(\Psi) \Big(\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)\,(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) \\
&+ g\,(1-\bar{A}(\Lambda))]\left[(1-\theta)(\bar{d}+d\Psi)\bar{B}(\zeta(\Psi)) + \tau\bar{D}(\zeta(\Psi))\left(\frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))}\right)\right. \\
&\left.+\theta(1-\mu)\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))\right]) - (1-\bar{A}(\Lambda))\zeta(\Psi)[\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) \\
&+ \Lambda\,[(1-g)pC(\Psi)\,(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g\,(1-\bar{A}(\Lambda))] \\
&[(1-\bar{B}\zeta(\Psi)) + \tau(1-\bar{B}\zeta(\Psi))(1-\bar{D}\zeta(\Psi)) + \theta\bar{B}\zeta(\Psi)(1-\bar{F}\zeta(\Psi))]] \\
&+ (1-\bar{A}(\Lambda))\zeta(\Psi)[\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) \\
&+ \Lambda\,[(1-g)pC(\Psi)\,(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g\,(1-\bar{A}(\Lambda))]] \\
&[(1-\bar{B}\zeta(\Psi)) + \tau(1-\bar{B}\zeta(\Psi))(1-\bar{D}\zeta(\Psi)) + \theta\bar{B}\zeta(\Psi)(1-\bar{F}\zeta(\Psi))]] \\
Dr(\Psi) &= \zeta(\Psi)\Lambda\Big\{\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)\,(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) \\
&+ g\,(1-\bar{A}(\Lambda))]\left((1-\theta)(\bar{d}+d\Psi)\bar{B}(\zeta(\Psi)) + \tau\bar{D}(\zeta(\Psi))\left(\frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))}\right)\right. \\
&\left.+\theta(1-\mu)\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))\right)\Big\}
\end{aligned}$$

5. Stability Condition

The PGF needs to meet P(1)=1. Applying the L'Hopital rules and equating the expression to 1 results in the result that satisfies the requirement.

$$\begin{aligned}
&b - [(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)] + p(1-g)(1-\bar{A}(\Lambda)) \\
&+ p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))[(1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B)) + \tau E(B) \\
&- \theta\Lambda(1-\mu)E(I)A_1] + \Lambda\theta E(I)A_1
\end{aligned}$$

Now we can determine the prob., that are unknown. P(1)=1 is therefore fulfilled if

$$\begin{aligned}
&\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)\,(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g\,(1-\bar{A}(\Lambda))] \\
&\left[(1-\theta)(\bar{d}+d\Psi)\bar{B}(\zeta(\Psi)) + \tau\bar{D}(\zeta(\Psi))\left(\frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))}\right) + \theta(1-\mu)\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))\right] > 0
\end{aligned}$$

$$\begin{aligned}
\rho = \frac{\begin{aligned}&[(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)] + [p(1-g)(1-\bar{A}(\Lambda)) \\ &+p\bar{A}(\Lambda) + g((1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B)) \\ &+ \tau E(B) - \theta\Lambda(1-\mu)E(I)A_1] + \Lambda\theta E(I)A_1\end{aligned}}{b}
\end{aligned} \tag{67}$$

then $\rho < 1$ is the condition to be satisfied for the existence of the SS for the model under consideration.

6. Performance Evaluation:

This section includes system performance metrics, a model stability study, and some unique system prob., while the system is in various states.
We obtain the following prob., if the system fulfills the stability requirement $\rho < 1$.

- Let P be the SS Prob., that the server is idle during the retrial time.

$$P = \lim_{\Psi \to 1} P(\Psi) = P(1) = \frac{(1-\theta)(1-\Lambda p_0)(1-\bar{A}(\Lambda))}{\Lambda(1-\theta) - [p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] \\ [(1-\theta)(d+\bar{d}) + \theta(1-\mu)]}$$

- If the server is busy, let Q be the SS Prob.,

$$\begin{aligned} Q &= \lim_{\Psi \to 1} Q(\Psi) \\ Q(1) &= E(B) \times \\ &\left\{ \begin{array}{c} \dfrac{(1-\Lambda p_0)[p(1-g)(1-\bar{A}(\Lambda))E(I)]}{b + \Lambda\theta E(I)A_1 - [(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)]} \\ +[p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] \\ (1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B)) + \tau E(B) - \Lambda\theta(1-\mu)E(I)A_1 \end{array} \right\} \end{aligned}$$

- R ought to indicate the SS Prob., that the server is being repaired.

$$\begin{aligned} R &= \lim_{\Psi \to 1} R(\Psi) \\ R(1) &= \tau E(B)E(D) \times \\ &\left\{ \begin{array}{c} \dfrac{(1-\Lambda p_0)[p(1-g)(1-\bar{A}(\Lambda))E(I)]}{b + \Lambda\theta E(I)A_1 - ([(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)]} \\ +[p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))] \\ [(1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B)) + \tau E(B) - \Lambda\theta(1-\mu)E(I)A_1]) \end{array} \right\} \end{aligned}$$

- Using V as the SS Prob., we may assume that the server is on vacation.

$$\begin{aligned} V &= \lim_{\Psi \to 1} V(\Psi) \\ V(1) &= \theta E(F)E(I) \times \\ &\left\{ \begin{array}{c} \dfrac{(1-\Lambda p_0)(-\Lambda E(B)p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda)) \\ + p(1-g)((1-\bar{A}(\Lambda))))}{b + \Lambda\theta E(I)A_1 - ([(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)]} \\ +[p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) \\ (1-\Lambda E(I)E(B)) + \tau E(B) - \Lambda\theta(1-\mu)E(I)A_1]) \end{array} \right\} \end{aligned}$$

6.1. **Average queue length:**

Computing at $\Psi = 1$ and differentiating (65) with regard to Ψ yields the mean number of users in the queue (L_q) under SS conditions.

$$L_q = \lim_{\Psi \to 1} \frac{d}{d\Psi} P(\Psi)$$

$$P'(1) = \frac{Nr''(1)Dr'(1) - Dr''(1)Nr'(1)}{2(Dr'(1))^2}$$

$$
\begin{aligned}
D^{'}(1) = &-\Lambda^2 E(I)\left\{\Psi^b - \theta\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi)) - [(1-g)pC(\Psi)\,(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda)\right.\\
&+g\,(1-\bar{A}(\Lambda))]\left[(1-\theta)(\bar{d}+d\Psi)\bar{B}(\zeta(\Psi)) + \tau\bar{D}(\zeta(\Psi))\left(\frac{1-\bar{B}(\zeta(\Psi))}{(\zeta(\Psi))}\right)\right.\\
&\left.\left.+\theta(1-\mu)\bar{F}(\zeta(\Psi))\bar{B}(\zeta(\Psi))\right]\right\}\\
D^{''}(1) = &-\Lambda^2\left\{E(I(I-1))[1-\theta-[p-g(p-1)(1-\bar{A}(\Lambda))][1-\theta\mu]\right.\\
&+2E(I)[b+\Lambda\theta E(I)A_1 - \left([(1-g)pE(I)(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d}) + \theta(1-\mu)\right]\\
&+[p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d})(1-\Lambda E(I)E(B))\\
&\left.+\tau E(B) - \Lambda\theta(1-\mu)E(I)A_1])]\right\}
\end{aligned}
$$

$$
\begin{aligned}
N'(1) = &-\Lambda^2 E(I)\left\{1-\theta-[p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))][(1-\theta)(d+\bar{d})\right.\\
&\left.+\theta(1-\mu)]\right\} + (1-\Lambda)\left\{-\Lambda E(I)(1-\theta)(1-\bar{A}(\Lambda)) + \Lambda^2 E(I)E(B)\right.\\
&\left.(p(1-g)(1-\bar{A}(\Lambda)) + p\bar{A}(\Lambda) + g(1-\bar{A}(\Lambda))) + \theta\Lambda E(I)E(F)\right\}\\
N''(1) = &-\Lambda^2\left\{E(I(I-1))\,(1-\theta-A_4(1-\theta\mu)) + 2E(I)\,(b+\theta\Lambda E(I)A_1 - A_2(1-\theta\mu)\right.\\
&\left.+A_4[(1-\theta)\,(1-\Lambda E(I)E(B))] + \tau E(B) - \theta\Lambda(1-\mu)E(I)A_1)\right\}\\
&+(1-\Lambda)\left\{(1-\bar{A}(\Lambda))[-(1-\theta)\Lambda E(I(I-1)) - \Lambda E(I)(b+\theta\Lambda E(I)A_1)]\right.\\
&+\Lambda^2 E(I)E(B)A_2 + \Lambda^2[E(I(I-1))E(B) + E(I)E^2(B)][p(1-g)((1-\bar{A}(\Lambda)))\\
&+p\bar{A}(\Lambda) + g((1-\bar{A}(\Lambda)))] - \theta\Lambda^2 E(B)E(F)E(I)^2\\
&\left.+\Lambda\theta[E(I(I-1))E(F) + E(I)E^2(F)]\right\}
\end{aligned}
$$

where,

$$
\begin{aligned}
A_1 =& E(B) + E(F)\\
A_2 =& p(1-g)E(I)((1-\bar{A}(\Lambda)))\\
A_4 =& p - g(p-1)(1-\bar{A}(\Lambda))
\end{aligned}
$$

- The Little's formula (W_q) is used to determine how long an average customer waits in queue.

$$W_q = \frac{L_q}{\Lambda E(I)}$$

7. Practical application of the model:

The field of telecommunications networks may be able to use the suggested model. This system manages a lot of consumer telephone communications. Call takers are referred to as servers and callers as customers in this context. A consumer may elect to exit the system if he calls and discovers that all the servers are occupied (impatience). Customers wait in orbit while the server is overloaded, out of commission, or undergoing maintenance. If a server has any questions or concerns that fall outside of their area of expertise, they may need to refer them to other servers who are available or speak with a senior in order to acquire the answers. A service failure can be used to represent this circumstance. The speed at which the agent receives responses from the expert in this case is known as the repair rate. Additionally, the server may do various maintenance procedures known as "vacations." Additionally, after each customer's service is finished, dissatisfied customers may re-join the line and be classified as feedback consumers.

8. Numerical Results

In this section, we'll use MATLAB to demonstrate how different parameters affect observations of system behavior. The batch size distance of the arrivals in this section is geometry; with a mean

of 2. Here, the exponential distance is followed by the service, vacation, and repair stages. By creating erroneous assumptions about the parameters, we make sure that the stability criterion is satisfied. Tables 1 to 3 present estimated values for our queueing system's utilization factor (ρ), average queue length (L_q), and average waiting time (W_q).

Table 1: *The effects of arrival rate* (Λ) *on* ρ, L_q, *and* W_q

$g = 0.5$, $p = 1.5$, $E = 0.6$, $G = 2.2$, $\theta = 3$, $d = 3$,
$e = 0.6$, $\mu = 0.9$, $B = 1.5$,$D = 1$, $F = 0.7$, $z = 1$,$b = 2$, $\tau = 1.8$

Arrival rate (Λ)	ρ	L_q	W_q
0.30	0.022680	3.505127	5.841879
0.31	0.096336	4.616984	7.446748
0.32	0.169992	6.010404	9.391256
0.33	0.243648	7.742148	11.730527
0.34	0.317304	9.878047	14.526540
0.35	0.390960	12.494169	17.848813
0.36	0.464616	15.678118	21.775164

Table 2: *The effects of the service rate* $\phi(\varsigma)$ *on* ρ, L_q, W_q

$g = 7.8$, $p = 0.7$, $E = 0.8$, $G = 6$, $\theta = 1$, $d = 3$, $e = 4.6$,
$\mu = 0.7$, $D = 1$, $F = 0.7$, $\Lambda = 0.3$, $z = 1$, $b = 2$, $\tau = 1$

service rate (B)	ρ	L_q	W_q
0.50	0.737200	0.223375	0.372292
0.51	0.687360	0.189247	0.315412
0.52	0.637520	0.160488	0.267480
0.53	0.587680	0.135989	0.226649
0.54	0.537840	0.114943	0.191572
0.55	0.488000	0.096748	0.161247
0.56	0.438160	0.080947	0.134912

Table 3: *The effects of the Breakdown rate* (τ) *on* ρ, L_q, W_q

$g = 0.2$, $p = 0.7$, $E = 2.9$, $G = 9$, $\theta = 1$,
$d = 7$, $e = 8.6$, $\mu = 0.2$, $B = 7$,$D = 2$, $F = 0.7$, $\Lambda = 0.4$,$z = 2$, $b = 4$

breakdown rate (τ)	ρ	L_q	W_q
1.0	0.264592	5.619523	7.024404
1.1	0.303092	6.095933	7.619916
1.2	0.341592	6.575234	8.219042
1.3	0.380092	7.057428	8.821784
1.4	0.418592	7.542516	9.428145
1.5	0.457092	8.030500	10.038125
1.6	0.495592	8.521381	10.651727

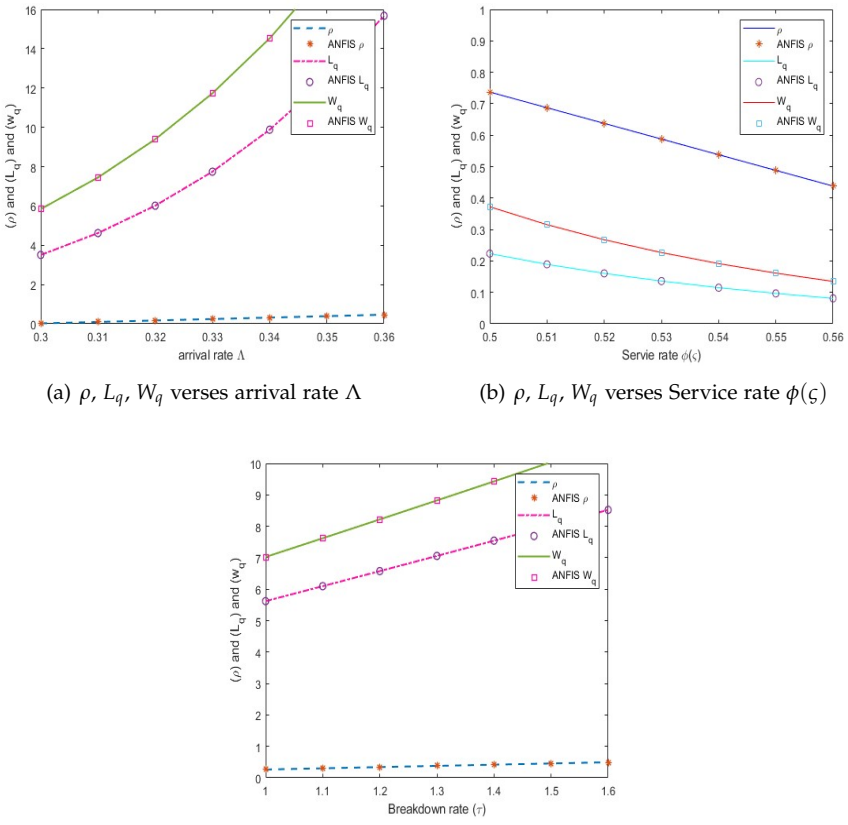
The two-dimensional graph that represents the system measurement of performance is shown in Figure 1 $(a-c)$.

- The figure 1 (a) demonstrates how the utilization factor (ρ), estimated queue length (L_q), and expected waiting time (W_q) all increase as the arrival rate (Λ) does.
- The figure 1 (b) shows that while the utilization factor (ρ) decreases, the service rate $\phi(\varsigma)$ rises. Expected waiting time (W_q) and queue length (L_q) decrease.
- The breakdown rate (τ), utilization factor (ρ), expected queue size (L_q), and expected waiting time (W_q) all show increasing trends in the figure 1 (c).

The three-dimensional graph of the system indicators of performance is shown in Figure 2 $(a-c)$.

- The surface in figure 2 (a) shows the growth of the arrival rate (Λ), estimated length of the line (L_q), and estimated wait time (W_q).
- Figure 2 (b) shows that as the service rate $\phi(\varsigma)$ rises, the estimated queue size (L_q) and waiting time (W_q) both decrease.
- Figure 2 (c) shows that as the breakdown rate τ rises, expected queue lengths (L_q) and waiting times (W_q) also rise.

The numerical results above allow us to determine the influence of attributes on the system's evaluation criteria, and we can be assured that they are representative of realistic conditions.

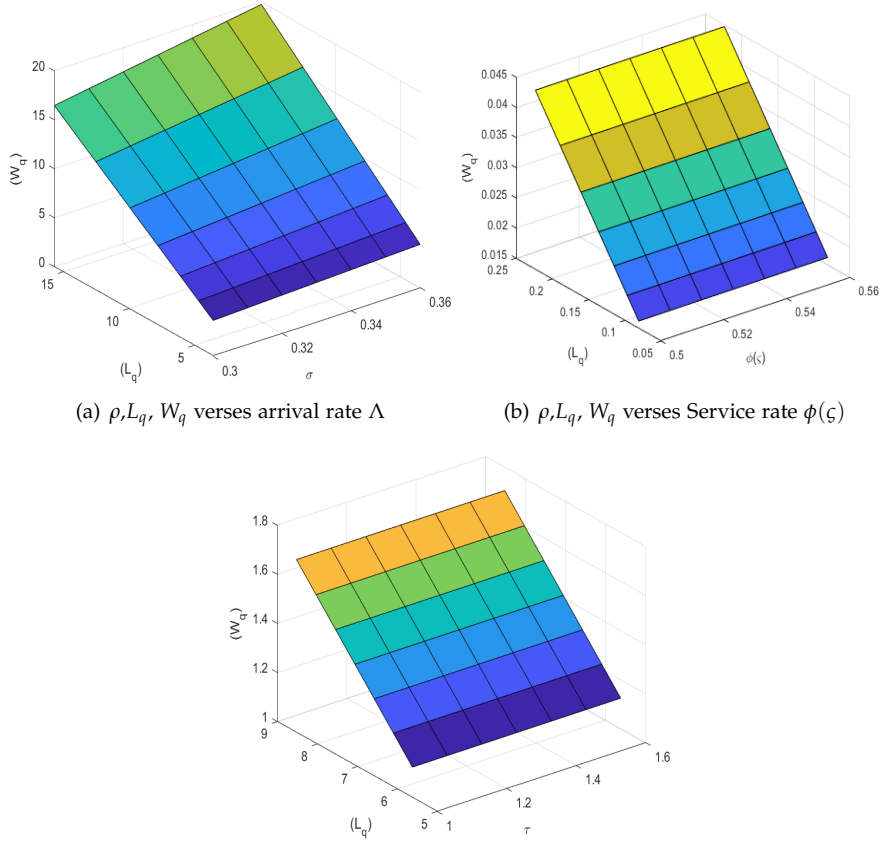


(a) ρ, L_q, W_q verses arrival rate Λ

(b) ρ, L_q, W_q verses Service rate $\phi(\varsigma)$

(c) ρ, L_q, W_q verses Breakdown rate τ

Figure 1: *2D representation effects*



(a) ρ,L_q, W_q verses arrival rate Λ

(b) ρ,L_q, W_q verses Service rate $\phi(\varsigma)$

(c) ρ,L_q, W_q verses Breakdown rate τ

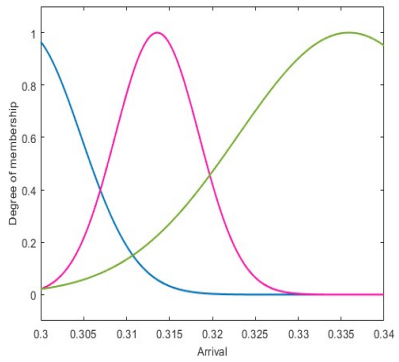
Figure 2: *3D representation effects*

9. Adaptive Neuro-Fuzzy Inference System (ANFIS)

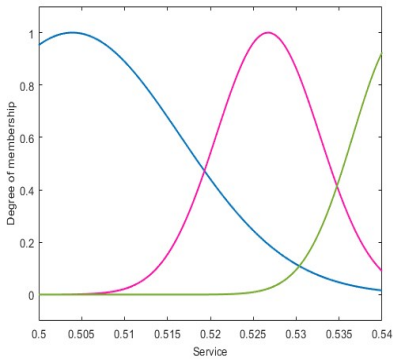
The ANFIS modal is actually applicable in a variety of fields, such as modes of transport, congestion, telecommuting, atmospheric research, etc. Artificial neural networks are used in communications networks to accomplish a variety of goals, including an increase in customers, expense reduction, shorter wait times, etc. With variations in arrival rates while on vacation, service rates, repair rates, and repair to busy rates, the current modal allows us to examine the impatience of the client while they wait for the service.

A very helpful approach for ANFIS is created by combining soft computing methods, artificial neural networks (ANNs), and fuzzy systems (FS). We are showing a simplified idea of the ANFIS architecture by using the fuzzy parameters. We can implement an ANFIS input-output function and input-output data pairs as fuzzy if-then logic. The fuzzy toolbox of MATLAB software can be utilized for contrasting the computational findings with the implementation of an ANFIS network.

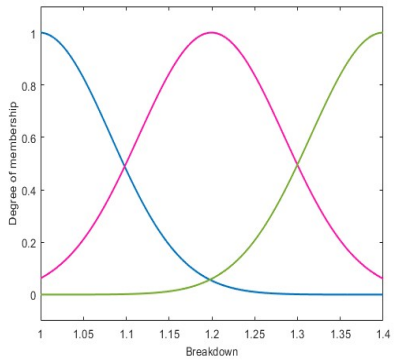
The input parameters and the membership function are assumed to be the Λ, $\phi(\varsigma)$, and τ Gaussian functions in order to produce computational results based on ANFIS. (see Fig. 3a, b, c). It is assumed that the linguistic values are low, moderate, or high. Tick marks are placed over the curves made for the results obtained analytically in Figure 1a, 1b and 1c to indicate the results produced by the ANFIS approach for the queue size. The figures show that the numerical outcomes produced using the Runge-Kutta method and the ANFIS results are nearly identical.



(a) ρ, L_q, W_q verses arrival rate Λ



(b) ρ, L_q, W_q verses Service rate $\phi(\varsigma)$



(c) ρ, L_q, W_q verses Breakdown rate τ

Figure 3: *ANFIS representation effects*

10. **Cost Optimization:**

The term "optimization" describes the method of determining the set of parameters for an objective function that produces the highest or lowest outcome. The continual, business-oriented activity known as "cost optimization" aims to reduce expenditures and costs while raising the organization's value. Standardizing, streamlining, and rationalizing platforms, application development, procedures, and services are all part of this process, along with establishing the most competitive possible terms and prices for all business transactions. The operating cost and profit of a system are closely tied in real-world situations. Therefore, the system's designers or managers place a lot of emphasis on reducing operational expenses per unit of time in order to enhance the system's earnings. Our objective is to identify the best cost per unit of time (TC) characteristics. In order to do this and increase the cost-effectiveness of our developed approach, we will build our competence in this field.

C_h - Holding expense for every user in the system per unit of time.
C_b - The cost for each unit of time the server is turned on and used.
C_v - The cost imposed on the server in vacation mode per unit of time.
C_r - The cost to repair the server after its failure, calculated per unit of time.
C_1 - The cost per unit over a busy time.
C_2 - Cost for each unit of time used over the vacation period.

$$TC = C_h L_q + C_v V + C_b Q + C_r R + C_1 \gamma_b + C_2 \gamma_v$$

The TC problem is solved using metaheuristic optimisation methods including PSO, ABC, and GA. In view of the importance of cost optimisation, this study was conducted using the global search optimisation algorithms particle swarm optimisation (PSO), artificial bee colony (ABC), and genetic algorithms (GA), each of which is separately described in three different subsections of this section. If the algorithm's assumptions are correct, local search techniques frequently offer the level of computer efficiency required to find the global optimal. Tables 5 to 7 display the effects of Λ, τ, and ϕ on TC* using PSO, ABC, and GA.

Table 4: ***Cost sets for optimal policy***

Cost sets	C_h	C_v	C_b	C_r	C_1	C_2
1	10	9	7	6	7	8
2	8	4	6	4	8	9
3	7	6	8	3	9	6

10.1. **Particle Swarm Optimization (PSO)**

One of the meta-heuristic methods used to solve optimization issues is the particle swarm optimization (PSO) technique, which has been employed successfully in a number of single objective optimization problems. Kennedy and Eberhart first proposed this algorithm. The PSO algorithm has the benefit of being simple to implement and apply for solving different function optimization problems, which can be categorized as function minimization or maximization problems.

Table 5: ***Effect of*** $\Lambda, \tau, \phi(\varsigma)$ ***on*** TC^* ***using PSO***

$g = 0.2$, $p = 0.7$, $G = 9$, $\theta = 0.95$, $d = 7$, $e = 8.6$,
$c = 0.2$, $B = 7$, $D = 2$, $\tau = 1.6$, $b = 4$

Cost sets	TC^*			
		Cost set 1	Cost set 2	Cost set 3
Λ	0.4	149.1752	133.0711	127.2781
	0.5	162.6882	143.5244	136.3697
	0.6	173.5857	152.1798	143.7058
τ	1.6	149.1752	133.0711	127.2781
	1.7	161.5959	141.8755	134.7960
	1.8	175.6139	151.7985	143.2466
$\phi(\varsigma)$	7	149.1752	133.0711	127.2781
	8	184.5033	159.0594	151.0219
	9	230.2302	192.6975	181.7546

10.2. **Artificial Bee Colony(ABC)**

One of Dervis Karaboga's most recent algorithms—created in 2005—is called the Artificial Bee Colony and was modeled after the cunning behaviour of honey bees. Basic process indicators like colonies and highest levels are essentially all that are used. Like PSO and differential evolutionary approaches, it is equally simple to comprehend. The search for huge areas of nectar-containing

food sources, and ultimately the one with the most nectar, is the bees' main goal. This population-based search approach is the main one used by ABC. The cost of the suggested structure is decreased through a process known as ABC.

Table 6: ***Effect of*** $\Lambda, \tau, \phi(\varsigma)$ ***on*** TC^* ***using ABC***

$g = 0.2$, $p = 0.7$, $G = 9$, $\theta = 0.95$, $d = 7$, $e = 8.6$,
$c = 0.2$, $\Lambda = 0.4$, $D = 2$, $\tau = 1.6$, $b = 4$

Cost sets		TC^*		
		Cost set 1	Cost set 2	Cost set 3
Λ	0.4	108.0030	108.8585	109.6965
	0.5	112.6458	113.7397	114.0666
	0.6	115.8805	117.2631	116.9601
τ	1.6	108.0030	108.8585	109.6965
	1.7	112.4344	113.2737	114.1395
	1.8	117.4656	117.8425	118.7674
$\phi(\varsigma)$	7	108.0030	108.8585	109.6965
	8	120.7245	120.9982	122.5394
	9	138.2173	134.5400	136.4184

10.3. **Genetic Algorithm (GA)**

The genetic algorithm, created in the 1960s and 1970s by Bremermann, Holland, and their colleagues, is a technique for addressing optimization problems brought on by natural selection, the mechanism that promotes evolution in biology. They are frequently employed to deliver superior solutions to stochastic search issues. The full procedure serves as a representation of the criteria for choice that were used to select the people who would make the best parents for the coming human generation.

Table 7: ***Effect of*** $\Lambda, \tau, \phi(\varsigma)$ ***on*** TC^* ***using GA***

$g = 0.2$, $p = 0.7$, $G = 9$, $\theta = 0.95$, $d = 7$, $e = 8.6$,
$c = 0.2$, $\Lambda = 0.4$, $D = 2$, $B = 7$, $b = 4$

Cost sets		TC^*		
		Cost set 1	Cost set 2	Cost set 3
Λ	0.4	152.5341	131.8364	124.6592
	0.5	163.2414	141.3111	131.8853
	0.6	170.0165	148.0781	136.4720
τ	1.6	152.5341	131.8364	124.6592
	1.7	167.5959	142.6723	133.7200
	1.8	184.4744	154.7915	143.8279
$\phi(\varsigma)$	7	152.5341	131.8364	124.6592
	8	192.8582	162.3320	151.7762
	9	243.6750	200.7627	185.9493

10.4. **Analogy of PSO, ABC and GA**

This section compares the three approaches—particle swarm optimization (PSO), artificial bee colony (ABC), and genetic algorithm (GA)—to determine which has the least expense using the corresponding MATLAB programs. Then, one by one, the MATLAB programs for each of the aforementioned algorithms are run. We found that all three programs generated values that were nearly identical. Because of this, the three solutions are nearly comparable in terms of their optimum results and the fewest associated costs. It proves the reliability (local) and potency of these three simple techniques. Any technique can be used to calculate the optimal cost; however, PSO outperforms all others in comparison to our model. Because PSO has so many advantages, we have found that it is the best approach out of all of them. It performs well in global queries, requires a small number of arguments, is easy to configure, and is unaffected by design variable scalability. In addition to suffering sluggish convergence in a concentrated searching region, PSO has a tendency to lead to swift and early convergence in mid-optimal locations (being able to impair local search capabilities).

10.5. **Convergence in PSO, ABC and GA**

After employing an optimization methodology like PSO, ABC, or GA, it is crucial to comprehend whether a particle recovers to normal or not and when it will roam around in search of a better solution. As a result, convergence is a significant component of cost evaluation. A statistical analysis (Fig. 4) of the outcomes demonstrates that ABC exceeds the PSO approach. For the whole standard optimization, ABC had fewer functional evaluations overall than PSO. The findings demonstrate that PSO converges more quickly. ABC cannot be employed if a speedy result is required for time-sensitive applications.

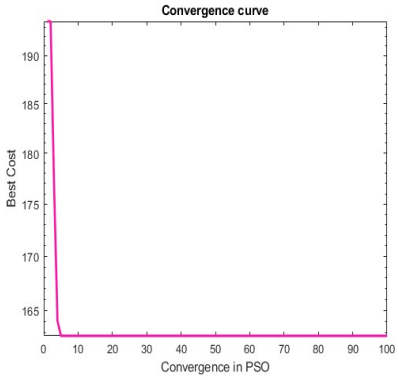
The study shows the applicability of our concept to real-world situations. Some of the analysts' financial issues will be partially overcome once they know how much the system will cost overall. The current situation may heavily rely on the cost-benefit assessment that was produced, which serves to illustrate the logic of our strategy and aid network administrators and specialists in lowering the issue of communications services that explicitly deal with blocking.

11. **Conclusion:**

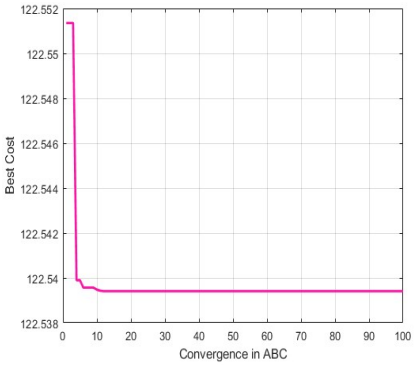
This paper investigates the $M^X/G(a,b)/1$ retrial queue with random failure and feedback under extended Bernoulli vacation with impatient customers. The SVT is utilized to determine indicators of efficiency for the various system stages. The efficiency of the system is then evaluated after considering the effects of various parameters. Finally, we gave a thorough explanation of the ANFIS. PSO, ABC, and GA are also used to compute the total cost. In an effort to find the best offer, these techniques compare and contrast the outcomes. The impetus for this study came from the prospective applications for the developed model, such as call centres, wireless networks, or telecommunication infrastructures, which might be powered by controlled precision test queueing systems to provide outstanding service at low prices. The simple mail transfer protocol utilizes a way to convey the messages between the mail servers. The recommended approach might be used in an email system's transfer model.

Declarations:

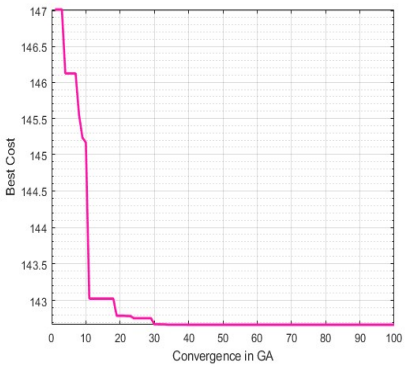
Acknowledgments: Not applicable
Funding information: This research did not receive any specific grant from funding agencies in



(a) TC verses Arrival rate (Λ) in PSO



(b) TC verses Service $\phi(\varsigma)$ in ABC



(c) TC verses Breakdown rate (τ) in GA

Figure 4: *Cost Optimization effects*

the public, commercial, or not-for-profit sectors.
Conflicts of interest: The authors declare no conflict of interest.
Data availability: Not applicable
Authors contribution: All the authors made substantial contributions to the conception or design of the work.
Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Chaudhry, M. and Templeton, J. (1983). A first course in bulk queues. *John Wiley & Sons*. http://dx.doi.org/10.1016/j.asej.2016.08.025.
[2] Bailey, N.T. (1954). On queueing processes with bulk service. *J R Stat Soc Ser A Stat Soc., Series B (Methodological)*. 80-87. https: //doi.org/10.1111/j.2517-6161.1954.tb00149.x
[3] Sumitha, D. and Udaya Chandrika, K. (2012). Retrial queuing system with starting failure, single vacation and orbital search.*Int.J. Comput. Appl.* 40.13, 29-33.
[4] Radha,J., Rajadurai, P., Indhira, K. and Chandrasekaran, V.M. (2014). A batch arrival retrial queue with K-optional stages of service, Bernoulli feedback, single vacation and random breakdown. *Glob. J. Pure Appl. Math.* 10.2, 265-283.
[5] Krishna Kumar,B., Pavai Madheswari,S. and Vijayakumar,A. (2002). The M/G/1 retrial queue with feedback and starting failures. *Appl.Math. Model*. 26.11, 1057-1075.https://doi.org/10.1016/S0307-904X(02)00061-6

[6] Yang, Tao, Hui Li. (1994). The M/G/1 retrial queue with the server subject to starting failures. *Queueing Syst.* 16.1, 83-96. https://doi.org/10.1007/BF01158950

[7] Mokaddis, G.S., Metwally,S.A. and Zaki,B.M. (2007). A feedback retrial queuing system with starting failures and single vacation. *J. Appl. Sci. Eng.* 10.3, 183-192. https://doi.org/10.6180/jase.2007.10.3.01

[8] Bouchentouf, Amina Angelika, Abdelhak Guendouzi. (2021). Single server batch Arrival Bernoulli feedback queueing system with waiting server, K-variant vacations and impatient customers. *Oper. Res.Forum.* Vol.2. No.1. Springer International Publishing. https://doi.org/10.1007/s43069-021-00057-0.

[9] Ayyappan, G. and Sathiya, K. (2013). Transient analysis of batch arrival feedback retrial queue with starting failure and Bernoulli vacation. *Math. Model. Anal.* 3.8, 60-67.

[10] Ayyappan, Govindhan, Udayageetha, J. (2020). Transient Analysis of M [X1], M [X2]/G1, G2/1 retrial queueing system with priority services, working breakdown, start up/close down time, Bernoulli vacation, reneging and balking. *Pak. J. Stat. Oper.* 203-216. https://doi.org/10.18187/pjsor.v16i1.2181.

[11] Kulkarni, Vidyadhar, G., Bong Dae Choi. (1990). Retrial queues with server subject to breakdowns and repairs. *Queueing syst.* 7.2, 191-208. https://doi.org/10.1007/BF01158474

[12] Ayyappan, G. and Shyamala,S. (2013). M[X]/G/1 with Bernoulli Schedule Server Vacation Random Break Down and second optional Repair. *J. Comp., & Model.* 3, 159-175.

[13] Ke, Jau-Chuan, Tzu-Hsin Liu, Siping Su, Zhe-George Zhang. (2022). On retrial queue with customer balking and feedback subject to server breakdowns. *Commun. Stat.* 51.17, 6049-6063. https://doi.org/10.1080/03610926.2020.1852432

[14] Takacs, L. (1963). A single-server queue with feedback. *Bell Syst.Tech.J.* 2, 505-519.

[15] Rajadurai, P., Saravanarajan, M., Chandrasekaran, V.M. and Indhira, K. (2015). An M/(G1, G2)/1 Feedback Retrial Queue with Two Phase Service, Variant Vacation Policy Under Delaying Repair for Impatient Customer. *Int. J. Fuzzy Math.* Arch 6, 45-55.

[16] D'Arienzo, M.P., Dudin, A.N., Dudin,S.A. and Manzo, R. (2020). Analysis of a retrial queue with group service of impatient customers. *J Ambient Intell Humaniz Comput.* 11.6, 2591-2599. https://doi.org/10.1007/s12652-019-01318-x

[17] Nila, M. and Sumitha, D. Batch Arrival Retrial Queueing Model with Starting Failures, Customer Impatience, Multi Optional Second Phase and Orbital Search.

[18] Ayyappan, G., Thamizhselvi, P., Somasundaram, B. and Udayageetha, J. (2021). Analysis of an retrial queueing system with priority services, working breakdown, Bernoulli vacation, admission control and balking. *Int. j. stat. manag. syst.* 24.4, 685-702. https://doi.org/10.1080/09720529.2020.1744812

[19] Ayyappan, G. and Nirmala, M. (2021). Analysis of customer's impatience on bulk service queueing system with unreliable server, setup time and two types of multiple vacations. *Int. J. Ind. Syst.* 38.2, 198-222. https://doi.org/10.1504/IJISE.2021.115321

[20] Sethi, R., Jain, M., Meena, R. K. and Garg, D. (2020). Cost optimization and ANFIS computing of an unreliable M/M/1 queueing system with customers' impatience under n-policy. *Int. J. Appl. Comput. Math.* 6, 1-14. https://doi.org/10.1007/s40819-020-0802-0

[21] Miss Vaishnawi, Shweta Upadhyaya and Rakhee Kulshrestha (2022). Optimal Cost Analysis for Discrete-Time Recurrent Queue with Bernoulli Feedback and Emergency Vacation. *Int. J. Appl. Comput. Math.* 8.5, 254. https://doi.org/10.1007/s40819-022-01445-8

APPLICATION OF AUTOMATIC MONITORING AND CONTROL SYSTEMS FOR RELIABILITY OF POWER TRANSMISSION LINES

S.V. Rzayeva[1], I.A. Guseynova[2]

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[1] sona.rzayeva@asoiu.edu.az ; [2] huseynova.ilduze@asoiu.edu.az

Abstract

The main aspects of the use of automatic monitoring and control systems in order to increase the reliability of power lines are considered. The article highlights modern technologies and techniques that make it possible to quickly identify and prevent possible emergency situations on power transmission lines. The advantages of automated systems compared to traditional monitoring and control methods are discussed, and examples of research and practical applications of such systems are presented. The results obtained can be useful for energy specialists and engineers involved in the design, operation and maintenance of power transmission lines, as well as for developers and manufacturers of automated monitoring and control systems.

Keywords: power lines, sensors, monitoring, automatic control

I. Introduction

The reliability of power transmission lines (PTLs) plays a key role in ensuring the stable operation of energy systems. Power transmission lines are the main channel for transmitting electricity from generating sources to end consumers, and their reliability directly affects the continuity of power supply. Failures or malfunctions in power transmission lines can lead to power outages, which have serious consequences for both industry and residential areas, including loss of production, financial losses and even threats to human life and health.

Moreover, modern energy systems are becoming increasingly complex and integrated, with an increasing share of distributed generation, including renewable energy sources such as solar and wind power. In this context, the reliability of transmission lines becomes even more critical, since even small power outages can have a cascading effect on the operation of the entire energy system. Therefore, ensuring the reliability of power lines is an essential condition for ensuring stable and safe operation of energy systems as a whole. Online monitoring and automatic control, which are provided by ASMU systems, are becoming necessary tools for the timely identification and elimination of possible problems on power lines, which helps to increase their reliability and ensures the efficient functioning of the entire energy system [1-3].

In recent years, there has been a significant increase in interest in automated monitoring and control systems (ASMU) in various fields, including energy. This increase in interest is due to several factors. Firstly, with the development of technology and the introduction of the Internet of Things (IoT) concept, it has become possible to create more efficient and intelligent monitoring systems that are capable of continuously collecting and analyzing equipment condition data in real

time. This allows operators to quickly respond to any anomalies and prevent possible accidents.

Secondly, the desire to improve the efficiency and reliability of technical systems also stimulates interest in ASMU. Automated monitoring and control systems make it possible to optimize maintenance and repair processes, reduce personnel and resource costs, and also increase the service life of equipment due to timely detection and elimination of faults.

Finally, the growing need for security and reliability of critical infrastructures such as power systems is also driving demand for ASMUs. Automated monitoring and control systems provide a higher level of control and protection, which helps prevent incidents and minimize their consequences.

Thus, automated monitoring and control systems are becoming an integral part of modern technical infrastructure, ensuring continuous operation and increased reliability of various technical systems, including energy ones.

II. Formulation of the problem

ASMUs (automatic monitoring and control systems) are an important component of modern energy systems, providing continuous monitoring and control of power transmission lines (PTLs). The technical characteristics and capabilities of ASMU include the use of various types of sensors and sensors, such as temperature, humidity, pressure, vibration and others, to continuously collect data on the condition of power transmission line equipment. This data is transmitted to a central control center via data networks such as the Internet, cellular or satellite communications, ensuring a rapid response to any changes or faults.

In addition, ASMUs are equipped with high-performance controllers and management software capable of analyzing sensor data and making appropriate decisions in real time. This allows monitoring and control systems to automatically respond to changes in the technical condition of power lines, preventing possible emergency situations and ensuring the uninterrupted operation of the energy system. In addition, ASMUs integrate with other control and automation systems to optimize the operation of the entire energy system and increase its efficiency.

Finally, one of the important characteristics of ASMUs is their data analytics and diagnostic capabilities. Using machine learning and data analytics techniques, ASMUs are able to identify hidden patterns, predict possible failures and determine the causes of failures based on the analysis of collected data. This allows operators of monitoring and control systems to make informed decisions on the maintenance and repair of power lines, minimizing the risk of emergency situations and ensuring stable operation of the energy system.

The use of automated monitoring and control systems (ASMU) is a key solution for increasing the reliability of power transmission lines (PTLs) and ensuring the safety of energy systems in general. The main benefits of ASMU for power line reliability include:

Firstly, increasing the efficiency and effectiveness of diagnostics and detection of possible faults or anomalies in the operation of power lines. ASMUs are capable of continuously monitoring the condition of equipment and automatically identifying any deviations from normal operation, which allows you to quickly respond to potential threats and prevent possible emergency situations.

Secondly, optimization of planned and emergency maintenance of power lines. ASMUs make it possible to carry out preventive maintenance and repair of equipment at more optimal time intervals based on real data about its condition, which reduces the likelihood of failures and increases the service life of power lines [4].

In addition, ASMUs provide the ability to automatically control and optimize the operation of power lines in real time depending on changing operating conditions, such as load, weather

conditions, etc. This allows for more efficient use of resources and ensures stable power supply in all conditions.

Finally, ASMUs help improve safety and protect transmission lines from external threats such as power outages, cyber-attacks or natural disasters. Automatic monitoring and control systems ensure a quick response to any threats and prevent the possible consequences of such incidents from spreading to other parts of the energy system.

III. Problem solution

When transporting electricity through a specific power transmission line, the permissible current loads are regulated. In this case, current limit values are used that determine the sag of the wires above the critical value. These data are taken for the most extreme conditions, which do not occur in more than 90% of the operating time of power lines [5]. Consequently, there is a resource for transmitting large capacities without violating the regulations. That is, it is possible to transmit additional power (15–30%) almost 90% of the operating time. The presence of a monitoring system allows you to use this additional resource without reducing the reliability regulations. To do this, it is necessary to monitor the current level and temperature of the wires along the entire route and, in accordance with the real state of the line, dynamically adjust the level of transmitted power (Fig. 1).

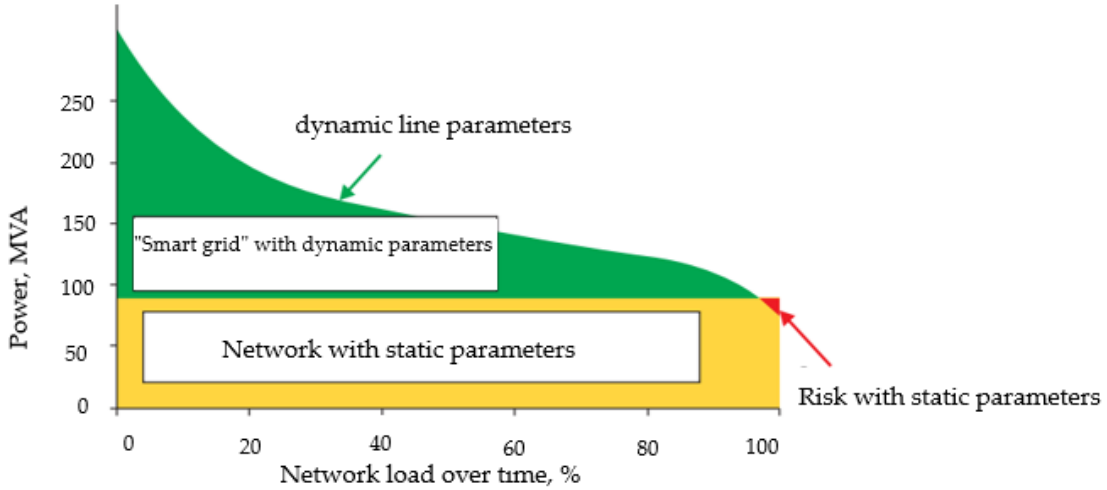


Figure 1: *Efficiency of energy transfer in power transmission lines with static and dynamic parameters*

Telemetric monitoring of power line wire parameters was first proposed more than 40 years ago. The first controlled parameter via a telemetric radio channel was the current in the wire. The appearance of the American patent Remote measuring system [6] dates back to this time ("Systems for remotely measuring current in a wire with transmission of the measured value via a radio channel"). The proposed solution used power to the measuring device from an induction transformer due to the current flowing in the wire. It was measured through a transformer current sensor. The signal modulated the grid circuit of the tube transmitter (Fig. 2). As can be seen in the figure, the current meter used measuring and current transformers to power the lamp circuit (anode and filament circuit). The transmitter is made on a single-tube stage. The AM RF signal is used by modulating the grid current of the transmitter oscillator. In the last 15 years, thanks to the development of information technology, the commercial implementation of power line wire monitoring systems has become possible.

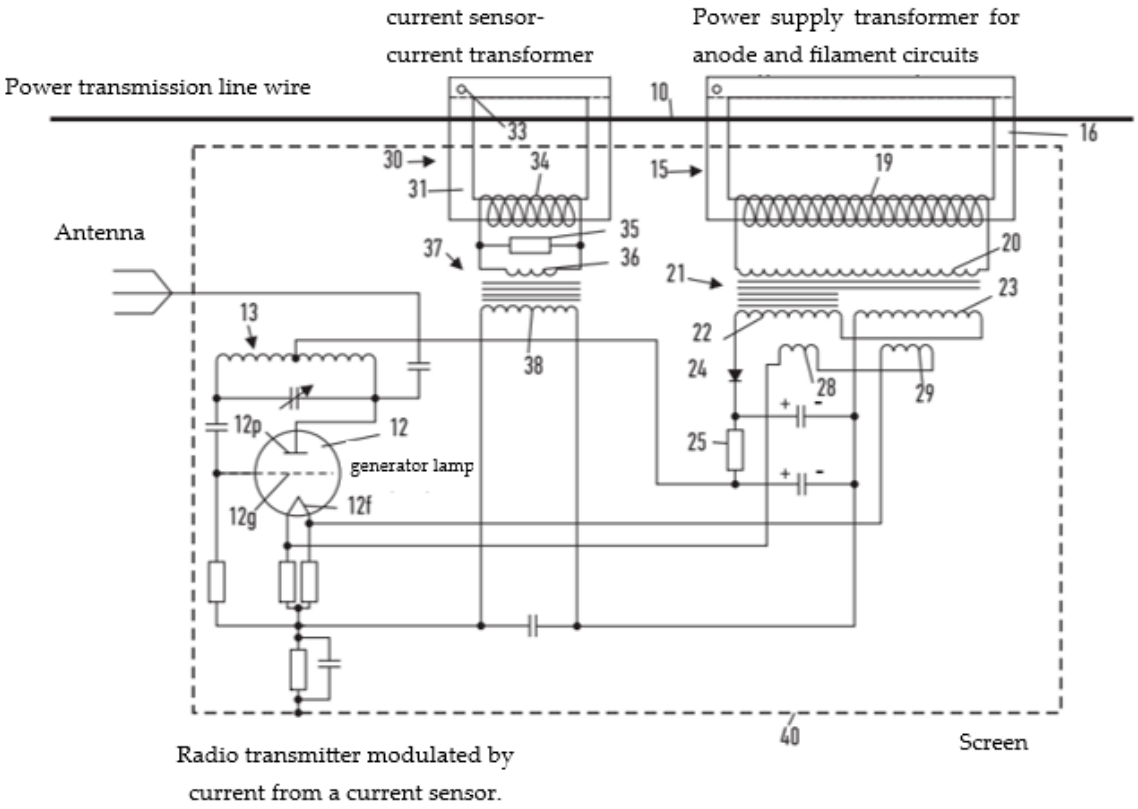


Figure 2: *Circuit diagram of a remote current meter with a radio channel*

Currently, various monitoring systems for overhead power lines are widely used throughout the world, providing the system operator with detailed information about the current state of overhead cable power supply networks. The monitoring system consists of a network of measuring units connected through a communication channel with equipment at the control center. Measuring units are distributed along the power line route and mounted on supports or directly on high-voltage wires. Figure 3 shows the structure of the transmission line capacity monitoring system.

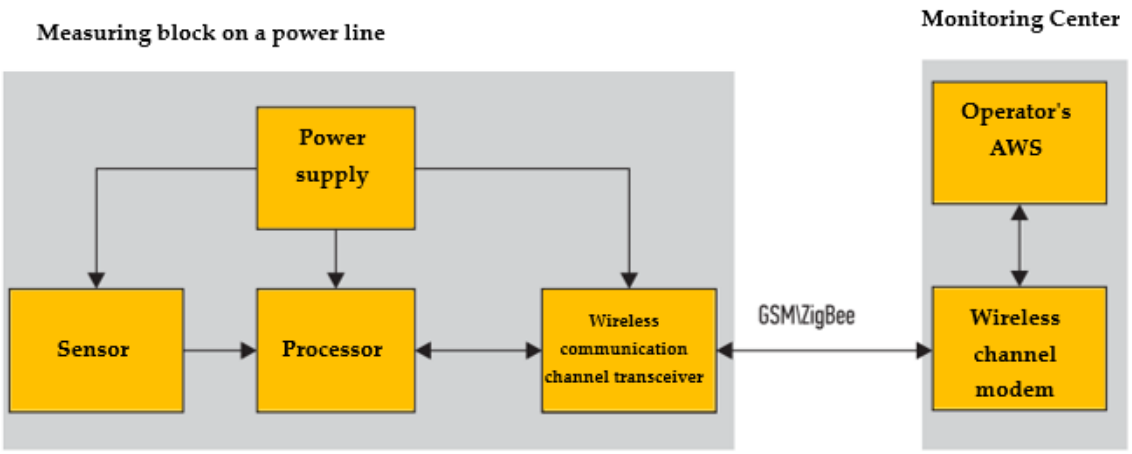


Figure 3: *Power line wire monitoring system*

Control rooms are located at the nodes of energy redistribution networks. Currently, they usually use SCADA systems that provide processing and interpretation of data received from measuring units (Fig. 4).

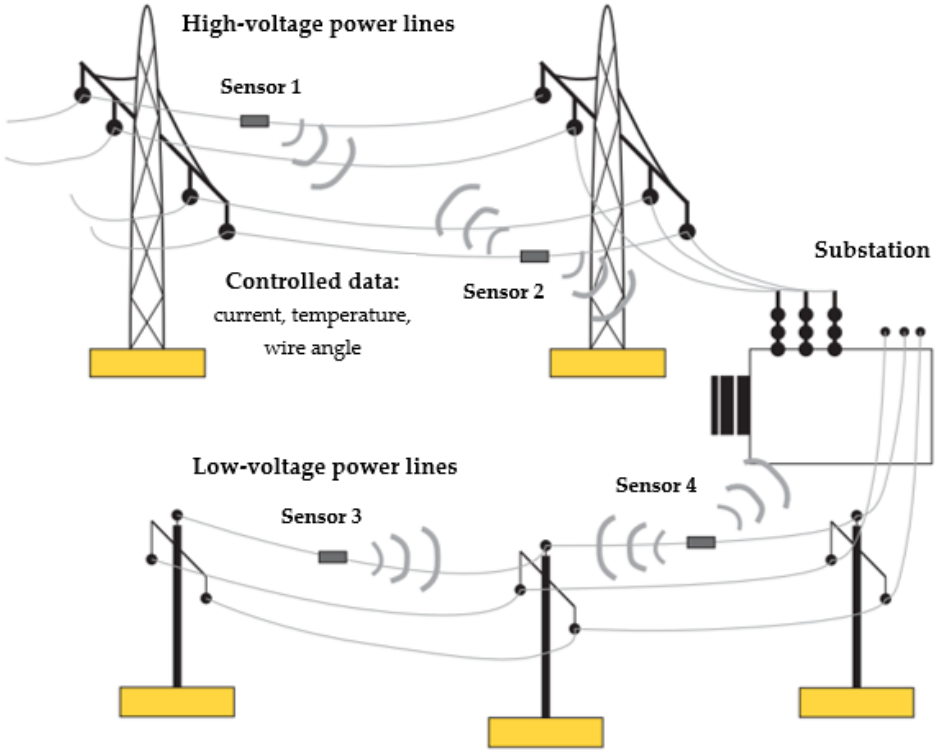


Figure 4: *Structure of the measuring unit and monitoring center*

The measuring unit includes the following basic components:

- a group of sensors for measuring the main current parameters of the wire line;
- processor module for processing measured data;
- data transmission system;
- autonomous power module.

Depending on the functional purpose, various types of sensors can be used in monitoring systems:

- for measuring current in a wire;
- temperature of the wire in the span;
- mechanical stress of the wire at suspension points (strain gauges);
- for measuring attenuation in optical fibers of ground wires or phase wires;
- to measure critical sag;
- climatic conditions (weather station);
- vibration characteristics of wires (accelerometers). Current measurement is carried out using a non-contact method, for which sensors based on the Hall effect or Rogowski coil are used.

Currently, wireless communication channels are mainly used to transmit data in overhead line monitoring systems - these are GSM or ISM radio modems operating at frequencies of 434, 868 MHz and 2.4 GHz. GSM modems have been used in the market of automated process control systems for more than ten years, including for data transmission in monitoring systems. The first models had limited capabilities for transmitting SMS messages and data in analog mode. The operation of such devices in analog modem mode provides a data transfer speed of only 9.5 kbaud, and payment is made in accordance with the time spent on the network. The GPRS system implements packet switching throughout the entire communication channel, significantly optimizing data transmission services in GSM networks. It establishes connections almost instantly, utilizes network resources, and occupies bandwidth only when data is actually being transmitted, ensuring extremely efficient use of available bandwidth. GPRS provides a multipoint transmission service (multicast) between a specific network provider and a group of mobile subscribers with GPRS terminals. GPRS requires traffic payment, which is charged only for the volume of transmitted and received information, and not for the time the modem is in the receiving/transmitting state. To transfer data from the measuring modules to the monitoring system server, a wireless network created on the basis of xBee radio modems from Digi can be used. Currently, transceivers are produced at frequencies of 868 MHz and 2.4 GHz. Transceivers provide a line-of-sight data transmission range of up to 4 km. Based on a network of ZigBee

transceivers with a backbone topology, it is possible to organize relay data transmission over the network between meters to the monitoring system data server. The direction of transmission in the transmission network along power lines is always set towards the server. To increase reliability, it is possible to alternatively bypass the problematic node that blocks communication along the chain.

IV. Conclusions

In conclusion, the application of automatic monitoring and control systems (ASMU) is an important step to ensure the reliability of power transmission lines (PTLs) and improve the efficiency of energy systems in general. ASMUs provide the ability to continuously monitor the condition of equipment, automatically identify and eliminate faults, optimize the operation of power lines and analyze data to predict possible failures. These systems play a key role in ensuring a stable and safe power supply, as well as reducing the cost of maintaining and repairing power lines.

However, further development of ASMU requires efforts to improve monitoring and data analytics technologies, as well as the creation of standards and regulations to ensure compatibility and interoperability of various monitoring and management systems. In the future, the development of ASMU will be aimed at improving automation and artificial intelligence, which will improve the efficiency and reliability of energy systems, ensuring stable and uninterrupted power supply in any conditions.

References

[1] S.A. Bazhanov, “Infrared Diagnostics of Electrical Equipment of Switchgears”, Mining Science and Technology, p. 76, Moscow, Russia, 2000.

[2] M. Sasic, C. Chan, “Using Measurement Results to Diagnose the Condition of High-Voltage Rotating Machines Part 1 and 2”, Industry Topics, Iris Power, pp. 1-9, Canada, 2019

[3] Ahmedov, E., Rzayeva, S., Ganiyeva, N., & Safiyev, E. Improving the lightning resistance of high-voltage overhead power line. Przeglad Elektrotechniczny, 2023, Vol 2023, Issue 11, p121

[4] SV, Rzayeva, Mammadov NS, and Ganiyeva NA. "Neutral grounding mode in the 6-35 kv network through an arcing reactor and organization of relay protection against single-phase ground faults." Deutsche Internationale Zeitschrift für Zeitgenössische Wissenschaft 2022. Issue 42, p31

[5] Mohamed ES. Development and analysis of a variable position thermostat for smart cooling system of a light duty diesel vehicles and engine emissions assessment during NEDC. Applied Thermal Engineering 2016; 99: pp.358–372. https://doi.org/10.1016/j.applthermaleng.2015.12.099

[6] Rzayeva SV, Ganiyeva NA, Piriyeva NM. Modern methods of diag-nostics of electric power equipment. The 19th International Conference on" Technical and Physical Problems of Engineering 2023; T. 31. pp.105-110.

THE TRIANGLE-G FAMILY OF DISTRIBUTIONS: PROPERTIES, SUB-MODELS, ESTIMATION AND APPLICATION IN LIFETIME STUDIES

Habibah Rahman
•
University of Science and Technology Meghalaya, Meghalaya, India
Email: umme.habibah.rahman17@gmail.com

Abstract

This paper pictures the importance and the generalization of a new family of distribution developed on Triangle distribution. The new family provides some useful expansions, properties and a suitable alternative to some of existing models with same and higher number of parameters. Exponential distribution (one parameter) and Inverse Weibull distribution (Two parameter) play the role of sub-models. This new family distribution is used as a statistical model to estimate the parameters using the maximum likelihood estimation method. A complete study of Percentage points has been tabled. Two real-world data sets are investigated, demonstrating the suggested model's capacity to fit a variety of data sets along with some other models.

Keywords: G-family of distribution, Maximum Likelihood estimation, Percentage Points, Lifetime data.

I. Introduction

Let us suppose a random variable $T \in (a, b)$ for $-\infty \leq a < b < \infty$ having a probability density function (pdf) $y(t)$ and $W[F(x)]$ be a function of a cumulative distribution function (cdf) of the random variable X which satisfies some statistical conditions such as $W[F(x)] \in (a, b)$, $W[F(x)]$ is differentiable and monotonically non-decreasing and $W[F(x)] \to a$ as $x \to -\infty$ and $W[F(x)] \to b$ as $x \to \infty$. Aljarrah et al. (2014) defined the T-X family cdf by $G(x) = \int_a^{W[F(x)]} y(t)dt = Y\{W[F(x)]\}$, where $W[F(x)]$ satisfied all the conditions. The corresponding pdf of T-X family of distribution is $g(x) = \left\{\frac{d}{dx} W[F(x)]\right\} y\{W[F(x)]\}$.

This study suggests a new distribution family that is inspired by the Triangle-G family. Below is a quick explanation of the Trianlge-G family. The pdf and cdf of Triangle distribution is as follows

$$g(x; p) = \frac{2x}{p};\ x \in \mathbb{R} \tag{1.1}$$

$$G(x; p) = \frac{x\hat{}2}{p};\ x \in \mathbb{R} \tag{1.2}$$

The simple form (putting $p = 1$) of the pdf and cdf of Triangle distribution is defined as

$$g(x) = 2x; x \in \mathbb{R} \tag{1.3}$$

$$G(x) = x^2; x \in \mathbb{R} \tag{1.4}$$

This distribution has a number of advantages, such as its simplicity and capacity for enhancing the flexibility of PDF and CDF while introducing new flexible models. Researchers may have more options when it comes to these distributions along with trigonometric functions.

Here a table of chronological review has been added for the recent G-families based on the trigonometric functions and their inverses techniques.

Table 1*: Literature reviews of some recent trigonometric functions and G-families*

Sl. No.	Authors	Years	Contributions in distribution family
1	Souza et al.	2022	Sec-G class of probability distribution
2	Sakthivel et al.	2022	transmuted Sin-G class of probability distribution
3	Mahmood et al.	2022	extended cosine-G class of probability distribution
4	Rahman M.	2021	Arcsine-G class of probability distribution
5	Eghwerido et al.	2021	Teissier-G class of probability distribution
6	Chesneau et al.	2021	distribution based on the arccosine function
7	Muhammad et al.	2021	exponentiated sine-G class of probability distribution
8	Ahmad et al.	2021	exponential T-X class of probability distribution
9	Liang Tung et al.	2021	arcsine-X class of probability distribution
10	Alkhiary et al.	2021	ArcTan Lomax distribution
11	Souza et al.	2021	Tan-G class of probability distribution
12	Muhammad et al.	2021	A New Extended Cosine—G distributions
13	He et al.	2020	arcsine exponentiated- X class of probability distribution
14	Al-Babtain et al.	2020	Sine Topp-Leone-G class of probability distribution
15	Chesneau and Jamal	2019	Sine Kumaraswamy-G class of probability distribution
16	Mahmood et al.	2019	A New Sine-G Family of Distributions: Properties and Applications
17	Chesneau et al.	2019	new class of probability distributions via cosine and sine functions
18	Mahmood et al.	2019	sine-G class of probability distribution

The Triangle-G family of distribution was introduced in this study. The Tr-G family's key benefit is that practitioners will have a one-parameter class that is adaptable to actual data in relevant disciplines. It may be a good substitute for other distributions with one, two, three, or four parameters. In some real-world circumstances, nevertheless, it might also exceed other kinds of distributions in terms of model fit, although this is not always assured. Additionally, a full account of some of its mathematical properties is provided.

The outline of rest of the paper is as follows. The derivation of the form for the Tr-G density function described in Section 2. Some of the general mathematical aspects of the proposed family that are

included in Section 3. In Section 4, one unique model of this family is presented, along with various plots of their pdfs and hrfs. The proposed model's percentage point results are discussed in Section 5.

In Section 6, we use two particular models of the proposed family on real data sets to demonstrate their applicability. In Section 7, some concluding remarks are presented.

II. Triangle- G (TR-G) family of distribution

The derivation of pdf and cdf of Triangle-G family of distribution is discussed in this section. Let us consider a random variable X that belongs to the Triangle-G family, the cdf and pdf can be written in the following form

$$G_{Triangle-G}(x;p;\phi) = \frac{F(x;\phi)^2}{p}; x \in \mathbb{R} \quad (2.1)$$

$$g_{Triangle-G}(x;p;\phi) = \frac{2F(x;\phi)f(x;\phi)}{p}; x \in \mathbb{R} \quad (2.2)$$

The simplest form of TR-G family of distribution is formed by putting $p=1$. The cdf and pdf are as follows

$$G_{Triangle-G}(x;\phi) = F(x;\phi)^2; x \in \mathbb{R} \quad (2.3)$$

$$g_{Triangle-G}(x;\phi) = 2F(x;\phi)f(x;\phi); x \in \mathbb{R} \quad (2.4)$$

Here $f\left(x;\Phi\right)$ and $F\left(x;\Phi\right)$ are considered as the pdf and cdf of baseline (or parent) random variable depending on the parameter vector. The complementary cdf (or survival function (srf)), instantaneous failure rate (or hazard rate function (hrf), retro hazard (or reversed hazard rate function), integrated hazard rate (or cumulative hazard rate function) can be written as below

$$S_{Triangle-G}(x;p;\phi) = 1 - \left[\frac{F(x;\phi)^2}{p}\right]; x \in \mathbb{R} \quad (2.5)$$

$$h_{Triangle-G}(x;p;\phi) = \frac{2F(x;\phi)f(x;\phi)}{p-[F(x;\phi)^2]}; x \in \mathbb{R} \quad (2.6)$$

$$r_{Triangle-G}(x;p;\phi) = \frac{2f(x;\phi)}{F(x;\phi)}; x \in \mathbb{R} \quad (2.7)$$

$$H_{Triangle-G}(x;p;\phi) = -\log\left[1 - \left\{\frac{F(x;\phi)^2}{p}\right\}\right]; x \in \mathbb{R} \quad (2.8)$$

III. Some Properties

I. Quantile function, Median, Bowley skewness and Moors kurtosis

The quantile function (also known as the inverse cdf) of the Triangle-G family follows by inverting the Triangle-G distribution function. Let us consider $u \sim U(0,1)$, the u^{th} quantile function of TR-G is defined as $Q_F(u)$ is the solution of $Q(u) > 0$. It may be written as follows in terms of the tangent trigonometric function as

$$x = Q_F(u) = G^{-1}(u) = F^{-1}\left[(pu)^{\frac{1}{2}}\right] \quad (3.1)$$

where $u \in (0,1)$. The quantile function expression may be used to generate random numbers from

TR-G distributions. The median of the TR-G family can be obtained by setting $u = 0.5$. The effects of the shape parameters on the skewness and kurtosis can be studied by using (3.1).
The Bowley skewness (S) and Moors kurtosis(K) can be formulated as

$$(S) = \frac{Q\left(\frac{3}{4}\right)+Q\left(\frac{1}{4}\right)-2Q\left(\frac{1}{2}\right)}{Q\left(\frac{3}{4}\right)-Q\left(\frac{1}{4}\right)} \text{ and } (K) = \frac{Q\left(\frac{3}{8}\right)-Q\left(\frac{1}{8}\right)+Q\left(\frac{7}{8}\right)-Q\left(\frac{5}{8}\right)}{Q\left(\frac{6}{8}\right)-Q\left(\frac{2}{8}\right)}$$

where $Q(.)$ represents the quantile function. When the distribution is symmetric, $S = 0$ and when the distribution is right (or left) skewed, $S > 0$(or $S > 0$). The tail of the distribution gets thicker as K expands. These metrics exist even for distributions without moments and are less subject to outliers.

II. Critical Points and Asymptotes

The critical points of $f(x)$ are the solution x_0 of the nonlinear equation $f'(x_0) = 0$ i.e.,

$$\frac{2f(x)^2 + F(x)f'(x)}{p} = 0$$

The critical points of $h(x)$ are the solution x_* of the nonlinear equation $h'(x_*) = 0$ i.e.,

$$\frac{2f(x_*)^2[F(x_*)^2 + p] + 2F(x_*)f'(x_*)[p - F(x_*)^2]}{[p - F(x_*)^2]^2} = 0$$

By identifying the sign of the second derivative of the function taken at this point, we are able to identify the type of the critical point.

IV. Ordinary and Incomplete moments, Moment generating Function and Mean Deviation

Moments are crucial in the fields of actuarial and financial science, especially in applications. It assists the researcher in taking important features and characteristics of the suggested distribution under perspective. The r^{th} moment of the TR-G family of distribution is given by

$$\mu'_r = \int_{-\infty}^{\infty} x^r g_{Tr-G}(x; p, \Phi)\, dx \qquad (3.4)$$

Using the pdf of TR-G family of distribution (2.1) in equation number (3.4), we have

$$\mu'_r = \int_{-\infty}^{\infty} \frac{x^r F(x)^2}{p} dx$$

Using Binomial Expansions

$$\mu_r{}' = \sum_{k=0}^{\infty} \alpha_k \psi_{2k+1}$$

Where $\alpha_k = \sum_{k=0}^{\infty} \frac{2(-1)^k}{p(2k+1)!}$ and $\psi_{2k+1} = \sum_{k=0}^{\infty} G(x)^{4k+2}$

The i^{th} incomplete moment is defined as $I\left(x; p, \Phi\right)$ and is given by

$$I(x, \theta, \Phi) = \int_{-\infty}^{t} x^i f(x, \theta, \Phi)\, dx$$

$$I(x) = \sum_{k=0}^{\infty} \alpha_k \Psi_{i,2k+1}$$

Getting the mean deviations is important for lifetime models as well. The following are the possible ways to express the mean deviations from the mean and median for a random variable $X \sim TR - G$.

$$\varepsilon_1 = \int_0^{\infty} |x - \mu_1'|\, g_{TR-G}(x, p, \Phi) dx = 2\mu_1' G(\mu_1') - 2I_{(1)}(\mu_1').$$

where $I_1(\mu_1)$ is the first incomplete moment of TR-G family.

$$\varepsilon_2 = \int_0^{\infty} |x - Q(0.5)|\, g_{TR-G}(x, p, \Phi) dx = \mu_1' - 2I_{(1)}\big(Q(0.5)\big)$$

The moment-generating function and cumulant-generating function for the TR-G family can be expressed in a general form as follows

$$M_x(t) = \sum_{r,n=0}^{\infty} \frac{t^r}{r!}\, \alpha_k \Psi_{r,2k+1}.$$

$$\Phi_x(t) = M_x(it) = \sum_{r,n=0}^{\infty} \frac{(it)^r}{r!}\, \alpha_k \Psi_{r,2k+1}.$$

V. Reliability function for parallel and series systems

Let us Consider an independent system with $n *$ TR-G family-equipped components. The reliability of the parallel system (P) and reliability of the series system (S) are provided by

$$R_p(x; p, \Phi) = \left[1 - \left\{\frac{2F(x)f(x)}{p}\right\}\right]^{\theta n *}.$$

$$R_s(x; p, \Phi) = \left[\left\{1 - \left\{\frac{2F(x)f(x)}{p}\right\}\right\}^{\theta}\right]^{n *}.$$

VI. Mean time to failure (MTTF), mean time between failure (MTBF) and availability (AvB)

The reliability signs MTTF, MTBF, and AvB are based on techniques and procedures for predicting a product's longevity. A failure rate and the subsequent time frame of expected performance may be quantified using metrics such as MTTF, MTBF, and AvB, which are techniques of delivering a numerical number based on a compilation of data.

If $X \sim TR - G(p_1, \Phi_1)$ then the MTBF is given as

$$MTBF = \frac{-x}{\ln(1-G(x;p_1,\Phi_1))};\ x > 0.$$

If $X \sim TR - G(p_2, \Phi_2)$ then the MTTF is given as

$$MTTF = E(X) = \mu_1'|(p_2, \Phi_2);\ x > 0.$$

The AvB is consider the probability that the component is successful at time X, i.e.

$$AvB = MTTF/MTBF = -\mu_1'|(p_2, \Phi_2)\frac{\ln(1 - G(x, p_1, \Phi_1))}{x}.$$

VII.Bonferroni and Lorenz curves

Bonferroni and Lorenz curves defined for a given probability π is given by

$$B(\pi) = I_1(q)/\pi\mu_1' \text{ and } L(\pi) = I_1(q)/\mu_1'.$$

Where $q = Q(\pi)$ is the quantile function of X at π.

IV. Special Members of TR-G family of distribution

This section carries certain cases of the intended family of distributions by using different base cumulative distribution functions.

I. TR-Inverse Weibull Distribution

Let us consider the cdf and pdf of Inverse Weibull distribution with positive parameter (α, β) given by $\alpha\beta^{\alpha}x^{-\alpha-1}e^{-\left(\frac{\beta}{x}\right)^{\alpha}}$ and $e^{-\left(\frac{\beta}{x}\right)^{\alpha}}$ respectively with the random variable X. Considering that $F(x;\alpha,\beta)$ and $f(x;\alpha,\beta)$ are the cdf and pdf of the two-parameter Inverse Weibull distribution.
The cdf of the three parameter TR-IW distribution (substituting in (2.2)), for $x > 0$, can be expressed as

$$G_{TR-IW}(x;\alpha,\beta,p) = \frac{e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}}{p};\quad x \in \mathbb{R}, p > 0. \tag{4.1}$$

The corresponding pdf and the complementary cdf (or survival function (srf)), instantaneous failure rate (or hazard rate function (hrf), retro hazard (or reversed hazard rate function), integrated hazard rate (or cumulative hazard rate function) (three parameter) can be written as below

$$g_{TR-IW}(x;\alpha,\beta,p) = \frac{2\alpha\beta^{\alpha}x^{-\alpha-1}e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}}{p};\quad x \in \mathbb{R}, p > 0. \tag{4.2}$$

$$S_{TR-IW}(x;\alpha,\beta,p) = 1 - \left[\frac{e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}}{p}\right]; x \in \mathbb{R} \tag{4.3}$$

$$h_{TR-IW}(x;\alpha,\beta,p) = \frac{4\alpha\beta^{\alpha}x^{-\alpha-1}e^{-\left(\frac{\beta}{x}\right)^{4\alpha}}}{p\left[p-\left\{e^{-\left(\frac{\beta}{x}\right)^{4\alpha}}\right\}\right]}; x \in \mathbb{R} \tag{4.4}$$

$$r_{TR-IW}(x;\alpha,\beta,p) = 4\alpha\beta^{\alpha}x^{-\alpha-1}; x \in \mathbb{R} \quad (4.5)$$

$$H_{TR-IW}(x;\alpha,\beta,p) = -\log\left[1-\left\{\left[\frac{e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}}{p}\right]\right\}\right]; x \in \mathbb{R} \quad (4.6)$$

By substituting $p = 1$, the two parameter TR-IW exists with the pdf and cdf

$$g_{TR-IW}(x;\alpha,\beta) = 2\alpha\beta^{\alpha}x^{-\alpha-1}e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}; \quad x \in \mathbb{R},\ (\alpha,\beta) > 0. \quad (4.7)$$

$$G_{TR-IW}(x;\alpha,\beta) = e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}; \quad x \in \mathbb{R},\ (\alpha,\beta) > 0. \quad (4.8)$$

$$S_{TR-IW}(x;\alpha,\beta) = 1 - e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}; \quad x \in \mathbb{R},\ (\alpha,\beta) > 0. \quad (4.9)$$

$$h_{TR-IW}(x;\alpha,\beta) = \frac{2\alpha\beta^{\alpha}x^{-\alpha-1}e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}}{1-e^{-\left(\frac{\beta}{x}\right)^{2\alpha}}}; \quad x \in \mathbb{R},\ (\alpha,\beta) > 0. \quad (4.10)$$

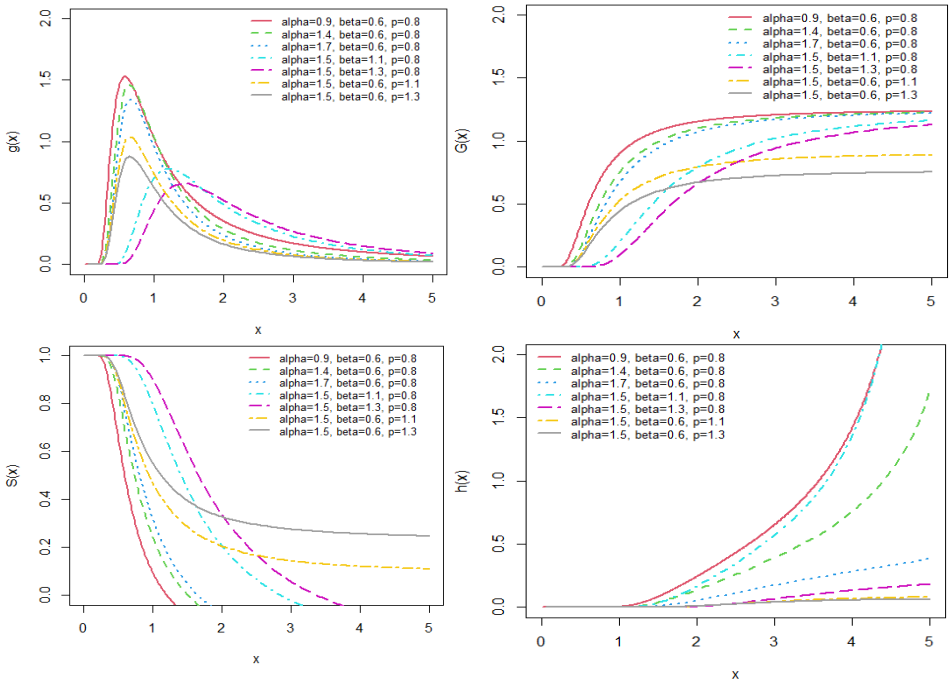
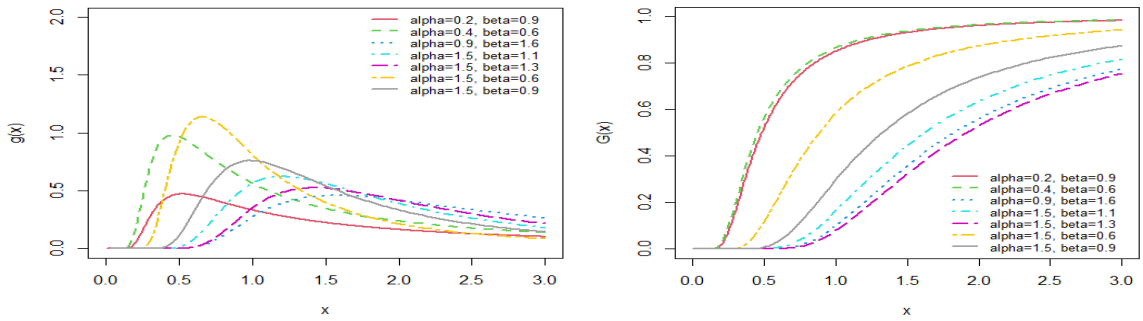


Figure 1: *pdf, cdf, survival and hazard plot of Tr-IW (three parameter) distribution*



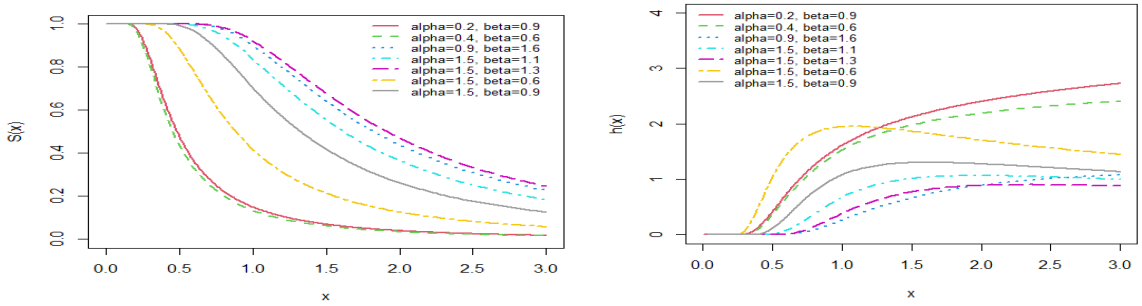


Figure 2: *pdf, cdf, survival and hazard plot of Tr-IW (two parameter) distribution*

The PDF in figures 1 and 2 can have different forms based on the values of the parameters. The shape of the proposed distribution is closed to bell shape by increasing the shape parameter. Furthermore, the hrf can be increasing or unimodal-bell shape, it increases the distribution's adaptability to fit various sets of lifespan data, as seen in figure 4.

Table 2: *New contributed special cases of the Triangle-G family*

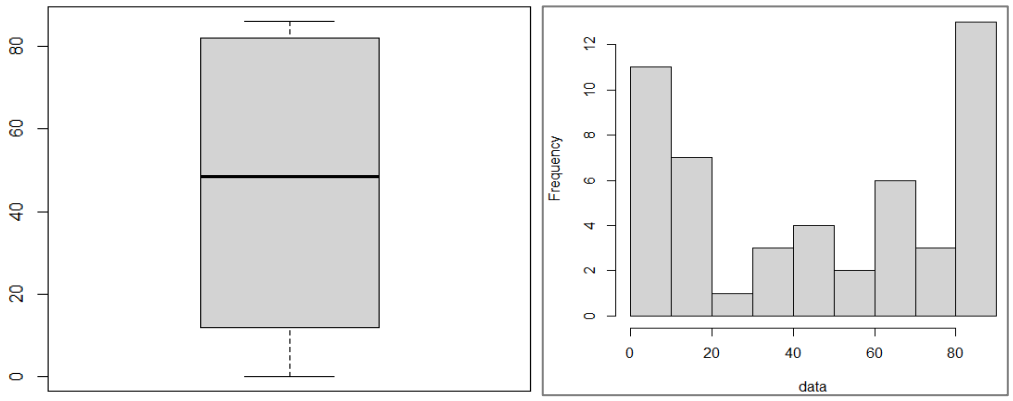
Sl. No.	Baseline mode	CDF form	Generated Model	Support
1	Exponential	$\frac{\left(1-\exp(-\lambda x)\right)^2}{p}$	Tr-E	$x \in \mathbb{R}^*$
2	Rayleigh	$\frac{\left(1-\exp\left(\frac{-x^2}{2\sigma^2}\right)\right)^2}{p}$	Tr-R	$x \in \mathbb{R}^*$
3	Frechet	$\frac{\exp\left(\frac{-x-m}{s}\right)^{-2\alpha}}{p}$	Tr-F	$x \in \mathbb{R}^*$
4	Gamma	$\frac{\left(\left(\frac{1}{\Gamma(\alpha)}\right)\gamma(\alpha,\beta x)\right)^2}{p}$	Tr-G	$x \in \mathbb{R}^*$
5	Lomax	$\frac{\left(1-\left(1+\frac{x}{\lambda}\right)\right)^{-2\alpha}}{p}$	Tr-L	$x \in \mathbb{R}^*$

V. Practical Illustration

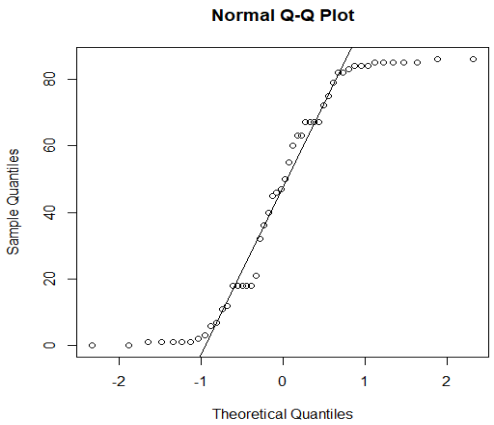
This section discusses the theoretical significance of the Tr-G model utilizing two applications to complete real data. The competitive distributions' best-fitting capabilities are determined using certain analytical metrics. To choose the most suited ones, the values of the Akaike Information Criterion (AIC),
Hannan-Quinn Information Criterion (HQIC), Corrected Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC) were taken into consideration. Other goodness-of-fit tests, such

as the Cramer-von Mises (W) distance value test, the Kolmogorov-Smirnov (K-S) statistic with accompanying p values, and the loglikelihood function, are also recorded in addition to discriminating tests. The AIC, BIC, CAIC, and HQIC values as well as the W and K-S tests are consistently should be lowest for the ideal model. To compare the competing distributions, the model with the highest p values for the K-S statistics is used. Two data sets have been taken into consideration.

Dataset 1: The Tr-IW (three parameter) distribution is analysed using the dataset contained the lifetimes of fifty devices. They were given by: 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 83, 84, 84, 84, 85, 85, 85, 85, 85, 86, 86, 0.1, 0.2, 1, 1, 79, 82, 82, 1, 1, 1, 2, 3, 6, 7, 11, 12, 18, 18, 18, 18, and 18.



(a) (b)



(c)

Figure 3: *Boxplot (a), Histogram (b) and Normal QQ plot (c) for Data set 1*

The competing models included the generalized modified Extended Cosine Power (ECSP) model [29], Weibull–Poisson (GMWP) model, generalized modified Weibull-Geometric (GMWG) model, generalized modified Weibull-logarithmic (GMWL) model [7], Poisson-odd generalized uniform (POGE-U) model [28], exponentiated generalized linear exponential (EGLE) model [35], gamma-uniform (GU) model [41], generalized linear failure rate (GLFR) model [36], beta Weibull (BW) model [21], generalized modified Weibull (GMW) model [9], modified Weibull distribution (MW) model [20], generalized linear exponential (GLE) model [25], beta-modified Weibull (BMW) model [37], power (P) model.

Table 3: *MLE's and other statistics value for dataset 1*

Model	Estimated Parameter					Model Comparison Method			
	$\hat{p}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$	$-L$	AIC	BIC	K-S (p-value)
Tr-IW3	0.10	0.23	0.44	-	-	128.30	262.60	268.33	0.05 (0.97)
ECSP	0.21	86.01	0.35	-	-	202.59	411.91	416.92	0.08 (0.86)
GU	0.27	51.94	0.09	86.71	-	207.33	418.65	426.30	0.15 (0.20)
TUq	-0.19	0.10	86.0	0.93	-	212.86	433.72	441.37	0.12 (0.42)
BMW	2.4×10^{-4}	0.05	0.20	0.17	1.4	220.28	450.56	460.12	0.13 (0.36)
BW	1.0×10^{-5}	0.13	0.07	3.32	-	223.11	454.22	461.87	0.12 (0.42)
MW	0.06	0.02	0.36	-	-	226.16	460.31	466.05	0.14 (0.33)
EGLE	3.3×10^{-3}	1.7×10^{-4}	4.56	0.11	-	224.34	456.67	464.32	0.15 (0.21)
GLE	9.9×10^{-3}	4.5×10^{-4}	0.73	-	-	235.93	477.85	483.59	0.16 (0.14)
GLFR	3.8×10^{-3}	3.1×10^{-4}	0.53	-	-	233.15	472.29	478.03	0.16 (0.13)
POGE-U	0.02	0.37	1.77	87.01	-	206.68	419.34	425.08	0.09 (0.75)
GMWP	5.4×10^{-8}	0.13	0.08	2.13	-	220.88	451.75	461.31	0.14 (0.28)
GMWL	2.13	2.68	0.01	0.28	1.00	217.77	445.53	455.09	0.13 (0.36)
GMWG	9.4×10^{-8}	0.12	0.08	2.23	0.46	220.78	451.55	461.11	0.13 (0.33)
GMW	1.0×10^{-5}	0.07	0.22	1.37	-	221.40	452.81	460.46	0.15 (0.23)
P	86.01	0.73	-	-	-	219.89	433.78	447.60	0.99 8.9×10^{-16}

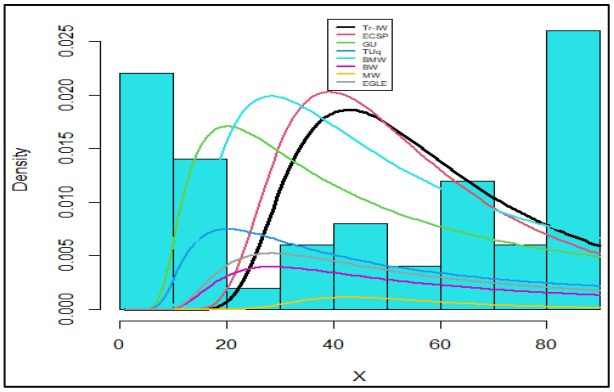
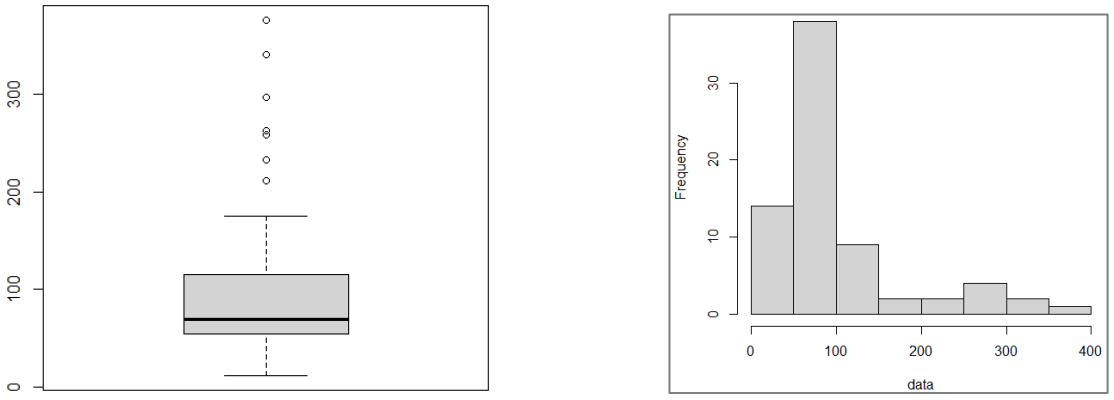
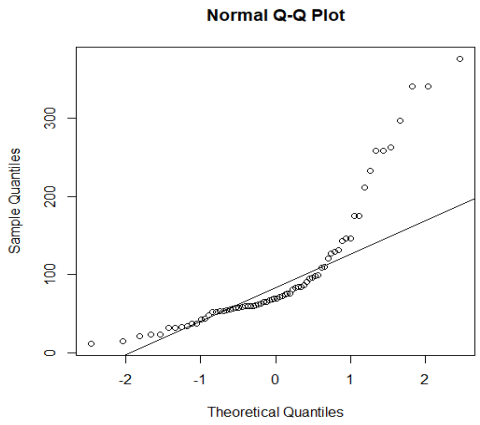


Figure 4: *Plots of the estimated pdfs for dataset 1*

Dataset 2: The dataset of 72 survival times in days of guinea pigs, voluntary contaminated with different doses of tubercle bacilli [8] is used for analysed the Tr-IW (two parameter). The data are listed as 12, 15, 22, 24, 24, 32, 32, 33, 34, 38, 38, 43, 44, 48, 52, 53, 54, 54, 55, 56, 57, 58, 58, 59, 60, 60, 60, 60, 61, 62, 63, 65, 65, 67, 68, 70, 70, 72, 73, 75, 76, 76, 81, 83, 84, 85, 87, 91, 95, 96, 98, 99, 109, 110, 121, 127, 129, 131, 143, 146, 146, 175, 175, 211, 233, 258, 258, 263, 297, 341, 341, 376.



(a) (b)



(c)

Figure 5: *Boxplot (a), Histogram (b) and Normal QQ plot (c) for Data set 2*

The comparative distributions are the New Sine Inverse Weibull (NSIW) model [23], sine inverse Weibull model (SIW) [19], inverse Weibull model (IW) [17], inverse Nadarajah-Haghighi model (INH) [40], inverse exponential model (IED) [18] and inverse Rayleigh model (IRD) [42].

Table 4: *MLE's and other statistics value for dataset 2*

Model	Estimated Parameter		Model Comparison			
	$\hat{\alpha}$	$\hat{\beta}$	$-L$	AIC	BIC	$KS\,(p-value)$
Tr-IW2	0.74	33.07	356.25	716.51	721.06	0.09 (0.85)
NSIW	1.19	59.28	391.11	786.23	790.78	0.12 (0.25)
SIW	1.09	78.68	391.82	787.66	792.21	0.13 (0.20)
IW	1.42	54.15	395.65	795.47	799.85	0.15 (0.07)
INH	1.84	25.78	400.47	804.94	809.49	0.14 (0.11)
IED	60.09	-	402.67	807.34	809.62	0.18 (0.01)
IRD	2124.00	-	406.77	815.53	817.81	0.26 (0.0001)

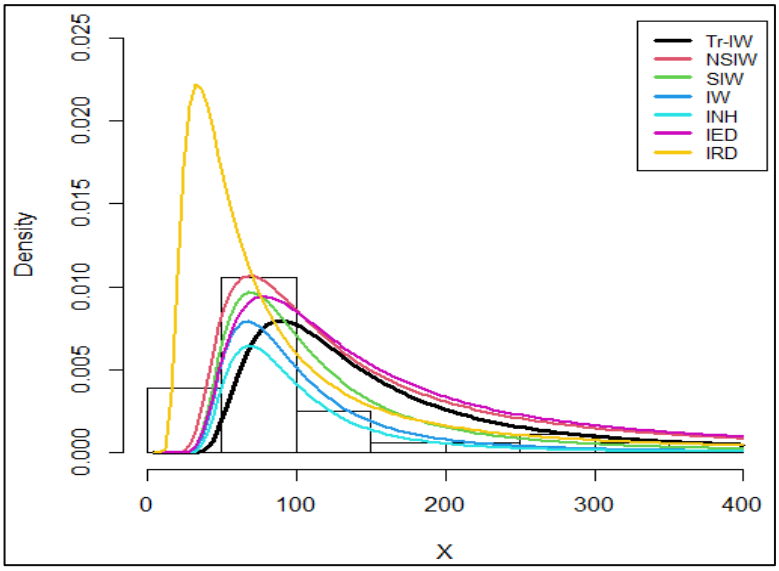


Figure 6: *Plots of the estimated pdfs for dataset 2*

VI. Results

The three parameter Tr-IW distribution is applied in Dataset1 and compared with three, four and more parameter distribution models. The Tr-IW model is fitted comfortably more flexible than other models with more parameters. The two parameter Tr-IW model is fitted better than the other equal parameter models.

VII. Conclusion and Remarks

The triangle family of distribution is introduced. Some of the important properties are discussed. Inverse Weibull model is taken as sub model distribution. The paper introduced two types of distribution with two and three parameters. Both of the distributions are discussed with various properties and real-life data fitting. Both of the distributions are fitted consistently better than the other models with equal and more parameter. This paper introduced one created family and two generated model with a hope that it will attract wider applications in several areas such as reliability engineering, insurance, hydrology, economics and survival analysis.

References

[1] Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Trans. Reliab,* 36:106–108.
[2] Abdel-Hamid, A. H. (2016). Properties, estimations and predictions for a Poisson-half-logistic distribution based on progressively type-II censored samples. *Appl. Math. Model*, 40:7164–7181.
[3] Ahmad, Z., Mahmoudi, E., Alizadeh, M., Roozegar, R. and Afify, Z. A. (2021). The Exponential T-X Family of Distributions: Properties and an Application to Insurance Data. *Journal of Mathematics*.
[4] Al-Babtain, 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.
[5] Alkhairy, I., Nagy, M., Muse, A. H. and Hussam, E. (2021). The Arctan-X Family of Distributions: Properties, Simulation, and Applications to Actuarial Sciences. *Complexity*.
[6] Alzaatreh, A., Lee, C. and Famoye, F. (2013). *A new method for generating families of continuous distributions. Metron*, **71**(1):63–79.
[7] Bagheri, S., Samani, E. B. and Ganjali, M. (2016). The generalized modified Weibull power series distribution: Theory and applications. *Comput. Stat. Data Anal*, 94:136–160.
[8] Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Amer J Hygiene*, 72:130-148.
[9] Carrasco, J. M., Ortega, E. M. and Cordeiro, G.M. (2008). A generalized modified Weibull distribution for lifetime modelling. *Comput. Stat. Data Anal*, 53:450–462.
[10] Chesneau, C., Bakouch, H. S. and Hussain, T. (2019). A new class of probability distributions via cosine and sine functions with applications. *Communications in Statistics - Simulation and Computation*, 48(8):2287-2300.
[11] Chesneau, C., Tomy, L. and Gillariose, J. (2021). On a new distribution based on the arccosine function. *Arab. J. Math,* 10:589–598.
[12] Chesneau, C. and Jamal, F. (2019). The Sine Kumaraswamy-G Family of Distributions. *HAL-02120197*.
[13] Cordeiro, G. M., de Andrade, T. A., Bourguignon, M. and Gomes-Silva, F. (2017). The exponentiated generalized standardized half-logistic distribution. *Int. J. Stat. Probab*. 6:24–42.
[14] Eghwerido, J. T., Nzei, L. C., Omotoye, Adebola, E. and Agu, F. I. (2022). The Teissier-G family of distributions: Properties and applications. *Mathematica Slovaca*. 72(5):1301-1318.
[15] He, W., Ahmad, Z., Afify, A. Z. and Goual, H. (2020). The Arcsine Exponentiated-X Family: Validation and Insurance Application. *Complexity*.

[16] Jose, J.K. and Manoharan, M. (2016). Beta half logistic distribution—A new probability model for lifetime data. J. *Stat. Manag. Syst*. 19:587–604.

[17] Keller, A.Z. and Kamath, A.R. (1982). Alternative reliability models for mechanical systems. Third international Conference on Reliability and Maintainability. Toulouse, France, 411-415.

[18] Keller, A.Z. and Kamath, A.R. (1982). Reliability analysis of CNC machine tools. *Reliability engineering*, 3:449-473.

[19] Kumar, D., Singh, U. and Singh, S. K. (2015). A New Distribution Using Sine Function Its Application to Bladder Cancer Patients Data. *J. Stat. Appl. Pro*, 4(3):417-427.

[20] Lai, C., Xie, M. and Murthy, D. (2003). A modified Weibull distribution. *IEEE Trans. Reliab*, 52: 33–37.

[21] Lee, C., Famoye, F. and Olumolade, O. (2007). Beta-Weibull distribution: Some properties and applications to censored data. *J. Mod. Appl. Stat. Methods*, 6 (17).

[22] Liang Tung, Y., Ahmad, Z. and Mahmoudi, E. (2021). The arcsine-x family of distributions with applications to financial sciences. *Computer Systems Science and Engineering*, 39(3):351–363.

[23] Mahmood, Z. and Chesneau, C. A. (2019). New Sine-G Family of Distributions: Properties and Applications. *HAL-02079224f*.

[24] Mahmood, Z., Jawa, T. M., Ahmed, N. S., Khalil, E. M., Muse, A. H. and Tolba, A. H. (2022). An Extended Cosine Generalized Family of Distributions for Reliability Modeling: Characteristics and Applications with Simulation Study. *Mathematical Problems in Engineering*.

[25] Mahmoud, M. A. and Alam, F. M. A. (2010). The generalized linear exponential distribution. *Stat. Probab. Lett*, 80:1005–1014.

[26] Muhammad M., Alshanbari H. M., Alanzi A. R. A., Liu L., Sami W., Chesneau C. and Jamal F. (2021). A New Generator of Probability Models: The Exponentiated Sine-G Family for Lifetime Studies. *Entropy*, 23(11):1394-1409.

[27] Muhammad, M. (2017). Generalized half-logistic Poisson distributions. *Commun. Stat. Appl. Methods*, 24:353–365.

[28] Muhammad, M. (2016). Poisson-odd generalized exponential family of distributions: Theory and applications. *Hacet. J. Math. Stat*, 47:1652–1670.

[29] Muhammad, M., Bantan, R.A.R., Liu, L., Chesneau, C., Tahir, M.H., Jamal, F. and Elgarhy, M. (2021). A New Extended Cosine—G distributions for Lifetime Studies. *Mathematics*. 9:2758-2771.

[30] Muhammad, M. and Liu, L. (2021). A New Three Parameter Lifetime Model: The Complementary Poisson Generalized Half Logistic Distribution. *IEEE Access*, 9:60089–60107.

[31] Muhammad, M. and Yahaya, M.A. (2017). The half logistic-Poisson distribution. *Asian J. Math. Appl*.

[32] Olapade, A. (2014). The type I generalized half logistic distribution. *J. Iran. Stat. Soc*. 13:69–82.

[33] Rahman, M. Arcsine-G Family of Distributions. *J. Stat. Appl. Pro. Lett*, 8(3): 169-179.

[34] Sakthivel, K. M. and Rajkumar, J. (2022). Transmuted Sine - G Family of Distributions: Theory and Applications. *Statistics and Applications*. 20(2):73–92.

[35] Sarhan, A. M., Abd EL-Baset, A. A. and Alasbahi, I. A. (2013). Exponentiated generalized linear exponential distribution. *Appl. Math. Model*, 37:2838–2849.

[36] Sarhan, A. M. and Kundu, D. (2009). Generalized linear failure rate distribution. *Commun. Stat.-Theory Methods*, 38:642–660.

[37] Silva, G.O., Ortega, E.M. and Cordeiro, G.M. (2010). The beta modified Weibull distribution. *Lifetime Data Anal*, 16:409–430.

[38] Souza L. (2015). New trigonometric classes of probabilistic distributions. Thesis, Universidade Federal Rural de Pernambuco.

[39] Souza, L., de Oliveira, W. R., de Brito, C. C. R., Chesneau, C., Fernandes, R. and Ferreira, T. A. E. (2022). Sec-G Class of Distributions: Properties and Applications. *Symmetry*. 14:299-312.

[40] Tahir, M. H., Cordeiro, G. M., Ali, S., Dey, S. and Manzoor, A. (2018). The inverted Nadarajah- Haghighi distribution: estimation methods and applications. *Journal of Statistical Computation and simulation,* 88(14):2775-2798.

[41] Torabi, H. and Hedesh, N. M. (2012). The gamma-uniform distribution and its applications. *Kybernetika*, 48:16–30.

[42] Voda, V. G. (1972). On the inverse Rayleigh random variable. *Rep Stat Appl Res*, 19:13-21.

RESEARCH OF ELECTROMECHANICAL DEVICES WITH LEVITATION ELEMENTS IN CONTROL SYSTEMS

G.S.Kerimzade[1], G.V.Mamedova[2]

•

Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[1] gulschen98@mail.ru ; [2] gulaya68@mail.ru

Abstract

The work examines the main technical indicators of electromechanical converters callers with levitation elements, a generalized design method has been developed tions, as well as design diagrams and functional dependencies of the main varieties of electromechanical devices with elements of levitation. Analytical expressions for the levitation coefficient as a function of the dimensions of the magnetic core and coefficient factor of force multiplicity, technical characteristics of levitation material element, set superheat temperature. A mathematical model has been compiled for based on the parameters of the current mode and forces from the equations of electrical, magnetic, mechanical and thermal circuits of the magnetic system. As a result, the main dimensions of the magnetic system and dimensionless quantities. Analytic expressions for the main dimensions, the specified values of the winding overheating temperature are taken into account, input and output parameters, the condition of uniformity of the magnetic field in the working air nom gap. The optimal values of the dimensions of the magnetic circuit, active resistors have been determined winding voltages are minimal, resulting in minimization of losses active capacities.

Key words: levitation element, electromechanical apparatus, research, method, magnetic system, mathematical model, current mode, force mode, stability, mechanism.

I. Introduction

The widespread use of electromechanical devices with levitation elements in automatic control systems ensures high reliability, accuracy, and stability during control and regulation of parameters and the technological process as a whole.

The designs of electromechanical devices with a levitation screen are more effectively involved in solving these problems, since in these devices there are no friction forces, the working stroke of the moving part is automatically controlled and additional elements are not required (for example, mechanical springs, guides, gearboxes, supports, etc.) [5-18].

Automation of technological processes requires automatic control of the vertical positions of the moving parts of working mechanisms using external force and alternating current voltage. In this case, there is a need to measure external force, stabilize the current on a variable load and obtain several nominal values of the current on the load. The design of a simple electromechanical device with a levitation element consists of a vertically located magnetic circuit 1, a stationary alternating current winding 2 and a levitation element 3 (fiqure 1). In the force mode, the levitation element is made in the form of a solid aluminum frame, and in the current mode - in the form of a short-circuited winding. When the device is turned on to the power source, the levitation element may

strike the upper yoke of the magnetic circuit.

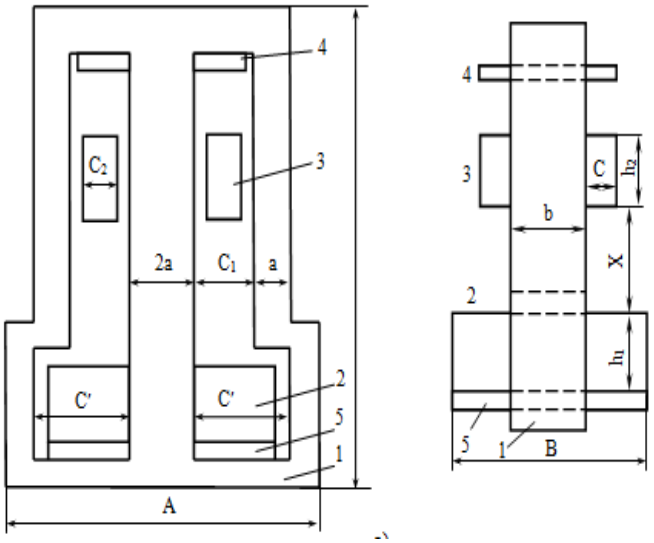


Figure 1. *Design of a simple electromechanical device with LE*

A magnetic system with a levitation element (straight and stepped forms) is shown in figure 2. To eliminate this undesirable phenomenon, a compensation winding 4 is placed near the yoke, which is connected in series with the power winding 1; a signal winding 5 is also provided. During operation of the device with the levitation element, the compensation winding is turned off. The AC winding (or excitation winding) is powered by an AC voltage source U_1 and is made of several sections, by switching which a family of control characteristics is achieved [1-4, 9].

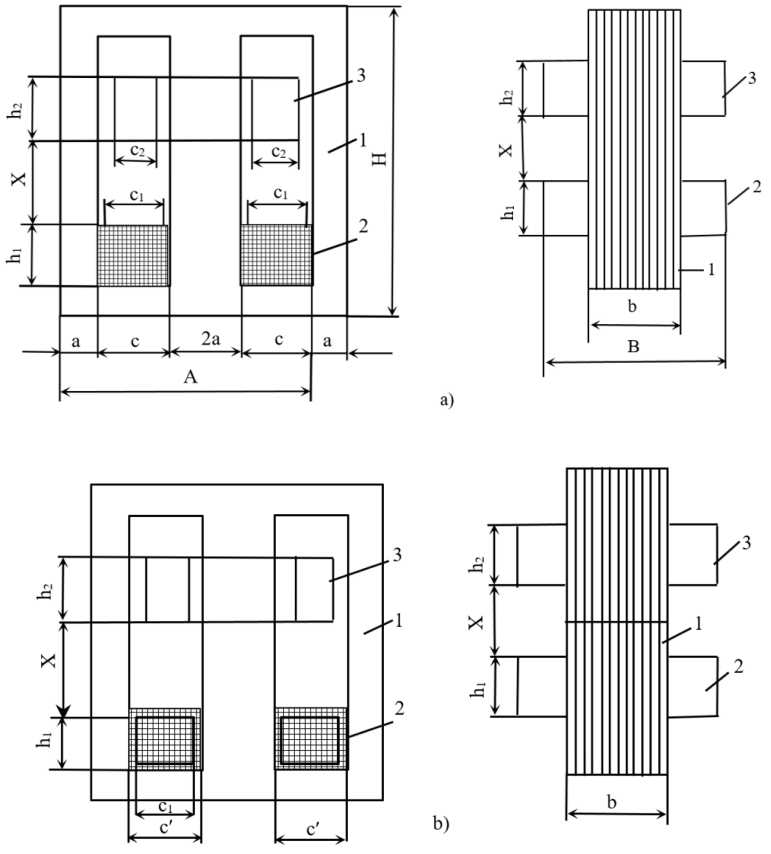


Figure 2. *Magnetic system with levitation element (straight (a) and step b))*

II. Formulation of the problem

When the excitation winding (EW) is connected to the power source, currents flow through the EW and the levitation element (LE), which significantly exceed the rated currents. As a result, losses in the LE sharply increase and all the energy released in it goes to heating it. Depending on its size, the temperature can reach a high value and, before reaching the steady-state value, the LE can melt. Therefore, the minimum dimensions of the LE must be limited.

In accordance with the insulation class, the permissible values of the LE overheating temperature are set. To reduce it, it is necessary to reduce the lateral heat transfer surface of the LE, which in turn is associated with an increase in the aspect ratio $n_{e2}=h_2/c_2$ under the condition $S_{02}=c_2 \cdot h_2=const$ and this can lead to an increase in the height of the magnetic system and a deterioration in the lateral stability of the LE. To determine the optimal relationship between the main dimensions of the LE c_2 and h_2, the dependence of the dimensionless quantity n_{e2} on the geometric dimensions of the magnetic circuit (a, b, c), the force multiplicity factor n_p, the physical and technical characteristics of the LE material and the given overheating temperature τ_2 is determined.

III. Problem solution

Based on the established dependencies of dimensions and parameters, a mathematical model is compiled for certain parameters of current and force modes:

$$\tau_2 = \frac{P_2}{k_T S_{T2}} = F_1^2 \frac{b_2^2 \rho_2}{k_{32} k_T} \left(\frac{l_2}{S_2 S_{T2}} \right) \tag{1}$$

$$F_1^2 = \frac{2}{\lambda} \left(P_x + P_T \right) = 2g\gamma k_{32} n_p \frac{S_2 l_2}{\lambda} \tag{2}$$

Here generally accepted designations. Force multiplicity factor:

$$n_p = 1 + \frac{P_x}{P_T} \tag{3}$$

As is known, for LE made of copper and aluminum:

$$(S_2 l_2) = 2c_2^3 n_{e2} k_u \tag{4}$$

Accordingly we have:

$$\left(\frac{l_2}{S_2 S_{T2}} \right)_m = \frac{1}{2n_{e2}^2 c_2^3 \left(1 - n_{e2}\right)} \tag{5}$$

By jointly solving equations (1)-(2) for LE made of copper and aluminum, we have:

$$n_{e2} = E_m M_m n_p \tag{6}$$

$$n_{e2} = E_a M_a n_p - 1 \tag{7}$$

where are the coefficients:

$$E_m = 2.2585 \cdot 10^{10} \left(\frac{\rho_2}{\tau_2} \right)_m \tag{8}$$

$$\left(\frac{\rho_2}{\tau_2} \right)_m = \frac{1.72 \cdot 10^{-8}}{\tau_2} (1.063 + 0.0042\tau_2) \tag{9}$$

$$M_m = \frac{\left(m_a + m_c + 0.5 m_a m_c\right)}{\left(0.909 m_a + m_a + m_c + m_a m_c\right)\left[m_a m_c + 2.92 m_a \lg\left(1 + \frac{\pi}{m_a}\right)\right]} \quad (10)$$

$$E_a = 0.3432 \cdot 10^{10}\left(\frac{\rho_2}{\tau_2}\right)_a \quad (11)$$

$$\left(\frac{\rho_2}{\tau_2}\right)_a = \frac{2.8 \cdot 10^{-8}}{\tau_2}(1.063 + 0.0042\tau_2) \quad (12)$$

Expressions (6)-(12) when implemented using the appropriate calculation program make it possible to compile tables 1-3, which show the parameter values for LE made of copper and aluminum. Figure 3 respectively show the graphical dependences of the coefficient $n_{e2} = f(\tau_2)$ on the current at the stroke values [7-16].

Table 1. Values of coefficients Mm, Ma

m_c	m_a					
	2	3	4	5	6	M_m, M_a
2	0.573464	0.53438	0.518343	0.510589	0.506504	M_m
	0.938378	0.916929	0.914149	0.91672	0.920897	M_a
3	0.594826	0.530443	0.49976	0.482146	0.470856	M_m
	0.953061	0.89554	0.870198	0.85682	0.848954	M_a
4	0.608226	0.528249	0.489108	0.466072	0.450967	M_m
	0.962093	0.882866	0.844812	0.822862	0.808747	M_a
5	0.617415	0.526858	0.482207	0.455742	0.43828	M_m
	0.96821	0.874482	0.828279	0.800992	0.783072	M_a
6	0.624107	0.525901	0.477373	0.448546	0.429485	M_m
	0.972627	0.868526	0.816656	0.785732	0.765256	M_a

Table 2. Values of levitation element parameters depending on overheating temperature

Parameter	ρ_{20},Om·m	τ_2,^{0}C			
		80	90	100	110
$\rho_{20}\cdot10^{-8}$Om·m	$1.72\cdot10^{-8}$	2.420	2.494	2.567	2.642
$\rho_{20}/\tau_2\cdot10^{-12}$		302.50	277.10	256.70	240.18
E_m		6.8323	6.2584	5.7977	5.4246
$\rho_{20}\cdot10^{-8}$Om·m	$2.87\cdot10^{-8}$	4.015	4.135	4.256	4.376
$\rho_{20}/\tau_2\cdot10^{-12}$		501.890	459.510	425.620	397.880
E_m		1.7226	1.5771	1.4608	1.3652

Table 3. n_{e2} values for a levitation element made of copper and aluminum at $\tau_2 = 80^0$C and n_p=1

m_c	Material	m_a				
		2	3	4	5	6
2	Al	3.918	3.651	3.541	3.488	3.460
	Cu	0.615	0.578	0.574	0.570	0.585
3	Al	4.064	3.624	3.414	3.294	3.217
	Cu	0.641	0.542	0.498	0.475	0.462
4	Al	4.155	3.609	3.341	3.184	3.081
	Cu	0.656	0.520	0.454	0.417	0.393
5	Al	4.218	3.599	3.294	3.114	2.994
	Cu	0.667	0.505	0.426	0.379	0.348
6	Al	4.264	3.593	3.261	3.065	2.934
	Cu	0.675	0.495	0.406	0.353	0.318

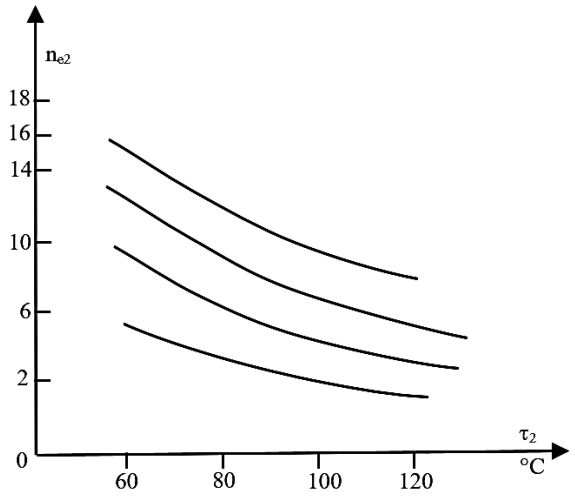


Figure 3. *Coefficient dependency $n_{e2}=f(\tau_2)$*

The patterns of changes in the values of the coefficients M_m, M_a from the dimensionless quantities m_a, m_c show that the smallest values of the coefficients M_m, M_a occur at $m_a = 6$, $m_c = 6$, and the largest values correspond to the values $m_a=2$, $m_c=6$. At $m_a=6$, $m_c=6$, the values of specific magnetic conductivity λ and cross-sectional area of the magnetic core rod S_c are minimal. At $m_a=2$, $m_c=6$, the values of these parameters, on the contrary, are maximum. Therefore, the overall dimensions in the first option are smaller than in the second. With increasing coefficients M_m, M_a and load capacity, the value of the coefficient n_{e2} increases, which leads to an increase in the overall size of the magnetic system. An increase in the overheating temperature leads to a decrease in the coefficients E_m, E_a, as a result, the overall size and dimensionless value n_{e2} decrease [6-14,18]. The given values of the coefficient n_{e2} in Table 3 allow you to pre-select the minimum values of the coefficient and the corresponding values of m_a, m_c. Next, the main dimensions of the magnetic system, other dimensions and parameters are determined, the minimum values of the coefficients, temperature, etc. are taken into account. in order to ensure the reliability of control of electromechanical devices in general.

IV. Conclusions

The obtained values take into account the minimum values of the coefficients n_{e1} and n_{e2}, the specified values of the overheating temperature τ_1 and τ_2, the working stroke x_p, the condition of uniformity of the magnetic field, the load current I_1, $m_a=2 \div 6$; $m_c=2 \div 6$. Electromechanical devices with levitation elements are low-current electrical devices, have simple designs, high accuracy, and stable performance characteristics. The active resistance of the excitation windings and the levitation element is minimal; as a result, losses of active power will also be reduced to a minimum.

Analytical expressions for the main parameters necessary for the design of electromechanical devices with levitation elements for various purposes are obtained. The calculation of electromechanical devices with levitation elements is significantly simplified by determining the optimal values for the height and thickness of the excitation winding and the levitation element.

References

[1] G.V.Mamedova, G.S. Kerimzade, N.M.Piriyeva."Electromagnetic calculation of tension devices for winding wires of small cross sections", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol. 14, No.4, pp. 80-85, December 2022.

[2] G.V.Mamedova, G.S. Kerimzade, N.M.Piriyeva."Issues of electrical devices with levitation elements", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 56, Vol. 15, No.3,

pp.120-125, September 2023.

[3] N.M.Piriyeva,G.S. Kerimzade."Methods for increasing electromagnetic efficiency in induction levitator", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT. ISSN0033-2097,R.99 NR 10/2023.Warszawa.pp192-195.

[4] N.M.Piriyeva,G.S. Kerimzade." Electromagnetic efficiency in induction levitators and ways to improve it", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097,R.99 NR 06/2023.Warszawa.pp204-207.

[5] N.M.Piriyeva,G.S. Kerimzade." Mathematical model for the calculation of electrical devices based on induction levitators", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 55, Vol.15, No.2, pp.274-280, Yune 2023.

[6] Ya.R. Abdullaev, G.S. Kerimzade, "Design of Electric Devices with LE", Electrical Engineering, No. 5, pp.16-22, 2015.

[7] G.S. Kerimzade, G.V. Mamedova, "Working Modes for Designing Electrical Apparatuses with Inductionmi with Levitation Elements", Bulletin Buildings, Vol. 17, No. 1, pp. 42-46, Baku, Azerbaijan, 2015.

[8] G.S. Kerimzade, G.V. Mamedova, "Analysis of the Parameters of Electric Devices with Levitational Elements", News of Universities Instrumentation, No. 12, Vol. 61, pp. 67-71, Sant Petersburg, Russia, 2018.

[9] G.S. Kerimzade, "Analytical Connections of the Parameters and sizes of the Precision Stabilizer of Alternating Current Using the Effect of Inductive Levitation", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 52, Vol. 14, No. 3, pp. 175-184, September 2022.

[10] G.S. Kerimzade, "Analysis of the Methodology for Calculation Current Stabilizer with Induction levitation", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 53, Vol.14, No. 4, pp. 170-174, December 2022.

[11] G.S. Kerimzade, "Analytical Expressions of the Relationship to the Calculation of the AC Stabilizer with Induction Levitation", International Journal on technical and Physical Problems of Engineering (IJTPE), Issue 54, Vol. 15, No. 1, pp. 135-139, March 2023.

[12] G.S. Kerimzade, "Controlled AC Stabilizers on the Principle of Induction Levitation", International Journal Universum, No. 2, Issue 107, Vol. 5, pp. 31-35, Moscow, Russai, 2023.

[13] G.S. Kerimzade, "High Precision Current Stabilizers with Induction Levitation", Journal of Renewable Energy, Electrical, and Computer Engineering, e-ISSN: 2776-0049, Vol. 3, No. 1, pp. 26-31, Indonesia, March 2023.

[14] G.S. Kerimzade, "Choice of Expressions of the Method of Calculation of the Current Stabilizer on the Principle of Induction Levitation", News of Azerbaijan High Technical Educational Institutions. 2023. Volume. 4, pp. 71-76.

[15] G.S. Kerimzade, "Optimization of parameters of AC stabilizer on the principle of induction levitation", The 19th International Conference on "TPE". ICTPE-2023. Number 06. Code 01PES06. 31 October 2023, pp.32-36.

[16] G.S.Kerimzade."Dependence of the overall dimensions of the control induction support", Vestnik nauki. Issue 2(71), Volume 1.February,04-2024.Tolyatti. pp.613-618.

[17] G.S. Kerimzade."Calculation of parameters of control induction support", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097,R.99NR 03/2024.Warszawa. pp192-195.

[18] G.V.Mamedova, G.S. Kerimzade."Design parameters for electromechanical devices with a levitation element", PRZEGLAD Elektrotechniczny Publishing house of magazines and technical literature SIGMA-NOT.ISSN0033-2097, 2024.Warszawa.(print in edidion)

A SOFTWARE RELIABILITY PREDICTION AND MANAGEMENT INCORPORATING CHANGE POINTS BASED ON TESTING EFFORT

ANUP KUMAR BEHERA[1] PRIYANKA AGARWAL[2]*

•

[1,2]Department of Mathematics, SRM Institute of Science and Technology, Delhi-NCR Campus, Ghaziabad, U.P. (INDIA)

[1]ab5656@srmist.edu.in, [2]priyankv@srmist.edu.in

*Corresponding Author

Abstract

This paper proposes a procedure for formulating the software reliability growth model using the non-homogeneous poisson process. We consider the software reliability growth model, which includes imperfect debugging, change points, and testing effort. Nevertheless, when formulating their software reliability models, the majority of scientists make the assumption of a constant detection rate per fault. When software is tested, they all suppose that each fault has an equal chance of being detected and that the rate is equal between generations. In practice, the fault detection rate varies depending on the test teams' abilities, program size, and software properties. Troubleshooting, even in the most realistic situations relevant to the error reintroduction rate due to incomplete debugging phenomena. In this case, changes in error detection and error introduction rates during software development program. Therefore, here we incorporate the generalized logic test workload function and change points. Parameters in software reliability modeling. Estimated using the least squares estimation method unknown parameters of the new model. Therefore, in our newly proposed model, we collect software testing data. use data from a practical application to illustrate the proposed model. Experimental results show that the proposed SRGM framework for imperfect debugging of integrated test jobs and change points has fairly accurate prediction capabilities.

Keywords: software reliability growth model, Non-homogeneous poison process(NHPP), Testing effort, Change Point

I. Introduction

Software reliability growth models (SRGMs) have a substantial historical background within the field of software engineering, serving as a pivotal tool for quantitatively evaluating and forecasting program reliability[1,2,3]. Over time, these models have developed to tackle the intricacies of software systems and the requirement for precise reliability evaluations. Many factors contribute to software failure, but mostly software fails from the design perspective. Software also fails whenever code is programmed or when changes are made to a project. Over the last few decades, numerous statistical models have been used to measure software reliability. As a result, we have discussed many established earlier models[4,5,6,7]. We believe that our new NHPP-based software reliability growth models have been proven quite efficient in practical software reliability engineering.

Researchers have examined several SRGMs throughout history to assess metrics like the number of remaining faults, software dependability, failure rate, failure intensity, and more. The literature has examined several classical models, focusing on factors such as time delay, correction procedure, fault severity, change point, and flawless debugging. Researchers have studied these models under specific assumptions. The researchers also incorporated the concept of perfect debugging, a process in which the testing team detects and fixes software errors, all while preventing the introduction of new errors during the testing process[8,9,10,11]. It is implausible that the statement is true, as the elimination process may introduce new defects that the testing team may be unaware of Several academics have suggested conducting experiments on faulty debugging.

There exist two distinct possibilities of incorrect debugging, specifically, i) imperfect fault removal and ii) generation of fault. When imperfect fault removal is present, the number of defects stays constant, signifying the elimination of the initial identified faults without introducing new ones. As new faults emerge in the system following the removal of the original problems, the overall fault content rises during the error generation scenario[12,13,14]. First introduced in 1985 was the concept of imperfect debugging. Subsequently, the concept of error production emerged, challenging the standard models' premise of complete flaw elimination upon detection. In our proposed model, we predict the existence of an imperfect debugging process that incorporates change points and testing efforts[15].

This paper fills this gap with this approach and is organized as follows. Sections II and III are discussed as NHPP software reliability growth models and software growth model change points respectively. In section IV, numerical description. Section V validates the analytical results and numerical interpretation. section VI cost model formulation and analysis of reliability and cost Section VII discusses the conclusion.

II. Non-homogeneous Poisson process uses software reliability growth models

The NHPP model is based on the assumption that software systems are subject to failures at random times due to the occurrence of residual errors. NHPP is often used to describe fault phenomena in the process testing phase. If N(t) follows a Poisson distribution with mean function m(t), then the counting process {N(t), t ≥ 0} is called NHPP with intensity function $\lambda(t)$, t ≥ 0 and is given as:

$$\Pr\{N(t)=k\}=\frac{[m(t)]^k}{k!}e^{-m(t)}, k=0,1,2,... \quad (1)$$

and

$$m(t)=\int_0^t \lambda(y)dy \quad (2)$$

Inversely,

$$\lambda(t)=\frac{dm(t)}{dt} \quad (3)$$

The failure intensity function $\lambda(t)$ or the mean function m(t) is the basic building block of all NHPP models.

The majority of reliability growth models for NHPP software operate under the assumption that the failure rate is directly proportional to the residual fault content. We derive a comprehensive category of NHPP-based SRGMs by solving the following differential equation:

$$\frac{dm(t)}{dt}=b(t)\left[a(t)-m(t)\right] \quad (4)$$

Where, a(t) represents the fault content function, which represents the entire number of faults in the software, including both initial and introduced faults at time t. b(t) represents the fault detection rate

per fault at time t, and m(t) represents the predicted number of faults detected by time t, which is the mean value function. The solution to the differential equation (1) can be expressed as follows:

$$m(t) = e^{-B(t)}\left[m_o + \int_{t_o}^{t} a(y)b(y)e^{-B(y)}dy \right] \tag{5}$$

We have $B(t) = \int_{t_o}^{t} b(y)dy$ and $m(t_0) = m_0$, where t_0 represents the initial time of the debugging procedure. Several NHPP models that now exist can be regarded as specific instances of the overarching model described in equation (2).

I. Software reliability growth model with change point

The NHPP SRGM, which combines imperfect debugging with a change-point problem, is based on the following assumptions:

- When faults that have been detected are removed at time t, there exists the potential for the introduction of additional faults at a rate denoted as $\alpha(t)$.

$$\alpha(t) = \begin{cases} \alpha_1, & 0 \le t \le y \\ \alpha_2, & t > y \end{cases}$$

- We express the rate of fault detection as a step function

$$b(t) = \begin{cases} b_1, & 0 \le t \le y \\ b_2, & t > y \end{cases}$$

- This study proposes an NHPP model to analyze the fault detection phenomenon in software systems.

Continuous monitoring of the testing strategy and resource allocation is possible throughout the fault detection process. It may be more justifiable to reassess the provided change point (y). Based on these assumptions, we can derive the new set of differential equations to generate the new mean value function.

$$\frac{dm(t)}{dt} = b(t)\left[a(t) - m(t)\right] \tag{6}$$

$$\frac{da(t)}{dt} = \alpha(t)\frac{dm(t)}{dt}$$

$$a(0) = a, m(0) = 0$$

The mean value function of the model is given as follows:

$$m(t) = \begin{cases} \frac{a}{1-\alpha_1}\left[1 - e^{-(1-\alpha_1)b_1 t}\right], & 0 \le t \le y \\ \frac{a}{1-\alpha_2}\left[1 - e^{-(1-\alpha_1)b_1 y - (1-\alpha_2)b_2 (t-y)}\right] + \frac{m(y)(\alpha_1 - \alpha_2)}{1-\alpha_2}, & t > y \end{cases} \tag{7}$$

III. Proposed model

In this part, we present a software reliability growth model that integrates flawed debugging practices with change-point and testing endeavors. Commencing with the imperative nature of

software reliability testing, we shall establish certain assumptions for the construction of our model. Inspection can identify a significant number of defects during the initial stage of the testing phase. Several factors, such as the efficiency of fault detection, the density of faults, the level of testing effort, and the rate of inspection, influence the pace of fault detection[17]. Subsequently, the rate of fault detection is contingent upon other supplementary characteristics, including the relationship between failures and faults, the factor of code expansion, the proficiency of test teams, the size of the program, and the testability of the software.

Assumptions of the proposed model

- According to the Non-Homogeneous Poisson Process (NHPP), the fault removal process is implemented.
- The software system may experience intermittent failures due to the presence of residual faults within the system.
- The average number of faults identified within the time interval (t, t + κt) by the present testing-effort expenditures is directly related to the average number of remaining defects in the system.
- A generalized logistic TEF model represents the testing-effort consumption curve.

$$W(t) = \frac{A}{1+n.e^{-\beta t}}$$

- When faults that have been detected are removed at time t, there exists the potential for the introduction of additional faults at a rate denoted as $\alpha(t)$.

$$\alpha(t) = \begin{cases} \alpha_1, & 0 \le t \le y \\ \alpha_2, & t > y \end{cases}$$

- We express the rate of fault detection as a step function

$$b(t) = \begin{cases} b_1, & 0 \le t \le y \\ b_2, & t > y \end{cases}$$

A generalized logistic TEF, which incorporates the fault introduction rate and change point, can characterize the software reliability growth model as follows:

$$\frac{dm(t)}{dt} * \frac{1}{w(t)} = b(t) * \left(a - m(t)\right) \tag{8}$$

$$\frac{da(t)}{dt} = \alpha(t)\frac{dm(t)}{dt}$$

$$a(0) = a, m(0) = 0$$

$W(t)$ can defined as follows:

$$W(t) = \int_0^t w(y)dy$$

The mean value function of the model is given as follows:

$$m(t) = \begin{cases} \frac{a}{1-\alpha_1}\left[1 - e^{-(1-\alpha_1)b_1(W(t)-W(0))}\right], & 0 \le t \le y \\ \frac{a}{1-\alpha_2}\left[1 - e^{-(1-\alpha_1)b_1(W(y)-W(0))-(1-\alpha_2)b_2(W(t)-W(y))}\right] + \frac{m(y)(\alpha_1-\alpha_2)}{1-\alpha_2}, & t > y \end{cases} \tag{9}$$

Software reliability R(x/t) refers to the probability that no software problem will be detected during the time interval (t, t + x), where $t \geq 0$, $x > 0$.

$$R(x\,f\ \ t) = e^{-\left[m(t+x)-m(t)\right]} \tag{10}$$

IV. Numerical Description

The existing models in the literature have employed the maximum likelihood estimation technique for parameter estimations. This study employs a nonlinear least square estimation (LSE) approach. We have utilized two historical data sets to substantiate the performance and conducted a comparative analysis between the presented model and current models. It is possible to find statistical measures like the sum of the square of error (SSE), root mean square error (RMSE), and adjusted R-square. We will now provide you with formulas for statistical measures that assess the model's fit to the data.

Root mean squared error (RMSE):

$$MSE = \frac{1}{N-n}\sum_{i=1}^{N}(y_i - m^*(t_i))^2$$

$$RMSE = \sqrt{MSE} \tag{11}$$

adjusted R^2 (Adjusted R2):

$$Adjusted\ R^2 = 1 - \frac{(1-R)(N-1)}{N-M-1} \tag{12}$$

sum of squared error (SSE):

$$SSE = \sum_{i=M}^{N}(y_i - m^*(t_i))^2 \tag{13}$$

V. Result analysis

Table 1 presents the overview of the data sets [16]. Table 2 represents the software reliability growth model and its mean value function. The unknown parameters in the proposed model are a, b_1, and b_2, and the unknown parameters in the testing effort function are β, n, and A. We have examined the proposed model with the given dataset. The results obtained from the TEF are as follows: a = 56.367, β = 12.054, and n = 0.256. The rate at which faults are introduced before the change point α_1 is evaluated to be 0.2, whereas the rate at which faults are introduced after the change point α_2 is evaluated to be 0.5. The results of the LSE reveal that the values of a, b_1, and b_2 are respectively 210.7, 0.2536, and 0.443. Table 3 shows that our proposed model is more accurate in RMSE, Adjust R^2, and SSE values than existing models. The fitting comparison of all models for using the data set is graphically illustrated in Figure 1. Figure 1, it can be seen that the proposed model fits the actual data better than all other models. Figure 2 a) RMSE, (b) Adjust R^2, (c) SSE graphically shows that compared to all models and Figure 2 shows a better fit proposed model.

Table 1: *summary of the data set*

Time	cumulative failure	testing effort consumption
1	15	2.45
2	44	4.90
3	66	6.86
4	103	7.84
5	105	9.52
6	110	12.89
7	146	17.10
8	175	20.47
9	179	21.43
10	206	23.35
11	233	26.23
12	255	27.67
13	276	30.93
14	298	34.77
15	304	38.61
16	311	40.91
17	320	42.67
18	325	44.66
19	328	47.6

Table 2: *Software reliability growth model and their mean value function*

No.	Name of the Model	MVF
1.	Goel-Okumoto (GOM)	$m(t) = a(1-e^{-bt})$
2.	Delay S-shaped (DSSM)	$m(t) = a(1-(1+bt)e^{-bt})$
3.	PNZ (PNZM)	$m(t) = \frac{a}{1+\beta e^{-bt}}\left\{[1-e^{-bt}][1-\frac{\alpha}{b}]+\alpha at\right\}$
4.	PZ (PZM)	$m(t) = \frac{1}{1+\beta e^{-bt}}\left\{[c+a][1-e^{-bt}]-\frac{a}{b-\alpha}[e^{-\alpha t}-e^{-bt}]\right\}$
5.	Proposed (PM)	Equation (9)

Table 3: *Comparison criteria*

Model	Parameter estimation	RMSE	Adjust R2	SSE
Goel-Okumoto	$a = 760.5, b = 0.03227$	12.53	0.9851	2656
Delay S-shaped	$a = 374.1, b = 0.1978$	13.73	0.9857	3205
PNZ	$a = 53.49, \alpha = 0.005551,$ $\beta = 0.38, b = 0.9413$	14.52	0.9807	3160
PZ	$a = 423.3, \alpha = 1.1, \beta = 0.8845,$ $b = 0.1167, c = 0.98$	12.24	0.9863	2246
Proposed	$a = 210.7, b_1 = 0.2536, b_2 = 0.443,$	6.1253	0.9889	1704

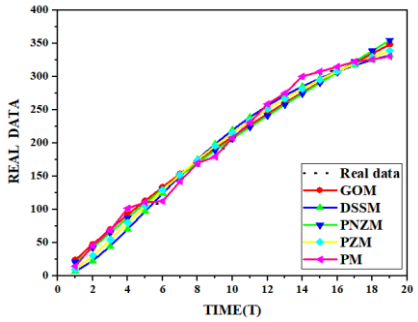
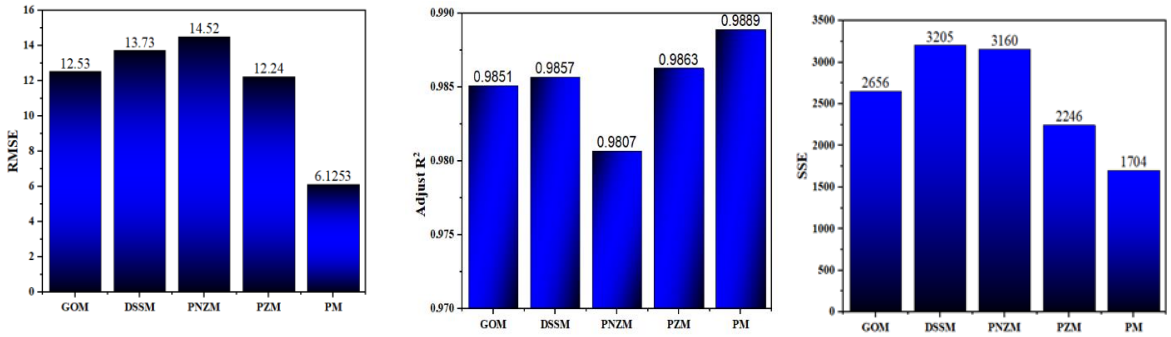


Figure 1: MVF curve for various model



(a) (b) (c)

Figure 2: *(a)RMSE, (b) Adjust R^2,(c) SSE*

VI. Cost model assumptions and formulation

Based on the above assumptions, we establish a cost function.

- During the first stages of the development process [18,19], there is an incurred cost for establishing the project. Designate this as C_0.
- The cost of testing is directly proportional to the duration of the testing process. Define $E_1(T)$ as the anticipated cost of testing.

$$\text{Therefore, } E_1(T) = C_1 T^{\delta}$$

Here δ represents the discount rate.

- During debugging phase, the fault removal cost is proportional to the total time spent on debugging. Let $E_2(T)$ represent the expected cost of reducing.

$$\text{Therefore, } E_2(T) = C_2 m(T) \lambda_y$$

Here, λ_y is the expected time for resolving each defect during the testing process.

- The cost of fixing faults is directly proportionate to the time needed for their fix during the warranty period. Let $E_3(T)$ denotes the projected cost of fixing all defects during the warranty duration.

$$\text{Therefore, } E_3(T) = C_3 \left[m(T_w) - m(T) \right] \lambda_w$$

Where, λ_w represents the expected repair time for each defect during the warranty period, and T_w denotes the duration of the warranty.

- Undiscovered faults that impact the software's reliability always result in a penalty cost after its release. Let E4 (T) denote the risk cost.

$$\text{Therefore, } E_4(T) = C_4 \left[1 - R(x/T_w) \right]$$

Let E(T) be the total software expenditure. E(T) can be expressed as:

$$E(T) = C_0 + C_1 T^{\delta} + C_2 m(T) \lambda_y + C_3 \left[m(T_w) - m(T) \right] \lambda_w + C_4 \left[1 - R(x/T_w) \right] \quad (14)$$

Where, C_1, C_2, C_3 and C_4 as weights for the following: the cost of testing, the cost of error removal during testing, the cost of error removal throughout the warranty term, and the penalty for software failure.

The cost coefficients $C_0, C_1, \cdots$, etc. are often established based on prior knowledge and the current state of the market. In this study, we can take C_0 = \$100, C_1 = \$150, C_2 = \$75, C_3 = \$200, C_4 = \$1000, λ_y = 0.1, λ_w = \$0.5, x = 0.04, T_w = 40, and δ = 0.9. We can determine 15 times with a development cost of \$172601.327 and reliability is 0.916. Figure 3 shows the minimum cost and this time reliability.

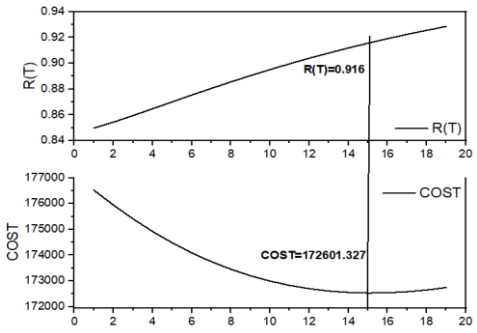


Figure 3 shows reliability and cost analysis with time

VII. Conclusion

This study introduces a novel change point software reliability growth model that incorporates the testing effort function, the NHPP framework, and imperfect debugging. In imperfect debugging, there exists a generalized logistic testing effort function and the impact of modification points.

We examine the new model and introduce the explicit mean value function. In addition, this model has been compared to several existing imprecise change point debugging models based on

the root mean squared error (RMSE), adjusted R^2, and sum square error (SSE) values on the data set. Quantitative findings indicate that the suggested model has a superior level of goodness-of-fit. This proposed model appears to be slightly more advanced, but imperfect debugging, testing effort, and change point impact results in a more robust property that accurately simulates the fluctuating fault detection rate.

This study additionally examines a software cost model that integrates warranty expenses, risk costs, and mistake removal costs. This approach facilitates the determination of the optimal testing cessation point for the product, reducing anticipated total expenses, and ensuring adherence to the software reliability growth model.

By incorporating this component, the model can provide a more accurate estimation of software reliability, hence improving its practical usefulness. In our future work, we will also engage in parameter estimation using the maximum likelihood estimation technique. This strategy will offer a strong and statistically reliable method to estimate the model parameters. This property accurately represents the real-world consequences of the testing process.

References

[1] Shrivastava, Avinash K., Ruchi Sharma, and Hoang Pham. "Software reliability and cost models with warranty and life cycle." Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 237.1 (2023): 166-179.

[2] Saraf, Iqra, and Javaid Iqbal. "Generalized multi-release modelling of software reliability growth models from the perspective of two types of imperfect debugging and change point." Quality and Reliability Engineering International 35.7 (2019): 2358-2370.

[3] Jain, M., S. C. Agrawal, and Priyanka Agarwal. "TESTING AND OPERATIONAL RELIABILITY FOR A DISTRIBUTED SOFTWARE SYSTEM WITH CORRECTION LAG CONSTRAINTS." Advances in Modeling, Optimization and Computing (2011): 630.

[4] Behera, Anup Kumar, and Priyanka Agarwal. "An SRGM using Fault Removal Efficiency and Correction Lag Function." International Journal of Reliability, Quality and Safety Engineering (2024).

[5] Rani, Sulekha, et al. "A software reliability growth model considering testing coverage subject to field environment." International Journal of Mathematics in Operational Research 18.2 (2021): 145-153.

[6] Shyur, Huan-Jyh. "A stochastic software reliability model with imperfect-debugging and change-point." Journal of Systems and Software 66.2 (2003): 135-141.

[7] Li, Qiuying, and Hoang Pham. "A testing-coverage software reliability model considering fault removal efficiency and error generation." PloS one 12.7 (2017): e0181524.

[8] Shrivastava, Avinash K., and P. K. Kapur. "Change-points-based software scheduling." Quality and Reliability Engineering International 37.8 (2021): 3282-3296.

[9] Sharma, Dinesh K., Deepak Kumar, and Shubhra Gautam. "Flexible software reliability growth models under imperfect debugging and error generation using learning function." *Journal of Management Information and Decision Sciences* 21.1 (2018): 1-12.

[10] Chatterjee, Subhashis, Deepjyoti Saha, and Akhilesh Sharma. "Multi-upgradation software reliability growth model with dependency of faults under change point and imperfect debugging." Journal of Software: Evolution and Process 33.6 (2021): e2344.

[11] Panwar, Saurabh, et al. "Software reliability prediction and release time management with coverage." International Journal of Quality & Reliability Management 39.3 (2022): 741-761.

[12] Yamada, Shigeru, Mitsuru Ohba, and Shunji Osaki. "S-shaped reliability growth modeling for software error detection." IEEE Transactions on reliability 32.5 (1983): 475-484.

[13] Tohma, Yoshihiro, et al. "The estimation of parameters of the hypergeometric distribution and its application to the software reliability growth model." IEEE Transactions on Software Engineering 17.5 (1991): 483.

[14] Goel, Amrit L., and Kazu Okumoto. "Time-dependent error-detection rate model for

software reliability and other performance measures." IEEE transactions on Reliability 28.3 (1979): 206-211.

[15] Rafi, Shaik Mohammad, and Shaheda Akthar. "Software reliability growth model with logistic-exponential testing effort function and analysis of software release policy." Proceedings of international conference on advances in computer science. 2010.

[16] Ohba, Mitsuru. "Software reliability analysis models." IBM Journal of research and Development 28.4 (1984): 428-443.

[17] Manjula, Taduru, Madhu Jain, and T. R. Gulati. "Cost optimization of a software reliability growth model with imperfect debugging and a fault reduction factor." ANZIAM Journal 55 (2013): C182-C196.

[18] Samal, Umashankar, and Ajay Kumar. "A software reliability model incorporating fault removal efficiency and it's release policy." Computational Statistics (2023): 1-19.

[19] Jain, M., P. Agarwal, and R. Solanki. "NHPP-based SRGM using time-dependent fault reduction factors (FRF) and Gompertz TEF." Decision Analytics Applications in Industry (2020): 81-89.

A COMPREHENSIVE ANALYSIS OF JUCHEZ DISTRIBUTION: EXPLORING STRUCTURAL PROPERTIES AND APPLICATIONS

Rashid A. Ganaie[1], R. Shenbagaraja[2], T. Vivekanandan[3] and Aafaq A. Rather[4,*], Manzoor A. Khanday[5], Asgar Ali[6]

•

[1]Department of Statistics, Annamalai University, Tamil Nadu-608002, India
[2] Department of Statistics, Dr. M. G. R. Government Arts and Science College for Women, Villupuram, Tamil Nadu-605602, India
[3]Department of Science and Humanities, Vel Tech Multi Tech Dr. Rangarajan Dr. Sagunthala Engineering College, Chennai, Tamil Nadu-600062, India
[4,*]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[5]Department of Statistics, Lovely Professional University, Phagwara, Punjab, India
[6]Department of Statistics, K. K. Das College, Garia, Kolkatta-700084, India
[1]rashidau7745@gmail.com, [2]statraja@gmail.com, [3]mtvivek2017@gmail.com, [4'*]aafaq7741@gmail.com, [5]manzoorstat@gmail.com, [6]ali.si2006@gmail.com

Abstract

This article introduces an innovative extension of the Juchez distribution, referred to as the length-biased Juchez distribution. This distribution, a specific instance of the broader weighted distribution, is thoroughly explored in terms of mathematical and statistical properties. Parameter estimation is accomplished through the application of maximum likelihood estimation techniques. To highlight the practical significance of this new distribution, a comprehensive analysis is conducted using two real-life lifetime datasets. The findings underscore the relevance and applicability of the proposed distribution in modeling and analyzing diverse datasets.

Keywords: Length biased distribution, Juchez distribution, Order statistics, Survival analysis, Maximum likelihood estimation.

1. Introduction

In statistics, the theory of weighted probability distributions has retained a reputed and prominent place because it provides a new shape to the existing classical distribution by introducing an additional parameter to it. This additional parameter brings more superiority and flexibility to a class of distribution functions and it should be very significant from data analysis point of view. This extra parameter can be introduced through various techniques. One of such technique is of weighted technique. The idea of weighted distribution was propounded firstly by Fisher [10] to study how the method of ascertainment can influence the form of distribution of recorded observation. Later, Rao [16] developed this concept in a collective way in association with modeling

statistical data when usual practice of using standard distributions was found to be unsuitable. The theory of weighted distributions plays a dominant and tremendous practical role in probability, statistics and mathematics. The concept of weighted distribution provides an integrative conceptualization for model stipulation and data representation problems. The weighted distributions also provide a collective approach for correction of biases that exists in unequally weighted sample data. The weighted distribution reduces to length biased distribution when the weight function considers only the length of units of interest. Length biased distribution have been applied in various biomedical areas such as survival analysis, family history, reliability analysis, clinical trials, intermediate events and population studies were a proper sampling frame is absent. In such situation items are sampled at a rate proportional to their lengths so that the larger value could be sampled with higher probability.

Many authors have described and developed some important length biased probability models along with their illustrations in various fields. Al-Omari and Alanzi [6] presented inverse length biased Maxwell distribution and obtain its statistical inference with an application. Andure (Yawale) and Ade [7] presented the new length biased Hamza distribution with statistical properties and applications. Saghir, Tazeem and Ahmad [24] discussed on the length biased weighted exponentiated inverted weibull distribution and introduce its necessary properties. Mustafa and Khan [12] developed the length biased power hazard rate distribution with some properties and applications. Abdullah et al. [2] presented the size biased Lomax distribution with applications. Alidamat and Al-Omari [5] described the length biased two parameter Mirra distribution with application to engineering data. Ahajeeth et al. [4] proposed the area biased Amarendra distribution with its application to model lifetime data. Sanat [25] derived the beta-length biased Pareto distribution and its properties. Ganaie and Rajagopalan [11] presented the length biased power quasi Lindley distribution with properties and applications of lifetime data. Abd-Elfattah et al. [1] studied the length biased Burr-XII distribution with properties and application. Reyad et al. [22] proposed the length biased weighted Erlang distribution. Rather and Subramanian [16] discussed on length biased Sushila distribution with properties and applications. Nanuwong and Bodhisuwan [14] executed the length biased beta-Pareto distribution with its structural properties and application. Saghir and Khadim [23] presented the mathematical properties of length biased weighted Maxwell distribution. Rather et al [21] enriched the research by offering a comprehensive overview, perspectives, and characterizations of a new size biased Ailamujia distribution with applications in engineering and medical science which shows more flexibility than classical distributions. Rather and Subramanian [19] discussed the characterization and estimation of length biased weighted generalized uniform distribution. Rather and Subramanian [18], obtained length biased sushila distribution with properties and its applications. Shenbagaraja et al. [26], discussed length biased Garima distribution. Rather and Subramanian [20], conducted a thorough examination of the length biased erlang-truncated exponential distribution with life time data. Subramanian and Rather [27], obtained a new extension of Shanker distribution with real life data. Rather and Ozel [17], explored a new length biased power lindley distribution with properties and its applications. Recently, Mustafa and Khan [13] developed the length biased powered inverse Rayleigh distribution with applications.

Juchez probability distribution is a recently developed one parametric continuous lifetime distribution studied by Echebiri and Mbegbu [9]. Some of statistical features including median, mode, mean, moments, coefficient of variation, skewness, kurtosis, mean residual life function, hazard function, bonferroni and lorenz curves, order statistics, stochastic ordering and Renyi entropy have been presented. Furthermore, its parameter has been estimated by using the maximum likelihood estimation.

2. Length Biased Juchez Distribution

The probability density function of Juchez distribution is given by

$$f(x;\theta)=\frac{\theta^4}{\theta^3+\theta^2+6}\left(1+x+x^3\right)e^{-\theta x};\ x>0,\ \theta>0 \tag{1}$$

and the cumulative distribution function of Juchez distribution is given by

$$F(x;\theta)=1-\left(1+\frac{\theta x[\theta^2+\theta^2x^2+3\theta x+6]}{\theta^3+\theta^2+6}\right)e^{-\theta x};\quad x>0,\ \theta>0 \tag{2}$$

Let X be the random variable represents non-negative condition with probability density function $f(x)$. Let its non-negative weight function be $w(x)$, then the probability density function of weighted random variable X_w is given by

$$f_w(x)=\frac{w(x)f(x)}{E(w(x))},\quad x>0.$$

When the non-negative weight function be $w(x)$ and $E(w(x))=\int w(x)f(x)dx<\infty$.

Depending upon the different weighted functions $w(x)$ obviously when $w(x) = x^c$, proposed distribution is known as weighted distribution. In this paper, we have to be considered the length biased version of Juchez distribution termed as length biased Juchez distribution. So the weight function considered at $w(x) = x$ by taking weight parameter c is 1 in weights x^c resulting distribution is known as length biased distribution and its probability density function given by

$$f_l(x)=\frac{x\,f(x)}{E(x)} \tag{3}$$

Where $E(x)=\int_0^\infty x\,f(x)dx$

$$E(x)=\frac{\left(\theta^3+2\theta^2+24\right)}{\theta(\theta^3+\theta^2+6)} \tag{4}$$

Now by using the equations (1) and (4) in equation (3), we will get required probability density function of length biased Juchez distribution as

$$f_l(x)=\frac{x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x} \tag{5}$$

and the cumulative distribution function of length biased Juchez distribution can be determined as

$$F_l(x)=\int_0^x f_l(x)dx$$

$$=\int_0^x\frac{x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}dx$$

$$=\frac{1}{\left(\theta^3+2\theta^2+24\right)}\int_0^x x\theta^5\left(1+x+x^3\right)e^{-\theta x}dx$$

$$= \frac{1}{\left(\theta^3 + 2\theta^2 + 24\right)}\left(\theta^5 \int_0^x x e^{-\theta x} dx + \theta^5 \int_0^x x^2 e^{-\theta x} dx + \theta^5 \int_0^x x^4 e^{-\theta x} dx\right) \tag{6}$$

$$\text{Put } \theta x = t \quad \Rightarrow \theta dx = dt \quad \Rightarrow dx = \frac{dt}{\theta}, \qquad \text{Also} \quad x = \frac{t}{\theta}$$

$$\text{When } x \to x,\ t \to \theta x \quad \text{and when} \quad x \to 0,\ t \to 0$$

After the simplification of equation (6), we will determine cumulative distribution function of length biased Juchez distribution as

$$F_l(x) = \frac{1}{\left(\theta^3 + 2\theta^2 + 24\right)}\left(\theta^3 \gamma(2,\ \theta x) + \theta^2 \gamma(3,\ \theta x) + \gamma(5,\ \theta x)\right) \tag{7}$$

Figure 1 and figure 2, shows the graphical representation of the pdf and cdf plot and has been R-core version [15] for this.

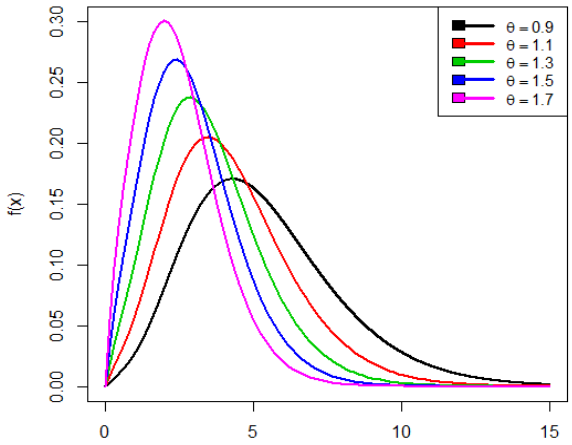


Figure 1: *Pdf plot of length biased Juchez distribution*

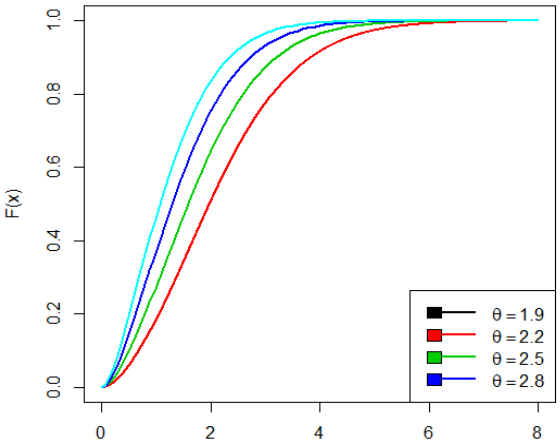


Figure 2: *Cdf plot of length biased Juchez distribution*

3. Survival Analysis

In this section, we will derive the survival function, hazard rate function, reverse hazard rate function and Mills ratio of the proposed length biased Juchez distribution. The survival or reliability function of length biased Juchez distribution can be obtained as

$$S(x) = 1 - F_l(x)$$

$$= 1 - \frac{1}{\left(\theta^3 + 2\theta^2 + 24\right)}\left(\theta^3 \gamma(2,\ \theta x) + \theta^2 \gamma(3,\ \theta x) + \gamma(5,\ \theta x)\right) \tag{8}$$

The hazard function is also known as hazard rate or failure rate or force of mortality and is given by

$$h(x) = \frac{f_l(x)}{1 - F_l(x)}$$

$$= \frac{x\theta^5 (1 + x + x^3)\ e^{-\theta x}}{(\theta^3 + 2\theta^2 + 24) - (\theta^3 \gamma(2,\ \theta x) + \theta^2 \gamma(3,\ \theta x) + \gamma(5,\ \theta x))} \tag{9}$$

The reverse hazard rate function is given by

$$h_r(x) = \frac{f_l(x)}{F_l(x)}$$

$$= \frac{x\theta^5(1+x+x^3)\ e^{-\theta x}}{(\theta^3\gamma(2,\ \theta x)+\theta^2\gamma(3,\ \theta x)+\gamma(5,\ \theta x))} \tag{10}$$

The Mills Ratio is given by

$$M.R = \frac{1}{h_r(x)} = \frac{(\theta^3\gamma(2,\ \theta x)+\theta^2\gamma(3,\ \theta x)+\gamma(5,\ \theta x))}{x\theta^5(1+x+x^3)\ e^{-\theta x}} \tag{11}$$

Figure 3 and figure 4, shows the graphical representation of the survival function and cdf plot.

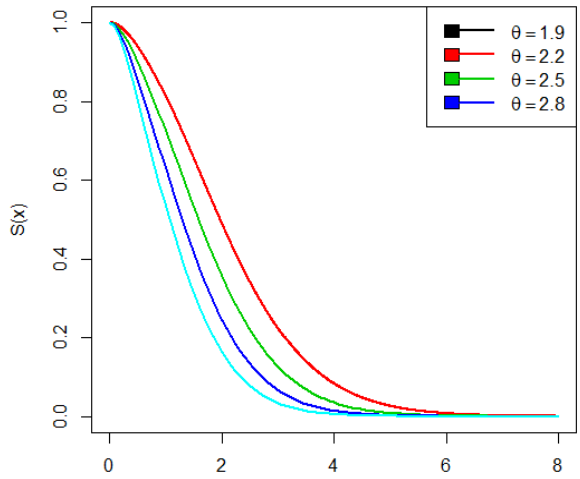


Figure 3: *Survival function of length biased Juchez distribution*

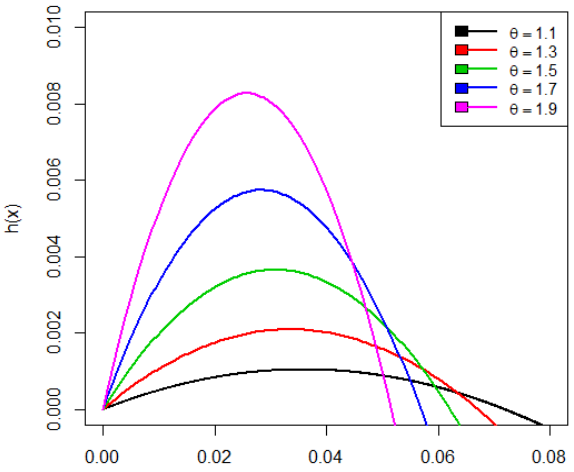


Figure 4: *Hazard function of length biased Juchez distribution*

4. Order Statistics

Order statistics is a very significant concept in statistical sciences and has wide range of applications in modeling auctions, insurance policies, car races, optimizing production processes and estimating parameters of distributions. Consider X(1), X(2),…, X(n) be the order statistics of a random sample X1, X2,…, Xn from a continuous population with probability density function fx (x) and cumulative distribution function FX(x), then the probability density function of rth order statistics X(r) is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!} f_X(x)\left(F_X(x)\right)^{r-1}\left(1-F_X(x)\right)^{n-r} \tag{12}$$

By using the equations (5) and (7) in equation (12), we will obtain the probability density function of r^{th} order statistics $X_{(r)}$ of length biased Juchez distribution as

$$f_{x(r)}(x)=\frac{n!}{(r-1)!(n-r)!}\left(\frac{x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}\right)$$
$$\times\left(\frac{1}{\left(\theta^3+2\theta^2+24\right)}\left(\theta^3\gamma(2,\theta x)+\theta^2\gamma(3,\theta x)+\gamma(5,\theta x)\right)\right)^{r-1}$$
$$\times\left(1-\frac{1}{\left(\theta^3+2\theta^2+24\right)}\left(\theta^3\gamma(2,\theta x)+\theta^2\gamma(3,\theta x)+\gamma(5,\theta x)\right)\right)^{n-r} \quad (13)$$

Therefore, the probability density function of higher order statistic $X_{(n)}$ of length biased Juchez distribution can be obtained as

$$f_{x(n)}(x)=\frac{n\,x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}$$
$$\times\left(\frac{1}{\left(\theta^3+2\theta^2+24\right)}\left(\theta^3\gamma(2,\theta x)+\theta^2\gamma(3,\theta x)+\gamma(5,\theta x)\right)\right)^{n-1} \quad (14)$$

and the probability density function of first order statistic $X_{(1)}$ of length biased Juchez distribution can be obtained as

$$f_{x(1)}(x)=\frac{n\,x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}$$
$$\times\left(1-\frac{1}{\left(\theta^3+2\theta^2+24\right)}\left(\theta^3\gamma(2,\theta x)+\theta^2\gamma(3,\theta x)+\gamma(5,\theta x)\right)\right)^{n-1} \quad (15)$$

5. Likelihood Ratio Test

Let the random sample X1, X2,…., Xn of size n drawn from the length biased Juchez distribution. To analyze its significance, the hypothesis is to be tested

$$H_o: f(x)=f(x;\theta) \quad \text{against} \quad H_1: f(x)=f_l(x;\theta) \quad (16)$$

In order to determine, whether the random sample of size n comes from Juchez distribution or length biased Juchez distribution, the following test statistic is employed

$$\Delta=\frac{L_1}{L_o}=\prod_{i=1}^{n}\frac{f_l(x;\theta)}{f(x;\theta)} \quad (17)$$

$$=\frac{L_1}{L_o}=\prod_{i=1}^{n}\left(\frac{x_i\,\theta(\theta^3+\theta^2+6)}{\left(\theta^3+2\theta^2+24\right)}\right) \quad (18)$$

$$= \frac{L_1}{L_o} = \left(\frac{\theta(\theta^3 + \theta^2 + 6)}{\left(\theta^3 + 2\theta^2 + 24\right)} \right)^n \prod_{i=1}^{n} x_i \tag{19}$$

We should refuse to retain the null hypothesis, if

$$\Delta = \left(\frac{\theta(\theta^3 + \theta^2 + 6)}{\left(\theta^3 + 2\theta^2 + 24\right)} \right)^n \prod_{i=1}^{n} x_i > k \tag{20}$$

Equivalently, we should also refuse to retain the null hypothesis, where

$$\Delta = \prod_{i=1}^{n} x_i > k \left(\frac{\left(\theta^3 + 2\theta^2 + 24\right)}{\theta(\theta^3 + \theta^2 + 6)} \right)^n \tag{21}$$

$$\Delta = \prod_{i=1}^{n} x_i > k^*, \text{where } k^* = k \left(\frac{\theta^3 + 2\theta^2 + 24}{\theta(\theta^3 + \theta^2 + 6)} \right)^n \tag{22}$$

such that $p\left(\Delta > k^*\right) = \alpha$, where α is the level of significance.

6. Structural Properties

In this section, we will derive several statistical properties of length biased Juchez distribution which include moments, harmonic mean, moment generating function and characteristic function.

6.1 Moments

Let X be the random variable following length biased Juchez distribution with parameter θ, then the r^{th} order moment $E(X^r)$ of introduced distribution can be obtained as

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_l(x) dx \tag{23}$$

$$= \mu_r' = \int_0^\infty x^r \frac{x\theta^5}{\left(\theta^3 + 2\theta^2 + 24\right)} \left(1 + x + x^3\right) e^{-\theta x} dx \tag{24}$$

$$= \mu_r' = \int_0^\infty \frac{x^{r+1}\theta^5}{\left(\theta^3 + 2\theta^2 + 24\right)} \left(1 + x + x^3\right) e^{-\theta x} dx \tag{25}$$

$$= \mu_r' = \frac{\theta^5}{\left(\theta^3 + 2\theta^2 + 24\right)} \int_0^\infty x^{r+1} \left(1 + x + x^3\right) e^{-\theta x} dx \tag{26}$$

$$= \mu_r' = \frac{\theta^5}{\left(\theta^3 + 2\theta^2 + 24\right)} \left(\int_0^\infty x^{(r+2)-1} e^{-\theta x} dx + \int_0^\infty x^{(r+3)-1} e^{-\theta x} dx + \int_0^\infty x^{(r+5)-1} e^{-\theta x} dx \right) \tag{27}$$

After the simplification of above equation, we obtain

$$E(X^r) = \mu_r' = \frac{\theta^3\Gamma(r+2)+\theta^2\Gamma(r+3)+\Gamma(r+5)}{\theta^r(\theta^3+2\theta^2+24)} \tag{28}$$

Now putting r = 1, 2, 3 and 4 in equation (28), we will obtain the first four moments of length biased Juchez distribution as

$$E(X) = \mu_1' = \frac{2\theta^3+6\theta^2+120}{\theta(\theta^3+2\theta^2+24)} \tag{29}$$

$$E(X^2) = \mu_2' = \frac{6\theta^3+24\theta^2+720}{\theta^2(\theta^3+2\theta^2+24)} \tag{30}$$

$$E(X^3) = \mu_3' = \frac{24\theta^3+120\theta^2+5040}{\theta^3(\theta^3+2\theta^2+24)} \tag{31}$$

$$E(X^4) = \mu_4' = \frac{120\theta^3+720\theta^2+40320}{\theta^4(\theta^3+2\theta^2+24)} \tag{32}$$

$$\text{Variance} = \frac{6\theta^3+24\theta^2+720}{\theta^2(\theta^3+2\theta^2+24)} - \left(\frac{2\theta^3+6\theta^2+120}{\theta(\theta^3+2\theta^2+24)}\right)^2 \tag{33}$$

$$S.D(\sigma) = \sqrt{\left(\frac{6\theta^3+24\theta^2+720}{\theta^2(\theta^3+2\theta^2+24)} - \left(\frac{2\theta^3+6\theta^2+120}{\theta(\theta^3+2\theta^2+24)}\right)^2\right)} \tag{34}$$

6.2 Harmonic mean

The harmonic mean for the executed length biased Juchez distribution can be determined as

$$H.M = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_l(x)dx \tag{35}$$

$$= \int_0^\infty \frac{1}{x} \frac{x\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{36}$$

$$= \int_0^\infty \frac{\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(1+x+x^3\right)e^{-\theta x}dx \tag{37}$$

$$= \frac{\theta^5}{\left(\theta^3+2\theta^2+24\right)}\int_0^\infty \left(1+x+x^3\right)e^{-\theta x}dx \tag{38}$$

$$= \frac{\theta^5}{\left(\theta^3+2\theta^2+24\right)}\left(\int_0^\infty x^{(2)-2}e^{-\theta x}dx + \int_0^\infty x^{(2)-1}e^{-\theta x}dx + \int_0^\infty x^{(4)-1}e^{-\theta x}dx\right) \tag{39}$$

After the simplification of above equation, we obtain

$$H.M = \frac{\theta(\theta^2+\theta^2+6)}{\left(\theta^3+2\theta^2+24\right)} \tag{40}$$

6.3 Moment generating function and characteristic function

Let X be the random variable following length biased Juchez distribution with parameter θ, then the moment generating function of proposed distribution can be obtained as

$$M_X(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx} f_l(x)dx \tag{41}$$

Using Taylor's series, we obtain

$$M_X(t) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + \ldots\right) f_l(x)dx \tag{42}$$

$$= \int_0^\infty \sum_{j=0}^\infty \frac{t^j}{j!} x^j f_l(x)dx \tag{43}$$

$$= \sum_{j=0}^\infty \frac{t^j}{j!} \mu_j'$$

$$= \sum_{j=0}^\infty \frac{t^j}{j!} \left(\frac{\theta^3\Gamma(j+2) + \theta^2\Gamma(j+3) + \Gamma(j+5)}{\theta^j(\theta^3 + 2\theta^2 + 24)}\right) \tag{44}$$

$$= \frac{1}{(\theta^3 + 2\theta^2 + 24)} \sum_{j=0}^\infty \frac{t^j}{j!\theta^j} \left(\theta^3\Gamma(j+2) + \theta^2\Gamma(j+3) + \Gamma(j+5)\right) \tag{45}$$

Similarly, the characteristic function of length biased Juchez distribution can be obtained as

$$\varphi_x(t) = M_X(it)$$

$$M_X(it) = \frac{1}{(\theta^3 + 2\theta^2 + 24)} \sum_{j=0}^\infty \frac{it^j}{j!\theta^j} \left(\theta^3\Gamma(j+2) + \theta^2\Gamma(j+3) + \Gamma(j+5)\right) \tag{46}$$

7. Bonferroni and Lorenz Curves

The bonferroni and Lorenz curves also termed as income distribution curves or classical curves are frequently being applied to measure the distribution of inequality in income or poverty. The bonferroni and Lorenz curves can be executed as

$$B(p) = \frac{1}{p\mu_1'} \int_0^q x\, f_l(x)dx \tag{47}$$

$$L(p) = pB(p) = \frac{1}{\mu_1'} \int_0^q x\, f_l(x)dx \qquad \text{and} \quad q = F^{-1}(p)$$

$$\text{where } \mu_1' = \frac{\left(2\theta^3 + 6\theta^2 + 120\right)}{\theta(\theta^3 + 2\theta^2 + 24)}$$

$$B(p) = \frac{\theta(\theta^3 + 2\theta^2 + 24)}{p(2\theta^3 + 6\theta^2 + 120)} \int_0^q \frac{x^2\theta^5}{\left(\theta^3 + 2\theta^2 + 24\right)} \left(1 + x + x^3\right) e^{-\theta x} dx \tag{48}$$

$$= \frac{\theta^6}{p(2\theta^3 + 6\theta^2 + 120)} \int_0^q x^2 \left(1 + x + x^3\right) e^{-\theta x} dx \tag{49}$$

$$= \frac{\theta^6}{p(2\theta^3+6\theta^2+120)}\left(\int_0^q x^{(3)-1}e^{-\theta x}dx + \int_0^q x^{(4)-1}e^{-\theta x}dx + \int_0^q x^{(6)-1}e^{-\theta x}dx\right) \quad (50)$$

After simplification, we obtain

$$B(p) = \frac{\theta^6}{p(2\theta^3+6\theta^2+120)}\left(\gamma(3,\ \theta q)+\gamma(4,\ \theta q)+\gamma(6,\ \theta q)\right) \quad (51)$$

$$L(p) = \frac{\theta^6}{(2\theta^3+6\theta^2+120)}\left(\gamma(3,\ \theta q)+\gamma(4,\ \theta q)+\gamma(6,\ \theta q)\right) \quad (52)$$

8. Maximum Likelihood Estimation and Fisher's Information Matrix

In this section, we will discuss the technique of maximum likelihood estimation to estimate the parameters of length biased Juchez distribution. Consider $X_1, X_2, \ldots, X_n$ be a random sample of size n from length biased Juchez distribution, then the likelihood function can be defined as

$$L(x) = \prod_{i=1}^{n} f_l(x)$$

$$L(x) = \frac{\theta^{5n}}{(\theta^3+2\theta^2+24)^n}\prod_{i=1}^{n}\left(x_i\left(1+x_i+{x_i}^3\right)e^{-\theta x_i}\right) \quad (53)$$

The log likelihood function is given by

$$\log L = 5n\log\theta - n\log(\theta^3+2\theta^2+24) + \sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log\left(1+x_i+{x_i}^3\right) - \theta\sum_{i=1}^{n}x_i \quad (54)$$

Now differentiating log likelihood equation (54) with respect to parameter θ, we establish following normal equation

$$\frac{\partial \log L}{\partial\theta} = \frac{5n}{\theta} - n\left(\frac{\left(3\theta^2+4\theta\right)}{\left(\theta^3+2\theta^2+24\right)}\right) - \sum_{i=1}^{n} x_i = 0 \quad (55)$$

The above likelihood equation is too complicated to solve it algebraically. Therefore, we use numerical technique like Newton Raphson method for estimating the required parameter of proposed distribution.

In order to use the asymptotic normality results for determining the confidence interval. We have that if $(\hat{\beta}=\hat{\theta})$ denotes the MLE *of* $(\beta=\theta)$. We can execute the results as

$$\sqrt{n}(\hat{\beta}-\beta) \to N(0,\ I^{-1}(\beta))$$

where $I^{-1}(\beta)$ is Fisher's informatio n matrix.i.e .,

$$I(\beta) = -\frac{1}{n}\left(E\left(\frac{\partial^2 \log L}{\partial\theta^2}\right)\right)$$

Here, we see that

$$E\left(\frac{\partial^2 \log L}{\partial\theta^2}\right) = -\frac{5n}{\theta^2} - n\left(\frac{(\theta^3+2\theta^2+24)\ (6\theta+4)-(3\theta^2+4\theta)^2}{(\theta^3+2\theta^2+24)^2}\right)$$

Since β being unknown, we estimate $I^{-1}(\beta)$ by ($I^{-1}(\hat{\beta})$ and this can be used to obtain asymptotic confidence interval for θ.

9. Applications

In this section, we have fitted two real lifetime data sets in length biased Juchez distribution to determine its goodness of fit and then comparison has been developed in order to reveal that the length biased Juchez distribution provides a better result over Juchez, exponential and Lindley distributions. The two real lifetime data sets are given below as.

The following first real data set reported by Bader and priest [8] represents the strength measured in GPA for single carbon fibres and impregnated 1000-carbon fibre tows. Single fibres were tested under tension at guage length of 10mm with sample size (n = 63) and the data set is given below in table 1

Table 1: *Data regarding the strength of carbon fibres measured in GPA reported by Bader & Priest (1982)*

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		

The second real lifetime data set represents the waiting time (in minutes) of 65 dental patients, waiting before OPD (out Patient Diagnosis) at Halibet hospital, Asmara, from 25th to 29th December, 2017 available in the master thesis of Abebe [3] and the data set is given below in table 2.

Table 2: *Data regarding waiting time (in minutes) of 65 dental patients waiting before OPD (out Patient Diagnosis)*

2(5)	6	7	8(3)	9	10	11	12(2)	13	14(4)
15	16	17(2)	18(2)	19	20(3)	22	23(2)	26	27
28	29(2)	30(3)	31	32	33	35	36(2)	37(2)	40(2)
41(2)	42	43	44	46	47	49	52	53	55
56	58	90							

To determine the model comparison criterions along with the estimation of unknown parameters, the technique of R software is used. In order to compare the performance of length biased Juchez distribution over Juchez, exponential and Lindley distributions, we consider the criterions like *AIC* (Akaike Information Criterion), *BIC* (Bayesian Information Criterion), *AICC* (Akaike Information Criterion Corrected), *CAIC* (Consistent Akaike Information Criterion), Shannon's entropy *H(X)* and *-2logL*. The better distribution is which corresponds to lesser values of *AIC*, *BIC*, *AICC*, *CAIC*, *H(X)* and -2logL. For determining the criterions AIC, BIC, AICC, CAIC, H(X) and -2logL given below following formulas are used.

$AIC = 2k - 2\log L,$ $BIC = k\log n - 2\log L,$ $AICC = AIC + \frac{2k(k+1)}{n-k-1}$

$CAIC = -2\log L + \frac{2kn}{n-k-1}$ and $H(X) = -\frac{2\log L}{n}$

Where n is the sample size, k is number of parameters in statistical model and $-2\log L$ is the maximized value of log-likelihood function under the considered model. Table 3 shows the parameter and standard error values and table 4 shows the comparison of distributions.

Table 3: *Shows MLE and S.E Estimates for the data set 1 and data set 2*

Data sets	Distributions	MLE	S.E
1	Length Biased Juchez	$\hat{\theta} = 1.45059991$	$\hat{\theta} = 0.07765314$
	Juchez	$\hat{\theta} = 1.07320917$	$\hat{\theta} = 0.06273572$
	Exponential	$\hat{\theta} = 0.32687301$	$\hat{\theta} = 0.04118174$
	Lindley	$\hat{\theta} = 0.53923226$	$\hat{\theta} = 0.04958387$
2	Length Biased Juchez	$\hat{\theta} = 2.4289412$	$\hat{\theta} = 0.1590837$
	Juchez	$\hat{\theta} = 1.6783037$	$\hat{\theta} = 0.1224368$
	Exponential	$\hat{\theta} = 0.6615390$	$\hat{\theta} = 0.1008835$
	Lindley	$\hat{\theta} = 0.9934028$	$\hat{\theta} = 0.1144634$

Table 4: *Shows Comparison and Performance of fitted distributions*

Data sets	Distributions	-2logL	AIC	BIC	AICC	CAIC	H(X)
1	Length Biased Juchez	186.03	188.03	190.1731	188.0955	188.0955	2.9528
	Juchez	211.9434	213.9434	216.0865	214.0089	214.0089	3.3641
	Exponential	266.8915	268.8915	271.0347	268.9570	268.9570	4.2363
	Lindley	242.7153	244.7153	246.8584	244.7808	244.7808	3.8526
2	Length Biased Juchez	100.0493	102.0493	103.8105	102.1127	102.1127	1.5392
	Juchez	118.4688	120.4688	122.23	120.5322	120.5322	1.8225
	Exponential	121.5341	123.5341	125.2953	123.5975	123.5975	1.8697
	Lindley	114.5096	116.5096	118.2708	116.5730	116.5730	1.7616

From results given above in table 4, it has been clearly realized and observed that the length biased Juchez distribution has lesser *AIC*, *BIC*, *AICC*, *CAIC*, *H(X)* and *-2logL* values as compared to the

Juchez, exponential and Lindley distributions. Hence, it can be concluded that the length biased Juchez distribution provides a better fit over Juchez, exponential and Lindley distributions.

10. Conclusion

In the present study, we have developed a new class of Juchez distribution termed as length biased Juchez distribution. The proposed new distribution is executed by using the length biased technique to its classical distribution. Its various structural properties those include moments, shape of the pdf and cdf, harmonic mean, order statistics, survival function, hazard rate function, reverse hazard function, moment generating function, bonferroni and Lorenz curves have been derived. The parameter of proposed new distribution is estimated by using the maximum likelihood estimation. Finally, a new distribution has been examined and analyzed with two real data sets to demonstrate its superiority and flexibility. Hence, it is revealed from the results that the proposed length biased Juchez distribution leads to a better fit over Juchez, exponential and Lindley distributions.

References

[1] Abd-Elfattah., A. M., Mahdy, M. and Ismail, G. (2021). Length biased Burr-XII distribution: Properties and Application, *International Journal of Sciences: Basic and Applied Research (*IJSBAR), 56(1), 222-244.

[2] Abdullah., M., Iqbal., Z., Ali., A., Zakria M., and Ahmad, M. (2016). Size biased Lomax distribution, *Journal of Statistics,* 23, 32-49.

[3] Abebe, B. (2018). Discretization of some continuous distribution and their Applications, Master thesis submitted to Department of statistics, College of Science, Eritrea Institute of technology, Asmara, Eritrea.

[4] Ahajeeth., M., Mohiuddin, M. and Kannan, R. (2021). The area biased Amarendra distribution with its application to model lifetime data, *International Journal of Statistics and Reliability Engineering*, 8(3), 428-438.

[5] Alidamat A. J. and Al-Omari, A. I. (2021). The extended length biased two parameter Mirra distribution with an application to engineering data, *Advanced Mathematical Models & Applications*, 6(2), 113-127.

[6] Al-Omari, A. I. and. Alanzi, A. R. A. (2021). Inverse length biased Maxwell distribution: Statistical inference with an application, *Computer Systems Science & Engineering*, 39(1), 147-164.

[7] Andure N. W., (Yawale) and Ade, R. B. (2021). The new length biased Hamza distribution: statistical properties & applications, *International Journal of Statistics and Reliability Engineering,* 8(1), 63-68.

[8] Badar, M. G. and Priest, A. M. (1982). Statistical aspects of fiber and bundle strength in hybrid composites. In T. Hayashi, K. Kawata & S. Umekawa (Eds.), *Progress in Science and Engineering Composites. ICCM-IV, Tokyo*, 1129-1136.

[9] Echebiri, U. V. and Mbegbu, J. L. (2022). Juchez probability distribution: Properties and Applications, *Asian Journal of Probability and Statistics*, 20(2), 56-71.

[10] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, 6, 13-25.

[11] Ganaie, R. A. and Rajagopalan, V. (2021). A New extension of Power quasi Lindley distribution with Properties and Applications of the Life time Data, *International Journal of Statistics and Reliability Engineering*, 8(1), 171-183.

[12] Mustafa A. and Khan, M. I. (2022). The length-biased powered inverse Rayleigh distribution with applications, *J. Appl. Math. & Informatics*, 40(1-2), 1-13.

[13] Mustafa, A. and Khan, M. I. (2022). The length-biased power hazard rate distribution: some properties and applications, *STATISTICS IN TRANSITION new series*, 23(2), 1-16.

[14] Nanuwong, N. and Bodhisuwan, W. (2014). Length biased beta-Pareto distribution and its structural properties with application, *Journal of Mathematics and Statistics*, 10(1), 49-57.

[15] R core Team, (2019). R version 3.5.3: A language and environment for statistical computing. R Foundation for statistical computing, *Vienna, Austria*.URL https:// www.R-project .org/.

[16] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; *Pergamum Press and Statistical Publishing Society*, Calcutta. 320-332.

[17] 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.

[18] Rather, A. A. & Subramanian, C. (2018). Length biased Susila distribution, *Universal Review*, vol 7, issue. XII, pp. 1010-1023.

[19] 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,* Vol.5, Issue.5, pp.72-76.

[20] Rather, A. A. and 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.

[21] Rather, A. A., Subramanian, C., Shafi, S., Malik, K. A., Ahmad, P. J., Para, B. A., and Jan, T. R. (2018). A new Size Biased Distribution with applications in Engineering and Medical Science, *International Journal of Scientific Research in Mathematical and Statistical Sciences,* Vol.5, Issue.4, 75-85.

[22] Reyad., M. H., Othman, A. S. and Moussa, A. A. (2017). The Length-biased Weighted Erlang distribution, *Asian Research Journal of Mathematics*, 6(3), 1-15.

[23] Saghir, A. and Khadim, A. (2016). The mathematical properties of length biased weighted Maxwell distribution, *Journal of Basic and Applied Research International*, 16(3), 189-195.

[24] Saghir., A., Tazeem, S. and Ahmad, I. (2016). The length-biased weighted exponentiated inverted Weibull distribution, *Cogent Mathematics*, 3(1), Article:1267299.

[25] Sanat, P. (2016). Beta-length biased Pareto distribution and its properties, *Journal of Emerging Technologies and Innovative Research (JETIR)*, 3(6), 553-557.

[26] Shenbagaraja, R., Rather, A. A. and Subramanian, C. (2019). On Some Aspects of Length Biased Technique with Real Life Data, *Science, Technology and Development*, Vol VIII, issue IX, pp 326-335.

[27] Subramanian, C. & Rather, A. A. (2020). A New Extension of Shanker distribution with Real Life Dat, *Journal of Xi'an University of Architecture & Technology*, vol. XII, Issue III, pp 2001-2009.

MISSING DATA IMPUTATION VIA OPTIMIZATION APPROACH: AN APPLICATION TO K-MEANS CLUSTERING OF EXTREME TEMPERATURE

Geovert John D. Labita[1], Bernadette F. Tubo[2]

•

[1]University of Science and Technology of Southern Philippines
[2]Mindanao State University – Iligan Institute of Technology
[1]geovertjohn.labita@g.msuiit.edu.ph

Abstract

This paper introduces an optimization approach to impute missing data within the K-means cluster analysis framework. The proposed method has been applied to Philippine climate data over the previous 18 years (2006-2023) with the goal of classifying the regions according to average annual temperature including the maximum and minimum. This dataset contains missing values which is the result of the weather stations' measurement failure for some time and there is no chance of recovery. As an effect, the regional groupings are greatly affected. This paper adapts a modified method of missing value imputation suitable for climate data clustering, inspired by the work of Bertsimas et al. (2017). The proposed methodology focuses on imputing missing values within observations by finding the value that minimizes the distance between the observation and a cluster centroid in which the Mahalanobis distance is used as the similarity measure. Consequently, the outcomes of clustering obtained through this optimization approach were compared with certain imputation techniques namely Mean Imputation, Expectation-Maximization algorithm, and MICE. The assessment of the derived clusters was conducted using the silhouette coefficient as the performance metric. Results revealed that the proposed imputation gave the highest silhouette scores which means that most of the observations were being clustered appropriately as compared to the results using other imputation algorithms. Moreover, it was found out that most of the areas showing the features of extreme condition are located in the middle part of the country.

Keywords: Optimization, K-Means, Mahalanobis

I. Introduction

The risk of extreme temperature most directly affects health by compromising the body's ability to regulate its internal temperature. Loss of internal temperature control can result in various illnesses including heat cramps, heat exhaustion, heatstroke, and hyperthermia from extreme heat events [7]. Thus, awareness of the climatic differences of a particular region of interest becomes a major concern for the safety of the individual.

In detecting weather phenomena like extreme temperature, it is important to classify or cluster the regions according to their climatic elements. However, the problem of missing climatic data is common in most weather stations which might result from damaged or failure of the weather equipment or instrument. Also, events such as sickness or vacation of the personnel in-charge can create daily missing data values which could affect the climate statistics. If this happens, there will

be no record of measurements for a particular time and could affect the clustering of weather data which is a valuable endeavor in multiple respects. For example, the results can be used in various ways within a larger weather prediction framework or could simply serve as an analytical tool for characterizing climatic differences [4].

From the study of Calvo et al. [6], a new clustering technique was shown aiming to generate a robust regionalization using climate datasets with incomplete information. Their method provided a new approach to cluster time series of different temporal lengths using most of the information contained in heterogeneous sets of climate records. Although they showed that their algorithm is able to generate a climatically consistent regionalization, it must be noted that there is no imputation happened on the missing information. In a sense, the clustering accuracy is somehow questionable.

A common practice for dealing with missing values in the context of clustering is to first impute the missing values, and then apply the clustering algorithm on the completed data [5]. From the study of Bertsimas et al. [3], a flexible framework based on formal optimization to impute missing data was proposed. Specifically, this framework can readily incorporate various predictive models like the k Nearest Neighbors (kNN) for data classification in which the missing data of an observation is imputed by determining the k nearest observations and getting the average of those k observations. However, the imputation for each observation is not based on the possibility that the point belongs to a particular cluster. Thus, the kNN imputation is based purely on the k neighbors without the involvement or intervention of the possible resulting clustering.

Trying to resolve the aforementioned issues or deficiencies, this paper creates an appropriate imputation technique for missing values when dealing with clustering problem. Specifically, this study aims to construct a two-step optimization approach for data imputation in K-means cluster analysis where K is the number of clusters. The first step is to determine the optimal initial cluster centroids which are the K most frequent nearest neighbors from all incomplete observations, that is, the K points with highest densities. The second step is then imputing the missing value of an observation by determining the value that gives the minimum distance from the observation to a cluster centroid. The outcomes of clustering achieved through this optimization approach will be compared with some imputation approaches namely Mean Imputation, Expectation-Maximization algorithm, and Multivariate Imputation by Chained Equations in which the assessment of the derived clusters will be conducted using the silhouette coefficient.

This paper is arranged as follows. Methodology is introduced and discussed in section 2. The model solution is presented and derived in section 3. In section 4, the application of the proposed imputation is illustrated while some concluding remarks are stated in section 5.

II. Methods

This section presents the derivation of the optimization models of the proposed method with imputation algorithm.

Let $X = \{x_i\}_{i=1}^{n}$ be the dataset given with p variables and assume that each data vector x_i contains continuous variables indexed by $q \in \{1, 2, \dots, p\}$. Now, the missing and known values are defined by the following sets:

$$\mathcal{M} = \{(i, q) : x_{iq}\ is\ missing\},$$
$$\mathcal{N} = \{(i, q) : x_{iq}\ is\ known\}.$$

Also, let J be the set of indices of all incomplete observations given by

$$J = \{i :\ x_i\ has\ at\ least\ 1\ missing\ coordinate\}.$$

Let $W \in \mathbb{R}^{n \times p}$ be the matrix of imputed values where w_{jq} is the imputed value for entry x_{jq} for $(j, q) \in \mathcal{M}$. The full imputation for observation x_j is referred to as w_j where $j \in J$. The idea is to consider the missing data problem as an optimization problem in which it optimizes the missing values in all incomplete data points. Thus, the key decision variables are the missing values

$\{w_{jq} : (j,q) \in \mathcal{M}\}$.

As a similarity measure, we can incorporate different types of distance metrics, but we prefer to use the Mahalanobis distance because it takes into account the variances and covariances amongst the variables which is very important in clustering multivariate data. In constructing the Mahalanobis metric, it involves the centroid of the whole dataset which means that the distance actually measures a point from the mean of the distribution. Specifically, according to Ghorbani [8], the Mahalanobis distance measures the number of standard deviations that an observation is from the mean of a distribution.

In using the Mahalanobis distance as a similarity measure, the nearest neighbors of incomplete data are formulated based on the differences of the squared Mahalanobis distances of the two observations. Thus, the nearest neighbor of each w_j, $j \in J$ is the smallest difference $M_j - M_i$ for all $i = 1, 2, \dots, n$, that is, the smallest deviations between w_j and w_i where the squared Mahalanobis distance M_i is given by

$$M_i(w_i, \mu) = [w_{i1} - \mu_1 \quad \dots \quad w_{ip} - \mu_p]\Sigma^{-1}\begin{bmatrix} w_{i1} - \mu_1 \\ \vdots \\ w_{ip} - \mu_p \end{bmatrix}$$

with $\mu = \{\mu_1, \dots, \mu_p\}$ and Σ are the mean and covariance matrix of the whole data respectively which are updated per iteration.

Imputation Model

To obtain the imputed values, the Mahalanobis distance between w_j, $j \in J$ and its appropriate centroid w_{i_l}, $l \in \{1, 2, \dots, K\}$ is minimized. Thus, for each $j \in J$, the goal is to solve the imputation model:

$$\min M_j - M_c \tag{1}$$

subject to

$$w_c \in \{w_{i_l}\} \qquad l = 1, 2, \dots, K \tag{2}$$

$$w_{jq} = x_{jq} \qquad (j,q) \in \mathcal{N} \tag{3}$$

The solution $\{w_{jq}\}$, $(j,q) \in \mathcal{M}$ are regarded as the imputed values for the corresponding $\{x_{jq}\}$. It must be noted that in the objective function (1), we assume that $M_j > M_c$. If $M_c > M_j$, we change the objective to $\max M_j - M_c$ in order to represent the same idea that the value of M_j should be near to M_c. In other words, the objective function ensures that whatever imputed values w_{jq} obtained, the observation w_j is very close to its appropriate cluster centroid w_c which is selected based on constraint (2). These centroids are determined in the assignment model discussed in the next section. The constraint (3) assures that all the observed data are preserved.

Assignment Model

Let K be the number of clusters specified by the analyst. Now, assume that the initial cluster centroids are given by $\{w_{i_l} : l = 1, 2, \dots, K\}$ which are the K most frequent nearest neighbors from all incomplete observations. To obtain the initial centroids, the immediate nearest neighbor for each w_j, $j \in J$ must be determined resulting to the following assignment model:

$$\min \textstyle\sum_{i=1}^{n} z_{ij}(M_j - M_i) \tag{4}$$

subject to

$$\textstyle\sum_{i=1}^{n} z_{ij} = 1 \tag{5}$$

$$z_{jj} = 0 \tag{6}$$

$$z_{ij} \in \{0, 1\}$$

The assignment model assigns each incomplete observation to its immediate nearest neighbor where $z_{ij} = 1$ if w_i is the nearest neighbor of w_j and 0 otherwise. The objective function (4) will

determine which w_i is the nearest neighbor of w_j among all observations. Because of constraint (5), there will only be one immediate nearest neighbor per incomplete observation and an incomplete observation cannot be the nearest neighbor of itself because of constraint (6).

From all of the nearest neighbors, the K most frequent observations can then be formulated as an optimization problem using the binary variables $y_i \in \{0,1\}$ as follows:

$$\max \sum_{i=1}^{n} y_i \sum_{j \in J} z_{ij} \qquad \text{subject to} \qquad \sum_{i=1}^{n} y_i = K \tag{7}$$

The solution $\{y_{i_1}, \dots, y_{i_K}\}$ of model (7) corresponds to the desired initial centroids $\{w_{i_1}, \dots, w_{i_K}\}$. It must be noted that the assignment model will work only on complete data with imputed values. For the first iteration with missing values, the model can be started with mean values as the warm start values for the optimization process. The imputed values from the imputation model are then based on the centroids obtained from the assignment model. In return, the centroids are updated based on the new imputed values making this procedure an iterative process.

Imputation Algorithm

The proposed data imputation algorithm is given in the following steps:

1. **Input:** $X \in \mathbb{R}^{n \times p}$, a data matrix with missing entries $\mathcal{M} = \{(i,q) : x_{iq} \text{ is missing}\}$, warm start $W^0 \in \mathbb{R}^{n \times p}$ and number of clusters K.
2. **Output:** W^*, a full matrix with imputed values, $\mu^* = \{w_{i_1}, \dots, w_{i_K}\}$ initial centroids.
3. **Initialize:** $W^{old} \leftarrow W^0$
4. **repeat**
5. Update mean μ and covariance matrix Σ based on W^{old}.
6. Update the auxiliary variables Z^* using the assignment model.
7. Update the initial centroids μ^* following:
$$\sum_{j \in J} z_{i_l j} > \sum_{j \in J} z_{ij} \quad \forall i \in \{1, 2, \dots, n\}$$
8. Update the imputation W^* using the imputation model.
9. $(Z^{old}, W^{old}, \mu^{old}) \leftarrow (Z^*, W^*, \mu^*)$
10. **until** $\mu^* = \mu^{old}$

III. Results

This section presents the solution of the proposed imputation method using Mahalanobis distance.

Proposition 1. Let $X = \{x_i\}_{i=1}^n$ be a dataset given with p variables where the missing and known values are specified by the sets $\mathcal{M} = \{(i,q) : x_{iq} \text{ is missing}\}$ and $\mathcal{N} = \{(i,q) : x_{iq} \text{ is known}\}$ respectively. If $(j,q) \in \mathcal{M}$, then the solution of the optimization problem (1-3) is given by

$$w_{jq} = \mu_q - \frac{1}{2\sigma_{qq}} \sum_{a:a \neq q}^{p} \sigma_{qa}(w_{ja} - \mu_a)$$

where $\mu_q, \sigma_{qa} \in \mathbb{R}$ and $\sigma_{qq} > 0$.

Proof. Let $(j,q) \in \mathcal{M}$ and consider the optimization problem (1-3). Suppose that $w_c = w_{i_l}$ such that $M_j - M_{i_l} < M_j - M_m$ for all $m \neq l$. Then by considering an unconstrained optimization where we plugin the values of the x_{jq} to the corresponding w_{jq} for all $(j,q) \in \mathcal{N}$ in objective function (1), we can use the concept of relative minimum in calculus to solve for w_{jq} that would minimize $M_j - M_{i_l}$. Since the missing variable w_{jq} is present only in M_j, the problem reduces to differentiating,

$$M_j = \begin{bmatrix} w_{j1}-\mu_1 & \cdots & w_{jq}-\mu_q & \cdots & w_{jp}-\mu_p \end{bmatrix} \Sigma^{-1} \begin{matrix} w_{j1}-\mu_1 \\ \vdots \\ w_{jq}-\mu_q \\ \vdots \\ \lfloor w_{jp}-\mu_p \rfloor \end{matrix}$$

with respect to w_{jq} where $\mu = \{\mu_1, \dots, \mu_p\}$ and Σ are the mean and covariance matrix respectively. Now, suppose that

$$\Sigma^{-1} = \begin{matrix} \sigma_{11} & \cdots & \sigma_{1q} & \dots & \sigma_{1p} \\ \vdots & & \vdots & & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} & \cdots & \sigma_{qp} \\ \vdots & & \vdots & & \vdots \\ \lfloor \sigma_{p1} & \cdots & \sigma_{pq} & \cdots & \sigma_{pp} \rfloor \end{matrix},$$

then we have

$$M_j = \sum_{b=1}^{p} \sum_{a=1}^{p} \sigma_{ab}(w_{ia} - \mu_a)(w_{jb} - \mu_b).$$

To differentiate M_j, we have to separate the terms containing w_{jq}, that is,

$$M_j = \sum_{a=1}^{p} \sigma_{qa}(w_{jq} - \mu_q)(w_{ja} - \mu_a) + \sum_{b:b\neq q}^{p} \sum_{a:a\neq q}^{p} \sigma_{ab}(w_{ja} - \mu_a)(w_{jb} - \mu_b)$$

$$D_{w_{jq}}(M_j) = 2\sigma_{qq}(w_{jq} - \mu_q) + \sum_{a:a\neq q}^{p} \sigma_{qa}(w_{ja} - \mu_a).$$

Finally, equating the derivative to zero will solve for the imputed value as follows

$$2\sigma_{qq}(w_{jq} - \mu_q) + \sum_{a:a\neq q}^{p} \sigma_{qa}(w_{ja} - \mu_a) = 0$$

$$2\sigma_{qq} w_{jq} = 2\sigma_{qq}\mu_q - \sum_{a:a\neq q}^{p} \sigma_{qa}(w_{ja} - \mu_a)$$

$$w_{jq} = \mu_q - \frac{1}{2\sigma_{qq}} \sum_{a:a\neq q}^{p} \sigma_{qa}(w_{ja} - \mu_a). \qquad \blacksquare$$

The following theorem will be used to prove the next proposition.

Theorem 1 (Andreasson et al.). Suppose that $f: \mathbb{R}^d \to \mathbb{R}$ is in C^2 on $\mathbb{R}^d$, that is, f is twice differentiable with continuous second partial derivatives. Then $\nabla f(w^*) = 0^{(d)}$ and $\nabla^2 f(w^*)$ is positive definite implies that w^* is a *strict local minimum* of f where $\nabla f(w) = \left(\frac{\partial f(w)}{\partial w_q}\right)_{q=1}^{d}$. For $d = 1$, $f'(w^*) = 0$ and $f''(w^*) > 0$ implies $w^* \in \mathbb{R}$ is a *strict local minimum*.

Proposition 2. The solution w_{jq} given in Proposition 1 is a strict local minimum of the optimization problem (1-3) in an unconstrained setting.

Proof (for the case when $d = 1$). Let $f: \mathbb{R} \to \mathbb{R}$ be defined by the objective function in the optimization problem (1-3) in an unconstrained setting. Following the same argument from the proof of Proposition 1, for any solution w^*, we have

$$f'(w^*) = 2\sigma_{qq}(w^* - \mu_q) + \sum_{a:a\neq q}^{p} \sigma_{qa}(w_{ja} - \mu_a) \quad \Rightarrow \quad f''(w^*) = 2\sigma_{qq}.$$

Since $f'(w^*)$ and $f''(w^*)$ are linear functions, then they are continuous. Also, $f''(w) = 2\sigma_{qq} > 0$ since the diagonal entries of a covariance matrix are positive assuming that the data samples are unique. Now,

$$f'(w_{jq}) = 2\sigma_{qq}\left(\mu_q - \frac{1}{2\sigma_{qq}}\sum_{a:a\neq q}^{p}\sigma_{qa}(w_{ja}-\mu_a) - \mu_q\right) + \sum_{a:a\neq q}^{p}\sigma_{qa}(w_{ja}-\mu_a)$$
$$= 2\sigma_{qq}\left(-\frac{1}{2\sigma_{qq}}\sum_{a:a\neq q}^{p}\sigma_{qa}(w_{ja}-\mu_a)\right) + \sum_{a:a\neq q}^{p}\sigma_{qa}(w_{ja}-\mu_a)$$
$$= -\sum_{a:a\neq q}^{p}\sigma_{qa}(w_{ja}-\mu_a) + \sum_{a:a\neq q}^{p}\sigma_{qa}(w_{ja}-\mu_a)$$
$$= 0.$$

Thus, by Theorem 1, the solution w_{jq} is a strict local minimum. ■

IV. Application

The proposed methodology is applied on the historical Philippine climate data (2006-2023) taken from the 52 weather stations around the country which can be downloaded at https://en.tutiempo.net/climate/philippines.html and shown in Table 2. This dataset of three continuous variables per year (52×54 data matrix) contains actual missing values. This study can be considered as a multivariate time series clustering with the goal of classifying the regions suspected to have extreme temperature conditions.

In doing the experiment, the missing elements among the data are firstly imputed using the different imputation methods, and then the traditional K-means algorithm is applied into the imputed dataset. The experiments with random centroid initialization (mean, MICE, EM) are repeated 100 times with different random seed to reduce the effect of randomness caused by the traditional K-means, and report the best result.

We use the R function "*silhouette()*" from the R package "*cluster*" for obtaining the silhouette scores of the clustering results. Silhouette coefficient or Silhouette score ranging from -1 to +1 is a measure of how similar an object is to its own cluster compared to other clusters. In other words, it is a metric used to calculate the goodness of a clustering [2]. A high value indicates that the object is well matched or having a high relationship to its own cluster. Thus, it acts as the accuracy in the case when the cluster labels are not known.

Table 1 shows the silhouette score results from different number of clusters where the numbers in red are the highest score per case.

Table 1: *Silhouette Scores (%) using different imputation algorithms*

# of Clusters	Proposed Imputation	Mean Imputation	MICE	Expectation-Maximization
K=2	84.78	75.26	61.51	70.98
K=3	72.6	62.7	52.64	58.66
K=4	58.02	36.03	29.39	21.75
K=5	58.02	22	26.38	19.12
K=6	57.99	20.5	33.33	19.04
K=7	41.11	20.09	17.89	18.17
K=8	38.06	17.89	17.66	17.16
K=9	36.56	17.59	17.03	16.65
K=10	36.27	16.88	15.74	17.75

Using the proposed imputation method, we can classify the extreme temperature areas. For example, if we set $K = 10$, results showed that there are two clusters exhibiting extreme temperature having an overall average of at least 28°C. These areas are shown in Figure 1.

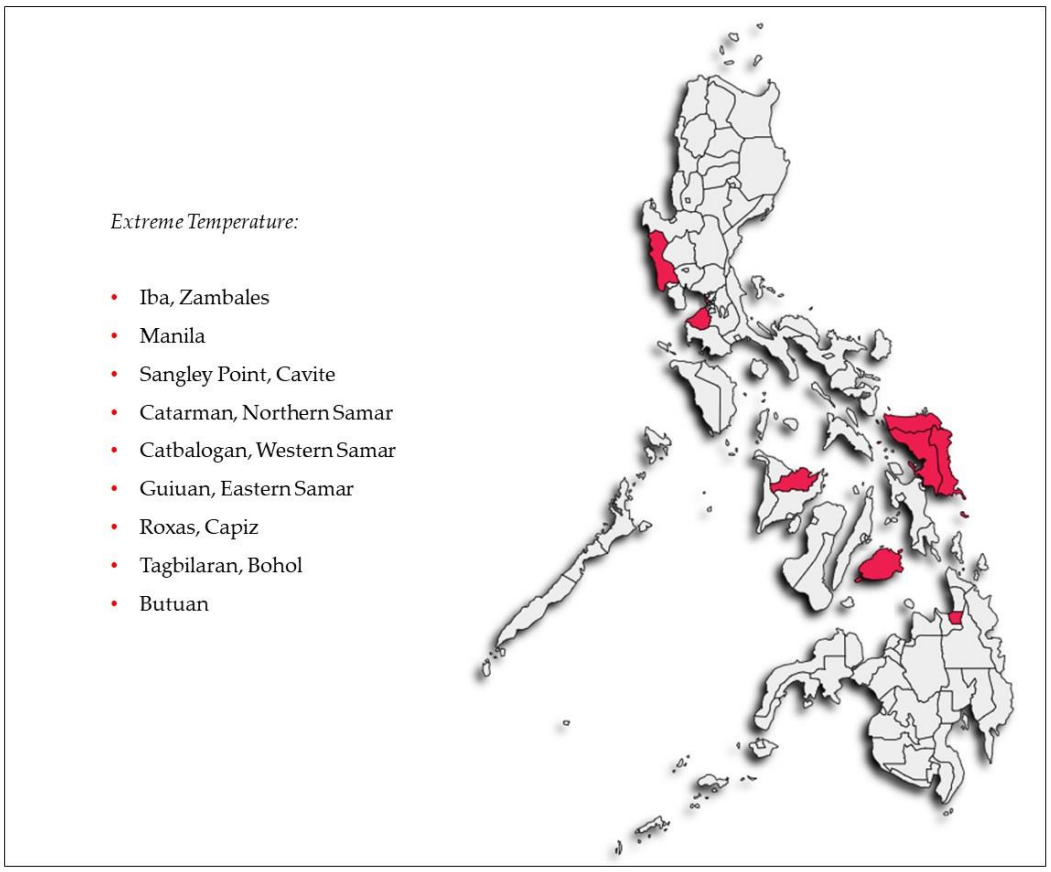


Figure 1: *Philippine map with clustering results from the proposed imputation*

From Figure 1, the areas with red spots are classified with extreme temperature. It can be observed that most of the areas are located in the middle part of the country.

V. Concluding Remarks

This paper presents a missing data imputation algorithm that can handle partitional clustering. It is created out of an optimization approach for imputing missing data and making use of the Mahalanobis distance metric as a similarity measure. Also, it avoids the problem of centroid initialization when performing K-means clustering because the initial cluster centroids are fixed based on the algorithm's generated centroids.

When clustering the Philippine Climate data with 21% actual missing values, we were able to identify 9 places with extreme temperature classification which means that these places must be considered when predicting extreme temperature occurrence. It was found out that the proposed imputation using Mahalanobis distance gave higher clustering performance and is consistent for different number of clusters which means that the proposed optimization approach using Mahalanobis distance is a suitable imputation algorithm in the context of partitional clustering.

Table 2: *Philippine Climate Dataset 2006-2023 (T-average temperature, TM-average maximum temperature, Tm-average minimum temperature)*

Weather	2006			2007			2008			2009			2010			2011			2012			2013			2014		
Stations	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm
Alabat																						27.4	31	24	27.3	30.8	23.8
Ambulong																						27.7	31.3	23.8	27.6	31.5	23.2
Aparri				27.8	32.1	23.8	27.4	31.5	23.4	27.1	30.8	23.5	27.6	31.6	23.7	27	30.6	24.2	27.6	31.1	24.4	27.4	31.1	24.3	26.8	30.7	24
Baguio	19	23.8	15.5	19	23.8	15.5	18.7	23.3	15.3	18.7	23.2	15.4	19.4	24.6	15.3	18.8	23.6	14.5	19.4	23.9	14.8	19.5	24.1	15.3	19	23.3	15
Baler Radar																			26.1	30.2	23.1				26.2	30.6	23
Basco Radar	25.4	27.7	22.1										25.9	28.1	22.2	24.7	26.8	21.1	25.2	27.4	21	25.3	27.3	20.8	25	27.2	22.1
Cabanatuan	27.4	33.2	23.1	27.4	33.3	23.1	27	32.9	23.1	26.9	33.1	23.1	27.2	34	23.4	26.7	32.6	23.4	27.5	32.8	23.6	27.7	33	23.9	27.6	33.1	23.7
Casiguran													26.9	31.7	22	26.1	31	20.3	26.5	31.5	22.4	26.9	30.8	23.3	26.8	31.3	22.8
Clark AB	26.4	31.2	22.3	26.5	31.3	22.2	26.1	30.8	22.3	26.2	30.7	22.2	26.8	32	22.4	26	30.7	22.2	26.5	31.3	22.5	26.6	31.3	22.6	26.7	31.1	23.1
Daet	27.8	31.1	25	27.8	31	24.9	27.4	30.8	24.8	27.4	30.4	24.8	27.7	31.1	24.8	27.3	30.5	24.7	27.6	31	24.4	27	30.9	23.9	27.3	30.7	23.8
Dagupan	27.7	32.2	23.1	27.6	32.1	22.8	27.4	32.1	22.9	27.3	31.9	22.7	28	32.8	23.4	27.3	31.5	23.3	27.8	32.3	23.6	28	32.7	23.8	27.6	32.4	23.5
Iba	27.7	31.5	23.6				27.7	31.5	23.8	27.4	31.3	23.5				27.6	31.3	24.1	28	32.1	24.3	28.3	32.3	24.5	27.8	31.6	24
Infanta	27.3	31.2	24.1				26.9	30.7	23.8	27	31	23.8	27.4	31.6	23.8	26.8	30.7	23.7	27.4	31.6	23.9	27.2	31.2	23.6	27	30.4	23.7
Itbayat										25.4	28.1	21.6				25.1	27.9	21.8	26.1	28.7	22.7	26	28.3	22.5	25.3	28	21.8
Laoag	28.9	32.9	23.1	28.8	32.5	22.8	28.7	32.3	22.9	28.4	32	22.8	28.4	32.7	22.9	28	31.8	23	28	31.6	23	28	31.7	23.1	27.6	31.3	22.5
Legaspi	27.7	30.6	24.9				27.3	30.3	24.6	27.4	30.5	24.8	27.9	31.5	24.7	27.3	30.4	22.7	27.6	30.8	23.3	27.7	31.3	24.2	27.5	30.8	24.5
Manila	28.6	31.7	25.7	28.7	31.8	25.7	28.3	31.4	25.5	28.2	31	25.4	29	32.2	26	28.2	31.2	25.5	28.4	31.6	25.5	28.4	31.8	25.3	28.1	31.8	24.6
Masbate	27.9	32.2	24.9	28.1	32.5	24.9	27.8	31.9	24.8	27.9	31.9	24.8	28.4	32.7	25	27.5	31.7	24.6	27.8	32.1	24.7	28.1	31.8	24.7	27.8	31.6	24.6
NAIA	28	31.6	24.7	28.1	32.1	24.7	27.7	31.5	24.6	27.6	31.3	24.5	28.5	32.7	25	27.7	31.5	24.5	28	31.8	24.7	28.2	32.1	24.8	28.1	32.1	24.7
Puerto Princesa	27.8	32.1	24.5	28	32.3	24.4	27.5	31.5	24.2	27.8	32	24.4	28.1	32.2	24.6	27.6	31.5	24.5	27.9	31.7	24.7	28	31.9	24.8	27.8	31.5	24.5
Romblon	27.7	30.2	25.4	27.6	30.2	25.1	27.4	29.9	25	27.2	29.7	25	28	31	25.4				26.3	28.7	24.4	26.3	28.8	24.3			
Sangley Point	28.8	32	26.1	28.6	32	25.7	28.3	31.5	25.5							28.3	31.6	25.6	28.7	32.1	25.8	28.9	32.3	25.9	28.2	32.1	25.1
Science Garden	27.4	31.8	23.5	27.5	32.1	23.6	27.1	31.9	23.2	27	31.7	23.4	27.7	32.9	23.7	26.8	31.5	23.3	27.2	32	23.3	27.4	32.2	23.4	27.2	32	23.1
Subic Bay	27.8	31.9	24.1	27.9	32.1	23.9	27.5	31.5	23.9	27.6	31.6	23.7	28	33.1	24	27.4	32	23.9	27.7	32.2	24	28.1	33.1	23.8	27.6	32.8	23.1
Tanay				23	26.6	20.1	22.6	25.8	20.2	22.5	25.7	20.1	23.2	27.2	20.4	22.4	25.8	20.2	22.7	26	19.8	22.9	26.3	19.9	22.8	26.4	19.7
Tayabas	26.6	30.4	23.4	26.6	30.7	23	26.2	30.2	22.9	26.2	30.1	22.8	26.7	30.9	23.2	26.2	29.7	22.8	26.5	30.3	23	26.6	30.4	23	26.2	30	22.7
Tuguegarao	27.2	32.5	22.6	27.1	32.6	21.2	26.6	31.8	20.6	26.5	31.4	20.9	27.5	33.3	21.4	26.5	31.6	21.4	26.9	32.2	21.9	26.9	32	22.2	26.8	32.3	22.5
Vigan	27.3	31.1	23.2	27.1	31.1	22.9	27.2	30.9	23	27.2	31	23.2	27.4	31.9	23	27	31.4	22.5	27.3	31.5	23.2	27.3	31.4	23.1	26.8	31.3	22.8
Borongan																									26.9	30.5	23.4
Calapan				27.2	31.2	23.6										26.7	30.5	23.1	27	31.5	23.4						
Catarman	27.6	31.3	23.8	27.6	31.5	23.3	27.2	31.2	23	27.5	31.4	23.3	27.5	31.9	23.2	27.3	30.9	23.4	27.6	31.3	23.7	27.7	31.4	23.7	27.2	30.9	23.5
Catbalogan	28.1	32.3	24.4				27.6	31.9	22.6	28	31.9	24.1	28.1	32.8	24.7	27.6	31.5	24.6	28	32.4	24.7	28.1	32.4	24.6	27.6	32.2	24.3
Dumaguete	27.6	30.6	24.3	27.6	30.8	24.4	27.4	30.7	24.2	27.5	30.9	24.3	27.9	31.3	24.4	27.5	31.1	24.3	27.7	31.4	24.4	28	31.6	24.7	27.6	31.3	24.4
Guiuan	27.6	30	25.4	27.7	30	25.5	27.4	29.5	25.2	27.6	30.1	25.1	27.9	30.7	24.9	27.4	29.9	24.3	27.9	30.5	24.5				27.4	29.8	24
IloIlo	27.9	31.7	23.9	27.9	32	23.4				27.2	30.7	23	28	32.1	23.5												
Maasin				27.3	31.6	24.1	26.9	31.1	23.7	27.1	31.6	23.9	27.1	32.3	23.5	26.5	30.5	21.9	27.2	31.5	21.5	27.3	31.3	23.3	27	30.9	23.7
Mactan	27.9	30.9	24.8	28	31.2	24.5	27.8	31	24.7	27.8	31.4	25	28.4	32.3	24.9	27.4	30.8	24.7	27.8	31.3	24.9	28.1	31.4	25.2	27.8	31.2	24.6
Roxas	27.8	30.8	25.1	28	31	25	27.7	30.6	25	27.8	30.7	25.2	28.2	31.3	25.4	27.6	30.3	25.2	28	30.9	25.5	28	30.9	25.5	27.9	30.8	25.3
San Jose																			28.1	32.3	23.8	28.2	32.4	23.8	27.8	32.4	23.4
Tacloban	27.7	31.3	24.9				27.3	31	24.7	27.6	31.3	24.7	27.6	31	24.9	27.2	30.3	24.8	27.5	31.1	24.8						
Tagbilaran	28.1	32.1	24.3				27.8	32.1	23.9	28.2	32.6	24.2	28.2	33.1	24.1	27.7	32.2	24.1	28.1	32.9	24.3	27.9	31.8	24.3	27.4	31.4	23.9
Virac				27.5	30.8	23.8	27.2	30.6	23.8	27.3	30.7	23.9	27.4	31.1	23.8	27	30.5	23.7	27.4	30.9	24	27.3	30.8	24	27	30.8	23.4
Butuan	27.8	31.9	24.2	27.8	32.2	23.9	27.5	32.1	24	27.8	32.1	24.3	27.9	32.6	24	27.5	31.7	24.1	27.6	32	24	27.5	32.1	23.8	27.6	31.8	23.9
Cotabato																									27.8	32.6	23.3
Davao Airport	28.2	31.8	24.2	28.3	31.7	24.1	28	31.6	23.9	28.5	32	24.2	28.7	32.4	24.3	28.2	31.7	24.1	28.4	32	24.2	28.6	32.3	24.4	28.5	32.4	24.3
Dipolog	27.9	32	24.6	27.8	31.6	24.7	27.3	31.3	24.4	27.5	31.7	24.5	27.7	31.8	24.7	27.4	31.3	24.2	27.5	31.2	24.2	27.6	31.4	24.1	27.6	31.8	23.9
Gen. Santos				27.6	32.7	23.1	27.1	32.1	22.7	27.6	32.8	23.2	27.8	33.1	23.1	27.1	31.9	22.6	27.1	32.1	22.7						
Hinatuan	27.9	31.8	24.2	27.5	31.7	23.9	27.2	31.2	23.7				27.4	31.3	23.8	27.1	31.2	23.6	27.3	31.5	23.9	27.6	31.3	24	27.3	31.3	23.8
Lumbia	26.4	31.6	22.3	26.4	31.6	22.2	26	31.2	21.7	26.4	31.8	22.4	26.6	32.3	22.6	26.1	31.7	22.5	26.4	32.2	22.4						
Malaybalay	24.2	29.2	19.1	24.3	29.6	18.9	24.1	29.5	18.7	24.1	29.3	18.4	24.4	30.8	16.6	23.8	29.8	17.2	23.7	29.7	17.7	23.9	30	18.2	23.9	29.8	18.7
Surigao	28.7	31.4	25.4	28.1	31.6	24.7	27.9	31.3	24.3	27.7	31.3	24.1	27.9	31.7	24	28	31.4	23.5	28.1	31.9	24.2	27.8	31.6	23.6	27.8	31.5	23.5
Zamboanga	28.3	32.4	24	28.3	32.3	24.1	28.2	31.9	23.8	28.5	32.4	23.8	28.7	33.1	21.9	28.4	32.8	23.7	28.7	32.5	24.3	28.5	32.6	24.1	28.8	33.3	23.9

Weather	2015			2016			2017			2018			2019			2020			2021			2022			2023		
Stations	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm	T	TM	Tm
Alabat	27.4	31.1	23.7	27.7	31.4	24	27.4	31.1	23.9	27.5	31.1	24	27.7	31.5	23.7	27.6	31.2	24	27.6	30.8	24.3						
Ambulong	27.8	32	23.2	28.1	32.8	23.7	27.9	32.4	23.8	27.9	32.1	23.8	28	32.4	23.4				27.7	31.8	23.9						
Aparri	27.1	30.4	24.4	27.2	30.2	24.5	26.9	29.4	24.6	27	29.6	24.7	27	29.7	24.8	27.3	30.2	25				27.3	30.4	24.8	27.5	30.6	24.7
Baguio	19	23.3	15.2	19.5	23.7	15.6	19.3	23.7	15.6	19	23.4	15.3	19.1	23.5	15.3	19.5	24.2	15.5				19.4	23.6	16	19.7	23.8	16.4
Baler Radar	26.4	31	23.3	26.7	31	23.6	26.2	30	23.3	26.3	30.5	23.4	26.4	30.8	23.4	26.3	30.4	23.4	26.5	30.6	23.5						
Basco Radar	25.3	27.5	22.4	25.7	27.8	22.9	25.5	27.5	22.5	24.6	27.3	22.3	25.4	27.7	22.9	25.6	27.9	22.7				25.3	27.6	22.7	25.1	27.7	21.9
Cabanatuan	27.5	33.2	23.3	28.1	33.2	24	28	33.1	23.8	27.9	33.2	23.7															
Casiguran	26.5	30.8	22.5	26.8	31.1	23.2	26.4	30.8	23.2	26.8	31.3	23.2	26.9	31.7	23	27	31.4	22.6				26.8	30.9	22.3	27	31.6	22.3
Clark AB	27	31.4	23.5	27.7	32.1	24.2	27.6	32.2	24	27.5	31.9	24	27.4	32.1	23.7	27.9	32.8	23.8	27.7	32.5	23.9	28.1	32.8	24.3	28.1	32.8	24.2
Daet	27.3	30.9	23.5	27.6	31.2	23.5	27.3	30.7	23.9	27.4	30.8	24.5	27.6	31.4	24.2	27.6	31	24				27.4	30	24.3	27.8	31.1	25
Dagupan	27.9	32.5	23.7	28.1	32.6	24.3	28	32.8	24.2	27.9	32.5	23.8	27.8	32	23.8	28.1	32.1	24.4				28.1	32.6	24.5	28.4	32.6	24.8
Iba	27.7	31.7	23.4	27.9	32.1	23.5	28	31.9	24.2	27.8	30.8	24.6	27.7	31.3	23.9	28.1	31.9	24.2	28	32	24.2						
Infanta	27.1	30.7	23.9	27.5	31.1	24	27.1	30.5	22.9	27.2	30.7	22.9	27.5	31.2	23.9	27.4	30.9	23.9									
Itbayat	25.6	28.2	22.2	26	28.5	22.8	25.7	28.3	22.4	25.4	27.9	21.9				26	29.1	22.1									
Laoag	27.6	31.2	22.5	27.9	31.5	23.2	28	31.6	23.2	27.9	31.4	23	27.7	31.1	23	28	31.6	23.3	28	31.7	23.3	28	31.6	23.4	28.2	32	23.7
Legaspi	27.6	30.9	24.7	28.1	31.4	25.3	27.9	31.2	25.1	27.8	31.2	24.9	28	31.6	24.9	28.2	31.8	25.3				27.9	31.2	25.2	27.9	31.3	25
Manila	28.3	32.3	24.8	28.7	32.8	25.4	28.4	32.1	25.2	28.4	31.7	25.4	28.6	32	25.4				28.7	32	25.2						
Masbate	28	32.1	24	28.2	32	24	27.9	31.8	23.6	28.1	31.4	23.6	28.2	31.7	24.8	28.2	31.9	25				28.2	32.2	25.2	28.3	31.8	25.5
NAIA	28.3	32.5	24.7	28.7	32.9	25.2	28.4	32.4	25	28.4	32.3	25	28.6	32.6	25.4	28.8	32.8	25.3	28.7	32.7	25.3	28.5	32.6	25.4	28.7	32.6	25.4
Puerto Princesa	28.1	31.7	24.7	28.9	32.3	25	28.6	31.9	24.9	28.7	31.8	24.9	28.7	32	24.6	28.9	32.3	25.1	28.3	31.7	25	28.5	31.6	25.2	28.8	32.1	25.1
Romblon				26.9	29.5	24.8	26.5	28.9	24.5	26.4	28.8	24.5	26.8	29.6	24.5	26.9	29.4	24.7				26.2	28.5	24.4	26.5	28.9	24.6
Sangley Point	28.9	32.5	25.9	29.1	32.4	26.3	28.9	32	26.3	28.9	32.1	26.2	28.8	32.6	25.9	28.9	32.8	25.9	28.8	32.5	26.1						
Science Garden	27.6	32.4	23.3	28.2	32.5	24.2	27.9	32.1	24	27.8	31.8	23.9	28	32.2	23.9	28.1	32.3	24.1				27.9	31.9	24.4	27.9	31.8	24.4
Subic Bay	27.7	32.8	23.2	28	33.2	23.9	27.7	32.4	24	27.8	32.3	24.1	27.5	31.9	23.7	27.8	32.2	24	27.7	32.2	24	27.7	32	24.1	27.9	32.1	24.2
Tanay	23	26.7	20	23.4	26.4	20.5	23.2	26.4	19.5	23.1	26.1	19.2	23.3	26.3	20.8	23.5	26.7	21	23.1	26.1	20.9						
Tayabas	26.1	30.5	22.7	26.4	30.8	23.3	26.2	30.3	22.7	26.3	30.6	23.3	26.5	31	23.3	26.5	30.8	23.2	26.1	29.9	23						
Tuguegarao	26.8	32.2	22.5	27.4	32.6	23.2	26.6	31.5	23	27	32.5	22.9	26.9	32.3	22.8	27.1	32.4	23.1									
Vigan	27	31.4	23.2	27.1	31.4	23.4	27	31.5	23.3	27	30.7	23.4	26.9	30.6	23.2	27.3	31.2	23.5	27.5	31.3	23.6						
Borongan	27.2	30.6	23.8	27.3	30.8	23.9	27.2	30.8	23.9	27.3	31.4	24.1				27.3	31.4	24.3	27.2	31.2	24.3						
Calapan	27.5	31.7	23.7	27.3	31.4	23.8	27	31.1	23.5	27.1	31.4	23.8	27.7	32.3	23.8	27.8	32.2	24				27.4	31	24	27.7	31.3	24.9
Catarman	27.1	30.9	23.4	27.3	31.3	23.6	27	30.8	23.7	27.3	31.3	24	27.3	31.4	23.6	27.2	31.1	23.9	27.2	30.8	24.1						
Catbalogan	28.2	32.8	24.5	28.5	33	24.9	28.1	32.2	24.8	28.2	32.5	24.8	28.5	33.2	24.6	28.3	33.1	24.8	28.2	32.3	24.8						
Dumaguete	28	31.6	24.6	29	31.8	24.8	28.3	31	24.4	28.9	31.4	24.8	28.9	31.4	24.8	28.9	31.2	25.3	28.6	31.1	25.3	28.8	31.1	25.3	28.8	31.1	25.1
Guiuan	27.6	30.7	24.5	27.7	31.2	24.3	27.3	31.5	24.1	27.3	31.4	24.1	27.6	31.9	24.6	27.7	32.5	24.5	27.4	31.2	24						
IloIlo																											
Maasin	27.2	30.6	24.3	27.5	31.1	24.4	26.9	29.9	24.2	27.1	29.9	24	27.1	30.4	24.1	27.2	30.6	24									
Mactan	28	31.6	24.7	28.3	32.1	25	27.8	31.4	24.8	28.1	31.8	24.9	28.2	32	24.9	28.3	32	25				28	31.2	25.1	28.3	31.4	25.5
Roxas	28.2	31.2	25.4	28.4	31.6	25.6	28.2	31.3	25.5	28.3	31.4	25.7	28.3	31.5	25.6	28.4	31.6	25.8	28.3	31.1	25.8						
San Jose	28.1	32.7	23.6	28.4	32.8	23.5	28.3	32.7	22.9	28.2	32.6	22.2	28.3	33.6	21.6	28.5	33.6	21.3				28.5	33.2	21.8	28.5	33.1	24.7
Tacloban													27.8	31.4	24.7	27.7	31.1	25.2									
Tagbilaran	27.3	31.3	23.8	27.8	31.7	24.3	27.2	31	24	27.1	31.4	23.6	27.1	31.4	23.4	27.6	31.7	23.8	27.6	31.5	23.9						
Virac	27.2	31	23.6	27.7	31	24.2	27.5	30.8	24.1	27.5	31.4	24	27.5	30.9	23.5	27.6	30.9	23.9	27.6	30.7	24.3						
Butuan	28	32.4	23.9	28.2	33.1	24.1	27.7	32.1	24.1	27.9	32.4	23.9	27.8	32.3	23.7	27.8	32.5	23.7	27.7	32.1	23.8						
Cotabato	28.1	32.9	23.3	28.7	33.8	23.5	28.2	33.4	23.1	28.4	34.1	23.1	28.4	34.1	23.2	28.3	34.6	23.5	27.8	34	23.4						
Davao Airport	28.2	32.7	24.2	28.6	33.3	24.5	27.8	32.2	24.2	28.1	32.9	24.3	27.9	32.6	24.2	28.1	32.9	24.4	27.8	32.3	24.3	27.7	32.1	24.3	28	32.9	24.4
Dipolog	28.2	32.3	23.9	28	32.4	23.6	27.6	31.9	23.9	27.9	32.3	24.4	28.1	33.2	24.2	28.5	33.2	24.5				27.9	32.6	24.4	28.1	32.8	24.5
Gen. Santos							28.7	32.6	22.9				28.9	32.7	22.7	28.9	32.9	22.9	27.3	31.5	22.3				28.5	32.3	22.9
Hinatuan	27.3	31.4	23.8	27.4	31.5	23.8	27.3	31.4	24	27.4	31.9	23.8	27.3	31.7	23.7	27.4	31.6	23.8				27.2	31.2	23.8			
Lumbia																											
Malaybalay	24	30.3	18.6	24.6	30.9	18.6	24.2	30.2	19.6	24.5	30.6	19.7	24.6	30.7	19.4	24.5	30.7	18.1	24.2	30.4	18.7						
Surigao	28.1	32.2	23.1	28.2	32.3	24.1	27.7	31.8	24.3	28	32.2	24.4	28.1	32.3	24.5	27.8	31.2	24.3				27.7	31	24.1	28	31.3	24.6
Zamboanga	29.1	33.6	23.6	29	33.3	24.2	28.6	32.7	23.9	28.6	33.2	23.6	28.5	33.4	23.7	28.5	33.7	23.9	28.4	33.7	23.9	28.1	33.3	24	28.6	33.9	24.2

References

[1] Andr´easson, N., Evgrafov, A., & Patriksson, M. (2005). An introduction to optimization: Foundations and fundamental algorithms. *Chalmers University of Technology Press: Gothenburg*, Sweden, 1;1-205.

[2] Bhardwaj, A. (2020). Silhouette coefficient validating clustering techniques. *Towards Data Science*.

[3] Bertsimas, D., Pawlowski, C., & Zhuo, Y.D. (2017). From predictive methods to missing data imputation: an optimization approach. *J. Mach. Learn. Res.*, 18(1);7133-7171.

[4] Beveridge, N.R. (2021). Deep learning for weather clustering and forecasting. *Air Force Institute of Technology*. https://scholar.afit.edu/etd/5082

[5] Boluki, S., Zamani Dadaneh, S., Qian, X., & Dougherty, E.R. (2018). Optimal clustering with missing values. *In Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 593-594.

[6] Carro-Calvo, L., Jaume-Santero, F., García-Herrera, R., & Salcedo-Sanz, S. (2021). k-Gaps: a novel technique for clustering incomplete climatological time series. *Theoretical and Applied Climatology*, 143(1-2);447-460.

[7] Ebi, K.L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., ... & Jay, O. (2021). Hot weather and heat extremes: health risks. *The Lancet*, 398(10301);698-708.

[8] Ghorbani, H. (2019). Mahalanobis distance and its application for detecting multivariate outliers. *Facta Universitatis, Series: Mathematics and Informatics*, 583-595.

WEIGHTED TRANSMUTED MUKHERJEE-ISLAM DISTRIBUTION WITH STATISTICAL PROPERTIES

Danish Qayoom[1], Aafaq A. Rather[2,*]

•

[1,2]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[1]danishqayoom11@gmail.com, [2,*]aafaq7741@gmail.com

Abstract

In this study, we employ a weighted transformation approach to introduce a novel model that generalises the Transmuted Mukherjee-Islam distribution. The resulting generalized distribution is referred to as the Weighted Transmuted Mukherjee-Islam (WTMI) distribution The paper thoroughly explores the probability density function (PDF) and the corresponding cumulative distribution function (CDF) associated with the WTMI distribution. A thorough investigation of the distinctive structural properties of the proposed model is conducted, including survival function, conditional survival function, hazard function, cumulative hazard function, mean residual life, moments, moment generating function (MGF), characteristics function (CF), cumulant generating function (CGF), likelihood ratio test, ordered statistics, entropy measures, and Bonferroni and Lorenz curves. The maximum likelihood estimation method is employed for the precise estimation of model parameters.

Key words: Transmuted Mukherjee-Islam distribution, Reliability analysis, Maximum likelihood estimator, Ordered statistics

1. Introduction

In numerous applied sciences, including, engineering, agricultural science, biological science, biomedicine, ecology and various social science fields such as economics, finance, and population science, the modelling and analysis of lifetime data holds paramount importance. Multiple lifetime distributions have been employed to characterize such data and the effectiveness of statistical analysis procedures relies significantly on the chosen probability model or distribution. Consequently, substantial efforts have been taken for creating extensive classes of standard probability distributions along with corresponding statistical methodologies. Despite these advancements, numerous significant challenges persist, as real-world data often deviates from classical or standard probability models. Thus, the development of new forms of probability distributions remains a common objective of statistical theory. To extend the applicability of probability distributions, the literature proposes several methods that introduce additional parameter(s) to established baseline probability models. This enhances the flexibility of the models to capture the complexity of the data, leading to several generalized classes such as the Pearson Family, Burr Family, Exponentiated Family, Marshall-Olkin Family, T − X Family, Transmuted Family, Weighted Family, and more.

In this article, we employ a weighted transformation approach to introduce a novel model that generalizes the Transmuted Mukherjee-Islam distribution. The new generalized distribution will be termed as WTMI distribution. The weighted family of distributions emerged from the pioneering work of Fisher in 1934 [10], subsequently refined and formalized by Rao in 1965 [16]. This concept

serves as an essential tool in statistical theory, particularly in scenarios where observations are derived from non-experimental, non-replicated, and non-random conditions. In practice, Observing and recording all events is not always possible due to various factors. Some events may not be observable by the method used, may only be observable with a certain probability, or may change randomly during observation. Additionally, events produced under different mechanisms with unspecified relative frequencies may be mixed up and added to the same record. Therefore, the original event specification may not be appropriate for the recorded data unless modified. The weighted transformation approach enhances the flexibility of standard probability distributions in such scenarios. Researchers and scholars have extensively explored weighted probability models, along with the application of these models across various domains. Notably, Ghitany et al. [14] conducted a comprehensive study on the two-parameter weighted Lindley distribution, focusing on its relevance in analyzing survival data. Jain et al. [11] introduced the weighted gamma distribution, while Dey et al. [8] contributed significantly by introducing the weighted exponential distribution and its estimation techniques. In 2016 Das and Kundu [7], have obtained weighted and length biased version of exponential distribution. Kilany [13] have obtained the weighted version of lomax distribution. Subramanian and Rather [23] studied the weighted version of the exponentiated Mukherjee-Islam distribution, derived its statistical properties. Further, in 2018 Rather et al [20] explored the size-biased Ailamujia distribution with applications in engineering and medical science. Para and Jan [15] introduced the Weighted Pareto type-II distribution as a new model for handling medical science data and unveiling its statistical properties and potential applications across various fields. Rather and Subramanian [17] extensively studied the weighted Sushila distribution with properties and applications which shows more flexibility then its baseline distribution. Rather and Subramanian [19] further explored weighted distributions by offering a comprehensive overview, perspectives, and characterizations of the weighted version of Akshaya distribution with applications in engineering science. In a recent contribution, Rather and Ozel [18] discussed the weighted power Lindley distribution, demonstrating its effectiveness in analyzing lifetime data.

2. Probability density function (PDF) and cumulative distribution function (CDF)

The Transmuted Mukherjee-Islam distribution had explored by Rather and Subramanian {21] using the quadratic rank transmutation map studied first by Shaw and Buckley in 2007 [22]. Loai M. A. Al-Zou'bi [4] also obtained various properties of Transmuted Mukherjee-Islam distribution and its applications. The probability density function of a random variable say Z following Transmuted Mukherjee- Islam distribution with parameters say $(\varepsilon, \nu, \omega)$ is given by

$$f(z;\varepsilon,\nu,\omega)=\frac{\varepsilon}{\nu^{\varepsilon}}z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right);0<z<\nu,\varepsilon>0,\nu>0,-1\le\omega\le 1 \quad (1)$$

And the corresponding cumulative distribution function is

$$F_Z(z)=\left(\frac{z}{\nu}\right)^{\varepsilon}\left(1+\omega-\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right) \quad (2)$$

Various researchers have attracted to the quadratic transmutation map and they have introduced the new members of this family for various choices of baseline distributions. Transmuted Weibull distribution by Aryal and Tsokos [5], transmuted inverse Rayleigh distribution by Ahmad et al. [3], transmuted Marshall-Olkin Frechet distribution by Afify et al. [2], transmuted generalized Lindley distribution by Elgarhy et al. [9], transmuted modified Weibull distribution by Cordeiro et al. [6], transmuted exponential Lomax distribution by Abdullahi and Ieren [1], transmuted Burr Type X distribution by Khan et al. [12].

Using the weighted transformation approach, the PDF $y(z)$ of a non-negative random variable Z is given by

$$y_w(z) = \frac{w(z)\,y(z)}{E\big(w(z)\big)}; \qquad z > 0$$

Where $w(z)$ be a non-negative weight function and $E\big(w(z)\big) = \int_{-\infty}^{\infty} w(z)\,y(z)\,dz < \infty$.

In this paper, we will consider the weight function as $w(z) = z^s$ and PDF of the random variable Z to be Transmuted Mukherjee- Islam distribution to derive the PDF of Weighted Transmuted Mukherjee- Islam distribution. The PDF of Weighted Transmuted Mukherjee- Islam distribution distribution is given by

$$g(z;\varepsilon,\nu,\omega,s) = \frac{z^s f(z;\varepsilon,\nu,\omega)}{E(z^s)} \tag{3}$$

Now

$$E(z^s) = \int_0^{\nu} z^s \frac{\varepsilon}{\nu^{\varepsilon}} z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)dz \tag{4}$$

$$E(z^s) = \frac{\varepsilon}{\nu^{\varepsilon}}\left((1+\omega)\int_0^{\nu} z^{s+\varepsilon-1}dz - \frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\nu} z^{s+2\varepsilon-1}dz\right) \tag{5}$$

After simplification we get

$$E(z^s) = \frac{\varepsilon\,\nu^s\big(s(1-\omega)+2\varepsilon\big)}{(s+\varepsilon)(s+2\varepsilon)} \tag{6}$$

Using (1) and (4) in (3) we get

$$g(z;\varepsilon,\nu,\omega,s) = \frac{z^s \frac{\varepsilon}{\nu^{\varepsilon}} z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\frac{\varepsilon\,\nu^s\big(s(1-\omega)+2\varepsilon\big)}{(s+\varepsilon)(s+2\varepsilon)}} \tag{7}$$

$$g(z;\varepsilon,\nu,\omega,s) = \frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\big(s(1-\omega)+2\varepsilon\big)} \tag{8}$$

The corresponding CDF of WTMI distribution is given by

$$G_Z(z) = \int_0^{z}\left(\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\big(s(1-\omega)+2\varepsilon\big)}\right)dz \tag{9}$$

$$G_Z(z) = \frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\big(s(1-\omega)+2\varepsilon\big)}\left((1+\omega)\int_0^{z} z^{s+\varepsilon-1}dz - \frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{z} z^{s+2\varepsilon-1}\right) \tag{10}$$

After simplification we get

$$G_Z(z)=\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{11}$$

3. Reliability Analysis

3.1 Survival function

The survival function of WTMI distribution is given by

$$R_T(t)=P_r(T>t)$$

$$R_T(t)=1-P_r(T\le t)$$

$$R_T(t)=1-\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}t^{s+\varepsilon}-2\omega(s+\varepsilon)t^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{12}$$

$$R_T(t)=\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}t^{s+\varepsilon}+2\omega(s+\varepsilon)t^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{13}$$

After simplification we get

$$R_T(t)=\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}t^{s+\varepsilon}+2\omega(s+\varepsilon)t^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{14}$$

After simplification we get

$$R_T(t)=\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-t^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)t\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{15}$$

3.2 Conditional survival function

For an item survived for at least t_0 time (years, $t_0>0$), the probability that the item will survive additional t years is known as conditional survival function. In case of WTMI distribution the conditional survival function is given by

$$R_T(t\mid t_0)=P_r(T>t_0+t\mid T>t_0)$$

$$R_T(t\mid t_0)=\frac{P_r(T>t_0+t)}{P_r(T>t_0)}$$

$$R_T(t\mid t_0)=\frac{R_T(t_0+t)}{R_T(t_0)}$$

$$R_T(t\mid t_0)=\frac{\dfrac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-(t_0+t)^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)(t_0+t)\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}{\dfrac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-t^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)t\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}$$

$$R_T(t\mid t_0)=\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-(t_0+t)^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)(t_0+t)\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-t^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)t\right)} \tag{16}$$

3.3 Hazard function

The hazard function of WTMI distribution is given by

$$H_T(t)=\frac{g(t;\varepsilon,\nu,\omega,s)}{1-G_T(t)}$$

$$H_T(t)=\frac{\dfrac{(s+\varepsilon)(s+2\varepsilon)t^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\dfrac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}{1-\dfrac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}t^{s+\varepsilon}-2\omega(s+\varepsilon)t^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}$$

After simplification we have

$$H_T(t)=\frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^{\varepsilon}\,t^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\dfrac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-t^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)t^{\varepsilon}\right)} \tag{17}$$

3.4 Cumulative hazard function

The cumulative hazard function of WTMI distribution is given by

$${}_C H_T(t)=-\ln\left(R_T(t)\right)$$

$${}_C H_T(t)=-\ln\left(\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-t^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)t\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right) \tag{18}$$

Similarly, the Conditional Cumulative hazard function of WTMI distribution is given by

$${}_C H_T(t\mid t_0)=-\ln\left(R_T(t\mid t_0)\right)$$

$${}_C H_T(t\mid t_0)=-\ln\left(\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-(t_0+t)^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)(t_0+t)\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-t^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)t\right)}\right)$$

3.5 Reverse Hazard function

The reverse hazard function of WTMI distribution is given by

$$H_r(t)=\frac{g(t;\varepsilon,\nu,\omega,s)}{G_T(t)}$$

$$H_r(t)=\frac{\dfrac{(s+\varepsilon)(s+2\varepsilon)t^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\dfrac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}{\dfrac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}t^{s+\varepsilon}-2\omega(s+\varepsilon)t^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}$$

After simplification we get

$$H_r(t)=\frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^{\varepsilon}\,t^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}\,t^{s+\varepsilon}-2\omega(s+\varepsilon)t^{s+2\varepsilon}} \tag{19}$$

3.6 Mills Ratio

The Mills ratio of WTMI distribution is given by

$$Mills\ ratio=\frac{1}{H_r(t)}$$

$$Mills\ ratio=\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}\,t^{s+\varepsilon}-2\omega(s+\varepsilon)t^{s+2\varepsilon}}{(s+\varepsilon)(s+2\varepsilon)(\nu)^{\varepsilon}\,t^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)} \tag{20}$$

3.7 Mean residual life

The mean residual life (MRL) in case of WTMI distribution is given by

$$MRL=\frac{1}{1-G_Z(z)}\int_z^{\nu} t\,g(t;\varepsilon,\nu,\omega,s)dt-z$$

$$MRL=\frac{1}{1-G_Z(z)}\int_z^{\nu} t\,\frac{(s+\varepsilon)(s+2\varepsilon)t^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dt-z \tag{21}$$

$$MRL=\frac{1}{1-G_Z(z)}\int_z^{\nu} \frac{(s+\varepsilon)(s+2\varepsilon)t^{s+\varepsilon}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dt-z \tag{22}$$

$$MRL=\frac{(s+\varepsilon)(s+2\varepsilon)}{\left(1-G_Z(z)\right)\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_z^{\nu}t^{s+\varepsilon}dt-\frac{2\omega}{(\nu)^{\varepsilon}}\int_z^{\nu}t^{s+2\varepsilon}dt\right)-z \tag{23}$$

After simplification we get

$$MRL=\frac{(s+\varepsilon)(s+2\varepsilon)}{\left(1-G_Z(z)\right)\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+1)(s+2\varepsilon+1)}\left\{\begin{matrix}(\nu)^{s+2\varepsilon+1}\left((1-\omega)(s+\varepsilon+1)+(1+\omega)\varepsilon\right)+\\ 2\omega(s+\varepsilon+1)z^{s+2\varepsilon+1}-(\nu)^{\varepsilon}(1+\omega)(s+2\varepsilon+1)z^{s+\varepsilon+1}\end{matrix}\right\}-z$$

4. Moments

The rth raw moment about origin of WTMI distributionis defined as

$$\mu_r'=\int_0^{\nu} z^r\,g(z;\varepsilon,\nu,\omega,s)dz$$

$$\mu_r' = \int_0^{\nu} z^r \frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz$$

$$\mu_r' = \frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_0^{\nu} z^{s+\varepsilon+r-1}dz - \frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\nu} z^{s+2\varepsilon+r-1}dz\right) \tag{24}$$

$$\mu_r' = \frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left((1+\omega)\frac{(\nu)^{s+\varepsilon+r}}{s+\varepsilon+r} - \frac{2\omega(\nu)^{s+2\varepsilon+r}}{(\nu)^{\varepsilon}(s+2\varepsilon+r)}\right) \tag{25}$$

After simplification we get

$$\mu_r' = \frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^r\left((1-\omega)(s+\varepsilon+r)+(1+\omega)\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+r)(s+2\varepsilon+r)} \tag{26}$$

Putting $r = 1, 2, 3, 4$ in (6) we get

$$\mu_1' = \frac{(s+\varepsilon)(s+2\varepsilon)(\nu)\left((1-\omega)(s+\varepsilon+1)+(1+\omega)\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+1)(s+2\varepsilon+1)} \tag{27}$$

$$\mu_2' = \frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^2\left((1-\omega)(s+\varepsilon+2)+(1+\omega)\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+2)(s+2\varepsilon+2)} \tag{28}$$

$$\mu_3' = \frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^3\left((1-\omega)(s+\varepsilon+3)+(1+\omega)\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+3)(s+2\varepsilon+3)} \tag{29}$$

$$\mu_4' = \frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^4\left((1-\omega)(s+\varepsilon+4)+(1+\omega)\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+4)(s+2\varepsilon+4)} \tag{30}$$

The variance and coefficient of variance $(C.V)$ respectively are given by

$$\sigma^2 = \mu_2' - (\mu_1')^2$$

and

$$C.V = \frac{\sigma}{\mu_1'}; \quad \text{where,} \quad \sigma = \sqrt{\mu_2' - (\mu_1')^2}$$

5. Harmonic mean

The harmonic mean of WTMI distribution can be obtained as

$$Harmonic\,mean = E\left(\frac{1}{Z}\right)$$

$$Harmonic\,mean = \int_0^{\nu} \frac{1}{z}\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz$$

$$Harmonic\,mean = \int_0^{\nu} \frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-2}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz \tag{31}$$

$$Harmonic\,mean = \frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_0^{\nu} z^{s+\varepsilon-2}dz - \frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\nu} z^{s+2\varepsilon-2}dz\right) \tag{32}$$

After simplification we get

$$Harmonic\,mean = \frac{(s+\varepsilon)(s+2\varepsilon)\left((1-\omega)(s+\varepsilon-1)+(1+\omega)\right)}{\nu\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon-1)(s+2\varepsilon-1)} \tag{33}$$

6. MGF, CF and CGF

The MGF of WTMI distributionis equal to

$$M_Z(t) = E(e^{tz})$$

$$M_Z(t) = \int_0^{\nu} e^{tz}\,\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz$$

$$M_Z(t) = \int_0^{\nu}\sum_{k=0}^{\infty}\frac{(tz)^k}{k!}\,\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz \tag{34}$$

$$M_Z(t) = \sum_{k=0}^{\infty}\frac{(t)^k}{k!}\int_0^{\infty} z^k g(z;\varepsilon,\nu,\omega,s)dz$$

$$M_Z(t) = \sum_{k=0}^{\infty}\frac{(t)^k}{k!}\mu_k'$$

$$M_Z(t) = \sum_{k=0}^{\infty}\frac{(t)^k}{k!}\,\frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^k\left((1-\omega)(s+\varepsilon+k)+(1+\omega)\,\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+k)(s+2\varepsilon+k)} \tag{35}$$

The CF of WTMI distribution can be obtained as

$$\phi_Z(t) = E(e^{itz})$$

$$\phi_Z(t) = \int_0^{\nu} e^{\iota tz}\,\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz$$

$$\phi_Z(t) = \int_0^{\nu}\sum_{k=0}^{\infty}\frac{(\iota tz)^k}{k!}\,\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}dz \tag{36}$$

$$\phi_Z(t) = \sum_{k=0}^{\infty}\frac{(\iota)^k(t)^k}{k!}\int_0^{\infty} z^k g(z;\varepsilon,\nu,\omega,s)dz \tag{37}$$

$$\phi_Z(t) = \sum_{k=0}^{\infty}\frac{(\iota\,t)^k}{k!}\mu_k'$$

$$\phi_Z(t)=\sum_{k=0}^{\infty}\frac{(\imath t)^k}{k!}\frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^k\left((1-\omega)(s+\varepsilon+k)+(1+\omega)\,\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+k)(s+2\varepsilon+k)} \tag{38}$$

The CGF of WTMI distribution is given by

$$\kappa_Z(t)=\log\left(M_Z(t)\right)$$

$$\kappa_Z(t)=\log\left(\sum_{k=0}^{\infty}\frac{(t)^k}{k!}\frac{(s+\varepsilon)(s+2\varepsilon)(\nu)^k\left((1-\omega)(s+\varepsilon+k)+(1+\omega)\,\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+k)(s+2\varepsilon+k)}\right) \tag{39}$$

7. Estimation of Parameters

Let $z_1, z_2, z_3, \ldots, z_n$ be a random sample of size n from WTMI distribution. Then The likelihood function is defined as the joint density of the random sample, which is given as

$$L(\varepsilon,\nu,\omega,s)=\prod_{l=1}^{n}g\left(z_l;\varepsilon,\nu,\omega,s\right)=\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{40}$$

$$L(\varepsilon,\nu,\omega,s)=\prod_{l=1}^{n}\frac{(s+\varepsilon)(s+2\varepsilon)z_l^{\,s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z_l}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{41}$$

$$L(\varepsilon,\nu,\omega,s)=\frac{(s+\varepsilon)^n\,(s+2\varepsilon)^n}{\nu^{n(s+\varepsilon)}\left(s(1-\omega)+2\varepsilon\right)^n}\left(\prod_{l=1}^{n}z_l^{\,s+\varepsilon-1}\right)\left(\prod_{l=1}^{n}\left(1+\omega-2\omega\left(\frac{z_l}{\nu}\right)^{\varepsilon}\right)\right) \tag{42}$$

Taking logarithm on both sides we get

$$\log L(\varepsilon,\nu,\omega,s)=n\log(s+\varepsilon)+n\log(s+2\varepsilon)-n(s+\varepsilon)\log(\nu)-n\log\left(s(1-\omega)+2\varepsilon\right)$$
$$+(s+\varepsilon-1)\sum_{l=1}^{n}\log z_l+\sum_{l=1}^{n}\log\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right) \tag{43}$$

Differentiating equation (43) partially with respect to ε and equating to zero we get

$$\frac{n}{s+\varepsilon}+\frac{2n}{s+2\varepsilon}-n\log(\nu)-\frac{2n}{s(1-\omega)+2\varepsilon}+\sum_{l=1}^{n}\log z_l-\sum_{l=1}^{n}\frac{2\omega}{\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right)}\left(\frac{z}{\nu}\right)^{\varepsilon}\log\left(\frac{z}{\nu}\right)=0 \tag{44}$$

Differentiating equation (43) partially with respect to ν and equating to zero we get

$$\sum_{l=1}^{n}\frac{2\,\omega\,(z_l)\,\varepsilon}{\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right)(\nu)^{\varepsilon+1}}-\frac{n(s+\varepsilon)}{\nu}=0 \tag{45}$$

Differentiating equation (43) partially with respect to ω and equating to zero we get

$$\frac{n\,s}{s(1-\omega)+2\varepsilon}+\sum_{l=1}^{n}\frac{1-2\left(\frac{z_l}{\nu}\right)^{\varepsilon}}{\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right)}=0 \tag{46}$$

Differentiating equation (43) partially with respect to s and equating to zero we get

$$\frac{n}{s+\varepsilon}+\frac{n}{s+2\varepsilon}-n\log(\nu)-\frac{n(1-\omega)}{s(1-\omega)+2\varepsilon}+\sum_{l=1}^{n}\log z_l=0 \tag{47}$$

On solving equation (44), (45), (46), and (47) simultaneously, we obtain the maximum likelihood estimators of parameters involved in the given distribution. However, the above system of non-linear equations cannot be evaluated directly. So, to get the maximum likelihood estimates for the distribution parameters, we have to solve these system of equations using Newton-Raphson method, Mathematica, or Secant method.

8. Distribution of ordered statistics

Let $z_1, z_2, z_3, ..., z_n$ be a random sample of size n from WTMI distribution. Then $Z_{(1)}, Z_{(2)}, Z_{(3)}, ..., Z_{(n)}$ be the ordered statistics associated with the given sample such that $Z_{(1)} \le Z_{(2)} \le Z_{(3)} \le ... \le Z_{(n)}$, Where

$Z_{(1)} = \min(z_1, z_2, z_3, ..., z_n)$ and $Z_{(n)} = \max(z_1, z_2, z_3, ..., z_n)$

The probability density function of k^{th} ordered statistics from the given distribution is given by

$$g_{Z_{(k)}}(z)=\frac{n!}{(k-1)!(n-k)!}g(z;\varepsilon,\nu,\omega,s)\left(G_Z(z)\right)^{k-1}\left(1-G_Z(z)\right)^{n-k}$$

$$g_{Z_{(k)}}(z)=\frac{n!}{(k-1)!(n-k)!}\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left(\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{k-1}$$
$$\times\left(1-\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n-k} \tag{48}$$

$$g_{Z_{(k)}}(z)=\frac{n!}{(k-1)!(n-k)!}\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left(\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{k-1}$$
$$\times\left(\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-z^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)z^{\varepsilon}\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n-k} \tag{49}$$

And the corresponding cumulative distribution function of k^{th} ordered statistics is

$$G_{Z_{(k)}}(z)=\sum_{j=k}^{n}\left({}^{n}C_{j}\right)\left(G_Z(z)\right)^{j}\left(1-G_Z(z)\right)^{n-j}$$

$$G_{Z_{(k)}}(z)=\sum_{j=k}^{n}\left(\begin{array}{l}\left({}^{n}C_{j}\right)\left(\dfrac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{j}\\ \times\left(\dfrac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-z^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)z^{\varepsilon}\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n-j}\end{array}\right) \tag{50}$$

On substituting $k=1, n$ in equation (49) we get the probability density functions of smallest and highest ordered statistics respectively and are given as

$$g_{Z_{(1)}}(z)=n\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\times\left(\frac{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)-z^{s+\varepsilon}\left((s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}-2\omega(s+\varepsilon)z^{\varepsilon}\right)}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n-1} \tag{51}$$

and

$$g_{Z_{(n)}}(z)=\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left(\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n-1} \tag{52}$$

Their corresponding cumulative density functions are obtained on substituting $k=1,n$ in equation (50) and are given by

$$G_{Z_{(1)}}(z)=1-\left(1-\left(\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)\right)^{n} \tag{53}$$

And

$$G_{Z_{(n)}}(z)=\left(\frac{(s+2\varepsilon)(1+\omega)(\nu)^{\varepsilon}z^{s+\varepsilon}-2\omega(s+\varepsilon)z^{s+2\varepsilon}}{\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n} \tag{54}$$

9. Likelihood ratio test

In the context of probability distributions, the likelihood ratio test is often employed to compare whether two distributions adequately describe the observed data or not. Suppose $z_1, z_2, z_3, ..., z_n$ be a random sample of size n from WTMI distribution.. To test the hypothesis

$H_0: g(z)=g(z;\varepsilon,\nu,\omega,)$ against $H_1: g(z)=g(z;\varepsilon,\nu,\omega,s)$

The likelihood ratio test is defined as

$$\ell=\prod_{k=1}^{n}\frac{g(z_k;\varepsilon,\nu,\omega,s)}{f(z_k;\varepsilon,\nu,\omega)}$$

$$\ell=\prod_{k=1}^{n}\frac{(s+\varepsilon)(s+2\varepsilon)z_k{}^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z_k}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}}{\frac{\varepsilon}{\nu^{\varepsilon}}z_k{}^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z_k}{\nu}\right)^{\varepsilon}\right)} \tag{55}$$

$$\ell=\left(\frac{(s+\varepsilon)(s+2\varepsilon)}{\varepsilon\nu^{s}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n}\prod_{k=1}^{n}z_k{}^{s} \tag{56}$$

So, we reject null hypothesis at α level of significance if $\ell>\mathrm{K}^{*}$ such that $P(\ell>\mathrm{K}^{*})=\alpha$, where K^{*} is the critical value at α level of significance of the given test statistics. That is,

$$\left(\frac{(s+\varepsilon)(s+2\varepsilon)}{\varepsilon\nu^{s}\left(s(1-\omega)+2\varepsilon\right)}\right)^{n}\prod_{k=1}^{n}z_k{}^{s}>\mathrm{K}^{*} \tag{57}$$

$$\prod_{k=1}^{n}z_k{}^{s}>\mathrm{K}^{*}\left(\frac{\varepsilon\nu^{s}\left(s(1-\omega)+2\varepsilon\right)}{(s+\varepsilon)(s+2\varepsilon)}\right)^{n} \tag{58}$$

For large sample size n, $-2\log(\ell)$ is distributed as Chi-square distribution with one degree of freedom. Also p-value is calculated from the chi-square distribution. On the basis of p-value, we reject the null hypothesis when the p-value is less than level of significance.

10. Entropy measures

10.1 Renyi entropy and Tsallis entropy

By definition, the Renyi entropy is given by

$$R(\tau)=\frac{1}{1-\tau}\log\left(\int_0^{\nu}\left(g\left(z_k;\varepsilon,\nu,\omega,s\right)\right)^{\tau}dz\right) \tag{59}$$

$$R(\tau)=\frac{1}{1-\tau}\log\left(\int_0^{\nu}\left(\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{\tau}dz\right) \tag{60}$$

$$R(\tau)=\frac{1}{1-\tau}\log\left(\left(\frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{\tau}\int_0^{\nu}\left(z^{\tau(s+\varepsilon-1)}\sum_{k=0}^{\tau}\left({}^{\tau}C_k\right)(-1)^k(1+\omega)^{\tau-k}\left(2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)^k\right)dz\right)$$

$$R(\tau)=\frac{1}{1-\tau}\log\left(\left(\frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\right)^{\tau}\sum_{k=0}^{\tau}\left({}^{\tau}C_k\right)(-1)^k(1+\omega)^{\tau-k}\left(\frac{2\omega}{(\nu)^{\varepsilon}}\right)^k\int_0^{\nu}z^{\varepsilon k+\tau(s+\varepsilon-1)}dz\right) \tag{61}$$

After simplification we get

$$R(\tau)=\frac{1}{1-\tau}\log\left(\frac{\left((s+\varepsilon)(s+2\varepsilon)\right)^{\tau}(\nu)^{1-\tau}}{\left(s(1-\omega)+2\varepsilon\right)^{\tau}}\sum_{k=0}^{\tau}\frac{\left({}^{\tau}C_k\right)(-1)^k(1+\omega)^{\tau-k}(2\omega)^k}{\varepsilon k+\tau(s+\varepsilon-1)+1}\right) \tag{62}$$

Similarly, the Tsallis entropy associated with the given distribution is given by

$$T_s(\xi)=\frac{1}{\xi-1}\left(1-\int_0^{\nu}\left(g\left(z_k;\varepsilon,\nu,\omega,s\right)\right)^{\xi}dz\right)$$

$$T_s(\xi)=\frac{1}{\xi-1}\left(1-\frac{\left((s+\varepsilon)(s+2\varepsilon)\right)^{\xi}(\nu)^{1-\xi}}{\left(s(1-\omega)+2\varepsilon\right)^{\xi}}\sum_{k=0}^{\xi}\frac{\left({}^{\xi}C_k\right)(-1)^k(1+\omega)^{\xi-k}(2\omega)^k}{\varepsilon k+\xi(s+\varepsilon-1)+1}\right) \tag{63}$$

11. Bonferroni and Lorenz curves

The Bonferroni curve of the given distribution is given by

$$\Psi(\zeta)=\frac{1}{\zeta\,\mu_1'}\int_0^{\varphi}z\,g\left(z;\varepsilon,\nu,\omega,s\right)dz$$

Where $\mu_1'=\dfrac{(s+\varepsilon)(s+2\varepsilon)(\nu)\left((1-\omega)(s+\varepsilon+1)+(1+\omega)\varepsilon\right)}{\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+1)(s+2\varepsilon+1)}$ and $\varphi=F^{-1}(\zeta)$

$$\Psi(\zeta)=\frac{1}{\zeta\,\mu_1'}\int_0^{\varphi}z\,\frac{(s+\varepsilon)(s+2\varepsilon)z^{s+\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)} \tag{64}$$

$$\Psi(\zeta)=\frac{1}{\zeta\ \mu_1'}\frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_0^{\varphi}z^{s+\varepsilon}dz-\frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\varphi}z^{s+2\varepsilon}dz\right) \tag{65}$$

$$\Psi(\zeta)=\frac{1}{\zeta\ \mu_1'}\frac{(s+\varepsilon)(s+2\varepsilon)}{\nu^{s+\varepsilon}\left(s(1-\omega)+2\varepsilon\right)}\left((1+\omega)\left(\frac{(\varphi)^{s+\varepsilon+1}}{s+\varepsilon+1}\right)-\frac{2\omega}{(\nu)^{\varepsilon}}\left(\frac{(\varphi)^{s+2\varepsilon+1}}{s+2\varepsilon+1}\right)\right) \tag{66}$$

After simplification we get

$$\Psi(\zeta)=\frac{(s+\varepsilon)(s+2\varepsilon)(\varphi)^{s+\varepsilon+1}\left((1+\omega)(s+2\varepsilon+1)(\nu)^{\varepsilon}-2\omega(s+\varepsilon+1)(\varphi)^{\varepsilon}\right)}{\zeta\ \mu_1'\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+1)(s+2\varepsilon+1)} \tag{67}$$

Also, the Lorenz curve of the given distribution is given by

$$\Phi(\zeta)=\zeta\ \Psi(\zeta)$$

$$\Phi(\zeta)=\zeta\left(\frac{(s+\varepsilon)(s+2\varepsilon)(\varphi)^{s+\varepsilon+1}\left((1+\omega)(s+2\varepsilon+1)(\nu)^{\varepsilon}-2\omega(s+\varepsilon+1)(\varphi)^{\varepsilon}\right)}{\zeta\ \mu_1'\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+1)(s+2\varepsilon+1)}\right) \tag{68}$$

$$\Phi(\zeta)=\frac{(s+\varepsilon)(s+2\varepsilon)(\varphi)^{s+\varepsilon+1}\left((1+\omega)(s+2\varepsilon+1)(\nu)^{\varepsilon}-2\omega(s+\varepsilon+1)(\varphi)^{\varepsilon}\right)}{\mu_1'\nu^{s+2\varepsilon}\left(s(1-\omega)+2\varepsilon\right)(s+\varepsilon+1)(s+2\varepsilon+1)} \tag{69}$$

12. Conclusion

In this research paper, we have explored an innovative extension of the Transmuted Mukherjee-Islam distribution, known as the weighted Transmuted Mukherjee-Islam distribution. This distribution is formulated by incorporating a weighted model, utilizing the three-parameter Transmuted Mukherjee-Islam distribution as the base distribution. We thoroughly examine and discuss the newly introduced weighted Transmuted Mukherjee-Islam distribution, exploring its mathematical and statistical properties. The parameters of this novel distribution are determined through the application of maximum likelihood estimation techniques.

References

[1] Abdullahi, U. K. and Ieren, T. G. (2018). On the inferences and applications of transmuted exponential Lomax distribution, *International Journal of Advanced Statistics and Probability*, 6:30–36, doi:10.14419/ijasp.v6i1.8129.

[2] Afify, A. Z., Hamedani, G. G., Ghosh, I., and Mead, M. E. (2015a). The transmuted Marshall-OlkinFrechet distribution: Properties and applications, *International Journal of Statistics and Probability*, 4:132–148.

[3] Ahmad, A., Ahmad, S. P., and Ahmad, A. (2014). Transmuted inverse Rayleigh distribution: A generalization of the inverse Rayleigh distribution, *Mathematical Theory and Modeling*, 4:90–98.

[4] Al-Zou'bi, L. M. (2017). Transmuted Mukherjee-Islam Distribution: A Generalization of Mukherjee-Islam Distribution, *Journal of Mathematics Research*, 9(4), 135-144. https://doi.org/10.5539/jmr.v9n4p135.

[5] Aryal, G. R. and Tsokos, C. P. (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution, *European Journal of Pure and Applied Mathematics*, 4:89–102.

[6] Cordeiro, G. M., Saboor, A., Khan, M. N., Provost, S. B., and Ortega, E. M. M. (2017). The transmuted generalized modified weibull distribution, *Filomat*, 31:1395–1412, doi:10.2298/FIL1705395C.

[7] Das, S., Kundu, D. (2016). On Weighted Exponential Distribution and Its Length Biased Version, *Journal of Indian Society of Probability and Statistics*, 17, 57-69. https://doi.org/10.1007/s41096-016-0001-9.

[8] Dey, S., Ali, S., & Park, C. (2015). Weighted exponential distribution: properties and different methods of estimation, *Journal of Statistical Computation and Simulation*, 85(18), 3641-3661.

[9] Elgarhy, M., Rashed, M., and Shawki, A. W. (2016). Transmuted generalized lindley distribution, *International Journal of Mathematics Trends and Technology*, 29:145–154, doi:10.14445/22315373/IJMTT–V29P520.

[10] Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, 6, 13- 25.

[11] K. Jain; N. Singla; R. D. Gupta ; A weighted version of gamma distribution, *Discussiones Mathematicae: Probability and Statistics*,34 89-111 (2014).

[12] Khan, M. S., King, R., and Hudson, I. L. (2019). Transmuted Burr Type X Distribution with Covariates Regression Modeling to Analyze Reliability Data, *American Journal of Mathematical and Management Sciences*, doi:10.1080/01966324.2019.1605320.

[13] Kilany, N. M. (2016). Weighted Lomax distribution, *SpringerPlus*, 5(1). doi:10.1186/s40064-016-3489-2.

[14] M. Ghitany; F. Alqallaf; D. Al-Mutairi;H. Husain; A two-parameter weighted Lindley distribution and its applications to survival data, *Mathematics and Computers in Simulation*, 81(6) 1190-1201 (2011). doi:10.1016/j.matcom.2010.11.005.

[15] Para, B. A., & Jan T. R., (2018), On Three Parameter Weighted Pareto Type II Distribution: Properties and Applications in Medical Sciences, *Applied Mathematics & Information Sciences Letters*, 6(1),13-26.

[16] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; *Pergamum Press and Statistical Publishing Society, Calcutta*. 320-332.

[17] Rather, A. A. & Subramanian C. (2019). On weighted sushila distribution with properties and its applications, *International Journal of Scientific Research In Mathematical and Statistical Science*, vol 6, issue. 1, pp. 105-117.

[18] Rather, A. A. and Ozel, G.(2020). The Weighted Power Lindley Distribution with Applications on the Life Time Data, *Pakistan Journal of Statistics and Operation Research*, Vol. 16, No. 2, pp 225-237.

[19] 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,* Vol-IX, Issue VI, pp 2175-2184.

[20] Rather, A. A., Subramanian, C., Shafi, S., Malik, K. A., Ahmad, P. J., Para, B. A. and Jan, T. R. (2018), A new Size Biased Distribution with applications in Engineering and Medical Science, *International Journal of Scientific Research in Mathematical and Statistical Sciences,* Vol.5, Issue.4, pp.75-85.

[21] Rather, A. A., and Subramanian, C., (2018), Transmuted Mukherjee-Islam failure model, *Journal of Statistics Applications & Probability*, 7(2), 343-347.

[22] Shaw, W.T. and Buckley, I.R.C. (2007). The alchemy of probability distributions: beyond gram-charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. *Research report.*

[23] Subramanian C. & Rather, A. A. (2018). Weighted exponentiated mukherjee-islam distribution, *International Journal of Management, Technology and Engineering*, vol 8, issue XI, pp. 1328-1339.

BAYESIAN ESTIMATION OF $P(X \leq Y)$ FOR POWER-SERIES MODEL

Soumik Halder[1], Sudhansu S. Maiti[2] and Mriganka Mouli Chowdhury[2]

•

[1]Department of Statistics, Vivekananda Mahavidyalaya
Purba Bardhaman-713103, West Bengal, India
[2]Department of Statistics, Visva-Bharati University
Santiniketan-731 235, West Bengal, India
soumikhalder987@gmail.com, dssm1@rediffmail.com, mmc.uttarpara@gmail.com

Abstract

When the probability distributions for the stress (X) and strength (Y) are different members of the power series family, the expressions of the stress-strength reliability function, $R = P(X \leq Y)$, are derived. Apart from stress-strength reliability, it has applications in statistical tolerancing, measurement of demand-supply system performance, genetic trait hereditary measure, bio-equivalence study, etc. The Bayes' estimates of R under squared error and Precautionary losses are derived for various combinations of distributions of X and Y like binomial, Poisson, negative binomial, and geometric. As in practice, the availability of prior parameters is difficult; the empirical Bayes estimation procedure has been adopted to get their estimates from observed data. Simulation results have been reported, and estimates of posterior risks are compared. In the context of real Soccer games, the Bayes estimates are enumerated and compared with their classical counterparts.

Keywords: Empirical Bayes' estimate, Estimated Posterior risk, Precautionary loss, Squared error loss, Stress-strength reliability.

1. Introduction

It is a satisfactory fact that the strength of a manufactured product is a variable quantity. When ascertaining the reliability of equipment or the viability of a material, it is also necessary to consider the stress conditions of the operating environment. The uncertainty of the stressful environment leads us to take stress as a random component. In the stress-strength model, X is the stress applied on the unit by the operating system, and Y is the unit's strength, which is the in-built capacity of the unit to withstand the applied stress. No doubt, any unit can perform its actual function if its strength is greater than the stress given to it. In this context, we consider the reliability (R) as

$$R \quad = \quad P(X \leq Y),$$

which is the probability that the unit performs its task satisfactorily. Also, we can say this is the probability of the unit overcoming the stress. In this paper, the estimation of R when X and Y are independently distributed but not necessarily identical follows the power-series distribution. The quantity R has many applications, for example, statistical tolerance, measurement of demand-supply system performance, stress-strength reliability, genetic trait hereditary measure, a study of bio-equivalence, etc.

Several authors explore different statistical properties concerning R for several members of

the power series distributions. A study on the various estimation methods of R when X and Y are independently distributed geometric random variables was made by [11]. [2] worked the Bayes estimation of it. A study on different estimations of the negative binomial distribution was considered by [8] and [15] in the context of a system reliability estimation. [5] and [4] used Poisson distribution to explore different estimates of R. [13] studied the application of log series distribution.

In most of the papers, the Bayes estimate for the stress-strength problems was found for the continuous distributions, like [16] found the estimate of R for the Gompertz case. [14] found for Weibull distribution, [1] used type-II censoring for Rayleigh distribution for finding Bayes estimate. [12] used exponential-Poisson distribution. [9] used generalized exponential distribution, and [7] used inverted gamma distribution to estimate it using the Bayes method.

In the current work, we consider the Bayesian estimate of R for the stress and strength distributions as general members of the power series family of distributions. Section 2 starts with the exact expression of R for the generalized form of power series for both X and Y. The Bayesian estimates of R for power series distributions under squared error loss and precautionary loss, as well as some prerequisites, are discussed in section 3. The Bayesian estimates and their estimated posterior risks for particular choices of distributions like binomial, Poisson, negative binomial, and geometric of X and Y are discussed in section 4. Section 5 is devoted to searching estimates of the hyper-parameters of the prior distributions and hence is engaged in finding the empirical Bayes estimates of R. Simulation study results have been reported in section 6. An application of R in real soccer games is discussed in section 7, where we find the R estimate with the estimated posterior risk. Section 8 draws concluding remarks.

2. The initial set-up

Let X and Y be two random variables belonging to the power-series family of distributions, which are defined as follows.

$$P(X=x)=\frac{a(x)\alpha^x}{f(\alpha)}\,, x=0,1,... \quad \alpha>0 \text{ and } f(\alpha)=\sum_{x=0}^{\infty} a(x)\alpha^x. \tag{1}$$

$$P(Y=y)=\frac{b(y)\beta^y}{g(\beta)}\,, y=0,1,... \quad \beta>0 \text{ and } g(\beta)=\sum_{x=0}^{\infty} b(y)\beta^y. \tag{2}$$

The stress-strength reliability R is given by

$$R=P(X\leq Y)=\frac{1}{f(\alpha)g(\beta)}\sum_{y=0}^{\infty}\Big\{\sum_{x=0}^{y}a(x)\alpha^x\Big\}b(y)\beta^y. \tag{3}$$

Let $x_1, x_2, \ldots, x_{n_1}$ be a sample of size n_1 from the distribution of X and $y_1, y_2, \ldots, y_{n_2}$ be a sample of size n_2 from that of Y. Then, the likelihood function is

$$\begin{aligned}
L(\alpha,\beta|\underline{x},\underline{y}) &= \prod_{i=1}^{n_1}P(X=x_i)\prod_{j=1}^{n_2}P(Y=y_j)\\
&= \frac{\alpha^{\sum_{i=1}^{n_1}x_i}}{f^{n_1}(\alpha)}\prod_{i=1}^{n_1}a(x_i)\frac{\beta^{\sum_{j=1}^{n_2}y_j}}{g^{n_2}(\beta)}\prod_{j=1}^{n_2}b(y_j)\\
&= \alpha^{t_x}\{f(\alpha)\}^{-n_1}\prod_{i=1}^{n_1}a(x_i)\beta^{t_y}\{g(\beta)\}^{-n_2}\prod_{j=1}^{n_2}b(y_j)\\
&= L(\alpha|\underline{x})L(\beta|\underline{y}),
\end{aligned}$$

where $t_x = \sum_{i=1}^{n_1} x_i$, $t_y = \sum_{j=1}^{n_2} y_j$.
The joint posterior density function of α, β corresponding to prior distributions h(α), k(β) is

$$\prod(\alpha, \beta | t_x, t_y) = \frac{L(\alpha, \beta | t_x, t_y) h(\alpha) k(\beta)}{\int_0^\infty \int_0^\infty L(\alpha, \beta | t_x, t_y) h(\alpha) k(\beta) d\alpha d\beta} = \frac{L(\alpha | t_x) h(\alpha)}{\int_0^\infty L(\alpha | t_x) h(\alpha) d\alpha} . \frac{L(\beta | t_y) k(\beta)}{\int_0^\infty L(\beta | t_y) k(\beta)}.$$

3. Bayes estimation of R for power-series distributions

The main objective of this section is to find Bayes' estimates of R, the stress-strength reliability. We have considered two different loss functions, and the Bayes estimates corresponding to each loss function are given for the general cases. The joint distributions, the prior distributions of the parameters, and some preliminaries, which are required to find the Bayes estimates of the different combinations of the power-series model, are discussed.

3.1. The Bayes Estimates under different loss functions

The squared error loss (SEL) and Pre-cautionary loss (PL) functions have been considered. If d is the estimate of the parameter θ, the SEL is given by

$$L(d, \theta) = (d - \theta)^2. \tag{4}$$

Under the SEL, the posterior mean is the Bayes estimate of the parameter θ and is denoted by

$$d = \hat{\theta} = E(\theta).$$

The Posterior risk, in this case, is

$$PR_{SEL} = E(\theta^2) - E^2(\theta).$$

The PL is defined as

$$L(\theta, d) = \frac{(d - \theta)^2}{d}.$$

The Bayes estimate under the PL is defined as

$$d = \hat{\theta} = \sqrt{E(\theta^2)}.$$

The Posterior risk, in this case, is

$$PR_{PL} = 2[\sqrt{E(\theta^2)} - E(\theta)].$$

So, the Bayes estimate of R under the SEL is given by

$$\begin{aligned} \hat{R}_{SEL} &= E(R | t_x, t_y) \\ &= \int_0^\infty \int_0^\infty \sum_{y=0}^{\infty} \sum_{x=0}^{y} \frac{a(x) b(y) \alpha^x \beta^y}{f(\alpha) g(\beta)} \prod(\alpha, \beta | t_x, t_y) d\alpha d\beta \\ &= \sum_{y=0}^{\infty} \sum_{x=0}^{y} \frac{\int_0^\infty \frac{a(x)\alpha^x}{f(\alpha)} L(\alpha) h(\alpha) d\alpha}{\int_0^\infty L(\alpha) h(\alpha) d\alpha} \frac{\int_0^\infty \frac{b(y)\beta^y}{g(\beta)} L(\beta) k(\beta) d\beta}{\int_0^\infty L(\beta) k(\beta) d\beta}, \end{aligned} \tag{5}$$

and the same under the PL is $\hat{R}_{PL} = \sqrt{E(R^2|t_x, t_y)}$, where

$$\begin{aligned} E(R^2|t_x,t_y) &= \int_0^\infty \int_0^\infty [\sum_{y=0}^{\infty} \sum_{x=0}^{y} \frac{a(x)b(y)\alpha^x\beta^y}{f(\alpha)g(\beta)}]^2 \prod(\alpha,\beta|t_x,t_y)d\alpha d\beta \\ &= \int_0^\infty \int_0^\infty \sum_{y=0}^{\infty} \sum_{x=0}^{y} \frac{a^2(x)\alpha^{2x}}{f^2(\alpha)} \frac{b^2(y)\beta^{2y}}{g^2(\beta)} \prod(\alpha,\beta|t_x,t_y)d\alpha d\beta + \\ &\quad \int_0^\infty \int_0^\infty \sum_{y_j=0}^{\infty} \sum_{x_i=0}^{y_j} \sum_{y_l=0}^{\infty} \sum_{x_k=0}^{y_l} \frac{a(x_i)a(x_k)\alpha^{x_i+x_k}}{f^2(\alpha)} \frac{b(y_j)b(y_l)\beta^{y_j+y_l}}{g^2(\beta)} \prod(\alpha,\beta|t_x,t_y)d\alpha d\beta. \end{aligned} \tag{6}$$

The estimated posterior risks of Bayes estimate of R under SEL and PL are $\hat{PR}_{SEL} = \hat{E}(R^2|t_x,t_y) - \hat{E}^2(R|t_x,t_y)$ and $\hat{PR}_{PL} = 2[\sqrt{\hat{E}(R^2|t_x,t_y)} - \hat{E}(R|t_x,t_y)]$, respectively.

3.2. Some particular distributions and joint distributions

Four different distributions have been considered; all belong to the power-series family. These are binomial (m,p), Poisson (λ), negative binomial (r,η), and geometric (μ) distributions. Let z_1,z_2,...,z_n be a sample from the distribution of Z. Then, t_z=$\sum_{i=0}^{n} z_i$ be the complete sufficient statistic for estimating the distribution parameter. The probability mass function (pmf) of different distributions and their joint distributions for a sample of size n are shown in Table 1.

Table 1: *The pmfs of the different distributions and their joint distributions*

Distribution (Z)	pmf, P(Z=z)	Joint distribution
Bin(m,p)	P(Z=z)=$\binom{m}{p}p^z(1-p)^{m-z}$	$p^{t_z}(1-p)^{mn-t_z}$
Poisson(λ)	P(Z=z)=$\frac{e^{-\lambda}\lambda^z}{z!}$	$\frac{e^{-n\lambda}\lambda^{t_z}}{\prod_{i=1}^{n} z_i!}$
Negative binomial (r,η)	P(Z=z)=$\binom{r+z-1}{z}\eta^z(1-\eta)^r$	$\eta^{t_z}(1-\eta)^{nr}$
Geometric (μ)	P(Z=z)=$\mu^z(1-\mu)$	$\mu^{t_z}(1-\mu)^n$

3.3. Prior distributions and their probability distribution functions

Each distribution characteristic depends on the value of the single parameter it evolves. The parameter itself follows a distribution. The probability distribution function (pdf) of the prior parameters and some forms related to the posterior distribution of the parameters have been found, and presented in Table 2.

Table 2: *The pdfs of the prior parameters and some forms related to the posterior distribution*

Distribution	Prior distribution	pdf of prior dist	Terms required for posterior dist
Bin(m,p)	Beta(α,β)	$\frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$	$I_{bin}(m,p)$= $\frac{p^{t_z+\alpha-1}(1-p)^{mn-t_z+\beta-1}}{B(t_z+\alpha,mn-t_z+\beta)}$
Poisson(λ)	Gamma(γ,δ)	$\frac{\delta^\gamma e^{-\gamma\lambda}\lambda^{\delta-1}}{\Gamma\delta}$	$I_{pois}(\lambda)$= $\frac{(n+\gamma)^{(t_z+\delta)}e^{-(n+\gamma)\lambda}\lambda^{t_z+\delta-1}}{\Gamma(t_z+\delta)}$
Neg bin (r,η)	Beta(c,d)	$\frac{\eta^{c-1}(1-\eta)^{d-1}}{B(c,d)}$	$I_{neg}(\eta)$= $\frac{\eta^{t_z+c-1}(1-\eta)^{nr+d-1}}{B(t_z+c,nr+d)}$
Geometric (μ)	Beta(a,b)	$\frac{\mu^{a-1}(1-\mu)^{b-1}}{B(a,b)}$	$I_{geo}(\mu)$= $\frac{\mu^{t_z+a-1}(1-\mu)^{n+b-1}}{B(t_z+a,n+b)}$

3.4. Some preliminaries

The expression of R is given in 3. The expression for Bayes estimates under SEL and PL are shown in 5 and 6, respectively. To find the explicit forms, we require some terms defined in this

section and presented in Tables 3 and 4. These are to be used for finding E(R) and E(R^2).

Table 3: *Prerequisite for finding E(R) and E(R^2)*

Distribution	Terms related to E(R) (R_{dist})	Terms for R^2 ($z_i = z_k$) (Rsq_{dist})
Bin(m,p)	$\binom{m}{z}\frac{B(t_z+z+\alpha,m(n+1)-t_z-z+\beta)}{B(t_z+\alpha,mn-t_z+\beta)}$	$\binom{m}{z}^2\frac{B(2z+t_z+\alpha,2m-2z+n_1m-t_z+\beta)}{B(t_z+\alpha,mn_1-t_z+\beta)}$
Poisson(λ)	$\frac{(n+\gamma)^{(t_z+\delta)}}{z!\Gamma(t_z+\delta)}\frac{\Gamma(t_z+z+\delta)}{(n+\gamma+1)^{(t_z+z+\delta)}}$	$\frac{(n_2+\gamma)^{(t_z+\delta)}\Gamma 2z+t_z+\delta}{(2+n_2+\delta)^{(2z+t_z+\delta)}\Gamma t_z+\delta}$
Negative binomial (r,η)	$\binom{r+z-1}{z}\frac{B(t_z+z+c,nr+r+d)}{B(t_z+c,nr+d)}$	$\binom{r+z-1}{y}^2\frac{B(2z+t_z+c,2r+n_2r+d)}{B(t_z+c,n_2r+d)}$
Geometric (μ)	$\frac{B(t_z+z+a,n+1+b)}{B(t_z+a,n+b)}$	$\frac{B(t_z+2z+a,n_2+b+2)}{B(t_z+a,n_2+b)}$

Table 4: *Prerequisite for finding E(R) and E(R^2)*

Distribution	Terms for R^2 ($z_j \neq z_k$) ($Rcov_{dist}$)
Bin(m,p)	$\binom{m}{z_i}\binom{m}{z_j}\frac{B(z_i+z_j+t_z+\alpha,2n-z_i-z_j+n_1m-t_z+\beta)}{B(t_z+\alpha,mn_1-t_z+\beta)}$
Poisson(λ)	$\frac{(n_2+\gamma)^{(t_z+\delta)}\Gamma(z_i+z_j+t_z+\delta)}{(2+n_2+\delta)^{(z_i+z_j+t_z+\delta)}\Gamma t_z+\delta}$
Negative binomial (r,η)	$\binom{r+z_i-1}{z_i}\binom{r+z_j-1}{z_j}\frac{B(z_i+z_j+t_z+c,2r+n_2r+d)}{B(t_z+c,n_2r+d)}$
Geometric(μ)	$\frac{B(t_z+z_i+z_j+a,n_2+b+2)}{B(t_z+a,n_2+b)}$

4. Theoretical expression of R for different stress-strength models

The theoretical expressions of R for general power series stress-strength models have been derived earlier. Such expressions are of theoretical interest because specific members of the power series family are used for modeling stress and/or strength. Being restricted to the family of power series distributions, we provide simplified expressions of such quantities for several stress and strength distribution choices. This section found the posterior distributions for different combinations of the parameters attached to the different distributions. In each subsection, we have mentioned the pmfs of the random variables used for stress and strength, the prior distributions of the parameters involved in the pmfs, the joint posterior distribution of the parameters, the $E(R|t_x,t_y)$ and $E(R^2|t_x,t_y)$ that are required for the Bayes estimates $\hat{R}$.

4.1. X and Y both follow binomial distributions

Let X$\sim$binomial(m_1,p_1) and Y$\sim$binomial(m_2,p_2), where p_1 $\sim$beta(α_1,β_1) and p_2 $\sim$beta(α_2,β_2). The joint prior distribution of p_1 and p_2 is

$$\pi(p_1,p_2) = g(p_1).h(p_2).$$

The posterior distribution of p_1 and p_2 is given by

$$\prod(p_1,p_2|t_x,t_y) = I_{bin}(m_1,p_1)I_{bin}(m_2,p_2). \tag{7}$$

The Bayes estimate under SEL is given by

$$\begin{aligned}
\hat{R}_{SEL} &= E(R_{BB}|t_x,t_y)\\
&= \int_0^1\int_0^1\sum_{y=0}^{m_2}\sum_{x=0}^{min(y,m_1)}\binom{m_1}{x}p_1^x(1-p_1)^{m_1-x}\binom{m_2}{y}p_2^y(1-p_2)^{m_2-y}\prod(p_1,p_2|t_x,t_y)dp_1dp_2\\
&= \sum_{y=0}^{m_2}\sum_{x=0}^{min(y,m_1)}R_{bin}(m_1,p_1)R_{bin}(m_2,p_2). \qquad (8)
\end{aligned}$$

The Bayes estimate under PL is $\hat{R}_{PL}= \sqrt{E(R^2)}$, where $E(R^2)$ is given by

$$\begin{aligned} E(R^2_{BB}|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{m_2}\sum_{x=0}^{min(y,m_1)} \binom{m_1}{x} p_1^x(1-p_1)^{m_1-x}\binom{m_2}{y}p_2^y(1-p_2)^{m_2-y}\right]^2 \\ &\quad \times \prod(p_1,p_2|t_x,t_y)dp_1dp_2 \\ &= \sum_{y=0}^{m_2}\sum_{x=0}^{min(y,m_1)} Rsq_{bin}(m_1,p_1)Rsq_{bin}(m_2,p_2) \\ &\quad + \sum_{y_j=0}^{m_2}\sum_{x_i=0}^{min(y_j,m_1)}\sum_{y_l=0}^{m_2}\sum_{x_k=0}^{min(y_l,m_1)} Rcov_{bin}(m_1,p_1)Rcov_{bin}(m_2,p_2). \end{aligned}$$

The estimated posterior risks are as follows.

$$\hat{PR}_{SEL} = \hat{E}(R^2_{BB}) - \hat{E}^2(R_{BB}).$$
$$\hat{PR}_{PL} = 2\big(\sqrt{\hat{E}(R^2_{BB})} - \hat{E}(R_{BB})\big).$$

4.2. X and Y both follow Poisson distributions

Let X∼Poisson(λ_1) and Y∼Poisson(λ_2), where λ_1 ∼Gamma(γ_1,δ_1) and λ_2 ∼Gamma(γ_2,δ_2).
The joint prior distribution of λ_1 and λ_2 is

$$\pi(\lambda_1,\lambda_2) = g(\lambda_1).h(\lambda_2).$$

The posterior distribution of λ_1 and λ_2 is given by

$$\prod(\lambda_1,\lambda_2|t_x,t_y) = I_{pois}(\lambda_1)I_{pois}(\lambda_2).$$

Hence, under the SEL

$$\begin{aligned} E(R_{PP}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{m_2}\sum_{x=0}^{y} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2}\lambda_2^y}{y!}\prod(\lambda_1,\lambda_2|t_x,t_y)d\lambda_1d\lambda_2 \\ &= \sum_{y=0}^{m_2}\sum_{x=0}^{y} R_{pois}(\lambda_1)R_{pois}(\lambda_2), \end{aligned}$$

and under the PL

$$\begin{aligned} E(R^2_{PP}|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{m_2}\sum_{x=0}^{y} \frac{e^{-\lambda_1}\lambda_1^x}{x!}\frac{e^{-\lambda_2}\lambda_2^y}{y!}\right]^2\prod(\lambda_1,\lambda_2|t_x,t_y)d\lambda_1d\lambda_2 \\ &= \sum_{y=0}^{m_2}\sum_{x=0}^{y} Rsq_{pois}(\lambda_1)Rsq_{pois}(\lambda_2) + \sum_{y_j=0}^{m_2}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{m_2}\sum_{x_k=0}^{y_l} Rcov_{pois}(\lambda_1)Rcov_{pois}(\lambda_2). \end{aligned}$$

4.3. X and Y both follow negative binomial distributions

Let X∼negative binomial(r_1,η_1) and Y∼negative binomial(r_2,η_2), where η_1 ∼beta(c_1,d_1) and η_2 ∼beta(c_2,d_2).
The joint prior distribution of η_1 and η_2 is

$$\pi(\eta_1,\eta_2) = g(\eta_1).h(\eta_2).$$

The posterior distribution of η_1 and η_2 is given by

$$\prod(\eta_1,\eta_2|t_x,t_y) = I_{neg}(\eta_1)I_{neg}(\eta_2).$$

Therefore, under SEL

$$\begin{aligned} E(R_{NN}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{r_1+x-1}{x}\eta_1^x(1-\eta_1)_1^r\binom{r_2+y-1}{y}\eta_2^y(1-\eta_2)_2^r \\ &\quad \times\prod(\eta_1,\eta_2|t_x,t_y)d\eta_1 d\eta_2 \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{neg}(\eta_1)R_{neg}(\eta_2), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{NN}^2|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{r_1+x-1}{x}\eta_1^x(1-\eta_1)_1^r\binom{r_2+y-1}{y}\eta_2^y(1-\eta_2)_2^r\right]^2 \\ &\quad \times\prod(\eta_1,\eta_2|t_x,t_y)d\eta_1 d\eta_2 \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} Rsq_{neg}(\eta_1)Rsq_{neg}(\eta_2) + \sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l} Rcov_{neg}(\eta_1)Rcov_{neg}(\eta_2). \end{aligned}$$

4.4. X and Y both follow geometric distributions

Let X∼geometric(μ_1) and Y∼geometric(μ_2), where μ_1 ∼beta(a_1, b_1) and μ_2 ∼beta(c_2, d_2).
The joint prior distribution of μ_1 and μ_2 is

$$\pi(\mu_1,\mu_2) = g(\mu_1).h(\mu_2).$$

The posterior distribution of μ_1 and μ_2 is given by,

$$\prod(\mu_1,\mu_2|t_x,t_y) = I_{geo}(\mu_1)I_{geo}(\mu_2).$$

Therefore, under SEL

$$\begin{aligned} E(R_{GG}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{\infty}\sum_{x=0}^{y}\mu_1^x(1-\mu_1)\mu_2^y(1-\mu_2)\prod(\mu_1,\mu_2|t_x,t_y)d\mu_1 d\mu_2 \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{geo}(\mu_1)R_{geo}(\mu_2), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{GG}^2|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\mu_1^x(1-\mu_1)\mu_2^y(1-\mu_2)\right]^2\prod(\mu_1,\mu_2|t_x,t_y)d\mu_1 d\mu_2 \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} Rsq_{geo}(\mu_1)Rsq_{geo}(\mu_2) + \sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l} Rcov_{geo}(\mu_1)Rcov_{geo}(\mu_2). \end{aligned}$$

4.5. X follows binomial distribution and Y follows Poisson distribution

Let X∼binomial(m,p) and Y∼Poisson(λ), where p∼beta(α, β) and λ ∼gamma(γ, δ).
The joint prior distribution of p and λ is

$$\pi(p,\lambda) = g(p).h(\lambda).$$

The posterior distribution of p and λ is given by

$$\prod(p,\lambda|t_x,t_y) = I_{bin}(p)I_{pois}(\lambda).$$

Under SEL

$$\begin{aligned} E(R_{BP}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{m}{x}p^x(1-p)^{m-x}\frac{e^{-\lambda}\lambda^y}{y!}\prod(p,\lambda|t_x,t_y)dpd\lambda \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{bin}(p)R_{pois}(\lambda), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{BP}^2|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{m}{x}p^x(1-p)^{m-x}\frac{e^{-\lambda}\lambda^y}{y!}\right]^2\prod(p,\lambda|t_x,t_y)dpd\lambda \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} Rsq_{bin}(p)Rsq_{pois}(\lambda) + \sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l} Rcov_{bin}(p)Rcov_{pois}(\lambda). \end{aligned}$$

4.6. X follows binomial distribution and Y follows negative binomial distribution

Let X∼binomial(m,p) and Y∼negative binomial(r,η), where p∼beta(α,β) and η ∼beta(c,d).
The joint prior distribution of p and η is

$$\pi(p,\eta) = g(p).h(\eta).$$

The posterior distribution of p and η is

$$\prod(p,\eta|t_x,t_y) = I_{bin}(p)I_{neg}(\eta).$$

So, under SEL

$$\begin{aligned} E(R_{BN}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{m}{x}p^x(1-p)^{m-x}\binom{r+y-1}{y}\eta^y(1-\eta)^r\prod(p,\eta|t_x,t_y)dpd\eta \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{bin}(p)R_{neg}(\eta), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{BN}^2|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{m}{x}p^x(1-p)^{m-x}\binom{r+y-1}{y}\eta^y(1-\eta)^r\right]^2\prod(p,\eta|t_x,t_y)dpd\eta \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} Rsq_{bin}(p)Rsq_{neg}(\eta) + \sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l} Rcov_{bin}(p)Rcov_{neg}(\eta). \end{aligned}$$

4.7. X follows binomial distribution and Y follows geometric distribution

Let X∼binomial(m,p) and Y∼geometric (μ), where p∼beta(α,β) and μ ∼beta(a,b).
The joint prior distribution of p and μ is

$$\pi(p,\mu) = g(p).h(\mu).$$

The posterior distribution of p and μ is given by

$$\prod(p,\mu|t_x,t_y) = I_{bin}(p)I_{geo}(\mu).$$

Under SEL,

$$\begin{aligned} E(R_{BG}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{m}{x}p^x(1-p)^{m-x}\mu^y(1-\mu)\prod(p,\mu|t_x,t_y)dpd\mu \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{bin}(p)R_{geo}(\mu), \end{aligned}$$

and under PL,

$$\begin{aligned} E(R_{BG}^2|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{m}{x}p^x(1-p)^{m-x}\mu^y(1-\mu)\right]^2\prod(p,\mu|t_x,t_y)dpd\mu \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} Rsq_{bin}(p)Rsq_{geo}(\mu) + \sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l} Rcov_{bin}(p)Rcov_{geo}(\mu). \end{aligned}$$

4.8. X follows Poisson distribution and Y follows binomial distribution

Let X∼Poisson(λ) and Y∼binomial (n,p), where λ ∼gamma(δ,γ) and p ∼beta(α,β).
The joint prior distribution of λ and p is

$$\pi(\lambda,p) = g(\lambda).h(p).$$

The posterior distribution of λ and p is given by

$$\prod(\lambda,p|t_x,t_y) = I_{pois}(\lambda)I_{bin}(p).$$

Therefore, under SEL

$$\begin{aligned} E(R_{PB}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{m}\sum_{x=0}^{y}\frac{e^{-\lambda}\lambda^x}{x!}\binom{m}{y}p^y(1-p)^{m-y}\prod(\lambda,p|t_x,t_y)d\lambda dp \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{pois}(\lambda)R_{bin}(p), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{PB}^2|t_x,t_y) &= \int_0^1\int_0^1 \left[\sum_{y=0}^{m}\sum_{x=0}^{y}\frac{e^{-\lambda}\lambda^x}{x!}\binom{m}{y}p^y(1-p)^{m-y}\right]^2\prod(\lambda,p|t_x,t_y)d\lambda dp \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} Rsq_{pois}(\lambda)Rsq_{bin}(p) + \sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l} Rcov_{pois}(\lambda)Rcov_{bin}(p). \end{aligned}$$

4.9. X follows Poisson distribution and Y follows negative binomial distribution

Let X∼Poisson(λ) and Y∼negative binomial (r,η), where λ ∼gamma(δ,γ) and η ∼beta(c,d).
The joint prior distribution of λ and η is

$$\pi(\lambda,\eta) = g(\lambda).h(\eta).$$

The posterior distribution of λ and η is

$$\prod(\lambda,\eta|t_x,t_y) = I_{pois}(\lambda)I_{neg}(\eta).$$

So, under SEL

$$\begin{aligned} E(R_{PN}|t_x,t_y) &= \int_0^1\int_0^1 \sum_{y=0}^{\infty}\sum_{x=0}^{y}\frac{e^{-\lambda}\lambda^x}{x!}\binom{r+y-1}{y}\eta^y(1-\eta)^r\prod(\lambda,\eta|t_x,t_y)d\lambda d\eta \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y} R_{pois}(\lambda)R_{neg}(\eta), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{PN}^2|t_x,t_y) &= \int_0^1\int_0^1\left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\frac{e^{-\lambda}\lambda^x}{x!}\binom{r+y-1}{y}\eta^y(1-\eta)^r\right]^2\prod(\lambda,\eta|t_x,t_y)d\lambda d\eta \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{pois}(\lambda)Rsq_{neg}(\eta)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{pois}(\lambda)Rcov_{neg}(\eta). \end{aligned}$$

4.10. X follows Poisson distribution and Y follows geometric distribution

Let X∼Poisson(λ) and Y∼geometric (μ), where λ ∼gamma(δ,γ) and μ ∼beta(a,b).
The joint prior distribution of λ and μ is

$$\pi(\lambda,\mu) = g(\lambda).h(\mu).$$

The posterior distribution of λ and μ is

$$\prod(\lambda,\mu|t_x,t_y) = I_{pois}(\lambda)I_{geo}(\mu).$$

Under SEL

$$\begin{aligned} E(R_{PG}|t_x,t_y) &= \int_0^1\int_0^1\sum_{y=0}^{\infty}\sum_{x=0}^{y}\frac{e^{-\lambda}\lambda^x}{x!}\mu^y(1-\mu)\prod(\lambda,\mu|t_x,t_y)d\lambda d\mu \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}R_{pois}(\lambda)R_{geo}(\mu), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{PG}^2|t_x,t_y) &= \int_0^1\int_0^1\left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\frac{e^{-\lambda}\lambda^x}{x!}\mu^y(1-\mu)\right]^2\prod(\lambda,\mu|t_x,t_y)d\lambda d\mu \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{pois}(\lambda)Rsq_{geo}(\mu)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{pois}(\lambda)Rcov_{geo}(\mu). \end{aligned}$$

4.11. X follows negative binomial distribution and Y follows binomial distribution

Let X∼negative binomial(r,η) and Y∼binomial (n,p), where η ∼beta(c,d) and p ∼beta(α,β).
The joint prior distribution of η and p is

$$\pi(\eta,p) = g(\eta).h(p).$$

The posterior distribution of η and p is given by

$$\prod(\eta,p|t_x,t_y) = I_{neg}(\eta)I_{bin}(p).$$

So, under SEL

$$\begin{aligned} E(R_{NB}|t_x,t_y) &= \int_0^1\int_0^1\sum_{y=0}^{m}\sum_{x=0}^{y}\binom{r+x-1}{x}\eta^x(1-\eta)^r\binom{m}{y}p^y(1-p)^{m-y}\prod(\eta,p|t_x,t_y)d\eta dp \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}R_{neg}(\eta)R_{bin}(p), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{NB}^2|t_x,t_y) &= \int_0^1\int_0^1\left[\sum_{y=0}^{m}\sum_{x=0}^{y}\binom{r+x-1}{x}\eta^x(1-\eta)^r\binom{m}{y}p^y(1-p)^{m-y}\right]^2 \\ &\quad\times\prod(\eta,p|t_x,t_y)d\eta dp \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{neg}(\eta)Rsq_{bin}(p)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{neg}(\eta)Rcov_{bin}(p). \end{aligned}$$

4.12. X follows negative binomial distribution and Y follows Poisson distribution

Let X∼negative binomial(r,η) and Y∼Poisson (λ), where η ∼beta(c,d) and λ ∼gamma(δ,γ). The joint prior distribution of η and λ is

$$\pi(\eta,\lambda) = g(\eta).h(\lambda).$$

The posterior distribution of η and λ is given by

$$\prod(\eta,\lambda|t_x,t_y) = I_{neg}(\eta)I_{pois}(\lambda).$$

Under SEL

$$\begin{aligned} E(R|t_x,t_y) &= \int_0^1\int_0^1\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{r+x-1}{x}\eta^x(1-\eta)^r\frac{e^{-\lambda}\lambda^y}{y!}\prod(\eta,\lambda|t_x,t_y)d\eta d\lambda \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}R_{neg}(\eta)R_{pois}(\lambda), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{NP}^2|t_x,t_y) &= \int_0^1\int_0^1\left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{r+x-1}{x}\eta^x(1-\eta)^r\frac{e^{-\lambda}\lambda^y}{y!}\right]^2\prod(\eta,\lambda|t_x,t_y)d\eta d\lambda \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{neg}(\eta)Rsq_{pois}(\lambda)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{neg}(\eta)Rcov_{pois}(\lambda). \end{aligned}$$

4.13. X follows negative binomial distribution and Y follows geometric distribution

Let X∼negative binomial(r,η) and Y∼geometric (μ), where η ∼beta(c,d) and μ ∼beta(a,b). The joint prior distribution of η and μ is

$$\pi(\eta,\mu) = g(\eta).h(\mu).$$

The posterior distribution of η and μ is given by

$$\prod(\eta,\mu|t_x,t_y) = I_{neg}(\eta)I_{geo}(\mu).$$

Under SEL

$$\begin{aligned} E(R_{NG}|t_x,t_y) &= \int_0^1\int_0^1\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{r+x-1}{x}\eta^x(1-\eta)^r\mu^y(1-\mu)\prod(\eta,\mu|t_x,t_y)d\eta d\mu \\ &= \sum_{y=0}^{\infty}\sum_{x=0}^{y}R_{neg}(\eta)R_{geo}(\mu), \end{aligned}$$

and under PL

$$E(R^2_{NG}|t_x,t_y) = \int_0^1\int_0^1\left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\binom{r+x-1}{x}\eta^x(1-\eta)^r\mu^y(1-\mu)\right]^2\prod(\eta,\mu|t_x,t_y)d\eta d\mu$$
$$= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{neg}(\eta)Rsq_{geo}(\mu)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{neg}(\eta)Rcov_{geo}(\mu).$$

4.14. X follows geometric distribution and Y follows binomial distribution

X∼geometric(μ) and Y∼binomial (n,p), where μ ∼beta(a,b) and p ∼beta(α,β).
The joint prior distribution of μ and p is given by

$$\pi(\mu,p) = g(\mu).h(p).$$

The posterior distribution of μ and p is given by

$$\prod(\mu,p|t_x,t_y) = I_{geo}(\mu)I_{bin}(p).$$

Under SEL

$$E(R_{GB}|t_x,t_y) = \int_0^1\int_0^1\sum_{y=0}^{m}\sum_{x=0}^{y}\mu^x(1-\mu)^r\binom{m}{y}p^y(1-p)^{m-y}\prod(\mu,p|t_x,t_y)d\mu dp$$
$$= \sum_{y=0}^{\infty}\sum_{x=0}^{y}R_{geo}(\mu)R_{bin}(p),$$

and under PL

$$E(R^2_{GB}|t_x,t_y) = \int_0^1\int_0^1\left[\sum_{y=0}^{m}\sum_{x=0}^{y}\mu^x(1-\mu)^r\binom{m}{y}p^y(1-p)^{m-y}\right]^2\prod(\mu,p|t_x,t_y)d\mu dp$$
$$= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{geo}(\mu)Rsq_{bin}(p)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{geo}(\mu)Rcov_{bin}(p).$$

4.15. X follows geometric distribution and Y follows negative binomial distribution

Let X∼geometric(μ) and Y∼negative binomial (r,η), where μ ∼beta(a,b) and η ∼beta(c,d).
The joint prior distribution of μ and η is

$$\pi(\mu,\eta) = g(\mu).h(\eta).$$

The posterior distribution of μ and η is given by

$$\prod(\mu,\eta|t_x,t_y) = I_{geo}(\mu)I_{neg}(\eta).$$

Under SEL

$$E(R_{GN}|t_x,t_y) = \int_0^1\int_0^1\sum_{y=0}^{\infty}\sum_{x=0}^{y}\mu^x(1-\mu)^r\binom{r+y-1}{y}\eta^y(1-\eta)^r\prod(\mu,\eta|t_x,t_y)d\mu d\eta$$
$$= \sum_{y=0}^{\infty}\sum_{x=0}^{y}R_{geo}(\mu)R_{neg}(\eta),$$

and under PL

$$E(R^2_{GN}|t_x,t_y) = \int_0^1\int_0^1\left[\sum_{y=0}^{\infty}\sum_{x=0}^{y}\mu^x(1-\mu)^r\binom{r+y-1}{y}\eta^y(1-\eta)^r\right]^2\prod(\mu,\eta|t_x,t_y)d\mu d\eta$$
$$= \sum_{y=0}^{\infty}\sum_{x=0}^{y}Rsq_{geo}(\mu)Rsq_{neg}(\eta)+\sum_{y_j=0}^{\infty}\sum_{x_i=0}^{y_j}\sum_{y_l=0}^{\infty}\sum_{x_k=0}^{y_l}Rcov_{geo}(\mu)Rcov_{neg}(\eta).$$

4.16. X follows geometric distribution and Y follows Poisson distribution

X∼geometric (μ) and Y∼Poisson (λ), where μ ∼beta(a,b) and λ ∼gamma(δ, γ).
The joint prior distribution of μ and λ is

$$\pi(\mu, \lambda) = g(\mu).h(\lambda).$$

The posterior distribution of μ and λ is given by

$$\prod(\mu, \lambda | t_x, t_y) = I_{geo}(\mu) I_{pois}(\lambda).$$

Under SEL

$$\begin{aligned} E(R|t_x, t_y) &= \int_0^1 \int_0^1 \sum_{y=0}^{\infty} \sum_{x=0}^{y} \mu^x (1-\mu) \frac{e^{-\lambda} \lambda^y}{y!} \prod(\mu, \lambda | t_x, t_y) d\mu d\lambda \\ &= \sum_{y=0}^{\infty} \sum_{x=0}^{y} R_{geo}(\mu) R_{pois}(\lambda), \end{aligned}$$

and under PL

$$\begin{aligned} E(R_{GP}^2|t_x, t_y) &= \int_0^1 \int_0^1 \left[\sum_{y=0}^{\infty} \sum_{x=0}^{y} \mu^x (1-\mu) \frac{e^{-\lambda} \lambda^y}{y!} \right]^2 \prod(\mu, \lambda | t_x, t_y) d\mu d\lambda \\ &= \sum_{y=0}^{\infty} \sum_{x=0}^{y} Rsq_{geo}(\mu) Rsq_{pois}(\lambda) + \sum_{y_j=0}^{\infty} \sum_{x_i=0}^{y_j} \sum_{y_l=0}^{\infty} \sum_{x_k=0}^{y_l} Rcov_{geo}(\mu) Rcov_{pois}(\lambda). \end{aligned}$$

5. Estimation of the hyper-parameters

Though the prior distributions are assumed to be known, most of the time, in practice, those need to be estimated based on observed data. In this section, the estimates of the hyper-parameters of the prior distributions have been found. These parameters could be estimated using the empirical Bayes procedure [see [10] and [3]]. Given the observations, the joint likelihood distributions have been compared with the joint prior distributions. The joint likelihood distributions are just the multiplication of the likelihood distribution of X and Y, and joint prior distributions are the multiplication of the prior distributions of the parameters. We can estimate the prior parameters by comparing them individually with their corresponding likelihood functions. The estimates of the hyper-parameters are shown in Table 5.

Table 5: *Estimate of the hyper-parameters*

Distribution	Joint distribution	Prior distribution	Estimate of hyper-parameters
Binomial(m,p)	$p^{t_z}(1-p)^{mn-t_z}$	$\frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}$	$\hat{\alpha} = t_x + 1,\ \hat{\beta} = mn - t_x + 1$
Poisson(λ)	$\frac{e^{-n\lambda}\lambda^{t_z}}{\prod_{i=1}^{n} z_i!}$	$\frac{\delta^{\gamma} e^{-\gamma\lambda} \lambda^{\delta-1}}{\Gamma\delta}$	$\hat{\delta} = t_x + 1,\ \hat{\gamma} = n$
Neg. binomial (r, η)	$\eta^{t_z}(1-\eta)^{nr}$	$\frac{\eta^{c-1}(1-\eta)^{d-1}}{B(c,d)}$	$\hat{c} = t_x + 1,\ \hat{d} = nr + 1$
Geometric (μ)	$\mu^{t_z}(1-\mu)^{n}$	$\frac{\mu^{a-1}(1-\mu)^{b-1}}{B(a,b)}$	$\hat{a} = t_x + 1,\ \hat{b} = n + 1$

6. Simulation study

The properties of the Bayesian estimations of R for the different combinations of stress-strength models within the power series family have been explored empirically. The estimated posterior risks for those different combinations have been found.

Table 6: *Geometric-Geometric*

Sample size		Distribution of μ_1			Distribution of μ_2		Actual R	E(R)	$E(R^2)$	Bayes estimate		Posterior risk	
n_1	n_2	c	d	$E(\mu_1)$	a	b	$E(\mu_2)$				SEL	PL	SEL	PL
15	15	3	7	0.3	2	8	0.2	0.745	0.727	0.532	0.727	0.729	0.00411	0.00564
15	15	3	7	0.3	5	5	0.5	0.824	0.827	0.686	0.827	0.828	0.00217	0.00263
15	15	3	7	0.3	8	2	0.8	0.921	0.934	0.873	0.934	0.935	0.00046	0.00049
15	15	5	5	0.5	2	8	0.2	0.556	0.561	0.319	0.561	0.565	0.00418	0.00742
15	15	5	5	0.5	5	5	0.5	0.667	0.672	0.455	0.672	0.675	0.00383	0.00569
15	15	5	5	0.5	8	2	0.8	0.833	0.830	0.691	0.830	0.831	0.00192	0.00232
15	15	8	2	0.8	2	8	0.2	0.238	0.233	0.056	0.233	0.236	0.00153	0.00652
15	15	8	2	0.8	5	5	0.5	0.333	0.341	0.119	0.341	0.345	0.00277	0.00807
15	15	8	2	0.8	8	2	0.8	0.556	0.554	0.311	0.554	0.558	0.00384	0.00691
15	30	3	7	0.3	2	8	0.2	0.745	0.747	0.563	0.747	0.750	0.00393	0.00525
15	30	3	7	0.3	5	5	0.5	0.824	0.829	0.689	0.829	0.830	0.00173	0.00208
15	30	3	7	0.3	8	2	0.8	0.921	0.922	0.851	0.922	0.922	0.00043	0.00047
15	30	5	5	0.5	2	8	0.2	0.556	0.583	0.344	0.583	0.586	0.00410	0.00702
15	30	5	5	0.5	5	5	0.5	0.667	0.671	0.454	0.671	0.674	0.00338	0.00502
15	30	5	5	0.5	8	2	0.8	0.833	0.849	0.723	0.849	0.850	0.00126	0.00149
15	30	8	2	0.8	2	8	0.2	0.238	0.238	0.058	0.238	0.241	0.00142	0.00591
15	30	8	2	0.8	5	5	0.5	0.333	0.335	0.115	0.335	0.339	0.00239	0.00708
15	30	8	2	0.8	8	2	0.8	0.556	0.585	0.346	0.585	0.588	0.00315	0.00537
30	30	3	7	0.3	2	8	0.2	0.745	0.735	0.542	0.735	0.736	0.00201	0.00274
30	30	3	7	0.3	5	5	0.5	0.823	0.823	0.678	0.823	0.824	0.00131	0.00159
30	30	3	7	0.3	8	2	0.8	0.921	0.928	0.861	0.928	0.928	0.00029	0.00032
30	30	5	5	0.5	2	8	0.2	0.556	0.550	0.305	0.550	0.552	0.00216	0.00391
30	30	5	5	0.5	5	5	0.5	0.667	0.659	0.436	0.659	0.661	0.00201	0.00305
30	30	5	5	0.5	8	2	0.8	0.833	0.847	0.718	0.847	0.847	0.00076	0.00090
30	30	8	2	0.8	2	8	0.2	0.238	0.229	0.053	0.229	0.231	0.00074	0.00322
30	30	8	2	0.8	5	5	0.5	0.333	0.322	0.105	0.322	0.324	0.00139	0.00431
30	30	8	2	0.8	8	2	0.8	0.556	0.534	0.287	0.534	0.536	0.00203	0.00380

Table 7: *Negative binomial-Poisson*

Sample size		Distribution of η			Distribution of λ			Actual R	E(R)	$E(R^2)$	Bayes estimate		Posterior risk	
n_1	n_2	c	d	$E(\eta)$	δ	γ	$E(\lambda)$				SEL	PL	SEL	PL
15	15	3	7	0.3	2	10	0.2	0.039	0.032	0.001	0.032	0.034	$7.61 * 10^{-5}$	$2.30 * 10^{-3}$
15	15	3	7	0.3	5	10	0.5	0.058	0.054	0.003	0.054	0.056	$2.14 * 10^{-4}$	$3.90 * 10^{-3}$
15	15	3	7	0.3	8	10	0.8	0.079	0.080	0.007	0.080	0.083	$4.36 * 10^{-4}$	$5.36 * 10^{-3}$
15	15	5	5	0.5	2	10	0.2	0.162	0.156	0.025	0.156	0.160	$1.06 * 10^{-3}$	$6.70 * 10^{-3}$
15	15	5	5	0.5	5	10	0.5	0.218	0.214	0.048	0.214	0.218	$1.89 * 10^{-3}$	$8.75 * 10^{-3}$
15	15	5	5	0.5	8	10	0.8	0.273	0.283	0.083	0.283	0.289	$2.94 * 10^{-3}$	$1.03 * 10^{-2}$
15	15	8	2	0.8	2	10	0.2	0.570	0.592	0.355	0.592	0.596	$4.00 * 10^{-3}$	$6.74 * 10^{-3}$
15	15	8	2	0.8	5	10	0.5	0.645	0.639	0.413	0.645	0.642	$4.39 * 10^{-3}$	$6.85 * 10^{-3}$
15	15	8	2	0.8	8	10	0.8	0.707	0.693	0.484	0.693	0.696	$4.17 * 10^{-3}$	$6.00 * 10^{-3}$
15	30	3	7	0.3	2	10	0.2	0.039	0.043	0.002	0.043	0.044	$7.27 * 10^{-5}$	$1.64 * 10^{-3}$
15	30	3	7	0.3	5	10	0.5	0.058	0.054	0.003	0.054	0.056	$1.93 * 10^{-4}$	$3.48 * 10^{-3}$
15	30	3	7	0.3	8	10	0.8	0.079	0.075	0.006	0.075	0.077	$3.42 * 10^{-4}$	$4.51 * 10^{-3}$
15	30	5	5	0.5	2	10	0.2	0.162	0.163	0.028	0.163	0.167	$1.35 * 10^{-3}$	$8.19 * 10^{-3}$
15	30	5	5	0.5	5	10	0.5	0.218	0.245	0.062	0.245	0.249	$2.01 * 10^{-3}$	$8.15 * 10^{-3}$
15	30	5	5	0.5	8	10	0.8	0.273	0.273	0.077	0.273	0.277	$2.37 * 10^{-3}$	$8.60 * 10^{-3}$
15	30	8	2	0.8	2	10	0.2	0.570	0.578	0.338	0.578	0.581	$3.77 * 10^{-3}$	$6.51 * 10^{-3}$
15	30	8	2	0.8	5	10	0.5	0.645	0.642	0.416	0.642	0.645	$3.81 * 10^{-3}$	$5.91 * 10^{-3}$
15	30	8	2	0.8	8	10	0.8	0.707	0.716	0.516	0.716	0.718	$3.43 * 10^{-3}$	$4.78 * 10^{-3}$
30	30	3	7	0.3	2	10	0.2	0.039	0.040	0.0017	0.040	0.041	$5.61 * 10^{-5}$	$9.27 * 10^{-4}$
30	30	3	7	0.3	5	10	0.5	0.058	0.063	0.004	0.063	0.064	$1.43 * 10^{-4}$	$2.24 * 10^{-3}$
30	30	3	7	0.3	8	10	0.8	0.079	0.078	0.006	0.078	0.079	$2.11 * 10^{-4}$	$2.69 * 10^{-3}$
30	30	5	5	0.5	2	10	0.2	0.162	0.169	0.029	0.169	0.171	$6.16 * 10^{-4}$	$3.61 * 10^{-3}$
30	30	5	5	0.5	5	10	0.5	0.218	0.240	0.059	0.240	0.243	$1.27 * 10^{-3}$	$5.27 * 10^{-3}$
30	30	5	5	0.5	8	10	0.8	0.273	0.282	0.081	0.282	0.284	$1.41 * 10^{-3}$	$4.97 * 10^{-3}$
30	30	8	2	0.8	2	10	0.2	0.570	0.565	0.321	0.565	0.567	$2.02 * 10^{-3}$	$3.57 * 10^{-3}$
30	30	8	2	0.8	5	10	0.5	0.645	0.637	0.407	0.637	0.638	$2.19 * 10^{-3}$	$3.44 * 10^{-3}$
30	30	8	2	0.8	8	10	0.8	0.707	0.700	0.491	0.700	0.701	$2.17 * 10^{-3}$	$3.10 * 10^{-3}$

For computation, we draw $n_1(n_2)$ independent observations from the prior distribution(s) of the parameters of the stress (strength) distribution(s). For different combinations of those hyper-parameters of the stress (strength) distributions considered, the hyper-parameters are so chosen that the expected value of the parameters is equal to the values of the parameters selected combinations in [6]. We draw a random sample from the distribution(s) of the stress (strength) distribution(s). Then we compute $\hat{R_{SEL}}$, $\hat{R_{PL}}$ and their estimated posterior risk, where $\hat{R_{SEL}}$ and $\hat{R_{PL}}$ are the Bayesian estimates under squared error and precautionary loss functions. All these calculations are done using R-Software, and two representative tables are reported in Tables $6-7$ for space limitation. Some more tables are prepared and may be available from the corresponding author on request. The figures of the tables compared to [6]. We can say that the estimates under the Bayesian method are closer than those under MLE and UMVUE. Also, the estimated posterior risks under the empirical Bayesian method are smaller than the variance of UMVUE and the MSE of the MLE. So, we can infer that the Bayesian estimation gives better estimates than the MLE and UMVUE.

7. Real Life Data Analysis

We have provided a real dataset as an application of the stress-strength reliability model, which is related to soccer matches where the defenders and the goalkeeper are responsible for not allowing the opposition to score goals. They protect the team from continuous attacks from the opponent team. The number of goals conceded by the team can be treated as the stress put on the system's defence, whereas the number of goals saved by the defence acts as the strength of the team's defence. We want to estimate the reliability of the team's defence, i.e., $R = P(X \leq Y)$, where X is the stress on the system, and Y is the system's strength. We have considered the same dataset used in [6], where Manchester United played 38 matches against various teams of the EPL in the season 2017-2018. The data is presented in the Table 8.

Among those 38 matches, in 33 matches, they have saved more goals than they conceded. So,

Table 8: *EPL data for Manchester United during 2017-2018*

Match	Goals conceded	Goals saved	Match	Goals conceded	Goals saved
1	0	1	20	2	0
2	0	0	21	0	3
3	0	4	22	0	0
4	2	3	23	0	5
5	0	3	24	0	3
6	0	4	25	2	4
7	0	1	26	0	0
8	0	5	27	1	2
9	2	1	28	1	6
10	0	4	29	2	6
11	1	7	30	1	1
12	1	4	31	0	2
13	0	2	32	2	4
14	2	1	33	1	3
15	1	0	34	0	2
16	2	5	35	1	2
17	0	7	36	1	3
18	1	4	37	0	2
19	2	1	38	0	3

in about 86% cases, the defence system has worked above the stress given by the opposition. The reliability of the system is estimated using the theory defined above. The data are discrete.

The authors have shown that the "number of goals conceded" follows a geometric distribution, whereas the "number of goals saved" follows the Poisson distribution. We have found the estimates of R under SEL and PL as $\hat{R}_{SEL}$= 0.84683 and $\hat{R}_{PL}$= 0.84737 with estimated posterior risks $5.91 * 10^{-4}$ and $6.98 * 10^{-4}$, respectively.

8. Concluding remarks

In this article, the Bayesian estimation of $R = P(X \leq Y)$ has been considered when X and Y belong to a power-series family of distributions under two loss functions: (i) squared error loss, a symmetric and (ii) precautionary loss, an asymmetric one. The conjugate prior distribution(s) of the parameter(s) are chosen for deriving the Bayes estimates of R. Though the prior distributions are assumed to be known, the practice is not such. Generally, the prior distribution(s) parameters are estimated based on observations. The empirical Bayes method has been used to estimate the prior parameters and hence get the estimate of R. Simulation study results slightly favour the Bayes estimate of R under SEL than that under PL concerning estimated posterior risk sense. Data analysis result also affirms this. Scientists and practitioners are recommended to use the proposed Bayes estimate of R.

Declarations

Disclosure of Conflicts of Interest/Competing Interests: The authors declare no conflict of interest.
Authors contributions: Each author has an equal contribution. All authors jointly write, review, and edit the manuscript.
Funding: The authors received no specific funding for this study.
Data Availability Statements: All cited data analyzed in the article are included in References. Data sets are also provided in the article.
Ethical Approval: This article does not contain any studies with human participants performed by authors.
Code availability: Codes are available on request.

References

[1] Abu-Moussa, M., Abd-Elfattah, A. and Hossam, E. (2021): Estimation of Stress-Strength Parameter for Rayleigh Distribution based on Progressive Type-II Censoring, *Information Sciences Letters*, 10, 101-110: 10.18576/isl/100112.

[2] Ahmad, K. E. and and Fakhry, M. E. and Jaheen, Z. F. (1995): Bayes estimation of $P(Y \geq X))$ in the geometric case, *Microlectronic Reliability*, 35(5), 817-820.

[3] Awad, A. M. and Gharraf, M. K. (1986): Estimation of $P(Y < X)$ in the Burr case: a comparative study, *Communications in Statistics-Simulation and Computation*, 15(2), 389-403.

[4] Barbiero, A. (2013): Inference on reliability of stress-strength models for Poisson data, *Journal of Quality and Reliability Engineering*, 1: 10.1155/2013/530530.

[5] Belyaev, Y. and Lumelskii, Y. (1988): Multidimensional Poisson Walks, *Journal of Mathematical Sciences*: https://doi.org/10.1007/BF01085105.

[6] Choudhury, Mriganka and Bhattacharya, R. and Maiti, S. S. (2019): On estimating reliability function for the family of power series distribution, *Communications in Statistics - Theory and Methods*, 50(10), 1-30.

[7] Iranmanesh, A., Vajargah, K. and Hasanzadeh, M. (2018): On the estimation of stress strength reliability parameter of inverted gamma distribution, *Mathematical Sciences*, 12(3), 71-77.

[8] Ivshin, V. V. and Lumelskii, Ya, P.(1995): Statistical estimation problems in "Stress-Strength" models, *Perm University Press, Perm, Russia.*

[9] Li, C. and Hao, H. (2016): Likelihood and Bayesian Estimation in Stress Strength Model from Generalized Exponential distribution containing Outliers, *IAENG International Journal of Applied Mathematics*, 46(2), 155-159.

[10] Lindley, D. V. (1969): Introduction to Probability and Statistics from a Bayesian Viewpoint, *Cambridge University Press, Cambridge, UK*, 1

[11] Maiti, S. S. (1995): Estimation of $P(X \leq Y)$ in the Geometric case, *Journal of Indian Statistical Association* 3391), 87-91.

[12] Nadarajah, S., Bagheri J. K., Fazel, S., and Sangtarashani, M. A. and Samani, E. (2017): Estimation of the Stress Strength Parameter for the Generalized Exponential-Poisson Distribution, *Journal of Testing and Evaluation*, 46: 10.1520/JTE20160650.

[13] Obradovic, M., Jovanovic, M., Milosevic, B. and Jevremovic, V. (2015). Estimation of $P(XY)$ for Geometric-Poisson model, *Hacettepe Journal of Mathematics and Statistics*, 44(4), 949-964.

[14] Raqab, M. (2021): Estimation of stress-strength reliability $R = P(X > Y)$ based on Weibull record data in the presence of inter-record times, *AEJ - Alexandria Engineering Journal*: 10.1016/j.aje.2021.07.025.

[15] Sathe, Y. S. and Dixit, U. J. (2001): Estimation of $P(X \leq Y)$ in the negative binomial distribution, *Journal of Statistical Planning and Inference*, 93(1-2), 83-92: 10.1016/S0378-3758(00)00206-8.

[16] Saracoglu, B., Kaya, M. F. and Abd-Elfattah, A. M. (2009): Comparison of estimators for stress"strength reliability in Gompertz case, *Hacettepe Journal of Mathematics and Statistics*, 38, 339-349.

IDENTIFYING THE PROBABILITY DISTRIBUTION MODELS OF ECONOMIC LOSSES DUE TO NATURAL DISASTERS

Ashish Jha, Vikas Kumar Sharma, Abhimanyu Singh Yadav
•
Department of Statistics, Banaras Hindu University, Varanasi, India
vikasstats@rediffmail.com

Abstract

Natural catastrophes have a tremendous influence on the environment and our economy, which has raised significant concerns and spurred scientific research. Several studies have been done to model the economic losses brought on by natural disasters. In this article, we primarily concentrate on examining the distributions of economic losses resulting from big catastrophes including wildfires, earthquakes, droughts, volcanic eruptions, and harsh weather. We recommend utilizing five well-known statistical distributions, including the Weibull, Log-logistics, Gamma, Generalized Pareto, and Lognormal distributions since we observe the skewed forms of the empirical distributions. We employ the maximum likelihood technique for each distribution for the available data sets in order to estimate the distributions. The parameter estimations are numerically computed using the PSO method. We select the distribution that best fits the economic losses using the Akaike Information Criterion and Kolmogorov-Smirnov statistics. We discovered that the Log-logistic distribution is the distribution that fits the total economic losses caused by all-natural disasters the best.

Keywords: Natural catastrophes, Economic losses, Probability distribution models, Maximum likelihood estimation, PSO Method, R-software, Goodness-of-fit tests

I. Introduction

Nature has been giving us gifts since the beginning of time. But we have also had to deal with the terrible things about it. Every year, a large number of different natural disasters, including floods, wildfires, earthquakes, extreme heat, cold, and volcanic activity, claim the lives of on average 60,000 people. Direct and indirect effects are distinguished by a recognised typology of disaster effects [1]. The destruction of fixed assets, raw materials, natural resources, high-yielding crops, and the loss of priceless lives are examples of direct repercussions. Indirect effects, which are frequently referred to as economic losses, are those that have an impact on economic activity over time, particularly in the goods and services sectors [2].

According to EM-DAT, catastrophes caused 0.1% of fatalities in the previous two decades. High-impact incidents accounted for 0.1% to 0.4% of all fatalities. Flood and drought were the deadliest natural calamities, but they no longer kill many. Earthquakes are the deadliest nowadays. Along with life, calamities also destroy resources. These risks affect economic activity, causing volatility and losses for the global economy; see [1,2,3,4].

Natural disasters have increased dramatically over the last three decades, posing a significant threat to the world's economies, particularly those of developing countries. The impact of economic losses on developing countries is far greater than that on developed countries. Between 1970 and 2002, 6436 natural disasters occurred, with developing countries bearing the brunt of the damage. It

demonstrates that developing countries are unable to combat these deadly disasters due to a lack of resources [2,3]. The relationship between natural disasters and economic losses is widely evident all over the world, and a number of studies have prompted the need for disaster mitigation strategies to reduce human and economic suffering. For such studies, we refer the readers to follow [5,6].

The preceding discussion highlights the necessity of evaluating the distribution of economic losses caused by natural disasters. Estimating economic losses due to disasters was a huge difficulty in the early days, and it was dependent on a hypothetical or singular historical occurrence rather than mathematical or statistical modelling. Few studies have been done to calculate the economic damages incurred by various natural disasters. [7] estimated economic losses for the whole spectrum of extreme weather, such as draught and flood, by combining stochastic hydro-meteorological crop-loss models with a regionalized computable general equilibrium model. [8] estimated the economic losses caused by natural disasters using the input-output model and associated modelling frameworks such as the social accounting matrix and the computable general equilibrium. Furthermore, [9] introduced a novel modelling framework known as the regional input-output model to explore the effects of natural catastrophes.

Coronese et al. [10] estimated the damage and mortality caused by natural catastrophes using a quantile regression model. They discovered an increasing trend in extreme natural catastrophe damages, which is consistent with a climate-change signal. Natural catastrophe casualties have reduced, despite an increase in economic damages. They also noticed an alarming increase in casualties associated with severe temperatures. [4] proposed using extreme value theory for modelling economic losses in a monograph, and they employed extreme value and extended pareto distributions for fitting heavy tailed distributions of economic losses. [11] discovered that generalised extreme value models and generalised Pareto distributions match well to the extreme losses of natural disasters and are helpful tools for calculating the tails of loss severity distributions. [12] used a generalised Pareto distribution to describe economic damages resulting from non-natural disasters.

The majority of academicians generally support the use of generalised extreme value distributions to describe economic damages brought on by natural disasters. Only one or two natural disasters are modelled using probability distribution models. Numerous probability models that are available in the literature could match such datasets more accurately than the generalised Pareto distribution. In this article, we use different probabilistic models to fit the economic losses caused by six significant calamities: drought, earthquake, extreme weather, extreme temperature, wildfire, and volcanic activity. For each of these data sets, we look at five three-parameter statistical distributions (size, shape, and placement). The new research completely contradicts the studies under consideration and discovers a very suitable distribution for natural disasters. For the purpose of calculating numerical maximum likelihood estimates of the unidentified model parameters, we use the particle swarm optimization approach (PSO). We employ goodness-of-fit tests like Kolmogorov-Smirnov and Akaike Information Criterion to choose the probability distribution model that best fits the distribution of natural catastrophes.

This paper is organized into different sections. In the first section, we review the literature on the economic loos due to natural disasters. Which is then followed by a discussion on data and variables viz. drought, earthquake, extreme temperature, extreme weather, volcanic activity, and wildfire. Which is then knowledge about the three parametric distribution (viz. Weibull, Log-logistics, Gamma, Gen. Pareto, and Log-normal) and estimation techniques. Analysis of results follows which is finally concluded by the discussion of the results with respect to the objectives of the study.

II. Data description

The Pro Vention Consortium of the World Bank Catastrophe Management Facility launched a coordinated effort to review the quality, accuracy, and completeness of three global disaster data sets after realising the need for higher quality data to enhance disaster preparedness and mitigation. These were EM-DAT managed by the Centre for Research on the Epidemiology of Disasters (CRED), Sigma

maintained by Swiss Reinsurance Company (Zürich), and Nat Cat maintained by Munich Reinsura nce Company (Munich).

Over 22,000 mega catastrophes have occurred throughout the world since 1900, and EM-DAT provides crucial core data on their incidence and consequences. The database is created using data from a variety of sources, including UN agencies, non-governmental organisations, insurance firms, research institutions, and press outlets. Our catastrophe information is taken from the EM-DAT International Disaster Database, which CRED and the US Office for Foreign Disaster Assistance both administer (OFDA). A disaster is one that meets at least one of the following criteria, according to the database: 10 or more fatalities, 2000 or more people impacted by hunger and drought, 100 or more by other calamities, a government disaster declaration, or an appeal for outside help.

We examine economic losses caused by draughts, earthquakes, volcanic activity, harsh weather, high temperature, and wildfires. Minimum value, maximum value, mean, variance, standard deviation, coefficient of variation, skewness, and kurtosis of economic losses due to natural disasters were computed. Table 1 shows an overview of the descriptive statistics for each economic variable.

Table 1: *The descriptive statistics of economic losses due to the natural disasters*

Variable	Mean	Variance	Coefficient Variation	Skewness	Kurtosis
Draught	340.33	264018.7	150.98	2.66	8.0
Earthquake	981.81	8821100.0	302.51	5.73	38.6
Extreme Tem.	196.46	185152.3	219.03	3.92	16.3
Extreme Weather	1363.40	7919309.0	206.40	3.60	16.4
wildfire	200.53	141181.8	187.37	4.23	21.9
Volcanic Activity	14.05	666.4	183.74	2.48	5.6

According to the overview, between 1900 and 2018, the minimum damages attributable to natural disasters ranged from $0.02 billion to $0.10 billion. Volcanic activity provided the lowest minimum losses, but drought and wildfire caused the highest minimum losses. Natural disasters can cause maximum losses ranging from $100 to $23030 billion. Volcanic activity has the lowest maximum losses, whereas earthquakes have the highest possible losses. The mean or average of all disaster-related losses from 1900 to 2018 ranges from $14.05 to $1363.40 billion, with extreme weather events having the highest mean value and volcanic activity having the lowest mean value. The range of the standard deviation for all six variables is $25.8 to $297 billion. Volcanic activity has the lowest standard deviation while earthquakes have the highest. The fact that some effects created by extreme value to the huge value in the raw data makes it evident that the standard deviation for each variable is always larger than the mean. The degree of asymmetries in a distribution around the mean is determined using the coefficient of skewness. All distributions are positively skewed as the skewness value is more than zero and lies in between 2.48 and 5.73. The skewness value for an earthquake is the highest at 5.73, clearly showing that it is very obviously skewed and that its asymmetric tail is extending to the right, while the skewness value for volcanic activity is the lowest at 2.48, showing that it is less obviously skewed and that its symmetric tail is also extending to the right.

The relative peak or flatness of a distribution can be assessed using the value of kurtosis. Kurtosis values range from 5.6 to 38.6, and they are always greater than 3. All distributions of economic losses have a greater peak than the typical normal distribution. Earthquake's height value is 38.6, which suggests the likelihood of a leptokurtic distribution, in which the data set tends to have a prominent peak close to the mean and a heavy tail. Volcanic activity obtains the lowest value and tends to have a flat peak close to the mean in the data set. The same can be deduced from Figure 1. It is possible for us to state here that one ought to employ the probability distribution models for the purpose of fitting such data sets which are positively skewed and have a frequency curve with high peaks.

III. Methodology

In this section, we will explain the statistical approaches that would be utilised to achieve the fitting of economic variables, which was covered in the previous section. Inclusion in this category includes that of probability distributions, parameter estimation, and the goodness-of-fit criterion.

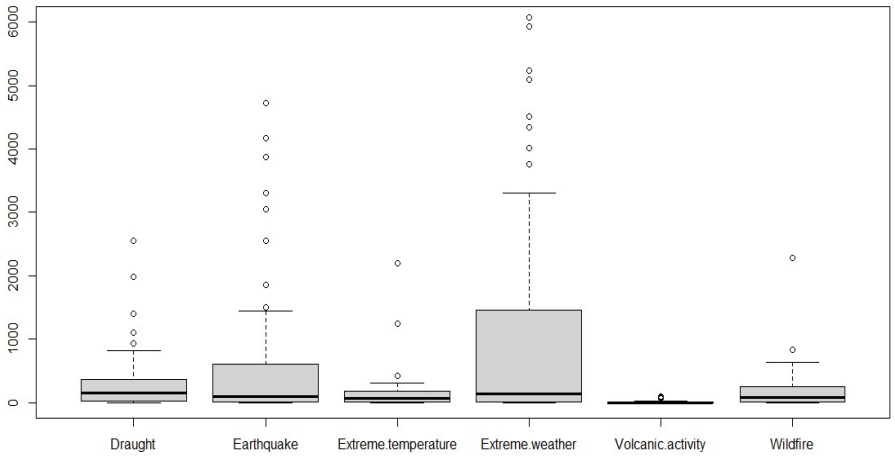


Figure 1: *Boxplots of economic losses due to six main natural disasters*

I. Probability distributions

The human mind is capable of incredible feats, and statistical modelling is one of them. It involves abstracting the results of observation in order to determine the similarities and differences between occurrences. When it comes to protecting ourselves against the effects of natural disasters, statistical models are a typical tool. The process of evaluating risks, making forecasts, and issuing warnings all depend heavily on modelling.

It is possible to draw the conclusion from the previous section that a statistical distribution with a right-skewed spread is the most accurate when it comes to modelling economic losses due to natural disasters. The current article makes use of five positively skewed distributions, namely the Weibull distribution, the Log-logistics distribution, the gamma distribution, the generalised Pareto distribution, and the Lognormal distribution, in order to fit the economic losses caused by natural disasters such as drought, earthquake, extreme temperature, extreme weather, volcanic activity, and wildfire. The probability density function (PDF), and cumulative distribution function (CDF) are given in Table 2.

The three-parameter Weibull distribution is commonly utilised in reliability and life data analysis [13]. Weibull distributions with $\beta=1$ have a constant failure rate, indicating usable life or random failures. Weibull distributions with $\beta > 1$ have a wear-out failure rate. Next is the Log-logistics three-parameter distribution, often known as the Fisk distribution in economics [14]. Characterizing the lifetime distributions log logistics distributions have property of a constant discrete Log-odds rate (LOR) with respect to t and ln t [15]. A random variable with a logistic logarithm has a log-logistic distribution. It resembles Lognormal but has heavier tails. Its cumulative distribution function is closed, unlike the lognormal. This distribution can exhibit a monotonically decreasing failure rate function for some parameter values. It is a survival analysis model for occurrences whose rate rises then falls. Some applications of the log-logistic distribution are discussed in economics to model wealth or income distribution [16] and in hydrology to estimate stream flow and precipitation [17].

Table 2: *The PDF & CDF of the distributions*

Model	PDF and its Support	CDF
Weibull	$f(x) = \frac{\beta}{\alpha}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}}, x \geq \gamma, \beta > 0, \alpha > 0.$	$F(x) = 1 - e^{\left(\frac{x-\gamma}{\alpha}\right)^{\beta}}$
Log-logistics	$f(x) = \frac{\left(1+\frac{\beta}{\alpha}(x-\gamma)\right)^{-\left(\frac{1}{\beta}+1\right)}}{\sigma\left(1+\left(1+\frac{\beta}{\alpha}(x-\gamma)\right)^{-\left(\frac{1}{\beta}\right)}\right)^2}$, $x \geq \gamma$, $\beta > 0$, $\alpha > 0$.	$F(x) = \left(1+\left(1+\frac{\beta}{\alpha}(x-\gamma)\right)^{-\left(\frac{1}{\beta}\right)}\right)^{-1}$
Gamma	$f(x) = \frac{1}{\Gamma\alpha\beta^{\alpha}}(x-\gamma)^{\alpha-1} e^{-\left(\frac{x-\gamma}{\beta}\right)}, \gamma < x < \infty, \beta > 0, \alpha > 0.$	$F(x) = \frac{\frac{\Gamma(x-\gamma)}{\beta^{\alpha}}}{\Gamma\alpha}$
Gen. Pareto	$f(x) = \frac{1}{\alpha}\left(1+\beta\left(\frac{x-\gamma}{\alpha}\right)\right)^{-1-\left(\frac{1}{\beta}\right)}$, $x \geq \gamma, \beta > 0.$	$F(x) = 1 - \left(1+\frac{\beta(x-\gamma)}{\sigma}\right)^{-\left(\frac{1}{\beta}\right)}$ for $\beta \neq 0$
Log-normal	$f(x) = \frac{1}{(x-\gamma)\alpha\sqrt{2\pi}} exp\left\{\frac{[ln(x-\gamma)-\beta]^2}{2\alpha^2}\right\}$, $x > \gamma \geq 0$, $-\infty < \beta < \infty$, $\alpha > 0$.	$F(x) = \Phi\left(\frac{ln(x-\gamma)-\beta}{\alpha}\right)$

It is positively skewed and the amount of skew depending inversely on the shape parameter. In gamma distribution median does not have a closed-form equation. Some applications of the gamma distribution are discussed in climatology to estimate the different behaviour of the natural climatic events [18] and in hydrological analysis [19]. Environmental studies use the Generalized Pareto distribution to model heavy-tailed data sets [4]. The distribution is called the "peaks over thresholds" model because it models flood control threshold exceedances. Generalized Pareto distribution models are used for extreme event [20]. The log-normal distribution is a function distributing a dependent variable in a normal or Gaussian fashion on a logarithmic scale of the independent variable (i.e., if the random variable X is log-normally distributed then Y=ln(X) has a normal distribution). A distribution that is log-normal in one of its moments will be log-normal in any of its moments with the same geometric standard deviation, describing the spread of the dependent variable [21]. The median size of any moment is connected to the median size of any other moment by an analytical relationship derived by [22]. One of the most common applications where log-normal distributions are used in finance is in the analysis of stock prices [23].

II. Maximum Likelihood Estimation

The most common method for obtaining estimators is by far the maximum likelihood approach. According to the MLE concept, the probability distribution that is "most likely" to accommodate for the observed data is the one that is wanted. As a result, one must look for the parameter vector value that maximises the likelihood function $L(\theta|x)$. The notion of maximum likelihood, which selects as the estimator that value of the parameter that maximises the PDF $f_{\theta}(x)$, effectively presupposes that the sample is representative of the population.

For each sample point x, let $\hat{\theta}(x)$ be a parameter value at which $L(\theta|x)$ attains its maximum as a function of θ, with x held fixed. Then $\hat{\theta}(x)$ is called the MLE of the parameter θ, (θ may be vector valued). Obtain n independent observations, x_1, x_2, …, x_n the estimates of parameters $\hat{\theta}_1, \hat{\theta}_2, \ldots\ldots\ldots, \hat{\theta}_k$ can be obtained by solving the differentiation of the logarithmic likelihood function as;

$$\frac{\partial logL(\hat{\theta}; x_1, x_2, \ldots, x_n)}{\partial \theta_j} = 0, j = 1, 2, \ldots, k. \quad (1)$$

Here, we discuss the complete producer of finding the MLEs of the Weibull distribution parameters. Consider the pdf and cdf of the Weibull distribution from Table 2. Assuming that the observations are independently distributed, the likelihood function is defined by,

$$L(\alpha,\beta,\gamma|data) = \prod_{i=1}^{n} f(x_i,\alpha,\beta,\gamma) \tag{2}$$

Our aim of estimation is to determine the three unknown parameter α,β,γ by the maximizing the likelihood (2) or equivalently log-likelihood function (3). The log-likelihood function is shown below.

$$logL\left(\hat{\theta};\, x_1, x_2, \ldots, x_n\right) = \sum_{i=1}^{n}\left[\ln(\beta) + (\beta-1)\ln(x-\gamma) - \beta\ln(\alpha) - \left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] \tag{3}$$

Using the conventional approach, we take the partial derivatives of the log-likelihood function (3) in terms of α,β,γ and set them equals to zero. We obtain the following equations,

$$\begin{cases} \dfrac{\partial \ln(L)}{\partial\alpha} = \displaystyle\sum_{i=1}^{n}\left[\left(\frac{1}{\beta}\right) + \ln(x-\gamma) - \ln(\alpha) - \left(\frac{x-\gamma}{\alpha}\right)^{\beta}\ln\left(\frac{x-\gamma}{\alpha}\right)\right] = 0, \\ \dfrac{\partial \ln(L)}{\partial\beta} = \displaystyle\sum_{i=1}^{n}\left[-\left(\frac{\beta}{\alpha}\right) + \left(\frac{\beta}{\alpha}\right)\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] = 0, \\ \dfrac{\partial \ln(L)}{\partial\gamma} = \displaystyle\sum_{i=1}^{n}\left[-\frac{(\beta-1)}{(x-\gamma)} + \left(\frac{\beta}{\alpha}\right)\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1}\right] = 0. \end{cases}$$

It is commonly understood that obtaining estimates of unknown parameters by solving equations given above numerically is challenging. The particle swarm optimization (PSO) approach is used to find unknown parameters estimates and is inspired by the notion of heuristic algorithms. Using this process, we can find the MLE for all of the distributions under consideration.

III. Particle Swarm Optimization Method

The biologically inspired approach known as particle swarm optimization, which was initially described by [24], is based on the flocking behaviour of birds. PSO is a population-based, self-adaptive search optimization method also referred to as an optimizer. All of the particles in the swarm move faster toward the best individual and overall position while continuously evaluating the value of their present location according to the same controlling principle. Each particle has a memory that aids it in remembering its most recent optimal location. Particle positions are classified as either personal best (pbest) or global best (gbest). Each particle has a unique pbest that is based on the journey it has taken. The particle compares the fitness value of its present position to that of pbest at each step along its route. The pbest is changed to the present location if the latter has a greater fitness value. Each particle also had a method of knowing where the swarm's greatest concentration of flowers had been located. The gbest, is the name given to this site of the best fitness ever found. There is a single gbest to which every particle is drawn throughout the whole swarm.

In a *n*-dimensional search space, the position and velocity of individual (particle or solution) *i* are represented as the vectors $Xi = (xi1, xi2, \ldots, xin)$ denote a particle's position (coordinate) and $Vi = (vi1, vi2, \ldots, vin,)$ denote the particle's flight velocity over a solution space in the PSO algorithm. Each individual x in the swarm is scored using a scoring function that obtains a score (fitness value) representing how good it solves the problem. Let $pbest_i$ and gbest = (x_1^{gbest},..., x^{gbest_n}) be the position of individual *i* and its neighbors' best position so far, respectively. Each particle records its own personal best position (pbest), and knows the best positions found by all particles in the swarm (gbest). Then, all particles that fly over the *n*-dimensional solution space are subject to updated rules for new positions, until the global optimal position is found. The modified velocity and position of each individual can be calculated using the current velocity and the distance from $pbest_i$ to gbest as follows:

$$V_i^{k+1} = \omega V_i^k + C_1 R_{ran\ 1}\left(pbest_i^k - X_i^k\right) + C_2 R_{ran\ 2}\left(gbest_i^k - X_i^k\right) \tag{4}$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \quad (5)$$

where V_i^k velocity of individual i at iteration k, ω weigh parameter (inertia weight), c_1, c_2 acceleration coefficients, R_{rand_1} and R_{rand_2} random numbers uniformly distributed between 0 and 1, X_i^k position of individual i at iteration k, $pbest_i^k$ best position of individual i until iteration k, $gbest_i^k$ best position of the group until iteration k.

The fundamental structure and pseudo-code of PSO algorithm

```
for each particle
        generate an initial particle
 end
 do
 for each particle
        calculate fitness value
        if  the fitness value is better than the best fitness value (pBest) in history
            set current value as the new pBest
        end
 end
 choose the particle with the best fitness values of all the particles as the gBest
 or each particle
      calculate particles velocity according eq (5)
      update particle position according eq (6)
 end
while maximum iteration criterion is not attained.
```

Marinho et al. [25] introduce the Adequacy Model computational library version 2.0.0 for the R statistical environment with two major contributions: a general optimization technique based on the PSO method (with a minor modification of the original algorithm) and a set of statistical measures for assessment of the adequacy of the fitted model. The goodness.fit() function provides some useful statistics to assess the quality of fit of probabilistic models. The function can also compute other measures such as AIC and KS test statistic. The general form for the function is given below:

Goodness.fit (pdf, cdf, starts=NULL, data, method="PSO", lim_inf, lim_sup, min(x), e, s, N, domain=c (0, inf))
where,

- pdf: probability density function (pdf);
- cdf: cumulative distribution function;
- starts: initial parameters to maximize the likelihood function;
- data: data vector;
- method: method used for minimization of the -log-likelihood function.
- method = "PSO", then all arguments of the PSO() function could be passed to the goodness.fit() function.
- lim_inf and lim_sup: define the inferior and superior boundaries of the search space, respectively;
- e: current error. The algorithm stops if the variance in the last iterations is less than or equal to e;
- S: number of considered particles
- domain: domain of the pdf. By default the domain of the pdf is the open interval (0, 1).

IV. Model selection criterion

The choice of the best probability distribution is a crucial step. The best distribution model for economic variables is determined using the goodness-of-fit (GoF) test and Akaike information criterion (AIC). The model with the lowest AIC value is the best fitting model. We discover the more accurate estimate for selecting the optimal model using the PSO approach. In addition, we carry out the same task as a probability plot using an empirical CDF plot. The empirical cumulative probability that is closest to the S-curve empirical one is chosen as the best fitting. The GoF test determines if a statistical model fits a collection of observations supplied in advance. Accordingly, the GoF measures are primarily used to summarise the discrepancy between observed values and predicted values under the specified statistical model. The minimal error produced, as assessed by the methods below, will be used to find the distribution that is best fitted:

The AIC, developed by [26], ranks models according to how well they fit the data and how little error they generate in their estimates. To move away from a solely inferential and limited approach to model selection, AIC has become part of a growing movement. It is defined as follows.

$$AIC = -2logL\left(\hat{\theta}_k\right) + 2k \tag{7}$$

Among all investigated distributions, the model with the lowest AIC value is regarded to be the best fitting model. Kolmogorov- Smirnov test compares empirical and theoretical distributions. Let us consider $F_0(x)$ is the population CDF and $S_N(x)$ the observed cumulative step function of a sample (i.e., $S_N(x) = k/N$, where k is the number of observation less or equal to x), then KS test statistic is defined as

$$T = \max_x |F_0(x) - S_N(x)|. \tag{8}$$

For implications, we reject the hypothesis at the level of significance, α, if T exceeds the $1-\alpha$ quantile as given by the table of quantile for the KS test statistic.

IV. Results and Discussion

The economic losses caused by six natural catastrophes (drought, earthquake, extreme weather, extreme temperature, wildfire, and volcanic activity) are examined in this part and fitted to the five probability distributions discussed in section 2. First, the investigation focuses on determining the best-fitting model using the AIC value. Among all the models evaluated, the model with the lowest AIC was deemed the best. However, the PSO technique in R was used to estimate the parameters of the five theoretical probability distributions using maximum likelihood estimation. Tables [3–8] provide the MLEs, KS statistic (along with p-value), and AIC value for each fitted model for each economic variable. The fitting results show that some PDF characteristics are more suited for some places while being less appropriate for others.

Table 3: *MLEs, KS statistics, p-value and AIC for all five distributions for Drought data*

Distribution	MLE			P - Value	Statistic	AIC
Log-Logistic	0.7277	94.8424	0.1000	0.5207	0.1131	653.0388
Weibull(3P)	0.3469	87.8642	0.0978	0.0005	0.2852	659.8648
Gamma	0.3742	0.0005	0.0629	0.0173	0.2161	654.3794
Gen.Pareto	1.0106	99.7400	0.0660	0.2976	0.1360	662.5274
Lognormal	2.6033	4.5910	0.0068	0.4443	0.1202	662.7647

Figure 3 depicts the PDF plot of all heavy-tailed variables. Our findings are closed in terms of log-logistic and Weibull distributions. The empirical investigations show that Weibull considerably fits the

greatest value whereas Generalised Pareto underestimates it. Meanwhile, of all competing models in the research, Generalised Pareto provides the weakest match. The log-logistic model is the second best. The KS test statistic are used to select the distribution at 95% confidences interval from the Tables [3-8] we compare the results for all the distribution.

To fit the distributions for economic losses owing to drought, Table 3 shows that the p-value of Weibull and Gamma is less than 0.05 so we reject the null hypothesis. Log-logistic distribution has the lowest AIC value (653.0338), highest likelihood estimates (0.7277, 94.8424, 0.1), and smallest p-value (0.5207). The gamma distribution has the second-lowest AIC value (654.3794) among all distributions. Table 4 shows that the p-value of Weibull, Gamma and General Pareto is less than 0.05 that implies we reject the hypothesis. Log-logistic has the lowest AIC for earthquake economic losses (1117.5730). Weibull has a higher AIC value than Lognormal (1122.6180).

Table 4: *MLEs, KS statistics, p-value and AIC for all five distributions for earthquake data*

Distribution		MLE		P - Value	Statistic	AIC
Log-Logistic	0.5823	68.3169	0.0650	0.3430	0.1029	1117.5730
Weibull(3P)	0.4213	89.6176	0.0616	0.0046	0.1912	1145.4830
Gamma	0.4980	0.0015	0.0603	0.0225	0.1644	1240.6280
Gen. Pareto	2.4820	25.8344	0.0565	0.0332	0.1571	1136.3800
Lognormal	2.6557	4.0074	0.0439	0.1001	0.1343	1122.6180

Table 5 shows that the p-value of Gamma is less than 0.05 so we reject the hypothesis. Log-logistics has the lowest AIC value (384.8178) across all distributions. Generalized Pareto has the second lowest AIC, whereas gamma has the highest. According to Table 6, the p-value of Weibull and Gamma is less than 0.05 that implies we reject the null hypothesis. The gamma distribution has the highest AIC value among all distributions for the distribution of economic losses brought on by extreme weather, while log-logistic has the lowest value.

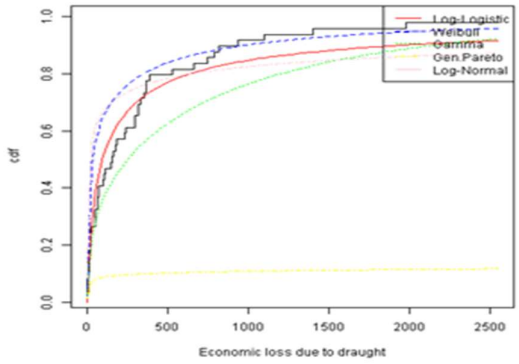
Table 5: *MLEs, KS statistics, p-value and AIC for all five distributions for Temperature data*

Distribution		MLE		P - Value	Statistic	AIC
Log-Logistic	0.7077	72.8837	0.0600	0.2893	0.1737	384.8178
Weibull(3P)	0.5810	89.9062	0.0303	0.4292	0.1546	388.0758
Gamma	0.6875	0.0025	0.0083	0.0068	0.2979	397.0683
Gen.Pareto	0.8057	62.7886	0.0392	0.9816	0.0824	386.1265
Lognormal	1.7809	3.8543	0.0110	0.6215	0.1332	390.8100

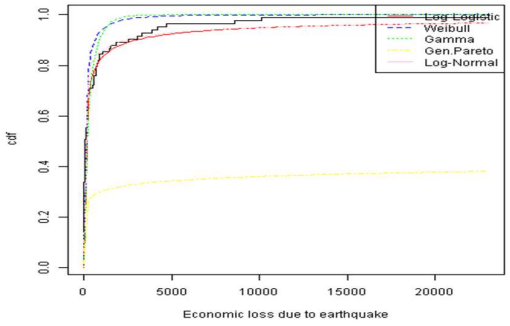
Table 6: *MLEs, KS statistics, p-value and AIC for all five distributions for Extreme Weather data*

Distribution		MLE		P - Value	Statistic	AIC
Log-Logistic	0.5061	70.7698	0.0500	0.2772	0.1025	1343.7310
Weibull	0.2981	83.7310	0.0389	0.0003	0.2180	1372.3730
Gamma	0.2195	0.0004	0.0172	0.0068	0.1738	1384.3180
Gen. Pareto	3.0613	23.8658	0.0496	0.0984	0.1266	1366.3450
Lognormal	3.4142	4.7270	0.0423	0.5663	0.0811	1349.6740

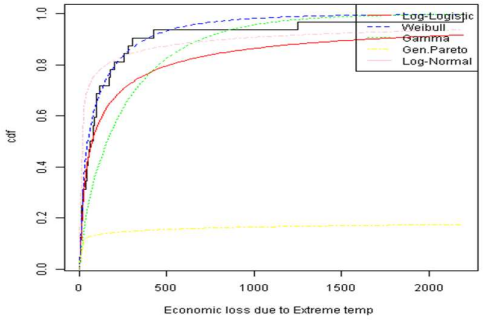
(a) Drought



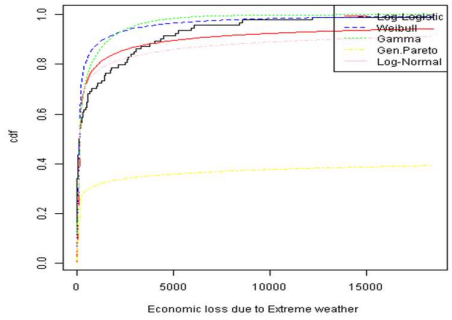
(b) Earthquakes



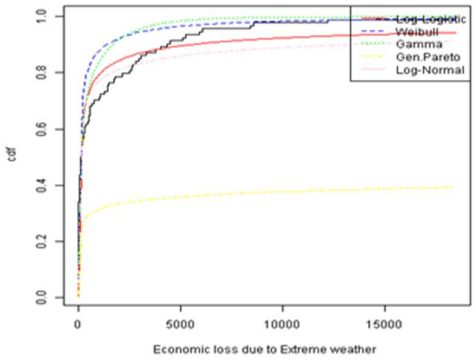
(c) Extreme Temperature



(d) Extreme Weather



(e) Volcanic Activity



(f) Wildfire

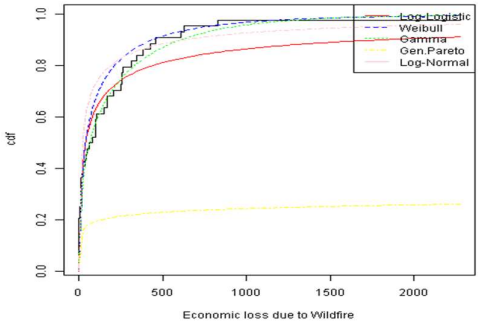
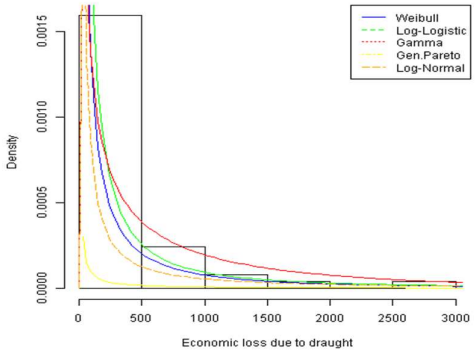
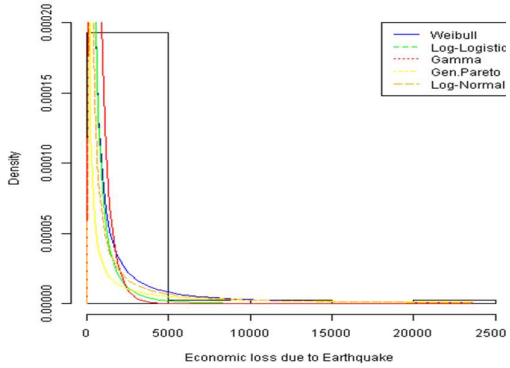


Figure 2: *Fitted CDF plots for economic losses.*

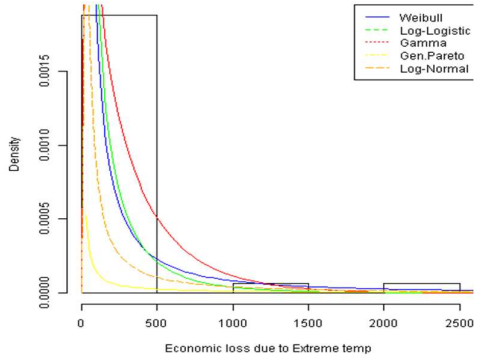
(a) Drought



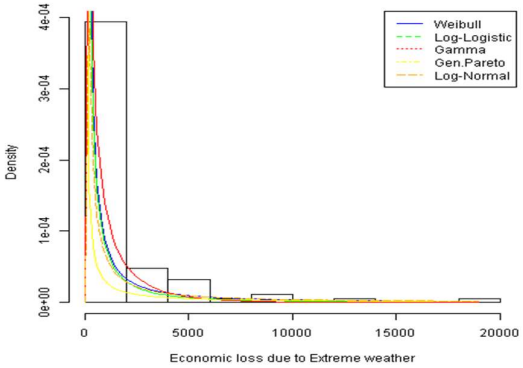
(b) Earthquakes



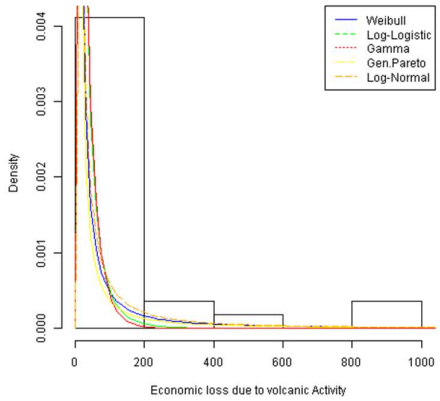
(c) Extreme Temperature



(d) Extreme Weather



(e) Volcanic Activity



(f) Wildfire

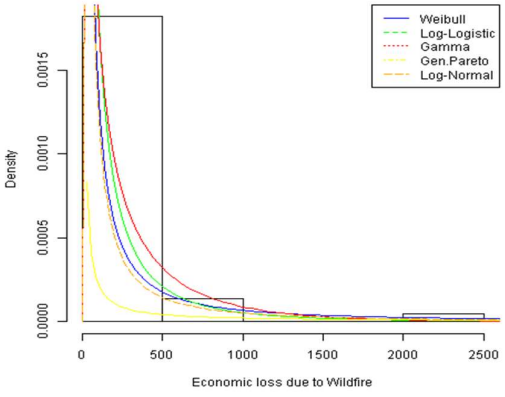


Figure 3: *Fitted PDF plots for economic losses*

According to Table 7, the economic losses as a result of volcanic activity we discovered out of all the distributions, log-logistic has the least AIC value, followed by Weibull. Among all, Generalized Pareto has the highest AIC values.

Table 7: *MLEs, KS statistics, p-value and AIC for all five distributions for Volcanic data*

Distribution	MLE			P - Value	KS	AIC
Log-Logistic	0.6367	2.2675	0.0200	0.9628	0.0948	168.3601
Weibull(3P)	0.5663	9.6857	0.0199	0.0817	0.2390	171.2018
Gamma	0.3550	0.0253	0.0165	0.2350	0.1955	174.1816
Gen.Pareto	1.0966	1.9923	0.0200	0.6410	0.14019,	177.8270
Lognormal	0.8888	2.8176	0.0199	0.9108	0.1061	174.4506

Next, from Table 8 we discovered that among all distributions, Weibull has a lower AIC value than log-logistic, which has the second lowest value for the distribution of economic losses caused by wildfire. The AIC value of Generalized Pareto is the highest of all.

Table 8: *MLEs, KS statistics, p-value and AIC for all five distributions for Wildfire data*

Distribution	MLE			P - Value	KS	AIC
Log-Logistic	0.5814	40.4627	0.1000	0.5036	0.1244	510.7080
Weibull(3P)	0.5027	82.5439	0.0998	0.3161	0.1446	504.5558
Gamma	0.3547	0.0017	0.0477	0.9977	0.0594	516.5590
Gen. Pareto	1.4600	36.3105	0.0794	0.3054	0.1460	534.6831
Lognormal	2.9891	2.6805	0.0896	0.0507	0.2044	525.7421

V. Conclusion and future work

The goal of the current work is to identify the most appropriate three parametric probability models for datasets of economic losses from natural catastrophes. For modelling economic losses, scholars have previously advocated using the Generalized Pareto or extreme value distribution. Both probability distributions are specified on the real line, and economic losses occur on the positive real line. As a result, these distributions can offer a negative lower bound on the economic losses. In this work, we take into account five significant probability distributions (Weibull, Log-logistics, Gamma, Generalized Pareto, and Lognormal) that are defined on the positive real line to describe the economic scenarios. Utilizing the KS-test, CDF plot, and AIC criterion, the best fitted probability distribution is determined for each dataset. Empirical CDF plots show that Weibull and log-logistic fit pretty well, whereas generalised Pareto fits poorly. According to the KS-test statistic, we discovered that the Log-logistic and Lognormal suit all economic losses resulting from natural catastrophe data for the stated level of significance, 5%. However, draught, earthquake, and extreme weather datasets cannot be fitted by Weibull or Gamma. The earthquake dataset does not match the generalised Pareto model. The Log-logistic distribution offers the greatest fit among all taken distributions for five datasets (drought, earthquake, extreme weather, extreme temperature, wildfire, and volcanic activity) according to the AIC criteria. It is, nonetheless, the second-best fitted distribution for the wildfire dataset. It should be noticed that the Weibull distribution is rejected by the KS-test yet has the minimum AIC for the wildfire dataset. As a result, we may also suggest log-logistic for modelling economic losses from wildfires. Finally, we advise using the log-logistic probability model to fit and analyse economic losses brought on by natural catastrophes in future research.

Regression analysis is employed when the assumption of normality is taken into consideration to develop and investigate the relationship between the response and explanatory variables. In certain applications, the assumption of normalcy is not valid practically; see [27]. Numerous examples are

given in [28] that demonstrate the usage of skewed or non-normal distributions for both random components and response variables. In such circumstances, we advise fitting parametric regression for the economic losses using the log-logistic probability model. The model may be defined as

$$Economic\ loss\ (Y) = \beta X' + \sigma\ \varepsilon$$

where X' are the features matrix (regressions of economic losses), β is a vector of regression coefficients, σ is a scale parameter and ε stands for random component that may follow the log-logistics distribution. Readers are encouraged to take this work into consideration while planning their own future projects on modelling of economic losses due to natural disasters.

Table 9: *Ranks of the fitted distributions based on AIC values.*

	Variables					
Distributions	Draught	Earthquake	Extreme Temperature	Extreme Weather	Volcanic Activity	Wildfire
Log-Logistic	I	I	I	I	I	II
Weibull (3P)	III	V	III	IV	II	I
Gamma	II	IV	V	V	III	III
Gen. Pareto	IV	III	II	III	V	V
Lognormal	V	II	IV	II	IV	IV

Acknowledgement

Authors thank editor in-chief and referees for their fruitful suggestions. Dr. Sharma greatly acknowledges the financial support from the Banaras Hindu University as seed grant under Institute of Eminence Scheme.

References

[1] Sedghi, A. (2013). Typhoon Haiyan: how does it compare with other tropical cyclones. *The Guardian*.

[2] Yamano, N., Kajitani, Y., & Shumuta, Y. (2007). Modeling the regional economic loss of natural disasters: the search for economic hotspots. *Economic Systems Research*, *19*(2), 163-181.

[3] Kellenberg, D. K., & Mobarak, A. M. (2008). Does rising income increase or decrease damage risk from natural disasters? *Journal of urban economics*, *63*(3), 788-802.

[4] Pisarenko, V. F., & Rodkin, M. V. (2014). *Statistical analysis of natural disasters and related losses* (p. 82). Dordrecht-Heidelberg-London-New York: Springer.

[5] Panwar, V., & Sen, S. (2019). Economic impact of natural disasters: An empirical re-examination. *Margin: The Journal of Applied Economic Research*, *13*(1), 109-139.

[6] Benson, C., & Clay, E. (2003). Economic and financial impacts of natural disasters: an assessment of their effects and options for mitigation: synthesis report. *Overseas Development Institute, London*.

[7] Pauw, K., Thurlow, J., Bachu, M., & Van Seventer, D. E. (2011). The economic costs of extreme weather events: a hydrometeorological CGE analysis for Malawi. *Environment and Development Economics*, *16*(2), 177-198.

[8] Okuyama, Y. (2007). Economic modeling for disaster impact analysis: past, present, and future. *Economic Systems Research*, *19*(2), 115-124.

[9] Hallegatte, S. (2008). An adaptive regional input-output model and its application to the assessment of the economic cost of Katrina. *Risk Analysis: An International Journal*, *28*(3), 779-799.

[10] Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., & Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. *Proceedings of the National Academy of Sciences*, *116*(43), 21450-21455.

[11] Jindrová, P., & Pacáková, V. (2016). Modelling of extreme losses in natural disasters. *International Journal of Mathematical Models and Methods in Applied Sciences, volume 10, issue: 2016*.
[12] Ibrahim, R. A., Sukono, S., & Riaman, R. (2021). Estimation of the Extreme Distribution Model of Economic Losses Due to Outbreaks Using the POT Method with Newton Raphson Iteration. *International Journal of Quantitative Research and Modeling*, *2*(1), 37-45.
[13] Moeini, A., Jenab, K., Mohammadi, M., & Foumani, M. (2013). Fitting the three-parameter Weibull distribution with Cross Entropy. *Applied Mathematical Modelling*, *37*(9), 6354-6363.
[14] Ramos, M. W. A., Cordeiro, G. M., Marinho, P. R. D., Dias, C. R. B., & Hamedani, G. G. (2013). The Zografos-Balakrishnan log-logistic distribution: Properties and applications. *Journal of Statistical Theory and Applications*, *12*(3), 225-244.
[15] Khorashadizadeh, M., Rezaei Roknabadi, A. H., & Mohtashami Borzadaran, G. R. (2013). Characterization of life distributions using Log-odds rate in discrete aging. *Communications in Statistics-Theory and Methods*, *42*(1), 76-87.
[16] Kleiber, C., & Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. John Wiley & Sons.
[17] Ashkar, F., & Mahdi, S. (2006). Fitting the log-logistic distribution by generalized moments. *Journal of Hydrology*, *328*(3-4), 694-703.
[18] Thom, H. C. (1958). A note on the gamma distribution. *Monthly weather review*, *86*(4), 117-122.
[19] Aksoy, H. (2000). Use of gamma distribution in hydrological analysis. *Turkish Journal of Engineering and Environmental Sciences*, *24*(6), 419-428.
[20] Holmes, J. D., & Moriarty, W. W. (1999). Application of the generalized Pareto distribution to extreme value analysis in wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics*, *83*(1-3), 1-10.
[21] Heintzenberg, J. (1994). Properties of the log-normal particle size distribution. *Aerosol Science and Technology*, *21*(1), 46-48.
[22] Hatch, T., & Choate, S. P. (1929). Statistical description of the size properties of non uniform particulate substances. *Journal of the Franklin Institute*, *207*(3), 369-387.
[23] Dufresne, D. (2004). The log-normal approximation in financial and other computations. *Advances in applied probability*, *36*(3), 747-773.
[24] Kennedy, J., & Eberhart, R. (1995). "Particle swarm optimization," Proceedings of ICNN'95–International Conference on Neural Networks, Perth, WA, Australia.
[25] Marinho, P. R. D., Silva, R. B., Bourguignon, M., Cordeiro, G. M., & Nadarajah, S. (2019). AdequacyModel: An R package for probability distributions and general purpose optimization. *PloS one*, *14*(8), e0221487.
[26] Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic bulletin & review*, *11*, 192-196.
[27] Dunteman, G. H., & Ho, M. H. R. (2006). An introduction to generalized linear models (Vol. 145). Sage.
[28] Kalbfleisch, J. D., & Prentice, R. L. (2011). The statistical analysis of failure time data. John Wiley & Sons.

STOCHASTIC ANALYSIS OF THE UTENSIL INDUSTRY SUBJECT TO REPAIR FACILITY

Hanumanolla Indrasena Reddy[1], Mohit Yadav[2] and Hemant Kumar[3*]

•

[1,2]Department of Mathematics, University Institute of Sciences
Chandigarh University, Mohali, Chandigarh, India
[3]SOET, Raffles University, Neemrana, Rajasthan
*Corresponding Author
22msm40224@cuchd.in, mohit.e15793@cumail.in, hemantkumar@rafflesuniversity.edu.in

Abstract

The availability and profit values of the utensil industry are analyzed using the regenerative point graphical technique. The utensil industry contains three different units where two units can work with reduced capacity. It is considered that units C and D may be in a complete failed state through partial failure but unit B is in only complete failed state. When a unit is completely failed then the system is in failed state. An expert technician is available to repair the failed unit. Failure and repair times are independent of each other. The distribution of the failure time is general and repair time is exponential. Various parameters such as mean time to system failure, availability, busy period of the server, expected number of server visits and profit values are calculated with the help of tables.

Keywords: Reliability, base-state, mean sojourn time, availability and profit.

I. Introduction

To satisfy the growing demand for products, manufacturers and industrialists must produce products continually which they can accomplish by optimizing their manufacturing processes. This paper discusses the MTSF, availability and profit values of the utensil industry with priority in repair using the regenerative point graphical technique under specified conditions. A large amount of research work has been done on repairable systems such that Kapur and Kapoor [8] described the stochastic nature of a two unit repairable system under one spare unit. Gnedenko and Igor [5] explored reliability and probability measures for engineering purposes. Jack and Murthy [6] discovered the role of limited warranty and extended warranty for the product. Wang and Zhang [16] examined the repairable system of two non identical components under repair facility using geometric distributions. Diaz et al. [4] threw light on the warranty cost management system. Kumar and Goel [12] explored the idea of an imperfect switch on redundant systems in banking industry. Kumar and Goel [11] analyzed the preventive maintenance in two unit cold standby system under general distributions. Malik and Rathee [14] threw light on the two parallel units system under preventive maintenance and maximum operation time. Kashid and Kumar [9] examined the availability of two unit system under degradation and subject to the repair facility. Chaudhary and Tomar [3] examined the behavior of a two unit cold standby system under inspection. Kumar et al. [10] evaluate the effects of washing unit in the paper industry by using the regenerative point graphical technique. Levitin et al. [13] explored the results of optimal

preventive replacement of failed units in a cold standby system by using the poisson process. Agarwal et al. [1] analyzed the performance and reliability of water treatment plant under repair facility. Jia et al. [7] explored the two unit system under demand and energy storage techniques. Sengar and Mangey [15] examined the performance of complicated systems under inspection using copula methodology.

II. System Assumptions

There are following system assumptions:

- The utensil industry consists of three distinct units such that cutting system, pressing system, spinning and buffing system.
- Unit *B* consists of a cutting system in which sheets are cut into circular sheets.
- Unit *C* has a pressing system that converts the sheets into the shape of utensils.
- Unit *D* has a spinning and buffing system that gives the final shape and polish to the utensils.
- It is considered that units *C* and *D* may be in a complete failed state through partial failure but unit B is in only complete failed state.
- Failure time follows general distribution whereas repair time follows the exponential distribution.
- A server is always available to repair the failed unit.
- The failed unit works like a new unit after repair.

III. System Notations

There are following system notations:

$i \xrightarrow{Sr} j$	r^{th} directed simple path from state 'i' to state 'j' where 'r' takes the positive integral values for different directions from state 'i' to state 'j'.
$\xi \xrightarrow{s\,f\,f} i$	A directed simple failure free path from state ξ to state 'i'.
$m-cycle$	A circuit (may be formed through regenerative or non regenerative / failed state) whose terminals are at the regenerative state 'm'.
$m-\overline{cycle}$	A circuit (may be formed through the unfailed regenerative or non regenerative state) whose terminals are at the regenerative 'm' state.
$U_{k,k}$	Probability factor of the state 'k' reachable from the terminal state 'k' of 'k' cycle.
$U_{\overline{k,k}}$	The probability factor of state 'k' reachable from the terminal state 'k' of $k\ \overline{cycle}$.
μ_i	Mean sojourn time spent in the state 'i' before visiting any other states.
μ_i'	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state 'i' at t=0.
η_i	Expected waiting time spent while doing a job given that the system entered to the regenerative state 'i' at t=0.
B/b	System first unit is in the operative state/failed state.
$C/\overline{C}/c$	System second unit is in the operative state/reduced state/failed state.
$D/\overline{D}/d$	System third unit is in the operative state/reduced state/failed state.
λ_1 / λ_3	The constant partial failure rate of the unit C/D respectively.

λ_2 / λ_4	The constant complete failure rate of the unit C/D respectively.
λ_5	The constant complete failure rate of unit B.
$f_1(t) / F_1(t)$	PDF/CDF of repair time of unit C from partial failed state.
$f_2(t) / F_2(t)$	PDF/CDF of repair time of unit C from complete failed state.
$f_3(t) / F_3(t)$	PDF/CDF of repair time of unit D from partial failed state.
$f_4(t) / F_4(t)$	PDF/CDF of repair time of unit D from complete failed state.
$f_5(t) / F_5(t)$	PDF/CDF of repair time of unit B from complete failed state.

IV. Transition Diagram and Their Descriptions

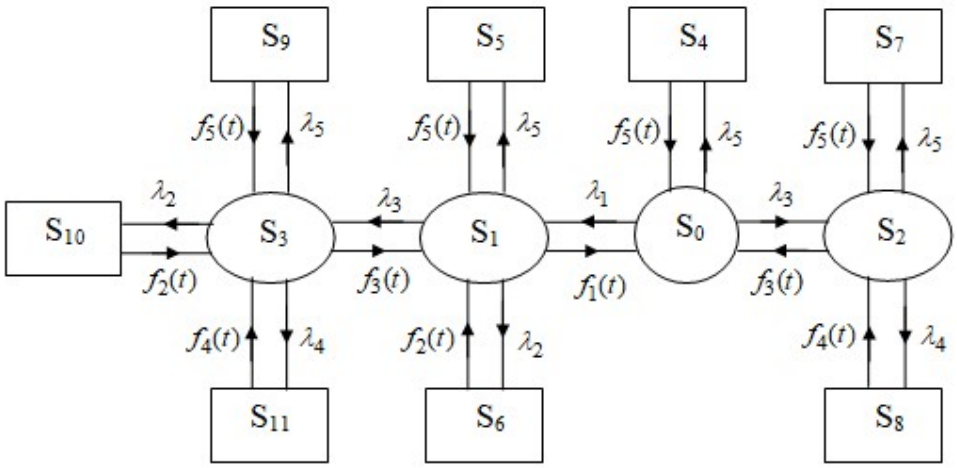


Figure 1: *State Transition Diagram*

In the system transition diagram, there are following states

where, $S_0 = BCD$, $S_1 = B\overline{C}D$, $S_2 = BC\overline{D}$, $S_3 = B\overline{CD}$, $S_4 = bCD$, $S_5 = b\overline{C}D$

$S_6 = BcD$, $S_7 = b\overline{C}\overline{D}$, $S_8 = BCd$, $S_9 = b\overline{CD}$, $S_{10} = Bc\overline{D}$, $S_{11} = B\overline{C}d$

V. Transition Probabilities

The transition probabilities are following

$p_{0,1} = \lambda_1 /(\lambda_1 + \lambda_3 + \lambda_5)$, $p_{0,2} = \lambda_3 /(\lambda_1 + \lambda_3 + \lambda_5)$

$p_{0,4} = \lambda_5 /(\lambda_1 + \lambda_3 + \lambda_5)$, $p_{1,0} = w_1 /(w_1 + \lambda_2 + \lambda_3 + \lambda_5)$

$p_{1,3} = \lambda_3 /(w_1 + \lambda_2 + \lambda_3 + \lambda_5)$, $p_{1,5} = \lambda_5 /(w_1 + \lambda_2 + \lambda_3 + \lambda_5)$

$p_{1,6} = \lambda_2 /(w_1 + \lambda_2 + \lambda_3 + \lambda_5)$, $p_{2,0} = w_3 /(w_3 + \lambda_4 + \lambda_5)$

$p_{2,7} = \lambda_5 /(w_3 + \lambda_4 + \lambda_5)$, $p_{2,8} = \lambda_4 /(w_3 + \lambda_4 + \lambda_5)$

$p_{3,1} = \lambda_3 /(w_2 + \lambda_3 + \lambda_4 + \lambda_5)$, $p_{3,9} = \lambda_5 /(w_2 + \lambda_3 + \lambda_4 + \lambda_5)$

$p_{3,10} = \lambda_2 /(w_2 + \lambda_3 + \lambda_4 + \lambda_5)$, $p_{3,11} = \lambda_4 /(w_3 + \lambda_2 + \lambda_4 + \lambda_5)$

$$p_{4,0} = p_{5,1} = p_{6,1} = p_{7,2} = p_{8,2} = p_{9,3} = p_{10,3} = p_{11,3} = 1 \tag{1}$$

It has been concluded that

$$p_{0,1} + p_{0,2} + p_{0,4} = 1,\ p_{1,0} + p_{1,3} + p_{1,5} + p_{1,6} = 1$$
$$p_{2,0} + p_{2,7} + p_{2,8} = 1,\ p_{3,1} + p_{3,9} + p_{3,10} + p_{3,11} = 1 \tag{2}$$

VI. Mean Sojourn Time

Let μ_i represents the mean sojourn time. Mathematically, the time taken by a system in a particular state becomes

$$\mu_i = \sum_j m_{i,j} = \int_0^\infty P(T > t)\,dt\ .$$

and $\mu_0 = 1/(\lambda_1 + \lambda_3 + \lambda_5)$, $\mu_1 = 1/(w_1 + \lambda_2 + \lambda_3 + \lambda_5)$, $\mu_2 = 1/(w_2 + \lambda_4 + \lambda_5)$

$$\mu_3 = 1/(w_2 + \lambda_3 + \lambda_4 + \lambda_5),\ \mu_4(t) = \mu_5(t) = 1/(w_5),\ \mu_6 = \mu_{10} = 1/(w_2)$$
$$\mu_7 = \mu_9 = 1/(w_5),\ \mu_8 = \mu_{11} = 1/(w_4) \tag{3}$$

VII. Evaluation of Parameters

All reliability parameters (such as mean time to system failure, availability, busy period of the server and expected number of visits) are determined by using the regenerative point graphical technique.

I. Mean Time to System Failure (MTSF)

The regenerative un-failed states (i=0, 1, 2, 3) to which the system can transit (with initial state 0) before entering to any failed state (using base state ξ=0) then MTSF becomes

$$T_0 = \left[\sum_{i=0}^{3} Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr(sff)} i) \right\} . \mu_i}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[1 - \sum Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr(sff)} 0) \right\}}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$T_0 = [\mu_0 + p_{0,1}\mu_1 + p_{0,2}\mu_2 + p_{0,1}p_{1,3}\mu_3] / [1 - p_{0,1}p_{1,0} - p_{0,2}p_{2,0}] \tag{4}$$

II. Availability of the System

The system is available for use at regenerative states j=0, 1, 2, 3 with ξ=0 then the availability of system is defined as

$$A_0 = \left[\sum_{j=0}^{3} Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\} . f_j . \mu_j}{\prod_{k_1 \neq 0} \left\{ 1 - V_{k_1 k_1} \right\}} \right\} \right] \div \left[\sum_{i=0}^{11} Sr \left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\} . \mu_i'}{\prod_{k_2 \neq 0} \left\{ 1 - V_{k_2 k_2} \right\}} \right\} \right]$$

$$A_0 = \frac{[U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3]}{\begin{bmatrix} U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 \\ + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 + U_{0,9}\mu_9 + U_{0,10}\mu_{10} + U_{0,11}\mu_{11} \end{bmatrix}} \tag{5}$$

III. Busy Period of the Server

The server is busy due to repair of the failed unit at regenerative states *j*= 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 with ξ = 0 then the fraction of time for which the server remains busy is defined as

$$B_0 = \left[\sum_{j=1}^{11} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\}.\eta_j}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}} \right\} \right] \div \left[\sum_{i=0}^{11} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}} \right\} \right]$$

$$B_0 = \frac{\begin{bmatrix} U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 + U_{0,6}\mu_6 \\ + U_{0,7}\mu_7 + U_{0,8}\mu_8 + U_{0,9}\mu_9 + U_{0,10}\mu_{10} + U_{0,11}\mu_{11} \end{bmatrix}}{\begin{bmatrix} U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 \\ + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 + U_{0,9}\mu_9 + U_{0,10}\mu_{10} + U_{0,11}\mu_{11} \end{bmatrix}} \tag{6}$$

IV. Estimated Number of Visits Made by the Server

The technician visits at regenerative states *j*= 1, 2, 3 with ξ=0 then the number of visits by the repairman is defined as

$$V_0 = \left[\sum_{j=1}^{3} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} j) \right\}}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}} \right\} \right] \div \left[\sum_{i=0}^{11} Sr\left\{ \frac{\left\{ pr(0 \xrightarrow{Sr} i) \right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}} \right\} \right]$$

$$V_0 = \frac{[U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3]}{\begin{bmatrix} U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 \\ + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 + U_{0,9}\mu_9 + U_{0,10}\mu_{10} + U_{0,11}\mu_{11} \end{bmatrix}} \tag{7}$$

V. Profit Analysis

The profit function may be used to do a profit analysis of the system and it is given by

$$P = E_0 A_0 - E_1 B_0 - E_2 V_0 \tag{8}$$

where, $E_0 = 5000$ (Pay per unit uptime of the system)

$E_1 = 1000$ (Charge per unit time for which technician is busy due to repair)

$E_2 = 500$ (Charge per visit of the technician)

VI. Particular cases

It is considered that

$$f_1(t) = w_1 e^{-w_1 t},\ f_2(t) = w_2 e^{-w_2 t},$$

$$f_3(t) = w_3 e^{-w_3 t},\ f_4(t) = w_4 e^{-w_4 t},\ f_5(t) = w_5 e^{-w_5 t}$$

and $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda$, $w_1 = w_2 = w_3 = w_4 = w_5 = w$.

$$T_0 = \frac{[(w+3\lambda)(w+4\lambda)+\lambda^2](w+2\lambda)}{(w+3\lambda)\left[3\lambda(w+3\lambda)+(w+2\lambda)-\lambda w(w+2\lambda)-\lambda w(w+3\lambda)\right]}$$

$$A_0 = \frac{w^3(w+2\lambda)[(w+3\lambda)^3+\lambda(w+2\lambda)(w+3\lambda)]}{\left[\begin{array}{l} w^3(w+2\lambda)[(w+3\lambda)^3+\lambda(w+2\lambda)(w+3\lambda)] \\ +\lambda w^2(w+2\lambda)(w+3\lambda)^3(w+2\lambda) \\ +[\lambda^2 w(2w+5\lambda)]w(w+2\lambda)(w+3\lambda)+3\lambda^3 w^2(w+2\lambda)^2 \end{array}\right]}$$

$$B_0 = \frac{\left[\begin{array}{l} w^3(w+2\lambda)[(w+3\lambda)^2\lambda+\lambda(w+2\lambda)(w+3\lambda)] \\ +\lambda w^2(w+2\lambda)(w+3\lambda)^3(w+2\lambda) \\ +[\lambda^2 w(2w+5\lambda)]w(w+2\lambda)(w+3\lambda)+3\lambda^3 w^2(w+2\lambda)^2 \end{array}\right]}{\left[\begin{array}{l} w^3(w+2\lambda)[(w+3\lambda)^3+\lambda(w+2\lambda)(w+3\lambda)] \\ +\lambda w^2(w+2\lambda)(w+3\lambda)^3(w+2\lambda) \\ +[\lambda^2 w(2w+5\lambda)]w(w+2\lambda)(w+3\lambda)+3\lambda^3 w^2(w+2\lambda)^2 \end{array}\right]}$$

$$V_0 = \frac{w^3(w+2\lambda)[(w+3\lambda)^2\lambda+\lambda(w+2\lambda)(w+3\lambda)]}{\left[\begin{array}{l} w^3(w+2\lambda)[(w+3\lambda)^3+\lambda(w+2\lambda)(w+3\lambda)] \\ +\lambda w^2(w+2\lambda)(w+3\lambda)^3(w+2\lambda) \\ +[\lambda^2 w(2w+5\lambda)]w(w+2\lambda)(w+3\lambda)+3\lambda^3 w^2(w+2\lambda)^2 \end{array}\right]}$$

VIII. Discussion

Table 1 describes the nature of the mean time to system failure of the utensil industry. It has an

Table 1: *MTSF vs. Repair Rate*

w ↓	λ=0.02	λ=0.035	λ=0.05
0.05	4.357262	4.200299	3.696809
0.10	4.600326	4.405797	3.903394
0.15	4.832536	4.599156	4.102564
0.20	5.054602	4.781421	4.29471
0.25	5.267176	4.953519	4.480198
0.30	5.470852	5.116279	4.659367
0.35	5.666179	5.27044	4.832536
0.40	5.853659	5.416667	5.057888
0.45	6.033755	5.555556	5.162037
0.50	6.206897	5.687646	5.318907

increasing trend corresponding to increment in repair rate (w) and has decreasing trend corresponding to an increment in failure rate (λ). In this table, the values of parameters are λ= 0.02, 0.035, 0.05 and w=0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50 respectively. When the value of repair rate enhances then MTSF values are also enhanced. When λ=0.02 changes into λ=0.035, 0.05 then MTSF values are declined.

Table 2 explores the increasing trends of availability with respect to increments in repair rate (w) and has decreasing trends corresponding to increments in failure rate (λ). When the value of the repair rate is enhanced then the availability values are also enhanced. Also, when the failure rate of unit changes λ=0.02 to 0.035, 0.05 then the availability of system declines.

Table 2: *Availability vs. Repair Rate*

w ↓	λ=0.02	λ=0.035	λ=0.05
0.05	0.665768	0.658036	0.627958
0.10	0.677838	0.668825	0.640662
0.15	0.688592	0.678383	0.652123
0.20	0.698233	0.686909	0.662516
0.25	0.706926	0.694563	0.671982
0.30	0.714804	0.70147	0.680641
0.35	0.721976	0.707736	0.688592
0.40	0.728534	0.713446	0.695917
0.45	0.734553	0.71867	0.702689
0.50	0.742585	0.728578	0.717855

Table 3 explores the trend of profit values with respect to repair rate (w) and its value increase corresponding to increments in repair rate (w) and decrease corresponding to increments in failure rate (λ). It is concluded that when the value of the repair rate enhances then profit values are also enhanced but when the failure rate of the unit changes λ=0.02 to 0.035, 0.05 then the profit of the system declines.

Table 3: *Profit vs. Repair Rate*

w ↓	λ=0.02	λ=0.035	λ=0.05
0.05	57529.56	56550.96	55338.67
0.10	58611.63	57814.53	56419.58
0.15	59875.52	58249.63	57660.33
0.20	60296.25	59831.35	58504.51
0.25	61854.87	60540.89	59545.69
0.30	62532.78	61361.63	60161.83
0.35	63316.93	62278.88	61875.72
0.40	64193.73	63028.72	62677.66
0.45	65153.92	64361.91	63559.86
0.50	66187.82	65507.87	64513.96

IX. Conclusion

The performance of the utensil industry is discussed using the regenerative point graphical technique. The above tables concluded that when the repair rate increases then the MTSF, availability of the system and profit values also increase but when the failure rate increases then these reliability measures decrease. It is clear that the regenerative point graphical technique is helpful for industries to analyze the behaviour of the products and components of a system.

`

References

[1] Agarwal, A., Garg, D., Kumar, A. and Kumar, R. (2021). Performance analysis of the water treatment reverses osmosis plant. *Reliability Theory and Applications*, 16(3): 16-25.

[2] Chaudhary, A., Jaiswal, S. and Sharma, N. (2023). Probabilistic analysis of a two unit cold standby system with repair and replacement policies. *Reliability Theory and Applications*, 18 (72): 56-64.

[3] Chaudhary, P. and Tomar, R. (2019). A two identical unit cold standby system subject to two types of failures. *Reliability Theory and Applications*, 14(1): 34-43.

[4] Diaz, V. G., Gomez, J. F., Lopez, M., Crespo, A. and Leon, P. M. (2009). Warranty cost models State-of-Art: A practical review to the framework of warranty cost management. *ESREL:* 2051–2059.

[5] Gnedenko B. and Igor A.U. (1995). Probabilistic reliability engineering. *John Wiley and Sons.*

[6] Jack, N. and Murthy, D. P. (2007). A flexible extended warranty and related optimal strategies. *Journal of the Operational Research Society*, 58(12): 1612–1620.

[7] Jia, H., Peng, R., Yang, L., Wu, T., Liu, D. and Li, Y. (2022). Reliability evaluation of demand based warm standby systems with capacity storage. *Reliability Engineering and System Safety,* 218: (108132)

[8] Kapur, P. K. and Kapoor, K. R. (1978). Stochastic behaviour of some 2-unit redundant systems. *IEEE Transactions on Reliability*, 27(5): 382-385.

[9] Kashid, D. U. and Kumar, R. (2017). Availability modeling of two units system subject to degradation when intermediate repair is feasible using RPGT. *International Multidisciplinary Conference on Commerce, Management, Technology, Engineering and Environmental Sciences & International Conference on Humanities and Social Sciences, Mumbai*.

[10] Kumar, A., Garg, D. and Goel, P. (2019). Mathematical modeling and behavioral analysis of a washing unit in paper mill. *International Journal of System Assurance Engineering and Management*, 10: 1639-1645.

[11] Kumar, J. and Goel, M. (2016). Availability and profit analysis of a two-unit cold standby system for general distribution. *Cogent Mathematics*, *3*(1): 1262937.

[12] Kumar, S. and Goel, P. (2014). Availability analysis of two different units system with a standby having imperfect switch over device in banking industry. *Arya Bhatta Journal of Mathematics and Informatics*, 6(2): 299-304.

[13] Levitin, G., Finkelstein, M. and Xiang, Y. (2020). Optimal preventive replacement for cold standby systems with reusable elements. *Reliability Engineering and System Safety,* 204: (107135).

[14] Malik, S. C. and Rathee, R. (2016). Reliability modelling of a parallel system with maximum operation and repair times. *International Journal of Operational Research*, 25(1): 131-142.

[15] Sengar S. and Mangey R. (2022). Reliability and performance analysis of a complex manufacturing system with inspection facility using copula methodology. *Reliability Theory & Applications,* 17(71): 494-508.

[16] Wang, G. J. and Zhang, Y. L. (2007). An optimal replacement policy for repairable cold standby system with priority in use. *International Journal of Systems Science*, *38*(12): 1021-1027.

STOCHASTIC ANALYSIS OF A COMPLEX REPAIRABLE SYSTEM WITH A CONSTRAIN ON THE NUMBER OF REPAIRS

Late Dr. K. Shankar Bhat

Miriam Kalpana Simon

•

Madras Christian College
miriamkalpana@mcc.edu.in

Abstract

Reliability characteristics of repairable systems have been studied in the past in great detail by numerous researchers. Their findings are based mainly on the significant assumption that the repairs are carried out by one or more repair facilities, and the process of repair renews the functional behavior of the components or units in the system. In other words, the statistical properties of the components or units can be restored by carrying out the repair upon failure. This means that failed units may be treated "as good as new" after each repair. In many practical situations we observe that in the process of making a unit as good-as-new, considerable damage will be done to the operational ability of the repair facility, which may reflect upon the repair rates of the units in subsequent repairs. Intuitively, we expect that the average repair time of a unit to increase after each repair. This paper makes an attempt to incorporate these concepts in a two unit warm standby redundant system in which the efficiency, equivalently, repair capacity of the repair facility decreases upon each repair. Subsequently, the process of repair may not contribute significantly in improving the system reliability. In order to increase the system reliability and that the system might be available in the long run, an optimum replacement of the repair facility in terms of the mean time to system failure (MTSF) is suggested.

Keywords: Reliability, repair facility, warm standby redundant system, optimum replacement, Mean time to system failure.

I. Introduction

A great majority of real systems are repaired after they fail rather than replaced in toto. Jensen and Petersen [8] identified Printed Circuit Board (PCB) as a good example for repairable systems. This is not particularly so since failed PCB's are often discarded and replaced by new ones. Nevertheless, it does emphasize the point that the systems such as sonar systems, radar systems or communication systems of which PCB's form a small proportion are certainly repaired rather than discarded. Repair maintenance is sought to increase the Mean Time To System Failure (MTSF) vis-à-vis system reliability. In addition to standby redundancy, repair maintenance is often resorted to improve the system reliability. System components or units are repaired upon their failure. Operable standbys are switched over to online for efficient functioning of the system, during the repair of the online failed units.

One of the important objectives of a system engineer is to resort to repair maintenance that increases the mean time to system failure by removing the bottle necks or constrains hindering on the improvement of the system reliability. System performance and reliability characteristics have been studied for such systems by great many researchers. Bhat, Gururajan and Nayak [1] provided the availability and reliability measures and the MTBF of a two-unit cold standby system supported by a single repair facility. Cao and Wu [3] obtained reliability quantities of the system and the repair facility of a two-dissimilar-unit cold standby system where the repair facility is subject to failure and can be replaced by a new one after it fails.

Chaudhary, Sharma and Gupta [4] deals with a system composed of two-non identical units and a single repairman when the joint distribution of failure and repair times for each unit is bivariate exponential distribution. The stochastic analysis of a two-identical unit cold standby system wherein a single repair facility appears in and disappears from the system randomly is considered by Gupta and Bhardwaj [5].

Gupta and Tyagi [6] discusses the stochastic analysis of a two identical unit cold standby system model with a single repairman depending upon the perfect and imperfect environment. Reliability, availability and interval reliability measures of a two-unit warm standby system with a single repair facility wherein the lifetime of the functioning unit has a general distribution, while the standby unit has a phase-type distribution is derived by Gururajan and Srinivasan [7].

Kumar, Malik and Nandal [9] described the stochastic analysis of a repairable system consisting of two non-identical units with a single repairman and the distribution for failure rates of the units has been considered as negative exponential while arbitrary distributions have been taken for repair and treatment rates. A warm standby repairable system including two dissimilar units, one repairman and imperfect switching mechanism is studied by Sadeghi and Roghanian [11].

A well accounted bibliography in this direction is also found in Osaki and Nakagawa [10], Srinivasan and Subramanian [12] and Bhat and Gururajan [2]. Their findings are based mainly on the noteworthy assumption that the repairs are carried out by one or more repair facilities, and the process of repair renews the functional behavior of the components or units in the system. In other words, the statistical properties of the components or units can be restored by carrying out the repair upon failure. This means that failed units may be treated "as-good-as-new" after each repair. In many practical situations we observe that in the process of making a unit as-good-as-new, considerable damage may be done to the operational ability of the repair facility, which may reflect upon the repair rates of the units in subsequent repairs. Instinctively, we expect that the average repair time of a unit to increase after each repair. At one stage the repair facility will have little contribution to our desire of increasing the system reliability. At this stage it is worthwhile to replace the repair facility by a new one. This paper incorporates the above mentioned ideas in a two unit warm standby repairable system and arrives at an optimum replacement policy, clearly indicating the stage of replacement as a function of MTSF.

II. System Description

Let us characterize the complex two unit standby redundant repairable system under study.

[01] The system consists of two dissimilar units having same statistical properties. The units are labeled as U_1 and U_2. Initially, U_1 is put online and U_2 is kept as a warm standby. Whenever a unit fails while functioning online the standby unit is switched over to online for functioning, and the online failed unit is sent for repair.

[02] The unit that is kept in standby is vulnerable to failure. This unit is sent to repair upon failure in the standby state, and is restored immediately after the completion of its repair.
[03] There are two repair facilities RF_1 and RF_2. Online failed units are repaired using RF_1 and standby failed units are repaired using RF_2.
[04] The repair time distribution of online failed unit (in RF1) is different on each failure. Furthermore, it is assumed that the repair rate of each unit increases as the number of repair increases. In other words, the efficiency of the repair facility decreases after each repair completion. On the other hand, a unit that is failed in the standby state is as-good-as-new after repair
[05] The operational ability of the repair facility is considered not satisfactory once it completes $2k$ number of repairs.
[06] A unit that has completed $2k^{th}$ repair may not help us in our objective of improvement of the system reliability. At this stage, a replacement policy for the repair facility may be considered feasible.
[07] All switchover times involved in the system operation are negligible and the switch that performs the switchover operation is immune to failure.

III. Notation

$f_i(.), F_i(.), \bar{F}_i(.)$ $p.d.f, c.d.f, s.f$ of $failure\ time\ of\ unit\ i,$ $i = 1,2$
$g_{ij}(.), G_{ij}(.)$ $p.d.f, c.d.f$ of $unit\ i\ while\ undertaking\ repair\ for\ the\ jth\ time,$ $j = 1,2,\ldots$
$i = 1,2$

IV. Stochastic Behaviour of the Standby Unit

During the failure free operation of the one unit online, the behavior of the standby unit can be completely described by a stochastic process $\{Q_{ij}(t), t > 0\}$.

$$Q_{ij}(t)\,\delta t = Pr\{the\ standby\ unit\ is\ in\ state\ j\ at\ t\ conditioned\ that\ it\ entered\ the\ state\ i\ in\ (-\delta t, 0)\}$$
$$i, j = O, F;\ O = operable, F = under\ repair \quad (1)$$

We observe that the time spent in the state O and F by the standby unit forms an alternating renewal process whose transition probabilities may be described through the functions $Q_{ij}(t)$.

Thus, the transition probabilities represented through $Q_{ij}(.)$ are evaluated as

$$Q_{OO}(t) = [\beta + \alpha e^{-(\alpha+\beta)t}][\alpha + \beta]^{-1}; Q_{OF}(t) = 1 - Q_{OO}(t) \quad (2)$$
$$Q_{OF}(t) = [\beta - \alpha e^{-(\alpha+\beta)t}][\alpha + \beta]^{-1}; Q_{FF}(t) = 1 - Q_{FO}(t) \quad (3)$$

We observe that the function $Q_{OO}(t)$ is the p-function of the Kingman's regenerative phenomenon. Since these functions find repeated usage in our discussion, we provide their Laplace Stieltjes Transforms.

$$Q^*_{OO}(s) = \frac{\beta}{[(\alpha+\beta)s]} + \frac{\alpha}{[(\alpha+\beta)(s+\alpha+\beta)]} \quad (4)$$
$$Q^*_{FO}(s) = \frac{\beta}{[(\alpha+\beta)s]} - \frac{\beta}{[(\alpha+\beta)(s+\alpha+\beta)]} \quad (5)$$
$$Q^*_{OF}(s) = \frac{\alpha}{[(\alpha+\beta)s]} - \frac{\alpha}{[(\alpha+\beta)(s+\alpha+\beta)]} \quad (6)$$
$$Q^*_{OO}(s) = \frac{\alpha}{[(\alpha+\beta)s]} + \frac{\beta}{[(\alpha+\beta)(s+\alpha+\beta)]} \quad (7)$$

V. Reliability Analysis

In our effort to characterize the system we define the following events:

$\omega_{i,0}$: *event that unit* i, *which has not gone through any repair till then, just begins to operate online* $i = 1,2$

$\omega_{i,1}$: *event that* j*th repair of unit* i *just begins; at this instant an operable standby unit is put online* $i = 1, 2;\ j = 1, 2, \dots, k.$

These events, constituting themselves into a regenerative process, facilitate us to trace the behavior of the system completely on a time horizon. For the system to be continuously operable in $(0, t]$, it is necessary that at the instant of failure of unit i, unit $(3 - i)$ should be in operable condition. To facilitate this, the following auxiliary, system down forbidding functions are defined.

$$P_r(j,t)\delta t = \Pr\{\omega_{1,j+1} \text{ occurs between } t \text{ and } t+\delta t \text{ and the system is operable in } (0,t] \,/\, \omega_{2,j} \text{ at } t = 0\}$$
$$j = 1, 2, \dots k-1 \qquad (8)$$

We observe that the function $P_r(j, t)\ \delta t$ represents the pdf of the time interval between $\omega_{2,j}$ and $\omega_{1,j+1}$ events. Further

$$Q_r(j,t)\delta t = Pr\{\omega_{2,j+1} \text{occurs between } t \text{ and } t+\delta t \text{and the system is operable in} (0,t]/\omega_{1,j+1} \text{at } t = 0\}$$
$$j = 1, 2, \dots k-1 \qquad (9)$$
$$H_r(j,t)\delta t = Pr\{\omega_{2,j} \text{occurs between } t \text{ and } t+\delta t \text{ and the system is operable in} (0,t]\ /\omega_{2,j-1} \text{ at } t = 0\}$$
$$j = 1, 2, \dots k \qquad (10)$$

We notice that $H_r(j,t)\ \delta t$ represents the pdf of the time interval between two successive ω_2 events. Similarly,

$$\phi_r(j,t)\delta t = Pr\{\omega_{2,j} \text{ occurs between } t \text{ and } t+\delta t \text{ and the system is operable in } (0,t]/\omega_{2,0} \text{ at } t = 0\}$$
$$j = 1, 2, \dots k \qquad (11)$$

We observe that the functions (9) and (10) are system down forbidding functions in the sense that a system down is not acceptable between the occurrences of two events. These functions are evaluated with the help of regenerative events ω_{ij} observing that, at the instant of failure of U_1 at which epoch U_2 has completed its j^{th} repair and is found in operable condition in its standby state.

Thus, the pdf's between two successive events are

$$P_r(j,t) = f_1(t)\{G_{2,j}(t) \copyright [\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\} \qquad j = 1, 2, \dots k \qquad (12)$$
$$Q_r(j,t) = f_2(t)\{G_{1,j+1}(t) \copyright [\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\} \qquad j = 1, 2, \dots k-1 \qquad (13)$$

The pdf between successive regenerative events are obtained using the forward recurrence relation between the P and Q functions and also between H and ϕ functions Thus

$$H_r(j,t) = P_r(j-1,t) \copyright Q_r(j-1,t) \qquad j = 2, 3, \dots k \qquad (14)$$
$$\phi_r(j,t) = \phi_r(j-1,t) \copyright H_r(j,t) \qquad j = 2, 3, \dots k \qquad (15)$$
$$\phi_r(1,t) = H_r(1,t) \qquad (16)$$

where $H_r(1,t) = f_2(t)\ \{G_{1,1}(t) \copyright [\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\}$.

Observing that a unit switched online after repair at which epoch the repair of the other unit commences, is a point of regeneration, we are in a position to write an expression for the reliability function of the system. The reliability function $R(t,k)$ of the system is given by

$$R(t,k) = \bar{F}_1(t) + f_1(t)\, Q_{oo}(t) \copyright \bar{F}_2(t) + f_1(t) Q_{oo}(t) \copyright \Sigma\, \phi_r(j,t) \copyright \{\bar{F}_1(t) + P_r(j,t) \copyright \bar{F}_2(t)\} \quad j = 1, 2, \ldots k \qquad (17)$$

The expression (17) is obtained by considering the following mutually exclusive and exhaustive cases:

(a) U_1, which is fresh and has not gone through any repairs, does not fail before t.

(b) U_2, that is instantaneously switched over from standby and has not gone through any repair till then, does not fail before t.

(c) U_i while operating online after j^{th} repair $(j = 1, 2, \ldots k)$ does not fail before t.

VI. Availability Analysis

The following auxiliary system-down allowing functions are defined to obtain p.d.f. of time intervals between $\omega_{i,j}$ events.

$$P_a(j,t)\delta t = Pr\{\omega_{1,j+1}\ occurs\ between\ t\ and\ t+\delta t\ in\ (0,t]/\omega_{2,j}\ at\ t = 0\} \quad j = 1, 2, \ldots k-1 \qquad (18)$$

$$Q_a(j,t)\delta t = Pr\{\omega_{2,j+1}\ occurs\ between\ t\ and\ t+\delta t\ in\ (0,t]/\omega_{1,j+1}\ at\ t = 0\} \quad j = 1, 2, \ldots k-1 \qquad (19)$$

$$H_a(j,t)\delta t = Pr\{\omega_{2,j}\ occurs\ between\ t\ and\ t+\delta t\ in\ (0,t]/\omega_{2,j-1}\ at\ t = 0\} \quad j = 1, 2, \ldots k \qquad (20)$$

$$\phi_a(j,t)\delta t = Pr\{\omega_{2,j} occurs\ between\ t\ and\ t+\delta t\ in\ (0,t]/\omega_{2,0} at\ t = 0\} \quad j = 1, 2, \ldots k \qquad (21)$$

Schematic representation of system behavior between $\omega_{2,j}$ and $\omega_{1,j+1}$ events is shown in Figure 1.

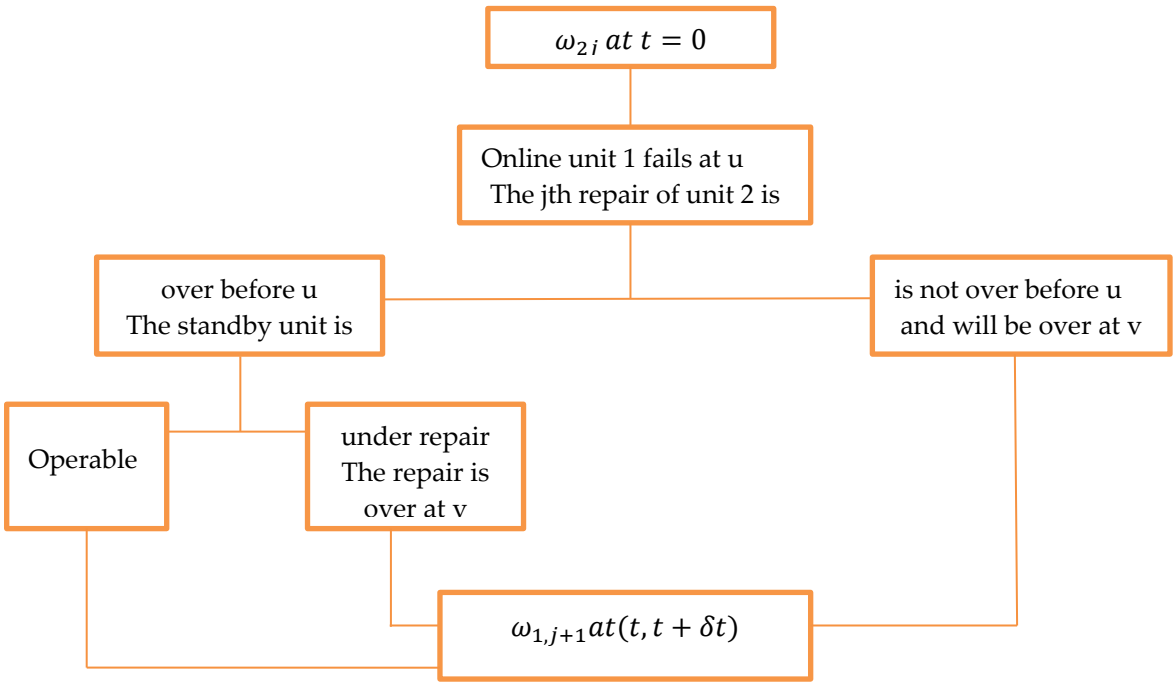


Figure 1: *A Schematic Representation of Evaluation of $P_a(j,t)$*

We observe that the functions [18] and [19] are system down allowing functions. We scrutinize the first possibility that, at the instant of failure of U_1, U_2 has complete its j^{th} repair and is found in operable condition in its standby state. The term that corresponds to this possibility is:

$$f_1(t)[G_{2,j}(t) \odot \{[\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\}] \qquad j = 1, 2, \dots k$$

The second possibility corresponds to the situation when the online U_1 fails at u, at this time point U_2 is found 'not in operable condition' and is undergoing repair. The unit is switched over to online for its operation at the epoch of the repair completion. Thus, the term that corresponds to this possibility is:

$$f_1(t)[G_{2,j}(t) \odot \{[\alpha - \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\} \odot e^{-\beta t}] \qquad j = 1, 2, \dots k$$

Thirdly, when online U_1 fails at u, the j^{th} repair of U_2 is not over before u and the same will be over at v, $v > u$. This probability is given by

$$F_1(t) g_{2,j}(t) \qquad j = 1, 2, \dots k$$

Thus, in its totality, the pdf of the time interval between a $\omega_{2,j}$ event and $\omega_{1,j+1}$ event is given by

$$P_a(j,t) = f_1(t)[G_{2,j}(t) \odot \{[\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\}] + f_1(t)\{G_{2,j}(t) \odot \{[\alpha - \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\} \odot e^{-\beta t}] + F_1(t)[g_{2,j}(t)] \qquad j = 1, 2, \dots k \qquad (22)$$

Similarly,

$$Q_a(j,t) = f_2(t)\{G_{1,j+1}(t) \odot \{[\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\}] + f_2(t)\{G_{1,j+1}(t) \odot \{[\alpha - \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\} \odot e^{-\beta t}] + F_2(t)[g_{1,j+1}(t)] \qquad j = 1, 2, \dots k-1 \qquad (23)$$

By means of (22) and (23) we obtain,

$$H_a(j,t) = P_a(j-1,t) \odot Q_a(j-1,t) \qquad j = 2,3, \dots k \qquad (24)$$

$$\phi_a(j,t) = \phi_a(j-1,t) \odot H_a(j,t) \qquad j = 2,3, \dots k \qquad (25)$$

$$\phi_a(1,t) = H_a(1,t) \qquad (26)$$

where

$$H_a(1,t) = f_2(t)[G_{1,1}(t) \odot \{[\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\}] + f_2(t)[G_{1,1}(t) \odot \{[\alpha - \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\} \odot e^{-\beta t}] + F_2(t)[g_{1,1}(t)]$$

The availability function $A(t,k)$ of the system is derived by taking into consideration the following mutually exclusive and exhaustive cases:

(a) U_1, which is fresh and has not gone through any repairs, does not fail before t.

(b) U_2, that is instantaneously switched over from standby and has not gone through any repair till then, does not fail before t.

(c) U_i while operating online after j^{th} repair ($j = 1, 2, \dots k$) does not fail before t.

$$A(t,k) = \bar{F}_1(t) + f_1(t)[Q_{OO}(t) + Q_{OF}(t) \odot e^{-\beta t}\} \odot \bar{F}_2(t) + f_1(t)[Q_{OO}(t) + Q_{OF}(t) \odot e^{-\beta t}] \odot \Sigma \phi_a(j,t) \odot \{\bar{F}_1(t) + Pa(j,t) \odot \bar{F}_2(t)\} \qquad j = 1, 2, \dots k \qquad (27)$$

VII. Mean Time To System Failure

In the analysis of the system we have assumed arbitrary failure time and repair time distributions for the units while working online. For the purpose of illustration we consider a model in which both the units are identical by virtue of their statistical properties and their failure time distributions are exponential. In addition to the assumptions made for the standby unit, we formalize the failure time and repair time distributions of the online units.

$$f_i(t) = \lambda e^{-\lambda t} \qquad \lambda > 0, \mathrm{i} = 1,2$$
$$g_{ij}(t) = \mu_j e^{-\mu_j t} \qquad \mu_j > 0, i = 1,2\,;\ j = 1,2,\ldots k$$

The integral equations given in (17) are solved using Laplace transform technique and $R^*(s,k)$, the Laplace transform of $R(t,k)$ is:

$$R^*(s,k) = \frac{1}{s+\lambda} + \frac{\lambda h(s)}{s+\lambda} + \lambda h(s) \sum_{j=1}^{k} L_1^R(s) \prod_{n=2}^{j} L_n^R(s) \left[\frac{1}{s+\lambda} + \frac{\lambda \mu_j h(s)}{(s+\lambda)^3(s+\lambda+\mu_j)}\right] \tag{28}$$

where $L_1^R(s) = \frac{\lambda \mu_1 h(s)}{(s+\lambda)(s+\lambda+\mu_j)}$, $L_n^R(s) = \frac{\lambda^2 \mu_n \mu_{n-1}[h(s)]^2}{(s+\lambda)^2(s+\lambda+\mu_n)(s+\lambda+\mu_{n-1})}$ and $h(s) = \frac{1}{(\alpha+\beta)}\left[\frac{\beta}{(s+\lambda)} + \frac{\alpha}{(s+\lambda+\alpha+\beta)}\right]$

We observe that $R^*(s,k)$ is a rational function of its arguments and can be easily inverted for small values of k. Thus, the reliability can be explicitly computed for small values of k.

The Mean Time To System Failure (MTSF) is given by

$$R^*(0,k) = \frac{1}{\lambda} + h(0) + h(0) \sum_{j=1}^{k} L_1^R(0) \prod_{n=2}^{j} L_n^R(0) \left[1 + \frac{h(0)}{1+\eta_j)}\right] \tag{29}$$

where $L_1^R(0) = \frac{h(0)}{1+\eta_1}$. $L_n^R(0) = \frac{[h(0)]^2}{(1+\eta_n)(1+\eta_{n-1})}$, $h(s) = \frac{1}{(\alpha+\beta)}\left[\frac{\beta}{\lambda} + \frac{\alpha}{\lambda+\alpha+\beta}\right]$ and $\eta_j = \frac{\lambda}{\mu_j}$

A coding is written for the precise evaluation of R*(0, k). The program evaluates the MTSF for specified values of the parameters. As a function of t and k, MTSF is evaluated for specific values of the parameters and are tabulated in Table 1.

Table 1: *MTSF of the system for the parameters(λ, μ, α, β) = (0.95, 50, 10, 40)*

k	MTSF = R*(0)	k	MTSF = R*(0)
1	3.1756	21	6.2046
2	4.0626	22	6.2056
3	4.6862	24	6.2028
4	5.1269	26	6.2075
6	5.6607	27	6.2077
8	5.9303	28	6.2078
10	6.0669	30	6.2079
12	6.1362	32	6.2080
13	6.1568	33	6.2081
15	6.1820	34	6.2081
16	6.1895	35	6.2081
18	6.1986	36	6.2081
19	6.2013	37	6.2081
20	6.2033	38	6.2081

The graphical representation of $R^*(0,k)$ for specific values of parameters is depicted in Figure 2.

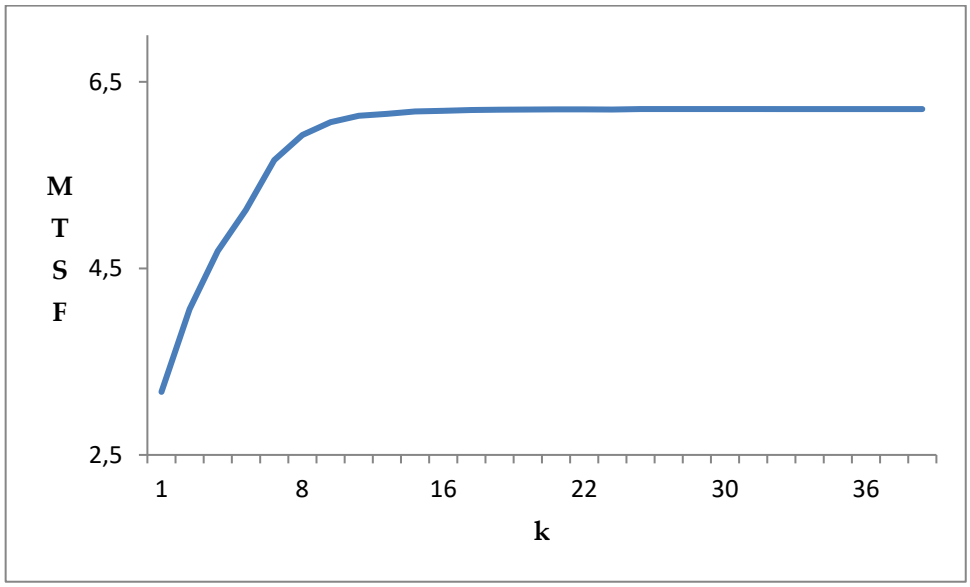


Figure 2: *Graphical Representation of the MTSF for (λ, μ, α, β) = (0.95, 50, 10, 40)*

The graph clearly indicates that there is no improvement in MTSF once it completes 2k = 34 repairs. Intuitively one would conclude that it is not worthwhile to retain the repair facility once it completes 2k = 34 repairs. Consequently, we suggest at this stage that the repair facility should be replaced by a new one in order to increase the system performance and to make the system to be available in the long run.

VIII. A Provision for Replacement of Repair Facility

A wise strategy suggests that when a repair facility is unable to perform its operation it should be scrapped. If one follows this strategy the system becomes unavailable in the long run. However, a prudent policy is to replace the repair facility by a new one so that the system might be available in the long run. When a repair facility completes 2k repairs, it is replaced by a similar new repair facility. The variable k realizes into a number at which MTSF stabilizes in the sense that

$$R^*(0,k) = R^*(0,k+r), \quad r = 1,2,3 \ldots \tag{30}$$

The policy of replacement is as follows:
"After nk-th repair completion of unit 1, the old repair facility is scrapped and a new repair facility is introduced. Here n denotes the number of such replacements, n ≥ 1. We suggest replacement of repair facility only and not operable units. When a unit, while operating online after nk-th repair, fails, it is switched over to the new repair facility; at this epoch an operable standby is instantaneously switched online."

I. Reliability Analysis of the Modified System

Let us define

$\phi_r(0,t)\delta t = Pr\{\omega_{1,1} occurs\ between\ t\ and\ t+\delta t and\ the\ system\ is\ operable\ in (0,t]/\omega_{1,1} at\ t=0\}$ (31)

The function $\phi_r(0,t)$ is the pdf of time interval between two successive $\omega_{1,1}$ events, during which the system being operable between these two events. Thus

$$\phi_r(0,t) = \phi_r(k,t) \copyright [g_{2,k}(t) \copyright \{[\beta + \alpha e^{-(\alpha+\beta)t}][\alpha+\beta]^{-1}\}] f_1(t) \quad (32)$$

The reliability function of the modified system is given by

$$R_1(t,k) =_1 R_1(t,k) + \left[\sum_{n=1}^{\infty}\{\phi_r(0,t)\}^n\right] \copyright \left\{\bar{F}_2(t) + \sum_{j=1}^{k} \phi_r(j,t)\{\bar{F}_1(t) + \{g_{2,k}(t) \copyright Q_{OO}(t)\} f_1(t) \copyright \bar{F}_2(t)\}\right\} \quad (33)$$

where ${}_1R_1(t, k)$ is the expression given in the right hand side of (17). The equation (33) is derived by considering the following mutually exclusive and exhaustive possibilities
(a) the interval (0, t] is not intercepted by an $\omega_{1,1}$ event.
(b) the interval (0, t] is intercepted by at least one $\omega_{1,1}$ event.

II. Availability Analysis of the Modified System

The pdf of time interval between system-down allowing regenerative events is evaluated through

$$\phi_a(0,t)\delta t = Pr\{\omega_{1,1} occurs\ between\ (t, t+\delta t)/\omega_{1,1}\ at\ t=0\} \quad (34)$$

and is given by

$$\phi_a(0,t) = \phi_a(k,t) \copyright [\{g_{2,k}(t) \copyright Q_{OO}(t)\} f_1(t) + \{g_{2,k}(t) \copyright Q_{OF}(t) \copyright e^{-\beta t}\} F_1(t)] \quad (35)$$

Arguments that lead to the derivation of (19) will give us the availability function of the system with a provision for a replacement of repair facility. Thus,

$$A_1(t,k) =_1 A_1(t.k) + \sum_{n=1}^{\infty}\left[\{\phi_a(0,t)\}^n\right] \copyright \{\bar{F}_2(t) + \sum_{j=1}^{k} \phi_a(j,t)\{\bar{F}_1(t) + [\{g_{2,k}(t) \copyright Q_{OO}(t)\} f_1(t) + \{g_{2,k}(t) \copyright Q_{OF}(t) \copyright e^{-\beta t}\} F_1(t)] \copyright \bar{F}_2(t)\}\} \quad (36)$$

where ${}_1A_1(t, k)$ is the expression given on the right hand side of (27).

The steady state availability of the system is given by

$$A_\infty = \lim_{t\to\infty} A(t) = \lim_{s\to 0} A^*(s)$$

References

[1] Bhat, K. S., Gururajan, M. and Nayak, P. (1988). A study of a 2-unit system with random breakdown'of the repair facility. *Microelectronics Reliability*, *28*(3):369-371.

[2] Bhat, K. S. and Gururajan, M. (1993). A two-unit cold standby system with imperfect repair and excessive availability period. *Microelectronics Reliability*, *33*(4):509-512.

[3] Cao, J. and Wu, Y. (1989). Reliability analysis of a two-unit cold standby system with a replaceable repair facility. *Microelectronics Reliability*, *29*(2):145-150.

[4] Chaudhary, P., Sharma, A. and Gupta, R. (2022). A Discrete Parametric Markov-Chain Model of a Two NonIdentical Units Warm Standby Repairable System with Two Types of Failure. *Reliability: Theory & Applications*, *17*(2 (68)):21-30.

[5] Gupta, R. and Bhardwaj, P. (2019). A discrete parametric Markov-chain model of a two unit cold standby system with appearance and disappearance of repairman. *Reliability: Theory & Applications*, *14*(1):13-22.

[6] Gupta, R. and Tyagi, A. (2019). A Discrete Parametric Markov-Chain Model Of A Two-Unit Cold Standby System With Repair Efficiency Depending On Environment. *Reliability: Theory & Applications*, *14*(1):23-33.

[7] Gururajan, M. and Srinivasan, B. (1995). A complex two-unit system with random breakdown of repair facility. *Microelectronics Reliability*, *35*(2):299-302.

[8] Jensen, F. and Petersen, N. E. Burn-in, A Wiley-Interscience Publication, 1982

[9] Kumar, N., Malik, S. C. and Nandal, N. (2022). Stochastic analysis of a repairable system of non-identical units with priority and conditional failure of repairman. *Reliability: Theory & Applications*, *17*(1 (67)):123-133.

[10] Osaki, S. and Nakagawa, T. (1976). Bibliography for reliability and availability of stochastic systems. *IEEE Transactions on Reliability*, *25*(4):284-287.

[11] Sadeghi, M. and Roghanian, E. (2017). Reliability analysis of a warm standby repairable system with two cases of imperfect switching mechanism. *Scientia Iranica*, *24*(2):808-822.

[12] Srinivasan, S. K. and Subramanian, R. Probabilistic analysis of redundant systems, Springer-Verlag, 1980.

RT&A, No 2 (78)
Volume 19, June, 2024

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

Nuzhat Ahad

•

University of Kashmir, Srinagar, India
nuzhatahad01@gmail.com

S.P.Ahmad

•

University of Kashmir, Srinagar, India
sprvz@yahoo.com

J.A.Reshi*

•

Govt. Degree College Pulwama, India
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, bathtub, increasing and constant behaviour. The probability density function (pdf) of KWMBD is given by:

$$f(x) = \frac{x^{2(k+1)}\alpha^{-(3+2k)}e^{-\frac{x^2}{2\alpha^2}}}{2^{k+\frac{1}{2}}\Gamma(k+\frac{3}{2})} \qquad x > 0, \alpha > 0, k \in R. \tag{1}$$

And, the corresponding cummulative distribution function (cdf) of KWMBD is given by:

$$F(x) = 1 - \frac{\Gamma\left((k+\frac{3}{2}), \frac{x^2}{2\alpha^2}\right)}{\Gamma(k+\frac{3}{2})} \qquad x > 0, \alpha > 0, k \in R. \tag{2}$$

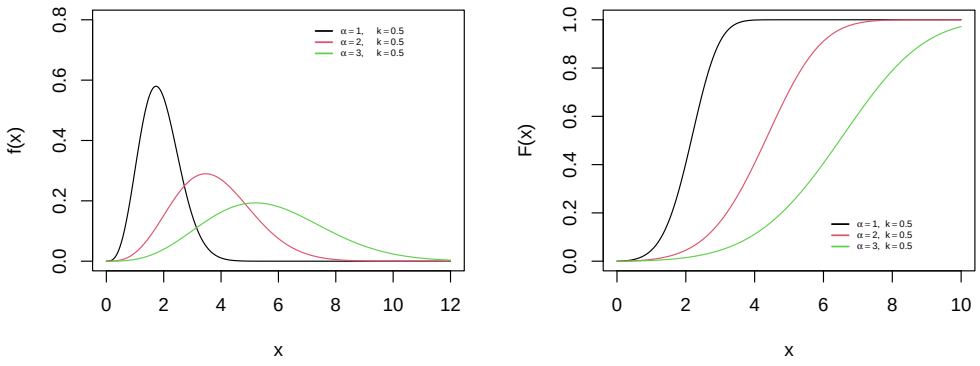


Figure 1: *Probability density plot and cumulative distribution plot of KWMBD for different combinations of parameters.*

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 a 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, $I(\alpha) = -nE\{\frac{d^2}{d^2(\alpha)} log f(x)\}$ is fisher-information matrix.
Thus, the resulting extension of Jeffrey's-prior for KWMBD will be:

$$g(\alpha) = \left[\frac{1}{\alpha^2}\right]^{c_1}; \quad c_1 \in R^+ \tag{3}$$

2.2. Inverse-gamma prior

The density of parameter α^2 on assuming it to follow Gamma(β, λ) distribution is given by:

$$g(\alpha^2) = \frac{\lambda^\beta}{\Gamma(\beta)} (\alpha^2)^{-\beta-1} e^{-\frac{\lambda}{\alpha^2}} \tag{4}$$

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 hyperparameters. 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:

$$l_{sq}(\hat{\alpha}, \alpha) = c(\hat{\alpha} - \alpha)^2; \quad c \in R^+ \tag{5}$$

2.3.2. Precautionary loss function

The Precautionary loss function is given by:

$$l_{pr}(\hat{\alpha}, \alpha) = \frac{c(\hat{\alpha} - \alpha)^2}{\hat{\alpha}} \tag{6}$$

2.3.3. Al-Bayyati's loss function

The Al-Bayyati's loss function is given by:

$$l_{Al}(\hat{\alpha}, \alpha) = \alpha^{c_2}(\hat{\alpha} - \alpha)^2; \quad c_2 \in R^+ \tag{7}$$

2.3.4. Stein's loss function

The Stein's loss function is given by:

$$l_{St}(\hat{\alpha},\alpha) = \frac{\hat{\alpha}}{\alpha} - log\left(\frac{\hat{\alpha}}{\alpha}\right) - 1 \tag{8}$$

3. Parametric Estimation of KWMBD

In this section, we discuss the various estimation methods for KWMB Distribution.

3.1. Maximum Likelihood Estimation

Let $x_1, x_2, x_3, ..., x_n$ be a random sample of size n from kth Order Weighted Maxwell-Boltzmann Distribution. Therefore the maximum likelihood estimator(MLE) of α is:

$$\hat{\alpha} = \sqrt{\frac{x_i^2}{n(2k+3)}} \tag{9}$$

3.2. Baye's Estimator under Extension of Jeffrey's Prior

The Joint Probability Density Function of x and given α is given by:

$$L(\underline{x}|\alpha) = \frac{\prod_{i=1}^{n} x_i^{2(k+1)} \alpha^{-n(3+2k)} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right)} \tag{10}$$

The posterior probability density function of α for given data x is given by:

$$\pi_1(\alpha|\underline{x}) \propto L(\underline{x}|\alpha) g(\alpha)$$

$$\pi_1(\alpha|\underline{x}) \propto \frac{\prod_{i=1}^{n} x_i^{2(k+1)} \alpha^{-n(3+2k)} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right)} \frac{1}{\alpha^{2c_1}}$$

$$\pi_1(\alpha|\underline{x}) = k\alpha^{-n(3+2k)-2c_1} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}}$$

where k is normalising constant independent of α and is given by:

$$k^{-1} = \int_0^{\infty} \alpha^{-n(3+2k)-2c_1} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}} d\alpha$$

$$k^{-1} = \frac{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{-n(3+2k)-2c_1+1}{2}} \Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}{2^{\frac{-n(3+2k)-2c_1+3}{2}}}$$

Therefore, the posterior probability density function is:

$$\pi_1(\alpha|\underline{x}) = \frac{2^{\frac{-n(3+2k)-2c_1+3}{2}} \alpha^{-n(3+2k)-2c_1} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}}}{\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{-n(3+2k)-2c_1+1}{2}} \Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)} \tag{11}$$

3.2.1. Baye's Estimator under squared error loss function

The Risk Function Under SELF is given by:

$$R_{(sq,ej)}(\hat{\alpha}) = \int_0^\infty c(\hat{\alpha}-\alpha)^2 \pi_1(\alpha|\underline{x})d\alpha$$

$$R_{(sq,ej)}(\hat{\alpha}) = c\hat{\alpha}^2 + \frac{\sum_{i=1}^{n} x_i^2}{(n(3+2k)+2c_1-3)} - 2\hat{\alpha}c\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2}}\frac{\Gamma\left(\frac{n(3+2k)+2c_1-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)} \tag{12}$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(sq,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(ej,sq)} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2}}\frac{\Gamma\left(\frac{n(3+2k)+2c_1-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)} \tag{13}$$

3.2.2. Baye's Estimator under precautionary Loss function

The Risk Function Under PLF is given by:

$$R_{(pre,ej)}(\hat{\alpha}) = \int_0^\infty c\frac{(\hat{\alpha}-\alpha)^2}{\hat{\alpha}}\pi_1(\alpha|\underline{x})d\alpha$$

$$R_{(pre,ej)}(\hat{\alpha}) = c\hat{\alpha} + c\frac{\sum_{i=1}^{n} x_i^2}{\hat{\alpha}(n(3+2k)+2c_1-3)} - 2c\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2}}\frac{\Gamma\left(\frac{n(3+2k)+2c_1-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)} \tag{14}$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(pre,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(pre,ej)} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{(n(3+2k)+2c_1-3)}} \tag{15}$$

3.2.3. Baye's Estimator under Al-Bayyati's loss function

The Risk Function Under Al-Bayyati's loss function is given by:

$$R_{(alb,ej)}(\hat{\alpha}) = \int_0^\infty \alpha^{c_2}(\hat{\alpha}-\alpha)^2 \pi_1(\alpha|\underline{x})d\alpha$$

$$R_{(alb,ej)}(\hat{\alpha}) = \hat{\alpha}^2 + \frac{\sum_{i=1}^{n} x_i^2}{(n(3+2k)+2c_1-c_2-3)} - 2\hat{\alpha}\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2}}\frac{\Gamma\left(\frac{n(3+2k)+2c_1-c_2-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-c_2-1}{2}\right)} \tag{16}$$

now, the Baye's estimator is obtained by solving

$$\frac{d(R_{(alb,ej)}(\hat{\alpha}))}{d\hat{\alpha}} = 0$$

and, is given by:

$$\hat{\alpha}_{(alb,ej)} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2}} \frac{\Gamma\left(\frac{n(3+2k)+2c_1-c_2-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-c_2-1}{2}\right)} \tag{17}$$

3.2.4. Baye's Estimator under combination of Stein's loss function

The Risk Function Under SLF is given by:

$$R_{(ste,ej)}(\hat{\alpha}) = \int_0^{\infty} \left(\frac{\hat{\alpha}}{\alpha} - log\left(\frac{\hat{\alpha}}{\alpha}\right) - 1\right) \pi_1(\alpha|\underline{x}) d\alpha$$

$$R_{(ste,ej)}(\hat{\alpha}) = \hat{\alpha} \sqrt{\frac{2}{\sum_{i=1}^{n} x_i^2}} \frac{\Gamma\left(\frac{n(3+2k)+2c_1}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)} - log(\hat{\alpha}) - m - 1 \tag{18}$$

where, 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:

$$\hat{\alpha}_{(ste,ej)} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2}} \frac{\Gamma\left(\frac{n(3+2k)+2c_1-1}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2c_1}{2}\right)} \tag{19}$$

3.3. Baye's Estimator under Inverse-Gamma Prior

The Joint Probability Density Function of x and given α^2 is given by:

$$L(\underline{x}|\alpha^2) = \frac{\prod_{i=1}^{n} x_i^{2(k+1)} (\alpha^2)^{\frac{-n(3+2k)}{2}} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right)} \tag{20}$$

The posterior probability density function of α^2 for given data x is given by:

$$\pi_2(\alpha^2|\underline{x}) \propto L(\underline{x}|\alpha^2) g(\alpha^2)$$

$$\pi_2(\alpha^2|\underline{x}) \propto \frac{\prod_{i=1}^{n} x_i^{2(k+1)} (\alpha^2)^{\frac{-n(3+2k)}{2}} e^{\frac{\sum_{i=1}^{n} x_i^2}{2\alpha^2}}}{\Gamma\left(k+\frac{3}{2}\right)} \frac{\lambda^{\beta}}{\Gamma(\beta)} (\alpha^2)^{-\beta-1} e^{-\frac{\lambda}{\alpha^2}}$$

$$\pi_2(\alpha^2|\underline{x}) = k(\alpha^2)^{\frac{-n(3+2k)-2\beta-2}{2}} e^{\left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)\frac{1}{\alpha^2}}$$

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:

$$\pi_2(\alpha^2|\underline{x}) = \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)^{\frac{n(3+2k)+2\beta}{2}} (\alpha^2)^{\frac{-n(3+2k)-2\beta-2}{2}} e^{\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)\frac{1}{\alpha^2}}}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} \tag{21}$$

3.3.1. Baye's Estimator under squared error loss function

The Risk Function Under SELF is given by:

$$R_{(sq,igp)}(\hat{\alpha^2}) = \int_0^\infty c(\hat{\alpha^2}-\alpha^2)^2 \pi_2(\alpha^2|\underline{x}) d(\alpha^2)$$

$$R_{(sq,igp)}(\hat{\alpha^2}) = c(\hat{\alpha}^2)^2 + \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)^2}{\left(\frac{n(3+2k)+2\beta-2}{2}\right)\left(\frac{n(3+2k)+2\beta-4}{2}\right)} - \hat{\alpha^2} c \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)}{\left(\frac{n(3+2k)+2\beta-2}{2}\right)} \tag{22}$$

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:

$$\hat{\alpha}_{(sq,igp)} = \sqrt{\frac{2\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)}{(n(3+2k)+2\beta-2)}} \tag{23}$$

3.3.2. Baye's Estimator under precautionary Loss function

The Risk Function Under PLF is given by:

$$R_{(pre,igp)}(\hat{\alpha^2}) = \int_0^\infty c\frac{(\hat{\alpha^2}-\alpha^2)^2}{\hat{\alpha^2}} \pi_2(\alpha^2|\underline{x}) d\alpha^2$$

$$R_{(pre,igp)}(\hat{\alpha^2}) = c\hat{\alpha}^2 + c\frac{1}{\hat{\alpha^2}} \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)^2}{\left(\frac{(n(3+2k)+2\beta-2)(n(3+2k)+2\beta-4)}{4}\right)} - 2c \frac{\left(\frac{\sum_{i=1}^n x_i^2}{2}+\lambda\right)}{\left(\frac{n(3+2k)+2\beta-2}{2}\right)} \tag{24}$$

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:

$$\hat{\alpha}_{(pre,igp)} = \sqrt{\frac{2\left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)}{\sqrt{(n(3+2k)+2\beta-2)(n(3+2k)+2\beta-4))}}} \tag{25}$$

3.3.3. Baye's Estimator under Al-Bayyati's loss function

The Risk Function Under Al-Bayyati's loss function is given by:

$$R_{(alb,igp)}(\hat{\alpha^2}) = \int_0^\infty (\alpha^2)^{c_2}(\hat{\alpha^2} - \alpha^2)^2 \pi_2(\alpha^2|\underline{x})d\alpha^2$$

$$R_{(alb,igp)}(\hat{\alpha^2}) = (\hat{\alpha^2})^4 \left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)^{c_2} \frac{\Gamma\left(\frac{n(3+2k)+2\beta-2c_2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} + \left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)^{c_2+2} \frac{\Gamma\left(\frac{n(3+2k)+2\beta-2c_2-4}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} - 2\hat{\alpha^2}\left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)^{c_2+1} \frac{\Gamma\left(\frac{n(3+2k)+2\beta-2c_2-2}{2}\right)}{\Gamma\left(\frac{n(3+2k)+2\beta}{2}\right)} \tag{26}$$

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:

$$\hat{\alpha}_{(alb,igp)} = \sqrt{\frac{2\left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)}{(n(3+2k)+2\beta-2)}} \tag{27}$$

3.3.4. Baye's Estimator under combination of Stein's loss function

The Risk Function Under SLF is given by:

$$R_{(s,igp)}(\hat{\alpha^2}) = \int_0^\infty \left(\frac{\hat{\alpha^2}}{\alpha^2} - log\left(\frac{\hat{\alpha^2}}{\alpha^2}\right) - 1\right)\pi_2(\alpha^2|\underline{x})d\alpha^2$$

$$R_{(s,igp)}(\hat{\alpha^2}) = \hat{\alpha^2}\frac{(n(3+2k)+2\beta)}{2\left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)} - log(\hat{\alpha}) - m - 1 \tag{28}$$

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:

$$\hat{\alpha}_{(ste,igp)} = \sqrt{\frac{2\left(\frac{\sum_{i=1}^{n} x_i^2}{2} + \lambda\right)}{(n(3+2k)+2\beta)}} \tag{29}$$

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 2kth 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	$\hat{\alpha}_{\text{mle}}$	$\hat{\alpha}_{\text{sq}}$	$\hat{\alpha}_{\text{pre}}$	$\hat{\alpha}_{\text{alb}}$	$\hat{\alpha}_{\text{ste}}$
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	0.23772	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	0.04642	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	0.02153	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	0.25445	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	0.05155	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	0.02379	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	$\hat{\alpha}_{\text{mle}}$	$\hat{\alpha}_{\text{sq}}$	$\hat{\alpha}_{\text{pre}}$	$\hat{\alpha}_{\text{alb}}$	$\hat{\alpha}_{\text{ste}}$
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	0.20474	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	0.04349	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	0.02165	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	0.25073	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	0.04957	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	0.02219	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	0.20498	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	0.04372	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	0.02024	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	0.23589	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	0.05069	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	0.02525	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	0.21821	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	0.04053	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	0.02360	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	0.25536	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	0.05101	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	0.02559	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, with square error and precautionary loss function, outperforms MLE estimation, while mle estimation outperforms Bayesian estimation

with Albayyati's and Stein's loss functions.

- In terms of MSE, the bayesian estimation under precautionary loss function and 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 datasets 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 baysian estimation.

Table 7: *Average estimate,Mean Squared Error, AIC, BIC for posterior distribution under different priors for dataset I.*

criterion	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.*

criterion	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.*

criterion	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 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 with the square error loss function and precautionary loss function 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 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.

RT&A, No 2 (78)
Volume 19, June, 2024

STUDY ON ACCEPTANCE SAMPLING PLAN BASED ON PERCENTILES FOR EXPONENTIATED GENERALIZED INVERSE RAYLEIGH DISTRIBUTION

S Jayalakshmi[1], Aleesha A[2]

•

1 Assistant Professor, Department of Statistics, Bharathiar University Coimbatore - 641 046
2 Research Scholar, Department of Statistics, Bharathiar University Coimbatore - 641 046
E-mail: [1]statjayalakshmi16@gmail.com,[2] aleeshaa992@gmail.com

Abstract

An acceptance sampling plan is a sampling procedure with a set of rules for making decisions about a lot of products. The decision is based on the number of defectives in a sample. The sampling inspection plans which are developed for taking decision about a lot based on lifetime of the product are called reliability sampling plans. In this paper, we have developed Acceptance sampling plan (ASP) based on truncated life tests when the lifetime of a product follows the exponentiated generalized inverse Rayleigh distribution (EGIR). The minimum sample sizes needed to ensure the specified life percentile are obtained for a fixed value of the consumer™s confidence level. The operating characteristic values according to the different quality levels are obtained and the minimum ratios of the mean life to the specified life are calculated. The important tables based on the suggested acceptance sampling plan are calculated and illustrated.

Keywords: Acceptance Sampling Plan, Truncated Life Tests, Percentiles, Exponentiated Generalized Inverse Rayleigh Distribution, Operating Characteristic function, producer's Risk

1. Introduction

Two important tools for ensuring quality are statistical quality control and acceptance sampling (AS). Acceptance sampling is concerned with inspection and decision-making regarding lots of products and constitutes one of the oldest techniques in quality control. If the quality characteristic is about the lifetime of the product, the acceptance sampling problem becomes a life test. To determine the sample size from a lot under consideration is the main issue in most acceptance sampling plans for a truncated life test.

If the life test indicates that the true mean lifetime of products exceeds the specified one, then the lot is accepted, otherwise it is rejected. For the purpose of reducing the test time and cost, a truncated life test may be conducted to determine the smallest sample size to ensure certain mean lifetime or percentile lifetime of products when the life test is terminated at a pre-determined time, and the number of failures observed does not exceed a given acceptance number c. The decision is to accept the lot if a pre-determined mean life time or percentile life can be reached with a pre-determined high probability which provides protection to the consumer. Therefore, the life test is ended at the time the failure is observed or at the pre-assigned time, whichever is earlier. For such a truncated life test and the associated decision rule, we are interested in determining the smallest sample size to arrive at a decision.

Epstein [3] was the first who considered truncated life tests in the exponential distribution. Truncated life tests are considered by many authors for various distributions. Sobel and Tischendrof [15], Gupta and Groll [5] using Gamma distribution, Kantam and Rosaiah [6] based on Half

logistic distribution, Ayman et al. [1] using Rayleigh model, Tsai and Wu [17] based on Generalized Rayleigh distribution, Balakrishnan et al. [2] discussed generalized Birnbaum"Saunders distribution. Kantam et al. [7] considered truncated life tests for log-logistic distribution, Rao [13] considered acceptance sampling plans for Marshall"Olkin extended Lomax distribution.

Percentiles provide more information about a life distribution than the mean life does. When the life distribution is symmetric, the 50^{th} percentile or the median is equivalent to the mean life. Hence, developing acceptance sampling plans based on percentiles of a life distribution can be treated as a generalization of developing acceptance sampling plans based on the mean life of it. Lio et al. [8] developed acceptance sampling plans for percentiles using Birnbaum-Saunders distribution. Rao and Kantam [12] considered acceptance sampling plans for truncated life tests based on the log-logistic distribution for percentiles. Rao et al. [11] developed acceptance sampling plans for percentiles based on the inverse Rayleigh distribution. Srinivasa Rao and Kantam [16] studied acceptance sampling plans for percentiles of half logistic distribution. Rao and Naidu [14] considered acceptance sampling plans for Percentiles based on the Exponentiated Half Logistic distribution. Pradeepa Veerakumari and Ponneeswari [10] designed acceptance double sampling plan for life test based on percentiles of Exponentiated Rayleigh distribution. Neena Krishna and Jayalakshmi [9] studied special type double sampling plan for life tests based on percentiles using Exponentiated Frechet Distribution. Percentiles are taken into account because lesser percentile provides more information than mean life regarding the life distribution.

2. Exponentiated Generalized Inverse Rayleigh Distribution

The Exponentiated Generalized Inverse Rayleigh distribution was developed by Fatima et al. in 2018. The CDF of the Exponentiated Generalized Inverse Rayleigh distribution is given by

$$F(t;\alpha,\sigma,\gamma) = [1-[1-e^{(}\frac{-\sigma^2}{t^2})]^{\alpha}]^{\gamma} \tag{1}$$

The PDF of the distribution is given as

$$f(t;\alpha,\sigma,\gamma) = \frac{2\alpha\gamma\sigma^2}{t^2}e^{\frac{-\sigma^2}{t^2}}[1-e^{(}\frac{-\sigma^2}{t^2})]^{\alpha-1}][1-[1-e^{(}\frac{-\sigma^2}{t^2}/)]^{\alpha}]^{\gamma-1}]; t>0 \tag{2}$$

For given 0 < q < 1 the $100q^{t`}$ actual percentile of the Exponentiated Generalized Inverse Rayleigh distribution can be given by

$$t_q = \frac{1}{\sigma}[-ln(1-(1-q^{\frac{1}{\gamma}})^{1/\alpha})]^{-1/2} \tag{3}$$

The t_q increase as q increases Let

$$\eta = [-ln(1-(1-q^{\frac{1}{\gamma}})^{1/\alpha})]^{-1/2} \tag{4}$$

Then from (3), $\sigma = \eta/t_q$
By letting $\delta = t/t_q$, F(t) becomes

$$F(t) = [1-[1-e^{-}(\delta\eta)^{-}2]^{\alpha}]^{\gamma} \tag{5}$$

Equation (5) gives the modified cdf and by partially differentiating the equation (5) w.r.t a we will get the modified pdf for percentiles of Exponentiated Generalized Inverse Rayleigh distribution where t_q is the 10^{th} percentile of the given distribution.

3. Proposed Acceptance Sampling Plan

Main goal of this study is to obtain minimum sample size required to ensure a percentile life when the life test is terminated at a pre-determined time t_q^0 and when the observed number of failures

does not exceed a given acceptance number. The operating procedure of the sampling plan is to accept a lot only if the specified percentile of lifetime is fixed with pre-specified probability λ, which is an indicator of consumer confidence. The life test experiment gets terminated at the time $(c+1)^{th}$ failure is observed or at quantile time t_q whichever is earlier.

A sampling plan in which a decision about the acceptance or rejection of a lot is based on a single sample that has been inspected is known as a single sampling plan. A single sampling plan requires the specification of two quantities which are known as parameters of the single sampling plan. These parameters are n " size of the sample and c " acceptance number for the sample. The Reliability Single Sampling Plan can be represented as $(n, c, t/(t_q^0))$. Here n and c are the sample size and acceptance number for the sampling plan. Assume that a life test is conducted and will be terminated at time t_q^0.

3.1. Operating Procedure

The acceptance sampling plan based on truncated life tests consists of the following:

1. Draw a random sample of size n from the lot received from the supplier. The maximum test duration time is t.
2. We inspect each and every unit of the sample and classify it as defective or non-defective. At the end of the inspection, we count the number of defective units found in the sample. Suppose the number of defective units found in the sample is d.
3. We compare the number of defective units (d) found in the sample with the stated acceptance number (c).
4. We take the decision of acceptance or rejection of the lot on the basis of the sample as follows: If the number of defective units (d) in the sample is less than or equal to the stated acceptance number (c), i.e., if $d \leq c$ defectives out of n occur at the end of the test period t_q^0, we accept the lot and if $d > c$, we reject the lot.

3.2. Minimum sample size

Given P^* and assuming that the lot size is large enough to be considered infinite, then the probability of accepting a lot can be evaluated by the binomial cdf up to c and the smallest sample size n required to assert that $t_q > t_q^0$ must satisfy

$$\sum_{i=0}^{c} p^i (1-p)^{(n-i)} \leq (1-P^*) \tag{6}$$

where p=F(t,δ_q), is the probability of a failure observed during the time t given a specified $100q^{th}$ percentile of lifetime t_q^0 and depends only on δ=$t/(t_q^0)$ since t_q^0 increases as q increases. Accordingly, we have

$$F(t,\delta) < F(t,\delta_0) \iff \delta \leq \delta_0$$

Or, equivalently

$$F(t;t_q) < F(t;t_q^0) \iff t_q \geq t_q^0$$

The smallest sample size n satisfying eq. (6) can be obtained for any given sampling plan $(n, c, t/t_q^0)$ is given in Table 1.

3.3. Operating Characteristic (OC) Function

The OC function L(p) of the acceptance sampling plan (n,c,t/t_q^0) is the probability of accepting a lot. It is given as

$$L(p) = \sum_{i=0}^{c} p^i (1-p)^{(n-i)} \tag{7}$$

where $p = F(t, \delta_q)$. It should be noticed that $F(t, \delta_q)$ can be represented as a function of $\delta_q = t/t_q$. Therefore, we have

$$p = F(t, \delta) = F(\frac{t}{t_q}\frac{1}{d_q})$$

where d_q=t_q/t_q^0
Using eq. (7) the OC values can be obtained for any sampling plan $(n, c, t/t_q^0)$. The OC values for the proposed sampling plan is presented in Table 2.

3.4. Producer's Risk

The producer's risk is defined as the probability of rejecting the lot when $t_q > t_q^0$. For a given value of the producer's risk, say λ , we are interested in knowing the value of d_q to ensure the producer's risk is less than or equal to λ if a sampling plan $(n, c, t/t_q^0)$ is developed at a specified confidence level P^*. Thus, one needs to find the smallest value d_q according to euation (7).

$$L(p) \geq 1 - \lambda$$

Based on the sampling plans (n,c,$t/{t_q}^0$) given in Table 1 the minimum ratios of $d_{0.10}$ at the producer's risk of $\lambda = 0.05$ are presented in Table 3.

4. Illustration

Assume that the life distribution is Exponentiated Generalized Inverse Rayleigh distribution, and the experimenter is interested in showing that the true unknown 10^{th} percentile life $t_{0.10}$ is at least 1000 hrs. Let $\alpha = 2$,$\gamma = 1$ and $\lambda = 0.05$. It is desire to stop the experiment at time t=1500 hrs. For the acceptance number c=1 from the Table 1 one can obtain the Single Sampling plan $(n, c, t/t_q^0)$= (9,1,1.5). The optimum sample sizes needed for the given requirement is found to be as n=9. The respective OC values for the proposed acceptance sampling plan $(n, c, t/t_q^0)$ with $P^* = 0.95$ for Exponentiated Generalized Inverse Rayleigh distribution from the Table 2 are given in below table. This shows that if the actual 10^{th} percentile is equal to the required 10^{th} percentile

$t_q/(t_q^0)$	0.75	1	1.25	1.5	1.75	2	2.25	2.5
L(p)	0.0002	0.0525	0.3297	0.7748	0.9629	0.9965	0.9998	1

$(t_{0.10}/t_{0.10^0} = 1)$, the producer's risk is approximately 0.9475 $(1 - 0.0525)$. The producer's risk almost equal to zero when the actual 10^{th} percentile is greater than or equal to 2.5 times the specified 10^{th} percentile. Table 3 gives the $d_{0.10}$ values for c=1 and $t/t_{0.10}^0 = 1.5$ to assure that the producer's risk is less than or equal to 0.05.
In this example, the value of $d_{0.10}$ is 1.7136 for c=1,$t/t_{0.10}^0 = 1.5$ and λ =0.05. This means the product can have a 10th percentile life of 1.7136 times the required 10^{th} percentile lifetime. That is under the above Single Sampling Plan the product is accepted with probability of at least 0.95.

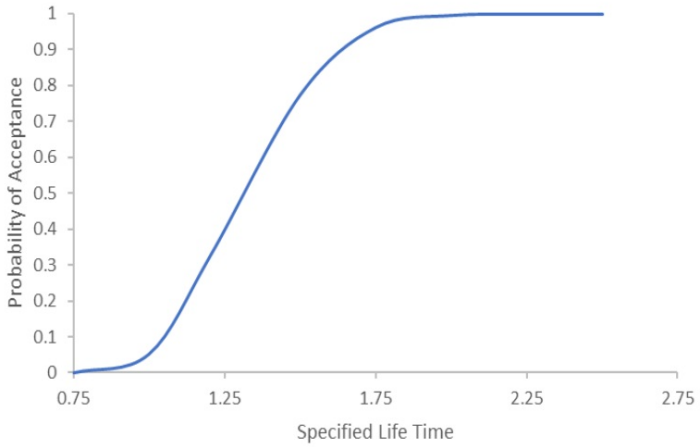


Figure 1: *OC curve for the sampling plan* $(n = 9, c = 1, t/t^0_{0.10} = 1.5)$

5. Construction of the Table

Step 1: Find the value of η for the fixed values of $\alpha = 2, \gamma = 1$ and q=0.10

Step 2: Set the value of $t/{t_q}^0$ =0.7, 0.9, 0.9, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5

Step 3: Find the sample size n by satisfying $L(p) \leq 1 - P^*$ when $P^* = 0.99, 0.95, 0.90$ and 0.75. Here P^* is the probability of rejecting a bad lot and

$$L(p) = \sum_{i=0}^{c} p^i (1-p)^{(n-i)}$$

Step 4: for the n value obtained find the $d_{0.10}$ value such that $L(p) \geq 1 - \lambda$ where $\lambda = 0.05$ and $p = F(t/t^0_q, 1/d_q)$; $d_q = t_q/t^0_q$

Table 1: *Minimum Sample Size values necessary to assure* 10^{th} *percentile for Exponentiated Generalized Inverse Rayleigh distribution*

p^*	c	$t/(t^0_q)$							
		0.7	0.9	1.0	1.5	2.0	2.5	3	3.5
0.75	1	576	53	27	5	5	3	3	2
	2	841	77	39	8	4	4	3	3
	3	1096	101	51	10	7	5	5	3
	4	1346	124	62	13	8	6	6	5
	5	1592	146	73	15	9	7	7	5
0.90	1	834	76	38	8	5	4	3	2
	2	1143	105	52	10	6	5	5	3
	3	1433	131	66	12	7	6	5	5
	4	1714	156	78	16	8	7	6	6
	5	1988	182	91	18	11	8	8	7
0.95	1	1017	93	46	9	5	4	3	3
	2	1349	123	61	12	7	5	4	4
	3	1662	152	76	15	8	7	5	5
	4	1960	180	90	18	10	8	7	7
	5	2253	206	103	20	12	9	8	7
0.99	1	1417	128	63	12	7	4	4	3
	2	1794	162	81	16	8	7	5	5
	3	2143	194	97	18	10	8	7	6
	4	2480	225	113	21	12	10	8	8
	5	2803	255	127	24	13	10	9	8

Table 2: *Operating characteristic values of the sampling plan* $(n, c = 1, t/(t_q^0))$ *for a given* P^* *under Exponentiated Genereralized Inverse Rayleigh Distribution distribution*

p^*	$t/({t_q}^0)$	n	$t_q/({t_q}^0)$							
			0.75	1	1.25	1.5	1.75	2	2.25	2.5
0.75	0.7	576	0	0.2508	0.9963	1	1	1	1	1
	0.9	53	0	0.2451	0.9533	0.9996	1	1	1	1
	1	27	0.0002	0.2326	0.9055	0.9979	1	1	1	1
	1.5	5	0.0222	0.2372	0.6578	0.9185	0.9887	0.9990	0.9999	1
	2	5	0.0008	0.0222	0.1537	0.446	0.7469	0.9185	0.9806	0.9964
	2.5	3	0.0087	0.0555	0.1848	0.3975	0.6303	0.8128	0.9212	0.972
	3	3	0.0024	0.0177	0.0705	0.1848	0.3584	0.5555	0.7306	0.8569
	3.5	2	0.0322	0.0905	0.189	0.3223	0.4738	0.6223	0.7497	0.8469
0.90	0.7	834	0	0.0997	0.9924	1	1	1	1	1
	0.9	76	0	0.0983	0.9123	0.9992	1	1	1	1
	1	38	0	0.0953	0.8345	0.9958	1	1	1	1
	1.5	8	0.0007	0.0546	0.3973	0.8131	0.9705	0.9972	0.9998	1
	2	5	0.0008	0.0222	0.1537	0.446	0.7469	0.9185	0.9806	0.9964
	2.5	4	0.0006	0.0105	0.0658	0.217	0.4541	0.6927	0.8601	0.9477
	3	3	0.0024	0.0177	0.0705	0.1848	0.3584	0.5555	0.7306	0.8569
	3.5	2	0.0322	0.0905	0.189	0.3223	0.4738	0.6223	0.7497	0.8469
0.95	0.7	1017	0	0.0498	0.9889	1	1	1	1	1
	0.9	93	0	0.0481	0.8773	0.9989	1	1	1	1
	1	46	0	0.048	0.7787	0.994	0.9999	1	1	1
	1.5	9	0.0002	0.0525	0.3297	0.7748	0.9629	0.9965	0.9998	1
	2	5	0.0008	0.0222	0.1537	0.446	0.7469	0.9185	0.9806	0.9964
	2.5	4	0.00006	0.0105	0.0658	0.217	0.4541	0.6927	0.8601	0.9477
	3	3	0.0024	0.0177	0.0705	0.1848	0.3584	0.5555	0.7306	0.8569
	3.5	3	0.0008	0.0062	0.0277	0.0827	0.1848	0.3311	0.4996	0.6606
0.99	0.7	1417	0	0.0102	0.0102	1	1	1	1	1
	0.9	128	0	0.0103	0.0107	0.9979	1	1	1	1
	1	63	0	0.0105	0.0109	0.9889	0.9999	1	1	1
	1.5	12	0	0.0065	0.0084	0.659	0.9366	0.9936	0.9996	1
	2	7	0	0.0023	0.0023	0.2496	0.5898	0.8503	0.9618	0.9927
	2.5	4	0.0006	0.0105	0.0105	0.217	0.4541	0.6927	0.8601	0.9477
	3	4	0.0001	0.0019	0.0019	0.0658	0.1844	0.3701	0.5791	0.7581
	3.5	3	0.0008	0.0062	0.0062	0.0827	0.1848	0.3311	0.4996	0.6606

Table 3: *Minimum ratio of true* $d_{0.10}$ *for the acceptability of a lot for the Exponentiated Generalized Inverse Rayleigh distribution and producer's risk of* $\lambda = 0.05$

p^*	$t/(t_q^0)$	n	$t_q/(t_q^0)$							
0.75	0.7	576	1.1554	1.1546	1.1551	1.1551	1.1552	1.1549	1.155	1.1553
	0.9	53	1.246	1.2456	1.246	1.2459	1.2459	1.2458	1.2461	1.2464
	1	27	1.2985	1.2983	1.2985	1.2981	1.2987	1.2987	1.299	1.2988
	1.5	5	1.5676	1.5677	1.5672	1.5665	1.5681	1.5683	1.5696	1.5688
	2	5	2.0896	2.0902	2.0899	2.09	2.0878	2.0886	2.0902	2.0909
	2.5	3	2.3645	2.3641	2.3639	2.3618	2.3618	2.3637	2.3626	2.3652
	3	3	2.8339	2.8365	2.8371	2.8367	2.8322	2.8374	2.8371	2.8336
	3.5	2	2.9629	2.962	2.9633	2.9627	2.9585	2.9636	2.9629	2.9631
0.90	0.7	834	1.1811	1.1809	1.1812	1.181	1.1813	1.1811	1.1814	1.1815
	0.9	76	1.2846	1.2847	1.2842	1.2841	1.2853	1.2854	1.285	1.2854
	1	38	1.3426	1.3426	1.3426	1.3432	1.342	1.3421	1.3422	1.3436
	1.5	8	1.6855	1.6857	1.6856	1.6856	1.6881	1.6864	1.688	1.6857
	2	5	2.0896	2.0902	2.0899	2.09	2.0878	2.0886	2.0902	2.0909
	2.5	4	2.5101	2.5081	2.507	2.5098	2.5098	2.5095	2.5074	2.5107
	3	3	2.8339	2.8365	2.8371	2.8367	2.8322	2.8374	2.8371	2.8336
	3.5	2	2.9629	2.962	2.9633	2.9627	2.9585	2.9636	2.9629	2.9631
0.95	0.7	1017	1.1956	1.1947	1.1951	1.195	1.1947	1.1951	1.1952	1.1945
	0.9	93	1.306	1.3057	1.3059	1.3057	1.3065	1.3064	1.3064	1.3065
	1	46	1.3663	1.3666	1.3663	1.3657	1.367	1.3667	1.3674	1.3657
	1.5	9	1.7136	1.7136	1.7132	1.7129	1.7139	1.7131	1.7135	1.7151
	2	5	2.0896	2.0902	2.0899	2.09	2.0878	2.0886	2.0902	2.0909
	2.5	4	2.5101	2.5081	2.507	2.5098	2.5098	2.5095	2.5074	2.5107
	3	3	2.8339	2.8365	2.8371	2.8367	2.8322	2.8374	2.8371	2.8336
	3.5	3	3.3055	3.3101	3.3104	3.3101	3.3094	3.3103	3.31	3.3079
0.99	0.7	1417	1.2173	1.2171	1.2178	1.2174	1.2176	1.218	1.2178	1.2169
	0.9	128	1.3392	1.3388	1.3391	1.3396	1.3394	1.3389	1.3395	1.3395
	1	63	1.4046	1.4044	1.4051	1.4056	1.4052	1.4054	1.4051	1.4063
	1.5	12	1.7802	1.7775	1.7782	1.7787	1.7778	1.78	1.7796	1.7804
	2	7	2.202	2.2049	2.2048	2.2034	2.2033	2.2021	2.2023	2.206
	2.5	4	2.5101	2.5081	2.507	2.5098	2.5098	2.5095	2.5074	2.5107
	3	4	3.0117	3.0072	3.0121	3.0084	3.0116	3.0102	3.0077	3.0082
	3.5	3	3.3055	3.3101	3.3104	3.3101	3.3094	3.3103	3.31	3.3079

6. CONCLUSION

In this paper, acceptance sampling plan based on percentiles is suggested assuming that the lifetime of the products follows the Exponentiated Generalized Inverse Rayleigh distribution. The suggested Acceptance Sampling Plan is studied for fixed consumer's confidence level, minimum sample sizes necessary to assert the specified percentile life. The required Acceptance Sampling Plan tables are calculated and illustrated.

References

[1] Ayman, B., Abedel-Qader, E. M., and Amjad, A. N. (2005)).Acceptance sampling plans in the Rayleigh model. *Communications for Statistical Applications and Methods* ,12(1), 11-18.

[2] Balakrishnan, N., Leiva, V., 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 .

[3] Epstein, B. (1954) Truncated life tests in the exponential case. *The Annals of Mathematical Statistics* ,25(3), 555-564 .

[4] Fatima, K., Naqash, S., and Ahmad, S. P. (2018). Exponentiated Generalized Inverse Rayleigh Distribution with Applications in Medical Sciences. *Pak. J. Statist* ,34(5), 425-439.

[5] S. S. Gupta and P. A. Groll, (1961) Gamma distribution in acceptance sampling based on life tests. *Journal of the American Statistical Association* ,56(296):942-970.

[6] Kantam, R. R. L., and Rosaiah, K. (1998) Half logistic distribution in acceptance sampling based on life tests*IAPQR TRANSACTIONS*,23, 117-126.

[7] Kantam, R. R. L., Rosaiah, K., Rao, G. S. (2001). Acceptance sampling based on life tests: Log-logistic model. *Journal of Applied Statistics*, 28,121-128.

[8] Lio, Y. L., Tsai, T. R., and Wu, S. J. (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.

[9] PK, N. K., and Jayalakshmi, S. (2021) . Designing of Special Type Double Sampling Plan for Life Tests Based on Percentiles Using Exponentiated Frechet Distribution *Reliability: Theory and Applications* ,16(1 (61)), 117-123.

[10] PradeepaVeerakumari, K., and Ponneeswari, P. (2017). Designing of acceptance double sampling plan for life test based on percentiles of exponentiated Rayleigh distribution. *International Journal of Statistics and Systems* ,12(3), 475-484.

[11] RRL, K. (2012). Acceptance sampling plans for percentiles based on the inverse Rayleigh distribution. *Electronic Journal of Applied Statistical Analysis*,5(2), 164-177.

[12] Rao, G. S., and Kantam, R. R. L. (2010). Acceptance sampling plans from truncated life tests based on the Log-Logistic distribution for percentiles. *Economic Quality Control*,25(2), 153-167.

[13] Rao, G. S. (2013). Acceptance sampling plans for percentiles based on the Marshall-Olkin extended Lomax distribution. *International Journal of Statistics and Economics*, 11(2), 83-96.

[14] S Rao, G., and Naidu, C. R. (2014). Acceptance Sampling Plans for Percentiles Based on the Exponentiated Half Logistic Distribution. *Applications and Applied Mathematics: An International Journal (AAM)*, 9(1), 4.

[15] Sobel, M., and Tischendrof, J. A. (1959, January). Acceptance sampling with new life test objectives. *In Proceedings of fifth national symposium on reliability and quality control*, vol. 108, pp. 118).

[16] Srinivasa Rao, B., and Kantam, R. R. L. (2013). Acceptance sampling plans for percentiles of half logistic distribution.*International Journal of Reliability, Quality and Safety Engineering,* , 20(05), 1350016.

[17] Tsai, T. R., and Wu, S. J. (2006). Acceptance sampling based on truncated life tests for generalized Rayleigh distribution. *Journal of Applied Statistics*,33(6), 595-600.

RT&A, No 2 (78)
Volume 19, June, 2024

A NEW BAYESIAN CONTROL CHART FOR PROCESS MEAN USING EMPIRICAL BAYES ESTIMATES

Souradeep Das[1] and Sudhansu S. Maiti[2]

•

[1]Department of Statistics, Charuchandra College
Kolkata-700029, West Bengal, India
[2]Department of Statistics, Visva-Bharati University
Santiniketan-731 235, West Bengal, India
dassouradeep6@gmail.com, dssm1@rediffmail.com,

Abstract

This article develops a new control chart for the mean using empirical Bayes estimates. We assume that the quality characteristic of the proposed control chart follows a normal distribution with unknown mean and variance. Both the parameters have known prior probability distributions. In practice, the parameters of priors are unknown and are estimated using the empirical Bayes approach. For the performance assessment of the new control chart, the Average Run Length (ARL) procedure is used while the process is in control and out of control. A real-life example is also considered to evaluate the performance of the proposed control chart.

Keywords: Average Run Length, Empirical Bayes, Mean Chart, Posterior, Statistical Process Control.

1. Introduction

Statistical Process Control (SPC) is a popular methodology for monitoring and assessing the quality of a manufacturing process. The main objective of SPC is to minimize the process variability. A control chart is the main technique SPC uses to measure whether a manufacturing process is in control. Dr. Walter Shewhart first proposed the control chart technique in the 1920s. If the quality characteristic under study is quantifiable, we use variable control charts like $\bar{X}$, R, and S charts, etc. For these control charts, it is assumed that the quality characteristic follows a normal distribution. Over the years, researchers have developed control charts for means by considering different aspects. [5] proposed a $\bar{X}$ chart when the quality characteristic follows a skewed distribution. [9] introduced a new $\bar{X}$ chart by considering variable sample size and sampling intervals, which can detect the shift in the process mean in less time than a traditional $\bar{X}$ chart. [8] gave an idea of the Max chart by combining the $\bar{X}$ chart and S chart. [18] proposed a new control chart for mean based on variable and attribute inspections.

The Bayesian approach has recently become very popular among researchers for constructing control charts. Using empirical Bayes, [11] developed a multivariate process control chart. [19] compared the effectiveness of different mean charts under the Bayesian approach. [13] have constructed a new control chart for the coefficient of variation using prior information when the mean is variable, and the variance is the function of the mean. [17] designed a two-sided $\bar{X}$ control chart for mean. [4] developed a new control chart for mean using posterior distribution. [10] measured the performance of a Bayesian Control Chart using empirical Bayes based on

Weibull data. [2] proposed a mean control chart using a uniform prior. [12] used Empirical Bayes methods based on loss functions for a sequential sampling plan. [6] used the Bayesian model for constructing predictive control charts. [1] designed a Bayesian Shewhart-type control chart for the Maxwell distributed process.

The entire article is arranged in the following way. Section 2 discusses the Shewhart $\bar{X}$ chart. In section 3, a discussion is made on the posterior mean control chart. Section 4 briefly describes the empirical Bayes method. In section 5, we explain the construction of the new control chart for mean using empirical Bayes estimates. In section 6, the performance of the proposed control chart is evaluated concerning the Average Run Length values. In section 7, a real-life dataset is taken to analyze the performance of the proposed control chart for mean. In the last section (section 8), concluding remarks are given.

2. X-bar Control Chart

Let $X_1, X_2, \ldots, X_n$ be n observations of a quality characteristic X following a normal distribution with mean, μ and variance, σ^2 of a manufacturing process. Then, according to W. Shewhart, the 3-sigma control limits of $\bar{X}$ chart are

$$UCL = \mu + 3\frac{\sigma}{\sqrt{n}}$$
$$LCL = \mu - 3\frac{\sigma}{\sqrt{n}}$$

3. Posterior Control Chart for Mean

[4] proposed a new posterior $\bar{X}$ control chart for process mean. Suppose $X_1, X_2, \ldots, X_n$ be n observations of a quality characteristic X. It is assumed that X_i's are independently and identically distributed normal variables with mean μ and variance σ^2(known). Here, the process average μ has normal prior with known parameters.Then X_i's $\sim N(\mu, \sigma^2)$ and $\mu \sim N(\theta, \lambda^2)$, where θ and λ are known. So, the posterior mean $\alpha_0 = \bar{x}\zeta_0 + \theta(1-\zeta_0)$ and the posterior variance is $\rho_0 = \frac{n}{\sigma^2\zeta_0}$ where $\zeta_0 = \frac{n\lambda^2}{n\lambda^2+\sigma^2}$. Hence, the three-sigma control limits of the posterior control chart for the mean are

$$UCL = \bar{x}\zeta_0 + \theta(1-\zeta_0) + 3\frac{\sigma}{\sqrt{n}}\sqrt{\zeta_0}$$
$$CL = \bar{x}\zeta_0 + \theta(1-\zeta_0)$$
$$LCL = \bar{x}\zeta_0 + \theta(1-\zeta_0) - 3\frac{\sigma}{\sqrt{n}}\sqrt{\zeta_0}$$

4. Empirical Bayes method

In the Bayesian method, the probability distribution function's unknown parameters are considered the random variables. Suppose $X_1, X_2, \ldots, X_n$ are n observations from f(θ). Here, the parameter θ has some prior information. , θ has the prior distribution $\pi(\theta|\omega)$, where ω is the hyperparameter. The Bayes' theorem states that the posterior distribution of θ can be expressed as proportional multiplication of the likelihood L(θ) and the prior distribution $\pi(\theta|\omega)$. Symbolically, $h(\theta|x) = \frac{L(\theta)\pi(\theta|\omega)}{\int L(\theta)\pi(\theta|\omega)} \propto L(\theta)\pi(\theta|\omega)$

The Bayesian method is different from the frequentist method. In the parametric empirical Bayes method, the prior distribution $\pi(\theta|\omega)$ takes parametric form, where the prior distribution parameters are unknown. [7] estimates the prior parameters using the observed data. These parameters could be estimated using the empirical Bayes procedure (see [14] and [3]). Given the observations, the joint likelihood distributions have been compared with the joint prior distributions. The joint likelihood distributions are just the multiplication of the likelihood distribution of X, and

joint prior distributions are the multiplication of the prior distributions of the parameters. We can estimate the prior parameters by comparing them individually with their corresponding likelihood functions.

5. Control Chart for Mean using Empirical Bayes

Using the empirical Bayes method, we propose a new control chart for the mean. Suppose X is a quality characteristic of a manufacturing process and is assumed to follow a normal distribution with mean, μ and variance, σ^2. The location parameter, μ, has a normal prior with unknown parameters, and σ follows an inverse gamma distribution with unknown parameters. Let X has n observations $X_1, X_2, \ldots, X_n$, such that $X_i \sim N(\mu, \sigma^2)$. here $\mu \sim N(\mu_0, \sigma^2)$ and $\sigma \sim InverseGamma(\alpha, \beta)$.

$$g(x|\mu,\sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2} \quad -\infty < \mu < \infty, \sigma^2 > 0$$
$$g(\mu|\sigma^2,x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\mu-\mu_0)} \quad -\infty < \mu_0 < \infty$$
$$g(\sigma^2|x) = \frac{\beta^\alpha}{\Gamma\alpha}(\sigma^2)^{-\alpha-1} e^{\frac{\beta}{\sigma^2}} \quad \alpha > 0, \beta > 0$$

Hence, the posterior distribution of (μ, σ^2) is given by,

$$g(\mu,\sigma^2|x) = \frac{g(x|\mu,\sigma^2)g(\mu|\sigma^2)}{\int_0^\infty \int_{-\infty}^\infty g(x|\mu,\sigma^2)g(\mu|\sigma^2)d\mu d\sigma^2}$$

So, posterior mean $E(\mu|x) = \frac{\sum x_i + \mu_0}{n+1}$. The empirical Bayes estimate of μ_0 is $\bar{X}$.
So, $E(\hat{\mu}|x) = \frac{(n+1)\bar{x}}{n+1} = \bar{x}$.
Now,

$$\begin{aligned}
E(\sigma^2|x) &= \int_0^\infty \sigma g(\mu,\sigma^2|x)d\sigma^2 \\
&= \int_0^\infty \sigma g(\sigma^2|x)g(\mu|\sigma^2,x)d\sigma^2 \\
&= \int_0^\infty \sigma g(\sigma^2|x)g(\mu|\sigma^2,x)d\sigma^2 \\
&= \frac{\Gamma(\frac{n}{2}+\alpha)}{\Gamma(\frac{n+1}{2}+\alpha)} \frac{w_1^{\frac{n+1}{2}+\alpha}}{w_1^{\frac{n}{2}+\alpha}} \\
&= \frac{\Gamma(\frac{n}{2}+\alpha)}{\Gamma(\frac{n+1}{2}+\alpha)} \sqrt{w_1} \\
&= \frac{2^{n+2\alpha-1}\Gamma(\frac{n}{2}+\alpha)}{\sqrt{\pi}\Gamma(n+2\alpha)} \sqrt{w_1}
\end{aligned}$$

here $w_1 = \sum x_i^2 + 2\beta + \mu_0^2 - \frac{x_0^2}{n+1}$.
The empirical Bayes procedure will be used to estimate the parameters. So the estimated values of the parameters of the likelihood function of σ^2 are $\hat{\alpha} = (n-3)/2$ and $\hat{\beta} = \sum_{i=1}^n (X - \bar{X})^2/2$ (see [15]). Therefore

$$\begin{aligned}
E(\hat{\sigma}|x) &= \frac{\Gamma(\frac{n}{2}+\frac{n-3}{2})}{\Gamma(\frac{n+1}{2}+\frac{n-3}{2})} \sqrt{2\sum_{i=1}^n (x_i-\bar{x})^2} \\
&= \frac{\Gamma(\frac{2n-3}{2})}{\Gamma(\frac{2n-2}{2})} \sqrt{2\sum_{i=1}^n (x_i-\bar{x})^2} \\
&= \frac{\Gamma(\frac{2n-3}{2})}{\Gamma(\frac{2n-2}{2}+\frac{1}{2})} \sqrt{2\sum_{i=1}^n (x_i-\bar{x})^2}
\end{aligned}$$

Hence, the control limits of the proposed control chart for the mean are

$$UCL = \bar{x} + L\frac{\Gamma(\frac{2n-3}{2})}{\Gamma(\frac{2n-2}{2}+\frac{1}{2})}\sqrt{2\sum_{i=1}^{n}(x_i-\bar{x})^2}$$
$$CL = \bar{x}$$
$$LCL = \bar{x} - L\frac{\Gamma(\frac{2n-3}{2})}{\Gamma(\frac{2n-2}{2}+\frac{1}{2})}\sqrt{2\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

Here, L is the control chart coefficient.

6. Evaluation of Performance and Comparisons

This section uses Monte Carlo Simulation to compute the proposed control chart's Average Run Length (ARL) for mean using empirical Bayes. We consider different sample values of n and compare the computed in-control-ARL (ARL_0) and out-of-control ARL (ARL_1) with the existing posterior mean control chart and Shewhart Control Chart. The decision is based on the value of ARL_1. The control chart with a smaller ARL_1 value is more efficient in detecting a shift in the process mean than other control charts. We consider the shift in process mean as $\mu^* = \mu + c\sigma$.

Algorithm for construction of UCL and LCL is as follows.

- **Step 1:** Select a random sample of size n, say, $x_1, x_2, \ldots, x_n$ from $N(\mu, \sigma^2)$ distribution. Here we assume that μ has a normal prior and σ has an Inverse Gamma prior with unknown parameters.
- **Step 2:** Estimate the posterior distribution parameters using the empirical Bayes procedure.
- **Step 3:** For given values of n and fixed in-control Average Run Length(ARL), say r_0, find the control chart coefficient L.
- **Step 4:** Find UCL and LCL for each i, $i = 1, 2, \ldots, n$. The process is in control if all the values of x_i fall within the UCL and LCL of the proposed mean chart.
- **Step 5:** Next, we find the ARL_0 value for a particular choice of the process mean μ.
- **Step 6:** We shift the process mean μ to a certain amount, say c and compute ARL_1 by repeating steps 1 to steps 5.

Here, we fixed the ARL_0 at 370. The ARL values of the proposed control chart are given in Table 1 - Table 3.

We can see the proposed control chart for mean using empirical Bayes estimators has the least ARL among all the control charts under consideration. The ARL_1 of the proposed chart decreases quickly for a small shift in the process mean. As we increase the sample size, the ARL_1 values of the control chart decrease. Therefore, we can conclude that the new control chart for using empirical Bayes estimators is more efficient than the posterior control chart and Shewhart $\bar{X}$ control chart.

7. Illustrative Example

In recent trends, SPC researchers use both simulated and real-life data to evaluate the performance of a control chart. In this study, we have considered a real dataset from [16] to evaluate the performance of the new control chart for mean using empirical Bayes estimates. The data set is given in the appendix section. Here, we have filled out height data in ten subgroups of size 10. It is assumed that the control chart statistic, fill height, follows normal distribution where the parameters μ and σ have unknown prior distribution. Using the empirical Bayes procedure, the

Table 1: *Comparison of average run lengths of Empirical Bayes control chart for Mean with Posterior Mean Control Chart and Shewhart $\bar{X}$ Control Chart for n = 10 and ARL = 370*

Shift	Empirical Bayes Mean Chart	Posterior Mean Chart	Shewhart $\bar{X}$ Chart
	L = 3	L = 3	L = 3
0.0	370.398	370.398	370.398
0.05	281.397	312.467	328.011
0.1	183.248	221.991	249.167
0.15	128.813	144.631	181.701
0.2	86.003	94.297	123.981
0.25	58.238	68.997	87.457
0.3	34.184	39.832	59.301
0.4	9.327	18.115	28.034
0.7	2.265	3.168	4.387
0.9	1.183	1.719	2.814

Table 2: *Comparison of average run lengths of Empirical Bayes control chart for Mean with Posterior Mean Control Chart and Shewhart $\bar{X}$ Control Chart for n = 20 and ARL = 370*

Shift	Empirical Bayes Mean Chart	Posterior Mean Chart	Shewhart $\bar{X}$ Chart
	L = 3	L = 3	L = 3
0.0	370.398	370.398	370.398
0.05	236.234	289.754	300.373
0.1	131.197	168.103	185.559
0.15	67.469	91.476	106.358
0.2	34.482	50.893	61.539
0.25	21.107	29.555	36.807
0.3	12.893	17.985	22.885
0.4	5.221	7.663	9.959
0.7	1.934	2.988	3.824
0.9	1.021	1.504	2.357

Table 3: *Comparison of average run lengths of Empirical Bayes control chart for Mean with Posterior Mean Control Chart and Shewhart $\bar{X}$ Control Chart for n = 30 and ARL = 370*

Shift	Empirical Bayes Mean Chart	Posterior Mean Chart	Shewhart $\bar{X}$ Chart
	L = 3	L = 3	L = 3
0.0	370.398	370.398	370.398
0.05	183.609	262.736	271.659
0.1	96.064	130.865	142.164
0.15	34.627	63.37	71.433
0.2	24.922	32.409	37.614
0.25	9.088	17.731	20.860
0.3	3.167	10.384	12.343
0.4	2.081	4.338	5.163
0.7	1.355	2.805	3.532
0.9	1.008	1.244	1.841

prior parameters are estimated. The UCL and LCL of the new control chart for mean based on empirical Bayes for the data set are 0.7288446 and -0.6888446, respectively. From figure 2, we can see that the proposed control chart based on empirical Bayes can detect an out-of-control observation more precisely than the posterior control chart for mean and Shewhart $\bar{X}$ chart. In figure 4, the Average Run Lengths of the proposed chart and other control charts are drawn.

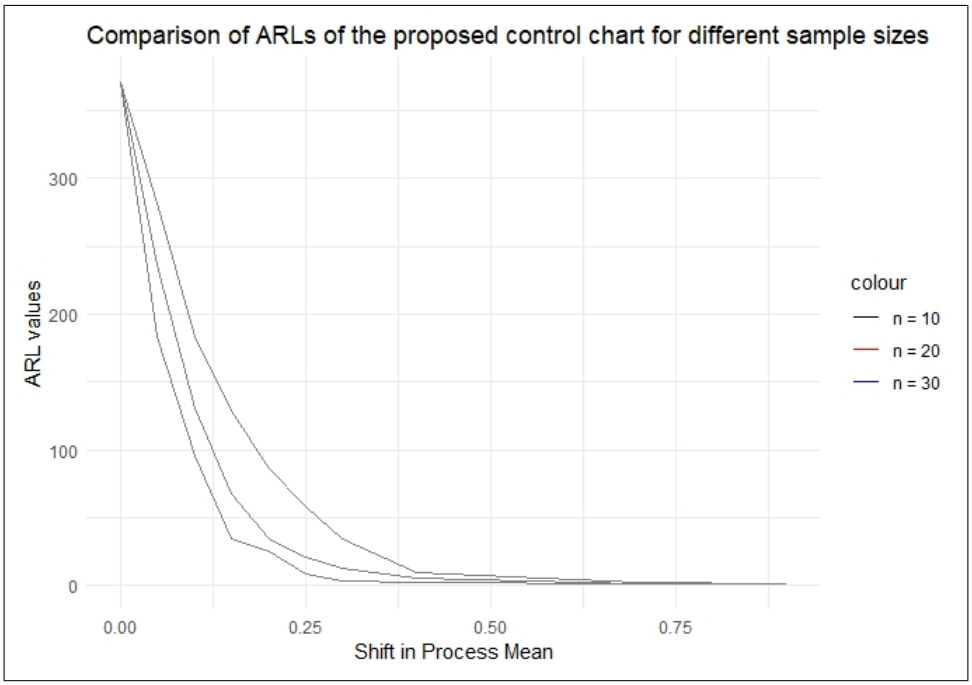


Figure 1: *Comparison of ARLs of Empirical Bayes Mean Control Chart for different sample values*

Table 4: *Comparison of average run lengths of Empirical Bayes control chart for Mean with Posterior Mean Control Chart and Shewhart $\bar{X}$ Control Chart*

Shift	Empirical Bayes Mean Chart	Posterior Mean Chart	Shewhart $\bar{X}$ Chart
	L = 3	L = 3	L = 3
0.0	370.398	370.398	370.398
0.05	297.351	322.097	334.916
0.1	180.398	227.721	257.719
0.15	101.842	147.533	182.071
0.2	58.247	94.044	124.894
0.25	34.533	60.687	85.584
0.3	21.331	40.032	59.301
0.4	9.217	18.786	29.912
0.7	1.867	3.538	5.911
0.9	1.198	1.829	2.822

From table 4 and figure 2, we can conclude that the proposed mean chart using the empirical Bayes estimator has smaller ARL values than the posterior mean control chart as well as Shewhart $\bar{X}$ control chart when there occurs a shift. In figure 1, we can see that the width of control limits for the proposed control chart is narrower than the posterior mean control chart and Shewhart $\bar{X}$ chart. This implies that the new mean chart based on empirical Bayes estimate can detect an 'out of control state' of a process mean earlier.

8. Conclusion

This article proposes a new control chart for mean using the empirical Bayes approach. For the mean, we compare the performance of the new control chart with that of the existing control charts. We used ARL values to measure the performance of the control charts for the mean. It is observed that the proposed control chart can detect the smaller shift in the process mean quickly than the posterior mean control chart and Shewhart $\bar{X}$ control chart. It was also noted that the

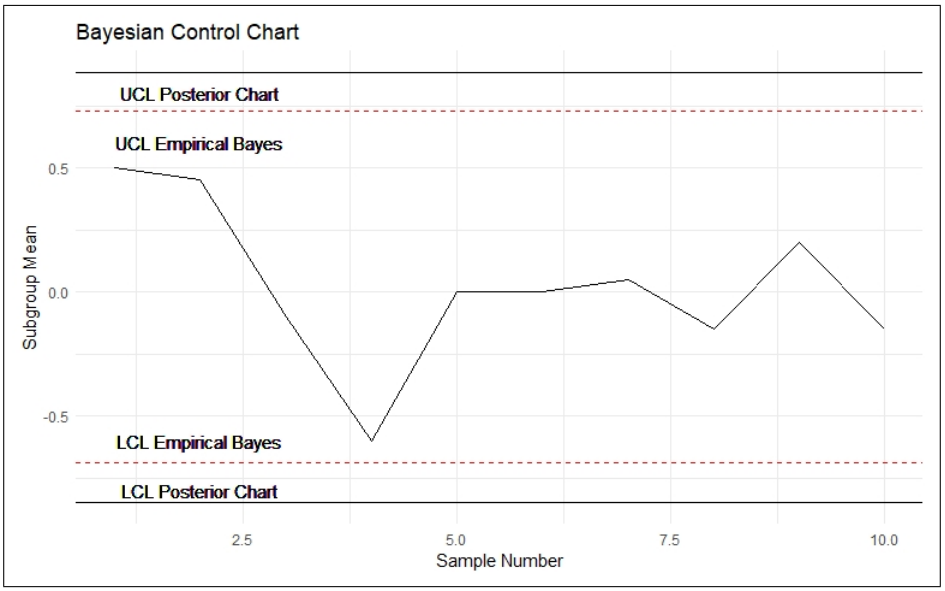


Figure 2: *Empirical Bayes control chart with that of Posterior Control Chart*

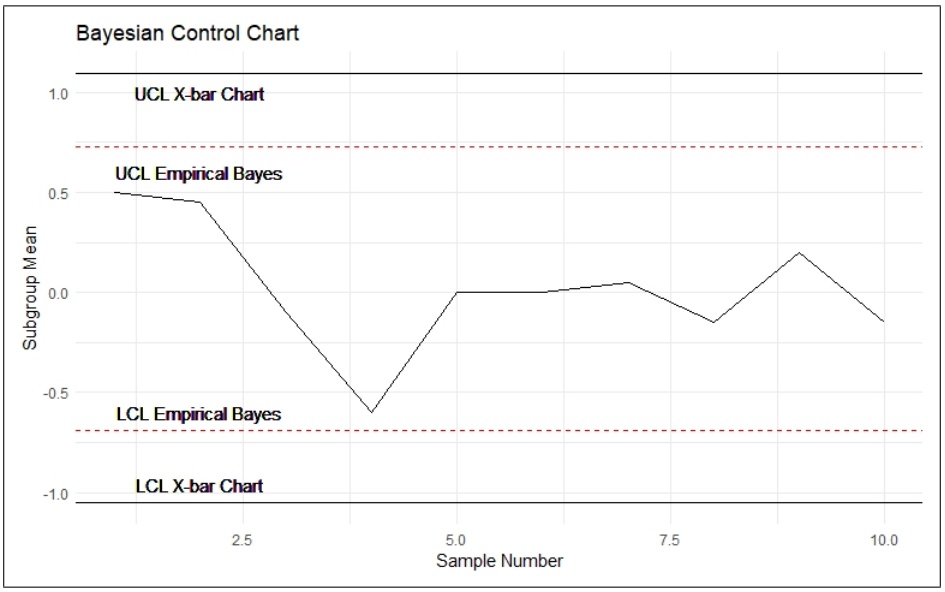


Figure 3: *Empirical Bayes Control Chart and Shewhart $\bar{X}$ control chart*

proposed control chart performed better for larger sample sizes.

Declarations

Disclosure of Conflicts of Interest/Competing Interests: The authors declare no conflict of interest.
Authors contributions: Each author has an equal contribution. All authors jointly write, review, and edit the manuscript.
Funding: The authors received no specific funding for this study.
Data Availability Statements: All cited data analyzed in the article are included in References. Data sets are also provided in the article.
Ethical Approval: This article does not contain any studies with human participants performed by authors.
Code availability: Codes are available on request.

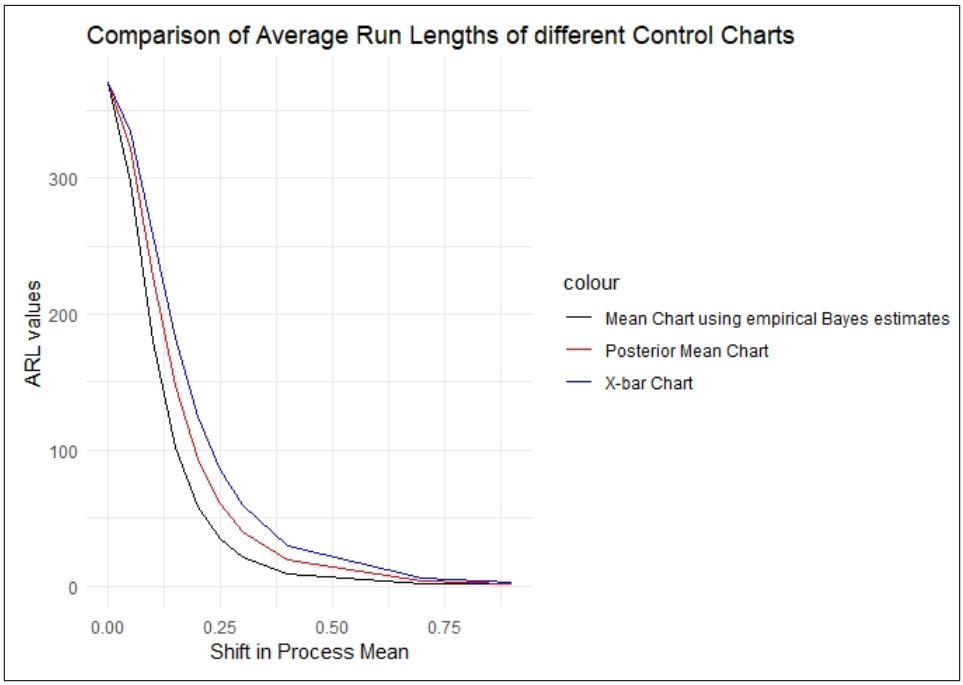


Figure 4: *Comparison of ARLs of Empirical Bayes Mean Control Chart with other Control Charts for the example dataset*

References

[1] Alshahrani, F., Almanjahie, I.M., Khan, M., Anwar, S.M., Rasheed, Z., Cheema, A.N. (2023): On Designing of Bayesian Shewhart-Type Control Charts for Maxwell Distributed Processes with Application of Boring Machine. *Mathematics*, 11(5),1126.

[2] Aunali, A. S., Venkatesan, D. (2019): Bayesian Control Charts Using Uniform Prior. *Journal of Information and Computational Science*, 9(11), 295-301.

[3] Awad, A. M. and Gharraf, M. K. (1986): Estimation of $P(Y < X)$ in the Burr case: a comparative study. *Communications in Statistics-Simulation and Computation*, 15(2), 389-403.

[4] Bhat, S. V., Gokhale, K. D. (2014): Posterior control chart for process average under conjugate prior distribution. *Economic Quality Control*, 29(1),19“27.

[5] Bai, D. S., Choi, T. S. (1995): $\bar{X}$ and R Control Charts for Skewed Populations. *Journal of Quality Technology*, 22(2),120-131.

[6] Bourazas, K., Kiagias, D., Tsiamyrtzis, P. (2022): Predictive Control Charts (PCC): A Bayesian approach in online monitoring of short runs. *Journal of Quality Technology*, 54(4), 367-391.

[7] Carlin, B. P., Louis T. A. (2000): Bayes and Empirical Bayes Methods for Data Analysis. *CRC Press*.

[8] Chen, G., Cheng, S. W. (1998): Max chart: combining X-bar chart and S chart. *Statistica Sinica*, 8(1),263-272.

[9] Costa, A. F. B. (1997): $\bar{X}$ chart with variable sample size and sampling intervals. *Journal of Quality Technology*. 29(2),197-204.

[10] Erto, P., Pallotta, G., Palumbo, B., Mastrangelo, C. M. (2018): The performance of semi-empirical Bayesian control charts for monitoring Weibull data. *Quality Technology & Quantitative Management*, 15(1), 69-86.

[11] Feltz, C. J., Shiau, J. J. H. (2001): Statistical process monitoring using an empirical Bayes multivariate process control chart. *Quality and Reliability Engineering International*, 17(2), 119-124.

[12] Jampachaisri, K., Tinochai, K., Sukparungsee, S., Areepong, Y. (2020): Empirical Bayes based on squared error loss and precautionary loss functions in sequential sampling plan. *IEEE Access*, 8, 51460-51469.
[13] Kang, C. W., Lee, M. S., Seong, Y. J., Hawkins, D. M. (2007): A control chart for the coefficient of variation. *Journal of Quality Technology*, 39(2),151-158.
[14] Lindley, D. V. (1969): Introduction to Probability and Statistics from a Bayesian Viewpoint. *Cambridge University Press, Cambridge, UK*.
[15] Maiti, S. S., Saha, M. (2012): Bayesian estimation of generalized process capability indices. *Journal of Probability and Statistics*.
[16] Montgomery, D. C. (2018): Introduction to statistical quality control. *John Wiley & Sons*.
[17] Nenes, G., Tagaras, G. (2007): The economically designed two-sided Bayesian $\bar{X}$ control chart. *European Journal of Operational Research*, 183(1), 263-277.
[18] Quinino, R. C., Cruz, F. R., Quinino, V. B. (2021). Control chart for process mean monitoring combining variable and attribute inspections. *Computers & Industrial Engineering*, 152, 106996.
[19] Tagaras, G., Nikolaidis, Y. (2002): Comparing the effectiveness of various Bayesian $\bar{X}$ control charts. *Operations Research*, 50(5), 878-888.

COMPARISON OF SINGLE SERVER RETRIAL QUEUING PERFORMANCE USING FUZZY QUEUING MODEL AND INTUITIONISTIC FUZZY QUEUING MODEL WITH INFINITE CAPACITY

S. Aarthi
•
Department of Mathematics & Statistics, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur-603203, Tamil Nadu, India.
aarthis4@srmist.edu.in

Abstract

A single server retrial fuzzy queuing model is presented in this study. An unreliable $FM/FM/1$ fuzzy retrial queue with a virtually unlimited retrial orbit and a standard queue is investigated. After an unspecified amount of time has elapsed and the server is workable and inactive, orbit patrons don't rejoin the regular queue, but instead, enter the server momentarily. Customers who arrive and discover the server is engaged or has struggled are placed in the regular queue, whereas customers who are disrupted are always placed in orbit. The model's prosecution proportions are also calculated in a hazy environment. The main goal of this investigation is to compare the efficacy of a single server retrial queuing system based on fuzzy queuing theory and intuitionistic fuzzy queuing theory. The arrival, service, failure, orbit, and repair rates are documented using triangular and triangular intuitionistic fuzzy numbers. The evaluation metrics for the fuzzy queuing theory model are proffered as a range of possible values, whereas the intuitionistic fuzzy queuing theory model encompasses a wide range of values. An approach is conducted to discover quality measures using a design protocol in which the fuzzy values are left alone and not repurposed to crisp values, allowing us to draw research findings in an ambiguous future. Two numerical problems are solved to emphasize the method's protracted survivability.

Keywords: queuing theory, retrial queues, fuzzy numbers, breakdown, repair

I. Introduction

We scrutinize an $FM/FM/1$ fuzzy retrial queue with an undependable server whose retrial orbit and standard queue both have inexhaustible capacity space in this manuscript. People can only access the retrial orbit if their service is thwarted due to an outage. Retrial patrons already don't resume the consistent backlog; instead, they try accessing the server explicitly at random intervals, independent of people arriving and perhaps other retrial clients. These hindered customers, on the other hand, can only regain entry access to the servers when it is fully functional and sedentary, and they just rehash the service until it is efficaciously processed. In the history of queuing systems, a variety of methods for placing fuzzy numbers has been developed. In this paper, we propose a method for solving the single server retrial queuing model in both fuzzy and intuitionistic fuzzy environments while sustaining their essence. Authors and researchers in the literature on fuzzy retrial queuing models used defuzzification methods, whereas here we keep the fuzziness until the end. Our paper is one-of-a-kind in this regard. This method applies to previous methods in that it is straightforward, configurable, and relatable. We can focus on the interplay between the retrial orbit and the standard queue in particular, which is excluded from the overwhelming bulk of retrial

concepts which does not include a standard queue with eternal or nontrivial capacity. The random variables are articulated as the combination of two independent random variables, one being a broad sweeping binomial random variable and the other of which conforms to the same estimation for a real-time constructive criticism prototype, that is, one with an unbounded retrial rate. Furthermore, an intriguing stabilization result will be demonstrated, namely that the standard queue can stay constant even though a whole system's (and, specifically, the orbit's) stability condition is contravened. There seem to be different sorts of breakdowns assumed here: engaged breakdowns that happen during a service delivery process and indolent breakdowns that happen when the server is not failing but is sluggish. The time between customer entrants, provider closure, shutdowns, retrials, and refurbishments are assumed to be a random variable with an exponential distribution.

The retrial queuing model with breakages and renovations is a queuing system with a broad array of applications in manufacturing technologies where a server can break down at any time, be repaired, and restarted. Retrial queues and queuing systems with malfunctions have both been intensively investigated in empirical studies. Authors in the antiquity of retrial queue literature considered an innumerable orbit retrial queue and a normal queue, but not a server that is prone to failure. Customers who arrive to seek the server preoccupied can enlist the retrial orbit or the regular queue, according to their model. Customers who arrive to seek the server down (hectic or ceased) are appended to the orbit in retrial concepts without any waiting area and server breakdowns. Some models oblige these customers to join the orbit, whereas others offer them the right to terminate the system. Except for two alternatives, some models also require or enable in-service customers who have been disrupted by a server's inability to enroll in the retrial orbit. Our prescribed concept is unique where orbit consumers need not re-join the standard queue but instead try to enter the server instantaneously after an unidentified amount of time has passed and the server is functional and idle. Customers who arrive and discover the server is overwhelmed or has struggled are placed in the regular queue because customers who are curtailed are always placed in orbit.

Starting failures, vacations, active shutdowns, and both active and idle breakdowns are all taken into account in the retrial fuzzy queue literature. Ramesh et al [1] with the incentre-based sorting method, convert the input rates to crisp numerals. By using retrial queuing models, this article proposes a ruse for perceiving bounteous exploration mission indicators of crisp values for a single server beauty salon using glycolic acid. In a fuzzy environment, the solitary server dual orbit retrial queuing model is probed with customer disparagement. Further, α −cut methodology is used to generate a series of parameterized nonlinear programming for evaluation metrics relying on Zadeh's extension principle, which is then remedied utilizing calculus concepts by S S Sanga et al [2]. Kannadasan et al [3] used hexagonal fuzzy numbers to the input parameters and solved retrial queues with a working vacation. Jain et al [4] looked at the performance of a machine repair system that operates in a fuzzy environment with an admission control F −policy. The steady-state governing equations are constructed using the auxiliary variable correlating to retrial times, and then overt derivations for the queue volume probability distributions are deduced by using the Laplace transform and iterative method, as well as defuzzification. Upadhyaya et al [5] analyzed the $M_x/G/1$ retrial queue with frustrated customers transformed the vacation policy and used Bernoulli feedback. The system size distribution and other key data points are determined using an auxiliary variable approach and the probability-generating function methodology. S S Sanga et al [6] dealt with the admittance control policy for a solo server countable space queuing system with disappointed consumers and dispersed retrial times. By introducing ancillary variables correlating to residual retrial times and interpreting Chapman–Kolmogorov formulations, the steady flow queue size characterization of the system size is reviewed. S S Sanga et al [7] in a dual orbit retrial queuing system with different types of customers, ordinary and premium class customers, the

behavior of balking customers was probed. The fuzzified indices are ascertained using a parameterized non-linear optimization framework that relies on the extension principle of Zadeh and the α −cut method is used to determine the fuzzified indices. nonlinear programming approach based on Zadeh's extension principle and α −cut method. Moreover, the performance targets are defuzzified using the ranking index method. Ebenesar Anna Bagyam et al [8] considered the state-dependent batch arrival two-phase retrial queue and used Zadeh's extension principle, the model is further examined in a fuzzy environment. Kalpana et al [9] proposed a numerical method to deduce the membership function of a fuzzy retrial queue with a solo server line model $FM_1, FM_2/FM_1, FM_2/1$ with priority and inequitable service rate. In this paper, fuzzy queues are transmogrified into classical queues using the α −cut methodology and Zadeh's principle. Sherman et al [12] presented several stochastic decomposability results as well as stability conditions where the customers in the retrial queue do not re-join the regular queue; meanwhile, they try to enter the server until it is found to be functional and idle. Mukeba [13] used a method named flexible α −cuts method to quantify the quality metrics of a solo server fuzzy retrial queue with malfunctions and repair work. Kulkarni et al [14] studied the limiting behavior of a solitary server retrial queue where the server is subject to malfunctions and repair work. He used Markov regenerative processes to deduce the convergence criteria and study the system's limiting behavior. Jau-Chuan Ke et al [16] used the α −cut method to turn a fuzzy into a group of traditional retrial queues. A sequence of parameterized non-linear programs is devised to explain the clan of crisp retrial queues using the membership functions of the system components. Artalejo et al [17] are concerned about the balking retrial queue. Using classical mean diffusion characteristics, the ergodicity condition is first researched. A recursive approach based on the theory of regenerative processes is used to ascertain the restricting distribution of the number of clients in the system. Kannadasan et al [18] examined finite capacity retrial queues using hexagonal fuzzy numbers. Rani Shobha et al [19] used ANFIS strategy and set of differential linear equations in the markovian retrial queue with double orbits.

Most previous research on fuzzy queuing models has concentrated on two or three fuzzy variables, with researchers employing ranking techniques or defuzzification processes to repurpose fuzzy variables into crisp. In this paper, we propose a way to collect information about system behavior for retrial queues by using five fuzzy variables. Throughout the paper, we keep the fuzzy values and don't change them to crisp values for the different membership functions (TFN and TIFN).

The remaining part of the article is configured as regards. Prelims and definitions are covered in section 2. The mathematical formalism, as well as the circumstances for stability, are described in Section 3. The layout method for dealing with the current model is detailed in Section 4. Standard queuing relevant factors are discussed in Section 5, and Mathematical descriptions and visual observations are provided in Section 6. This work wraps up with Section 7.

II. Preliminaries

The motive of this division is to give some basic definitions, annotations, and outcomes that are used in our further calculations.

Definition 2.1. [10] A fuzzy set $\tilde{A}$ is defined on R, the set of real numbers is called a **fuzzy number** if its membership function $\mu_{\tilde{A}}: R \to [0,1]$ has the following conditions:

(a) $\tilde{A}$ is convex, which means that there exists $x_1, x_2 \in R$ and $\lambda \in [0,1]$, such that$\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq min\{\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)\}$
(b) $\tilde{A}$ is normal, which means that there exists an $x \in R$ such that $\mu_{\tilde{A}}(x) = \tilde{1}$
(c) $\tilde{A}$ is piecewise continuous.

Definition 2.2. [10] A fuzzy number $\tilde{A}$ is defined on R, the set of real numbers is said to be a **triangular fuzzy number (TFN)** if its membership function $\mu_{\tilde{A}}: R \to [0,1]$ which satisfies the following conditions:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-\tilde{a}_1}{\tilde{a}_2-\tilde{a}_1} \text{ for } \tilde{a}_1 \le x \le \tilde{a}_2 \\ 1 \text{ for } x = \tilde{a}_2 \\ \frac{\tilde{a}_3-x}{\tilde{a}_3-\tilde{a}_2} \text{ for } \tilde{a}_2 \le x \le \tilde{a}_3 \\ 0 \text{ otherwise} \end{cases}$$

The triangular fuzzy number is illustrated in Figure 1.

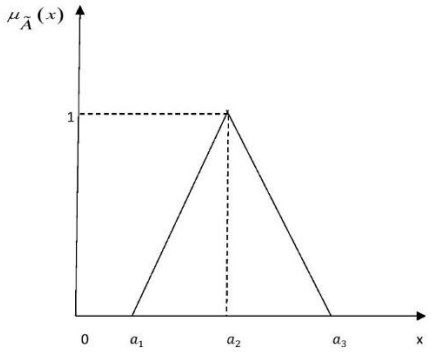


Figure 1: *Triangular fuzzy number*

Definition 2.3. Let the two triangular fuzzy numbers be $\tilde{P} \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3)$ and $\tilde{Q} \approx (\tilde{b}_1, \tilde{b}_2, \tilde{b}_3)$ and then the **arithmetic operations on TFN** be given as follows:

(A)Addition

$$\tilde{P} + \tilde{Q} \approx \left(\tilde{m}_1 + \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}\right) \quad (1)$$

(B)Subtraction

$$\tilde{P} - \tilde{Q} \approx \left(\tilde{m}_1 - \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}\right) \quad (2)$$

(C) Multiplication

$$\tilde{P}.\tilde{Q} \approx \left(\tilde{m}_1.\tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}\right) \quad (3)$$

(D) Division

$$\frac{\tilde{P}}{\tilde{Q}} \approx \left(\frac{\tilde{m}_1}{\tilde{m}_2}, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}\right) \quad (4)$$

Definition 2.4. For every triangular fuzzy number $\tilde{P} \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3) \in F(R)$ **ranking function** $\Re: F(R) \to R$ is defined by graded mean as

$$\Re(\tilde{P}) = \frac{(\tilde{a}_1 + 4\tilde{a}_2 + \tilde{a}_3)}{6}$$

For any two TFN $\tilde{P} \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3)$ and $\tilde{Q} \approx (\tilde{b}_1, \tilde{b}_2, \tilde{b}_3)$ we have the following comparisons,

$$(a)\tilde{P} \succ \tilde{Q} \Leftrightarrow \Re(\tilde{P}) > \Re(\tilde{Q})$$
$$(b)\tilde{P} \prec \tilde{Q} \Leftrightarrow \Re(\tilde{P}) < \Re(\tilde{Q})$$
$$(c)\tilde{P} \approx \tilde{Q} \Leftrightarrow \Re(\tilde{P}) = \Re(\tilde{Q})$$
$$(d)\tilde{P} - \tilde{Q} \approx 0 \Leftrightarrow \Re(\tilde{P}) - \Re(\tilde{Q}) = 0$$

A triangular fuzzy number $\tilde{P} \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3) \in F(R)$ is known to be **positive** if $\Re(\tilde{P}) > 0$ and defined by $\tilde{P} \succ 0$

Definition 2.5. [11] Let a non–empty set be X. An **Intuitionistic fuzzy set (IFS)** $\tilde{A}'$ is defined as $\tilde{A}' = \{(x, \mu_{\tilde{A}'}(x), \gamma_{\tilde{A}'}(x)/x \in X)\}$, where $\mu_{\tilde{A}'}: X \to [0,1]$ and $\gamma_{\tilde{A}'}: X \to [0,1]$ denotes the degree of membership

and degree of non–membership functions respectively where $x \in X$, for every $x \in X, 0 \leq \mu_{\tilde{A}'}(x) + \gamma_{\tilde{A}'}(x) \leq 1$.

Definition 2.6 [11] An intuitionistic fuzzy set described on R, the real numbers are said to be an **Intuitionistic fuzzy number (IFN)** if its membership function $\mu_{\tilde{A}'}: R \to [0,1]$ and its non–membership function $\gamma_{\tilde{A}'}: R \to [0,1]$ should be agreeable to the following conditions:

i) $\tilde{A}'$ is normal, which means that there exists an $x \in R$, such that $\mu_{\tilde{A}'}(x) = 1, \gamma_{\tilde{A}'}(x) = 0$

ii) $\tilde{A}'$ is convex for the membership functions $\mu_{\tilde{A}'}$, which means that there exists $x_1, x_2 \in R$ and $\lambda \in [0,1]$ such that $\mu_{\tilde{A}'}(\lambda x_1 + (1-\lambda)x_2) \geq min\{\mu_{\tilde{A}'}(x_1), \mu_{\tilde{A}'}(x_2)\}$.

iii) $\tilde{A}'$ is concave for the non–membership function $\gamma_{\tilde{A}'}$, which means that there exists $x_1, x_2 \in R$ and $\lambda \in [0,1]$ such that $\gamma_{\tilde{A}'}(\lambda x_1 + (1-\lambda)x_2) \leq max\{\gamma_{\tilde{A}'}(x_1), \gamma_{\tilde{A}'}(x_2)\}$.

Definition 2.7. [11] A fuzzy number $\tilde{A}'$on R is said to be a **triangular intuitionistic fuzzy number (TIFN)** if its membership function $\mu_{\tilde{A}'}: R \to [0,1]$ and non-membership function $\gamma_{\tilde{A}'}: R \to [0,1]$ has the following conditions:

$$\mu_{\tilde{A}'}(x) = \begin{cases} \frac{x-\tilde{a}_1}{\tilde{a}_2-\tilde{a}_1} \text{ for } \tilde{a}_1 \leq x \leq \tilde{a}_2 \\ 1 \text{ for } x = \tilde{a}_2 \\ \frac{\tilde{a}_3-x}{\tilde{a}_3-\tilde{a}_2} \text{ for } \tilde{a}_2 \leq x \leq \tilde{a}_3 \\ 0 \text{ otherwise} \end{cases}$$

and

$$\gamma_{\tilde{A}'}(x) = \begin{cases} 1 \text{ for } x < \tilde{a}_1', x > \tilde{a}_3' \\ \frac{\tilde{a}_2-x}{\tilde{a}_2-\tilde{a}_1'} \text{ for } \tilde{a}_1' \leq x \leq \tilde{a}_2 \\ 0 \text{ for } x = \tilde{a}_2 \\ \frac{x-\tilde{a}_2}{\tilde{a}_3-\tilde{a}_2} \text{ for } \tilde{a}_2 \leq x \leq \tilde{a}_3' \end{cases}$$

and is given by $\tilde{A}' = (a_1, a_2, a_3; a_1', a_2, a_3')$ where $a_1' \leq a_1 \leq a_2 \leq a_3 \leq a_3'$.
The triangular intuitionistic fuzzy number is illustrated in Figure 2.

Cases: Let $\tilde{A}' = (a_1, a_2, a_3; a_1', a_2, a_3')$ be a TIFN then the following cases arise.

Case:1 If $\tilde{a}_1' = \tilde{a}_1, \tilde{a}_3' = \tilde{a}_3$ then $\tilde{A}'$ represent a triangular fuzzy number.

Case:2 If $\tilde{a}_1' = \tilde{a}_1 = \tilde{a}_2 = \tilde{a}_3' = \tilde{a}_3 = \tilde{m}$ then $\tilde{A}'$ represent a real number $\tilde{m}$. The parametric form of TIFN $\tilde{A}'$ is represented as $\tilde{A}' = \left(\tilde{\alpha}, \tilde{m}, \tilde{\beta}; \tilde{\alpha}', \tilde{m}, \tilde{\beta}'\right)$ where $\tilde{\alpha}, \tilde{\alpha}'$&$\tilde{\beta}, \tilde{\beta}'$represents the left spread and right spread of membership functions and non–membership functions respectively.

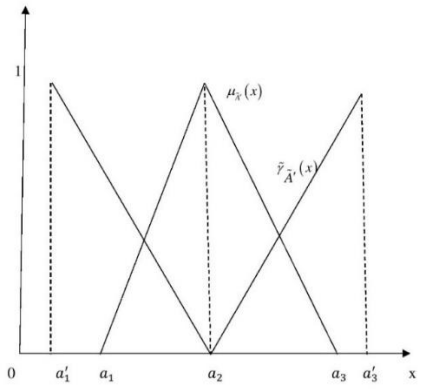


Figure 2: *Triangular intuitionistic fuzzy number*

Definition 2.8. The extension of fuzzy arithmetic operations of Ming Ma et al [10] to the set of triangular intuitionistic fuzzy numbers based upon both location indices and functions of fuzziness indices. The location indices number is taken in the regular arithmetic while the functions of

fuzziness indices are assumed to follow the lattice rule which is the least upper bound in the lattice $\tilde{I}'$.

For any two arbitrary TIFN $\tilde{P}' \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1')$ and $\tilde{Q}' \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$ and $* \in \{+, -, \times, \div\}$, then the **arithmetic operations on TIFN** are defined by $\tilde{P}' * \tilde{Q}' = (\tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1 \vee \tilde{\alpha}_2, \tilde{\beta}_1 \vee \tilde{\beta}_2; \tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1' \vee \tilde{\alpha}_2', \tilde{\beta}_1' \vee \tilde{\beta}_2')$

In particular, for any two TIFN $\tilde{P}' \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1')$ and $\tilde{Q}' \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$ the arithmetic operations are defined as

$\tilde{P}' * \tilde{Q}' = (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1') * (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$

$\tilde{P}' * \tilde{Q}' = (\tilde{m}_1 * \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}; \tilde{m}_1 * \tilde{m}_2, max\{\tilde{\alpha}_1', \tilde{\alpha}_2'\}, max\{\tilde{\beta}_1', \tilde{\beta}_2'\})$

$\tilde{P}' * \tilde{Q}' = (\tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1 \vee \tilde{\alpha}_2, \tilde{\beta}_1 \vee \tilde{\beta}_2; \tilde{m}_1 * \tilde{m}_2, \tilde{\alpha}_1' \vee \tilde{\alpha}_2', \tilde{\beta}_1' \vee \tilde{\beta}_2')$

In particular, for any two TIFN $\tilde{P}' \approx (\tilde{a}_1, \tilde{a}_2, \tilde{a}_3; \tilde{a}_1', \tilde{a}_2', \tilde{a}_3') \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1')$, $\tilde{Q}' \approx (\tilde{b}_1, \tilde{b}_2, \tilde{b}_3; \tilde{b}_1', \tilde{b}_2', \tilde{b}_3') \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_2', \tilde{\beta}_2')$we define:

Addition

$$\tilde{P}' + \tilde{Q}' = (\tilde{m}_1 + \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}; \tilde{m}_1 + \tilde{m}_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (5)$$

Subtraction

$$\tilde{P}' - \tilde{Q}' = (\tilde{m}_1 - \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}; \tilde{m}_1 - \tilde{m}_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (6)$$

Multiplication

$$\tilde{P}' \times \tilde{Q}' = (\tilde{m}_1 \times \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}; \tilde{m}_1 \times \tilde{m}_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (7)$$

Division

$$\tilde{P}' \div \tilde{Q}' = (\tilde{m}_1 \div \tilde{m}_2, max\{\tilde{\alpha}_1, \tilde{\alpha}_2\}, max\{\tilde{\beta}_1, \tilde{\beta}_2\}; \tilde{m}_1 \div \tilde{m}_2, max\{\alpha_1', \alpha_2'\}, max\{\beta_1', \beta_2'\}) \quad (8)$$

Definition 2.9. Consider an arbitrary TIFN $\tilde{A}' = (a_1, a_2, a_3; a_1', a_2, a_3') = (m, \alpha, \beta; m, \alpha', \beta')$ and the magnitude of TIFN $\tilde{A}'$ is given by

$$mag(\tilde{A}') = \frac{1}{2}\int_0^1 (\tilde{\beta} + \tilde{\beta}' + 6\tilde{m} - \tilde{\alpha} - \tilde{\alpha}')f(r)dr$$

In real-life scenarios, decision-makers select the value of $\tilde{f}(\tilde{r}')$ based on their circumstances. Here for our ease, we choose $f(r) = r^2$

$$\therefore\ mag(\tilde{A}') = \left(\frac{\tilde{\beta} + \tilde{\beta}' + 6\tilde{m} - \tilde{\alpha} - \tilde{\alpha}'}{6}\right)$$

For any two TIFN $\tilde{P}' \approx (\tilde{m}_1, \tilde{\alpha}_1, \tilde{\beta}_1; \tilde{m}_1, \tilde{\alpha}_1', \tilde{\beta}_1',)$ & $\tilde{Q}' \approx (\tilde{m}_2, \tilde{\alpha}_2, \tilde{\beta}_2; \tilde{m}_2, \tilde{\alpha}_1', \tilde{\beta}_1')$ in $F(R)$, we define

$$(a)\tilde{P}' \geq \tilde{Q}' \Leftrightarrow mag(\tilde{P}') \geq mag(\tilde{Q}')$$
$$(b)\tilde{P}' \leq \tilde{Q}' \Leftrightarrow mag(\tilde{P}') \leq mag(\tilde{Q}')$$
$$(c)\tilde{P}' \approx \tilde{Q}' \Leftrightarrow mag(\tilde{P}') = mag(\tilde{Q}')$$

III. Model Description and Stability Conditions

Presume that a single type of customer enters the queue through a Poisson process with a fuzzy parameter $\tilde{\lambda}$. They form a queue to receive an exponentially distributed service with a fuzzy rate $\tilde{\mu}$ from an unreliable server whose failure times are independent and exponentially distributed with a fuzzy rate $\tilde{\omega}$. When a customer's service is obstructed due to a server failure, the customer can exit

the zone and enter the retrial orbit, where a rate is a fuzzy number $\tilde{\theta}$. During this time, the server is delegated to be repaired at a variable rate $\tilde{\psi}$. When the server is functional and idle, orbit consumers do not rejoin the standard queue and instead try to enter the server explicitly after an unspecified period. All processes in the system are hypothesized to be self-contained and distributed uniformly. The queue and orbit sizes are assumed to be infinite, and the service discipline is FIFO (first in first out).

Customers enter the system through a Poisson process with a rate $\tilde{\lambda} > 0; \tilde{\lambda}' > 0$ and response times are independent and identically distributed exponential random variables with rate $\tilde{\mu} > 0; \tilde{\mu}' > 0$. Server faults happen at a stable level $\tilde{\omega} > 0; \tilde{\omega}' > 0$, and server repair occurs at a constant rate of $\tilde{\psi} > 0; \tilde{\psi}' > 0$. A customer whose service is disrupted by a server outage joins orbit and spends an accelerating span with a rate $\tilde{\theta} > 0; \tilde{\theta}' > 0$, whereby it arrives service (if available) or persists in orbit for a supplemental period with rate $\tilde{\theta}$ and exponentially distributed time.

The number of clients/messages in the line at the time $\tilde{t}$ is signified by $\tilde{N}_{q\tilde{t}}$. $\tilde{N}_{o\tilde{t}}$ stands for the number of clients/messages in the orbit at the time $\tilde{t}$. The random process $\tilde{X}_{\tilde{t}}$ is the invasion status of the site supplied by

$$\tilde{X}_{\tilde{t}} = \begin{cases} 1 & \text{if the site is overloaded at the time period } \tilde{t} \\ 0 & \text{if the site is not occupied at the time period } \tilde{t} \end{cases}$$

whereas $\tilde{Y}_{\tilde{t}}$ exemplifies the site's operational capability at the time $\tilde{t}$ categorized by

$$\tilde{Y}_{\tilde{t}} = \begin{cases} 1 & \text{if the site is up and running at the time period } \tilde{t} \\ 0 & \text{if the site is down at the time period } \tilde{t} \end{cases}$$

Then $\{(\tilde{N}_{q\tilde{t}}, \tilde{X}_{\tilde{t}}, \tilde{N}_{o\tilde{t}}, \tilde{Y}_{\tilde{t}}): \tilde{t} \geq 0\}$ is a continuous-time Markov process of explaining the system's state at the time $\tilde{t}$. Let $\tilde{N}_{s\tilde{t}}$ symbolised the total number of clients/messages in the system at a time $\tilde{t}$ which means it represents the number in orbit, queue, and in service. The procedure $\{\tilde{N}_{s\tilde{t}}: \tilde{t} \geq 0\}$ exemplifies how the system size varies over time. The server is operational for a proportion of time, and $\frac{\tilde{\psi}}{(\tilde{\psi}+\tilde{\omega})}$; thus, the excellent service rate is $\frac{\tilde{\psi}\tilde{\mu}}{(\tilde{\psi}+\tilde{\omega})}$ and $\frac{\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})}{\tilde{\psi}\tilde{\mu}} < 1$ is a necessary and sufficient condition for stability analysis.[15]

Specifying $\tilde{\pi}_{m,n,o,p}$ as the restricting probability that the system is in the state (m, n, o, p), i.e.,

$$\tilde{\pi}_{m,n,o,p} = \tilde{t} \xrightarrow{\lim} \infty P\left(\tilde{N}_{q\tilde{t}}=m, \tilde{X}_{\tilde{t}}=n, \tilde{N}_{o\tilde{t}}=o, \tilde{Y}_{\tilde{t}}=p\right),$$

Where the index m represents the queue size, the index n represents the invasion status (0 or 1), the index o represents the size of the orbit and the index p represents the operational capability of the server (0 or 1). The orbit and queue size are depicted by the morph variables $\tilde{v}_1$ and $\tilde{v}_2$.

Let the generating function of $\tilde{\pi}_{m,n,o,p}$ concerning the orbit size as follows
$\varepsilon_{m,n,p}(\tilde{v}_1) = \sum_{o=0}^{\infty} \tilde{v}_1^o \, \tilde{\pi}_{m,n,o,p}$ and

Let the generating function of $\tilde{\varepsilon}_{m,n,p}(\tilde{v}_1)$ concerning the queue size as follows

$$\tilde{\chi}_{n,p}(\tilde{v}_1, \tilde{v}_2) = \sum_{m=0}^{\infty} \tilde{v}_2^m \, \varepsilon_{m,n,p}(\tilde{v}_1)$$

Consider the probability-generating function as $\tilde{\varepsilon}_{0,0,1}(\tilde{v}_1)$, $\tilde{\chi}_{0,0}(\tilde{v}_1, \tilde{v}_2)$and $\tilde{\chi}_{1,1}(\tilde{v}_1, \tilde{v}_2)$ when the server is sluggish, ceased, and strenuous respectively. Define

$$E(\tilde{v}_1, \tilde{v}_2) = \sum_{m=0}^{\infty}\sum_{o=0}^{\infty} p(o,m)\, \tilde{v}_1^o \tilde{v}_2^m = \tilde{\varepsilon}_{0,0,1}(\tilde{v}_1) + \tilde{\chi}_{0,0}(\tilde{v}_1, \tilde{v}_2) + \tilde{\chi}_{1,1}(\tilde{v}_1, \tilde{v}_2)$$

is the joint probability generating function of the orbit and queue size where $p(o,m)$ is the joint probability mass function of $\widetilde{N}_q$ and $\widetilde{N}_o$. And

$$F(\tilde{v}) = \sum_{o=0}^{\infty} q(o)\, \tilde{v}^o = \tilde{\varepsilon}_{0,0,1}(\tilde{v}) + \tilde{\chi}_{0,0}(\tilde{v}, \tilde{v}) + \tilde{v}\tilde{\chi}_{1,1}(\tilde{v}, \tilde{v})$$

is the probability-generating function of system size where $q(o)$ denote the probability mass function of $\widetilde{N}_s$.

IV. Single Server Retrial Queues $(FM/FM/1)$: $(\infty/FIFO)$ in Fuzzy and Intuitionistic Fuzzy Environment

We assume a solitary-server retrial fuzzy queuing system with limitless capacity. The inter-arrival rates $\tilde{\lambda}$, service rate $\tilde{\mu}$, retrial rate $\tilde{\theta}$, failure rate $\widetilde{\omega}$ and repair rate $\tilde{\psi}$ are nearly comprehended and depicted by a fuzzy set,

$$\tilde{\lambda} = \{a, \mu_{\tilde{\lambda}}(a)/a \in A\}$$
$$\tilde{\mu} = \{s, \mu_{\tilde{\mu}}(s)/s \in S\}$$
$$\tilde{\theta} = \{o, \mu_{\tilde{\theta}}(o)/o \in O\}$$
$$\widetilde{\omega} = \{f, \mu_{\widetilde{\omega}}(f)/f \in F\}$$
$$\tilde{\psi} = \{r, \mu_{\tilde{\psi}}(r)/r \in R\}$$

In this, A, S, O, F & R are a traditional universal set of arrival rate, service rate, orbit rate, failure rate, and repair rate respectively and their corresponding membership functions are given as $\mu_{\tilde{\lambda}}(a), \mu_{\tilde{\mu}}(s), \mu_{\tilde{\theta}}(o), \mu_{\widetilde{\omega}}(f) \& \mu_{\tilde{\psi}}(r)$ respectively. In addition to that, assume a solitary server retrial intuitionistic fuzzy queuing system with limitless capacity. The inter-arrival rates $\tilde{\lambda}'$, service rate $\tilde{\mu}'$, retrial rate $\tilde{\theta}'$, failure rate $\widetilde{\omega}'$ and repair rate $\tilde{\psi}'$ are nearly comprehended and depicted by an intuitionistic fuzzy set,

$$\tilde{\lambda}' = \{a, \mu_{\tilde{\lambda}'}(a), \gamma_{\tilde{\lambda}'}(a)/a \in A\}$$
$$\tilde{\mu}' = \{s, \mu_{\tilde{\mu}'}(s), \gamma_{\tilde{\mu}'}(s)/s \in S\}$$
$$\tilde{\theta}' = \{o, \mu_{\tilde{\theta}'}(o), \gamma_{\tilde{\theta}'}(o)/o \in O\}$$
$$\widetilde{\omega}' = \{f, \mu_{\widetilde{\omega}'}(f), \gamma_{\widetilde{\omega}'}(f)/f \in F\}$$
$$\tilde{\psi}' = \{r, \mu_{\tilde{\psi}'}(r), \gamma_{\tilde{\psi}'}(r)/r \in R\}$$

In this, A, S, O, F & R are a traditional set of arrival, service, orbit, failure, and repair rate respectively and their corresponding membership and non-membership functions are given as $\mu_{\tilde{\lambda}'}(a), \mu_{\tilde{\mu}'}(s), \mu_{\tilde{\theta}'}(o), \mu_{\widetilde{\omega}'}(f), \mu_{\tilde{\psi}'}(r)$ & $\gamma_{\tilde{\lambda}'}(a), \gamma_{\tilde{\mu}'}(s), \gamma_{\tilde{\theta}'}(o), \gamma_{\widetilde{\omega}'}(f), \gamma_{\tilde{\psi}'}(r)$ respectively.

V. Solo Server Retrial Queuing Model with Infinite Capacity

Let the following assumptions $\tilde{\lambda}$ and $\tilde{\lambda}'$ be the fuzzy and intuitionistic fuzzy arrival rates respectively; $\tilde{\mu}$ and $\tilde{\mu}'$ be the fuzzy and intuitionistic fuzzy service rates respectively; $\tilde{\theta}$ and $\tilde{\theta}'$ be the fuzzy and intuitionistic fuzzy retrial(orbit) rate; $\widetilde{\omega}$ and $\widetilde{\omega}'$ be the fuzzy and intuitionistic fuzzy failure rates respectively; $\tilde{\psi}$ and $\tilde{\psi}'$ be the fuzzy and intuitionistic fuzzy repair rates respectively. At the steady-state, the $FIFO$ discipline is upheld and the capacity is unlimited.

The following are the fabrication characteristics of the above model:

i) The number of customers in the queue is given as

$$\tilde{N}_q = \frac{\tilde{\lambda}\left[\tilde{\mu}\tilde{\omega}(\tilde{\mu}+\tilde{\omega})+\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})^2\right]}{\tilde{\mu}(\tilde{\psi}+\tilde{\omega})[\tilde{\psi}(\tilde{\mu}+\tilde{\omega})-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} \quad (9)$$

ii) The sojourn time of customers in the queue is given as

$$\tilde{T}_q = \frac{\left[\tilde{\mu}\tilde{\omega}(\tilde{\mu}+\tilde{\omega})+\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})^2\right]}{\tilde{\mu}(\tilde{\psi}+\tilde{\omega})[\tilde{\psi}(\tilde{\mu}+\tilde{\omega})-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} \quad (10)$$

iii) The number of customers in the orbit is given as

$$\tilde{N}_o = \frac{\tilde{\psi}\tilde{\lambda}\tilde{\omega}[\tilde{\mu}(\tilde{\mu}+\tilde{\omega}-\tilde{\lambda})+\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]}{\tilde{\mu}[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})][\tilde{\psi}(\tilde{\mu}+\tilde{\omega})-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} + \frac{\tilde{\lambda}\tilde{\omega}(\tilde{\psi}+\tilde{\omega})}{\tilde{\theta}[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} \quad (11)$$

iv) The sojourn time of customers in the orbit is given as

$$\tilde{T}_o = \frac{\tilde{\psi}\tilde{\omega}[\tilde{\mu}(\tilde{\mu}+\tilde{\omega}-\tilde{\lambda})+\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]}{\tilde{\mu}[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})][\tilde{\psi}(\tilde{\mu}+\tilde{\omega})-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} + \frac{\tilde{\omega}(\tilde{\psi}+\tilde{\omega})}{\tilde{\theta}[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} \quad (12)$$

v) The number of customers in the system is given as

$$\tilde{N}_s = \frac{\tilde{\lambda}\left[\tilde{\mu}\omega+(\tilde{\psi}+\tilde{\omega})^2\right]}{(\tilde{\psi}+\tilde{\omega})[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} + \frac{\tilde{\lambda}\tilde{\omega}(\tilde{\psi}+\tilde{\omega})}{\tilde{\theta}[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} \quad (13)$$

vi) The waiting time of customers in the system is given as

$$\tilde{T}_s = \frac{\left[\tilde{\mu}\omega+(\tilde{\psi}+\tilde{\omega})^2\right]}{(\tilde{\psi}+\tilde{\omega})[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} + \frac{\tilde{\omega}(\tilde{\psi}+\tilde{\omega})}{\tilde{\theta}[\tilde{\psi}\tilde{\mu}-\tilde{\lambda}(\tilde{\psi}+\tilde{\omega})]} \quad (14)$$

VI. Mathematical Description

We considered a communications network with a cohort of streaming server devices linked to an interface message microprocessor as a bandwidth network. Messages arrive in a Poisson stream at the webserver. If the web host wants to transmit information to someone else host controller, one must deliver the data along with the node to the interface message processing unit with which it is hooked up. The message is acknowledged if the processor is free; alternatively, it is assumed to be a failure and the message is returned to the streaming server computer hard disk in a barrier to be transcoded at a later point and is considered a repair rate. In queuing terminology, the buffer in the host controller, the interface processing, and the transcoded policy correlate to the orbit, server, and retrial discipline, respectively. The above system can be modeled using an $FM/FM/1$ retrial queuing model. For consideration of performance and efficiency, the organization wants to learn more about the platform's characteristics, such as the expected wait time and the number of messages in orbit, queue, and system. Interpret the entry, departure, retrial, failure, and repair rate as both TFNs and TIFNs symbolized by $\tilde{\lambda}$, $\tilde{\lambda}'$; $\tilde{\mu}$, $\tilde{\mu}'$; $\tilde{\theta}$, $\tilde{\theta}'$; $\tilde{\omega}$, $\tilde{\omega}'$ and $\tilde{\psi}$, $\tilde{\psi}'$ respectively.

6.1 Solo Server Retrial Fuzzy Queuing Model with Unlimited Capability

Let $\tilde{\lambda} = (4,5,6)$, is the arrival rate, $\tilde{\mu} = (26,27,28)$ is the service rate, $\tilde{\theta} = (15,16,17)$ is the retrial rate, $\tilde{\omega} = (37,38,39)$ is the failure rate, $\tilde{\psi} = (47,48,49)$ is the repair rate.

Determine the TFN in the form of $(\tilde{m}, \tilde{\alpha}, \tilde{\beta})$ as $\tilde{\lambda} = (5,1,1)$, $\tilde{\mu} = (27,1,1)$, $\tilde{\theta} = (16,1,1)$, $\tilde{\omega} = (38,1,1)$, and $\tilde{\psi} = (48,1,1)$.

To determine the values of a number of messages and their sojourn time in the queue, orbit as well as a system using suitable formulas among (9), (10), (11), (12), (13) & (14). It is necessary to use the appropriate arithmetic operations described in (1), (2), (3), and (4) for add, sub, multiply, and divide, respectively.

The metrics of performance are calculated and tabulated as follows:

Table 1: *Performance Measures using Triangular Fuzzy Numbers*

Components	Number of Messages$(\widetilde{N})$	Waiting Time $(\widetilde{T})$
Queue	$\widetilde{N}_q = (-0.9171, 0.0829, 1.0829)$	$\tilde{T}_q = (-0.98342, 0.01658, 1.01658)$
Orbit	$\widetilde{N}_o = (0.4764, 1.4764, 2.4764)$	$\tilde{T}_o = (-0.7048, 0.2952, 1.2952)$
System	$\widetilde{N}_s = (0.7446, 1.7446, 2.7446)$	$\tilde{T}_s = (-0.6511, 0.3489, 1.3489)$

6.2 Solo Server Retrial Intuitionistic Fuzzy Queuing Model with Unlimited Capability

Let $\tilde{\lambda}' = (4,5,6;3,5,7)$, is the arrival rate, $\tilde{\mu}' = (26,27,28;25,27,29)$ is the service rate, $\tilde{\theta}' = (15,16,17;14,16,18)$ is the retrial rate, $\tilde{\omega}' = (37,38,39;36,38,40)$ is the failure rate, $\tilde{\psi}' = (47,48,49;46,48,50)$ is the repair rate.

Determine the TIFN in the form of $(\tilde{m}, \tilde{\alpha}, \tilde{\beta}; \tilde{m}, \tilde{\alpha}', \tilde{\beta}')$ as $\tilde{\lambda}' = (5,1,1;5,2,2)$, $\tilde{\mu}' = (27,1,1;27,2,2)$, $\tilde{\theta}' = (16,1,1;16,2,2)$, $\tilde{\omega}' = (38,1,1;38,2,2)$, and $\tilde{\psi}' = (48,1,1;48,2,2)$.

To determine the values of a number of messages and their sojourn time in the queue, orbit as well as a system using suitable formulas among (9), (10), (11), (12), (13) & (14). It is necessary to use the appropriate arithmetic operations described in (5), (6), (7), and (8) for add, sub, multiply, and divide, respectively.

The metrics of performance are calculated and tabulated as follows:

Table 2: *Performance Measures using triangular intuitionistic fuzzy numbers*

Components	Number of Messages$(\widetilde{N}')$	Waiting Time $(\widetilde{T}')$
Queue	$\widetilde{N}_q' = \begin{pmatrix} -0.9171, 0.0829, 1.0829; \\ -1.9171, 0.0829, 2.0829 \end{pmatrix}$	$\tilde{T}_q' = \begin{pmatrix} -0.98342, 0.01658, 1.01658; \\ -1.98342, 0.01658, 2.01658 \end{pmatrix}$
Orbit	$\widetilde{N}_o' = \begin{pmatrix} 0.4764, 1.4764, 2.4764; \\ -0.5236, 1.4764, 3.4764 \end{pmatrix}$	$\tilde{T}_o' = \begin{pmatrix} -0.7048, 0.2952, 1.2952; \\ -1.7048, 0.2952, 2.2952 \end{pmatrix}$
System	$\widetilde{N}_s' = \begin{pmatrix} 0.7446, 1.7446, 2.7446; \\ -0.2554, 1.7446, 3.7446 \end{pmatrix}$	$\tilde{T}_s = \begin{pmatrix} -0.6511, 0.3489, 1.3489; \\ -1.6511, 0.3489, 2.3489 \end{pmatrix}$

The following figures 3 – 14 depict the visualizations of Tables 1 and 2.

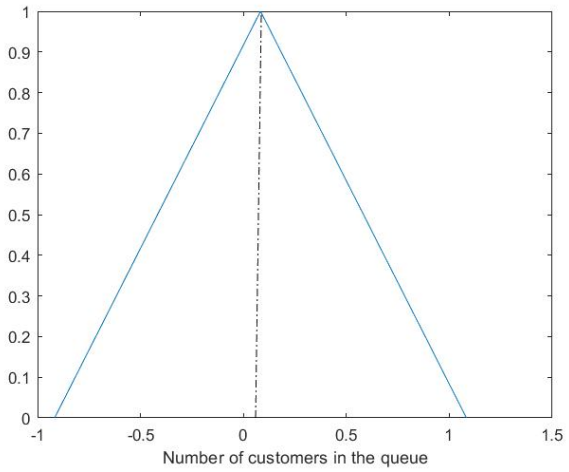


Figure 3: *The number of messages* $(\widetilde{N}_q)$ *in the queue*

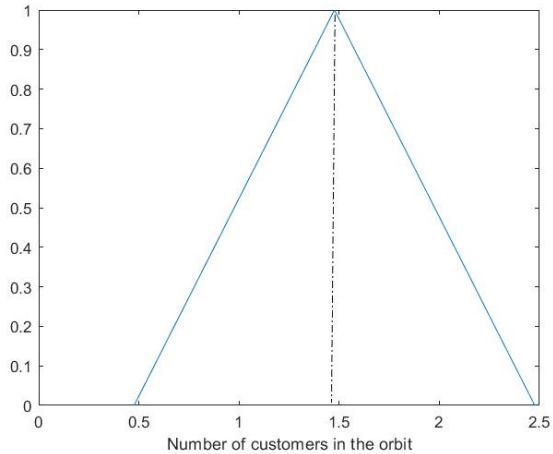


Figure 4: *The number of messages* $(\widetilde{N}_o)$*in the orbit*

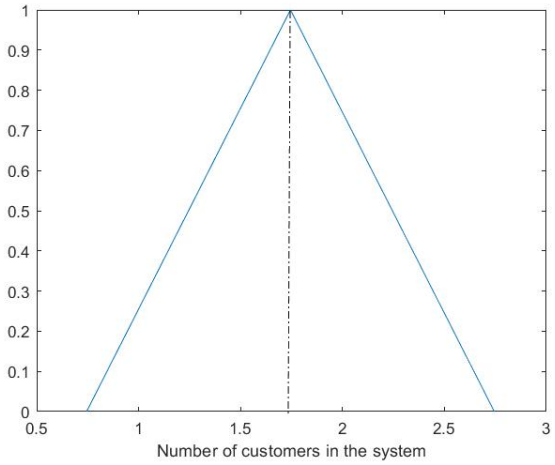


Figure 5: *The number of messages* $(\tilde{N}_s)$ *in the system*

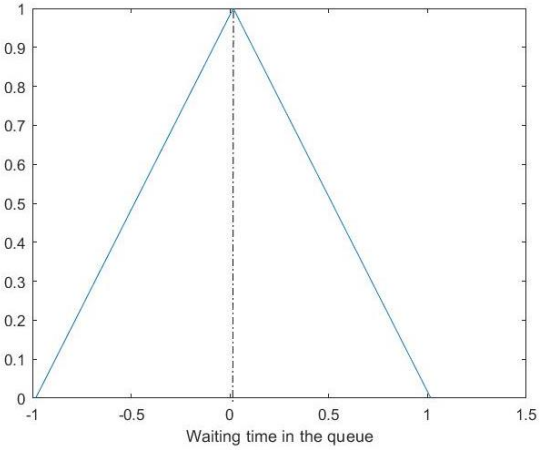


Figure 6: *The waiting time of messages* $(\tilde{T}_q)$*in the queue*

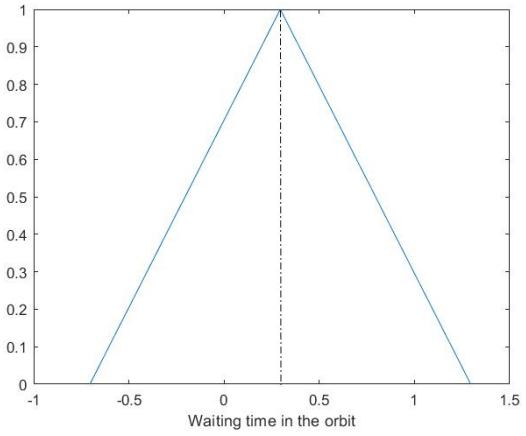


Figure 7: *The waiting time of messages* $(\tilde{T}_o)$ *in the orbit*

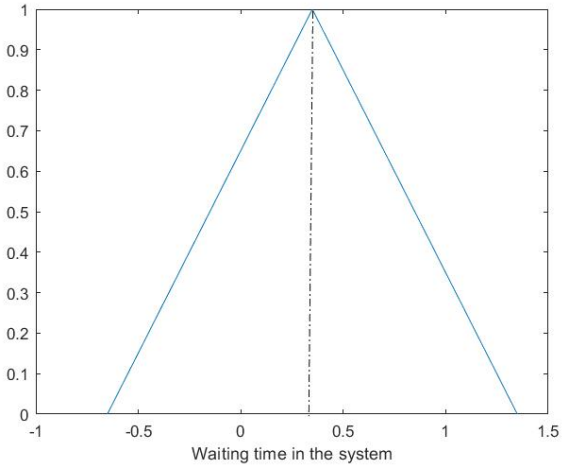


Figure 8: *The waiting time of messages* $(\tilde{T}_s)$ *in the system*

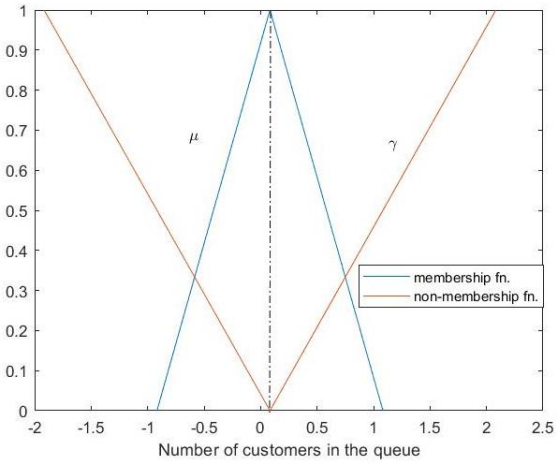


Figure 9: *The membership*$(\tilde{\mu})$ *and the non-membership*$(\tilde{\gamma})$ *functions of the number of messages in the queue* $\tilde{N}_q'$

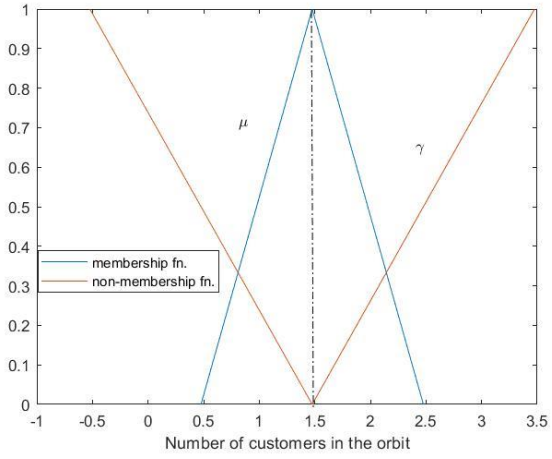


Figure 10: *The membership*$(\tilde{\mu})$ *and the non-membership*$(\tilde{\gamma})$ *functions of the no. of messages in the orbit* $\tilde{N}_o'$

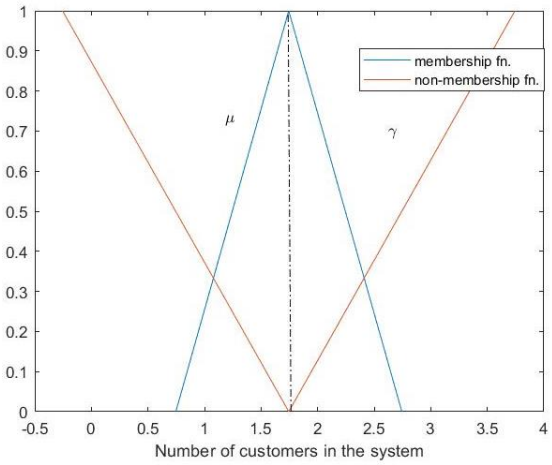


Figure 11: *The membership($\tilde{\mu}$) and the non-membership($\tilde{\gamma}$) functions of the number of messages in the system* $\tilde{N}_s'$

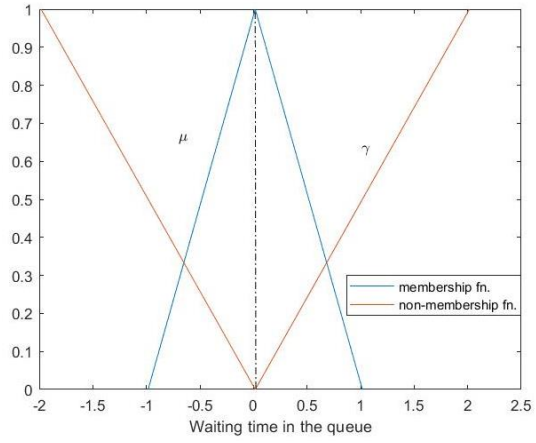


Figure 12: *The membership($\tilde{\mu}$) and the non-membership($\tilde{\gamma}$) functions of the waiting time of messages in the queue* $\tilde{T}_q'$

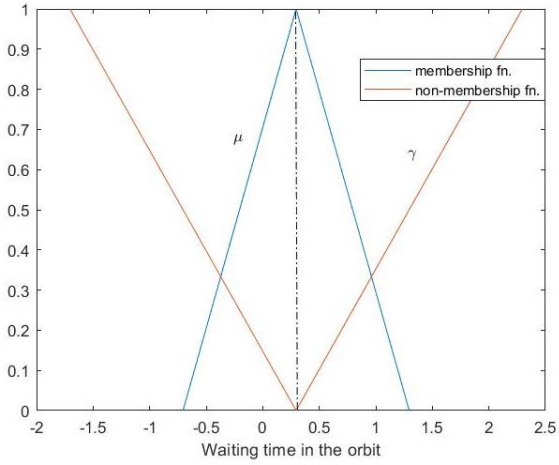


Figure 13: *The membership($\tilde{\mu}$) and the non-membership($\tilde{\gamma}$) functions of the waiting time of messages in the orbit* $\tilde{T}_o'$

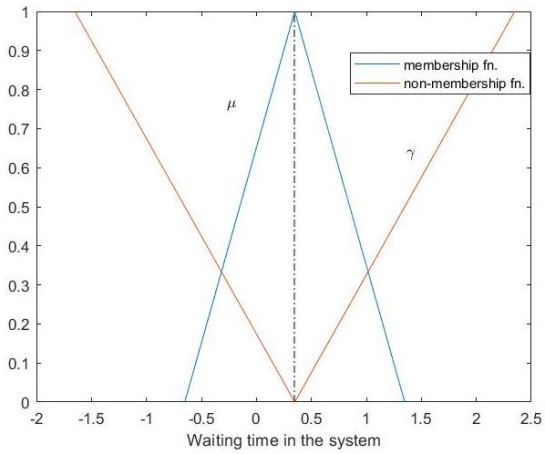


Figure 14: *The membership($\tilde{\mu}$) and the non-membership($\tilde{\gamma}$) functions of the waiting time of messages in the system* $\tilde{T}_s'$

VII. Conclusion

The retrial model with breakdowns and repairs has been studied for a large number of fuzzy parameters in the fuzzy queuing theory literature. Out of the existing methods for computing its characteristics, such as nonlinear programming, alpha cut, left-right method, interval arithmetic method, and so on, the present article shows that the suggested method is also suitable for dealing with this model, as evidenced by the example outlined in the previous section. Another advantage of the proposed method is that we solve the problems using the fuzzy value as is, instead of transitioning it to crisp, so it has a broad spectrum of applications in real-world situations. The predicted number of messages and their turnaround time in the queue, orbit, and system are efficaciously tabulated in this example, and the outcome is achieved in both a fuzzified and intuitionistic fuzzy environment. The TFN and TIFN arithmetical representations are used to compare the proposed queuing system's correctness. According to the research findings, the fuzzy queuing model's quality standards are within the range of the intuitionistic fuzzy queuing model's aggregated performance indicators. Because the intuitionistic fuzzy theory is more configurable, the intuitionistic fuzzy queuing model is significantly more efficient and appropriate for evaluating the dimensions of queuing models. As a result, intuitionistic fuzzy queuing, according to this

investigation, is one of the most efficient positions of computing evaluation criteria because the evidence gathered from the functionality is simpler to implement and discern. This strategy appears to be more pliable than the others because all estimations are fuzzy. As a result, fuzzy queuing models with a complicated structure benefit from it.

VIII. Acknowledgments

We are sincerely grateful to the referees for their insightful comments, as well as the authors of the journal/book that we used as a source of information.

References

[1] Ramesh, R. and Seenivasan, M. Performance calibrations of a single server glycolic acid-based beauty parlour by fuzzy retrial queuing models, International Conference on Advances in Materials Science, ICAMS 2021 (virtual), Materials Today: Proceedings, Vol. 51, 2021, pp 2422-2426.

[2] Sanga, Sudeep Singh. and Jain, Madhu. Fuzzy modelling of single server double orbit retrial queue, Journal of Ambient Intelligence and Humanized Computing, in press.,2022. Doi:10.1007/s12652-022-03705-3.

[3] Kannadasan, G. and Padmavathi, V. Classical fuzzy retrial queue with working vacation using hexagonal fuzzy numbers, Journal of Physics: Conference Series, International Conference on Advances in Physical Sciences and Materials, ICAPSM 2021(virtual), Vol. 2070, No. 1, 2021, code:174656. Doi: 10.1088/1742-6596/2070/1/012018.

[4] Jain, Madhu. and Sanga, Sudeep Singh. Fuzzy cost optimization and admission control for machine interference problem with general retrial, Journal of Testing and Evaluation, Vol. 48, No. 6, 2020. Doi: 10.1520/JTE20180882.

[5] Upadhyaya, Shweta. and Kushwaha, Chetna. Performance prediction and ANFIS computing for unreliable retrial queue with delayed repair under modified vacation policy, International Journal of Mathematics in Operational Research, Vol. 17, No. 4, 2020, pp 437-466. Doi: 10.1504/IJMOR.2020.110843.

[6] Sanga, Sudeep Singh. and Jain, Madhu. Cost optimization and ANFIS computing for admission control $M/M/1/k$ queue with general retrial times and discouragement, Applied Mathematics and Computation, Vol. 363, No. 124624, 2019. Doi: 10.1016/j.amc.2019.124624.

[7] Sanga, Sudeep Singh. and Jain, Madhu. $FM/FM/1$ double orbit retrial queue with messages' joining strategy: A parametric nonlinear programming approach, Applied Mathematics and Computation, Vol. 362, No. 124542, 2019. Doi: 10.1016/j.amc.2019.06.056.

[8] Ebenesar Anna Bagyam, J. and Udaya Chandrika, K. Fuzzy analysis of bulk arrival two phase retrial queue with vacation and admission control, Journal of Analysis, Vol. 27, No. 1, 2019, pp 209-232. Doi: 10.1007/s41478-018-0118-1.

[9] Kalpana, B. and Anusheela, N. A single server retrial fuzzy queue with priority and unequal service rates, Journal of Intelligent and Fuzzy Systems, Vol. 37, No. 1, 2019, pp 811-820. Doi: 10.3233/JIFS-181433.

[10] Ming Ma., Menahem Friedman. and Abraham Kandel. A new fuzzy arithmetic, Fuzzy sets and systems, Vol. 108, No. 1, 1999, pp 83-90.

[11] Shaw, A. K. and Roy, T. K. Some Arithmetic Operations on Triangular Intuitionistic Fuzzy Number and its Application on reliability evaluation, International Journal of Fuzzy Mathematics and Systems, Vol. 2, No. 4, 2012, pp 363-382.

[12] Sherman, Nathan. P. and Kharoufeh, Jeffrey. P. An $M/M/1$ retrial queue with unreliable server', Operations Research Letters, 34, 2006, pp 697-705. Doi: 10.1016/j.orl.2005.11.003.

[13] Mukeba Kanyinda, J. P. Analysis of retrial fuzzy queue with single server subject to breakdowns and repairs by flexible alpha-cuts method, Journal of Pure and Applied Mathematics:

Advances and Applications, Vol. 22, No. 1, 2020, pp 41-58. Doi: http://dx.doi.org/10.18642/jpamaa_7100122109.

[14] Kulkarni, V. G. and Choi, B. D. Retrial queues with server subject to breakdowns and repairs, Queueing Systems, Vol. 7, No. 2, 1990, pp 191-208. Doi: https://doi.org/10.1007/BF01158474.

[15] Pakes, A. G. Some conditions for ergodicity and recurrence of Markov chains, Operations Research, Vol. 17, No. 6, 1969, pp 1058-1061. Doi: https://doi.org/10.1287/opre.17.6.1058.

[16] Ke, J. C., Huang, H. I. and Lin, C. H. On retrial queueing model with fuzzy parameters, Physica A: Statistical Mechanics and its Applications, Vol. 374, No. 1, 2007, pp 272-280. https://doi.org/10.1016/j.physa.2006.05.057.

[17] Artalejo, J. R. and Lopez-Herrero, M. J. On the single server retrial queue with balking, INFOR: Information Systems and Operational Research, Vol. 38, No. 1, 2000, pp 33-50. Doi: 10.1080/03155986.2000.11732399.

[18] Kannadasan, G. and Padmavathi, V. Analysis of $FM/FG/1/N$ finite retrial queue with hexagonal fuzzy numbers, 2nd International Conference on Mathematical Techniques and Applications, ICMTA 2021, AIP Conference Proceedings, Vol. 2516, No. 360001, 2022, Doi: 10.1063/5.0108636

[19] Rani, Shobha. Madhu, Jain. and Sibasish, Dhibar. Analysis of Markovian retrial queue with double orbits, vacation, orbital search and disaster using ANFIS approach, Iranian Journal of Science, Vol. 47, No. 5-6, 2023, pp 1751 – 1764.

[20] Upadhyaya, S. Sharma, R. Agarwal, D. and Malik, G. Convexity analysis and cost optimization of a retrial queue with Bernoulli vacation and delayed phase mending, International Journal of System Assurance Engineering and Management, Vol. 14, No. 5, 2023, pp 1671 – 1690.

THE LENGTH-BIASED WEIGHTED WILSON HILFERTY DISTRIBUTION AND ITS APPLICATIONS

Shivendra Pratap Singh
•
Department of Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow, India
shivendra15.07@gmail.com

Surinder Kumar
•
Department of Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow, India
surinderntls@gmail.com

Naresh Chandra Kabdwal*
•
Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan, India
nareshchandra@banasthali.in

Abstract

In this article, we propose a new length-biased weighted form of Wilson Hilferty distribution named as Length-Biased Weighted Wilson Hilferty Distribution. The various Statistical properties of the proposed distribution like, reliability function, hazard rate function, reverse hazard rate function, moment generating function, quantile function, the coefficient of variation etc. are considered to understand its nature. Furthermore, we have used the method of maximum likelihood for estimation of the parameters of proposed distribution. Also, we obtain the Shannon's entropy, stochastic ordering, Lorenz and Bonferroni curves. The performance of the proposed distribution is compared with competitive distributions using two real data sets.

Keywords*:* Wilson Hilferty distribution, length-biased weighted Wilson Hilferty distribution, hazard function, reversed hazard function, maximum likelihood estimation.

1. Introduction

The weighted distribution arises when the observations are recorded from random process, the probability of recorded observations are not equal, and instead they are recorded according to some weighted function. The concept of weighted distributions was first given by [2]. Subsequently, [3] introduced a general form for model specification and data interpretation problems and identified that many situations can be modelled by weighted distributions.

Let T is as non-negative random variable with the probability density function (pdf) $f(t)$, then the weighted distribution is given by

$$g^w(t) = \frac{w(t)\, f(t)}{\Omega} \tag{1}$$

on the support of T, where $w(t) > 0$ and $\Omega = \int w(t) f(t) dt$ is considered as a normalizing factor that forces $g^w(t)$ to integrate to unity. When we replace $w(t) = t$ $(i.e.$ the length of units) in equation (1), we get a special case of the weighted distribution called length biased distribution see, [5]. A $r.v.\ T$, is said to have a length biased weighted distribution if its pdf is defined as

$$g^L(t) = \frac{w_L(t)\, w_j(t)\, f(t)}{\Omega} \tag{2}$$

Where, $\Omega = \int w_L(t)\, w_j(t)\, f(t) dt$ and $w_j(t) > 0$, provided thet $w_L(t) = t$.

Weighted distribution in general and length biased distributions are specifically very useful and convenient for the analysis of life time data. Weighted distributions are commonly used in study related to reliability, biomedicine, ecology, analysis of family data, branching process and various other fields of research, see [6], [8], [10] and [12].

Many length-biased weighted distributions with applications in different fields have been presented in the literature, see [9] discussed the characterization of inverse Gaussian and gamma distribution through their length biased distributions, [13] introduced a new class of weighted exponential distribution, [14], [15] discussed the length-biased weighted generalized Rayleigh distribution and also studied the length biased weighted Weibull distribution for rainfall data in India, a weighted Lindley distribution for survival data is given by [16], the length-biased lognormal distribution with application in the analysis of oil field exploration data is discussed by [17], [18] proposed the length-biased weighted exponential and Rayleigh distribution and its properties, different methods of estimation of parameters are applied for weighted exponential distribution by [19] which introduced by [13], [20] presented the length-biased weighted Lomax distribution and application with cancer data, [21] proposed inverted weighted exponential distribution and its properties, [24] discussed the length-biased exponential distribution for Bayesian reliability estimation, [23] proposed the length-biased weighted Lindley distribution, [25] proposed weighted exponentiated inverted exponential distribution and its properties, time and failure censoring schemes for Marshall Olkin alpha power extended Weibull distribution is presented by [26]. Recently, [28] proposed power weighted Sujatha distribution and application to survival times of patients head and neck cancer data, [27] proposed a weighted intervened exponential distribution as a lifetime model.

[1] introduced a Wison Hilferty distribution. For some recent developments of the Wilson Hilferty destitution the readers may, see [22]. Its probability density function (pdf), and cumulative distribution function (cdf), respectively as

$$\varphi(t) = \frac{3}{\Gamma(\alpha)} t^{3\alpha-1} \left(\frac{\alpha}{\beta}\right)^{\alpha} \exp\left\{-\frac{\alpha}{\beta} t^3\right\}; \qquad t,\ \alpha,\ \beta > 0 \tag{3}$$

$$\Phi(t) = \frac{\gamma\left(\alpha,\ \frac{\alpha}{\beta} t^3\right)}{\Gamma(\alpha)}; \qquad t,\ \alpha,\ \beta > 0 \tag{4}$$

where, α and β are the shape and scale parameters, respectively, $\Gamma(\alpha) = \int_0^{\infty} t^{\theta-1} e^{-t} dt$ and $\gamma(x, y) = \int_0^x w^{y-1} e^{-w} dw$ the complete and lower incomplete gamma functions, respectively.

In this paper, we propose a new length-biased distribution, called length-biased weighted Wilson-Hilferty distribution. Rest of the paper is structured as follows, in Section 2, the proposed distribution is introduced and its properties and reliability characteristics are discussed. In Section 3, the method of maximum likelihood is discussed for estimating the model parameters. Stochastic ordering and entropy are discussed in Section 4. Bonferroni and Lorenz curves and random number generation & Quantiles are discussed in Section 5 and 6, respectively. The applications of two real data sets are presented in Section 7. Finally, the conclusion is summarized in section 8.

2. Length-Biased Weighted Wilson Hilferty Distribution, its Properties and Reliability Characteristics

In this section, we develop the Length-Biased Weighted Wilson Hilferty distribution. For this proposed new distribution, we present the pdf, cdf, reliability function, hazard function, moments, skewness and discuss some properties.

Consider the weight function as $w_j(t) = \frac{n}{\alpha} t^k, k > 0$, $w_L(t) = t$. Hence using considered weight function, unit length and equation (3) into the equation (2), we obtain the density of Length-biased weighted Wilson Hilferty distribution (LBWWHD) of the form

$$f(t) = \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta}t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}; \quad t,\ \alpha,\ \beta,\ k > 0 \tag{5}$$

Where, α, β and k are the shape, scale and weighted parameters, respectively.

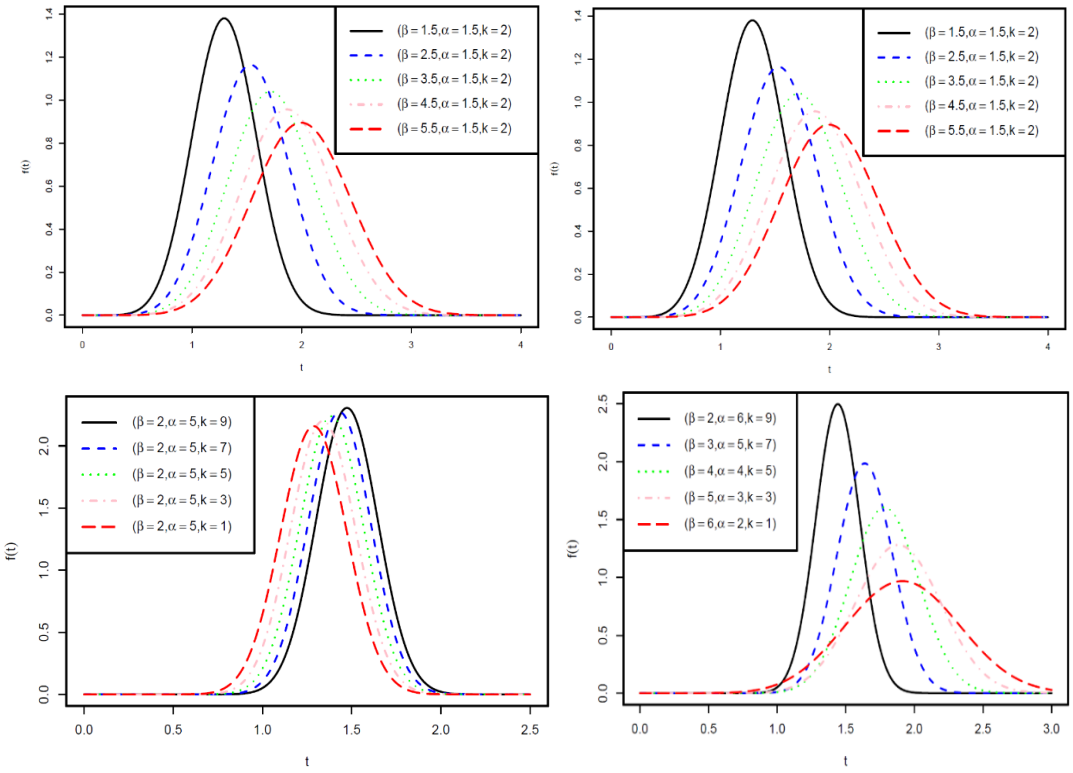


Figure 1: *pdf plots of LBWWHD*

Figure 1, clearly shows that LBWWH distribution is positively skewed.
The cdf of LBWWHD is given by

$$F(t) = \frac{\gamma\left(\alpha+\frac{k+1}{3}, \frac{\alpha}{\beta}t^3\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}; \quad t,\ \alpha,\ \beta,\ k > 0 \tag{6}$$

where, $\int_0^x t^{s-1} e^{-t} dt = \gamma(s,x)$ is a lower incomplete gamma function.
The reliability function $R(t)$ is given by

$$R(t) = 1 - \frac{\gamma\left(\alpha + \frac{k+1}{3},\ \frac{\alpha}{\beta}\ t^3\right)}{\Gamma\left(\alpha + \frac{k+1}{3}\right)}$$

On using some basic concept of an upper incomplete gamma integral's, it reduces to

$$R(t) = \frac{\Gamma\left(\alpha+\frac{k+1}{3},\ \frac{\alpha}{\beta}t^3\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)};\quad t,\ \alpha,\ \beta,\ k > 0 \tag{7}$$

where, $\int_x^\infty t^{s-1}\, e^{-t} dt = \Gamma(s,x)$ is an upper incomplete gamma function.

Table 1: Reliability function $R(t)$ of LBWWHD for $\alpha = 2$ *and* $\beta = 3$

t	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
0.5	0.9999827	0.9999965	0.9999993	0.9999999	1.0000000
0.6	0.9999268	0.9999822	0.9999958	0.999999	0.9999998
0.7	0.9997537	0.9999303	0.9999809	0.9999949	0.9999987
0.8	0.9993024	0.9997745	0.9999294	0.9999785	0.9999936
0.9	0.9982712	0.9993722	0.999779	0.9999244	0.9999748
1.0	0.9961542	0.9984503	0.9993945	0.99977	0.9999149

Table 2: *Reliability function* $R(t)$ *of LBWWHD for* $k = 2$ *and* $\beta = 3$

t	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$	$\alpha=5$
0.5	0.9999044	0.9999965	0.9999999	1.0000000	1.0000000
0.6	0.9997166	0.9999822	0.9999989	0.9999999	1.0000000
0.7	0.999292	0.9999303	0.9999935	0.9999994	0.9999999
0.8	0.9984419	0.9997745	0.9999691	0.9999958	0.9999994
0.9	0.9968914	0.9993722	0.9998803	0.9999773	0.9999957
1.0	0.9942659	0.9984503	0.9996054	0.9999000	0.9999746

Table 3: *Reliability function* $R(t)$ *of LBWWHD for* $k = 2$ *and* $\alpha = 3$

t	$\beta=1$	$\beta=2$	$\beta=3$	$\beta=4$	$\beta=5$
0.5	0.9993884	0.999997	0.9999999	1.0000000	1.0000000
0.6	0.9955998	0.9999748	0.9999989	0.9999999	1.0000000
0.7	0.979182	0.9998513	0.9999935	0.9999993	0.9999999
0.8	0.9297582	0.9993323	0.9999691	0.9999967	0.9999994
0.9	0.8219026	0.9975858	0.9998803	0.999987	0.9999977
1.0	0.6472319	0.9927078	0.9996054	0.9999557	0.9999921

From Table 1, 2 & 3, we conclude that, for the different values of α, β and k the reliability of the

distribution decreases with increase in the value of t.

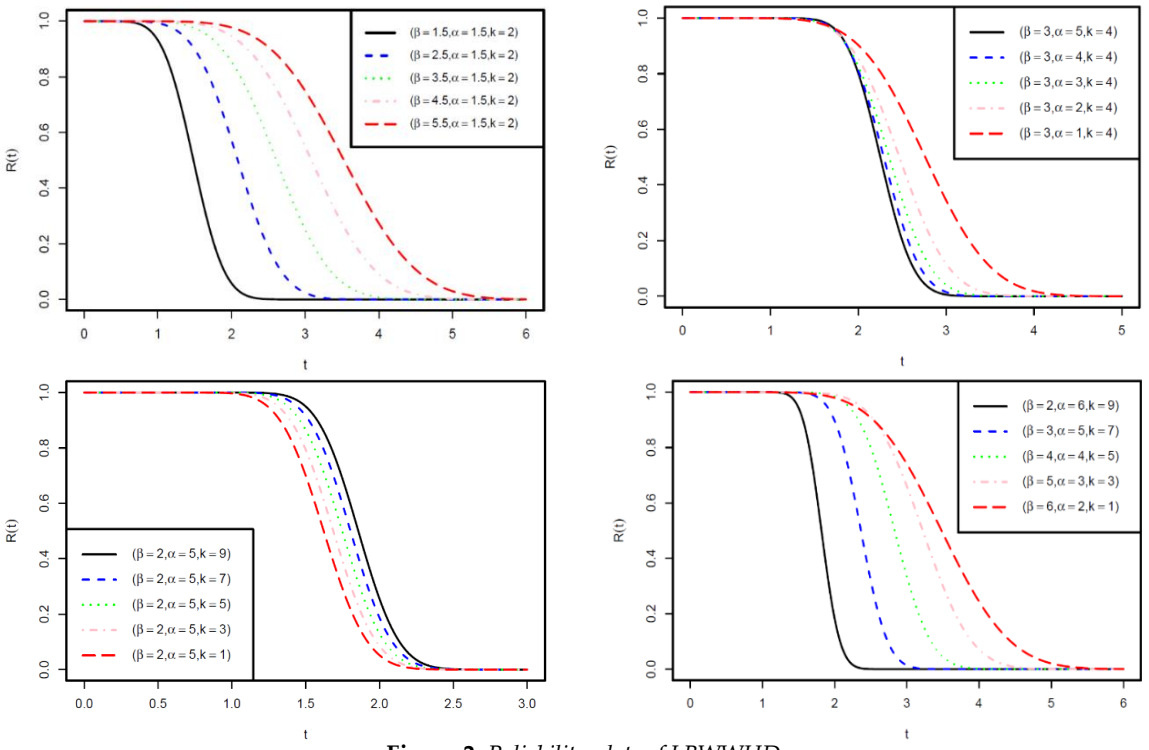


Figure 2: *Reliability plots of LBWWHD*

Figure 2, shows the reliability behavior of the LBWWHD for varying values of shape parameter α, scale parameter β and weighted parameter k. Reliability function behaves like decreasing function.

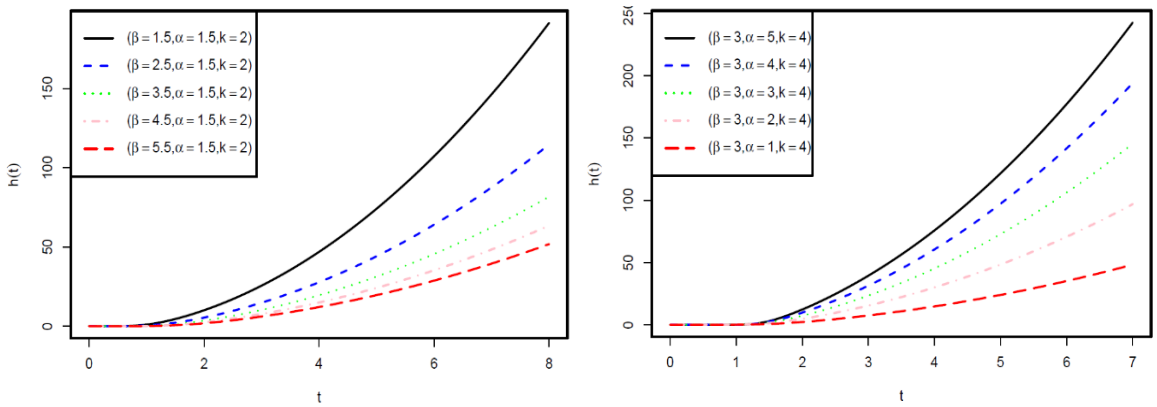
The hazard function is defined as

$$h(t) = \frac{f(t)}{R(t)}$$

$$= \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta}t^3\right\} \Gamma\left(\alpha+\frac{k+1}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right) \Gamma\left(\alpha+\frac{k+1}{3}, \frac{\alpha}{\beta}t^3\right)}$$

On simplifying, we get

$$h(t) = \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta}t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}, \frac{\alpha}{\beta}t^3\right)}; \quad t,\ \alpha,\ \beta,\ k > 0 \qquad (8)$$



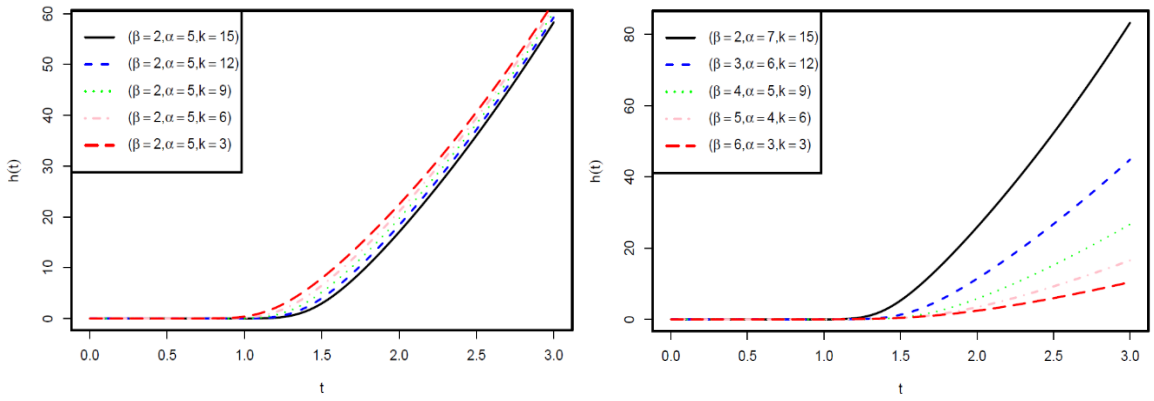


Figure 3: *Hazard rate plots of LBWWHD*

Figure 3, shows the behaviour of hazard function for distinct values of α, β and k. Clearly, it shows that the hazard function of LBWWH behaves increasing hazard rate.
The reverse hazard rate is defined as

$$R_h(t) = \frac{f(t)}{F(t)}$$

$$R_h(t) = \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta}\, t^3\right\} \Gamma\left(\alpha+\frac{k+1}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right) \gamma\left(\alpha+\frac{k+1}{3},\ \frac{\alpha}{\beta}\, t^3\right)}$$

On simplifying, we get

$$R_h(t) = \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta} t^3\right\}}{\gamma\left(\alpha+\frac{k+1}{3},\ \frac{\alpha}{\beta} t^3\right)}; \quad t,\ \alpha,\ \beta,\ k > 0$$

Theorem 2.1. For $r = 0, 1, 2, 3, \ldots$ r^{th} moment of random variable T is given by

$$\mu_r' = E(T^r) = \left(\frac{\beta}{\alpha}\right)^{r/3} \frac{\Gamma\left(\alpha+\frac{k+r+1}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \tag{9}$$

Proof. It T is a random variable with $pdf\ f(t)$ from equation (5), then the rth moment is

$$E(T^r) = \mu_r' = \int_0^\infty t^r \frac{3\left(\frac{\alpha}{\beta}\right)^{\left(\alpha+\frac{k+1}{3}\right)} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta}\, t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} dt$$

$$= \frac{3\left(\frac{\alpha}{\beta}\right)^{\left(\alpha+\frac{k+1}{3}\right)}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \int_0^\infty t^{(3\alpha+k+r)} \exp\left\{-\frac{\alpha}{\beta}\, t^3\right\} dt \tag{10}$$

Theorem follows on taking $y = \left(\frac{\alpha}{\beta}\, t^3\right)$, and using the gamma function in equation (10).

Lemma 2.1. If a random variable T follows Length-biased weighted Wilson Hilferty distribution then on substituting $r = 1, 2$ in equation (10), we obtain the mean and variance, respectively.

$$E(T) = \left(\frac{\beta}{\alpha}\right)^{1/3} \frac{\Gamma\left(\alpha+\frac{k+2}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}$$

$$E(T^2) = \left(\frac{\beta}{\alpha}\right)^{2/3} \frac{\Gamma\left(\alpha+1+\frac{k}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}$$

and,

$$\text{Variance}\,(T) = \left(\frac{\beta}{\alpha}\right)^{2/3} \frac{\Gamma\left(\alpha+1+\frac{k}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} - \left\{\left(\frac{\beta}{\alpha}\right)^{1/3} \frac{\Gamma\left(\alpha+\frac{k+2}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}\right\}^2$$

$$= \left(\frac{\beta}{\alpha}\right)^{2/3} \left[\frac{\Gamma\left(\alpha+1+\frac{k}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} - \left\{\frac{\Gamma\left(\alpha+\frac{k+2}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}\right\}^2\right]$$

Lemma 2.2. If a random variable T follows Length-biased Weighted Wilson Hilferty distribution then the coefficient of variation (C.V) is given by

$$\frac{\left[\Gamma\left(\alpha+1+\frac{k}{3}\right)\Gamma\left(\alpha+\frac{k+1}{3}\right)-\left\{\Gamma\left(\alpha+\frac{k+2}{3}\right)\right\}^2\right]^{1/2}}{\Gamma\left(\alpha+\frac{k+2}{3}\right)} \tag{11}$$

Proof. Coefficient of variation is given by,

$$C.V. = \frac{\sqrt{var(T)}}{E(T)} = \frac{\left(\frac{\beta}{\alpha}\right)^{1/3}\left[\frac{\Gamma\left(\alpha+1+\frac{k}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} - \left\{\frac{\Gamma\left(\alpha+\frac{k+2}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}\right\}^2\right]^{1/2}}{\left(\frac{\beta}{\alpha}\right)^{1/3}\frac{\Gamma\left(\alpha+\frac{k+2}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}}$$

$$= \frac{\left[\Gamma\left(\alpha+1+\frac{k}{3}\right)\Gamma\left(\alpha+\frac{k+1}{3}\right) - \left\{\Gamma\left(\alpha+\frac{k+2}{3}\right)\right\}^2\right]^{1/2}}{\Gamma\left(\alpha+\frac{k+2}{3}\right)}$$

Lemma 2.2, follows on using Lemma 2.1.

Table 4: *Coefficients of LBWWHD for* $\beta = 3$

α	k	*Mean*	*Variance*	*CV*	*Skewness*	*Kurtosis*
2	1	1.521644	0.10416361	0.212102093	0.03819147	2.895791
	2	1.590099	0.10032135	0.199192288	0.03175884	2.908649
	3	1.65319	0.09697548	0.188368465	0.02692476	2.918847
	4	1.71185	0.09402637	0.179126206	0.0231893	2.927098
	5	1.766776	0.09139999	0.17111638	0.02023536	2.933893
4	1	1.482214	0.05474587	0.157857336	0.01590133	2.944386
	2	1.519149	0.05351908	0.152283885	0.01427804	2.94851
	3	1.554379	0.05239432	0.147260131	0.01291154	2.95208
	4	1.588086	0.05135774	0.14270163	0.01174866	2.955199
	5	1.620426	0.05039801	0.138540736	0.01074952	2.957945

Table 5: *Coefficients of LBWWHD for* $\alpha = 3$

β	k	*Mean*	*Variance*	*CV*	*Skewness*	*Kurtosis*
2	1	1.306386	0.05475973	0.179126208	0.0231893	2.927098
	2	1.348303	0.05323016	0.171116302	0.02023536	2.933893
	3	1.387782	0.05185586	0.164088321	0.01785365	2.939574
	4	1.425148	0.05061155	0.157857366	0.01590133	2.944386
	5	1.460661	0.04947741	0.152283919	0.01427804	2.94851
4	1	1.645943	0.08692566	0.179126239	0.0231893	2.927098
	2	1.698755	0.08449762	0.171116343	0.02023536	2.933893
	3	1.748496	0.08231604	0.164088291	0.01785365	2.939574
	4	1.795574	0.08034083	0.157857365	0.01590133	2.944386
	5	1.840318	0.07854049	0.152283878	0.01427804	2.94851

According to the Table 4 & 5, skewness decreases and kurtosis increases whenever the values of k increases.

Lemma 2.3. If a random variable T follows Length-biased Weighted Wilson Hilferty distribution then harmonic mean (H) is given by

$$\frac{1}{H} = \left(\frac{\alpha}{\beta}\right)^{1/3} \frac{\Gamma\left(\alpha+\frac{k}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \tag{12}$$

Proof. The harmonic mean (H) is defined as

$$\frac{1}{H} = E\left(\frac{1}{T}\right) = \int_0^\infty \frac{1}{t} f(t)dt \tag{13}$$

Using equation(5), we get

$$\frac{1}{H} = \int_0^\infty \frac{1}{t} \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta} t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} dt$$

$$= \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \int_0^\infty t^{3\alpha+k-1} \exp\left\{-\frac{\alpha}{\beta} t^3\right\} dt$$

Lemma 2.3, follows on using the transformation $y = \left(\frac{\alpha}{\beta}\right) t^3$, and the gamma function.

Lemma 2.4. If a random variable T follows Length-biased Weighted Wilson Hilferty distribution then moment generating function (MGF) and characteristic function (CF) of T are respectively, given by

$$M_T(x) = \sum_{r=0}^\infty \frac{x^r}{r!} \left(\frac{\beta}{\alpha}\right)^{r/3} \frac{\Gamma\left(\alpha+\frac{k+r+1}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \tag{14}$$

$$\emptyset_T(x) = \sum_{r=0}^\infty \frac{(ix)^r}{r!} \left(\frac{\beta}{\alpha}\right)^{r/3} \frac{\Gamma\left(\alpha+\frac{k+r+1}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \tag{15}$$

Proof: On using equation (5) and Taylor's series expansion the Lemma 2.4, follows.

3. Parameter Estimation

In this section, we estimate the parameters of the LBWWHD by using the maximum likelihood technique. Let $T_1, T_2 \ldots T_n$ be the random sample of size n follows the LBWWHD(α, β, k), then the likelihood function given as

$$L(t) = \frac{3^n \left(\frac{\alpha}{\beta}\right)^{n\left(\alpha+\frac{k+1}{3}\right)}}{\left(\Gamma\left(\alpha+\frac{k+1}{3}\right)\right)^n} \prod_{i=1}^{n} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta}\, t^3\right\}$$

The log-likelihood function can be written as

$$\log L(t) = n\log 3 + n\left(\alpha+\tfrac{k+1}{3}\right)\log\alpha - n\left(\alpha+\tfrac{k+1}{3}\right)\log\beta - n\log\Gamma\left(\alpha+\tfrac{k+1}{3}\right) - \frac{\alpha}{\beta}\sum t_i{}^3 + (3\alpha+k)\sum\log t_i \quad (16)$$

Differentiating equations (16) partially with respect to α, β and k then equate to zero, we get normal equations on the following form

$$\frac{\partial \log L(t)}{\partial\beta} = 0 \;\Rightarrow\; \hat{\beta} = \frac{\alpha\sum t_i{}^3}{n\left(\alpha+\frac{k+1}{3}\right)} \quad (17)$$

$$\frac{\partial \log L(t)}{\partial\alpha} = 0 \;\Rightarrow\; n\log\alpha + \frac{n}{\alpha}\left(\alpha+\tfrac{k+1}{3}\right) - n\log\beta - \frac{1}{\beta}\sum t_i{}^3 - n\,\psi\left(\Gamma\left(\alpha+\tfrac{k+1}{3}\right)\right) + 3\sum\log t_i = 0 \quad (18)$$

$$\frac{\partial \log L(t)}{\partial k} = 0 \;\Rightarrow\; \frac{n}{3}\log\alpha - \frac{n}{3}\log\beta - \frac{n}{3}\,\psi\left(\Gamma\left(\alpha+\tfrac{k+1}{3}\right)\right) + \sum\log t_i = 0 \quad (19)$$

where, $\psi(z) = \frac{d}{dz}\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ is a logarithmic derivative of gamma function. As it seems, from equations (17), (18) and (19), the analytical solution of α, β and k are not available. Consequently, we have to use to non-linear estimation of the parameters using iterative method.

4. Stochastic Ordering and Entropy

Let X and Y be two independent random variables follows LBWWHD with shape parameter α, weighted parameter k and the scale parameters β_1 and β_2, respectively.

When $f_X(t)$ and $f_Y(t)$ be the density functions of X and Y, then X less than Y in likelihood order ($X \leq_{lr} Y$) if $\frac{f_Y(t)}{f_X(t)}$ is an increasing function of t. Here,

$$\frac{f_Y(t)}{f_X(t)} = \frac{3\left(\frac{\alpha}{\beta_2}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k}\exp\left\{-\frac{\alpha}{\beta_2}\,t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} \frac{\Gamma\left(\alpha+\frac{k+1}{3}\right)}{3\left(\frac{\alpha}{\beta_1}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k}\exp\left\{-\frac{\alpha}{\beta_1}\,t^3\right\}}$$

$$\frac{f_Y(t)}{f_X(t)} = \frac{\left(\frac{\alpha}{\beta_2}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k}\exp\left\{-\frac{\alpha}{\beta_2}\,t^3\right\}}{\left(\frac{\alpha}{\beta_1}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k}\exp\left\{-\frac{\alpha}{\beta_1}\,t^3\right\}}$$

$$\frac{f_Y(t)}{f_X(t)} = \left(\frac{\beta_1}{\beta_2}\right)^{\alpha+\frac{k+1}{3}} \exp\left\{\alpha\left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right)t^3\right\} \quad (20)$$

Differentiating equation (20), with respect to t, we get

$$\frac{d}{dt}\left(\frac{f_Y(t)}{f_X(t)}\right) = 3\alpha\, t^2 \left(\frac{\beta_1}{\beta_2}\right)^{\alpha+\frac{k+1}{3}} \left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right) \exp\left\{\alpha\left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right)t^3\right\} \geq 0$$

$$\text{hence, } \frac{d}{dt}\left(\frac{f_Y(t)}{f_X(t)}\right) \geq 0 \text{ when, } \left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right) \geq 0 \text{ i.e. } \beta_2 \geq \beta_1.$$

Therefore, $X \leq_{lr} Y$ for, $\beta_2 \geq \beta_1$. If $X \leq_{lr} Y$, then the following ordering shall also holds for the LBWWHD ([11]) ,

$$\begin{array}{c} X \leq_{lr} Y \Longrightarrow X \leq_{hr} Y \Longrightarrow X \leq_{mrl} Y \\ \Downarrow \\ X \leq_{st} Y \end{array}$$

for $\beta_2 \geq \beta_1$.

The idea of entropy has significant importance in several academic disciplines, including probability and statistics, physics, communication theory, and economics. Entropy is a measure that quantifies the level of variety, uncertainty, or unpredictability shown by a given system. The entropy of a random variable T may be defined as a quantitative measure of the level of uncertainty or variation associated with it.
The Shannon's entropy defined by

$$S(x) = -E[\log f(t)]$$

Using equation (5), we get

$$\begin{aligned} S(x) &= -E\left[\log\left(\frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k}\exp\left\{-\frac{\alpha}{\beta}\,t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}\right)\right] \\ &= -\log\left(\frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}\right) - (3\alpha+k)E(\log(t)) + \frac{\alpha}{\beta}E(t^3) \end{aligned} \qquad (21)$$

By solving the value of $E(\log(t))$ and $E(t^3)$ and put in equation (21), we get

$$S(x) = -\log\left(\frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}\right) - (3\alpha+k)\left(\log\left(\frac{\beta}{\alpha}\right)+\Psi\left(\alpha+\frac{k+1}{3}\right)\right)+\frac{\Gamma\left(\alpha+\frac{k+4}{3}\right)}{\Gamma\left(\alpha+\frac{k+1}{3}\right)}$$

5. Bonferroni and Lorenz Curves

Let's assume that the random variable T is a non-negative with a continuous and twice differentiable cumulative distribution function. The Bonferroni curve of the random variable T is defined as

$$B(p) = \frac{1}{\mu p}\int_0^q tf(t)dt$$

where, $p = F(t)$, $q = F^{-1}(p)$ and $\mu = E(t)$

$$B(p) = \frac{\Gamma\left(\alpha+\frac{k+1}{3}\right)}{p\left(\frac{\beta}{\alpha}\right)^{1/3}\Gamma\left(\alpha+\frac{k+2}{3}\right)} \int_0^q t \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}} t^{3\alpha+k} \exp\left\{-\frac{\alpha}{\beta} t^3\right\}}{\Gamma\left(\alpha+\frac{k+1}{3}\right)} dt \tag{22}$$

$$= \frac{3\left(\frac{\alpha}{\beta}\right)^{\alpha+\frac{k+1}{3}}}{p\left(\frac{\beta}{\alpha}\right)^{1/3}\Gamma\left(\alpha+\frac{k+2}{3}\right)} \int_0^q t^{3\alpha+k+1} \exp\left\{-\frac{\alpha}{\beta}\ t^3\right\} dt$$

Substituting, $y = \left(\frac{\alpha}{\beta}\right) t^3$ in equation (22), we get Bonferroni curve

$$B(p) = \frac{\gamma\left(\alpha+\frac{k+2}{3},\ \frac{\alpha}{\beta}q^3\right)}{p\,\Gamma\left(\alpha+\frac{k+2}{3}\right)} \tag{23}$$

Lorenz curve is defined as

$$L(p) = \frac{1}{\mu}\int_0^q tf(t)dt = pB(p)$$

Using equation (23), we get

$$L(p) = \frac{\gamma\left(\alpha+\frac{k+2}{3},\ \frac{\alpha}{\beta}q^3\right)}{\Gamma\left(\alpha+\frac{k+2}{3}\right)} \tag{24}$$

6. Random Number Generation and Quantiles

Random numbers of LBWWH can be easily generate by using the following function

$$t = \left[\left(\frac{\beta}{\alpha}\right) Q^{-1}\left(\alpha + \frac{k+1}{3},\ 1-U\right)\right]^{1/2}$$

where, $U \sim U(0,1)$ and $Q^{-1}(a,z)$ is inverse of regularized incomplete gamma function, where regularized incomplete gamma function is defined as $Q(a,z) = \frac{\Gamma(a,z)}{\Gamma(a)}$.
Quantiles are given by

$$t_q = \left[\left(\frac{\beta}{\alpha}\right) Q^{-1}\left(\alpha + \frac{k+1}{3},\ 1-q\right)\right]^{1/2} \tag{25}$$

By putting $q = 0.5$ in equation (25), we get the median of LBWWHD

$$t_{0.5} = \left[\left(\frac{\beta}{\alpha}\right) Q^{-1}\left(\alpha + \frac{k+1}{3},\ 0.5\right)\right]^{\frac{1}{2}}$$

7. Applications

In this section, we have considered two real data sets to check the suitability of the proposed distribution. Further, we have compared the distribution with the Length-Biased weighted Lindley distribution (LBWLD), Length-Biased Susheela distribution ($LBSD_1$) and length-Biased Suja distribution ($LBSD_2$), for suitability of proposed distribution. For this, we have used Akaike information criterion (AIC), Bayesian information criteria (BIC), Akaike Information Criterion Corrected (AICC) and Hannan-Quinn Information Criterion (HQIC), respectively. The AIC, BIC, AICC and HQIC are defined as:

$$\text{AIC} = 2K - 2\log L,\ \text{BIC} = K\log n - 2\log L,$$

$$\text{AICC} = \text{AIC} + \frac{2K(K+1)}{n-K-1}\text{, and } \text{HQIC} = 2K\log(\log(n)) - 2\log L.$$

where n is the sample size, K is the number of parameters, and L denotes the likelihood function. Any probability model having smaller value of AIC, BIC and $-LogL$ being the best model to fit the data set.

Dataset 1: The first dataset is taken from [7] which represent the tensile strength, measured in GPa, of 69 carbon fibers tested under tension at gauge lengths of 20mm.

Table 6: *MLE, AIC and BIC for gauge lengths of 20 mm*

Distribution	α	β	k	Log L	AIC	BIC	AICC	HQIC
LBWWHD	0.540	3.014	6.259	−48.895	103.790	110.492	104.148	106.449
LBWLD	--	--	--	−87.8984	179.796	184.265	179.973	181.569
LBSD1	--	--	--	−92.3037	188.607	193.075	188.78	190.380
LBSD2	--	--	--	−134.128	186.607	188.841	186.66	187.493

According to the results in Table 6, the LBWWHD has the smallest values of these statistics, followed by LBWLD, LBSD_1 and LBSD_2. Therefore, the suggested distribution is the best choice for the tensile strength data. The fitted $pdfs$ and empirical $cdfs$ plots of the four models are sketched in Figure 4. Therefore, we assert that the LBWWHD fitting successfully the empirical plots of the data set.

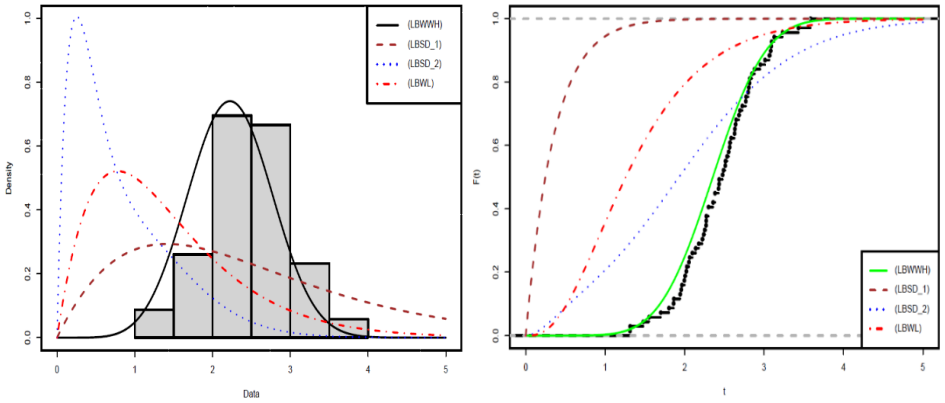


Figure 4: *Estimated densities and cdf plot of the models based on the real dataset 1.*

Dataset 2- The second dataset used by [7], which represent the tensile strength, measured in GPa, of 63 carbon fibers tested under tension at gauge lengths of 10mm.

Table 7: *MLE, AIC and BIC for gauge lengths of 10 mm*

Distribution	α	β	k	Log L	AIC	BIC	AICC	HQIC
LBWWHD	0.1587	1.730	7.402	−58.7320	123.464	129.893	123.857	125.992
LBWLD	--	--	--	−93.4265	190.85	195.133	191.046	192.538
LBSD_1	--	--	--	−97.9972	199.994	204.280	200.188	201.680
LBSD_2	--	--	--	−65.9256	133.851	135.995	133.914	134.694

For the dataset 2, we infer from the Table 7, the LBWWHD has the lowest values of AIC, BIC, AICC and HQIC followed by LBWLD, LBSD_1 and LBSD_2. Therefore, we conclude that LBWWHD is the most suitable choice for this dataset among the considered distributions. The fitted *pdfs* and empirical cdf plots of the four models are presented in Figure 5 and see that the LBWWHD fitting successfully the empirical plots of the data set.

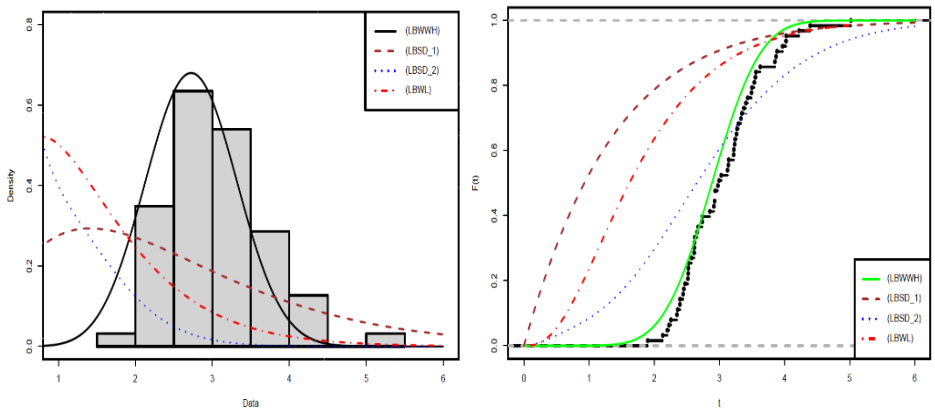


Figure 5: *Estimated densities and* cdf *plot of the models based on the real dataset 2.*

8. Conclusion

In the present article a new distribution known as the Length-biased Weighted Wilson Hilferty distribution, has been proposed. The distribution in discussion is characterized by three parameters called shape, scale and weighted parameter. Through the use of certain formulae, the properties and characteristics of this distribution such as its moments, failure rate, reliability function *etc.*, comprehensively examined and, the parameters estimation and stochastic comparison is also done. The examination and subsequent comparison of the criteria for AIC, BIC, AICC and HQIC have been conducted in relation to the Length-biased weighted Lindley distribution, Length-biased Sushila distribution and Length-biased Suja distribution. The actual lifetime of two sets of data has been successfully modelled and the resulting fit has been determined to be satisfactory.

Author Contributions: All authors have equal contribution. All authors reviewed the results and approved the final version of the manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors

Conflict of interest: The Authors declares that there is no conflict of interest.

References

[1] Wilson E. B., Hilferty M. M. (1931). The distribution of chi-square. *Proc Natl Acad. Sci USA.*;17:684-688.

[2] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13-25.

[3] Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment. *Sankhyā*: *The Indian Journal of Statistics, Series A*, 311-324.

[4] Patil, G. P. and Rao, C. R. The weighted distributions: a survey of their applications. In: applications of statistics, Ed. by P. R. Krishiaiah, 383-405. North Holland Publishing Company, Netherlands, 1977.

[5] Patil G. P and Rao, C. R. (1978): Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, 179-189.

[6] Stene, J. (1981). Probability distributions arising from the ascertainment and the analysis of data on human families and other groups. *Statistical distributions in scientific work, applications in physical, Social and life Sciences*, 6, 233-244.

[7] Bader, M. and Priest, A. (1982). Statistical Aspects of Fiber and Bundle Strength in Hybrid Composites. In: Hayashi, T., Kawata, S. and Umekawa, S., Eds., *Progress in Science and Engineering Composites, ICCM-IV*, Tokyo, 1129-1136.

[8] Gupta, R.C. and Keating, J.P. (1985). Relations for reliability measures under length biased sampling. *Scan. J. Statist.*, 13, 49-56.

[9] Khattree, R. (1989). Characterization of inverse Gaussian-and gamma distribution through their length-biased distributions. *IEEE Trans. Reliab.*, 38: 610-611.

[10] Gupta, R. C., and S. N. V. A. Kirmani (1990). The role of weighted distribution in stochastic modeling. *Commun. Stat. Theory Methods*, 19(9), 3147–3162.

[11] Shaked, M., Shanthikumar, J. G. Stochastic Orders and Their Applications. *Boston: Academic Press 1994*.

[12] Oluyede, B.O. and George, E.O. (2002). On Stochastic Inequalities and Comparisons of Reliability Measures for Weighted Distributions. *Mathematical problems in Engineering*, 8, 1-13.

[13] Gupta, R. D. and Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, 46(6), 621-634.

[14] Das, K. K., Roy, T. D. (2011a). Applicability of length-biased weighted generalized Rayleigh distribution. *Advances in Applied Science Research*, 2(4), 320-327.

[15] Das, K. K., Roy, T. D. (2011b). On some length-biased weighted Weibull distribution. *Advances in Applied Science Research*, 2(5), 465-475.

[16] Ghitany, M. E., Alqallaf, F., Al-Mutairi, H.A. (2011). A two parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in Simulation*, 81, 1190-1201.

[17] Ratnaparkhi, M. V. and Naik-Nimbalkar, U. V. (2012). The length-biased lognormal distribution and its application in the analysis of data from oil field exploration studies. *Journal of Modern Applied Statistical Methods*, 11, 225-260.

[18] Al-Kadim, K. A., Ali Hussein, N. (2014). New proposed length-biased weighted exponential and Rayleigh distribution with application. *Mathematical Theory and Modelling*, 4(7).

[19] Alqallaf, F., Ghitany, M. E. and Agostinelli, C. (2015). Weighted exponential distribution: Different methods of estimations. *Applied Mathematics & Information Sciences*, 9(3), 1167.

[20] Ahmad, A., Ahmad, S. P. and Ahmad, A. (2016). Length-biased weighted Lomax distribution: statistical properties and application. *Pakistan Journal of Statistics and Operation Research*, 12, 245-255.

[21] Oguntunde, P.E., Ilori, K. A. and Okagbue, H. I. (2018). The Inverted weighted exponential distribution with applications. *International Journal of Advanced and Applied Sciences*, 5(11), 46-50.

[22] Ramos PL, Almeida MP, Tomazella VL, Louzada F. (2019). Improved Bayes estimators and prediction for the Wilson-Hilferty distribution. *Anais da Academia Brasileira de Ciências, 91(3)*.

[23] Atikankul, Y., Thongteeraparp, A., Bodhisuwan, W., Volodin, A. (2020). The length-biased weighted Lindley distribution with applications, *Lobachevskii Journal of Mathematics*, 41, 308-319.

[24] Mathew, J. and George, S. (2020). Length biased exponential distribution as a reliability model: a Bayesian approach, Reliability: *Theory & Applications*, 3(58):84-91.

[25] Ilori, A. K., and Jolayemi, E. T. (2021). The weighted exponentiated inverted exponential distribution. *International Journal of Statistics and Applied Mathematics*, 6(1), 45-50.

[26] Almetwally, E. M. (2022). Marshall Olkin alpha power extended Weibull distribution: Different methods of estimation on type-I and type-II censoring. *Gazi University Journal of Science*, 35(1), 293-312.

[27] Bhat, V. A. and Pundir, S. (2023): Weighted intervened exponential distribution as a lifetime distribution, *Reliability: Theory & Applications*, 2(73), 253-266.

[28] Shanker, R. and Shukla, K. K. (2023). Power weighted Sujatha distribution with properties and application to survival times of patients of head and neck cancer data, *Reliability: Theory & Applications*, 3(74), 568-581.

A TWO NON-IDENTICAL UNIT STANDBY SYSTEM WITH CORRELATED PREVENTIVE MAINTENANCE TIME AND TIME TO PREVENTIVE MAINTENANCE AND INVERSE GAUSSIAN REPAIR TIME DISTRIBUTION

Anju Rani, *Rakesh Gupta, Pradeep Chaudhary

•

Department of Statistics,
Ch. Charan Singh University, Meerut-250004(India)
taliyan53anju@gmail.com; *smprgccsu@gmail.com; pc25jan@gmail.com

Abstract

The paper deals with the cost benefit analysis of a two non-identical unit cold standby system model with the implementation of preventive maintenance (PM) on the priority unit after it has operated for a random duration. The objective is to evaluate the economic viability and performance of such system. A single repairman is consistently available within the system, responsible for both PM and repair of each failed unit. The priority in repair is given to priority (p) unit over ordinary (o) unit. The failure time distribution of each unit is assumed to be exponential while the repair time distribution of both the unit is taken as inverse Gaussian. The PM time and time to PM of the priority unit are correlated having their joint distributions as bivariate exponential. By considering the regenerative point technique, various measures of system effectiveness are obtained.

Keywords: Transition probabilities, bivariate exponential distribution, regenerative point, reliability, MTSF, availability, busy period, net expected profit.

I. Introduction

The purpose of reliability engineering is to identify probable failures, implement appropriate actions to enhance reliability and identify the consequences of those failures. The manufacturers as well as consumer of a system always desire a high reliability. High reliability ensures that the system performs its intended function consistently and meets the expectations of its users over time. One way of improving a system's reliability is by incorporating additional or duplicate units into the system. This strategy is known as redundancy. Another crucial way is by providing regular repair and maintenance to the system when they are needed, ensuring its reliability and longevity. Maintenance strategies aim to prevent failures, detect potential issues, and rectify any existing problems to ensure the system operates optimally. Repair and maintenance strategies play a crucial role in improving system reliability, minimizing disruptions and reducing related expenses. These strategies focus on proactive and measures to keep the system in optimal working condition and address potential issues rather than simply responding to problems after they manifest.

Employing redundancies is one of the important aspects of enhancing the system's effectiveness and reliability. Redundant components or resources are intended to serve as backups or fail-safe mechanisms that are ready to take over the functions of primary components if they fail or experience issues. A significant number of authors including [1, 2, 5, 7, 8, 9] have analyzed the two non-identical units cold standby redundant system models due to their vital existence in ensuring uninterrupted operations and minimizing downtime in modern organizations and industries. These system models are particularly relevant in critical systems and industries where the stakes are high and failures can lead to severe consequences, such as aerospace and aviation, healthcare, telecommunications, power distributions, industrial control systems and other mission-critical applications to ensure high reliability and continuity of operations. In practice, planned maintenance activities performed on the system to improve its working capability, prevent potential failures and extend its overall lifespan is called preventive maintenance (PM). PM is a proactive maintenance strategy that involves scheduled inspections, adjustments and repair with the aim of keeping the system in optimal condition and preventing unexpected breakdowns. For example, PM for HVAC (Heating, Ventilation and Air Conditioning) systems. PM tasks may include inspecting electrical connections, calibrating controls, lubricating moving parts and changing filters on a regular basis. By performing these tasks according to a predetermined schedule, potential problems can be found and addressed before they escalate into serious issues, which guarantees the HVAC system will operate effectively and reliably. A number of authors, including [2, 4, 5, 10] have explored the concept of preventive maintenance (PM) i.e. after operating for an arbitrary amount of time, a unit goes for its preventive maintenance. In most of the studies and models related to maintenance and reliability analysis, it's commonly assumed that the working time and PM time of a unit are uncorrelated random variables. However in reality, there is some sort of positive correlation between the failure time and preventive maintenance time of a unit. The concept of correlation between failures times and repair times has been analyzed by various authors including [1, 2, 3, 4, 6].

This paper explores the concept of correlation between time to PM and PM time. The purpose of the present paper is to investigate a two non-identical unit cold standby system model with correlated PM time and time to PM of priority unit having their joint distribution as bivariate exponential. It is also assumed that a single repairman is consistently available with the system for both for PM and repair of each failed unit. Here are some economic related measures of system effectiveness that can be obtained using regenerative point techniques:

- Transition probabilities and sojourn times in various states.
- Reliability analysis and mean time to system failure (MTSF).
- Availability analysis of the system during (0, t).
- Expected busy period of repairman during time interval (0, t) that the repairman is busy in PM and in the repair of p-unit and o-unit.
- Net expected profit earned by the system in the time interval (0, t)

Graphical representations depicting the MTSF and Profit function with respect to different parameters have also been made.

II. System Description and Assumptions

The following are some assumptions about the system model under study:

- The system comprises of two non-identical units. One unit is designated as priority (p) unit while the other is referred as non-priority or ordinary (o) unit.
- Each unit of the system has two possible modes- Normal (N) and Total failure (F).
- Only p-unit is scheduled for preventive maintenance (PM) after working for its random period of time.

- A single repairman is consistently available with the system for PM and repair of a failed unit. The priority in repair and PM is given to p-unit.
- The switching device is used to switch on the standby unit into operation promptly and seamlessly only when the operative unit fails completely. The switching device is assumed to be perfect, independent and instantaneous.
- The failure time distribution of each unit is taken as exponential while the repair time distribution is taken as inverse Gaussian. The time to PM (X) and PM time (Y) are correlated random variables having their joint distribution as bivariate exponential with the density as follows:-

$$f(x,y)=\lambda\mu(1-r)e^{-\lambda x-\mu y}I_0\left(2\sqrt{\lambda\mu rxy}\right);x,y,\lambda,\mu>0;\ 0\le r<1$$

where,

$$I_0\left(2\sqrt{\lambda\mu rxy}\right)=\sum_{j=0}^{\infty}\frac{(\lambda\mu rxy)^j}{(j!)^2}$$

- Each repaired unit ideally functions as good as new.

III. Notations and States of the System

I. Notations:

E : Set of regenerative states $\equiv\{S_0,S_1,S_2,S_3\}$.

E : Set of non-regenerative states $\equiv\{S_4,S_5\}$.

α_1,α_2 : Constant failure rate of p-unit and o-unit respectively.

$G_i(.)/g_i(.)$: c.d.f./ p.d.f. of time to repair of failed p-unit and o-unit respectively i.e.

$$g_i(t)=\frac{1}{\sqrt{2\pi}}t^{-3/2}\exp\left\{\frac{(t-\beta_i)^2}{2\beta_i^2 t}\right\}dt\quad ;t>0,\beta_i>0;\{i=1,2\}$$

X : Time to PM of an operating unit when other unit is in standby state.

Y : Time taken in PM of a unit.

f(x,y) : Joint p.d.f. of (X,Y).

$$f(x,y)=\lambda\mu(1-r)e^{-\lambda x-\mu y}I_0\left(2\sqrt{\lambda\mu rxy}\right)\quad ;x,y,\lambda,\mu>0;\ 0\le r<1$$

where,

$$I_0\left(2\sqrt{\lambda\mu rxy}\right)=\sum_{j=0}^{\infty}\frac{(\lambda\mu rxy)^j}{(j!)^2}$$

k(y|x) : Conditional p.d.f. of Y given X=x.

$$=\mu e^{-\lambda rx-\mu y}I_0\left(2\sqrt{\lambda\mu rxy}\right)\quad ;x,y,\lambda,\mu>0;\ 0\le r<1$$

K(y|x) : Conditional c.d.f. of Y given X=x.

g(x) : Marginal p.d.f. of X i.e.

$$=\lambda(1-r)\exp\{-\lambda(1-r)x\}$$

II. Symbols for the states of the system

N_0^1, N_s^2 : Unit-1/Unit-2 in Normal (N) mode and operative/ standby state.

N_{pm}^1 : Unit-1 in normal mode and under preventive maintenance.

F_r^1, F_w^2 : Unit-1/Unit-2 in failure (F) mode and under repair/waiting for repair.

F_r^2 : Unit-2 in failure (F) mode and under repair.

By considering these symbols according to assumptions stated earlier, we have the following states of the system:

Up states : $S_0 \equiv (N_o^1, N_s^2)$ $S_1 \equiv (N_{pm}^1, N_o^2)$, $S_2 \equiv (F_r^1, N_o^2)$, $S_3 \equiv (N_o^1, F_r^2)$

Down states : $S_5 \equiv (N_{pm}^1, F_w^2)$

Failed states : $S_4 \equiv (F_r^1, F_w^2)$

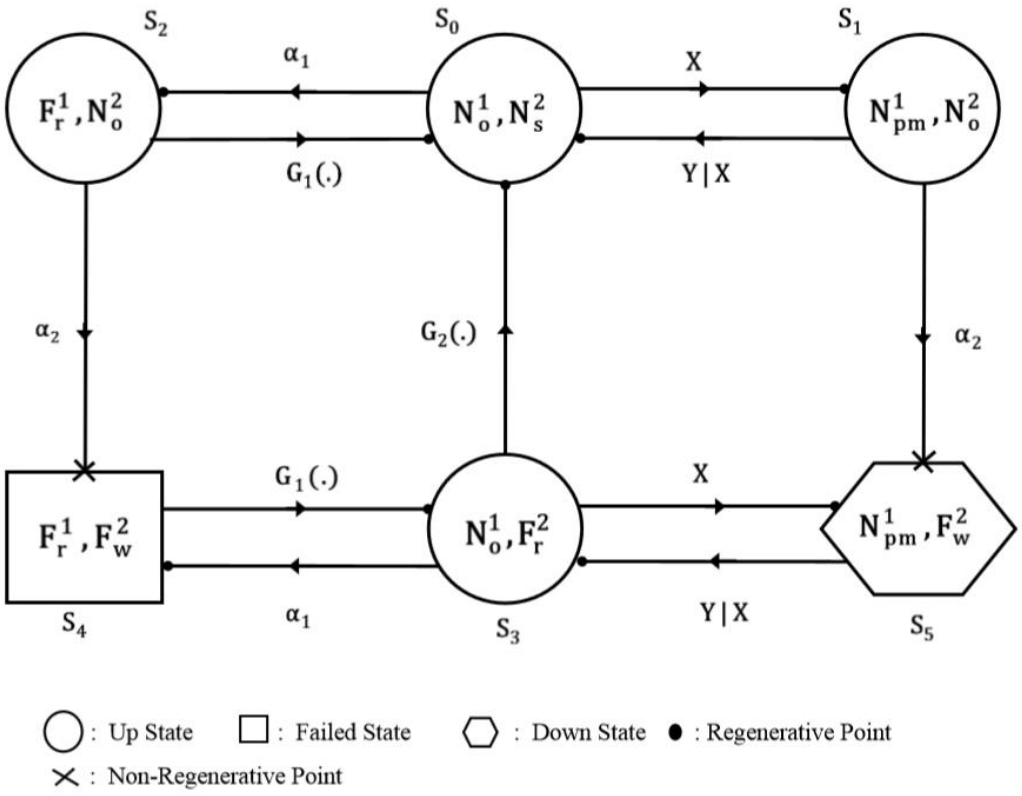


Figure 1: *Transition diagram*

The transition diagram depicting the system model along with failure rates/repair time c.d.f's is shown in Figure 1. From the transition diagram we find that the epochs of transitions into the states S_5 from S_1 and S_4 from S_2 are non-regenerative while all other entrance epochs are regenerative.

IV. Transition Probabilities and Sojourn Times

By using simple probabilistic arguments, the conditional and unconditional transition probabilities are given as:

$$p_{01} = \frac{\lambda(1-r)}{\alpha_1 + \lambda(1-r)} \qquad p_{02} = \frac{\alpha_1}{\alpha_1 + \lambda(1-r)}$$

$$p_{10|x} = \mu' \exp\left\{-\lambda\left(1-\mu'\right)rx\right\} \text{ ; where } \mu' = \frac{\mu}{\mu+\alpha_2} \qquad p_{15} = p_{13|x}^{(5)} = 1-\mu' \exp\left\{-\lambda\left(1-\mu'\right)rx\right\}$$

$$p_{20} = \exp\left[\left\{1-\sqrt{1+2\beta_1^2\alpha_2}\right\}\Big/\beta_1\right] \qquad p_{24} = p_{23}^{(4)} = 1-\exp\left[\left\{1-\sqrt{1+2\beta_1^2\alpha_2}\right\}\Big/\beta_1\right]$$

$$p_{30} = \exp\left[\left\{1 - \sqrt{1 + 2\beta_2^2\{\alpha_1 + \lambda(1-r)\}}\right\}\Big/\beta_2\right]$$

$$p_{34} = \frac{\alpha_1}{\alpha_1 + \lambda(1-r)}\left(1 - \exp\left[\left\{1 - \sqrt{1 + 2\beta_2^2\{\alpha_1 + \lambda(1-r)\}}\right\}\Big/\beta_2\right]\right)$$

$$p_{35} = \frac{\lambda(1-r)}{\alpha_1 + \lambda(1-r)}\left(1 - \exp\left[\left\{1 - \sqrt{1 + 2\beta_2^2\{\alpha_1 + \lambda(1-r)\}}\right\}\Big/\beta_2\right]\right)$$

$$p_{43} = \int dG_1(t) = 1 \qquad p_{53|x} = \int dK(t \mid x) = 1 \qquad (1\text{-}11)$$

It can be easily verified that

$$p_{01} + p_{02} = 1 \qquad p_{10|x} + p_{13|x}^{(5)} = 1$$

$$p_{20} + p_{23}^{(4)} = 1 \qquad p_{30} + p_{34} + p_{35} = 1$$

$$p_{43} = 1 \qquad p_{53|x} = 1$$

Unconditional transitional probabilities are as follows-

$$p_{10} = \frac{\mu'(1-r)}{(1-r\mu')} \qquad p_{13}^{(5)} = 1 - \frac{\mu'(1-r)}{(1-r\mu')}$$

$$p_{53} = 1 \qquad (12\text{-}14)$$

Thus, we observe the following relations-

$$p_{01} + p_{02} = 1 \qquad p_{10} + p_{13}^{(5)} = 1$$

$$p_{20} + p_{23}^{(4)} = 1 \qquad p_{30} + p_{34} + p_{35} = 1$$

$$p_{43} = 1 \qquad p_{53} = 1 \qquad (15\text{-}20)$$

Let T_i be the sojourn time in state $S_i \in E$, then the mean sojourn time in state S_i is given by

$$\Theta_i = \int P(T_i > t)dt$$

Therefore

$$\Theta_0 = \frac{1}{\alpha_1 + \lambda(1-r)} \qquad \Theta_{1|x} = \frac{1}{\alpha_2}\left[1 - \mu'\exp\left\{-\lambda\left(1-\mu'\right)rx\right\}\right]$$

$$\Theta_1 = \frac{\left(1-\mu'\right)}{\alpha_2\left(1-r\mu'\right)} \qquad \Theta_2 = \frac{1}{\alpha_2}\left(1 - \exp\left[\left\{1 - \sqrt{1 + 2\beta_1^2\alpha_2}\right\}\Big/\beta_1\right]\right)$$

$$\Theta_3 = \frac{1}{\alpha_1 + \lambda(1-r)}\left(1 - \exp\left[\left\{1 - \sqrt{1 + 2\beta_2^2\{\alpha_1 + \lambda(1-r)\}}\right\}\Big/\beta_2\right]\right)$$

$\Theta_4 = \beta_1$ = mean repair time of p-unit.

$$\Theta_{5|x} = \frac{1 + \lambda rx}{\mu} \qquad \Theta_5 = \frac{1}{\mu(1-r)} \qquad (21\text{-}28)$$

V. Analysis of results

I. Reliability and MTSF

Let the random variable T_i be the time to system failure (TSF), when at time t=0, the system starts its operation from state $S_i \in E$. Then, the reliability of the system is given by

$$R_i(t) = \int P(T_i > t)dt$$

To determine $R_i(t)$, we regard the failed state (S_4) of the system as an absorbing state. By employing simple probabilistic arguments, we observe the following relations:

$$R_0(t) = Z_0(t) + q_{01}(t) \odot R_1(t) + q_{02}(t) \odot R_2(t)$$
$$R_1(t) = Z_1(t) + q_{15}(t) \odot Z_5(t) + q_{10}(t) \odot R_0(t) + q_{13}^{(5)}(t) \odot R_3(t)$$
$$R_2(t) = Z_2(t) + q_{20}(t) \odot R_0(t)$$
$$R_3(t) = Z_3(t) + q_{30}(t) \odot R_0(t) + q_{35}(t) \odot R_5(t)$$
$$R_5(t) = Z_5(t) + q_{53}(t) \odot R_3(t) \qquad (29\text{-}32)$$

Where

$$Z_0(t) = \exp\left[-\{\alpha_1 + \lambda(1-r)\}t\right] \qquad Z_1(t) = \int \exp(-\alpha_2 t)\,\bar{K}(t \mid x)g(x)dx$$
$$Z_2(t) = \exp(-\alpha_2 t)\,\bar{G}_1(t) \qquad Z_3(t) = \exp\left[-\{\alpha_1 + \lambda(1-r)\}t\right]\bar{G}_2(t)$$
$$Z_5(t) = \int \bar{K}(t \mid x)g(x)dx \qquad (33\text{-}37)$$

Taking the Laplace transform of the relations (29-32) and simplifying the resulting set of equations for $R_0^*(s)$ we obtain;

$$R_0^*(s) = \frac{N_1(s)}{D_1(s)} \text{ (say)}$$
$$= \frac{\left(Z_0^* + q_{01}^* Z_1^* + q_{02}^* Z_2^* + q_{01}^* q_{15}^* Z_5^*\right)\left(1 - q_{35}^* q_{53}^*\right) + \left(Z_3^* + q_{35}^* Z_5^*\right) q_{01}^* q_{13}^{(5)*}}{\left(1 - q_{01}^* q_{10}^* - q_{02}^* q_{20}^*\right)\left(1 - q_{35}^* q_{53}^*\right) - q_{01}^* q_{13}^{(5)*} q_{30}^*} \qquad (38)$$

The mean time to system failure (MTSF) can be determined by using the formula;

$$E_0(t) = \int R_0(t)dt = \lim_{s \to 0} R_0^*(s) \qquad (39)$$
$$= \frac{\left(\Theta_0 + p_{01}\Theta_1 + p_{02}\Theta_2 + p_{01}p_{15}\Theta_5\right)\left(1 - p_{35}\right) + \left(\Theta_3 + p_{35}\Theta_5\right) p_{01} p_{13}^{(5)}}{\left(1 - p_{01}p_{10} - p_{02}p_{20}\right)\left(1 - p_{35}\right) - p_{01}p_{13}^{(5)}p_{30}} \qquad (40)$$

II. Availability Analysis

Let $A_i^p(t)$ and $A_i^o(t)$ be the probabilities that the system is up at epoch 't' due to p-unit and o-unit respectively, when the system initially starts from state $S_i \in E$. By using simple probabilistic laws we get the following relation among $A_i^p(t)$.

$$A_0^p(t) = Z_0(t) + q_{01}(t) \odot A_1^p(t) + q_{02}(t) \odot A_2^p(t)$$
$$A_1^p(t) = q_{10}(t) \odot A_0^p(t) + q_{13}^{(5)}(t) \odot A_3^p(t)$$
$$A_2^p(t) = q_{20}(t) \odot A_0^p(t) + q_{23}^{(4)}(t) \odot A_3^p(t)$$
$$A_3^p(t) = Z_3(t) + q_{30}(t) \odot A_0^p(t) + q_{34}(t) \odot A_4^p(t) + q_{35}(t) \odot A_5^p(t)$$
$$A_4^p(t) = q_{43}(t) \odot A_3^p(t)$$
$$A_5^p(t) = q_{53}(t) \odot A_3^p(t) \qquad (41\text{-}46)$$

Where, $Z_0(t)$ and $Z_3(t)$ has already been defined in equations (33) and (36).

Taking the Laplace transform of the relations (41-46) and simplifying the resulting set of equations for $A_0^{p*}(s)$ we obtain;

$$A_0^{p*}(s) = \frac{Z_0^*\left(1 - q_{34}^* q_{43}^* - q_{35}^* q_{53}^*\right) + Z_3^*\left(q_{01}^* q_{13}^{(5)*} + q_{02}^* q_{23}^{(4)*}\right)}{\left(1 - q_{34}^* q_{43}^* - q_{35}^* q_{53}^*\right)\left(1 - q_{01}^* q_{10}^* - q_{02}^* q_{20}^*\right) - q_{30}^*\left(q_{01}^* q_{13}^{(5)*} + q_{02}^* q_{23}^{(4)*}\right)} \qquad (47)$$

Similarly, employing the same probabilistic reasoning as in case of $A_i^o(t), (i = 0\text{-}5)$ the recurrence relations among can be determined as follows:-

$$A_0^o(t) = q_{01}(t) \copyright A_1^o(t) + q_{02}(t) \copyright A_2^o(t)$$

$$A_1^o(t) = Z_1(t) + q_{10}(t) \copyright A_0^o(t) + q_{13}^{(5)}(t) \copyright A_3^o(t)$$

$$A_2^o(t) = Z_2(t) + q_{20}(t) \copyright A_0^o(t) + q_{23}^{(4)}(t) \copyright A_3^o(t)$$

$$A_3^o(t) = q_{30}(t) \copyright A_0^o(t) + q_{34}(t) \copyright A_4^o(t) + q_{35}(t) \copyright A_5^o(t)$$

$$A_4^o(t) = q_{43}(t) \copyright A_3^o(t)$$

$$A_5^o(t) = q_{53}(t) \copyright A_3^o(t) \qquad (48\text{-}53)$$

Taking the L.T. of the relations (48-53) and simplifying the resulting sets of algebraic equations for $A_0^{o*}(s)$, we obtain;

$$A_0^{o*}(s) = \frac{\left(1 - q_{34}^* q_{43}^* - q_{35}^* q_{53}^*\right)\left(q_{01}^* Z_0^* + q_{02}^* Z_2^*\right)}{\left(1 - q_{34}^* q_{43}^* - q_{35}^* q_{53}^*\right)\left(1 - q_{01}^* q_{10}^* - q_{02}^* q_{20}^*\right) - q_{30}^*\left(q_{01}^* q_{13}^{(5)*} + q_{02}^* q_{23}^{(4)*}\right)} \qquad (54)$$

For brevity, we have omitted the argument's' from $q_{ij}^*(s)$ and $Z_i^*(s)$. Now the steady state availabilities of the system when p-unit and o-unit are operative, respectively given by;

$$A_0^{p*} = \lim_{t\to\infty} A_0^p(t) = \lim_{s\to 0} sA_0^{p*}(s) = N_2/D_2 \qquad (55)$$

$$A_0^{o*} = \lim_{t\to\infty} A_0^o(t) = \lim_{s\to 0} sA_0^{o*}(s) = N_3/D_2 \qquad (56)$$

Where

$$N_2 = p_{30}\Theta_0 + \left\{p_{01}p_{13}^{(5)} + p_{02}p_{23}^{(4)}\right\}\Theta_3 \qquad (57)$$

$$N_3 = p_{30}\left(p_{01}\Theta_1 + p_{02}\Theta_2\right) \qquad (58)$$

We observe that

$$D_2(0) = 0$$

Therefore by using L. Hospital rule, we get

$$A_0^p = \lim_{s\to 0}\frac{N_2(s)}{D_2'(s)} = \frac{N_2}{D_2'} \quad \text{(say)}$$

$$A_0^o = \lim_{s\to 0}\frac{N_3(s)}{D_2'(s)} = \frac{N_3}{D_2'} \quad \text{(say)}$$

Thus, we have

$$D_2' = p_{30}\Theta_0 + (1 - p_{01}p_{10} - p_{02}p_{20})\Theta_3 + \left\{p_{02}p_{30} + (1 - p_{01}p_{10} - p_{02}p_{20})p_{34}\right\}\Theta_4 + \left\{p_{01}p_{30} + (1 - p_{01}p_{10} - p_{02}p_{20})p_{35}\right\}\Theta_5 \qquad (59)$$

The mean up time of the system due to p-unit and o-unit during time interval (0, t) are respectively given by;

$$\mu_{up}^p(t) = \int_0^t A_0^p(u)du \quad \text{and} \quad \mu_{up}^o(t) = \int_0^t A_0^o(u)du \qquad (60\text{-}61)$$

Thus

$$\mu_{up}^{p*}(s) = \frac{A_0^{p*}(s)}{s} \quad \text{and} \quad \mu_{up}^{o*}(s) = \frac{A_0^{o*}(s)}{s} \qquad (63\text{-}63)$$

III. Busy period analysis

Let $B_i^1(t), B_i^2(t)$ and $B_i^3(t)$ be the respective probabilities that the repairman is busy in PM, in the repair of a failed p-unit and in the repair of failed o-unit at time 't', when system initially starts from state $S_i \in E$. Using simple probabilistic arguments the system of integral equations for $B_i^1(t), B_i^2(t)$ and $B_i^3(t)$

can be easily developed and by the technique of L.T. the values of $B_0^{1*}(s), B_0^{2*}(s)$ and $B_0^{3*}(s)$ can be easily determined.

The steady state probabilities B_0^1, B_0^2 and B_0^3 are given respectively as follows:

$$B_0^1 = N_4/D_2', \qquad B_0^2 = N_5/D_2' \qquad \text{and} \qquad B_0^3 = N_6/D_2' \tag{64-66}$$

Where

$$N_4 = p_{01}p_{30}(\Theta_1 + p_{15}\Theta_5) + \left\{p_{01}p_{13}^{(5)} + p_{02}p_{23}^{(4)}\right\}p_{35}\Theta_5$$

$$N_5 = p_{02}p_{30}(\Theta_2 + p_{24}\Theta_4) + \left\{p_{01}p_{13}^{(5)} + p_{02}p_{23}^{(4)}\right\}p_{34}\Theta_4$$

$$N_6 = \left\{p_{01}p_{13}^{(5)} + p_{02}p_{23}^{(4)}\right\}\Theta_3 \tag{67-69}$$

And D_2' is same as in case of availability analysis.

The expected busy periods of the repairman in PM, in repair of failed p-unit and in the repair of failed o-unit respectively, during time interval (0,t) are given by-

$$\mu_b^1(t) = \int_0^t B_0^1(u)du, \qquad \mu_b^2(t) = \int_0^t B_0^2(u)du \qquad \text{and} \qquad \mu_b^3(t) = \int_0^t B_0^3(u)du \tag{70-72}$$

So that

$$\mu_b^{1*}(s) = \frac{B_0^{1*}(s)}{s}, \qquad \mu_b^{2*}(s) = \frac{B_0^{2*}(s)}{s} \qquad \text{and} \qquad \mu_b^{3*}(s) = \frac{B_0^{3*}(s)}{s} \tag{73-75}$$

IV. Profit Function Analysis

The net expected gain incurred in time interval (0,t) is given by-

$$P(t) = K_0\mu_{up}^p(t) + K_1\mu_{up}^o(t) - K_2\mu_b^1(t) - K_3\mu_b^2(t) - K_4\mu_b^3(t) \tag{76}$$

Where

K_0 = revenue per unit time when system is operative due to p-unit.

K_1 = revenue per unit time when system is operative due to o-unit.

K_2 = cost per unit time for PM of p-unit.

K_3 = cost per unit time for repair of failed p-unit.

K_4 = cost per unit time for repair of failed o-unit.

Now the expected profit (gain) per-unit time in steady state is given by-

$$P = K_0A_0^p + K_1A_0^o - K_2B_0^1 - K_3B_0^2 - K_4B_0^3 \tag{77}$$

VI. Graphical Representation

In order to carry out a detailed analysis of the behavior of the system, we plot the MTSF and Profit curves with respect to multiple values of the failure rate (α_1), three distinct values of mean repair time of o-unit (β_2) and two distinct correlation coefficient (r) values.

The MTSF curves w.r.t. "α_1" are displayed in Figure 2 with three distinct values of mean repair time of o-unit (β_2), i.e., 0.25, 0.55, and 0.85, as well as two distinct values of correlation coefficient (r), i.e., 0.2 and 0.7. The other parameters remain constant at $\beta_1 = 0.9$, $\lambda = 0.7$, $\mu = 0.18$, and $\alpha_2 = 0.05$. From the observations provided in the figure, we observe that MTSF decreases uniformly as the value of failure rate 'α_1' increases. Furthermore, the observation indicates that as the values of the mean repair time of o-unit 'β_2' increase, the expected life of the system decreases. Moreover, with the increase in the value of the correlation coefficient 'r', MTSF tends to increase as well.

From Figure 3, we observe that the profit decreases as failure rate 'α_1' increases with varying three different values of 'β_2' i.e., 0.25, 0.55 and 0.85 and two different values of correlation coefficient

'r' i.e., 0.01 and 0.02, when values of other parameters are kept fixed as $\beta_1 = 0.5$, $\lambda=0.25$, $\mu=0.4$, $\alpha_2 = 0.1$, $K_0 = 60$, $K_1 = 190$, $K_2 = 200$, $K_3 = 300$ and $K_4 = 250$. From the curves, the linear trends in Figure 3 indicate that there is a constant rate of decrease in profit as the values of the failure rate 'α_1' increases.

From Figure 2, the dotted curves depict that to achieve MTSF at least 3000 units, the failure rate 'α_1' of unit-1 must be less than 0.012, 0.016 and 0.021, respectively, for $\beta_2 = 0.25$, 0.55 and 0.85 when r = 0.2. From smooth curves, we observe that to achieve MTSF at least 3500 units, the values of 'α_1' must be less than 0.011, 0.019 and 0.023, respectively, for $\beta_2 = 0.25$, 0.55 and 0.85 when r = 0.7.

From Figure 3, the dotted curves reveal that the system is profitable only if 'α_1' is less than 0.10, 0.18 and 0.28, respectively, for $\beta_2 = 0.25$, 0.55 and 0.85 when r = 0.01. From smooth curves, we conclude that the system is profitable only if 'α_1' is less than 0.11, 0.19 and 0.29, respectively, for $\beta_2 = 0.25$, 0.55 and 0.85 when r = 0.02.

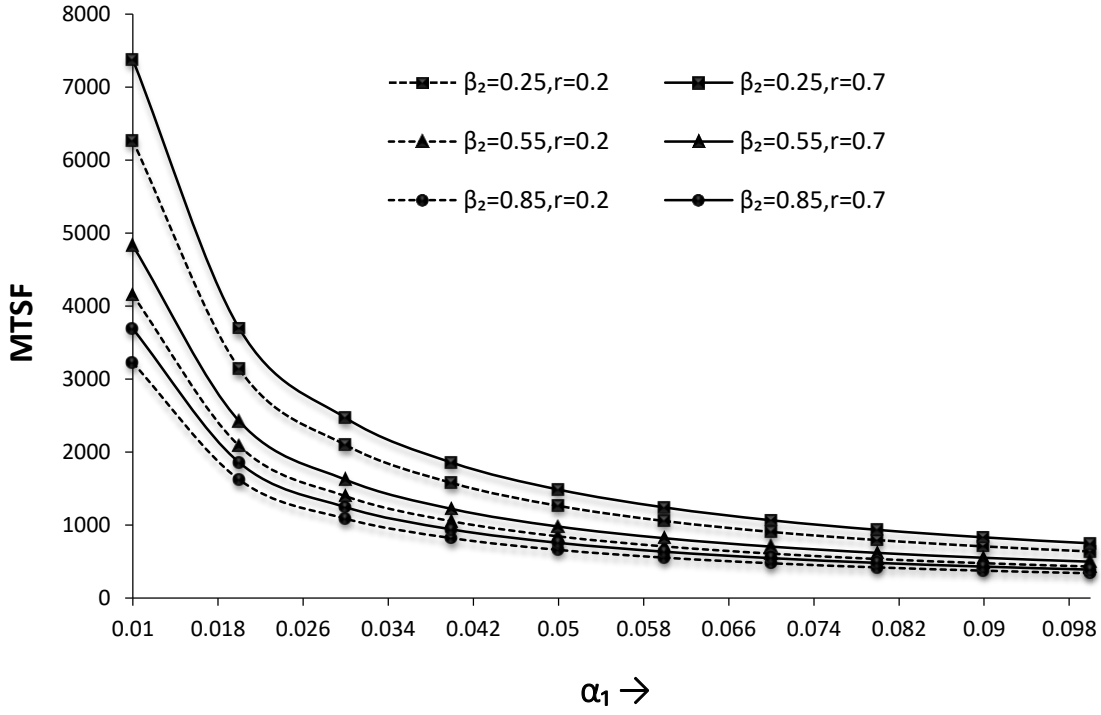


Figure 2: *Behaviour of MTSF with respect to α_1 for different values of β_2 and r*

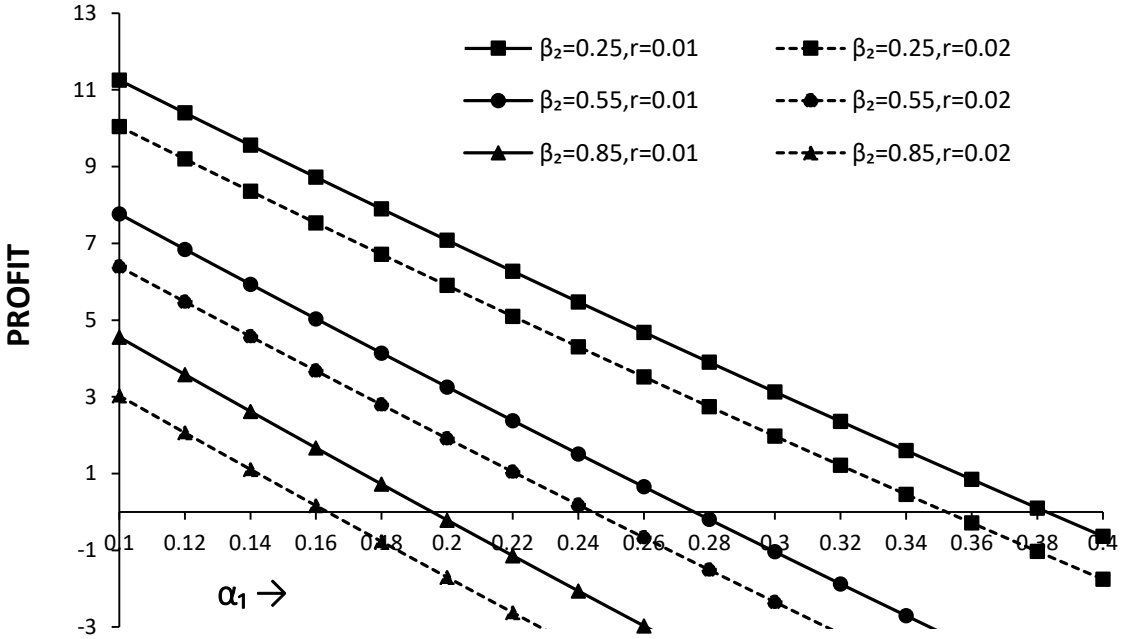


Figure 3: *Behaviour of Profit (P) with respect to α_1 for different values of β_2 and r.*

ACKNOWLEDGEMENT

One of the authors Dr. Pradeep Chaudhary is thankful to Chaudhary Charan Singh University, Meerut for awarding a minor research project to him vide letter no. DEV/URGS/2022-23/26 Dated-22/07/2022.

References

[1] Chaudhary, A., Sharma, S., & Sharma, A. (2023). A two non identical units cold standby system with correlated failure time and repair machine failure. *Reliability: Theory & Applications*, *18*(4 (76)):252-262.

[2] Goel, L. R., and Shrivastava, P. (1992). A two-unit standby system with imperfect switch, preventive maintenance and correlated failures and repairs. *Microelectronics Reliability*, *32*(12):1687-1691.

[3] Gupta, R. (2018). A two non-identical unit parallel system with correlated failure and repair times of repair machine. *International Journal of Agricultural & Statistical Sciences*, *14*(2).

[4] Gupta, R., and Sharma, V. (2010). A two unit standby system with preventive maintenance and inverse Gaussian repair time distribution. *Journal of Informatics and Mathematics Sciences*, *2*(2-3):183-191.

[5] Gupta, R., Kishan, R., & Kumar, D. (2012). Cost Benefit Analysis of a Two Non-Identical Unit Standby System with Preventive Maintenance. *Journal of Combinatorics, Information & System Sciences*, *37*(1):21.

[6] Gupta, R., Mahi, M., and Sharma, V. (2008). A two component two unit standby system with correlated failure and repair times. *Journal of Statistics and Management Systems*, *11*(1):77-90.

[7] Kaur, G., and Vinodiya, P. (2017). Reliability analysis of a two-non-identical units cold standby repairable system with switching of units by using linear first order differential equations. *Journal for innovative research in multidisciplinary field*, 3:92-98.

[8] Pundir, P.S., Patawa, R., and Gupta, P.K. (2021). Analysis of two non-identical unit cold standby system in presence of prior information. *American Journal of Mathematical and Management Sciences,* 40(4):320-335.

[9] Raghuvanshi, L., Gupta, R., and Chaudhary, P. (2021). A two non-identical unit standby system with helping unit of the priority unit. *International Journal of Agriculture & Statistical Sciences,* 3:92-98.

[10] Vilarinho, S., Lopes, I., and Oliveira, J.A. (2017). Preventive Maintenance decisions through maintenance optimization models: a case study. *Procedia Manufacturing,* 11:1170-1177.

RT&A, No 2 (78)
Volume 19, June, 2024

ON ESTIMATION AND PREDICTION FOR THE XLINDLEY DISTRIBUTION BASED ON RECORD DATA

F. Zanjiran[1], S.M.T.K. MirMostafaee[2,*]

•

[1,2]Department of Statistics, University of Mazandaran, Iran
[1]zanjiran@irstat.ir, [2]m.mirmostafaee@mail.umz.ac.ir

Abstract

This paper investigates the estimation of the unknown parameter in the XLindley distribution using record values and inter-record times, both in classical and Bayesian frameworks. It also delves into Bayesian prediction of a future record value. We also study the problem of estimation and prediction for the XLindley distribution based on lower records alone. A simulation study, as well as an analysis of a real data example, are conducted for comparison and illustration. The numerical findings underline that including the inter-record times in the study may enhance the performance of the estimators and predictors.

Keywords: XLindley distribution, lower record values, inter-record times, Bayesian estimation and prediction.

1. Introduction

The XLindley distribution was first proposed by [8] as an effective new distribution in modeling lifetime data. Suppose that X is a random variable following the one-parameter XLindley distribution. The probability density function (PDF) and cumulative distribution function (CDF) of X are given by

$$f(x;\theta) = \frac{\theta^2}{(1+\theta)^2}(2+\theta+x)\mathrm{e}^{-\theta x}, \tag{1}$$

$$F(x;\theta) = 1 - \left(1 + \frac{\theta x}{(1+\theta)^2}\right)\mathrm{e}^{-\theta x}. \tag{2}$$

respectively.

We write $X \sim XL(\theta)$ if the PDF of X is given by (1). The XLindley distribution enjoys an increasing hazard rate function. Chouia and Zeghdoudi [8] demonstrated that the XLindley distribution can fit better than some other one-parameter distributions such as the exponential, xgamma and Lindley distributions. Due to the flexibility of the XLindley model, several inferential researches have been accomplished by authors since its inception, for example, Alotaibi et al. [2] addressed the estimation problem for the XLindley distribution using an adaptive Type-II progressively hybrid censored data, Nassar et al. [31] investigated the reliability estimation of the XLindley constant-stress partially accelerated life tests using progressively censored samples and Alotaibi et al. [3] worked on the reliability estimation under normal operating conditions

*Corresponding Author

for progressively Type-II XLindley censored data. Moreover, Metiri et al. [29] focused on the characterization of XLindley distribution using the relation between the truncated moment and failure rate function or reverse failure rate function.

Suppose that $\{X_n, n = 1, 2, \cdots\}$ is a sequence of identical and independent random variables. Let $\{X_n, n = 1, 2, \cdots\}$ be a sequence of identically distributed and independent random variables. If an observation X_j is less than all its preceding observations, then it is termed a lower record value. Similarly, upper record values can be defined based on the comparisons with preceding observations in the sequence. The sequence of lower record values along with the inter-record times can be denoted by $(\boldsymbol{R}, \boldsymbol{T}) = \{R_1, T_1, R_2, T_2, \cdots, R_{m-1}, T_{m-1}, R_m\}$ where R_i represents the i-th record value and T_i is the i-th inter-record time, which is the number of observations needed after occurrence of R_i to obtain a new record value R_{i+1}. Record data play a crucial role in various practical scenarios, see for example [5]. Record values and the related subjects have been studied by many authors; see for example [1, 11, 12, 28]. For instance, Samaniego and Whitaker [38] explored the estimation problem of the mean parameter of the exponential distribution using records and inter-record times. Doostparast [9] delved into the Bayesian and non-Bayesian estimation of the two parameters of the exponential distribution based on records and inter-record times. In a similar study, Doostparast et al. [10] investigated the Bayesian estimation of the parameters of the Pareto distribution utilizing records and inter-record times. Kzlaslan and Nadar [21] estimated the parameter of the proportional reversed hazard rate model based on records and inter-record times. Nadar and Kzlaslan [30] discussed inferential methods for the Burr type XII distribution using record values and inter-record times. Additionally, Kzlaslan and Nadar [22, 23] centered their research on inferential procedures for the generalized exponential and Kumaraswamy distributions based on record values and inter-record time statistics, respectively. Amini and MirMostafaee [4] examined interval prediction of future order statistics from the exponential distribution based on records given the inter-record times. Pak and Dey [32] developed inferential procedures for the estimation of parameters and prediction of future record values for the power Lindley model using lower record values and inter-record times. Kumar et al. [24] directed their attention towards the estimation and prediction for the unit-Gompertz distribution based on records and inter-record times. Bastan and MirMostafaee [6] explored inferential problems for the Poisson-exponential distribution based on record values and inter-record times. Khoshkhoo Amiri and MirMostafaee [19] studied estimation and prediction issues for the xgamma distribution based on lower records and inter-record times. Most recently, Khoshkhoo Amiri and MirMostafaee [20] addressed the estimation and prediction problems for the Chen distribution, utilizing lower records and inter-record times.

In this paper, we intend to discuss estimation and prediction for the XLindley distribution based on lower records and inter-record times, as well as based on lower records alone. In what follows, first, we obtain maximum likelihood (ML) estimates and asymptotic confidence intervals (ACIs) for the parameter of the XLindley distribution in Section 2. In Section 3, we go through the Bayesian estimation method and find the Bayes estimates of the parameter under a symmetric loss function and an asymmetric loss function. The Bayes estimates do not seem to be expressible in closed forms, so we become inclined to use an approximation method such as the Metropolis-Hastings algorithm. Section 4 is devoted to the Bayesian prediction of a lower future record value. A simulation study and a real data example are given in Section 5. The numerical outcomes highlight the effect of incorporating inter-record times in the study on the performance of estimators and predictors. The paper is concluded with several remarks in Section 6.

2. Maximum Likelihood Estimation

In this section, we proceed to obtain the ML estimates, as well as ACIs, for the unknown parameter θ for the XLindley model based on record data. The record data are obtained through an inverse sampling scheme, where the units are sequentially observed until the mth record occurs. Additionally, for ease of computation, the mth inter-record time is assumed to be one.

2.1. ML Estimation Based on Records and Inter-Record Times

In this subsection, our attention shifts towards the ML estimate and an ACI for the parameter. Let $\boldsymbol{r} = (r_1, \cdots, r_m)$ and $\boldsymbol{t} = (t_1, \cdots, t_{m-1})$ be the observed sets of $\boldsymbol{R} = \{R_1, \cdots, R_m\}$ and $\boldsymbol{T} = \{T_1, \cdots, T_{m-1}\}$ respectively coming from $XL(\theta)$ distribution. Then, the likelihood function of θ, given the observed lower records and inter-record times, becomes

$$L(\theta; \boldsymbol{r}, \boldsymbol{t}) = \prod_{i=1}^{m} f(r_i)[1 - F(r_i)]^{t_i - 1} = \left(\frac{\theta}{1+\theta}\right)^{2m} \mathrm{e}^{-\theta \sum_{i=1}^{m} r_i} \prod_{i=1}^{m} \left[(2 + \theta + r_i)[\xi(r_i, \theta)]^{t_i - 1}\right], \quad (3)$$

where

$$\xi(x, \theta) = \left(1 + \frac{\theta x}{(1+\theta)^2}\right)\mathrm{e}^{-\theta x}, \qquad \theta > 0. \quad (4)$$

It is important to note that t_m is set to one for the sake of simplifying the equations. Therefore, the resulting log-likelihood function can be expressed as

$$l(\theta; \boldsymbol{r}, \boldsymbol{t}) = 2m \ln \theta - 2m \ln(1+\theta) - \theta \sum_{i=1}^{m} r_i + \sum_{i=1}^{m} \ln(2 + \theta + r_i) + \sum_{i=1}^{m} (t_i - 1) \ln \xi(r_i, \theta).$$

Upon taking the partial derivative of the log-likelihood function with respect to (w.r.t.) θ and setting it equal to zero, we get

$$\frac{\partial l(\theta; \boldsymbol{r}, \boldsymbol{t})}{\partial \theta} = \frac{2m}{\theta(1+\theta)} - \sum_{i=1}^{m} r_i + \sum_{i=1}^{m} \frac{1}{2 + \theta + r_i} + \sum_{i=1}^{m} (t_i - 1) \frac{\psi(r_i, \theta)}{\xi(r_i, \theta)} = 0,$$

where

$$\psi(x, \theta) = \frac{\partial \xi(x, \theta)}{\partial \theta} = -x\mathrm{e}^{-\theta x}\left(1 + \frac{\theta x}{(1+\theta)^2} + \frac{\theta - 1}{(1+\theta)^3}\right), \quad \theta > 0. \quad (5)$$

The ML estimate of θ may be determined through solving the above equation. However, it appears that there is no explicit form for the equation presented above, which necessitates the use of a numerical method. Subsequently, our focus shifts to constructing an ACI for the parameter θ. In this context, Fisher's information is defined as follows $I(\theta) = -E\left(\frac{\partial^2 l \ln f_\theta(\boldsymbol{R}, \boldsymbol{T})}{\partial \theta^2}\right)$, if the integral exists, where $f_\theta(\boldsymbol{r}, \boldsymbol{t})$ denotes the joint probability function of $R_1, T_1, R_2, T_2, \cdots, R_{m-1}, T_{m-1}, R_m$.

The second partial derivative of the log-likelihood function w.r.t. θ is given by

$$\frac{\partial^2 l(\theta; \boldsymbol{r}, \boldsymbol{t})}{\partial \theta^2} = -\frac{2m(1+2\theta)}{[\theta(1+\theta)]^2} - \sum_{i=1}^{m} \frac{1}{(2 + \theta + r_i)^2} + \sum_{i=1}^{m} (t_i - 1)\left(\frac{\psi'(r_i, \theta)\xi(r_i, \theta) - [\psi(r_i, \theta)]^2}{[\xi(r_i, \theta)]^2}\right),$$

where

$$\psi'(x, \theta) = \frac{\partial \psi(x, \theta)}{\partial \theta} = x\mathrm{e}^{-\theta x}\left(x + \frac{\theta x^2}{(1+\theta)^2} + \frac{2x(\theta - 1)}{(1+\theta)^3} - \frac{2(2-\theta)}{(1+\theta)^4}\right), \quad \theta > 0. \quad (6)$$

Let $\hat{\theta}_{ML}$ denote the ML estimator (MLE) of θ. Then, the $100(1-\alpha)\%$ modified asymptotic two-sided equi-tailed confidence interval (MATE CI) for θ can be given by (see for example [25])

$$\left(\max\left\{0, \hat{\theta}_{ML} - \frac{z_{\frac{\alpha}{2}}}{\sqrt{\tilde{I}(\hat{\theta}_{ML})}}\right\}, \hat{\theta}_{ML} + \frac{z_{\frac{\alpha}{2}}}{\sqrt{\tilde{I}(\hat{\theta}_{ML})}}\right),$$

where z_γ represents the γ-th upper quantile of the standard normal distribution and

$$\tilde{I}(\hat{\theta}_{ML}) = -\left.\frac{\partial^2 l(\theta | \boldsymbol{R}, \boldsymbol{T})}{\partial \theta^2}\right|_{\theta = \hat{\theta}_{ML}}.$$

2.2. ML Estimation Based on Record Values

The likelihood function of θ given the lower records $\boldsymbol{r}$ (without considering the inter-record times) is given by

$$L^*(\theta;\boldsymbol{r}) = f(r_m)\prod_{i=1}^{m-1}\frac{f(r_i)}{F(r_i)} = (\frac{\theta}{1+\theta})^{2m}\mathrm{e}^{-\theta\sum_{i=1}^m r_i}\frac{\prod_{i=1}^m(2+\theta+r_i)}{\prod_{i=1}^{m-1}(1-\xi(r_i,\theta))}, \tag{7}$$

where $\xi(x,\theta)$ is defined in (4).

The corresponding log-likelihood function of θ is then given by

$$l^*(\theta;\boldsymbol{r}) = 2m\ln\theta - 2m\ln(1+\theta) - \theta\sum_{i=1}^m r_i + \sum_{i=1}^m \ln(2+\theta+r_i) - \sum_{i=1}^{m-1}\ln[1-\xi(r_i,\theta)]. \tag{8}$$

Taking the first partial derivative of the log-likelihood (8) w.r.t. θ and equating it with zero, we have

$$\frac{\partial l^*(\theta;\boldsymbol{r})}{\partial\theta} = \frac{2m}{\theta(1+\theta)} - \sum_{i=1}^m r_i + \sum_{i=1}^m \frac{1}{2+\theta+r_i} + \sum_{i=1}^{m-1}\frac{\psi(r_i,\theta)}{1-\xi(r_i,\theta)} = 0,$$

where $\psi(x,\theta)$ is defined in (5).

So the ML estimate may be obtained by solving the above equation with the help of a numerical technique.

The second partial derivate of (8) w.r.t. θ is obtained to be

$$\frac{\partial^2 l^*(\theta;\boldsymbol{r})}{\partial\theta^2} = -\frac{2m(1+2\theta)}{[\theta(1+\theta)]^2} - \sum_{i=1}^m \frac{1}{(2+\theta+r_i)^2} + \sum_{i=1}^{m-1}\left(\frac{\psi'(r_i,\theta)[1-\xi(r_i,\theta)] + [\psi(r_i,\theta)]^2}{[1-\xi(r_i,\theta)]^2}\right),$$

where $\psi'(x,\theta)$ is defined in (6).

Let $\hat{\theta}^*_{ML}$ denote the MLE of θ based on lower records. Following the same approach described in the previous subsection, the $100(1-\alpha)\%$ MATE CI for θ can be given by

$$\left(\max\left\{0, \hat{\theta}^*_{ML} - \frac{z_{\frac{\alpha}{2}}}{\sqrt{\tilde{I}^*(\hat{\theta}^*_{ML})}}\right\}, \hat{\theta}_{ML} + \frac{z_{\frac{\alpha}{2}}}{\sqrt{\tilde{I}^*(\hat{\theta}^*_{ML})}}\right),$$

where

$$\tilde{I}^*(\hat{\theta}^*_{ML}) = -\frac{\partial^2 l^*(\theta|\boldsymbol{R},\boldsymbol{T})}{\partial\theta^2}\bigg|_{\theta=\hat{\theta}^*_{ML}}.$$

3. Bayesian Estimation

In the context of Bayesian estimation, the experimenter's information can be conveyed through a probability distribution for the parameter, referred to as the prior distribution. Due to the constraint that the parameter of the XLindley distribution must be positive, we use the popular gamma prior for θ, whose PDF is given by

$$\Pi(\theta) = \frac{b^a\theta^{a-1}\mathrm{e}^{-b\theta}}{\Gamma(a)}, \tag{9}$$

where, the positive hyperparameters a and b can be set based on the prior information available to the experimenter. In what follows, we focus on the Bayesian estimation of θ based on records and inter-record times and based on records alone.

3.1. Bayesian Estimation Based on Records and Inter-Record Times

Using (3) and the prior (9), we can derive the posterior density of θ given $\boldsymbol{r}$ and $\boldsymbol{t}$ as follows

$$\Pi(\theta|\boldsymbol{r},\boldsymbol{t}) = \frac{\theta^{2m+a-1}}{D(1+\theta)^{2m}} \mathrm{e}^{-\theta(b+\sum_{i=1}^{m} r_i)} \prod_{i=1}^{m} \Big[(2+\theta+r_i)[\xi(r_i,\theta)]^{t_i-1}\Big],$$

where $\xi(x,\theta)$ is defined in (4) and

$$D = \int_0^\infty \frac{\theta^{2m+a-1}}{(1+\theta)^{2m}} \mathrm{e}^{-\theta(b+\sum_{i=1}^{m} r_i)} \prod_{i=1}^{m} \Big[(2+\theta+r_i)[\xi(r_i,\theta)]^{t_i-1}\Big] \mathrm{d}\theta.$$

The squared error loss function (SELF) is widely used in Bayesian analyses. However, the SELF may not be appropriate for many real-world scenarios due to its equal weighting of overestimation and underestimation. An alternative asymmetric loss function is the linear-exponential loss function (LELF), proposed by [42], which is given by

$$L_{LE}(\theta,\hat{\theta}) = b[\exp\{c(\hat{\theta}-\theta)\} - c(\hat{\theta}-\theta) - 1], \qquad b>0, \qquad c\neq 0,$$

where $\hat{\theta}$ denotes an estimator of θ.

Without loss of generality, we assume $b = 1$. The appropriate determination of c involves considering both its sign and magnitude. When $c > 0$, then overestimation is more serious than underestimation and vice versa, see [43] for more details. The Bayes estimates of θ under the SELF and LELF become

$$\hat{\theta}_{SE} = \int_0^\infty \theta \Pi(\theta|\boldsymbol{r},\boldsymbol{t})\mathrm{d}\theta = \frac{1}{D}\int_0^\infty \frac{\theta^{2m+a}}{(1+\theta)^{2m}} \mathrm{e}^{-\theta(b+\sum_{i=1}^{m} r_i)} \prod_{i=1}^{m} \Big[(2+\theta+r_i)[\xi(r_i,\theta)]^{t_i-1}\Big] \mathrm{d}\theta,$$

and

$$\begin{aligned}\hat{\theta}_{LE} &= -\frac{1}{c}\ln M(-c|\boldsymbol{r},\boldsymbol{t}) = -\frac{1}{c}\ln[\int_0^\infty \mathrm{e}^{-c\theta}\Pi(\theta|\boldsymbol{r},\boldsymbol{t})\mathrm{d}\theta] \\ &= -\frac{1}{c}\ln\left(\frac{1}{D}\int_0^\infty \frac{\theta^{2m+a-1}}{(1+\theta)^{2m}} \mathrm{e}^{-\theta(c+b+\sum_{i=1}^{m} r_i)} \prod_{i=1}^{m} \Big[(2+\theta+r_i)[\xi(r_i,\theta)]^{t_i-1}\Big] \mathrm{d}\theta\right),\end{aligned}$$

respectively, provided that the integrals exist.

It appears that the above Bayes estimates of θ may not be expressible in closed forms. Therefore, we resort to an approximation method, called the Metropolis-Hastings (M-H) algorithm [27, 15]. An M-H algorithm suitable for our scenario can be outlined as follows.

Algorithm 1

Step1. Start with an initial guess $\theta_0 = \hat{\theta}_{ML}$ and set $t = 1$.

Step2. Given θ_{t-1}, generate θ^* from a truncated-normal distribution, $N(\theta_{t-1},\sigma^2)I_{\{\theta>0\}}$. Then, assign $\theta_t = \theta^*$ with the following probability

$$P = \min\left\{\frac{\Pi(\theta^*|\boldsymbol{r},\boldsymbol{t})q(\theta_{t-1}|\theta^*)}{\Pi(\theta_{t-1}|\boldsymbol{r},\boldsymbol{t})q(\theta^*|\theta_{t-1})}, 1\right\},$$

where $q(x|b)$ represents the density of $N(b,\sigma^2)I_{\{x>0\}}$, otherwise set $\theta_t = \theta_{t-1}$.

Step3. Set $t = t+1$ and repeat Step 2, T times, where T is a considerably large number. So, $\{\theta_{M+1},\theta_{M+2},\cdots,\theta_T\}$ constitutes the generated sample, where M denotes the burn-in period.

The approximate Bayes point estimates of θ under the SELF and LELF are then given by

$$\hat{\theta}_{SM} = \frac{1}{M^*}\sum_{t=M+1}^{T}\theta_t, \qquad \text{and} \qquad \hat{\theta}_{LM} = -\frac{1}{c}\ln\left(\frac{1}{M^*}\sum_{t=M+1}^{T}\mathrm{e}^{-c\theta_t}\right),$$

respectively, with $M^* = T - M$. In Section 5, we have taken $\sigma^2 = 1$.

Let $\theta_{(1)} \cdots \theta_{(M^*)}$ denote the ordered values of $\theta_{M+1}, \cdots, \theta_T$. Define the intervals $L_j(M^*) = [\theta_{(j)}, \theta_{(j+\lceil(1-\alpha)M^*\rceil)}]$ for $j = 1, 2, \cdots, M^* - \lceil(1-\alpha)M^*\rceil$. Consequently, the $100(1-\alpha)\%$ Chen and Shao short width credible interval (CSSW CrI) for θ can be represented as $L_q(M^*)$, where q is determined such that [7]

$$\theta_{(q+[(1-\alpha)M^*])} - \theta_{(q)} = \min_{1\le j\le M^*-[(1-\alpha)M^*]} \theta_{(j+[(1-\alpha)M^*])} - \theta_{(j)}.$$

3.2. Bayesian Estimation Based on Record Values

Using (7) and the prior (9), the posterior density of θ given $\boldsymbol{r}$ is derived to be

$$\Pi^*(\theta|\boldsymbol{r}) = \frac{\theta^{2m+a-1}}{D^*(1+\theta)^{2m}} \mathrm{e}^{-\theta(b+\sum_{i=1}^m r_i)} \frac{\prod_{i=1}^m (2+\theta+r_i)}{\prod_{i=1}^{m-1}[1-\xi(r_i,\theta)]},$$

where

$$D^* = \int_0^\infty \frac{\theta^{2m+a-1}}{(1+\theta)^{2m}} \mathrm{e}^{-\theta(b+\sum_{i=1}^m r_i)} \frac{\prod_{i=1}^m (2+\theta+r_i)}{\prod_{i=1}^{m-1}[1-\xi(r_i,\theta)]} \mathrm{d}\theta.$$

The Bayes estimates of θ under the SELF and LELF become

$$\hat{\theta}^*_{SE} = \frac{1}{D^*} \int_0^\infty \frac{\theta^{2m+a}}{(1+\theta)^{2m}} \mathrm{e}^{-\theta(b+\sum_{i=1}^m r_i)} \frac{\prod_{i=1}^m (2+\theta+r_i)}{\prod_{i=1}^{m-1}[1-\xi(r_i,\theta)]} \mathrm{d}\theta,$$

and

$$\hat{\theta}^*_{LE} = -\frac{1}{c} \ln\left(\frac{1}{D^*} \int_0^\infty \frac{\theta^{2m+a-1}}{(1+\theta)^{2m}} \mathrm{e}^{-\theta(c+b+\sum_{i=1}^m r_i)} \frac{\prod_{i=1}^m (2+\theta+r_i)}{\prod_{i=1}^{m-1}[1-\xi(r_i,\theta)]} \mathrm{d}\theta \right),$$

respectively, provided that the related integrals exist.

It appears that the above Bayes estimates of θ may not have closed-form expressions. So we may use the M-H algorithm (similar to that described in Algorithm 1) to approximate these Bayes estimates, see Subsection 3.1. We can also obtain the $100(1-\alpha)\%$ CSSW CrI for θ using a similar approach detailed in Subsection 3.1.

4. Bayesian Prediction

Let $R_1, T_1, R_2, T_2, \cdots, R_{m-1}, T_{m-1}, R_m$ be the first m of lower record values and their corresponding inter-record times from $XL(\theta)$. Let further $\boldsymbol{r} = (r_1, \cdots, r_m)$ and $\boldsymbol{t} = (t_1, \cdots, t_{m-1})$ be the observed sets of $\boldsymbol{R} = \{R_1, \cdots, R_m\}$ and $\boldsymbol{T} = \{T_1, \cdots, T_{m-1}\}$. We intend to predict the s-th unobserved lower record value, R_s, where $s > m$. Using the Morkovian property of records, the conditional PDF of R_s given $\boldsymbol{R} = \boldsymbol{r}$ and $\boldsymbol{T} = \boldsymbol{t}$, denoted by $f(r_s|\theta, \boldsymbol{r}, \boldsymbol{t})$ is identical to the conditional PDF of R_s given $R_m = r_m$, denoted by $f(r_s|\theta, r_m)$ (see for example [5, 20]). So, we have

$$\begin{aligned} f(r_s|\theta, \boldsymbol{r}, \boldsymbol{t}) \equiv f(r_s|\theta, r_m) &= \frac{f(r_s;\theta)[Q(r_s,\theta) - Q(r_m,\theta)]^{s-m-1}}{F(r_m,\theta)\Gamma(s-m)} \\ &= [Q(r_s,\theta) - Q(r_m,\theta)]^{s-m-1} \frac{\left(\frac{\theta}{1+\theta}\right)^2 (\theta+2+r_s)}{[1-\xi(r_m,\theta)]\Gamma(s-m)} \mathrm{e}^{-\theta r_s}, \end{aligned} \tag{10}$$

where $0 < r_s < r_m$, $Q(x,\theta) = -\ln(F(x;\theta))$ and $\xi(x,\theta)$ is defined in (4).

The Bayes predictive density of R_s given the lower records and inter-record times is obtained to be

$$h(r_s|\boldsymbol{r}, \boldsymbol{t}) = \int_0^\infty f(r_s|\theta, r_m) \Pi(\theta|\boldsymbol{r}, \boldsymbol{t}) \mathrm{d}\theta.$$

It can be easily seen that the associated posterior predictive density may not be obtained analytically. Thus, we estimate $h(r_s|,\boldsymbol{r},\boldsymbol{t})$ by means of a sample generated using the M-H algorithm. Let $\{\theta_v, v=1,\cdots,M^*\}$ be the generated sample using Algorithm 1, where $M^*=T-M$. Then, an estimate of $h(r_s|\boldsymbol{r},\boldsymbol{t})$ is given by

$$\tilde{h}(r_s|\boldsymbol{r},\boldsymbol{t})=\frac{1}{M^*}\sum_{v=1}^{M^*}f(r_s|\theta_v,r_m).$$

The approximate predictions of R_s under the SELF and LELF (provided that they exist) can be obtained as

$$\tilde{R}_s^{SEM}=\int_0^{r_m}r_s\tilde{h}(r_s|\boldsymbol{r},\boldsymbol{t})\mathrm{d}r_s=\frac{1}{M^*}\sum_{v=1}^{M^*}\int_0^{r_m}r_s f(r_s|\theta_v,r_m)\mathrm{d}r_s, \tag{11}$$

and

$$\tilde{R}_s^{LEM}=\frac{-1}{c}\ln\left[\int_0^{r_m}\mathrm{e}^{-cr_s}\tilde{h}(r_s|\boldsymbol{r},\boldsymbol{t})\mathrm{d}r_s\right]=\frac{-1}{c}\ln[\frac{1}{M^*}\sum_{v=1}^{M^*}\int_0^{r_m}\mathrm{e}^{-cr_s}f(r_s|\theta_v,r_m)\mathrm{d}r_s], \tag{12}$$

respectively.

A $100(1-\alpha)\%$ two-sided Bayesian prediction interval for R_s is given by $(L(\boldsymbol{r},\boldsymbol{t}),U(\boldsymbol{r},\boldsymbol{t}))$, where $L(\boldsymbol{r},\boldsymbol{t})$ and $U(\boldsymbol{r},\boldsymbol{t})$ satisfy the following equations at the same time

$$\int_0^{L(\boldsymbol{r},\boldsymbol{t})}h(r_s\mid\boldsymbol{r},\boldsymbol{t})\mathrm{d}r_s=\frac{\alpha}{2},\qquad\text{and}\qquad\int_0^{U(\boldsymbol{r},\boldsymbol{t})}h(r_s\mid\boldsymbol{r},\boldsymbol{t})\mathrm{d}r_s=1-\frac{\alpha}{2}.$$

A $100(1-\alpha)\%$ approximate two-sided Bayesian prediction interval (ATB PI) for R_s is given by (L,U), where L and U satisfy the following equations at the same time

$$\frac{1}{M^*}\sum_{v=M+1}^{T}\int_0^{L}f(r_s\mid\theta_v,r_m)\mathrm{d}r_s=\frac{\alpha}{2},\qquad\text{and}\quad\frac{1}{M^*}\sum_{v=M+1}^{T}\int_0^{U}f(r_s\mid\theta_v,r_m)\mathrm{d}r_s=1-\frac{\alpha}{2}.$$

4.1. Special Case: $s=m+1$

For the special case, when $s=m+1$, then $Y=R_{m+1}$ given $R_m=r_m$ follows the truncated XLindley distribution on interval $(0,r_m)$. So, we have

$$\begin{aligned}f(r_{m+1}|\theta,\boldsymbol{r},\boldsymbol{t})\equiv f(r_{m+1}|\theta,r_m)&=\frac{f(r_{m+1};\theta)}{F(r_m,\theta)}\\&=\frac{\left(\frac{\theta}{1+\theta}\right)^2(\theta+2+r_{m+1})}{1-\xi(r_m,\theta)}\,\mathrm{e}^{-\theta r_{m+1}},\quad 0<r_{m+1}<r_m.\end{aligned} \tag{13}$$

Moreover, we have following two relations

$$\begin{aligned}\int_0^{r_m}r_{m+1}f(r_{m+1}|\theta,r_m)\mathrm{d}r_{m+1}&=-\frac{(\theta+2)(\theta r_{m+1}+1)+r_{m+1}(\theta r_{m+1}+2)+2/\theta}{(1+\theta)^2[1-\xi(r_m,\theta)]}\mathrm{e}^{-\theta r_{m+1}}\Big]_0^{r_m}\\&=\frac{\theta+2+2/\theta-[(\theta+2)(\theta r_m+1)+r_m(\theta r_m+2)+2/\theta]\mathrm{e}^{-\theta r_m}}{(1+\theta)^2[1-\xi(r_m,\theta)]},\end{aligned}$$

$$\begin{aligned}\int_0^{r_m}\mathrm{e}^{-cr_{m+1}}f(r_{m+1}|\theta,r_m)\mathrm{d}r_{m+1}&=-\frac{\theta^2[1+(\theta+c)(\theta+2+r_{m+1})]}{(\theta+c)^2(1+\theta)^2[1-\xi(r_m,\theta)]}\mathrm{e}^{-(\theta+c)r_{m+1}}\Big]_0^{r_m}\\&=\frac{\theta^2\{1+(\theta+c)(\theta+2)-[1+(\theta+c)(\theta+2+r_m)]\mathrm{e}^{-(\theta+c)r_m}\}}{(\theta+c)^2(1+\theta)^2[1-\xi(r_m,\theta)]}.\end{aligned}$$

Therefore, from (11) and (12), the approximate predictions of R_s under the SELF and LELF can be obtained as

$$\begin{aligned}\tilde{R}_{m+1}^{SEM} &= \frac{1}{M^*}\sum_{v=1}^{M^*}\int_0^{r_m} r_{m+1} f(r_{m+1}|\theta_v, r_m)\mathrm{d}r_{m+1}\\ &= \frac{1}{M^*}\sum_{v=1}^{M^*}\frac{\theta_v+2+2/\theta_v-[(\theta_v+2)(\theta_v r_m+1)+r_m(\theta_v r_m+2)+2/\theta_v]\mathrm{e}^{-\theta_v r_m}}{(1+\theta_v)^2[1-\xi(r_m,\theta_v)]},\end{aligned}$$

and

$$\begin{aligned}\tilde{R}_{m+1}^{LEM} &= \frac{-1}{c}\ln\left[\frac{1}{M^*}\sum_{v=1}^{M^*}\int_0^{r_m}\mathrm{e}^{-cr_s} f(r_s|\theta_v, r_m)\mathrm{d}r_s\right]\\ &= \frac{-1}{c}\ln\left[\frac{1}{M^*}\sum_{v=1}^{M^*}\frac{\theta_v^2\{1+(\theta_v+c)(\theta_v+2)-[1+(\theta_v+c)(\theta_v+2+r_m)]\mathrm{e}^{-(\theta_v+c)r_m}\}}{(\theta_v+c)^2(1+\theta_v)^2[1-\xi(r_m,\theta_v)]}\right]\end{aligned}$$

respectively.

Additionally, A $100(1-\alpha)\%$ ATB PI for R_{m+1} is given by (L, U), where L and U satisfy the following nonlinear equations

$$\frac{1}{M^*}\sum_{v=M+1}^{T}\frac{1-\xi(L,\theta_v)}{1-\xi(r_m,\theta_v)}=\frac{\alpha}{2},\qquad\text{and}\quad\frac{1}{M^*}\sum_{v=M+1}^{T}\frac{1-\xi(U,\theta_v)}{1-\xi(r_m,\theta_v)}=1-\frac{\alpha}{2},$$

where $\xi(x,\theta)$ is defined in (4).

Remark 1. Using the Morkovian property of records, the conditional PDF of R_s given $\boldsymbol{R}=\boldsymbol{r}$ is identical to the conditional PDF of R_s given $R_m = r_m$ (see for example [5, 20]). Therefore, the approximate Bayesian point predictions and a $100(1-\alpha)\%$ ATB PI for R_s based on record values can be obtained using a similar procedure described above, with this difference that the M-H sample, $\{\theta_v, v=1,\cdots,M^*\}$, must be generated based on only records.

5. Numerical Illustration

This section involves a simulation study, as well as a real data analysis.

5.1. A Simulation Study

Here, we conduct a Monte Carlo simulation to evaluate the accuracy of the point and interval estimators and approximate predictors that are mentioned in this paper. In this simulation study, we set the number of replications to $N^*=1000$. For each replication, we generate $(m+1)$ records and their associated inter-record times from $XL(\theta)$. We consider the values of m to be $m=3,4,5$ and the values of the parameter to be $\theta=0.5,1$ and 2. In the context of the Bayesian estimation, we use the approximate non-informative prior with $a=b=0.1$. A few replications for which the predictions became negative were removed from the simulation.

We obtain the ML estimates and the approximate Bayes estimates based on the first m records and their corresponding $(m-1)$ record times and based on the first m records alone. Furthermore, we use Geweke's test [13], Raftery and Lewiss diagnostic [36, 37] and Heidelberger and Welch's convergence diagnostic [18] to assess the convergence of the generated M-H Markov chains. It is worth noting that Heidelberger and Welch [18] made use of or referenced the findings of [39, 16, 17, 40, 41]. In some cases, we have taken every second sampled value (and adjusted the number of sampled values accordingly) to ensure a convergent M-H Markov chain. All the final chains have sizes equal to 10000. Figure 1 shows the M-H Markov chains (the figure is for $m=4$ and $\theta=1$), from which the convergence of the M-H algorithm may be confirmed.

The performance of the different estimators is compared based on their estimated biases (biases for short) and estimated risks (ERs). Additionally, we evaluate the interval estimators

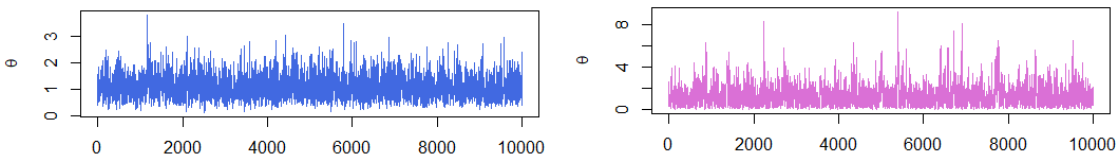


Figure 1: *Plots of Markov chains for θ, the left panel is for the case based on records and inter-records times, whereas the right panel is for the case based on lower records alone ($m = 4$ and $\theta = 1$).*

and predictors using the average width (AW) and coverage probability (CP) criteria. If $\widehat{\theta}$ is an estimator of θ and $\widehat{\theta}_i$ is the corresponding estimate obtained in the i-th replication, then the bias and ERs of $\widehat{\theta}$ w.r.t. the SELF and LELF are given by

$$Bias(\widehat{\theta}) = \frac{1}{N^*}\sum_{i=1}^{N^*}(\widehat{\theta}_i - \theta), \tag{14}$$

$$ER_S(\widehat{\theta}) = \frac{1}{N^*}\sum_{i=1}^{N^*}(\widehat{\theta}_i - \theta)^2, \tag{15}$$

and

$$ER_L(\widehat{\theta}) = \frac{1}{N^*}\sum_{i=1}^{N^*}\Big(\exp[c(\widehat{\theta}_i - \theta)] - c(\widehat{\theta}_i - \theta) - 1\Big), \tag{16}$$

respectively.

The point and interval predictions for the $(m+1)$-th record value, namely R_{m+1}, are also calculated. In terms of prediction assessment, we consider the estimated bias (bias for short) and the estimated prediction risks (EPRs) w.r.t. to the SELF and LELF for the point predictors, which are defined similarly to (14), (15) and (16), respectively. The simulation results are given in Table 1 for point estimation, Table 2 for point prediction and Table 3 for interval estimation and prediction. The results for point estimation and prediction in Tables 1 and 2 are provided for $m = 4$ and 5 for the sake of brevity, whereas the results presented in Table 3 are provided for $m = 3, 4$ and 5.

Based on Tables 1-3, we draw the following conclusions:

- The point estimators based on records and inter-record times outperform the corresponding point estimators based on record alone in terms of bias and ER in the most cases. Additionally, the biases and EPRs of the approximate point predictors based on records and inter-record times are smaller than those of approximate point predictors based on records alone in the most cases, as well.
- The ERs of the point estimators for $\theta = 1$ and 2 decrease w.r.t. to m in the most cases, whereas the EPRs of the point predictors decrease w.r.t. m for all selected values of θ without any exception.
- The AWs of the 95% approximate interval estimators and predictors based on records and inter-record times are less than those of the 95% approximate interval estimators and predictors based on records alone (except for one case for which they are equal up to 5 decimals).
- The CPs of the 95% approximate interval estimators and predictors are all equal to or close to the nominal value 0.95, as expected.

Table 1: *The biases and ERs of the point estimators of θ based on records and inter-record times (first row) and based on records alone (second row).*

	$m=4$				$m=5$			
$\theta=0.5$	bias	ER_S	ER_L $c=0.5$	ER_L $c=-0.5$	bias	ER_S	ER_L $c=0.5$	ER_L $c=-0.5$
MLE	0.0868	0.0971	0.0156	0.0101	0.0669	0.0483	0.0066	0.0056
	1.6656	99.041	>100	0.6617	0.1215	>100	>100	0.8858
Bayes (SELF)	0.0943	0.0991	0.0156	0.0104	0.0729	0.0510	0.0070	0.0059
	0.6746	2.7953	1.7739	0.1732	0.6977	3.0185	2.6072	0.1802
Bayes (LELF)	0.0795	0.0821	0.0123	0.0089	0.0634	0.0457	0.0062	0.0053
$c=0.5$	0.3751	0.7616	0.1494	0.0684	0.3854	0.7803	0.1546	0.0697
Bayes (LELF)	0.1112	0.1246	0.0212	0.0126	0.0831	0.0573	0.0079	0.0066
$c=-0.5$	1.7028	25.912	>100	0.6366	1.7503	26.783	>100	0.6545
$\theta=1$								
MLE	0.2293	0.5745	0.2497	0.0466	0.1564	0.3371	0.1164	0.0300
	7.4661	>100	>100	3.4686	4.8365	>100	>100	2.1844
Bayes (SELF)	0.2394	0.5045	0.1371	0.0443	0.1688	0.3186	0.0806	0.0297
	1.2180	6.3531	43.106	0.3665	1.0927	5.8714	9.0584	0.3326
Bayes (LELF)	0.1652	0.3243	0.0602	0.0317	0.1186	0.2277	0.0412	0.0230
$c=0.5$	0.4764	1.1666	0.2182	0.1077	0.4107	1.0520	0.1980	0.0973
Bayes (LELF)	0.3505	1.2209	14.838	0.0718	0.2316	0.5204	0.5563	0.0406
$c=-0.5$	3.9390	75.039	>100	1.6089	3.4741	67.636	>100	0.0481
$\theta=2$								
MLE	0.5934	2.9392	2.0688	0.1996	0.4055	1.7905	0.9308	0.1353
	9.4667	>100	>100	4.4454	9.9441	>100	>100	4.6788
Bayes (SELF)	0.5169	2.1067	0.7134	0.1628	0.3695	1.4141	0.4473	0.1166
	1.5136	9.0327	7.3213	0.5359	1.5155	8.9470	6.8274	0.5322
Bayes (LELF)	0.2013	0.9833	0.1922	0.0943	0.1448	0.7758	0.1505	0.0759
$c=0.5$	0.1222	1.0090	0.1527	0.1139	0.1291	1.0102	0.1549	0.1128
Bayes (LELF)	1.0939	6.8180	59.371	0.3483	0.7305	3.6409	17.917	0.2150
$c=-0.5$	6.5314	>100	>100	2.7891	6.6037	>100	>100	2.8190

Table 2: *The biases and EPRs of the approximate Bayes point predictors of θ based on records and inter-record times (first row) and based on records alone (second row)..*

	$m=4$				$m=5$			
$\theta=0.5$	bias	ER_S	ER_L $c=0.5$	ER_L $c=-0.5$	bias	ER_S	ER_L $c=0.5$	ER_L $c=-0.5$
SELF	0.002842	0.017065	0.002167	0.002128	0.000548	0.006948	0.000912	0.000843
	0.005136	0.017106	0.002211	0.002098	0.001393	0.007154	0.000958	0.000854
LELF	-0.001371	0.017407	0.002159	0.002222	-0.000844	0.006759	0.000865	0.000837
$c=0.5$	0.000895	0.017163	0.002166	0.002155	-0.000003	0.006893	0.000899	0.000841
LELF	0.007078	0.017195	0.002235	0.002097	0.001943	0.007233	0.000975	0.000858
$c=-0.5$	0.009345	0.017514	0.002317	0.002102	0.002780	0.007493	0.001031	0.000874
$\theta=1$								
SELF	0.002136	0.002491	0.000312	0.000312	0.001369	0.000653	0.000082	0.000081
	0.002754	0.002453	0.000309	0.000305	0.001667	0.000663	0.000084	0.000082
LELF	0.001516	0.002487	0.000310	0.000313	0.001114	0.000645	0.000081	0.000080
$c=0.5$	0.002130	0.002437	0.000305	0.000305	0.001411	0.000651	0.000082	0.000081
LELF	0.002760	0.002504	0.000315	0.000312	0.001625	0.000664	0.000084	0.000082
$c=-0.5$	0.003379	0.000248	0.000313	0.000307	0.001924	0.000678	0.000086	0.000084
$\theta=2$								
SELF	0.000476	0.000378	0.000047	0.000047	-0.000272	0.000105	0.000013	0.000013
	0.000689	0.000397	0.000050	0.000049	-0.000208	0.000106	0.000013	0.000013
LELF	0.000381	0.000374	0.000047	0.000047	-0.000302	0.000105	0.000013	0.000013
$c=0.5$	0.000594	0.000392	0.000049	0.000049	-0.000239	0.000106	0.000013	0.000013
LELF	0.000572	0.000382	0.000048	0.000048	-0.000242	0.000105	0.000013	0.000013
$c=-0.5$	0.000785	0.000401	0.000050	0.000050	-0.000177	0.000106	0.000013	0.000013

Table 3: *The AWs and CPs of* 95% *approximate interval estimators and predictors based on records and inter-record times (first row) and based on records alone (second row).*

	$m=3$		$m=4$		$m=5$	
$\theta=0.5$	AW	CP	AW	CP	AW	CP
MATE CI	1.09579	0.962	0.83556	0.963	0.70605	0.956
	6.92765	0.964	5.89956	0.959	7.21595	0.960
CSSW CrI	1.03103	0.955	0.80036	0.952	0.68460	0.951
	2.98580	0.957	2.93030	0.952	3.00196	0.957
ATB PI	0.47067	0.945	0.23426	0.938	0.11708	0.948
	0.47077	0.942	0.23428	0.938	0.11708	0.949
$\theta=1$						
MATE CI	2.55093	0.953	1.95860	0.954	1.61990	0.951
	11.8789	0.966	24.2695	0.977	16.4109	0.965
CSSW CrI	2.29057	0.954	1.82186	0.955	1.54003	0.946
	5.46901	0.962	5.77409	0.975	5.43254	0.960
ATB PI	0.18213	0.955	0.08794	0.946	0.04852	0.950
	0.18224	0.955	0.08797	0.946	0.04853	0.950
$\theta=2$						
MATE CI	5.81158	0.946	4.58167	0.956	3.75708	0.956
	29.2559	0.962	32.4821	0.955	33.8740	0.967
CSSW CrI	4.66783	0.952	4.00381	0.958	3.39805	0.953
	8.67913	0.959	9.29240	0.952	9.29188	0.962
ATB PI	0.08007	0.935	0.03415	0.954	0.01698	0.950
	0.08012	0.937	0.03416	0.952	0.01699	0.950

5.2. Real Data Example

Here, we consider the following data on the amount of rainfall (in inches) recorded at the Los Angeles Civic Center in February from 1999 to 2018; visit the website of Los Angeles Almanac: www.laalmanac.com/weather/we08aa.php.

```
0.56, 5.54,  8.87, 0.29,  4.64, 4.89, 11.02, 2.37, 0.92,  1.64,
3.57, 4.27,  3.29, 0.16,  0.20, 3.58, 0.83,  0.79, 4.17,  0.03.
```

We have used the Kolmogorov-Smirnov (K-S) test to check if the XLindley model fits the data. The K-S test statistic confirms that the XLindley distribution is quite suitable for fitting the above data (p-value greater than 0.5). We have extracted the lower records and the corresponding inter-record times as follows:

i	1	2	3	4
r_i	0.56	0.29	0.16	0.03
k_i	3	10	6	1

Here, we have used the approximate non-informative prior with $a = b = 0.1$. We have computed the ML and approximate Bayes point estimates, along with the 95% approximate interval estimates of the parameter for the XLindley distribution. Additionally, we have derived the point predictions and 95% ATB PIs for the next future record, namely R_5. The numerical results of this example are given in Table 4, where Case I denotes the case based on records and inter-record times, whereas Case II denotes the case based on records alone. Our findings suggest that the subsequent lowest rainfall amount (after 2018) is expected to be around 0.015 inches, which is the predicted 5-th lower record value since 1999.

Table 4: *The numerical results of the real data example.*

Estimation	MLE	SELF	LELF ($c = 0.5$)	LELF ($c = -0.5$)	95% MATE CI	95% CSSW CrI
Case I	0.9535	0.9729	0.9405	1.0087	(0.2470, 1.6601)	(0.3781, 1.7523)
Case II	1.8809	1.8679	1.4782	2.9436	(0, 4.9336)	(0.0400, 4.6741)
Prediction		SELF	LELF ($c = 0.5$)	LELF ($c = -0.5$)	95% ATB PI	
Case I		0.01495	0.01493	0.01497	(0.00074, 0.02924)	
Case II		0.01488	0.01486	0.01490	(0.00073, 0.02923)	

6. Concluding Remarks

Recently, the XLindley distribution has been introduced by [8] aiming at proposing a flexible distribution for lifetime phenomena. In our study, first, we obtained the ML estimates of the XLindley parameter based on record values and inter-record times, as well as solely based on records. Then, we considered the Bayesian estimation of the parameter, and we employed both symmetric and asymmetric loss functions. The Bayesian point estimates involve integrals that seem to lack closed forms, so we have utilized the M-H method to evaluate them. Our study extended to predicting future records, especially the immediate subsequent lower record value as a special case has been explored in detail. A simulation study has been conducted to evaluate the point and interval estimators of the unknown parameter of the XLindley distribution along with the approximate point and interval predictors of a future lower record value. The simulation study revealed the impact of including the inter-record times on the performance of the estimators and predictors. Furthermore, a real data set containing the rainfall data was analyzed, where a lower record value could serve as an indicator of an impending drought. The predicted values of the 5-th lower record have been obtained in the example. Summing up, the results

of this paper are anticipated to offer practical utility in the estimation and prediction in real phenomena. All the computations of the paper were carried out using the statistical software R [35], and the packages coda [33, 34], nleqslv [14] and truncnorm [26] therein.

Data Availability Statement

The data set used in this paper is provided in the manuscript.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

Funding Details

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] Ahmadi, J., & MirMostafaee, S. M. T. K. (2009). Prediction intervals for future records and order statistics coming from two parameter exponential distribution. *Statistics & Probability Letters*, 79:977–983.

[2] Alotaibi, R., Nassar, M. and Elshahhat, A. (2022). Computational analysis of XLindley parameters using adaptive Type-II progressive hybrid censoring with applications in chemical engineering. *Mathematics*, 10:3355.

[3] Alotaibi, R., Nassar, M. and Elshahhat, A. (2023). Reliability estimation under normal operating conditions for progressively Type-II XLindley censored data. *Axioms*, 2023, 12:352.

[4] Amini, M. and MirMostafaee, S. M. T. K. (2016). Interval prediction of order statistics based on records by employing inter-record times: A study under two parameter exponential distribution. *Metodološki Zvezki*, 13:1–15.

[5] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. Records. John Wiley & Sons, 1998.

[6] Bastan, F. and MirMostafaee, S. M. T. K. (2022). Estimation and prediction for the Poisson-exponential distribution based on records and inter-record times: A comparative study. *Journal of Statistical Sciences*, 15:381-405.

[7] Chen, M.-H. and Shao, Q.-M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics*, 8:69–92.

[8] Chouia, S. and Zeghdoudi, H. (2021). The XLindley distribution: Properties and application. *Journal of Statistical Theory and Applications*, 20:318–327.

[9] Doostparast, M. (2009). A note on estimation based on record data. *Metrika*, 69:69–80.

[10] Doostparast, M., Akbari, M. G. and Balakrishna, N. (2011). Bayesian analysis for the two-parameter Pareto distribution based on record values and times. *Journal of Statistical Computation and Simulation*, 81:1393–1403.

[11] Etemad Golestani, B., Ormoz, E., MirMostafaee, S. M. T. K. (2024). Statistical inference for the inverse Lindley distribution based on lower record values: Accepted-January 2024. *REVSTAT-Statistical Journal*.

[12] Fallah, A., Asgharzadeh, A. and MirMostafaee, S. M. T. K. (2018). On the Lindley record values and associated inference. *Journal of Statistical Theory and Applications*, 17:686–702.

[13] Geweke, K. N. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bayesian Statistics 4, Eds. J. M. Bernardo, J. O. Berger, A.P. Dawid and A. F. M. Smith, Clarendon Press, Oxford, UK, pp. 169–193.

[14] Hasselman, B. (2018). nleqslv: Solve systems of nonlinear equations. R package version 3.3.2, https://CRAN.R-project.org/package=nleqslv.

[15] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109.
[16] Heidelberger, P. and Welch, P. D. (1981). A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM*, 24:233–245.
[17] Heidelberger, P. and Welch, P. D. (1981). Adaptive spectral methods for simulation output analysis. *IBM Journal of Research and Development*, 25:860–876.
[18] Heidelberger, P. and Welch, P. D. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, 31:1109–1144.
[19] Khoshkhoo Amiri, Z. and MirMostafaee, S. M. T. K. (2023). Analysis for the xgamma distribution based on record values and inter-record times with application to prediction of rainfall and COVID-19 records. *Statistics in Transitions new series*, 24:89–108.
[20] Khoshkhoo Amiri, Z. and MirMostafaee, S. M. T. K.(2024). Statistical inference for a two-parameter distribution with a bathtub-shaped or increasing hazard rate function based on record values and inter-record times with an application to COVID-19 data. *Journal of Statistical Computation and Simulation*, DOI:10.1080/00949655.2024.2310682.
[21] Kızılaslan, F. and Nadar, M. (2014). Estimations for proportional reversed hazard rate model distributions based on upper record values and inter-record times. *İstatistik: Journal of The Turkish Statistical Association*, 7:55–62.
[22] Kızılaslan, F. and Nadar, M. (2015). Estimation with the generalized exponential distribution based on record values and inter-record times. *Journal of Statistical Computation and Simulation*, 85:978–999.
[23] Kızılaslan, F. and Nadar, M. (2016). Estimation and prediction of the Kumaraswamy distribution based on record values and inter-record times. *Journal of Statistical Computation and Simulation*, 86:2471–2493.
[24] Kumar, D., Dey, S., Ormoz, E. and MirMostafaee, S. M. T. K. (2020). Inference for the unit-Gompertz model based on record values and inter-record times with an application. *Rendiconti del Circolo Matematico di Palermo Series 2*, 69:1295–1319.
[25] Lehmann, E. L. and Casella, G. Theory of Point Estimation. Second Edition, Springer, 1998.
[26] Mersmann, O., Trautmann, H., Steuer, D. and Bornkamp, B. (2018). truncnorm: Truncated normal distribution, R package version 1.0-8, https://CRAN.R-project.org/package=truncnorm.
[27] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21:1087–1092.
[28] MirMostafaee, S. M. T. K., Asgharzadeh, A., & Fallah, A. (2016). Record values from NH distribution and associated inference. *Metron*, 74:37–59.
[29] Metiri, F., Zeghdoudi, H. and Ezzebsa, A. (2022). On the characterisation of X-Lindley distribution by truncated moments. Properties and application. *Operations Research and Decisions*, 32:97-109.
[30] Nadar, M. and Kızılaslan, F. (2015). Estimation and prediction of the Burr type XII distribution based on record values and inter-record times. *Journal of Statistical Computation and Simulation*, 85:3297–3321.
[31] Nassar, M., Alotaibi, R. and Elshahhat, A. (2023). Reliability estimation of XLindley constant-stress partially accelerated life tests using progressively censored samples. *Mathematics*, 11:1331.
[32] Pak, A. and Dey, S. (2019). Statistical inference for the power Lindley model based on record values and inter-record times. *Journal of Computational and Applied Mathematics*, 347:156–172.
[33] Plummer, M., Best, N., Cowles, K. and Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6:7–11.
[34] Plummer, M., Best, N., Cowles, K., Vines, K., Sarkar, D., Bates, D., Almond, R. and Magnusson, A. (2018). coda: Output analysis and diagnostics for MCMC, R package version 0.19-2, https://CRAN.R-project.org/package=coda.

[35] R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

[36] Raftery, A. E. and Lewis, S. M. (1992). Comment: One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science*, 7:493–497.

[37] Raftery, A. E. and Lewis, S. M. (1996). Implementing MCMC. In Markov Chain Monte Carlo in Practice, Eds. W. R. Gilks, S. Richardson and D. J. Spiegelhalter, Chapman and Hall/CRC, Boca Raton, pp. 115–130.

[38] Samaniego, F. J. and Whitaker, L. R. (1986). On estimating population characteristics from recordbreaking observations. i. parametric results. *Naval Research Logistics Quarterly*, 33:531–543.

[39] Schruben, L. W. (1982). Detecting initialization bias in simulation output. *Operations Research*, 30:569–590.

[40] Schruben, L., Singh, H. and Tierney, L. (1980). A test of initialization bias hypotheses in simulation output. Technical Report 471, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 14853.

[41] Schruben, L., Singh, H. and Tierney, L. (1983). Optimal tests for initialization bias in simulation output. *Operations Research*, 31:1167–1178.

[42] Varian, H. R. (1975). Bayesian approach to real estate assessment. In Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage, Eds. S. E. Fienberg and A. Zellner, North-Holland Pub. Co., Amsterdam, pp. 195–208.

[43] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, 81:446–451.

IMPROVED DEGRADATION TEST USING INVERSE GAUSSIAN PROCESS FOR SIMPLE STEP-STRESS MODEL

G. Sathya Priyanka[1*] S. Rita[2], M. Iyappan[3]

•

[1*]Ph. D Research Scholar, Department of Statistics, Periyar University, Salem-11, India
sathyapriyankstat@gmail.com
[2]Associate Professor and Head, Department of Statistics, Periyar University, Salem-11, India
ritasamikannu@gmail.com
[3]Assistant Professor, Department of Statistics, St. Francis College, Bengaluru-34, India
iyappastat@gmail.com

Abstract

The accelerated Degradation testing (ADT) experiments are important technical methods in reliability studies. Different type of accelerating degradation models has developed with the time and can be used in different types of situations. However, it has become necessary for the manager to test how many numbers of unit should be tested at a particular stress level so that the cost of testing is less. Accelerated Degradation testing (ADT) is preferred to be used in mechanized industries to obtain the required information about the reliability of product components and materials in a short period of time. Accelerated test conditions involve higher than usual pressure, temperature, voltage, vibration or any other combination of them. Data collected at such accelerated conditions are extrapolated through a physically suitable statistical model to estimate the lifetime distribution at design condition stress the life data collected from the high stresses the need to be extrapolated to estimate the life distribution under the normal-use condition. A special class of the ADT is the step-stress testing which regularly increases the stress levels at some pre-fixed time points until the test unit fails. Such experiments allow the experimenter to run the test units at higher-than-usual stress conditions in order to secure failures more quickly. The Inverse Gaussian process is flexible in incorporating random effects and explanatory variables. The different types of models based on IG process are random drift model, random volatility model and random drift-volatility model. In this paper we have considered random drift model for the study on stochastic degradation models for simple step-stress model using inverse Gaussian process observed in degradation problems.

Keywords: Degradation problem, random volatility model, accelerated life testing, inverse Gaussian process, and random drift-volatility model

I. Introduction

In automated industries, Accelerated Degradation Testing (ADT) is the ideal method for quickly obtaining the necessary information regarding the dependability of product components and materials [4]. Higher than normal pressure, temperature, voltage, vibration, etc., or any combination of these, are examples of accelerated test conditions. In order to estimate the lifetime distribution at design condition stress, data collected under such accelerated conditions are

extrapolated using a physically appropriate statistical model. The life data collected from the high stresses must also be extrapolated in order to estimate the life distribution under normal-use conditions. Step-stress testing is a unique type of ADT in which the stress level is gradually increased at predetermined intervals until the test unit malfunctions.

Such tests are mostly conducted in order to obtain dependability data as soon as possible or to save both time and money. Since many pressures tend to accelerate the deterioration process, we can employ accelerated degradation tests (ADT) to acquire degradation phenomena more quickly [5]. To assess the life characteristics of interest under use conditions, a basic constant stress ADT experiment allocates a number of units to different stress levels. The deterioration level of these units is then measured, analyzed, and extrapolated to the failure threshold. ADTs have garnered a lot of attention because to their ability to significantly reduce the testing length. For ADT data, there are two types of models [9].

Since Brownian motion's first passage time has an inverse Gaussian distribution, using it as a life time model makes sense. It is helpful for researching the dependability and life testing of a gadget, product, or subcomponent. In order to shorten the product's life or hasten its performance decline, engineers use accelerated testing to estimate the reliability of recently developed products. The items are subjected to severe conditions during this test, including a mix of random vibrations, increases in temperature, voltage, or pressure [11]. The inverse Gaussian process is a helpful model for repair time. Additionally, in the subject of reliability, the inverse Gaussian distribution has been applied in numerous fields, including hydrology, cardiology.

II. Methods

I. Gaussian Process Model Inverse

An inverse Gaussian process $\{Y(t);\ t \geq 0\}$ with mean function $\Lambda(t)$ and scale parameter λ has the following properties:

- $Y(t)$ has independent increments for every pair of disjoint intervals $(t_1, t_2), (t_3, t_4)$ with $t_1 < t_2 < t_3 < t_4$ the random variables $Y(t_2) - Y(t_1)$ and $Y(t_4) - Y(t_3)$ are independent.
- Each increment $Y(t) - Y(s)$ has an inverse Gaussian distribution $IG(\Delta\Lambda(t), \lambda\,\Delta\Lambda(t)^2)$ where $\Delta\Lambda = \Lambda(t) - \Lambda(s)$ and the PDF of an inverse Gaussian distribution random variable $IG(\mu, \lambda)$ with mean μ and variance $\frac{\mu^3}{\lambda}$ has discussed by Chikkara and Folks (1989) is

$$f(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi}} x^{-\frac{3}{2}} exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right) X > 0 \qquad (1)$$

- $Y(0) = 0$ With probability one. When the amount of degradation reaches a pre-specified critical level D, failure occurs. Let $T = Inf\{t: Y(t) = D\}$ denote the failure time. Since the inverse Gaussian process has a failure time distribution by [16]

$$P(T < t) = P(Y(t) > D) = 1 - G(D;\ \Lambda(t), \lambda\,\Lambda(t)^2)$$
$$= \Phi\left[\sqrt{\frac{\lambda}{D}}(\Lambda(t) - D)\right] - e^{2\lambda\Lambda(t)}\Phi\left[\sqrt{\frac{\lambda}{D}}(\Lambda(t) + D)\right][-\sqrt{\lambda D}(\Lambda(t) + D)] \qquad (2)$$

where, $G(.;\ \Lambda, \lambda)$ is a cumulative distribution function (CDF) of $IG(\Lambda, \lambda)$ and Φ is the standard normal CDF. From above equation we can write the CDF of the failure time distribution as

$$H_\lambda(t) = \Phi\left[\sqrt{\frac{\lambda}{D}}(t-D)\right] - e^{2\lambda t}\Phi\left[\sqrt{\frac{\lambda}{D}}(t+D)\right] \quad (3)$$

It is an increasing function. Thus, within this class of models, there is a one-to-one relationship between$\Lambda(t)$ and the cdf of the failure time distribution $H_\lambda(t)$ for a fixed scale parameter λ.

$$f(x;\mu,\lambda) = \sqrt{\frac{\lambda}{2\Lambda}}x^{\frac{3}{2}}exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right) \quad (4)$$

Where $\mu > 0\ and\ \lambda > 0$ the parameter μ is the mean of the distribution and λ is a scale parameter. (Tweedie) gives three form of above pdf, which he obtained by replace the set of parameters $(\mu,\lambda)\ by\ (\propto,\lambda)\ or\ (\mu,\phi), or\ (\phi,\lambda)$ using the relationship given by [13]

$$\mu = \frac{\lambda}{\phi} = (2 \propto)^{\frac{-1}{2}} \quad (5)$$

Both μ and λ are of the same physical extent as the random variable X itself; but the parameter $\mu = \frac{\lambda}{\phi}$ is invariant under a scale transformation of Xas can be seen from the following relationship:

$$f(x;\ \mu,\lambda) = \mu^{-1}f\left(\frac{x}{\mu};1,\phi\right) = \lambda^{-1}f\left(\frac{x}{\mu};\phi,1\right) \quad (6)$$

The probability density can be numerically computed using any of the three forms in above equation as shown above the cumulative distribution function depends fundamentally on only two variables, which might be taken as $x\mu$ and ϕ. According, the case $\mu = 1$ for the (μ,ϕ) parametric form of above equation could be adopted as a standard form [18]. This has also been obtained as a limiting form of the distribution of the sample size in a Wald's sequential probability ratio test and is sometimes referred to as the standard Wald's distribution of the density function model is

$$\mu\left[\left(1+\frac{9}{4\phi}\right)^{\frac{1}{2}} - \frac{3}{2\phi}\right] \quad (7)$$

II. Random Effects Inverse Gaussian Process

Random effects are needed in Inverse Gaussian process to account for inexplicable heterogeneous degradation rates within the product population. By linking to the Weiner process this investigates different options to incorporate the random effects in the IG process model. Consider the wiener process $W(x) = \mu x + \lambda B(x)$ where $\mu > 0$ is the drift parameter and $\lambda > 0$ is the volatility parameter and B(x) is the standard Brownian motion [12]. Given a fix threshold $\Lambda > 0$, it is well known that the first passage time $T_A = \inf\{x > 0 \mid W(x) \geq \Lambda\}$ follows $IG\left(\frac{\Lambda}{\mu}, \frac{\Lambda^2}{\lambda^2}\right)$ going one step further, we consider a series of the thresholds $\Lambda(t)$ indexed by t with $\Lambda(0) = 0$ and $\Lambda(t)$ increasing in t, and define the first passage time process $Y(t) = T_{\Lambda(t)}$ It is easily verified that the induced $\{Y(t); t > 0\}$ is an IG process with the mean function $\frac{\Lambda(t)}{\mu}$ and variance function $\frac{\Lambda(t)}{\lambda^2}$ by asset of the stationary and independent increment property of the Wiener process W(x).

The inverse relation between the IG and the Wiener processes motivates investigation of the IG process from a new perspective. Existing results on the Wiener processes can let somebody use support to the development of IG process model with the random effects [10].

III. Random Volatility Model

Consider a Wiener process $W(x) = \mu^{-1}x + \lambda^{\frac{-1}{2}}B(x)$ with the induced IG process other way of introducing unit-specific random effects is to assume that each unit possesses a separate realization of the volatility parameter. Accordingly, volatility parameter in the Inverse Gaussian process is random [17]. With the random volatility parameter in the Inverse Gaussian process all units have the same mean degradation path, even though they will have different variance functions. The Inverse Gaussian process with random volatility parameter was originally proposed by Wang and Xu (2010).

Shortcoming of random volatility model is unusual to use the volatility parameter to control heterogeneity in the Weiner process thus application of random volatility model is limited. Thus, random drift model was proposed which overcome inadequacy of random volatility model [13].

IV. Random Drift Model

An effective way to incorporate random effect in the IG process is to let μ be a random variable. To avoid the negative values of μ (Whitmore 1986) and ensure mathematical tractability, we assume $\mu - 1$ follows a truncated normal distribution $TN\ (\omega, k^{-2}), k > 0$ with PDF

$$g(\mu^{-1};\ \omega, k^{-2}) = \frac{k.\phi[k(\mu-1-\omega)]}{1-\Phi(-k\omega)}\mu > 0 \tag{8}$$

Where $(.)$ is a standard normal PDF. In a degradation test, if the degradation of the i[th] testing unit is observed at time $t_{io} < t_{i1} < \cdots . < t_{ini}$ with observations $Y_i(t_{ij}), j = 0,1,2, \ldots . , n_i$ the joint PDF of $Y_i = [Y_i(t_{i1}), Y_i(t_{i2}), \ldots . Y_i(t_{ini})\ ]$ is computed by first conditioning on the random drift parameter μi and then marginalizing it, which yields the following equation is

$$f_{IG}(Y_i) = \frac{1-\phi(-\tilde{\omega}_i\tilde{k}_i)}{1-\phi(-k\omega)}\frac{k}{\tilde{k}_i}\Pi_{j=1}^{ni}\sqrt{\frac{\lambda\Lambda_{ij}{}^2}{2\Lambda yij^3}}\frac{\tilde{k}_i{}^2\tilde{\omega}_i - k^2\omega^2}{2} - \lambda\sum_{j=1}^{ni}\frac{\Lambda ij^2}{2yij} \tag{9}$$

Where, $Y_{ij} = Y_i(t_{ij}) - Y_i(t_{ij} - 1)$ is the observed increment $\Lambda_{ij} = \Lambda(t_{ij}) - \Lambda(t_{ij} - 1)$

$$\tilde{k_{ij}} = \sqrt{\lambda Y_{ij}(t_{ij}k_j) + k^2} \tag{10}$$

$$\tilde{\omega_{ij}} = \frac{[\lambda\Lambda(t_{ij}k_j) + k^2 exp\propto_0 + \propto_1 x_j]}{(\lambda Y_{ij}(t_{ij}k_j) + k^2)} \tag{11}$$

Then the log-likelihood function is given by

$$l(\theta) = \sum_{i=1}^{j}\sum_{j=1}^{N_j}\left[ln\frac{k}{\tilde{k}_{ij}} + \frac{\tilde{k}_{ij}{}^2\tilde{\omega}_{ij}{}^2 - k^2 exp(2\propto_0 + 2\propto_1 x_j)}{2} + \frac{1}{2}\sum_{k=1}^{k_j}\left[In(\lambda\partial\Lambda_{ijk}) - \frac{\lambda\Lambda_{ijk^2}}{y_{ijk}}\right]\right] \tag{12}$$

$l(\theta)$ is the likelihood function up to a constant can be expressed by the above equation. Where θ is a parameter vector include $\propto_0, \propto_1, \lambda, \beta, and\ k$

V. Accelerated Degradation Test Assumptions

Let total N number of units is put into test. Suppose S_0 be the usage stress S_H being the maximum acceptable stress. To collect the degradation data timely we allocate these units J stress level $S_1 < 47\ S_2 < \cdots \ldots \ldots \ldots < S_J$ with $S_0 < S_1$and $S_J = S_H$ consider N_j units to be allocated to j[th] stress level. $j = 1, 2, 3, \ldots . J$. The degradation of these units is affected by the stress. Here, we have

assumed $\mu_i = h(s)$, and $\lambda$is constant over$s$, where $h(s)$ is a link function reflecting the effect of the stress on the degradation process [17].

Due to the above assumption the degradation speed and drift changes with the stress. Another alternative is that $\lambda = h(s)$ while μ is constant which is not valid for random drift model sinceμ is changing from unit to unit. For simplicity and without loss of generality, the additional assumptions are, the measurement time interval, and the number of measurements $K_j$under the j[th] stress level, where $j = 1,2, \ldots\ldots\ldots J$, are pre-determined and the link function follows one of the following acceleration relations:

- Power law relations $h(s) = \varphi_0 . s^{\propto}$
- Arrhenius relation $h(s) = \varphi_0 . e^{\frac{-\propto}{s}}$
- Exponential relation $h(s) = \varphi_0 . e^{\propto s}$

In real time applications the time approved for the test is often given by the manager and time intervals at which the units are measured are predetermined because of the working time of experimenters [10]. Thus, we assume that τ_j and k_j are given. In our model we delight these two variables as decision variables, and then we optimally determine their values. When the assumed stress-degradation relation i.e., is correct we can use a two-stress ADT, i.e., $J = 2$ in our model. But, in this minimum variance plan we are unable to check the validity of the assumed stress-degradation relationship. Thus, we prefer to use three-stress ADT planning taking $J = 3$ to check the validity of the assumed model. In our settings, the purpose of ADT planning is to optimally determine the stress levels (S_j), and the number of samples for each stress level (N_j) are be investigated in our proposed work [4].

VI. Normalizing the Stress Level

We standardize the stress levels depending on the acceleration relationship of the stress on the rate of degradation as follows:

$$Z_j = \frac{lnS_j - \ln S_0}{lnS_H - \ln S_0} \qquad \text{For the power law relation}$$

$$Z_j = \frac{\frac{1}{S_0} - \frac{1}{S_j}}{\frac{1}{S_0} - \frac{1}{S_H}} \qquad \text{For the Arrhenius relation}$$

$$Z_j = \frac{S_j - S_0}{S_H - S_0} \qquad \text{For the exponential relation}$$

From the above consistency, it is readily seen that $x_0 = 0, x_j = 1$, and $0 < Z_j \leq 1$ for $j = 1,2 \ldots, J$ then.

$$h(x) = exp\ (\propto_0 + \propto_1 Z_j)$$

$$h(x) = \varphi_0 . e^{\frac{-\propto}{s}}$$

$$\ln h(x) = \ln \varphi_0 - \frac{\propto}{s}$$

Were, $\propto_0 = \ln \varphi_0 - \frac{\propto}{S_0}$, $\propto_1 = \propto (\frac{1}{S_0} - \frac{1}{S_H})$ For the Arrhenius function, $\propto_0 = \ln \varphi_0 + \propto lnS_0$, $\propto_1 = \propto (\ln S_H - lnS_0)$ For the power law function and $\propto_0 = \ln \varphi_0 + \propto S_0$, $\propto_1 = \propto (S_H - S_0)$ For the exponential function.

VII. Inferential Procedure

We suppose that the i[th] unit under the j[th] stress level is measured at time $t_{ijk} = k_{\tau j}$ with observations $Y_{ij}\ (t_{ijk}), k = 0,1, \ldots\ldots\ldots, k_j$. Let $Y_{ijk} = Y_{ij}\ (t_{ijk}) - Y_{ij}\ (t_{ij}, k-1)$ be the observed increments, and $\Lambda_{ijk} = \Lambda(t_{ijk}) - \Lambda(t_{ijk}, k-1)$. Now, the log-likelihood function up to a constant can be expressed by the equation above 1. The Fisher information matrix $I(\theta)$ for the element $\propto_0, \propto_1, k, \omega, \Lambda(.)$ canbe developed as below [5]. We assume nonlinear function for $\Lambda(.)$, i.e.,

$\Lambda(t) = t_\beta$ and then $\theta = (k, \omega, \propto_0, \propto_1, \beta)'$ detailed expression for the elements along with the elements of the fisher information matrix can be developed as follows.

$$\frac{\partial l(\theta)}{\partial \omega_j} = \sum_{J=1}^{J}\sum_{i=1}^{N_j}\left[0 + \frac{1}{2}\left\{\frac{2(\lambda\Lambda(t_{ij}k_j)+k^2\omega_j)k^2}{(\lambda Y_{ij}(t_{ij}k_j)+k^2)} - 2k{\omega_j}^2\right\} + \frac{1}{2}\sum_{k=1}^{k_j}(0-0)\right] \tag{13}$$

$$\frac{\partial l(\theta)}{\partial \omega_j} = \sum_{J=1}^{J}\sum_{i=1}^{N_j}\left[\left\{\frac{k^2(\lambda\Lambda(t_{ij}k_j)+k^2\omega_j)}{(\lambda Y_{ij}(t_{ij}k_j)+k^2)}2k^2\omega_j\right\}\right] \tag{14}$$

$$\frac{\partial^2 l(\theta)}{\partial {\omega_j}^2} = \sum_{J=1}^{J}\sum_{i=1}^{N_j}\left[\left\{\frac{-k^2(0+k^2)}{(\lambda Y_{ij}(t_{ij}k_j)+k^2)} - k^2\right\}\right] \tag{15}$$

$$\frac{\partial^2 l(\theta)}{\partial {\omega_j}^2} = \sum_{J=1}^{J}\sum_{i=1}^{N_j}\left(\frac{-k^2\left(\lambda Y_{ij}(t_{ij}k_j)\right)}{\lambda Y_{ij}(t_{ij}k_j)+k^2} - k^2\right) \tag{16}$$

$$\frac{\partial l(\theta)}{\partial \beta} = \sum_{j=1}^{J}\sum_{i=1}^{N_j}\left[\left(\frac{\Lambda_{ijk}}{\partial\beta}\left\{\frac{\lambda(\lambda\Lambda_{ijk}+k^2\omega_j)}{(\lambda Y_{ij}(t_{ij}k_j)+k^2)}\right\}\right) + \sum_{k=1}^{k_j}\left(\frac{1}{\Lambda_{ijk}} - \frac{2\lambda\Lambda_{ijk}}{Y_{ijk}}\right)\frac{\partial\Lambda_{ijk}}{\partial\beta}\right] \tag{17}$$

$$\frac{\partial^2 l(\theta)}{\partial k\partial\beta} = \sum_{J=1}^{J}\sum_{i=1}^{N_j}\left[\frac{\left(\lambda\frac{\partial\Lambda(t_{ij}k_j)}{\partial\beta}\right)\left\{\left(2k\omega\lambda Y_{ij}(t_{ij}k_j)\right)+2k\omega_j-k\lambda\Lambda(t_{ij}k_j)+k^3\omega\right\}}{\left(\lambda Y_{ij}(t_{ij}k_j)+k^2\right)^2} - \frac{k\lambda\frac{\partial\Lambda(t_{ij}k_j)}{\partial\beta}}{(\lambda Y_{ij}(t_{ij}k_j)+k^2)}\right] \tag{18}$$

$$\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\propto_1} = \sum_{j=1}^{J}\left[Z_j exp\left(\propto_0+\propto_1 x_j\right)\frac{\partial l(\theta)}{\partial\omega_j} + exp\left(\propto_0+\propto_1 x_j\right)\frac{\partial^2 l(\theta)}{\partial {\omega_j}^2}x_j\right] \tag{19}$$

$$\frac{\partial^2 l(\theta)}{\partial {\propto_0}^2} = \sum_{j=1}^{J}\left[exp\left(\propto_0+\propto_1 Z_j\right)\frac{\partial l(\theta)}{\partial\omega_j} + exp\left(\propto_0+\propto_1 Z_j\right)\frac{\partial^2 l(\theta)}{\partial {\omega_j}^2}\right] \tag{20}$$

$$\frac{\partial^2 l(\theta)}{\partial\lambda\partial\beta} = \sum_{j=1}^{J}\sum_{j=1}^{N_j}\left[\frac{1}{2}\left\{\frac{2\left(2\lambda\Lambda(t_{ij}k_j)\right)\frac{\partial\Lambda(t_{ij}k_j)}{\partial\beta}+k^2\omega\frac{\partial\Lambda(t_{ij}k_j)}{\partial\beta}}{\lambda Y_{ij}(t_{ij}k_j)+k^2} - \frac{Y_{ij}(t_{ij}k_j)\lambda\frac{\partial\Lambda(t_{ij}k_j)}{\partial\beta}}{\left(\lambda Y_{ij}(t_{ij}k_j)+k^2\right)^2}\right\} + \frac{1}{2}\sum_{k=1}^{k_j}\left(-\frac{\Lambda_{ijk}}{Y_{ijk}}\frac{\partial\Lambda_{ijk}}{\partial\beta}\right)\right] \tag{21}$$

And then the fisher information matrix can be developed as given below:

$$\begin{matrix}
E\left[-\frac{\partial^2 l(\theta)}{\partial {\propto_0}^2}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\propto_1}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial k}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\lambda}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\beta}\right] \\
E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\propto_1}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial {\propto_1}^2}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_1\partial k}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_1\partial\lambda}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_1\partial\beta}\right] \\
E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial k}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_1\partial k}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial k^2}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial k\partial\lambda}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial k\partial\beta}\right] \\
E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\lambda}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_1\partial\lambda}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial k\partial\lambda}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\lambda^2}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\lambda\partial\beta}\right] \\
E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_0\partial\beta}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\propto_1\partial\beta}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial k\partial\beta}\right] & \left[-\frac{\partial^2 l(\theta)}{\partial\lambda\partial\beta}\right] & E\left[-\frac{\partial^2 l(\theta)}{\partial\beta^2}\right]
\end{matrix} \tag{22}$$

The log-likelihood function can be maximized to obtain maximum likelihood estimator MLEs [9]. The direct maximization of log-likelihood function gives equations which are computationally difficult to solve. Under the truncated normal distribution, direct maximization of the likelihood function often yields a solution far away from the MLE.

III. Results

I. Numerical Study

Utilizing the methodology of G. Yang et al. (2007), the suggested process is demonstrated here. In a case study, 30 samples at the electrical connector were found to have failed if the data were collected under one of three temperature levels: 55°C, 75°C, or 100°C. The resistors in the MEMS LAB at the Faculty of Engineering and Technology were all part of a constant stress ADT. The normal use temperature and threshold value for the percent increase in resistance were assumed to be l=6, where observed at different times during the measurement. The samples are tabulated in Table 1 with the 7th point of the second unit under 55°C labeled blank, as suggested by Yang et al. (2007), to maintain the monotone behavior of the stress.

Table 1: *Stress relaxation data under the temperature level*

Temperature	S. No	Stress loss	Mean Time
55^0c	1	2.13, 2.06, 3.43, 4.36, 5.86, 6.24, 6.63, 7.34, 7.58, 8.42, 9.57	7.60
	2	2.34, 3.65, 4.69, 4.85, 5.36, 0, 6.59, 8.48, 9.35, 10.95	
	3	2.8, 3.56, 4.65, 5.89, 6.3, 7.65, 8.95, 9.21, 10.45, 11.32	
	4	2.96, 3.58, 5.38, 5.32, 7.68, 8.27, 8.61, 9.854, 10.97, 11.57	
	5	3.65, 4.55, 5.33, 7.58, 8.39, 9.37, 9.33, 10.24, 11.89, 12.54, 13.59	
	6	3.59, 5.69, 5.87, 6.29, 8.98, 10.25, 11.00, 12.69, 13.69, 15.91	
75^0c	7	2.98, 4.98, 5.87, 6.38, 8.56, 10.21, 11.98, 11.00, 13.24, 15.38	10.65
	8	3.65, 4.27, 6.29, 8.91, 9.54, 10.14, 12.69, 14.32, 16.90	
	9	3.69, 4.28, 6.72, 8.34, 8.64, 10.81, 11.20, 14.57, 16.90, 18.18	
	10	3.58, 4.92, 6.91, 7.34, 9.38, 11.78, 12.98, 13.92, 15.39, 18.29	
	11	3.58, 4.87, 7.96, 8.64, 10.94, 12.61, 13.94, 15.38, 17.82, 19.34	
	12	5.96, 5.89, 8.91, 9.67, 12.67, 13.54, 15.98, 17.51, 20.64, 23.94	
100^0c	13	4.89, 5.91, 8.47, 9.38, 11.84, 13.57, 15.94, 16.97, 18.54, 19.82	14.09
	14	4.94, 6.85, 7.95, 9.64, 10.87, 12.67, 15.47, 16.32, 18.94, 21.98	
	15	5.97, 6.31, 8.57, 10.91, 12.97, 14.51, 16.78, 18.96, 19.49, 21.34	
	16	4.25, 7.58, 9.34, 10.64, 13.95, 15.27, 16.97, 19.84, 20.46, 22.7	
	17	5.94, 6.28, 8.94, 12.73, 14.61, 16.37, 18.39, 21.78, 22.96, 24.75	
	18	4.18, 8.91, 10.94, 12.71, 15.67, 17.64, 19.78, 21.64, 24.97, 28.45	

Table 2: *Measurement time under different temperatures*

Temperature	Measurement time epochs (in hours)
55°C	107, 238, 540, 838, 1063, 1249, 1536, 1789, 2164, 2414, 1812
75°C	45, 109, 247, 411, 641, 758, 1017, 1232, 1621, 249
100°C	44, 110, 204, 322, 457, 684, 847, 1041, 1204

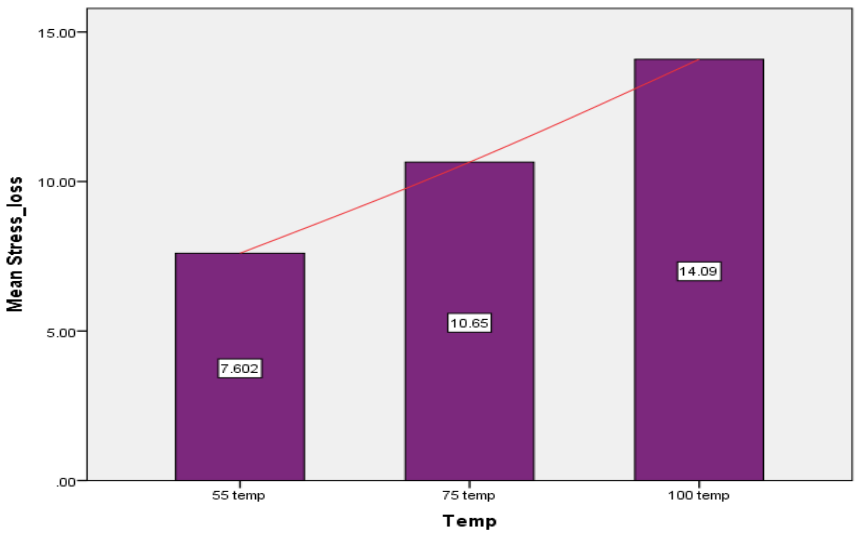


Figure 1: *Measurement temperatures*

In the following, we will determine the optimal ADT plans based on both models. Suppose 10 units are available for the ADT test. In the ADT, we set $\tau_j = 24$, and $k_j = 14$ for all$j = 1,2, \ldots . J$. this setting means that we measure the degradation level once every day, and the test lasts two weeks [19]. Our planning involves selecting the stress level, $(x_1, x_2, \ldots, x_{J-1})$, and the proportion of samples allocated to each testing level, $(N_1, N_2, \ldots, N_{J-1}))$. Consider a two-level ADT plan. Suppose we are interested in minimizing the asymptotic variance of B10, the 0.1-quantile of the failure time distribution at use conditions. When $J = 2$ yields the optimal ADT design

The elements of fisher matrix by solving through mat lab are:

$$\begin{bmatrix} -1.258 \times 10^8 & -1.269 \times 10^8 & -1.6891 \times 10^9 & -20.91 \times 10^8 & -1.62 \times 10^5 \\ -1.18 \times 10^8 & -8.94510 \times 10^7 & -1.6541 \times 10^9 & -15.7351 \times 10^8 & -1.127 \times 10^8 \\ -1.26578 \times 10^9 & -1.3298 \times 10^9 & -6.791 \times 10^9 & -8.734 \times 10^{12} & -1.339 \times 10^9 \\ -21.32 \times 10^8 & -15.761 \times 10^8 & -8.458 \times 10^{12} & -4.9780 \times 10^9 & -1.38 \times 10^7 \\ -1.29 \times 10^5 & -1.113 \times 10^8 & -1.325 \times 10^8 & -1.39 \times 10^7 & -5.69 \times 10^7 \end{bmatrix}$$

Table 3: *Optimization table for random drift model*

Process	x_1	x_2	N_1	N_2	$Std(\varphi p)$
Random drift model	0	1	1	9	4216

The table above displays the ideal ADT design. The fact that 0 is the ideal lower stress value is visually appealing. This outcome is accurate since, even when testing the unit under real-world conditions, the degradation under typical use conditions happens quickly enough to minimize the inaccuracy brought on by extrapolating to the failure threshold.

Table 4: *Optimization table for simple IG process*

Process	x_1	x_2	N_1	N_2	$Std(\varphi p)$
Simple Inverse Gaussian model	0	1	1	9	17450

IV. Discussion

Due to its ability to account for variance in sample product results from unit to unit, the random drift model was chosen for this paper's investigation. With time, many techniques for testing the product are developed. Accelerated deterioration testing, however, is more beneficial in the electronics sector than other approaches. Testing the product quickly is necessary because the corporation creates huge samples of comparable products. To study deterioration performance more effectively, an accelerated degradation test is more appropriate since it increases the stress value during life testing, causing the part to fail faster, and it gathers degradation data to forecast product reliability.

With time, several accelerating degradation models have emerged that can be applied in various contexts. However, to reduce testing costs, it has become imperative for the management to test the number of units that should be tested at a certain stress level. The development of the Simple Stress Accelerated Degradation Test technique considered a number of necessary criteria, including tightening the value of constraints, robustness, and optimality of design. Therefore, the number of units and stress value are optimized using the inverse Gaussian process. This paper presents a proposed model that minimizes the asymptotic variance value to estimate the number of units required for the optimal stress level. A helpful tool for evaluating the value of vectors required to estimate the asymptotic variance is the Fisher information matrix.

References

[1] Bagdonavicius, V., & Nikulin, M. S. (2001). Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, 7(1): 85-103.

[2] Balakrishna, N., Rahul, T. (2014). Inverse Gaussian distribution for Modelling Conditional Durations in Finance. *Communications in Statistics - Simulation and Computation*, 43(3): 476-486.

[3] Bhattacharyya G. K. and ZanzawiSoejoeti (1989). A tampered failure rate model for step-stress accelerated life test, *Communications in Statistics - Theory and Methods*, 18(5):1627-1643.

[4] Chhikara, R.S., Folks, J.L. The inverse gaussian distribution. Statistics: textbook and monographs,1989.

[5] Chien-Yu Peng (2015): Inverse Gaussian Processes with Random Effects and Explanatory Variables for Degradation Data, *Technometrics,* 57(1): 100-111.

[6] Folks, J. L., & Chhikara, R. S. (1978). The inverse Gaussian distribution and its statistical application--a review. *Journal of the Royal Statistical Society. Series B (Methodological)*,40: 263-289.

[7] 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.

[8] Joseph .C, Lu and William. Q, Meeker (1993). Using Degradation Measures to Estimate a Time-to- Failure Distribution, *Technometrics,* 35(2): 161-174.

[9] 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.

[10] Lawless, J. F., & Crowder, M. J. (2010). Models and estimation for systems with recurrent events and usage processes. *Lifetime data analysis, IEEE*. 16(4): 547-570.

[11] Li Sun, Xiaohui Gu, and Pu Song (2016). Accelerated Degradation Process Analysis Based on the Nonlinear Wiener Process with Covariates and Random Effects, *Mathematical Problems in Engineering*, 2016: 1-13.

[12] 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.

[13] Tang, L.C. and Shang, C.D. (1995). Reliability prediction using nondestructive accelerated-degradation data: case study on power supplies. *IEEE Transactions on Reliability*, 44(4):562-566.

[14] Tweedie M.C.K., (1957). Statistical properties of inverse Gaussian distributions I, II, *Annals of Mathematical Statistics,* 28: 362-377.

[15] Leydold, J., Hörmann, W. (2011). Generating generalized inverse Gaussian random variates by fast inversion. *Computational Statistics and Data Analysis,* 55: 213-217.

[16] Park, C. and W. J. Padgett, (2005). New cumulative damage models for failure using Geometric Brownian motion and gamma processes. *Lifetime Data Analysis*, 11(4): 511–527.

[17] Robert Miller and Wayne Nelson, (1983). Optimum Simple Step-Stress Plans for Accelerated Life Testing, IEEE *Transactions on Reliability*, 32(1): 59- 65.

[18] Y Yang, G., & Yang, K. (2002). Accelerated degradation-tests with tightened critical values. *IEEE Transactions* on *Reliability*, 51(4):463-468.

[19] G Yang, Life cycle Reliability Engineering. Hoboken, NJ, USA Wiley, 2007.

[20] Wang Y, Zhang C, Chen X, Tan Y. (2014). Lifetime prediction method for electron multiplier based on accelerated degradation test. *Maintenance and Reliability* 16 (3): 484–490.

CONSTRUCTION OF DOUBLE SAMPLING INSPECTION PLANS FOR LIFE TESTS BASED ON LOMAX DISTRIBUTION

A. Pavithra[1] and R. Vijayaraghavan[2]

•

(1). Assistant Professor, Department of Statistics, PSG College of Arts and Science, Coimbatore 641 014, Tamil Nadu, INDIA
(2). Senior Professor, Department of Statistics, Bharathiar University, Coimbatore 641 046, Tamil Nadu, INDIA
[1]pavistat95@gmail.com, [2]vijaystatbu@gmail.com

Abstract

A life test is a random experiment performed on manufactured products such as electrical and electronic components to estimate their life period based on a randomly chosen components. The lifespan of a component is considered as a random variable that follows a certain continuous-type distribution, called the lifetime distribution. Reliability sampling is one of the decision-making methodologies in product control and deals with inspection procedures for sentencing one or more lots or batches of items submitted for inspection. The concept of sampling plans for life tests involving with two random samples is employed in the present study under the assumption that the lifetime random variable is described by the Lomax distribution. A procedure based on mean / median life criterion is developed for designing the optimum plans with minimum sample sizes when two points on the desired operating characteristic curve are prescribed to ensure protection to the producer and the consumer.

Keywords: Consumer's risk, Double sampling plan, Lomax distribution, Operating characteristic function, Producer's risk, Reliability sampling.

1. Introduction

Sampling inspection is a product control strategy that decides whether a lot should be accepted or rejected based on the information obtained by the inspection of random sample(s) drawn from the submitted lot(s). Sampling inspection procedures are generally classified according to the nature of the quality characteristics, namely, measurable and non-measurable. When the quality characteristics are non-measurable, but are classified into go or no-go basis, such as good or bad, non-conforming or conforming, etc., the sampling inspection procedures are termed as attribute sampling. When the quality characteristics are measurable on a continuous scale, the corresponding sampling inspection procedures are called variables sampling, which are devised under the implicit assumption that the quality characteristic is a continuous random variable following a specific probability distribution. Reliability sampling plans, also termed as life test sampling plans, are operationally attributes sampling procedures, but involve lifetime of the components or items as a random variable which is distributed according to a specific continuous type probability distribution, such as the exponential, Weibull, lognormal, gamma distributions, etc. The lifetime of the components or items is observed by putting the sampled items under the test, called life test, which is defined as the process of evaluating the lifetime of the items through experiments. The

literature in product control provides the importance of various continuous probability distributions like exponential, Weibull, lognormal and gamma distributions as well as several compound distributions for modeling lifetime data in the studies relating to the design and evaluation of reliability sampling plans.

The earlier works, which laid the foundation for the expansion of various types of sampling plans, would include the theory of reliability sampling proposed and developed from [1] - [8]. Significant contributions in the development of life test sampling plans employing exponential, Weibull, lognormal and gamma distributions as well as several compound distributions for modeling lifetime data have also been made in the past four decades. A detailed account of such plans was provided in [9]. The recent advances in the theory of life test sampling plans provided in [10] – [28].

Lomax distribution, introduced in [29], is a heavy-tailed probability distribution and is considered as Pareto Type II distribution. It has a wide range of applications in many fields which include business, economics, actuarial, medical and biological sciences. It has been proved to be much useful in reliability and life testing studies and in survival analysis. Properties of Lomax distribution and its extended form can be seen in [30] – [33]. In this paper, a specific life-test sampling plan is devised with reference to the life-time quality characteristic, which is modeled by Lomax distribution. A procedure for the selection of such plans indexed by acceptable and unacceptable mean life ensuring protection to the producer and consumer is described with illustrations. Tables yielding optimum Double sampling plans for life tests are constructed for a set of fixed values of shape parameter of the Lomax distribution.

2. Lomax Distribution

Let T be a random variable representing the lifetime of the components. Assume that T follows Lomax distribution. The probability density function and the cumulative distribution function of T are, respectively, defined by

$$f(t;\theta,\lambda) = \frac{\lambda}{\theta}\left(1+\frac{t}{\theta}\right)^{-(\lambda+1)}, t>0, \theta>0, \lambda>0 \tag{1}$$

and $$F(t;\theta,\lambda) = 1-\left(1+\frac{t}{\theta}\right)^{-\lambda}, t>0, \theta>0, \lambda>0, \tag{2}$$

where λ and θ are the shape and scale parameters, respectively.

The mean life, the median life, the reliability function and hazard function for specified time t under Lomax distribution are. Respectively, given by

$$\mu = \frac{\theta}{\lambda-1},\ for\ \lambda>1, \tag{3}$$

$$\mu_d = \theta\left(\sqrt[\lambda]{2}-1\right), \tag{4}$$

$$R(t;\theta,\lambda) = \left(1+\frac{t}{\theta}\right)^{-\lambda}, t>0, \theta>0, \lambda>0 \tag{5}$$

and $$Z(t;\theta,\lambda) = \frac{\lambda}{\theta}\left(1+\frac{t}{\theta}\right)^{-1}, t>0, \theta>0, \lambda>0. \tag{6}$$

The reliability life is the life beyond which some specified proportion of items in the lot will survive. The reliability life associated with Lomax Distribution is defined and denoted by

$$\rho(t;\theta,\lambda) = \theta\left(R^{-1/\lambda}-1\right), \tag{7}$$

Where $\underline{R}$ is the proportion of items surviving beyond lifeρ.

The proportion,p, of product failing before time t, is defined by the cumulative probability distribution of T and is expressed by

$$p = P(T \leq t) = F(t; \theta, \lambda). \tag{8}$$

3. Operating Characteristic Function of Life Test Sampling Plan

The performance of a single sampling plan adopted in life testing is measured by the associated operating characteristic (OC) function, denoted by $P_a(p)$, which gives the probability of accepting a lot as a function of the failure probability p. Under the conditions for the application of binomial and Poisson models, the expressions for $P_a(p)$ are, respectively, expressed by

$$P_a(p) = \sum_{x:0}^{c} \binom{n}{x} p^x (1-p)^{n-x} \tag{9}$$

and $$P_a(p) = \sum_{x:0}^{c} e^{-np} \frac{(np)^x}{x!}. \tag{10}$$

Associated with a specific value of p, there exists a unique value of t/θ, which can be derived as a function of p and $\lambda$from the cumulative distribution function by virtue of expressions (2) and (8) as

$$\frac{t}{\theta} = (1-p)^{-1/\lambda} - 1. \tag{11}$$

The expression for $t/\mu$is then derived using (3) and (11) as

$$\frac{t}{\mu} = (\lambda - 1)\left[(1-p)^{-1/\lambda} - 1\right], \tag{12}$$

which indicates that associated with any specific value of p, there exists a unique value of the dimensionless ratio t/μ. As the value of p is associated witht/μ, the operating characteristic function of a life test sampling plan can be considered as a function of t/μ rather than p, and, hence, the OC curve of the plan could be obtained by plotting the acceptance probabilities against the values oft/μ.

4. Double Sampling Plans for Life Tests with Zero or One Failure

Often in practice sampling inspection plans for life tests are required to be constructed for product characteristics that involve costly or destructive testing. Industrial situations sometimes may warrant small samples to be used for inspection. In such cases, sampling inspection plans allowing either zero failures or a fewer number of failures in the samples are often employed for sentencing the submitted lots. According to [34], a single sampling plan by attributes with zero acceptance number (zero failures) is undesirable as it does not provide protection to the producer and fails to safeguard the primary interests of the producer.

Figure 1 depicts that a single sampling plan for life tests with zero failures or zero acceptance number, designated by $SSP-(n,\ 0)$, is not desirable as it fails to provide protection to the producer against the acceptable mean life of the product. It can be realized in general that the *OC* curves of any such single sampling plans having zero failures would be uniquely in poor shape, which obviously does not ensure protection to producers, but safeguard the interests of consumers against unacceptable mean life of the product.

It can be demonstrated that single sampling plans allowing one failure or more number of failures in a sample of items do not possess the undesirable properties or characteristics of $SSP-(n,\ 0)$, but would require larger sample sizes rather than small sample sizes. This shortcoming can be prevailed, to some extent, if double sampling plans allowing a maximum of one failure in the random samples drawn from the lot are effectively adopted for sentencing the lot submitted.

In small sample situations, single sampling plans with a fewer number of failures such as $c = 0$and $c = 1$ can be used. But, the *OC* curves of $c = 0$ and $c = 1$plans reveal the fact that there would always be a conflicting interest between the producer and the consumer as $c = 0$ plans would provide protection to the consumer with lesser amount of risk of accepting the lot against the unacceptable mean life of the product while $c = 1$ plans offer protection to the producer with lesser risk of rejecting the lot having acceptable mean life. Such conflict can be annulled if one is able to design a life test plan having its OC curve lying between the *OC* curves of $c = 0$ and $c = 1$plans.

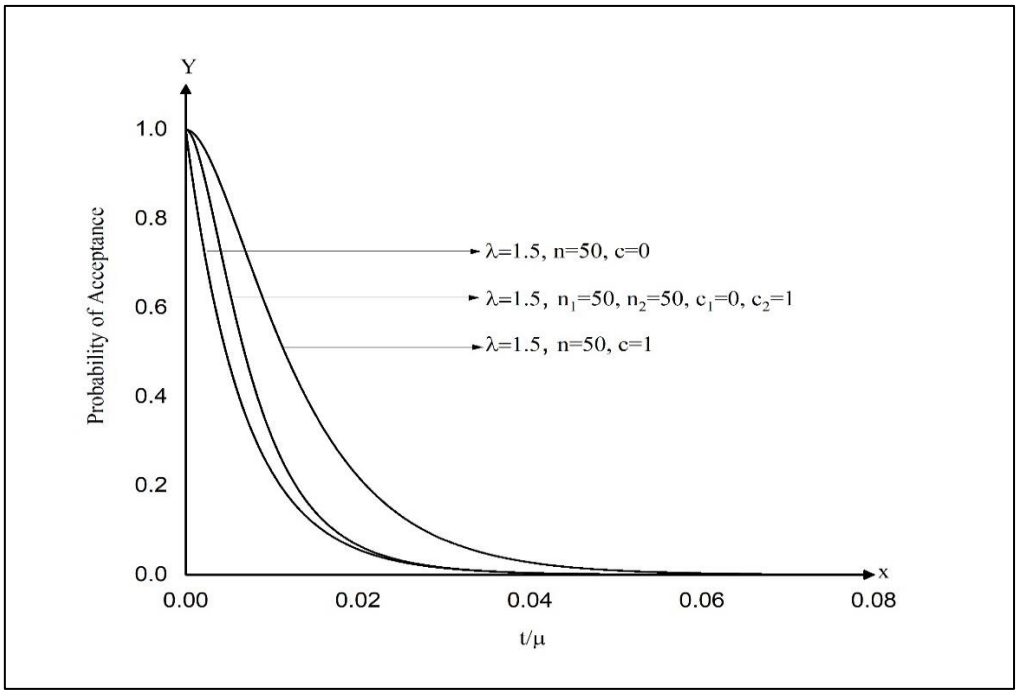


Figure 1: *Operating Characteristic Curves of Single and Double Sampling Plans for Life Tests Based on Lomax Distribution Having smaller acceptance number*$c_1 = 0$ *and*$c_2 = 1$

From Figure 1, it can also be observed that there is a wide gap between the *OC* curves of $c = 0$ and $c = 1$plans. Hence, it is desirable to bridge the gap by determining a suitable plan such that its *OC* curve is expected to lie between the *OC* curves of $c = 0$ and $c = 1$ plans.

A double sampling plan with $c_1 = 0$ and $c_2 = 1$, designated by $DSP - (n_1, n_2)$, overcomes the shortcoming of $c = 0$ plans to a greater extent by providing a desirable shape of the *OC* curve, which is considered as favorable to both producer and consumer. It can also be shown that the *OC* curves of $DSP - (n_1, n_2)$ would lie between the *OC* curves of $c = 0$ and $c = 1$ plans.

One can observe that the *OC* curve of $DSP - (n_1, n_2)$ coincides with the *OC* curve of $c = 1$ single sampling plan at the upper portion and coincides with the *OC* curve of $c = 0$ single sampling plan at the lower portion. This salient feature would be of much help in determining an optimum $DSP - (n_1, n_2)$ providing protection to the producer and consumer against rejection of the lot for the specified acceptable mean life and against acceptance of the lot for the specified unacceptable mean life. Detailed discussion on the significance and construction of sampling plans with the utilization of the conditions c = 0 and c = 1 together as an alternative to single sampling plans with either c = 0 or with c = 1 are found in the literature of acceptance sampling and especially from [35] – [37]. The operating procedure of $DSP - (n_1, n_2)$ is as follows:

Step 1: Draw a random sample of n_1 items from a given lot and put them for a life test.

Step 2: Observe the number,m_1, of failures before reaching the predetermined time t. If $m_1 = 0$, while testing n_1 items, then accept the lot; if $m_1 > 1$, reject the lot; if $m_1 = 1$, draw a second random sample of $n_2$items and put them for a life test.

Step 3: Observe the number, m_2, of failures while testing $n_2$items. If $m_2 = 0$, then accept the lot; if

$m_2 \geq 1$, then reject the lot.

Associated with $DSP-(n_1, n_2)$ are the performance measures, called *OC* and *ASN* functions, which are, respectively, expressed by

$$P_a(p) = p(0|n_1, p) + p(1|n_2, p)p(0|n_2, p) \tag{13}$$

And $ASN(p) = n_1 + n_2 p(1|n_1, p),$ (14)

Where p is the proportion,p, of product failing before time t, and $p(0|n_1, p)$, $p(0|n_2, p)$ and$p(1|n_1, p)$ are defined either from the binomial distribution or from the Poisson distribution whose probability functions are given as expressions (9) and (10). Under the conditions of binomial distribution, the expressions for $P_a(p)$and $ASN(p)$ are, respectively, given by

$$P_a(p) = (1-p)^{n_1} + n_1 p(1-p)^{n_1+n_2-1} \tag{15}$$

and $ASN(p) = n_1 + n_1 n_2 p(1-p)^{n_1-1}.$ (16)

Similarly, under the conditions of Poisson distribution, the expressions for $P_a(p)$and $ASN(p)$ are, respectively, given by

$$P_a(p) = exp(-n_1 p) + n_1 p\, exp(-(n_1+n_2)p) \tag{17}$$

and $ASN(p) = n_1 + n_1 n_2 p\, exp(-n_1 p).$ (18)

It is known that, under the assumption of Lomax distribution for a lifetime quality characteristic, p is defined by the cumulative probability distribution of the lifetime random variable, T, and is expressed by

$$p = P(T \leq t) = F(t; \theta, \lambda)$$

It can be noted that the double sampling plan for life tests allowing a maximum of one failure based on Lomax distribution is specified by the parameters $n_1, n_2, \theta$and λ, where n_1 and n_2 are the sample sizes under the plan, and $\theta$and λ are the parameters of Lomax distribution. As discussed earlier, the failure probability p is associated with t/θ, through the distribution function of Lomax distribution, and the acceptance probabilities can be computed when the sets of values of $n_1, n_2$and λ are specified. The probabilities of acceptance of the submitted lot under the double sampling plan for life tests can be computed against the dimensionless ratioμ/μ_0 based on the procedure described in the following Subsection for different combinations of parameters n_1, n_2 and λ, where μ/μ_0 is the ratio of the actual mean life to the assumed mean life. It is to be noted that any change in the values of these parameters would have some impact in the nature of the OC curve.

While selecting a sampling inspection plan for its application, it is the conventional practice to define the *OC* curve in accordance with the desired discrimination and to select the corresponding sampling plan. It is known that the operating ratio, defined as the ratio of the limiting quality level to the acceptable quality level, is one of the widely used measures of discrimination in sampling plans, and is, in general, used to fix the *OC* curve.

Further, a smaller value of μ/μ_0 would indicate that the actual mean life is relatively much smaller than the acceptable mean life whereas a larger value, which is nearer to one would indicate that the difference between μ and μ_0 is less. When the actual mean life is much smaller than the acceptable mean life, smaller the values of λ, greater is the protection to the consumer, whereas protection to the producer is more for larger values of λ. As the actual mean life increases, acceptance probabilities would increase, which indicate that the lots having items with higher mean life that is close to the acceptable mean life will most often have a greater chance of acceptance.

5. Procedure for the Selection of $DSP - (n_1, n_2)$ for Life Tests

Sampling inspection plans for attributes or variables are constructed based on a general approach that the operating characteristic curves of the desired plans should pass through two prescribed points, namely, the acceptable quality level,p_0, and the limiting quality level,p_1, which are associated with the producer's risk, α, and the consumer's risk,β, respectively. The specification of these points is required for the purpose of ensuring protection to the producer as well as the consumer and is considered for fixing the *OC* curve in accordance with a desired degree of discrimination. The operating ratio, *R*, defined as the ratio of $p_1$top_0, is often used as the measure of discrimination.

As discussed in the earlier sections, a specific sampling plan for life tests can be determined by specifying the requirements that the OC curve should pass through two prescribed points, namely,(μ_0, α) and (μ_1, β), where$\mu_0$and$\mu_1$are the acceptable and unacceptable mean life associated with the risks $\alpha$and β, respectively. In such a case, the operating ratio, $R = \mu_0/\mu_1$, which is the ratio of acceptable mean life to unacceptable mean life, can be used as the measure of discrimination just similar to the operating ratio of the limiting quality level to the acceptable quality level. It is obvious to note that $\mu_1<\mu_0$, and hence, R >1. An optimum double sampling plan for life tests can be determined by satisfying the following two conditions so that the maximum producer's and consumer's risks would be fixed at $\alpha$and β, respectively:

$$P_a(\mu_0) \geq 1 - \alpha \tag{19}$$

and

$$P_a(\mu_1) \leq \beta. \tag{20}$$

It is to be noted that the *OC* function given as (15) or (17) is not directly related to the mean life; but it can be expressed as a function of t/μ, which corresponds to *p*, *i.e.*, the proportion of lot failing before time *t*. The procedure is appropriately used to compute the operating characteristics while searching for the optimum values of the sampling plan satisfying the conditions (19) and (20).

For the specified values of t/μ_0 andt/μ_1, the optimum values of n_1 and n_2 of $\boldsymbol{DSP - (n_1, n_2)}$ under the conditions of Lomax distribution satisfying the conditions (19) and (20) can be determined by using the following procedure:

Step 1: Specify the value of the shape parameter λ or its estimate.

Step 2: Specify the values of t/μ_0 andt/μ_1, with the associated risks $\alpha = 0.05$and$\beta = 0.10$, respectively, so that the operating ratio is defined by$R = \mu_0/\mu_1$.

Step 4: Using the relationship between *p* andμ, from (3) and (8), obtain$p_0$and$p_1$corresponding to t/μ_0 and t/μ_1.

Step 5: Search for the values of n_1 and n_2 for the specified strength $(\mu_0,\ 1 - \alpha)$ and $(\mu_1,\ \beta)$ with the values of p_0 andp_1, or equivalently with the values of t/μ_0 andt/μ_1, by using either the expression (15) or the expression (17), such that the conditions (19) and (20) are satisfied.

Based on the above procedure, the optimum double sampling plans for life tests under the assumption of Lomax distribution are obtained for a set of five values of λ, given as 1.25, 1.5, 1.75, 2 and 3, and for various sets of combinations of $R = \mu_0/\mu_1$ and t/μ_0. These plans are provided in Tables 1 through to 5 along with the values of minimum *ASN* att/μ_0. The optimum plans given in the tables are obtained under the conditions of binomial distribution by a search procedure using the expression (15) for the *OC* function and the expression (8) for the proportion of product failing in an appropriate manner. The parameters of the optimum plans would have a maximum of 5 percent producer's risk and a maximum of 10 percent consumer's risk.

5.1. Numerical Illustration

In an electronic device manufacturing industry, a quality control practitioner wishes to adopt a suitable sampling inspection plan under isolated lot conditions. Though the practitioner is interested to have only zero failures in the random sample items which are placed under the life test, keeping

in mind the manufacturer's capabilities of producing long survival items, he wishes to allow a maximum of one failure item under the sampling plan. Hence, he desires to adopt a double sampling plan allowing a maximum of one failure in the randomly sampled items which are to be considered for a life test.

The past history in the industry reveals that the life time random variable is distributed according to Lomax distribution, whose shape parameter is specified to be λ = 1.5. It is expected that the plan shall provide the desired degree of discrimination, which is measured in terms of the operating ratio,R, ensuring protection to the producer in terms of the acceptable mean life $\mu_0 = 2000$ hours with the associated risk of 5 percent and protection to the consumer against the unacceptable mean life $\mu_1 = 110$ hours with the associated risk of 10 percent.

The practitioner would like to terminate the life test within 1 hour. Based on the given information, one gets $R = \mu_0/\mu_1 = 18.2 \approx 18$, and$t/\mu_0 = 0.0005$. The value of shape parameter λ = 1.5, with the ratio $R = \mu_0/\mu_1 = 18$ and $t/\mu_0 = 0.0005$, the optimum double sampling plan is chosen with the sample sizes $n_1 = 88$ and $n_2 = 178$, which yield the minimum$ASN = 109$ att/μ_0. Thus, the desired plan for the given conditions is implemented as given below:

1. Draw a random sample of $n_1 = 88$ items from a submitted lot and place them for life test.
2. Observe the number of failures before reaching the termination time of 1 hour.
3. Terminate the life test once the termination time, *i.e.*, t = 1 hour, is reached.
4. If no failures are observed in the 88 items tested or until time t is reached, accept the lot; if one failure is observed in the 88 items tested, select a random sample of $n_2 = 178$ items and place them for a life test.
5. Accept the lot, when no failures are observed while testing 178 items; reject the lot, if one or more failures are observed.
6. Treat the items which survive beyond time t = 1 hour as accepted.

5.2. Numerical Illustration

It is assumed that the lifetime of the components in an electronic device follows a Lomax distribution with shape parameter λ = 3.0. It is desired to implement a double sampling plan for life tests to sentence a submitted lot of manufactured components. The experimenter involved in the decision-making process fixes the test termination time as t = 75 hours. The acceptable and unacceptable proportions of the lot failing before time t are, respectively, prescribed as $p_0 = 0.003$ and $p_1 = 0.055$with the associated risks fixed at the levels $\alpha = 0.05$ and $\beta = 0.10$. The values of t/μ corresponding to $p_0 = 0.003$ and $p_1 = 0.055$ are determined as $t/\mu_0 = 0.002$and $t/\mu_1 = 0.038$. Hence, the desired operating ratio is obtained as $R = \mu_0/\mu_1 = 19$.

The value of shape parameter λ = 3, with the ratio R = 19 and the index $t/\mu_0 = 0.002$, one obtains the optimum double sampling plan having its parameters specified as $n_1 = 42$ and $n_2 = 120$ which yield the minimum$ASN = 56$ at $t/\mu_0 = 0.002$. These parameters satisfy the conditions (9) and (10). The acceptable mean life and unacceptable mean life are, then, determined as $\mu_0 = t/0.002 = 37500$ hours and $\mu_1 = t/0.038 = 1973.7 \approx 1974$ hours, respectively.

6. Conclusion

A double sampling inspection plans for life-tests which involve two samples and allows a maximum of one failure is proposed when the lifetime quality characteristic is modeled by a Lomax distribution. A procedure for the selection of the proposed plan is discussed through numerical illustrations. The sampling plan which could be derived by the procedure discussed in this paper will ensure protection to the producer and consumer as the plans are indexed by acceptable and unacceptable proportion of product failing before the specified time, t. The practitioners can generate the required sampling plans for various choices of shape parameter λ, adopting the procedure.

7. Acknowledgment

The authors are grateful to the Editor and Reviewers for making significant suggestions for improving the paper's substance. The authors are indebted to their respective institutions, namely, PSG College of Arts & Science, Coimbatore, India and Bharathiar University, Coimbatore, India for providing necessary facilities to carry out this research work.

References

[1] Epstein, B. (1960a). Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential, Part I, *Technometrics*, 2, pp. 83 - 101.

[2] Epstein, B. (1960b). Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential, Part II, *Technometrics*, 2, pp. 167 - 183.

[3] Handbook H-108. (1960). Sampling Procedures and Tables for Life and Reliability Testing, Quality Control and Reliability, Office of the Assistant Secretary of Defense, US Department of Defense, Washington, D.C.

[4] Goode, H. P., and Kao, J. H. K. (1961). Sampling Plans Based on the Weibull Distribution, *Proceedings of the Seventh National Symposium on Reliability and Quality Control,* Philadelphia, PA, pp. 24 - 40.

[5] Goode, H. P., and Kao, J. H. K. (1962). Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (Hazard Rate Criterion), *Proceedings of the Eight National Symposium on Reliability and Quality Control*, Washington, DC, pp. 37 - 58.

[6] Goode, H. P., and Kao, J. H. K. (1964). Hazard Rate Sampling Plans for the Weibull Distribution, *Industrial Quality Control*, 20, pp. 30 - 39.

[7] Gupta, S.S. and Groll, P. A. (1961). Gamma Distribution in Acceptance Sampling Based on Life Tests, *Journal of the American Statistical Association*, 56, pp. 942 – 970.

[8] Gupta, S. S. (1962). Life Test Sampling Plans for Normal and Lognormal Distributions, *Technometrics*, 4, pp. 151 - 175.

[9] Schilling, E. G., and Neubauer, D. V. (2009). *Acceptance Sampling in Quality Control*, Chapman and Hall, New York, NY.

[10] Wu, J. W., and Tsai, W. L. (2000). Failure Censored Sampling Plan for the Weibull Distribution, *Information and Management Sciences*. 11, pp. 13 - 25.

[11] Wu, J. W., Tsai, T. R., and Ouyang, L. Y. (2001). Limited Failure-Censored Life Test for the Weibull Distribution, *IEEE Transactions on Reliability*, 50, pp. 197 - 111

[12] Kantam, R.R.L., Rosaiah, K., and Rao, G.S. (2001). Acceptance Sampling Based on Life Tests: Log-Logistic Models, *Journal of Applied Statistics,* 28, pp. 121 - 128.

[13] Jun, C.-H., Balamurali, S., and Lee, S.-H. (2006). Variables Sampling Plans for Weibull Distributed Lifetimes under Sudden Death Testing, *IEEE Transactions on Reliability*, 55, pp. 53 - 58.

[14] Tsai, T.-R., and Wu, S.-J. (2006). Acceptance Sampling Based on Truncated Life-tests for Generalized Rayleigh Distribution, *Journal of Applied Statistics*, 33, pp. 595 - 600.

[15] Balakrishnan, N., Leiva, V., and Lopez, J. (2007). Acceptance Sampling Plans from Truncated Life-test Based on the Generalized Birnbaum - Saunders Distribution, *Communications in Statistics - Simulation and Computation*, 36, pp. 643 - 656.

[16] Aslam, M., and Jun, C.-H. (2009a). A Group Acceptance Sampling Plan for Truncated Life Test having Weibull Distribution, *Journal of Applied Statistics*, 39, pp. 1021 - 1027.

[17] Aslam, M., and Jun, C.-H. (2009b). Group Acceptance Sampling Plans for Truncated Life Tests Based on the Inverse Rayleigh Distribution and Log-logistic Distribution, *Pakistan Journal of Statistics,* 25, pp. 107 - 119.

[18] Aslam, M., Kundu, D., Jun, C.-H., and Ahmad, M. (2011). Time Truncated Group Acceptance Sampling Plans for Generalized Exponential Distribution, *Journal of Testing and Evaluation,* 39, pp. 968 - 976.

[19] Kalaiselvi, S., and Vijayaraghavan, R. (2010). Designing of Bayesian Single Sampling Plans for Weibull-Inverted Gamma Distribution, *Recent Trends in Statistical Research*, Publication Division, M. S. University, Tirunelveli, pp. 123 - 132.

[20] Kalaiselvi, S., Loganathan, A., and Vijayaraghavan, R. (2011). Reliability Sampling Plans under the Conditions of Rayleigh – Maxwell Distribution – A Bayesian Approach, *Recent Advances in Statistics and Computer Applications*, Bharathiar University, Coimbatore, pp. 280 - 283.

[21] Loganathan, A., Vijayaraghavan, R., and Kalaiselvi, S. (2012). Recent Developments in Designing Bayesian Reliability Sampling Plans – An Overview, *New Methodologies in Statistical Research*, Publication Division, M. S. University, Tirunelveli, pp. 61 - 68.

[22] Hong, C. W., Lee, W. C., and Wu, J. W. (2013). Computational Procedure of Performance Assessment of Life-time Index of Products for the Weibull Distribution with the Progressive First-failure Censored Sampling Plan, *Journal of Applied Mathematics,* Article ID 717184, 2012, pp. 1 - 13.

[23] Vijayaraghavan, R., Chandrasekar, K., and Uma, S. (2012). Selection of sampling inspection plans for life test based on Weibull-Poisson Mixed Distribution, *Proceedings of the International Conference on Frontiers of Statistics and its Applications,* Coimbatore, pp. 225 - 232.

[24] Vijayaraghavan, R., and Uma, S. (2012). Evaluation of Sampling Inspection Plans for Life Test Based on Exponential-Poisson Mixed Distribution, *Proceedings of the International Conference on Frontiers of Statistics and its Applications*, Coimbatore, pp. 233 - 240.

[25] Vijayaraghavan, R., and Uma, S. (2016). Selection of Sampling Inspection Plans for Life Tests Based on Lognormal Distribution, *Journal of Testing and Evaluation*, 44, pp. 1960 - 1969.

[26] Vijayaraghavan, R., Sathya Narayana Sharma, K., and Saranya, C. R. (2019a). Evaluation of Sampling Inspection Plans for Life-tests Based on Generalized Gamma Distribution, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 6, pp. 138 – 146.

[27] Vijayaraghavan, R., and Saranya. C.R., Sathya Narayana Sharma, K. (2019b). Life Test Sampling Plans Based on Marshall – Olkin Extended Exponential Distribution, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, 6, pp.131-139.

[28] Saranya, C.R., Vijayaraghavan, R. & Sathya Narayana Sharma, K. (2022). Design of double sampling inspection plans for life tests under time censoring based on Pareto type IV distribution. *Sci. Rep* 12, 7953.

[29] Lomax, K. S. (1954). Business Failures: Another Example of the Analysis of Failure Data, *Journal of the American Statistical Association*, 49, 847 – 852.

[30] Abdul, M. I. B. (2012). Recurrence Relations for Moments of Lower Generalized Order Statistics from Exponentiated Lomax Distribution and Its Characterization, *International Journal of Mathematical Archive* 3, pp. 2144 - 2150.

[31] Lemonte, A. J. and Cordeiro, G. M. (2013). An Extended Lomax Distribution, *Statistics*, 47, 800–816.

[32] Ramos, M. W. A., Marinho, P. R. D., *de* Silva R. V., and Cordeiro, G. M. (2013). The Exponentiated Lomax Poisson Distribution with an Application to Lifetime Data, *Advances and Applications in Statistics*, 34, 107 - 135.

[33] Oguntunde, *et al.*, (2017). A New Generalization of the Lomax Distribution with Increasing, Decreasing, and Constant Failure Rate, *Modelling and Simulation in Engineering*, Article ID 6043169, 2017, pp. 1- 6.

[34] Dodge HF. Chain Sampling Inspection Plans. *Ind. Quality Control.* 1955; 11:10–13.

[35] Govindaraju, K. (1991). Fractional Acceptance Number Single Sampling Plan, *Communications in Statistics – Simulation and Computation,* 20(1): 173 - 190.

[36] Soundararajan, V., and Vijayaraghavan, R. (1992). Sampling Inspection Plans with Desired Discrimination, *IAPQR Transactions*, 17(2): 19 – 24.

[37] Vijayaraghavan, R. (2007). Minimum Size Double Sampling Plans for Large Isolated Lots, *Journal of Applied Statistics*, 34(7): 799 - 806.

A COMPREHENSIVE STUDY OF LENGTH-BIASED TRANSMUTED DISTRIBUTION

Danish Qayoom[1], Aafaq A. Rather[2,*]

[1,2]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[1]danishqayoom11@gmail.com, [2,*]Corresponding author: aafaq7741@gmail.com

Abstract

In this study, we explore a new probability distribution termed as the Length-Biased Transmuted Mukherjee-Islam (LBTMI) distribution. This exploration enhances the conventional Transmuted Mukherjee-Islam distribution by integrating a weighted transformation approach. The paper examines the probability density function and the corresponding cumulative distribution function associated with the LBTMI distribution. A comprehensive examination of the unique structural properties of the proposed model is carried out, including the survival function, conditional survival function, hazard function, cumulative hazard function, mean residual life, moments, moment generating function (MGF), characteristic function (CF), cumulant generating function (CGF), likelihood ratio test, ordered statistics, entropy measures, and Bonferroni and Lorenz curves. To ensure precise estimation of model parameters, the study employs the maximum likelihood estimation method, contributing significantly to the advancement of statistical modelling in this domain.

Key words: Transmuted Mukherjee-Islam distribution, Weighted transformation, Reliability analysis, Maximum likelihood estimator, Ordered statistics

1. Introduction

One fundamental aim of statistics is to develop precise predictive models for real-world phenomena. However, due to the complex nature of these phenomena, conventional modelling approaches may prove inadequate. As a result, an array of probability distributions has been devised to address these challenges by using various transformation approaches. In our study, we introduce a novel extension of Transmuted Mukherjee-Islam distribution termed as LBTMI distribution. This distribution is crafted through the Fisher's [6] weighted transformation technique, initially introduced in 1934, and further elucidated by Rao [17] in 1965. This method allows us to construct weighted models of observations based on predefined weighted functions. 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 [4] and Zelen [26]. Scholars and researchers have extensively delved into weighted probability models and their wide-ranging applications across diverse domains. Modi and Gill [13] discussed the length-biased weighted Maxwell distribution, while Sanat [23] derived the beta-length biased Pareto distribution. Reyad et al. [22] examined the length-biased weighted Frechet distribution, elucidating its properties and practical applications. Rather and Subramanian [20] introduced a method for characterizing and estimating the length-biased weighted generalized uniform distribution. Mudasir and Ahmad [14] provided an in-depth

discussion on the characterization and estimation of the length-biased Nakagami distribution, and Khan et al. [7] discussed the weighted modified Weibull distribution. In subsequent years, Rather and Subramanian [19] explored the length-biased Erlang truncated exponential distribution, highlighting its practical applications. Mathew and Chesneau [12] studied the Marshall-Olkin length-biased Maxwell distribution. In recent developments, Rather and Ozel [18] introduced a new length-biased power Lindley distribution with applications. Al-Omari and Alanzi [1] presented the inverse length-biased Maxwell distribution and conducted statistical inference with an illustrative application. Mustafa and Khan [15] developed the length-biased powered inverse Rayleigh distribution with practical applications.

The Transmuted Mukherjee-Islam distribution had explored by Rather and Subramanian [21] using the quadratic rank transmutation map studied first by Shaw and Buckley [24] in 2007. Loai M. A. Al-Zou'bi [2] also obtained various properties of Transmuted Mukherjee-Islam distribution and its applications. The probability density function of a random variable say Z following Transmuted Mukherjee- Islam distribution with parameters say $(\varepsilon, \nu, \omega)$ is given by

$$f(z;\varepsilon,\nu,\omega)=\frac{\varepsilon}{\nu^{\varepsilon}}z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right);0<z<\nu,\varepsilon>0,\nu>0,-1\le\omega\le 1 \quad (1)$$

And the corresponding cumulative distribution function is

$$F_Z(z)=\left(\frac{z}{\nu}\right)^{\varepsilon}\left(1+\omega-\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right) \quad (2)$$

Several researchers have explored the quadratic transmutation mapping approach, introducing new members to this family across various baseline distributions. These include the transmuted extreme value distribution by Aryal and Tsokos [3], the transmuted Frechet distribution by Mahmoud and Mandouh [10], the transmuted generalized linear exponential distribution by Elbatal et al. [5], the transmuted additive Weibull distribution by Mansour et al. [11], the transmuted Gompertz distribution by Khan et al. [8], and the transmuted generalized inverse Weibull distribution by Khan et al. [9]. Additionally, Subramanian and Rather [25] examined the weighted version of the exponentiated Mukherjee-Islam distribution, deriving its statistical properties. Furthermore, Otiniano et al. [16] delved into the transmuted generalized extreme value distribution.

2. Probability density function (PDF) and cumulative distribution function (CDF)

Using the weighted transformation approach, the PDF $y_w(z)$ of a non-negative random variable Z is given by

$$y_w(z)=\frac{w(z)\,y(z)}{E\big(w(z)\big)};\qquad z>0$$

Where $w(z)$ be a non-negative weight function and

$$E\big(w(z)\big)=\int_{-\infty}^{\infty}w(z)\,y(z)\,dz<\infty.$$

Note that different choices of the weight function $w(z)$ give different weighted distributions. Consequently for weight function $w(z)=z$, the resulting distribution is called length-biased distribution. Let us assume that the PDF of the random variable Z to be Transmuted Mukherjee-Islam distribution, so the PDF of Length-Biased Mukherjee-Islam(LBTMI) distribution is given by

$$g(z;\varepsilon,\nu,\omega)=\frac{z\,f(z;\varepsilon,\nu,\omega)}{E(z)} \quad (3)$$

Now

$$E(z)=\int_0^{\nu} z\,\frac{\varepsilon}{\nu^{\varepsilon}}\,z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)dz \tag{4}$$

$$E(z)=\frac{\varepsilon}{\nu^{\varepsilon}}\left((1+\omega)\int_0^{\nu} z^{\varepsilon}\,dz-\frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\nu} z^{2\varepsilon}\,dz\right) \tag{5}$$

After simplification we get

$$E(z^{s})=\frac{\varepsilon\nu\big((1-\omega)+2\varepsilon\big)}{(\varepsilon+1)(2\varepsilon+1)} \tag{6}$$

Using (1) and (6) in (3) we get

$$g\big(z;\varepsilon,\nu,\omega\big)=\frac{z\,\frac{\varepsilon}{\nu^{\varepsilon}}\,z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\frac{\varepsilon\nu\big((1-\omega)+2\varepsilon\big)}{(\varepsilon+1)(2\varepsilon+1)}} \tag{7}$$

$$g\big(z;\varepsilon,\nu,\omega\big)=\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\big((1-\omega)+2\varepsilon\big)} \tag{8}$$

The corresponding CDF of LBTMI distribution is given by

$$G_Z(z)=\int_0^{z}\left(\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\big((1-\omega)+2\varepsilon\big)}\right)dz \tag{9}$$

$$G_Z(z)=\frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\big((1-\omega)+2\varepsilon\big)}\left((1+\omega)\int_0^{z} z^{\varepsilon}\,dz-\frac{2\,\omega}{(\nu)^{\varepsilon}}\int_0^{z} z^{2\varepsilon}\,dz\right) \tag{10}$$

After simplification we get

$$G_Z(z)=\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}\,z^{\varepsilon+1}-2\omega(\varepsilon+1)\,z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\big((1-\omega)+2\varepsilon\big)} \tag{11}$$

3. Reliability Analysis

3.1 Survival function

The survival function of LBTMI distribution is given by

$$R_T(t)=P_r(T>t)$$

$$R_T(t)=1-P_r(T\le t)$$

$$R_T(t)=1-\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}\,t^{\varepsilon+1}-2\omega(\varepsilon+1)t^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\big((1-\omega)+2\varepsilon\big)} \tag{12}$$

$$R_T(t)=\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}t^{\varepsilon+1}+2\omega(\varepsilon+1)t^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \tag{13}$$

After simplification we get

$$R_T(t)=\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}t^{\varepsilon+1}+2\omega(\varepsilon+1)t^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \tag{14}$$

After simplification we get

$$R_T(t)=\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-t^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)t\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \tag{15}$$

3.2 Conditional survival function

In case of LBTMI distribution the conditional survival function is given by

$$R_T(t\mid t_0)=P_r(T>t_0+t\mid T>t_0)$$

$$R_T(t\mid t_0)=\frac{P_r(T>t_0+t)}{P_r(T>t_0)}$$

$$R_T(t\mid t_0)=\frac{R_T(t_0+t)}{R_T(t_0)}$$

$$R_T(t\mid t_0)=\frac{\dfrac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-(t_0+t)^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)(t_0+t)\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}}{\dfrac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-t^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)t\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}} \tag{16}$$

$$R_T(t\mid t_0)=\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-(t_0+t)^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)(t_0+t)\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-t^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)t\right)} \tag{17}$$

3.3 Hazard function

The hazard function of LBTMI distribution is given by

$$H_T(t)=\frac{g(t;\varepsilon,\nu,\omega)}{1-G_T(t)}$$

$$H_T(t)=\frac{\dfrac{(\varepsilon+1)(2\varepsilon+1)t^{\varepsilon}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}}{1-\dfrac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}t^{\varepsilon+1}-2\omega(\varepsilon+1)t^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}}$$

After simplification we have

$$H_T(t)=\frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^{\varepsilon}t^{\varepsilon}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-t^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)t^{\varepsilon}\right)} \tag{18}$$

3.4 Cumulative hazard function

The cumulative hazard function of LBTMI distribution is given by

$$ {}_C H_T(t) = -\ln\left(R_T(t)\right) $$

$$ {}_C H_T(t) = -\ln\left(\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right) - t^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon} - 2\omega(\varepsilon+1)t\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \right) \tag{19} $$

Similarly, the Conditional Cumulative hazard function of LBTMI distribution is given by

$$ {}_C H_T(t \mid t_0) = -\ln\left(R_T(t \mid t_0)\right) $$

$$ {}_C H_T(t \mid t_0) = -\ln\left(\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right) - (t_0+t)^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon} - 2\omega(\varepsilon+1)(t_0+t)\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right) - t^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon} - 2\omega(\varepsilon+1)t\right)} \right) $$

3.5 Reverse Hazard function

The reverse hazard function of LBTMI distribution is given by

$$ H_r(t) = \frac{g\left(t;\varepsilon,\nu,\omega\right)}{G_T(t)} $$

$$ H_r(t) = \frac{\dfrac{(\varepsilon+1)(2\varepsilon+1)t^{\varepsilon}\left(1+\omega-2\omega\left(\dfrac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}}{\dfrac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}t^{\varepsilon+1} - 2\omega(\varepsilon+1)t^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}} $$

After simplification we get

$$ H_r(t) = \frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^{\varepsilon}t^{\varepsilon}\left(1+\omega-2\omega\left(\dfrac{t}{\nu}\right)^{\varepsilon}\right)}{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}t^{\varepsilon+1} - 2\omega(\varepsilon+1)t^{2\varepsilon+1}} \tag{20} $$

3.6 Mills Ratio

The Mills ratio of LBTMI distribution is given by

$$ Mills\ ratio = \frac{1}{H_r(t)} $$

$$ Mills\ ratio = \frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}t^{\varepsilon+1} - 2\omega(\varepsilon+1)t^{2\varepsilon+1}}{(\varepsilon+1)(2\varepsilon+1)(\nu)^{\varepsilon}t^{\varepsilon}\left(1+\omega-2\omega\left(\dfrac{t}{\nu}\right)^{\varepsilon}\right)} \tag{21} $$

3.7 Mean residual life

The mean residual life (MRL) in case of LBTMI distribution is given by

$$ MRL = \frac{1}{1-G_Z(z)}\int_z^{\nu} t\, g\left(t;\varepsilon,\nu,\omega\right)dt - z $$

$$MRL = \frac{1}{1-G_Z(z)}\int_z^{\nu} t\,\frac{(\varepsilon+1)(2\varepsilon+1)t^{\varepsilon}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}dt - z \tag{22}$$

$$MRL = \frac{1}{1-G_Z(z)}\int_z^{\nu} \frac{(\varepsilon+1)(2\varepsilon+1)t^{\varepsilon+1}\left(1+\omega-2\omega\left(\frac{t}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}dt - z \tag{23}$$

$$MRL = \frac{(\varepsilon+1)(2\varepsilon+1)}{\left(1-G_Z(z)\right)\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_z^{\nu} t^{\varepsilon+1}dt - \frac{2\omega}{(\nu)^{\varepsilon}}\int_z^{\nu} t^{2\varepsilon+1}dt\right) - z \tag{24}$$

After simplification we get

$$MRL = \frac{(\varepsilon+1)(2\varepsilon+1)}{\left(1-G_Z(z)\right)\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)(\varepsilon+2)(2\varepsilon+2)}\left\{\begin{matrix}(\nu)^{2\varepsilon+2}\left((1-\omega)(\varepsilon+2)+(1+\omega)\varepsilon\right)+\\ 2\omega(\varepsilon+2)z^{2\varepsilon+2}-(\nu)^{\varepsilon}(1+\omega)(2\varepsilon+2)z^{\varepsilon+2}\end{matrix}\right\} - z$$

4. Moments

The rth raw moment about origin of LBTMI distributionis defined as

$$\mu_r' = \int_0^{\nu} z^r\, g(z;\varepsilon,\nu,\omega)dz$$

$$\mu_r' = \int_0^{\nu} z^r\,\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}dz$$

$$\mu_r' = \frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_0^{\nu} z^{\varepsilon+r}dz - \frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\nu} z^{2\varepsilon+r}dz\right) \tag{25}$$

$$\mu_r' = \frac{(1+\varepsilon)(1+2\varepsilon)}{\nu^{1+\varepsilon}\left((1-\omega)+2\varepsilon\right)}\left((1+\omega)\frac{(\nu)^{\varepsilon+r+1}}{\varepsilon+r+1} - \frac{2\omega(\nu)^{2\varepsilon+r+1}}{(\nu)^{\varepsilon}(2\varepsilon+r+1)}\right) \tag{26}$$

After simplification we get

$$\mu_r' = \frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^r\left((1-\omega)(\varepsilon+r+1)+(1+\omega)\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+r+1)(2\varepsilon+r+1)} \tag{27}$$

Putting $r = 1, 2, 3, 4$ in (26) we get

$$\mu_1' = \frac{(\varepsilon+1)(2\varepsilon+1)(\nu)\left((1-\omega)(\varepsilon+2)+(1+\omega)\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+2)(2\varepsilon+2)} \tag{28}$$

$$\mu_2' = \frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^2\left((1-\omega)(\varepsilon+3)+(1+\omega)\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+3)(2\varepsilon+3)} \tag{29}$$

$$\mu_3' = \frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^3\left((1-\omega)(\varepsilon+4)+(1+\omega)\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+4)(2\varepsilon+4)} \tag{30}$$

$$\mu_4' = \frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^4\left((1-\omega)(\varepsilon+5)+(1+\omega)\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+5)(2\varepsilon+5)} \tag{31}$$

The variance and coefficient of variance $(C.V)$ respectively are given by

$$\sigma^2 = \mu_2' - (\mu_1')^2$$

And

$$C.V = \frac{\sigma}{\mu_1'}; \qquad \text{where } \sigma = \sqrt{\mu_2' - (\mu_1')^2}$$

5. Harmonic mean

The harmonic mean of LBTMI distributionis defined as

$$Harmonic\ \ mean = E\left(\frac{1}{Z}\right)$$

$$Harmonic\ mean = \int_0^{\nu} \frac{1}{z} \frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} dz$$

$$Harmonic\ mean = \int_0^{\nu} \frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon-1}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} dz \tag{32}$$

$$Harmonic\ mean = \frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_0^{\nu} z^{\varepsilon-1}dz - \frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\nu} z^{2\varepsilon-1}dz\right) \tag{33}$$

After simplification we get

$$Harmonic\ mean = \frac{(\varepsilon+1)(2\varepsilon+1)\left((1-\omega)\varepsilon+(1+\omega)\right)}{\nu\left((1-\omega)+2\varepsilon\right)(2\varepsilon^2)} \tag{34}$$

6. MGF, CF, and CGF

The moment generating function of LBTMI distribution is

$$M_Z(t) = E(e^{tz})$$

$$M_Z(t) = \int_0^{\nu} e^{tz} \frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} dz$$

$$M_Z(t) = \int_0^{\nu}\sum_{k=0}^{\infty}\frac{(tz)^k}{k!} \frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} dz \tag{35}$$

$$M_Z(t) = \sum_{k=0}^{\infty}\frac{(t)^k}{k!}\int_0^{\infty} z^k g(z;\varepsilon,\nu,\omega)dz$$

$$M_Z(t)=\sum_{k=0}^{\infty}\frac{(t)^k}{k!}\mu_k'$$

$$M_Z(t)=\sum_{k=0}^{\infty}\frac{(t)^k}{k!}\frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^k\left((1-\omega)(\varepsilon+k+1)+(1+\omega)\,\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+k+1)(2\varepsilon+k+1)} \tag{36}$$

The characteristics function of LBTMI distribution is

$$\phi_Z(t)=E(e^{itz})$$

$$\phi_Z(t)=\int_0^{\nu}e^{\iota t z}\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}dz$$

$$\phi_Z(t)=\int_0^{\nu}\sum_{k=0}^{\infty}\frac{(\iota tz)^k}{k!}\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}dz \tag{37}$$

$$\phi_Z(t)=\sum_{k=0}^{\infty}\frac{(\iota)^k(t)^k}{k!}\int_0^{\infty}z^k g(z;\varepsilon,\nu,\omega)dz \tag{38}$$

$$\phi_Z(t)=\sum_{k=0}^{\infty}\frac{(\iota\,t)^k}{k!}\mu_k'$$

$$\phi_Z(t)=\sum_{k=0}^{\infty}\frac{(\iota\,t)^k}{k!}\frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^k\left((1-\omega)(\varepsilon+k+1)+(1+\omega)\,\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+k+1)(2\varepsilon+k+1)} \tag{39}$$

The cumulant generating function of LBTMI distribution is

$$\kappa_Z(t)=\log\left(M_Z(t)\right)$$

$$\kappa_Z(t)=\log\left(\sum_{k=0}^{\infty}\frac{(t)^k}{k!}\frac{(\varepsilon+1)(2\varepsilon+1)(\nu)^k\left((1-\omega)(\varepsilon+k+1)+(1+\omega)\,\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+k+1)(2\varepsilon+k+1)}\right) \tag{40}$$

7. Estimation of Parameters

Let $z_1, z_2, z_3, \ldots, z_n$ be a random sample of size n from LBTMI distribution. Then the likelihood function is defined as the joint density of the random sample, which is given as

$$L(\varepsilon,\nu,\omega)=\prod_{l=1}^{n}g(z_l;\varepsilon,\nu,\omega)=\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \tag{41}$$

$$L(\varepsilon,\nu,\omega)=\prod_{l=1}^{n}\frac{(\varepsilon+1)(2\varepsilon+1)z_l^{\varepsilon 1}\left(1+\omega-2\omega\left(\frac{z_l}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \tag{42}$$

$$L(\varepsilon,\nu,\omega)=\frac{(\varepsilon+1)^n(2\varepsilon+1)^n}{\nu^{n(\varepsilon+1)}\left((1-\omega)+2\varepsilon\right)^n}\left(\prod_{l=1}^{n}z_l^{\varepsilon}\right)\left(\prod_{l=1}^{n}\left(1+\omega-2\omega\left(\frac{z_l}{\nu}\right)^{\varepsilon}\right)\right) \tag{43}$$

Taking logarithm on both sides we get

$$\log L(\varepsilon,\nu,\omega)=n\log(\varepsilon+1)+n\log(2\varepsilon+1)-n(\varepsilon+1)\log(\nu)-n\log\left((1-\omega)+2\varepsilon\right)$$
$$+(\varepsilon)\sum_{l=1}^{n}\log z_l+\sum_{l=1}^{n}\log\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right) \tag{44}$$

Differentiating equation (43) partially with respect to ε and equating to zero we get

$$\frac{n}{\varepsilon+1}+\frac{2n}{2\varepsilon+1}-n\log(\nu)-\frac{2n}{(1-\omega)+2\varepsilon}+\sum_{l=1}^{n}\log z_l-\sum_{l=1}^{n}\frac{2\omega}{\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right)}\left(\frac{z}{\nu}\right)^{\varepsilon}\log\left(\frac{z}{\nu}\right)=0 \tag{45}$$

Differentiating equation (43) partially with respect to ν and equating to zero we get

$$\sum_{l=1}^{n}\frac{2\omega(z_l)\varepsilon}{\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right)(\nu)^{\varepsilon+1}}-\frac{n(\varepsilon+1)}{\nu}=0 \tag{46}$$

Differentiating equation (43) partially with respect to ω and equating to zero we get

$$\frac{n}{(1-\omega)+2\varepsilon}+\sum_{l=1}^{n}\frac{1-2\left(\frac{z_l}{\nu}\right)^{\varepsilon}}{\left(1+\omega-\frac{2\omega}{(\nu)^{\varepsilon}}(z_l)^{\varepsilon}\right)}=0 \tag{47}$$

Simultaneously solving equation (44), (45), and (46), gives the maximum likelihood estimators of parameters involved in the given distribution. However, direct evaluation of the aforementioned system of nonlinear equations is unfeasible. To obtain maximum likelihood estimates for the distribution parameters, it is necessary to employ iterative methods such as the Newton-Raphson method, Mathematica, or the Secant method to solve this system effectively.

8. Distribution of ordered statistics

Suppose we draw a random sample $z_1, z_2, z_3, ..., z_n$ of size n from LBTMI distribution. Then the ordered statistics corresponding to the given sample is $Z_{(1)}, Z_{(2)}, Z_{(3)}, ..., Z_{(n)}$ such that $Z_{(1)} \le Z_{(2)} \le Z_{(3)} \le ... \le Z_{(n)}$, Where

$$Z_{(1)}=\min(z_1, z_2, z_3, ..., z_n)$$

and $$Z_{(n)}=\max(z_1, z_2, z_3, ..., z_n)$$

The PDF of k^{th} ordered statistics from LBTMI distribution is given by

$$g_{Z_{(k)}}(z)=\frac{n!}{(k-1)!(n-k)!}g(z;\varepsilon,\nu,\omega)(G_Z(z))^{k-1}(1-G_Z(z))^{n-k}$$

$$g_{Z_{(k)}}(z)=\frac{n!}{(k-1)!(n-k)!}\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left(\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{k-1} \tag{48}$$
$$\times\left(1-\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{n-k}$$

$$g_{Z_{(k)}}(z)=\frac{n!}{(k-1)!(n-k)!}\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left(\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{k-1} \tag{49}$$
$$\times\left(\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-z^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)z^{\varepsilon}\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{n-k}$$

And the corresponding CDF of k^{th} ordered statistics is

$$G_{Z_{(k)}}(z)=\sum_{j=k}^{n}({}^{n}C_{j})\left(G_{Z}(z)\right)^{j}\left(1-G_{Z}(z)\right)^{n-j}$$

$$G_{Z_{(k)}}(z)=\sum_{j=k}^{n}\left(\begin{array}{l}({}^{n}C_{j})\left(\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{j}\\ \times\left(\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-z^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)z^{\varepsilon}\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{n-j}\end{array}\right) \tag{50}$$

On substituting $k=1,n$ in equation (48) we get the PDF of smallest and highest ordered statistics respectively and are given as

$$g_{Z_{(1)}}(z)=n\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\times\left(\frac{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)-z^{\varepsilon+1}\left((2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}-2\omega(\varepsilon+1)z^{\varepsilon}\right)}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{n-1} \tag{51}$$

And

$$g_{Z_{(n)}}(z)=\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left(\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{n-1} \tag{52}$$

Their corresponding CDFs are obtained on substituting $k=1,n$ in equation (49) and are given by

$$G_{Z_{(1)}}(z)=1-\left(1-\left(\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)\right)^{n} \tag{53}$$

And

$$G_{Z_{(n)}}(z)=\left(\frac{(2\varepsilon+1)(1+\omega)(\nu)^{\varepsilon}z^{\varepsilon+1}-2\omega(\varepsilon+1)z^{2\varepsilon+1}}{\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{n} \tag{54}$$

9. Likelihood ratio test

The likelihood ratio test is a statistical technique designed to explore the adequacy of fit between two models. Specifically, in the realm of probability distributions, it is utilized to check the suitability of two distinct distributions in explaining observed data and its purpose is to ascertain whether the inclusion of additional parameters in a statistical model substantially enhances its ability to accurately represent the data or not. Suppose $z_1, z_2, z_3, ..., z_n$ be a random sample of size n from LBTMI distribution.. To test the hypothesis

$$H_0 : g(z) = f(z;\varepsilon,\nu,\omega,) \qquad \text{against} \qquad H_1 : g(z) = g(z;\varepsilon,\nu,\omega)$$

The likelihood ratio test is defined as

$$\ell = \prod_{k=1}^{n} \frac{g(z_k;\varepsilon,\nu,\omega)}{f(z_k;\varepsilon,\nu,\omega)}$$

$$\ell = \prod_{k=1}^{n} \frac{\dfrac{(\varepsilon+1)(2\varepsilon+1)z_k{}^{\varepsilon}\left(1+\omega-2\omega\left(\dfrac{z_k}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}}{\dfrac{\varepsilon}{\nu^{\varepsilon}} z_k{}^{\varepsilon-1}\left(1+\omega-2\omega\left(\dfrac{z_k}{\nu}\right)^{\varepsilon}\right)} \tag{55}$$

$$\ell = \left(\frac{(\varepsilon+1)(2\varepsilon+1)}{\varepsilon\nu\left((1-\omega)+2\varepsilon\right)}\right)^{n} \prod_{k=1}^{n} z_k \tag{56}$$

So, we reject null hypothesis at α level of significance if $\ell > \mathrm{K}^*$ such that $P(\ell > \mathrm{K}^*) = \alpha$, where K^* is the critical value at α level of significance of the given test statistics. That is,

$$\left(\frac{(\varepsilon+1)(2\varepsilon+1)}{\varepsilon\nu\left((1-\omega)+2\varepsilon\right)}\right)^{n} \prod_{k=1}^{n} z_k > \mathrm{K}^* \tag{57}$$

$$\prod_{k=1}^{n} z_k > \mathrm{K}^* \left(\frac{\varepsilon\nu\left((1-\omega)+2\varepsilon\right)}{(\varepsilon+1)(2\varepsilon+1)}\right)^{n} \tag{58}$$

For large sample size n, $-2\log(\ell)$ is distributed as Chi-square distribution with one degree of freedom. Also p-value is calculated from the chi-square distribution. On the basis of p-value, we reject the null hypothesis when the p-value is less than level of significance.

10. Entropy measures

10.1 Renyi entropy and Tsallis entropy

By definition, the Renyi entropy is given by

$$R(\tau)=\frac{1}{1-\tau}\log\left(\int_0^{\nu}\left(g\left(z_k;\varepsilon,\nu,\omega\right)\right)^{\tau}dz\right) \tag{59}$$

$$R(\tau)=\frac{1}{1-\tau}\log\left(\int_0^{\nu}\left(\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{\tau}dz\right) \tag{60}$$

$$R(\tau)=\frac{1}{1-\tau}\log\left(\left(\frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{\tau}\int_0^{\nu}\left(z^{\tau\varepsilon}\sum_{k=0}^{\tau}\left({}^{\tau}C_k\right)(-1)^k(1+\omega)^{\tau-k}\left(2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)^k\right)dz\right)$$

$$R(\tau)=\frac{1}{1-\tau}\log\left(\left(\frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\right)^{\tau}\sum_{k=0}^{\tau}\left({}^{\tau}C_k\right)(-1)^k(1+\omega)^{\tau-k}\left(\frac{2\omega}{(\nu)^{\varepsilon}}\right)^k\int_0^{\nu}z^{\varepsilon(k+\tau)}dz\right) \tag{61}$$

After simplification we get

$$R(\tau)=\frac{1}{1-\tau}\log\left(\frac{\left((\varepsilon+1)(2\varepsilon+1)\right)^{\tau}(\nu)^{1-\tau}}{\left((1-\omega)+2\varepsilon\right)^{\tau}}\sum_{k=0}^{\tau}\frac{\left({}^{\tau}C_k\right)(-1)^k(1+\omega)^{\tau-k}(2\omega)^k}{\varepsilon(k+\tau)+1}\right) \tag{62}$$

Similarly, the Tsallis entropy associated with the given distribution is given by

$$T_s(\xi)=\frac{1}{\xi-1}\left(1-\int_0^{\nu}\left(g\left(z_k;\varepsilon,\nu,\omega\right)\right)^{\xi}dz\right)$$

$$T_s(\xi)=\frac{1}{\xi-1}\left(1-\frac{\left((\varepsilon+1)(2\varepsilon+1)\right)^{\xi}(\nu)^{1-\xi}}{\left((1-\omega)+2\varepsilon\right)^{\xi}}\sum_{k=0}^{\xi}\frac{\left({}^{\xi}C_k\right)(-1)^k(1+\omega)^{\xi-k}(2\omega)^k}{\varepsilon(k+\xi)+1}\right) \tag{63}$$

11. Bonferroni and Lorenz curves

The Bonferroni curve of the given distribution is given by

$$\Psi(\zeta)=\frac{1}{\zeta\,\mu_1'}\int_0^{\varphi}z\,g\left(z;\varepsilon,\nu,\omega\right)dz$$

Where $\mu_1'=\dfrac{(\varepsilon+1)(2\varepsilon+1)(\nu)\left((1-\omega)(\varepsilon+2)+(1+\omega)\,\varepsilon\right)}{\left((1-\omega)+2\varepsilon\right)(\varepsilon+2)(2\varepsilon+2)}$ and $\varphi=F^{-1}(\zeta)$

$$\Psi(\zeta)=\frac{1}{\zeta\,\mu_1'}\int_0^{\varphi}z\,\frac{(\varepsilon+1)(2\varepsilon+1)z^{\varepsilon}\left(1+\omega-2\omega\left(\frac{z}{\nu}\right)^{\varepsilon}\right)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)} \tag{64}$$

$$\Psi(\zeta)=\frac{1}{\zeta\,\mu_1'}\frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left((1+\omega)\int_0^{\varphi}z^{\varepsilon+1}dz-\frac{2\omega}{(\nu)^{\varepsilon}}\int_0^{\varphi}z^{2\varepsilon+1}dz\right) \tag{65}$$

$$\Psi(\zeta)=\frac{1}{\zeta\,\mu_1'}\frac{(\varepsilon+1)(2\varepsilon+1)}{\nu^{\varepsilon+1}\left((1-\omega)+2\varepsilon\right)}\left((1+\omega)\left(\frac{(\varphi)^{\varepsilon+2}}{\varepsilon+2}\right)-\frac{2\omega}{(\nu)^{\varepsilon}}\left(\frac{(\varphi)^{2\varepsilon+2}}{2\varepsilon+2}\right)\right) \tag{66}$$

After simplification we get

$$\Psi(\zeta)=\frac{(\varepsilon+1)(2\varepsilon+1)(\varphi)^{\varepsilon+2}\left((1+\omega)(2\varepsilon+2)(\nu)^{\varepsilon}-2\omega(\varepsilon+2)(\varphi)^{\varepsilon}\right)}{\zeta\,\mu_1'\nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)(\varepsilon+2)(2\varepsilon+2)} \tag{67}$$

Also, the Lorenz curve of the given distribution is given by

$$\Phi(\zeta) = \zeta\, \Psi(\zeta)$$

$$\Phi(\zeta) = \zeta\left(\frac{(\varepsilon+1)(2\varepsilon+1)(\varphi)^{\varepsilon+2}\left((1+\omega)(2\varepsilon+2)(\nu)^{\varepsilon} - 2\omega(\varepsilon+2)(\varphi)^{\varepsilon}\right)}{\zeta\ \mu_1' \nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)(\varepsilon+2)(2\varepsilon+2)} \right) \tag{68}$$

$$\Phi(\zeta) = \frac{(\varepsilon+1)(2\varepsilon+1)(\varphi)^{\varepsilon+2}\left((1+\omega)(2\varepsilon+2)(\nu)^{\varepsilon} - 2\omega(\varepsilon+2)(\varphi)^{\varepsilon}\right)}{\mu_1' \nu^{2\varepsilon+1}\left((1-\omega)+2\varepsilon\right)(\varepsilon+2)(2\varepsilon+2)} \tag{69}$$

12. Conclusion

In this paper, we have introduced a novel extension of the Transmuted Mukherjee-Islam distribution. This extension incorporates a weighted transformation approach and the existing three-parameter Transmuted Mukherjee-Islam distribution and generates four parametric innovative model known as the Length-Biased Transmuted Mukherjee-Islam distribution. We conduct a thorough analysis of the Length-Biased Transmuted Mukherjee-Islam distribution, investigating its mathematical formulation and statistical properties in detail. Parameter estimation for this new distribution is performed using maximum likelihood estimation techniques. Additionally, to assess the goodness of fit between these two models, we employ the likelihood ratio test.

References

[1] Al-Omari, A. I. and. Alanzi, A. R. A. (2021). Inverse length biased Maxwell distribution: Statistical inference with an application, *Computer Systems Science & Engineering*, 39(1), 147-164.

[2] Al-Zou'bi, L. M. (2017). Transmuted Mukherjee-Islam Distribution: A Generalization of Mukherjee-Islam Distribution, *Journal of Mathematics Research*, 9(4), 135-144. https://doi.org/10.5539/jmr.v9n4p135

[3] Aryal, G. R. and Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods and Applications*, 71:1401–1407, doi:10.1016/j.na.2009.01.168.

[4] Cox D. R. (1969). Some sampling problems in technology, In New Development in Survey Sampling, Johnson, N. L. and Smith, H., *Jr.(eds.) New York Wiley Interscience*, 506-527.

[5] Elbatal, I., Diab, L. S., and Alim, N. A. A. (2013). Transmuted generalized linear exponential distribution. *International Journal of Computer Applications*, 83:29–37.

[6] Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, 6, 13- 25.

[7] Khan, M. N., Saeed, A., & Alzaatreh, A. (2018). Weighted Modified Weibull distribution. *Journal of Testing and Evaluation*, 47(5), 20170370.

[8] Khan, M. S., King, R., and Hudson, I. L. (2016a). Transmuted Gompertz distribution: Properties and estimation. *Pak. J. Statist*., 32:161–182.

[9] Khan, M. S., King, R., and Hudson, I. L. (2017b). Transmuted new generalized inverse Weibull distribution. *Pak.j.stat.oper.res.*, 13:277–296, doi:10.18187/pjsor.v13i2.1523.

[10] Mahmoud, M. R. and Mandouh, R. M. (2013). On the transmuted frechet distribution. *Journal of Applied Sciences Research*, 9:5553–5561.

[11] Mansour, M. M., Elrazik, E. M. B., Hamed, M. S., and Mohamed, S. M. (2015). A new transmuted additive Weibull distribution: Based on a new method for adding a parameter to a family of distribution. *International Journal of Applied Mathematical Sciences*, 8:31–54.

[12] Mathew, J. &Chesneau, C. (2020), Marshall–olkin length-biased Maxwell distribution and its applications, *Mathematical and Computational Applications*, 25(4), 65. doi:10.3390/mca25040065.

[13] Modi, K. and Gill, V., (2015), Length-biased Weighted Maxwell Distribution, *Pak.j.stat.oper.res.* Vol.XI No.4, pp465-472.

[14] Mudasir, S. & Ahmad, S. P., (2018), Characterization and estimation of length biased nakagami distribution, *Pak.j.stat.oper.res.* Vol.698 XIV No.3, pp697-715.

[15] Mustafa A. and Khan, M. I. (2022). The length-biased powered inverse Rayleigh distribution with applications, *J. Appl. Math. & Informatics*, 40(1-2), 1-13.

[16] Otiniano, C. E. G., de Paiva, B. S., Daniele, S. B., and Neto, M. (2019). The transmuted generalized extreme value distribution: properties and application. *Communications for Statistical Applications and Methods*, 26:239–259.

[17] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment, in classical and Contagious Discrete, G.P. Patiled; *Pergamum Press and Statistical Publishing Society, Calcutta*. 320-332.

[18] 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.

[19] 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.

[20] 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,* Vol.5, Issue.5, pp.72-76.

[21] Rather, A. A., and Subramanian, C., (2018), Transmuted Mukherjee-Islam failure model, *Journal of Statistics Applications & Probability, 7(2),* 343-347.

[22] Reyad, M. H., Hashish, M. A., Othman, A. S. & Allam, A. S. (2017), The length-biased weighted frechet distribution: properties and estimation, *International journal of statistics and applied mathematics*, 3(1), pp 189-200.

[23] Sanat, P. (2016). Beta-length biased Pareto distribution and its properties, *Journal of Emerging Technologies and Innovative Research (JETIR)*, 3(6), 553-557.

[24] Shaw, W.T. and Buckley, I.R.C. (2007). The alchemy of probability distributions: beyond gram-charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. Research report.

[25] Subramanian C. & Rather, A. A. (2018). Weighted exponentiated mukherjee-islam distribution, *international journal of management, technology and engineering*, vol 8, issue XI, pp. 1328-1339.

[26] Zelen, M. (1974). Problems in cell kinetic and the early detection of disease, in Reliability and Biometry, *F. Proschan & R. J. Sering, eds, SIAM, Philadelphia,* 701-706.

THE EXPECTED FISHER INFORMATION MATRIX OF POISSON HALF LOGISTIC MODEL

Ibrahim Abdullahi[1], Aminu Suleiman Mohammed[2*] and Sani Musa[3]

•

[1]Department of Mathematics, Yusuf Maitama Sule University, Kano-Nigeria.
[2]Department of Statistics, Ahmadu Bello University, Zaria.
[3]Department of Mathematics and Computer Science, Sule Lamido, University Kafin-Hausa, Jigawa, Nigeria

ibraabdul@googlemail.com
mohammedas@abu.edu.ng
musasani1010@gmail.com

*Corresponding author: Email: aminusmohammed@gmail.com

Abstract

This study delves into the computation and evaluation of the expected Fisher information matrix within the context of the Poisson-type I half logistic (PHL) distribution. Leveraging confidence intervals and their associated coverage probabilities, our investigation aimed to study the performance of information matrix by the maximum likelihood method in estimating parameters. Our results unveiled a consistent trend: as the sample size expanded, a reduction in the length of the confidence interval was observed, and the 95% asymptotic confidence interval's coverage probability aligned within the expected nominal size. This serves as a testament to the accuracy and robustness of the information matrix's performance within the PHL distribution framework. Also, tested using some real data set.

Keywords: Poisson half logistic, Maximum likelihood estimation, Fisher information matrix, confidence interval.

1. Introduction

The information matrix in maximum likelihood estimation is crucial as it quantifies the precision of parameter estimates, aiding in the construction of confidence intervals. It reflects the inverse of the variance-covariance matrix of the score function, providing insights into the asymptotic behavior of the maximum likelihood estimator.

Confidence intervals, a key component of statistical inferences, leverage the information matrix to quantify the uncertainty surrounding parameter estimates. They offer a range of plausible values for the parameters, enhancing the interpretability and reliability of statistical analyses. In essence, the information matrix and confidence intervals together form integral tools for understanding the robustness and precision of maximum likelihood estimates in statistical inference.

In engineering, the information matrix is crucial for assessing the precision of parameter estimates in various models. For example, in structural engineering, when estimating parameters related to material properties or structural components, the information matrix helps engineers

understand how well their estimates capture the underlying characteristics of the system. This is vital for designing structures with optimal safety margins. In social science, the information matrix is essential for understanding the reliability of parameter estimates in models describing human behavior or societal trends. For instance, in economics, when estimating the coefficients of a model describing consumer behavior, the information matrix helps economists gauge the precision of their estimates, informing policy decisions.

There have been several contributions in the literature regarding the applications of the information matrix. for example, [1] provided a discussion on deriving the information matrix for a logistic distribution. [2] derived the asymptotic expansions of the information matrix test statistic. Small-sample performance of the information matrix test was discussed by [3]. The performance evaluation of track fusion with information matrix filter was studied by [4]. The approximate Fisher information matrix to characterize the training of deep neural networks was used by [5]. The Fisher information matrix in gravitational-wave data analysis was extended by [6]. The general expressions for the quantum Fisher information matrix with its applications to discrete quantum imaging was provided by [7].

The rest of the paper follows: Section 2, discussed about Fisher information matrix. Section 3, provided the expected Fisher information matrix of Poisson-type I half logistic (PHL), and some simulation studies with real data example. Section 4, is the conclusions.

2. On the Fisher information matrix

This section elucidates the significance of Fisher information by delving into fundamental concepts in statistics, including unbiased estimators, the information inequality (Cramer-Rao inequality), and the asymptotic normality of maximum likelihood estimation (MLE). For details (see, [8]).

2.1. Asymptotic characteristics of MLE

To comprehend the significance of Fisher information, this section elucidates fundamental statistics concepts, including unbiased estimators, the Cramer-Rao inequality (information inequality), and the asymptotic normality of MLE.

Estimation in statistics involves mapping real values from observed data. Various methods exist for estimation. Let $\xi(\chi)$ be an estimator where χ represents the observation pattern. For instance, a constant function that maps a specific value, irrespective of observed data, can serve as an estimator. Evaluation of estimation methods is crucial.

Mean Square Error (MSE) is one criterion for evaluation. Assuming a random variable x_i is generated by a distribution with a probability density function $P(x|\xi^*)$ and true parameter value ξ^*, MSE is defined as;

$$MSE = E\lfloor (\xi(\chi) - \xi^*)^2 \rfloor \quad (1)$$

Where $E(X) = \int_x xp(x)dx$ and $Var(X) = \int_x x\left(x - E(x)\right)^2 p(x)dx$. MSE can be decomposed into the variance of the estimator and the square of the bias between the expectation of the estimator and the true parameter value.

Focusing on unbiased estimators with zero bias $\left(E(\xi(\chi) - \xi^*) = 0\right)$, the variance of the estimator becomes crucial. The Cramer-Rao inequality establishes a lower bound for the variance of an unbiased estimator $\hat{\xi}$:

$$Var(\hat{\xi}) \geq \frac{1}{F(\xi^*)} \tag{2}$$

where is the Fisher information. An efficient estimator achieves this lower bound. For unidimensional parameters, the inverse of the Fisher information sets a bound on the estimator's lower variance.

Returning to the Maximum Likelihood (ML) estimator, ML exhibits desirable properties, including asymptotic normality and asymptotic efficiency. In regularity conditions, the ML estimator $\hat{\xi}$ satisfies:

$$\hat{\xi} \to N\left(\xi^*, 1/F(\xi^*)\right) \tag{3}$$

This implies that the ML estimator asymptotically follows a normal distribution with the mean being the true parameter value and the variance (covariance matrix) being the inverse of the Fisher information.

In summary, the asymptotic efficiency of the ML estimator, characterized by the variance being the inverse of the Fisher information, makes it the best choice from the Mean Square Error perspective, given the restriction to unbiased estimators. This property, known as asymptotic efficiency, allows psychology researchers to optimize not only the estimation method but also experimental design and stimuli for variance reduction (increasing Fisher information) in their studies.

2.2 Definition of the Fisher Information Matrix

Let $\xi = (\xi_1, ..., \xi_k)$ represent k-dimensional parameters. The Fisher information matrix for the i-th participant (or trial) concerning parameter ξ is defined as

$$F_i(\xi) = E\left[\left(\frac{\delta}{\delta\xi}\log L(\xi|y_i)\right)\left(\frac{\delta}{\delta\xi}\log L(\xi|y_i)\right)^T\right] \tag{4}$$

where $\frac{\delta}{\delta\xi}\log L(\xi|y_i)$ is a $k\times 1$ column vector, and T denotes the transpose operation. In other words, $F_i(\xi)$ is a $k\times k$ matrix. The (m, n) element of the Fisher information matrix is given by $F_i(\xi)_{(m,n)} = E\left[\frac{\delta}{\delta\xi_m}\log L(\xi|y_i)\frac{\delta}{\delta\xi_n}\log L(\xi|y_i)\right]$. The expectation is over dependent variables y, assuming the model $P(y_i)$ is true. The Fisher information depends on parameter values ξ and stimuli (as well as the model). Although conventionally, the stimuli symbol is omitted in the Fisher information matrix representation, it's important to note that the Fisher information is dependent on experimental design and stimuli.

Additionally, when the true model is known, the following equation holds:

$$F_i(\xi) = E\left[\left(\frac{\delta}{\delta\xi}\log L(\xi|y_i)\right)\left(\frac{\delta}{\delta\xi}\log L(\xi|y_i)\right)^T\right] = -E\left(\frac{\delta^2}{\delta\xi^2}\log L(\xi|y_i)\right) \tag{5}$$

This means that researchers can calculate the Fisher information matrix using either the square of the score function or the second derivatives of the log-likelihood function. The choice between methods depends on the characteristics of the models. In the definition using second derivatives, the (m, n) element of the Fisher information matrix is given by

$$F_i(\xi)_{(m,n)} = -E\left(\frac{\delta^2}{\delta\xi_m\delta\xi_n}\log L(\xi|y_i)\right) \tag{6}$$

3. The Fisher Information Matrix of PHL

Here, we derive the expected Fisher information matrix of the Poisson half logistic distribution, and applied it to study the confidence interval of the maximum likelihood estimators using simulation studies and real data example.

3.1 On the PHL

The Poisson half logistic (PHL) distribution was introduced by [9], using the convolution of half logistic (HL) and Poisson distributions. The PHL was applied to right censored data in [9]. The probability density, and cumulative distribution are respectively given by;

$$f(x)=\frac{2\alpha\lambda e^{-\alpha x}e^{\lambda\left(\frac{1-e^{-\alpha x}}{1+e^{-\alpha x}}\right)}}{\left(e^{\lambda}-1\right)\left(1+e^{-\alpha x}\right)^{2}} \tag{7}$$

and

$$F(x)=\frac{e^{\lambda\left(\frac{1-e^{-\alpha x}}{1+e^{-\alpha x}}\right)}-1}{e^{\lambda}-1} \tag{8}$$

Where $\alpha>0$ and $\lambda\in R-\{0\}$.

The quantile function of the PHL distribution can be used to obtain a random data distributed according to CHLP(α, λ): if U is a uniform (0, 1), then,

$$X=-\frac{1}{\alpha}\left\{\ln\left[1-\frac{\ln\left(u\left(e^{\lambda}-1\right)+1\right)}{\lambda}\right]-\ln\left[1+\frac{\ln\left(u\left(e^{\lambda}-1\right)+1\right)}{\lambda}\right]\right\} \tag{9}$$

is a random variable distributed PHL.

There have been several contributions regarding the HL distribution, one can see, complementary Poisson generalized half logistic [10], generalized half-logistic Poisson [11], extension of the generalized half logistic [12], estimation of the reliability of a stress‐strength system from Poisson half logistic distribution [13], type I half-logistic family [14], new extended cosine generalized half logistic [15], for more details see, [16].

3.2 The expected Fisher Information Matrix of PHL

The maximum likelihood of α and λ can be obtain numerically by simultaneously solving (11) and (12) when set equal to zero using mathematical packages such as *nlminb* in R-software. Let a vector of parameter be $\Theta=\left(\alpha,\lambda\right)^{T}$, then the total log likelihood function of the PHL is given by

$$\begin{aligned}log(l(\alpha\lambda))&=n\log 2+n\log\lambda+n\log\alpha-\alpha\sum_{i=1}^{n}x_i-n\log\left(e^{\lambda}-1\right)\\&-2\sum_{i=1}^{n}\left(1+e^{-\alpha x_i}\right)+\lambda\sum_{i=1}^{n}\left(\frac{\left(1-e^{-\alpha x_i}\right)}{\left(1+e^{-\alpha x_i}\right)}\right)\end{aligned} \tag{10}$$

The first partial derivative of $log(l(\alpha\lambda))$ that is $\partial\log\ell(\alpha,\lambda)/\partial\alpha$ and $\partial\log\ell(\alpha,\lambda)/\partial\lambda$ are computed as;

$$\frac{\delta log(l(\alpha,\lambda))}{\delta\alpha}=\frac{n}{\alpha}-\sum_{i=1}^{n}x_i-2\sum_{i=1}^{n}\frac{x_i e^{-\alpha x_i}}{1+e^{-\alpha x_i}}+2\lambda\sum_{i=1}^{n}\frac{x_i e^{-\alpha x_i}}{\left(1+e^{-\alpha x_i}\right)^{2}} \tag{11}$$

$$\frac{\delta log(l(\alpha,\lambda))}{\delta\lambda}=\frac{n}{\lambda}-\frac{ne^{\lambda}}{e^{-\alpha x_i}-1}-\sum_{i=1}^{n}\left(\frac{1-e^{-\alpha x_i}}{1+e^{-\alpha x_i}}\right) \tag{12}$$

For a very large sample we apply the usual approximation that the MLEs of the CHLP can be approximated as bivariate normal with mean zero and variance covariance matrix $I^{-1}(\alpha,\lambda)$, where $I(\alpha,\lambda)$ is the expected information matrix. Alternatively, we can use $J^{-1}(\alpha,\lambda)$) evaluated at $\widehat{\alpha}$ and $\widehat{\alpha}$ to construct the asymptotic variance - covariance matrix of the MLEs.
Where,

$$J(\alpha,\lambda)=\left(\frac{\delta^2 log(l(\alpha,\lambda))}{\delta(\alpha,\lambda)\delta(\alpha,\lambda)^T}\right)=\begin{pmatrix}\frac{\delta^2\log l}{\delta\alpha^2} & \frac{\delta^2\log l}{\delta\alpha\delta\lambda}\\ \frac{\delta^2\log l}{\delta\alpha\delta\lambda} & \frac{\delta^2\log l}{\delta\lambda^2}\end{pmatrix} \tag{13}$$

Thus, we compute the element of the $J(\alpha,\lambda)$ as

$$\frac{\delta^2 log(l(\alpha,\lambda))}{\delta\lambda^2}=-\frac{n}{\lambda}-\frac{ne^{\lambda}}{\left(e^{-\alpha x_i}-1\right)^2}$$

$$\frac{\delta^2 log(l(\alpha,\lambda))}{\delta\alpha\delta\lambda}=2\sum_{i=1}^{n}\frac{x_i e^{-\alpha x_i}}{\left(e^{-\alpha x_i}-1\right)^2}$$

$$\frac{\delta^2 log(l(\alpha,\lambda))}{\delta\alpha^2}=-\frac{n}{\alpha^2}-2\sum_{i=1}^{n}\frac{x_i^2 e^{-\alpha x_i}}{\left(1+e^{-\alpha x_i}\right)^2}-2\lambda\sum_{i=1}^{n}\frac{x_i^2 e^{-\alpha x_i}}{\left(1+e^{-\alpha x_i}\right)^3}+2\lambda\sum_{i=1}^{n}\frac{x_i^2 e^{-2\alpha x_i}}{\left(1+e^{-\alpha x_i}\right)^3}$$

To construct the asymptotic distribution for the maximum likelihood estimate we consider the following Lemma and Theorem.

Lemma 1. For $k\in R$ and $q,\theta\in N$, let

$$\xi(k,q,\theta)=\int_0^{\infty}\frac{2\alpha\lambda x^q e^{-\alpha(k+1)x}e^{\lambda\left(\frac{1-e^{-\alpha x}}{1+e^{-\alpha x}}\right)}}{\left(e^{\lambda}-1\right)\left(1+e^{-\alpha x_i}\right)^{\theta+2}}dx \tag{14}$$

Then,

$$\xi(k,q,\theta)=\frac{2}{\left(e^{\lambda}-1\right)}\sum_{j=0}^{\infty}\sum_{w=0}^{\infty}\binom{-j-\theta-2}{w}\frac{(-1)^q\lambda^{j+1}}{\alpha^q\, j!}B_{o,p}\left(j+1,k+w+1\right) \tag{15}$$

Where $B_{o,p}(.,.)$ is a partial derivative of a beta function with respect to p.

Proof. By expanding the exponential expression $e^{\lambda\left(\frac{1-e^{-\alpha x}}{1+e^{-\alpha x}}\right)}$ then applying generalized binomial expansion to $\left(1+e^{-\alpha x}\right)^{(.)}$. Finally, some algebraic transformation of $u=\left(1-e^{-\alpha x}\right)$ and obtain the integral.

Theorem 1. The maximum likelihood estimators $\left(\widehat{\alpha}_n,\widehat{\lambda}_n\right)$ of (α,λ) are consistent estimators and $\sqrt{n}\left(\widehat{\alpha}_n-\alpha,\widehat{\lambda}_n-\lambda\right)^T$ is asymptotically normal with mean vector *0* and the variance covariance matrix I^{-1}, where $I=-E\left(\frac{\delta^2\log l}{\delta(\alpha,\lambda)\delta(\alpha,\lambda)^T}\right)$ and the elements of the Fisher information matrix *I* are;

$$E\left(\frac{\delta^2(logl(\alpha,\lambda))}{\delta\lambda^2}\right)=\frac{1}{\lambda^2}+\frac{e^{\lambda}}{(e^{\lambda}-1)^2} \tag{16}$$

$$E\left(\frac{\delta^2(logl(\alpha,\lambda))}{\delta\alpha\delta\lambda}\right)=-2\xi(1,1,2) \tag{17}$$

$$E\left(\frac{\delta^2(logl(\alpha,\lambda))}{\delta\alpha\delta\lambda}\right)=\frac{1}{\alpha^2}+2\xi(1,2,2)+2\lambda\xi(1,2,3)-2\lambda\xi(2,2,3) \quad (18)$$

Where $\xi(.,.,.)$ is given in Lemma 1.

For r = 1, 2., let $\widehat{\Theta}=(\widehat{\alpha},\widehat{\lambda})^T$ be the estimates of Θ, and Θ_r be the rth component of Θ. Then, a $100(1-\in)\%$ asymptotic confidence interval for Θ_r is given

$$ACI_r=\left(\widehat{\Theta}_r w_{\in/2}\sqrt{\widehat{j}_{rr,}}\widehat{\Theta}_r+w_{\in/2}\sqrt{\widehat{j}_{rr,}}\right) \quad (19)$$

Where $\widehat{\Theta}_r$ is the rth component of $\widehat{\Theta}$, $\widehat{j}_{rr}$ is the $(r,r)^{th}$ diagonal elements of $\widehat{I}^{-1}$, and $w_{\in/2}$ is the quantile $1-\epsilon/2$ of the standard normal distribution.

3.3 Simulation study

In this part, we evaluate the Information matrix through the confidence interval and its coverage probability, by the performance of the maximum likelihood estimates using simulation study. We generate 10,000 samples from the PHL (α, λ), each of sample sizes n=20, 50,…,300 using some selected values of $\alpha>0$ and $\lambda\in R-\{0\}$.

The resulting simulations are displayed in Figures 1, 2,3 and 4. The result shows that the method of maximum likelihood performed consistently and length of the confidence interval decrease as sample size increase, also, the coverage probability (PC) of the 95% asymptotic confidence interval is within the nominal size, showing how accurate the information matrix performed. Below is the simulation algorithm.

1. Choose the sample size n, replication number M,
2. Choose the values of parameters α and λ,
3. Generate random Pi ~ Uniform(0, 1) distribution, i = 1, 2, 3, ⋯, n,
4. Generate random Xi, i = 1, 2, 3, ・ ・ ・ n, from (3),
5. Calculate the MLEs from the simulated data,
6. Compute the expected information matrix
7. Compute the 95% asymptotic confidence interval for $\Theta=(\alpha,\lambda)$ using $ACI_r=\left(\widehat{\Theta}_r w_{\in/2}\sqrt{\widehat{j}_{rr,}}\widehat{\Theta}_r+w_{\in/2}\sqrt{\widehat{j}_{rr,}}\right), r=1,2$
8. Compute the length of ACI
9. Repeat steps 2-4, M times.
10. Compute the average ACI and coverage probability (CP)

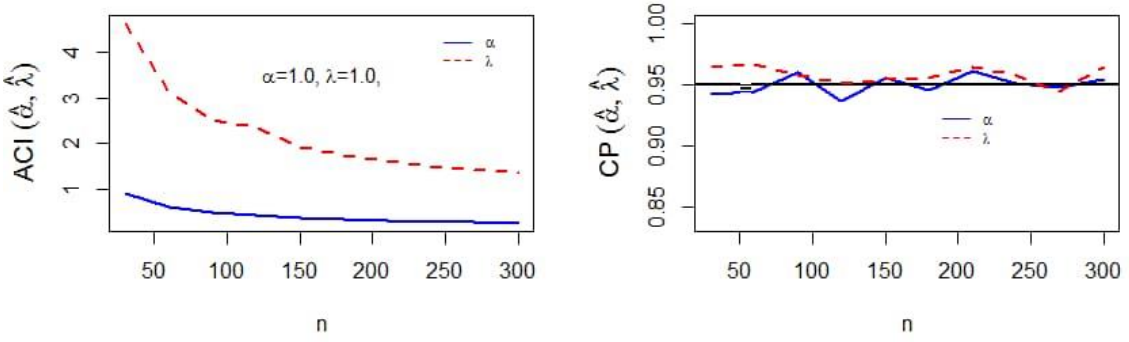


Figure 1: *Plots of the estimated average length of ACL and CP for the simulated data for α = 1.0 and λ= 1.0*

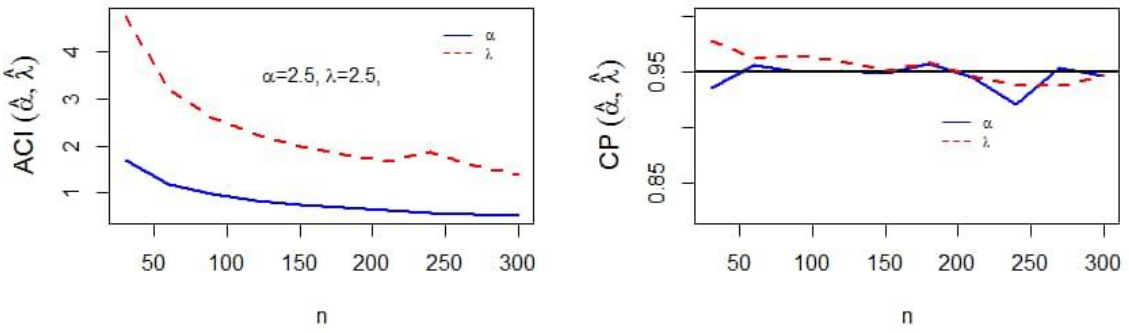


Figure 2: *Plots of the estimated average length of ACL and CP for the simulated data for α = 2.5 and λ= 2.5*

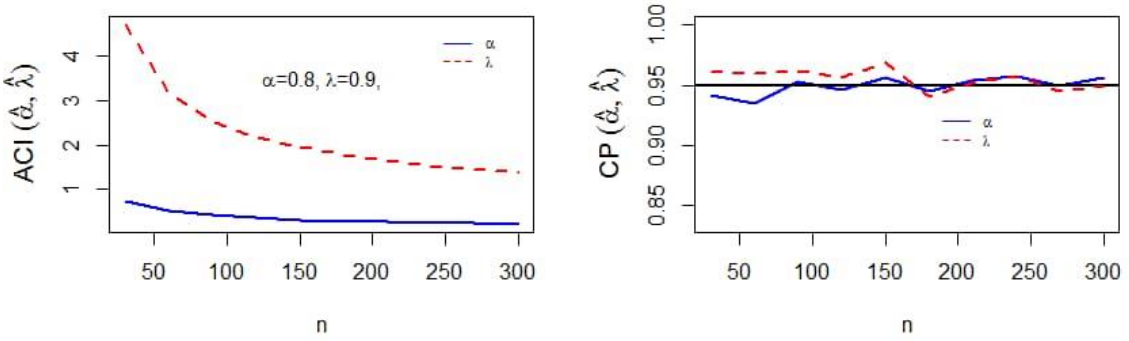


Figure 3: *Plots of the estimated average length of ACL and CP for the simulated data for α = 0.8 and λ= 0.9*

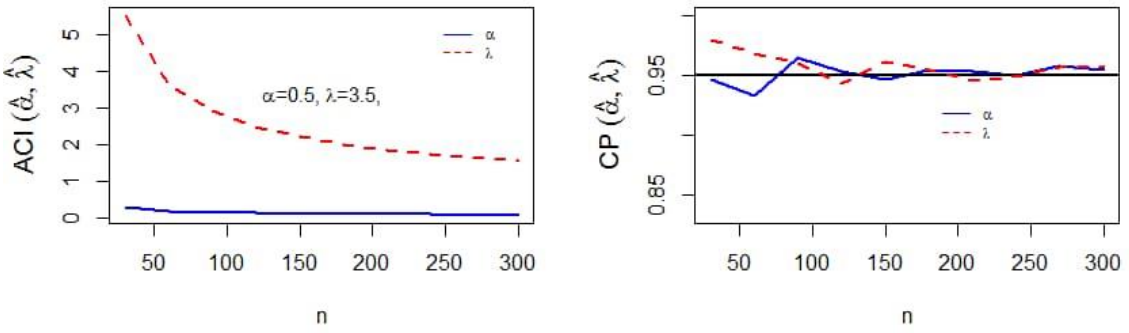


Figure 4:*Plots of the estimated average length of ACL and CP for the simulated data for α = 0.5 and λ= 3.5*

3.4 Application to Real Dataset

This subsection consists of illustration of the PHL expected information matrix to obtain confidence interval for the parameters using a real data with good fit by KS (Kolmogorov Smirnov) test statistic. The data set is given by [17] the data set are the 100 observations on breaking stress of carbon fibers (in Gba): The data set analyzed are: 194, 3.7, 2.74, 2.73, 2.5, 3.6, 3.11, 3.27, 2.87, 1.47, 3.11,4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.9, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53,2.67, 2.93, 3.22, 3.39, 2.81, 4.2, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56, 3.15, 2.35, 2.55, 2.59,2.38, 2.81, 2.77, 2.17,

2.83, 1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59, 3.19,1.57, 0.81, 5.56, 1.73, 1.59, 2, 1.22, 1.12, 1.71, 2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69,1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.7, 2.03, 1.8, 1.57, 1.08, 2.03, 1.61, 2.12,1.89, 2.88, 2.82, 2.05, 3.65.

We estimated the parameters by maximum likelihood and tested the good fit by KS test as α = 1.2004, β = 7.2595, with KS = 0.0823. The asymptotic Fisher information matrix is computed and the asymptotic confidence intervals are computed to verify the performance of the derived information matrix as $ACI\alpha$ = (1.0391, 1.3618) and $ACI\lambda$ = (4.8419, 9.6772). The confidence intervals are very good indicating the accuracy of the computed information matrix. Figure 5 shows the plot of the fitted PHL density and cumulative distribution function showing the good fit.

$$I = \begin{pmatrix} 401.9967 & -21.3447 \\ -21.3447 & 1.7906 \end{pmatrix} \text{ and } I^{-1} = \begin{pmatrix} 0.00677717 & 0.08078816 \\ 0.08078816 & 1.52152823 \end{pmatrix}$$

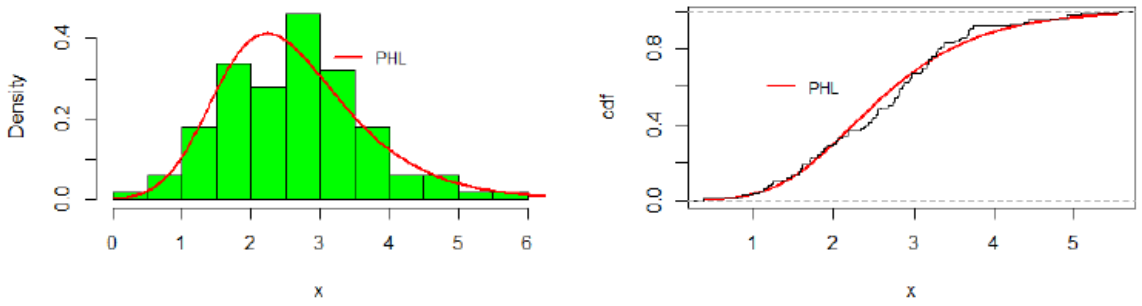


Figure 5: *Plots of the estimated density and cumulative distribution function of PHL*

4. Conclusion

In this study, we computed and assessed the expected information matrix for the PHL distribution, utilizing confidence intervals and their coverage probabilities. Our findings underscore the consistent performance of the maximum likelihood method, showcasing: as sample size increased, the confidence interval length consistently decreased. Moreover, the 95% asymptotic confidence interval' s coverage probability (PC) remained within the expected nominal size, affirming the accuracy and reliability of the information matrix's performance.

Acknowledgments: I will like to thank the editor, and referees for their useful comments which improve the paper.

References

[1] Decani, J. S., & Stine, R. A. (1986). A note on deriving the information matrix for a logistic distribution. *The American Statistician*, *40*(3), 220-222.

[2] Chesher, A., & Spady, R. (1991). Asymptotic expansions of the information matrix test statistic. *Econometrica: Journal of the Econometric Society*, 787-815.

[3] Orme, C. (1990). The small-sample performance of the information-matrix test. *Journal of Econometrics*, *46*(3), 309-331.

[4] Chang, K. C., Zhi, T., & Saha, R. K. (2002). Performance evaluation of track fusion with information matrix filter. *IEEE Transactions on Aerospace and Electronic Systems*, *38*(2), 455-466.

[5] Liao, Z., Drummond, T., Reid, I., & Carneiro, G. (2018). Approximate fisher information matrix to characterize the training of deep neural networks. *IEEE transactions on pattern analysis and machine intelligence*, *42*(1), 15-26.

[6] Wang, Z., Liu, C., Zhao, J., & Shao, L. (2022). Extending the Fisher information matrix in

gravitational-wave data analysis. *The Astrophysical Journal*, *932*(2), 102.

[7] Fiderer, L. J., Tufarelli, T., Piano, S., & Adesso, G. (2021). General expressions for the quantum Fisher information matrix with applications to discrete quantum imaging. *PRX Quantum*, *2*(2), 020308.

[8] Miura, K. (2011). An introduction to maximum likelihood estimation and information geometry. *Interdisciplinary Information Sciences*, *17*(3), 155-174.

[9] Abdel-Hamid, A. H. (2016). Properties, estimations and predictions for a Poisson-half-logistic distribution based on progressively type-II censored samples. *Applied Mathematical Modelling*, *40*(15-16), 7164-7181.

[10] Muhammad, M., & Liu, L. (2021). A new three parameter lifetime model: The complementary Poisson generalized half logistic distribution. *IEEE Access*, *9*, 60089-60107.

[11] Muhammad, M. (2017). Generalized half-logistic Poisson distributions. Communications for Statistical Applications and Methods, 24(4), 353-365.

[12] Muhammad, M., & Liu, L. (2019). A new extension of the generalized half logistic distribution with applications to real data. *Entropy*, *21*(4), 339.

[13] Muhammad, I., Wang, X., Li, C., Yan, M., & Chang, M. (2020). Estimation of the reliability of a stress–strength system from Poisson half logistic distribution. *Entropy*, *22*(11), 1307.

[14] Cordeiro, G. M., Alizadeh, M., & Diniz Marinho, P. R. (2016). The type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation*, *86*(4), 707-728.

[15] Muhammad, M., Bantan, R. A., 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.

[16] Muhammad, M., Tahir, M. H., Liu, L., Jamal, F., Chesneau, C., & Abba, B. (2023). On the Type-I Half-logistic Distribution and Related Contributions: A Review. *Austrian Journal of Statistics*, *52*(5), 34-62.

[17] Nichols, M. D., & Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and reliability engineering international*, *22*(2), 141-151.

RT&A, No 2 (78)
Volume 19, June, 2024

BULK ARRIVING RETRIAL QUEUE WITH G-QUEUE AND RENEGING CLIENTS

J. Bharathi[1], S. Nandhini[2,*], Nur Aisyah Abdul Fataf[3]

•

[1,2] Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore.
[3] Cyber Security and Digital Industrial Revilution Centre,
Universiti Pertahanan Nasional Malaysia, Malaysia.
bharathij2022@gmail.com, n.aisyah@upnm.edu.my, nandhini.s@vit.ac.in

Abstract

We consider a server queue with negative clients (G-Queue) in this effort, where clients are serviced one after the other in batches in a system of variable size. Additionally, we presumptively have a general distribution for the service times, delay times, and repair times. For various states, we concrete the probability-generating functions for the number of customers in the orbit. We scrutinize a single server queue with batches of reneging or balking clients in a system of variable size in this work. Different performance measures and unique situations are examined. The outcomes of this work have applications in satellite communication, software-design for various computer-communication systems and mailing systems among other things.

Keywords: G-Queue, Retrial Queue, Bulk, Reneging Clients, Sudden Breakdown

1. Introduction

The concept of positive and negative consumers coming in a queueing system received further interest and was researched due to its usage in organizations, industry, manufacturing, computer field, and network systems. This study [7] proposed such queues (G-Queues) for the first time to simulate neural networks. In [10] tremendous improvements have been made to the wait times for retrials and vacations. Adapted from [2], in [15] discussed the $M/G/1$ retrial queueing system, which has two service phases and immediate feedback. In this system, the regular busy server is impacted by the arrival of negative customers. By incorporating the idea of G-queues with immediate feedback used by [17]. A finite-source retrial queueing system is considered in [14] along with impatient clients and catastrophic failures. In [13] considered a modified Bernoulli vacation schedule with negative arrivals, reneging and starting failure.

Many authors have examined the queueing issues caused by different combinations of server vacations. A literature review on queues with server vacations can always be found in [6]. Consider this reliability modelling with G-queues in [8], when server failures are described by the arrival of negative clients that cause certain clients to lose service. It was taken into account by [12] to evaluate the queue containing feedback and server vacations (optional) utilizing an SVP (single vacation policy). In [9] examined a bulk- arriving with a server(starting) and more J service options. A batch arrival queue with an additional service channel was researched by [5] under -policy. Retrial queueing technique and balking clients are delved by [3, 4].

Both [1, 11] provided retry queues that take into consideration server faults and repair. The queueing indices and reliability features of an RRQM (repairable retrial queueing model) were

investigated by [16] in terms of reliability. On the M/G/1 retrial queue model with service, we have reviewed a variety of academic publications. This work is motivated by the Retrial Queue model (RQM), which includes service and repair.

The remainder of this article is categorized as follows. We provide a brief mathematical overview and its application of the model is specific in Section 2. The notations and the number of consumers in the orbit/system at a steady state are shown in Section 3 and Section 4. The system performance metrics and numerical outcomes are presented in Section 5 and Section 6. The work's conclusion is stated in section 7.

2. The Mathematical Model's descriptions

Consider an *SSRQM* (single server retry queueing system) with negative and positive independent arrivals. Assume that both categories of customers enter the system using separate Poisson processes with rates of λ and δ respectively. The bulk size Y is a RV (random variable) with *df* (distribution function) $P(\tilde{\tilde{Y}} = k) = \tilde{\tilde{T}}_k, k = 1, 2, \ldots,$.

If a huge proportion of positive consumers discover the server free upon arrival, any newly incoming customer begins his service, and others join the orbit. When positive customer enter the service with *prob.*, (probability) 1-$\tilde{\tilde{b}}$ and exit with probability $\tilde{\tilde{b}}$, balking (or reneging) may occur.

One of the arrivals starts his service, and others join the orbital, if a batch of affirmative clients finds the server unoccupied upon arrival. The generic distribution for the retrial queue is *DF*(distribution function) $A_l(\tilde{x}_1)$ with associated it LST(Laplace Stieltjes transform) $A_l^*(s)$ and *HR*(Hazard rate) $\Upsilon(\tilde{x}_1)d\tilde{x} = \frac{dA_l(\tilde{x}_1)}{1-A_l(\tilde{x}_1)}$.

This service-time also follows a generic distribution with *DF* $B_l(\tilde{x}_1)$, *LST* $B_l^*(s)$, n^{th} factorial moments $\tilde{\tilde{i}}_n$ and its *HR*, $\mu(\tilde{x}_1)d\tilde{x} = \frac{dB_l(\tilde{x}_1)}{1-B_l(\tilde{x}_1)}$. By the Poisson process, negative consumers come individually at a rate of δ. A server breakdown occurs when a negative client gets into the system, removing the server's functioning positive client. The server stops service and waits for repairs to begin whenever it fails.

This waiting time of the server is known as delay time. The Delay time follows a general distribution with *DF* $E_l(\tilde{x}_1)$, *LST* $E_l^*(s)$, n^{th} factorial moments $\tilde{\tilde{k}}_n$ and its *HR*, $\chi(\tilde{x_1})d\tilde{x_1} = \frac{dE_l(\tilde{x_1})}{1-E_l(\tilde{x}_1)}$. When a negative customer comes up, the system no longer has the positive customer in service, which forces the server to breakdown. When a server breaks down, it stops service and waits for repair to begin. The server's waiting period of time is known as the delay time. Furthermore, the repair time has a general distribution with *DF* $F_l(\tilde{x}_1)$, $LSTF_l^*(s)$, n^{th} factorial moments $\tilde{\tilde{l}}_n$ and its *HR*, $\xi(\tilde{x_1}) = \frac{dF_l(\tilde{x}_1)}{1-F_l(\tilde{x}_1)}$.

2.1. Application of the Model in Real Life

The size of the message buffer (orbit) of a CPS (computer processing system), where messages (customers) are received at a time. The work of processing communications falls on the processor (server). A virus infection (a negative customer) might affect the active mail server, and electronic failures (breakdowns) could occur at any time throughout the service term and require urgent repair. At that time, if the processor is not available, FCFS temporarily stores the messages in a buffer to be served later (retrial time). When all messages have been treated (processed) and there are no pending new messages, the processor will carry out several maintenance procedures, such as virus scanning, to improve the computer's performance. The processor checks the messages after each maintenance process is done before deciding whether to restore the rate of the standard services. If the system is currently empty of messages, the processor may decide to do another maintenance task.

3. Probability Notations

The system's stochastic processes are all considered to be independent from one another. We now introduce some more notations that will be utilized in this model's mathematical formulation. Let $C(\tilde{t})$ be the server state, where $C(\tilde{t})$

$$C(\tilde{t}) = \begin{cases} 0 & \longrightarrow \text{ idle (server)} \\ 1 & \longrightarrow \text{ busy (server)} \\ 2 & \rightarrowtail \text{ server is repair(waiting process)} \\ 3 & \longrightarrow \text{ server is repair(under process)} \end{cases}$$

Then the process $\{C(\tilde{t}), \tilde{\tilde{N}}(\tilde{t}); \tilde{t} \geq 0\}$ is a Markov Process.
Define the following probabilities are, for $\tilde{t} \geq 0$
$\tilde{\tilde{I}}_0(\tilde{t}) = P\{\{C(\tilde{t}) = 0, \tilde{\tilde{N}}(\tilde{t}) = 0\}$

$\tilde{\tilde{I}}_{\tilde{n}}(\tilde{x}_1, \tilde{t})d\tilde{x}_1 = P\{\{C(\tilde{t}) = 0, \tilde{\tilde{N}}(\tilde{t}) = \tilde{n}\}, \tilde{x}_1 < A_l^0(\tilde{t}) \leq \tilde{x}_1 + d\tilde{x}_1, \ \ \tilde{n} \geq 1,$

$\tilde{\tilde{M}}_{\tilde{n}}(\tilde{x}_1, \tilde{t})d\tilde{x}_1 = P\{\{C(\tilde{t}) = 1, \tilde{\tilde{N}}(\tilde{t}) = \tilde{n}\}, \tilde{x}_1 < B_l^0(\tilde{t}) \leq \tilde{x}_1 + d\tilde{x}_1, \ \ \tilde{n} \geq 0$

$\tilde{\tilde{Q}}_{\tilde{n}}(\tilde{x}_1, \tilde{t})d\tilde{x}_1 = P\{\{C(\tilde{t}) = 2, \tilde{\tilde{N}}(\tilde{t}) = \tilde{n}\}, \tilde{x}_1 < E_l^0 \leq \tilde{x}_1 + d\tilde{x}_1$

$\tilde{\tilde{R}}_{\tilde{n}}(\tilde{x}_1, \tilde{t})d\tilde{x}_1 = P\{\{C(\tilde{t}) = 3, \tilde{\tilde{N}}(\tilde{t}) = \tilde{n}\}, \tilde{x}_1 < F_l^0 \leq \tilde{x}_1 + d\tilde{x}_1$

4. Steady State Equations

The collection of equations governing the dynamics of the system behaviour in steady state is obtained using the *SVM*(supplementary variable method) as follows:

$$\tilde{\tilde{b}}\lambda\tilde{\tilde{I}}_0 = \int_0^\infty \tilde{\tilde{I}}_0(\tilde{x}_1)\Theta(\tilde{x}_1)d\tilde{x}_1 + \int_0^\infty \tilde{\tilde{R}}_0\xi(\tilde{x}_1)d\tilde{x}_1 \tag{1}$$

$$\frac{d\tilde{\tilde{I}}_{\tilde{n}}(\tilde{x}_1)}{d\tilde{x}_1} + (\lambda + \Upsilon(\tilde{x}_1))\tilde{\tilde{I}}_{\tilde{n}}(\tilde{x}_1) = 0, \ \tilde{n} \geq 1 \tag{2}$$

$$\frac{d\tilde{\tilde{M}}_{\tilde{n}}(\tilde{x}_1)}{d\tilde{x}_1} + (\tilde{\tilde{b}}\lambda + \delta + \mu(\tilde{x_1}))\tilde{\tilde{M}}_{\tilde{n}}(\tilde{x}_1) = \tilde{\tilde{b}}\lambda \sum_{k=0}^{\infty} \tilde{\tilde{T}}_k \tilde{\tilde{M}}_{\tilde{n}-k}(\tilde{x}_1) \tag{3}$$

$$\frac{d\tilde{\tilde{Q}}_{\tilde{n}}(\tilde{x}_1)}{d\tilde{x}_1} + (\tilde{\tilde{b}}\lambda + \chi(\tilde{x}_1))\tilde{\tilde{Q}}_{\tilde{n}}(\tilde{x}_1) = \tilde{\tilde{b}}\lambda \sum_{k=0}^{\infty} \tilde{\tilde{T}}_k \tilde{\tilde{Q}}_{\tilde{n}-k}(\tilde{x}_1) \tag{4}$$

$$\frac{d\tilde{\tilde{R}}_{\tilde{n}}(\tilde{x}_1)}{d\tilde{x}_1} + (\tilde{\tilde{b}}\lambda + \xi(\tilde{x}_1))\tilde{\tilde{R}}_{\tilde{n}}(\tilde{x}_1) = \tilde{\tilde{b}}\lambda \sum_{k=0}^{\infty} \tilde{\tilde{T}}_k \tilde{\tilde{R}}_{\tilde{n}-k}(\tilde{x}_1) \tag{5}$$

The *B.c* (boundary conditions) are

$$\tilde{\tilde{I}}_{\tilde{n}}(0) = \int_0^\infty \tilde{\tilde{I}}_0(\tilde{x}_1)\Theta(\tilde{x}_1)d\tilde{x_1} + \int_0^\infty \tilde{\tilde{R}}_0\xi(\tilde{x}_1)d\tilde{x_1} - \tilde{\tilde{b}}\lambda\tilde{\tilde{I}}_0 \tag{6}$$

$$\tilde{\tilde{M}}_0(0) = \lambda\tilde{\tilde{T}}_1\tilde{\tilde{I}}_0 + \int_0^\infty \tilde{\tilde{I}}_1(\tilde{x}_1)\Upsilon(\tilde{x}_1)d(\tilde{x}_1) \tag{7}$$

$$\tilde{\tilde{M}}_{\tilde{n}}(0) = \lambda\tilde{\tilde{T}}_{\tilde{n}+1}\tilde{\tilde{I}}_0 + \int_0^\infty \tilde{\tilde{I}}_{\tilde{n}+1}(\tilde{x}_1)\Upsilon(\tilde{x}_1)d(\tilde{x}_1) + \lambda \sum_{k=0}^{\infty} \tilde{\tilde{T}}_k \int_0^\infty \tilde{\tilde{M}}_{\tilde{n}-k+1}(\tilde{x}_1)d(\tilde{x}_1) \tag{8}$$

$$\tilde{\tilde{Q}}_{\tilde{n}}(0) = \delta \int_0^\infty \tilde{\tilde{M}}_{\tilde{n}}(\tilde{x}_1) d(\tilde{x}_1) \tag{9}$$

$$\tilde{\tilde{R}}_{\tilde{n}}(0) = \int_0^\infty \tilde{\tilde{Q}}_{\tilde{n}}(\tilde{x}_1) \chi(\tilde{x}_1) d(\tilde{x}_1) \tag{10}$$

Normalization Condition is

$$\tilde{\tilde{I}}_0 + \sum_{\tilde{n}=1}^{\infty} \int_0^\infty \tilde{\tilde{I}}_{\tilde{n}}(\tilde{x}_1) d(\tilde{x}_1) + \sum_{\tilde{n}=1}^{\infty} \int_0^\infty \tilde{\tilde{M}}_{\tilde{n}}(\tilde{x}_1) d(\tilde{x}_1) + \sum_{\tilde{n}=0}^{\infty} \int_0^\infty \tilde{\tilde{Q}}_{\tilde{n}}(\tilde{x}_1) d(\tilde{x}_1) + \sum_{\tilde{n}=0}^{\infty} \int_0^\infty \tilde{\tilde{R}}_{\tilde{n}}(\tilde{x_1}) d(\tilde{x}_1) \tag{11}$$

The following findings are obtained by multiply equ (2) - (10) by $\tilde{z_1}^{\tilde{n}}$ and adding all values(possible) of $\tilde{n}$:

$$\frac{d\tilde{\tilde{I}}(\tilde{x_1}, \tilde{z}_1)}{d\tilde{x}_1} + (\lambda + \Upsilon(\tilde{x}_1))\tilde{\tilde{I}}(\tilde{x}_1, \tilde{z}_1) = 0 \tag{12}$$

$$\frac{d\tilde{\tilde{M}}(\tilde{x}_1, \tilde{z}_1)}{d\tilde{x}_1} + (\tilde{\tilde{b}}\lambda(1 - \tilde{\tilde{T}}(\tilde{z_1})) + \delta + \mu(\tilde{x}_1)\tilde{\tilde{M}}(\tilde{x_1}, \tilde{z}_1) = 0 \tag{13}$$

$$\frac{d\tilde{\tilde{Q}}(\tilde{x}_1, \tilde{z}_1)}{d\tilde{x}_1} + (\tilde{\tilde{b}}\lambda(1 - \tilde{\tilde{T}}(\tilde{z}_1)) + \chi(\tilde{x}_1))\tilde{\tilde{Q}}(\tilde{x}_1, \tilde{z}_1) = 0 \tag{14}$$

$$\frac{d\tilde{\tilde{R}}(\tilde{x}_1, \tilde{z}_1)}{d\tilde{x}_1} + (\tilde{\tilde{b}}\lambda(1 - \tilde{\tilde{T}}(\tilde{z}_1)) + \xi(\tilde{x}_1))\tilde{\tilde{R}}(\tilde{x}_1, \tilde{z}_1) = 0 \tag{15}$$

Equations (12) to (15), using to solve partial differential

$$\tilde{\tilde{I}}(\tilde{x}_1, \tilde{z}_1) = \tilde{\tilde{I}}(0, \tilde{z}_1)[1 - A_l(\tilde{x}_1)]e^{-\lambda \tilde{x}_1} \tag{16}$$

$$\tilde{\tilde{M}}(\tilde{x}_1, \tilde{z}_1) = \tilde{\tilde{M}}(0, \tilde{z}_1)[1 - B_l(\tilde{x}_1)]e^{-\tilde{N}(\tilde{z}_1)\tilde{x}_1} \tag{17}$$

$$\tilde{\tilde{Q}}(\tilde{x}_1, \tilde{z}_1) = \tilde{\tilde{Q}}(0, \tilde{z}_1)[1 - E_l(\tilde{x}_1)]e^{-O(\tilde{z}_1)\tilde{x}_1} \tag{18}$$

$$\tilde{\tilde{R}}(\tilde{x}_1, \tilde{z}_1) = \tilde{\tilde{R}}(0, \tilde{z}_1)[1 - F_l(\tilde{x}_1)]e^{O(\tilde{z}_1)\tilde{x}_1} \tag{19}$$

where $N(\tilde{z}_1) = O(\tilde{z_1}) + \delta$, and $O(\tilde{z_1}) = \tilde{\tilde{b}}\lambda(1 - \tilde{\tilde{T}}(\tilde{z}_1))$

$$\tilde{\tilde{I}}(0, \tilde{z}_1) = \int_0^\infty \tilde{\tilde{M}}(\tilde{x}_1, \tilde{z}_1)\Theta(\tilde{x}_1) d\tilde{x}_1 + \int_0^\infty \tilde{\tilde{R}}(\tilde{x}_1, \tilde{z}_1)\xi(\tilde{x}_1) d\tilde{x}_1 - \tilde{\tilde{b}}\lambda\tilde{\tilde{I}}_0 \tag{20}$$

$$\tilde{\tilde{M}}(0, \tilde{z}_1) = \frac{1}{\tilde{z}_1}\int_0^\infty \tilde{\tilde{M}}(\tilde{x}_1, \tilde{z}_1)\Upsilon(\tilde{x}_1) d\tilde{x_1} + \frac{\lambda\tilde{\tilde{T}}(\tilde{z}_1)}{\tilde{z}_1}\left[\int_0^\infty \tilde{\tilde{I}}(\tilde{x}_1, \tilde{z}_1) d\tilde{x_1} + \tilde{b}\lambda\tilde{\tilde{I}}_0\right] \tag{21}$$

$$\tilde{\tilde{Q}}(0, \tilde{z}_1) = \delta \int_0^\infty \tilde{\tilde{M}}(\tilde{x_1}, \tilde{z}_1) d\tilde{x_1} \tag{22}$$

$$\tilde{\tilde{R}}(0, \tilde{z}_1) = \int_0^\infty \tilde{\tilde{D}}(\tilde{x_1}, \tilde{z}_1)\xi(\tilde{x}_1) d\tilde{x_1} \tag{23}$$

The orbital size partial *PGF* (probability generating function) while the server is inactive, active, waiting for repair, under repair

$$\tilde{\tilde{I}}(\tilde{z}_1) = \frac{\begin{bmatrix} \tilde{\tilde{I}}_0(1 - A_l^*(\lambda))\tilde{\tilde{b}}\{N(\tilde{z}_1)B_l^*(N(\tilde{Z}_1)) \\ +\delta\tilde{\tilde{T}}(z_1)(1 - B_l^*(N(\tilde{z}_1)))E^*O(\tilde{z}_1)F^*(O(\tilde{z_1}))\} \end{bmatrix}}{\begin{bmatrix} \tilde{z_1}N(\tilde{z}_1) - \{[A_l^*(\lambda) + \tilde{\tilde{T}}_{\tilde{z}_1}(1 - A_l^*(\lambda))]N(\tilde{z}_1)B_l^*N(\tilde{z}_1) \\ +\delta(1 - B_l^*(N(\tilde{z}_1)))E_l^*(O(\tilde{z}_1))F_l^*(O(\tilde{z_1}))\} \end{bmatrix}} \tag{24}$$

$$\tilde{\tilde{M}}(\tilde{z}_1) = \frac{\tilde{\tilde{I}}_0 A_l^*(\lambda)(B_l^*(N(\tilde{z}_1)) - 1)}{\left[\begin{array}{r} \tilde{z_1} N(\tilde{z}_1) - \{[A_l^*(\lambda) + \tilde{\tilde{T}}_{\tilde{z}_1}(1 - A_l^*(\lambda))]N(\tilde{z}_1)B_l^* N(\tilde{z}_1) \\ +\delta(1 - B_l^*(N(\tilde{z}_1)))E_l^*(O(\tilde{z}_1))F_l^*(O(\tilde{z_1}))\} \end{array}\right]} \tag{25}$$

$$\tilde{\tilde{Q}}(\tilde{z}_1) = \frac{\tilde{\tilde{I}}_0 A_l^*(\lambda)\delta(E_l^*(N(\tilde{z}_1)) - 1)(1 - B_l^*(N(\tilde{z}_1)))}{\left[\begin{array}{r} \tilde{z_1} N(\tilde{z}_1) - \{[A_l^*(\lambda) + \tilde{\tilde{T}}_{\tilde{z}_1}(1 - A_l^*(\lambda))]N(\tilde{z}_1)B_l^* N(\tilde{z}_1) \\ +\delta(1 - B_l^*(N(\tilde{z}_1)))E_l^*(O(\tilde{z}_1))F_l^*(O(\tilde{z_1}))\} \end{array}\right]} \tag{26}$$

$$\tilde{\tilde{R}}(\tilde{z}_1) = \frac{\tilde{\tilde{I}}_0 A_l^*(\lambda)\delta(F_l^*(N(\tilde{z}_1)) - 1)(1 - B_l^*(N(\tilde{z})))E_l^*(N(\tilde{z}_1))}{\left[\begin{array}{r} \tilde{z_1} N(\tilde{z}_1) - \{[A_l^*(\lambda) + \tilde{\tilde{T}}_{\tilde{z}_1}(1 - A_l^*(\lambda))]N(\tilde{z}_1)B_l^* N(\tilde{z}_1) \\ +\delta(1 - B_l^*(N(\tilde{z}_1)))E_l^*(O(\tilde{z}_1))F_l^*(O(\tilde{z_1}))\} \end{array}\right]} \tag{27}$$

Since $\tilde{\tilde{I}}_0$ can be calculated using the normalization condition and represents the probability that the server would be idle while there are no customers in the orbit,

$$\tilde{\tilde{I}}_0 = \left\{ \begin{array}{r} \delta(1 - \tilde{\tilde{j}}_1 + \tilde{\tilde{j}}_1 {A_l}^*(\lambda)) - \lambda\tilde{\tilde{j}}_1(1 - B_l^*(\delta)) \\ \hline (1 + \delta\tilde{\tilde{k}}_1 + \delta\tilde{\tilde{l}}_1) \\ \delta(1 - \tilde{\tilde{b}})(1 - \tilde{\tilde{j}}_1 + \tilde{\tilde{j}}_1 {A_l}^*(\lambda)) + \tilde{\tilde{b}}\delta {A_l}^*(\lambda) \\ -(1 - \tilde{\tilde{b}})({A_l}^*(\lambda))\lambda\tilde{\tilde{j}}_1(1 - B_l^*(\delta))(1 + \delta\tilde{\tilde{k}}_1 + \delta\tilde{\tilde{l}}_1) \end{array} \right\}$$

We establish the following definitions for the *PGF* for the system's customers:

$\tilde{\tilde{K}}_l(\tilde{z}_1) = \tilde{\tilde{I}}_0 + \tilde{\tilde{I}}(\tilde{z}_1) + \tilde{\tilde{M}}(\tilde{z}_1) + \tilde{\tilde{Q}}(\tilde{z}_1) + \tilde{\tilde{R}}(\tilde{z}_1)$
$\tilde{\tilde{K}}_l(\tilde{z}_1) = \tilde{\tilde{I}}_0\{\frac{N1}{D1}\}$

$$N1 = \left\{ \begin{array}{r} \tilde{z_1} N(\tilde{z_1})[1 - \tilde{\tilde{b}}(1 - {A_l}^*(\lambda))] + (\tilde{\tilde{b}} - 1)(1 - {A_l}^*(\lambda))\tilde{\tilde{T}}(\tilde{z})[N(\tilde{z})B^*(N(\tilde{z_{11}})) \\ +\delta(1 - B_l^*(N(\tilde{z}_1)))\tilde{E}^*(O(\tilde{z}_1))F_l^*(O(\tilde{z}_1))] - A^*(\lambda)[B_l^*(N(\tilde{z}_1)) \\ +(1 - B_l^*(N(\tilde{z}_1)))(\tilde{z_1}(O(\tilde{z}_1)) + \delta)] \end{array} \right\}$$

$$D1 = \left\{ \begin{array}{r} \tilde{z_1} N(\tilde{z}_1) - [A_l^*(\lambda) + \tilde{\tilde{T}}_{\tilde{z_1}}(1 - A_l^*(\lambda))]\{N(\tilde{z_1})B_l^*(N(\tilde{z}_1)) + \delta(1 - B_l^*(N(\tilde{z_1})))) \\ E_l^*(O(\tilde{z_1}))F_l^*(O(\tilde{z}_1))\} \end{array} \right\}$$

We define the probability generating functions of the number of customers in the orbit, where $\tilde{H}_l(\tilde{z}_1) = \tilde{\tilde{I}}_0 + \tilde{\tilde{I}}(\tilde{z_1}) + \tilde{z_1}\tilde{\tilde{M}}(\tilde{z}_1) + \tilde{\tilde{Q}}(\tilde{z_1}) + \tilde{\tilde{R}}(\tilde{z}_1)$ and $\tilde{\tilde{H}}_l(\tilde{z}_1) = \tilde{\tilde{I}}_0 \frac{N2}{D1}$

$$N2 = \left\{ \begin{array}{r} \tilde{z_1} N(\tilde{z_1})[1 - \tilde{\tilde{b}}(1 - {A_l}^*(\lambda))] - {A_l}^*(\lambda)[B_l^*(N(\tilde{z_1})) \\ +(1 - B_l^*(N(\tilde{z_1})))(\tilde{z_1}(O(\tilde{z}_1)) + \delta)] \\ +(\tilde{\tilde{b}} - 1)(1 - {A_l}^*(\lambda))\{\tilde{\tilde{T}}(\tilde{z}_1)[N(\tilde{z}_1)B_l^*(N(\tilde{z}_1)) \\ +\delta(1 - B_l^*(N(\tilde{z}_1)))E_l^*(O(\tilde{z}_1))F_l^*(O(\tilde{z}_1))]\} \end{array} \right\}$$

5. Performance Measures

We derive the system performance of our model.
By differentiating $\tilde{\tilde{K}}_l(\tilde{z}_1)$ with respect to $\tilde{z_1}$ and evaluating at $\tilde{z_1} = 1$, the average number of consumers in the system $\tilde{\tilde{L}}_s$ in steady-statestate conditions may be determined.

$$\tilde{\tilde{L}}_s = \lim_{\tilde{z} \to 1} \tilde{\tilde{K}}'(\tilde{z})$$

$$\tilde{L_s} = \tilde{I_0}\frac{(Dr'Nr_1'' - Nr_1'Dr'')}{2(Dr')^2}$$

$$Nr_1' = \left\{\begin{array}{r}(-\tilde{b}\lambda\tilde{j_1} + \delta)[1 - \tilde{b}(1 - A_l^*(\lambda))] + A_l^*(\lambda)\tilde{b}\lambda\tilde{j_1} + (\tilde{b} - 1)(1 - A_l^*(\lambda)) \\ \{\tilde{b}\lambda\tilde{j_1}[-B_l^*(\delta) + \delta(1 - B_l^*(\delta))(\tilde{k_1} + \tilde{l_1})] + \delta\tilde{j_1}\}\end{array}\right\}$$

$$Nr_1'' = -\left\{\begin{array}{r}[\tilde{b}\lambda\tilde{j_2} + 2\tilde{b}\lambda\tilde{j_1}][1 - \tilde{b}(1 - A_l^*(\lambda))] + A_l^*(\lambda)B_l^*(\delta)\tilde{b}\lambda\tilde{j_2} \\ +(\tilde{b} - 1)(1 - A_l^*(\lambda))\{2\tilde{j_1}(-\tilde{b}\lambda\tilde{j_1}B_l^*(\delta) \\ +\delta(1 - B_l^*(\delta))\tilde{b}\lambda\tilde{j_1}(\tilde{k_1} + \tilde{l_1})) + \delta\tilde{j_2} - (\tilde{b}\lambda\tilde{j_2}B_l^*(\delta) \\ +2B_l^*(\delta)\tilde{b}\lambda\tilde{j_i}\tilde{i_1} + 2\delta(B_l^*\lambda\tilde{j_1})^2\tilde{i_1}(\tilde{k_1} + \tilde{l_1})) \\ +\delta(1 - B_l^*(\delta))[(\tilde{b}\lambda\tilde{j_1})^2(\tilde{l_2} + \tilde{k_2}) \\ +\tilde{b}\lambda\tilde{j_2}(\tilde{l_1} + \tilde{k_1}) + 2(\tilde{b}\lambda\tilde{j_1})^2\tilde{k_1}\tilde{l_1}]\}\end{array}\right\}$$

$$Dr' = \delta(1 - \tilde{j_1} + \tilde{j_1}A_l^*(\lambda)) - \lambda\tilde{j_1}(1 - B_l^*(\delta))[1 + \delta(\tilde{l_1} + \tilde{k_1})]$$

$$Dr'' = \left\{\begin{array}{r}-\lambda\tilde{j_2}(1 - B_l^*(\delta)) - 2\lambda\tilde{j_2} - \delta\tilde{j_2}(1 - A_l^*(\lambda)) \\ -2\tilde{j_1}(1 - A_l^*(\lambda))(-\lambda\tilde{j_1}B_l^*(\delta) + \delta\lambda\tilde{j_1}(1 - B_l^*(\delta))(\tilde{k_1} + \tilde{l_1})) \\ +2(\lambda\tilde{j_1})^2\tilde{i_1}(\delta + (\tilde{k_1} + \tilde{l_1})) - 2\delta(\lambda\tilde{j_1})^2(1 - B_l^*(\delta))\tilde{k_1}\tilde{l_1} \\ -\delta\lambda\tilde{l_2}(\tilde{k_1} + \tilde{l_1})(1 - B_l^*(\delta)) - (\lambda\tilde{j_1})^2(1 - B_l^*(\delta))(\tilde{k_2} + \tilde{l_2})\end{array}\right\}$$

By differentiating $\tilde{H}_l(\tilde{z}_1)$ with respect to $\tilde{z_1}$ and evaluating at $\tilde{z_1} = 1$, the average number of consumers in the orbit $\tilde{L_q}$ in steady-statestate conditions may be determined

$$\tilde{L_q} = \lim_{\tilde{z}\to 1}\tilde{H_l}'(\tilde{z})$$

$$\tilde{L_q} = \tilde{I_0}\frac{(Dr'Nr_2'' - Nr_2'Dr'')}{2(Dr')^2}$$

$$Nr_2' = \left\{\begin{array}{r}[-\tilde{b}\lambda\tilde{j_1} + \delta][1 - \tilde{b}(1 - A_l^*(\lambda))] + A_l^*(\lambda)\tilde{b}\lambda\tilde{j_1}A_l^*(\lambda) \\ +(\tilde{b} - 1)(1 - A_l^*(\lambda))\{\ -\tilde{b}\lambda\tilde{j_1}B_l^*(\delta) \\ +\delta[1 - B_l^*(\delta)]\tilde{b}\lambda\tilde{j_1}(\tilde{k_1} + \tilde{l_1}) + \delta\tilde{j_1}\}\end{array}\right\}$$

$$Nr_2'' = \left\{\begin{array}{r}-[\tilde{b}\lambda\tilde{j_2} + 2\tilde{b}\lambda\tilde{j_1}][1 - \tilde{b}(1 - A_l^*(\lambda))] + A_l^*(\lambda)[\tilde{b}\lambda\tilde{j_2}B_l^*(\delta) + \delta\tilde{b}\lambda\tilde{j_1}\tilde{i_1} \\ +\delta(\tilde{b}\lambda\tilde{j_1})^2\tilde{i_2} + \delta\tilde{b}\lambda\tilde{j_2}\tilde{i_1} - \tilde{b}\lambda\tilde{j_2}(1 - B^*(\delta))] \\ +(\tilde{b} - 1)(1 - A_l^*(\lambda))2[\tilde{j_1}(-\tilde{b}\lambda\tilde{j_1})B_l^*(\delta) + \delta(1 - B_l^*(\delta))\tilde{b}\lambda\tilde{j_1}(\tilde{k_1} \\ +\tilde{l_1}) - ((\tilde{b}\lambda\tilde{j_2})B_l^*(\delta)) + 2(\tilde{b}\lambda\tilde{j_1})^2[\tilde{i_1} + \delta\tilde{i_1}(\tilde{k_1} + \tilde{l_1})] \\ +[1 - B_l^*(\delta)]((\tilde{b}\lambda\tilde{j_1})^2[\tilde{k_2} + \tilde{l_2} + 2\tilde{k_1}\tilde{l_1}] + \tilde{b}\lambda\tilde{j_2}(\tilde{l_1} + \tilde{k_1}))]\end{array}\right\}$$

6. Numerical Results

We demonstrate the various settings on system behavior measurements in this section using MATLAB. We examine retrial times, service times, service times with a reduced speed, vacation times, delayed repair times, and exponentially distributed repair times. To meet the stability condition, the numerical measurements are chosen at random. Regarded predicted values of our model's varying metrics, such as the typical queue size and the probability that the server isn't active while retries, and the likelihood that the server is idle overall. and the likelihood that the server is idle overall. $\lambda = 1, \delta = 1, \tilde{\tilde{j}}_1 = 1, \tilde{\tilde{j}}_1 = 0, \mu = 15, \chi = 0.9, \xi = 0.8, \tilde{\tilde{b}} = 0.6$.

In Table 1 represents the effect of retrial rate (Υ) on $\tilde{\tilde{I}}_0$, $\tilde{\tilde{I}}(1)$, $\tilde{\tilde{M}}(1)$, $\tilde{\tilde{Q}}(1)$ and L_q

Table 1: *The effect of retrial rate (Υ) on $\tilde{\tilde{I}}_0$, $\tilde{\tilde{I}}(1)$, $\tilde{\tilde{M}}(1)$, $\tilde{\tilde{Q}}(1)$ and L_q*

Υ	$\tilde{\tilde{I}}_0$	$\tilde{\tilde{I}}(1)$	$\tilde{\tilde{M}}(1)$	$\tilde{\tilde{Q}}(1)$	$\tilde{\tilde{R}}(1)$	L_q
6	0.9169	0.0079	0.0628	0.0055	0.0105	1.2886
7	0.9171	0.0067	0.0627	0.0054	0.0104	1.3339
8	0.9176	0.0059	0.0626	0.0053	0.0103	1.3692
9	0.9180	0.0052	0.0625	0.0052	0.0102	1.3976
10	0.9185	0.0047	0.0624	0.0051	0.0101	1.4208

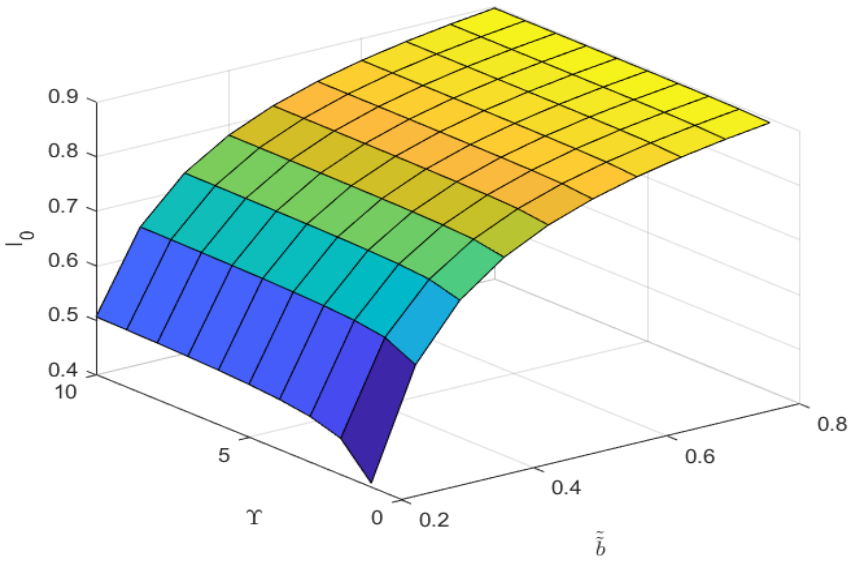


Figure 1: *I_0 versus $\tilde{\tilde{b}}$ and Υ*

Figures provide illustrations of three-dimensional graphs (1–5).
Figure 1 demonstrates $\tilde{\tilde{b}}$ and Υ increases I_0 also increases, Figure 2 demonstrates $\tilde{\tilde{b}}$ and μ increases I_0 also increases , Figure 3 demonstrates $\tilde{\tilde{b}}$ and Υ increases I_0 also increases , Figure 4 demonstrates shows μ increases $\tilde{\tilde{L}}_q$and $\tilde{\tilde{W}}_q$ also increases , Figure 5 demonstrates Υ increases $\tilde{\tilde{L}}_q$and $\tilde{\tilde{W}}_q$ also increases. Through the aforementioned numerical examples, we got able to see how parameters influenced the system's performance metrics and determine that the findings were accurate for real-world applications.

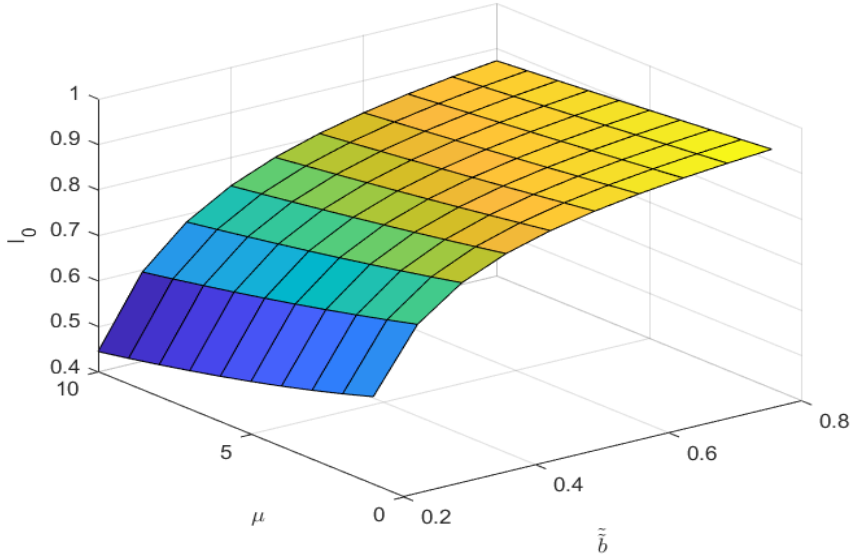


Figure 2: *I_0 versus $\tilde{\tilde{b}}$ and μ*

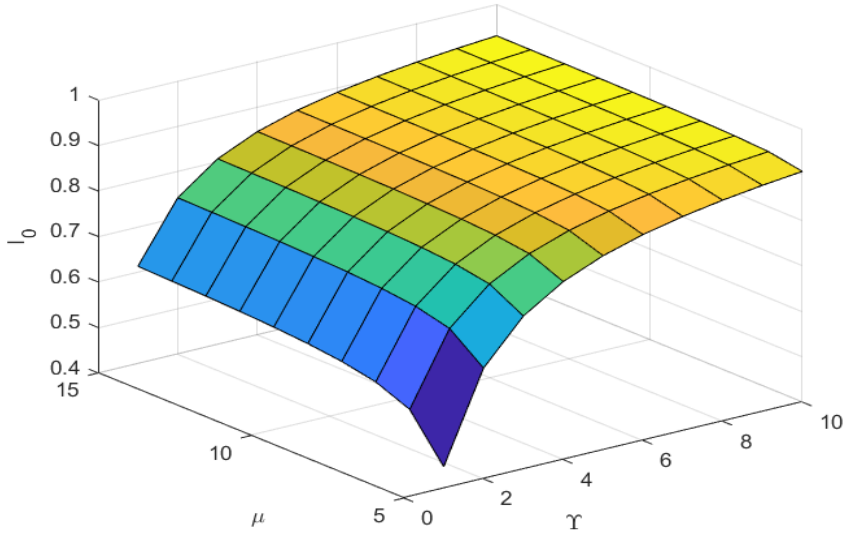


Figure 3: *I_0 versus μ and Υ*

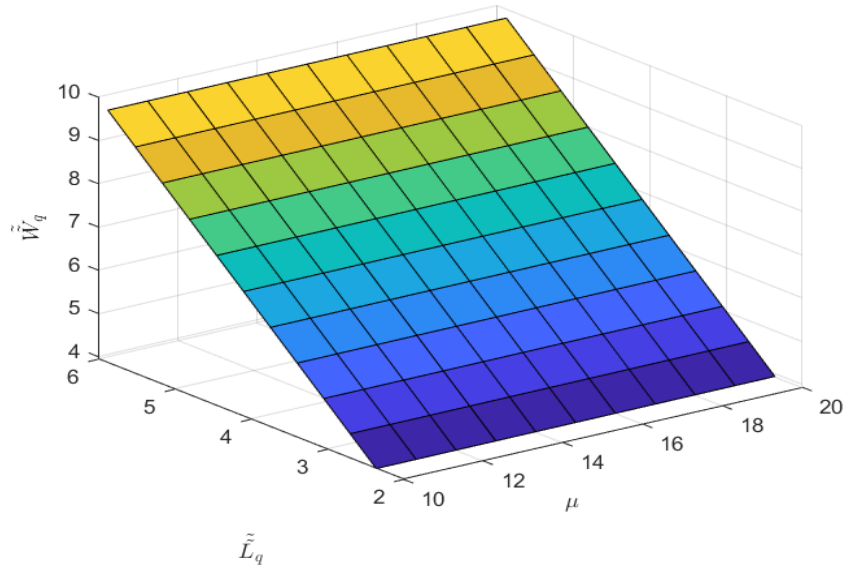


Figure 4: *μ versus $\tilde{L}_q$ and $\tilde{W}_q$*

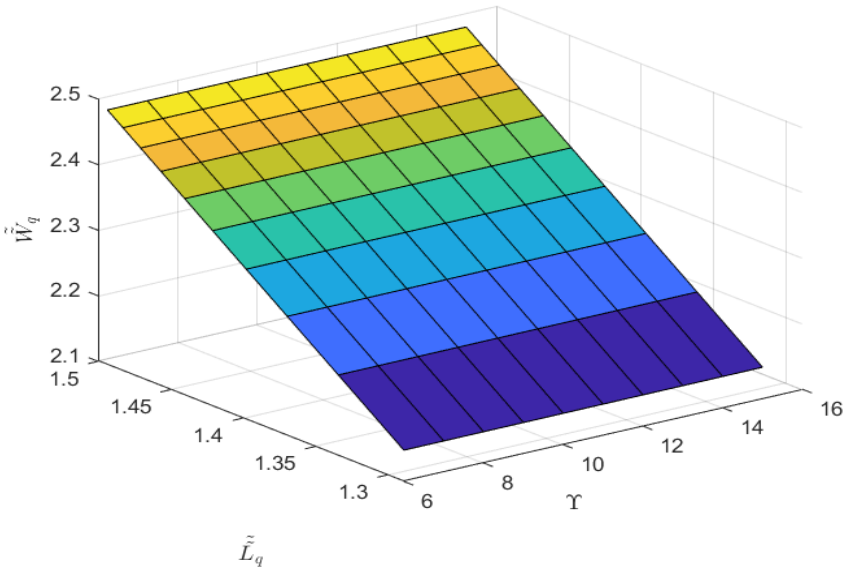


Figure 5: *Υ versus $\tilde{L}_q$ and $\tilde{W}_q$*

7. Conclusion

We discussed a server queue with negative clients in this effort, where customers are serviced one after the other in batches in a system of variable size. Additionally, we presumptively have a general distribution for the service times, delay times, and repair times. For various states, we derived the probability generating functions for the number of customers in the orbit. We have explored a single server queue with batches of reneging or balking clients in a system of variable size in this work. Different performance measures and unique situations have been examined. The outcomes of this work have applications in satellite communication, software-design for various computer-communication system and mailing systems among other things. By including orbit search, starting failure, and working vacation policies, this work can also be expanded.

References

[1] Aissani, A. (1988). On the M/G/1 queueing system with repeated orders and unreliable server. *Journal of Technology*, 6:93–123.

[2] Artalejo, J and Gomez-Corral, A. (2008). Retrial Queueing Systems. *Berlin Germany, Springer*.

[3] Bharathi, J and Nandhini, S. (2024). A single server Non-Markovian with non-compulsory re-service and balking under Modified Bernoulli Vacation. *Journal of King Saud University Science*, 36(1): 103007.

[4] Bharathi, J. and Nandhini, S. (2023). Unreliable server with Non-Markovian Retrial Queueing System, Bernoulli Vacation and Fortuitous breakdown. *Journal of Intelligent and Fuzzy Systems*, 45(6):10089—10098.

[5] Choudhury, G. and Paul, M. (2004). A batch arrival queue with an additional service channel under N-policy, *Applied Mathematics and Computation*, 156: 115–130.

[6] Doshi, B. T. (1986). Queueing systems with vacations: A survey. *Queueing Systems*, 1:29–66.

[7] Gelenbe, E. (1989). Random neural networks with negative and positive signals and product form solution. *Neural computation*, 1(4):502–510.

[8] Harrison, P. G. (2000). Patel, N. M. and Pitel, E. Reliability modelling using G-queues. *European Journal of Operational Research*, 126: 273–287.

[9] Ke, J. (2008). An $M^{[X]}/G/1$ system with startup server and J additional options for service. *Applied Mathematical Modelling*, 32:443-458, .

[10] Ke, J., 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.

[11] Kulkarni, V. G. and Choi, B. D. (1990). Retrial queues with server subject to breakdowns and repairs Queueing Systems. *Queueing Systems* 7:191–208.

[12] Madan, K. C. and Al-Rawwash, M. (2005). On the $M^X/G/1$ queue with feedback and optional server vacations based on a single vacation policy, *Applied Mathematics and Computation*, 160:909-919.

[13] Raj, L. F., Revathi, C. and Saravanarajan, M. C. (2022). An $M/G/1$ retrial G-queue for balking, reneging, subject to modified Bernoulli vacation, starting failure, *In AIP Conference Proceedings* 2385: 130015.

[14] Sztrik, J., Toth, A., Pinter, A. and Bacs, Z. (2022). Simulation of Finite-Source Retrial Queueing Systems with Impatient Customers Using Different Failure Modes. *In International Conference on Information Technologies and Mathematical Modelling,Springer, Cham* 1:16–27.

[15] Varalakshmi, M., Chandrasekaran, V. M. and Saravanarajan, M. C. (2017). A study on $M/G/1$ retrial G-queue with two phases of service, immediate feedback and working vacations. *In IOP conference series: materials science and engineering* 263:042156.

[16] Wang, J. Cao, J. and Li Q (2001). *Reliability analysis of the retrial queue with server breakdowns and repairs Queueing Systems*, 38: 363–380, .

[17] Zhang, M. and Liu, Q. (2015). An $M/G/1$, G-queue with server breakdown, working vacations and vacation interruption. *Opsearch*, 52(2): 256–270.

RT&A, No 2 (78)
Volume 19, June, 2024

A BAYESIAN APPROACH FOR CHRIS-JERRY DISTRIBUTION USING VARIOUS LOSS FUNCTIONS

Dr. G Meenakshi[1] and Balachandar.B[2]

•

(1). Professor, Department of Statistics, Annamalai University
(2). Research Scholar, Department of Statistics, Annamalai University
Email: progmau@gamil.com , balachandarb03061998@gmail.com

Abstract

The paper introduces a Bayesian approach for estimating parameters of the Chris-Jerry distribution, focusing on the use of a conjugate prior, specifically the gamma prior. The Bayesian estimation method is developed with a various loss function, offering a robust framework for parameter estimation. symmetric loss function and Linex loss functions are commonly used in Bayesian statistics to balance the trade-off between bias and variance. The central idea is to derive the Bayes estimate of the distribution parameter by leveraging the properties of the conjugate gamma prior. Conjugate priors simplify the Bayesian analysis by ensuring that the posterior distribution belongs to the same family as the prior, facilitating analytical calculations. The proposed methodology is implemented and validated through numerical illustrations using. This involves applying the developed Bayesian estimation framework to real-world data or simulated scenarios, demonstrating its effectiveness and practical applicability. The numerical and simulation studies are done by using r software

.

Keywords: prior, posterior distribution, posterior mean loss function, linex loss function and symmetric loss function

I. Introduction

In the realm of statistical inference, the Bayesian approach stands as a formidable paradigm, offering a unique perspective that seamlessly integrates prior knowledge with observed data to yield more robust and nuanced estimates. At the heart of Bayesian estimation lies the elegant concept of conjugate priors, a fact that not only simplifies the computational complexity but also enriches the analytical insights.

This article aims to unravel the significance of conjugate priors and their pivotal role in streamlining the inference process. From their foundational principles to practical applications, this paper will explore how these priors provide a harmonious bridge between prior beliefs and empirical evidence, creating a coherent framework for making informed decisions. The Bayesian paradigm, with its emphasis on updating beliefs in light of new information, has found extensive applications across various fields, from finance and engineering to medicine and machine learning. Within this framework, the choice of prior distributions can profoundly impact the outcome of Bayesian analyses. Conjugate priors, by virtue of their mathematical properties, offer an elegant

solution, simplifying the computations involved in posterior distribution calculations. This article will delve into the conceptual underpinnings of Bayesian estimation, shedding light on the fundamental principles that distinguish it from frequentist approaches. Then transition to the concept of conjugate priors, explaining how these specially chosen prior distributions yield posterior distributions of the same family, facilitating analytical tractability.

Moreover, the research showcases real-world examples where Bayesian estimation with conjugate priors has proven to be a powerful tool, enhancing decision-making processes in situations ranging from medical diagnostics to quality control in manufacturing. By illustrating the versatility and efficiency of this methodology, we aim to empower readers to harness the full potential of Bayesian analysis in their own pursuits.

The Linex loss function, short for Linear Exponential loss function, is a variant of the asymmetric loss functions commonly used in regression analysis. Unlike traditional symmetric loss functions like Mean Squared Error (MSE) or Mean Absolute Error (MAE), Linex loss asymmetrically penalizes overestimation and underestimation differently. It is particularly useful when the cost of underestimation is not the same as the cost of overestimation, making it suitable for scenarios where errors in one direction are more critical than errors in the other. The Linex loss function is defined as

The LINEX loss function you provided is:

$$\mathrm{L}(\theta,\hat{\theta}) = \mathrm{a}.\,\mathrm{e}^{\mathrm{b}(\theta-\hat{\theta})} - \mathrm{b}(\theta-\hat{\theta}) - 1$$

where:

- $\mathrm{L}(\theta,\hat{\theta})$ is the LINEX loss function.
- θ is the true parameter value.
- $\hat{\theta}$ is the estimated parameter value.
- a and b are parameters that control the shape of the loss function.

This type of LINEX loss function is sometimes used in Bayesian estimation, and Zellner is indeed associated with Bayesian methods. In Bayesian statistics, the choice of a loss function is crucial in constructing a suitable posterior distribution. The LINEX loss function, as you've written it, is a combination of linear and exponential terms, and the parameters a and b determine the weight given to these terms.

A symmetric loss function is a mathematical function used to measure the discrepancy or error between predicted and actual values in a regression problem. Unlike asymmetric loss functions, which penalize overestimation and underestimation differently, symmetric loss functions treat overestimation and underestimation equally

The symmetric loss function $\mathrm{L}(\theta,\mathrm{d}) = \mathrm{C}(\mathrm{d}-\theta)^{2f}$ penalizes the deviation between the decision d and the unknown parameter θ. Here, C is a scaling constant, and f is a parameter that controls the sensitivity of the loss function to deviations. This loss function is symmetric because it penalizes deviations equally on both sides of the decision d. The exponent 2 f controls the curvature of the loss function around d. larger values of f make the loss function more sensitive to deviations from d, leading to sharper penalties.

In the context of Bayesian estimation, our exploration centers on the symmetric loss function, elegantly expressed as

$$\mathrm{L}(\theta,\mathrm{d}) = \mathrm{C}(\mathrm{d}-\theta)^{2f}$$

with C serving as a constant. The transformation into the quadratic loss function (QLF) occurs when f assumes the value of 1, resulting in the concise form

$$L(\theta, d) = C(d - \theta)^2$$

By streamlining the equation through the abstraction of C to 1, we seamlessly transition to the squared error loss function (SELF). Introducing an alternative, the absolute loss function takes the form L (θ, d) = | (d − θ) |. Notably, the squared error loss function (SELF)

$$L(\theta, d) = (d - \theta)^2$$

The goal in Bayesian estimation is to find the posterior distribution of the parameter θ given the observed data. This involves combining the likelihood function with a prior distribution and the loss function. The posterior distribution is then obtained by maximizing the posterior expected loss (also known as the Bayes risk) with respect to θ or using other Bayesian decision theoretic criteria.

II. Review of literature

Box, G. E. P., Tiao, G. C., and Jenkins, G. M [6] done a foundational work in the field of Bayesian statistics. Zellner introduces the concept of Bayesian estimation and prediction with asymmetric loss functions and computational methods for Bayesian estimation and prediction using asymmetric loss functions [20]. Parsian introduces the concept of Bayes estimation using a LINEX loss function and explained its uses in Bayesian estimation and its advantages in decision-making under uncertainty [18]. Feroze, N. and Aslam, M. discusses [12] Bayesian analysis of the error function distribution using various loss functions and examine how different loss functions impact Bayesian estimation in the context of the error function distribution. Zaka, A. and Akhter, A. S. compares [19] various methods for estimating parameters of the power function distribution and done a simulation study and real-world. Chrisogonus K. Onyekwere and Okechukwu J. Obulezi [8] have proposed a new one-parameter distribution named Chris-Jerry is suggested from a two-component mixture of Exponential (θ) distribution and Gamma (3, θ) distribution with mixing proportion p = θ/ θ+2 having a flexibility advantage in modeling lifetime data. In this paper the posterior mean of Chris-jerry distribution is derived with various loss function.

III. A Bayesian approach for Chris-jerry distribution

The probability distribution function of Chris-jerry distribution is given by

$$f(x) = \frac{\theta^2}{\theta+2}.(1 + \theta x^2).e^{-\theta x} \qquad x, \theta > 0 \tag{1}$$

In this section the posterior distribution of Chris-jerry distribution is obtained. Let $X_1, X_2, \ldots$ be a sequence of random variables from Chris-jerry distribution, then the likelihood function is given by

$$\pi(x_i) = \prod_{i=1}^{n} \frac{\theta^2}{\theta+2}.(1 + \theta x_i^2).e^{-\theta x_i} \tag{2}$$

The prior is gamma prior (conjugate prior)

$$p(\theta) = \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} \qquad r > 0, \theta > 0 \tag{3}$$

The posterior distribution is given by

$$p(\theta/x) = \frac{1}{k} * \frac{\theta^{2n}}{(\theta+2)^{n}} * \prod_{i=1}^{n}(1+\theta x_i) \quad * \left(e^{-\theta \sum_1^n x_i}\right) * \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} \tag{4}$$

Were

$$k = \int_0^{\infty} \frac{\theta^{2n}}{(\theta+2)^{n}}(1+\theta x_i)^{n}.\left(e^{-\theta \sum_1^n x_i}\right) * \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} d\theta$$

$$K = \sum_{j=0}^{\infty}(-1)^{j} c(n+j-1,j) \sum_{l=0}^{\infty} c(n,i)(x)^{l} \frac{\{\Upsilon(2n+l+j+r)\}}{2^{n+j}[\Upsilon r][\sum_1^n x_i + 1]^{2n+l+j+r}}$$

The posterior mean is given by

$$E[\theta] = \int_0^{\infty} \theta p(\theta/x) d\theta$$

$$= \int_0^{\infty} \theta \frac{1}{k} * \frac{\theta^{2n}}{(\theta+2)^{n}} * \prod_{i=1}^{n}(1+\theta x_i) \quad * \left(e^{-\theta \sum_1^n x_i}\right) * \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} d\theta$$

Case: I

Bayesian estimation of θ under linex loss function by Zellner [**Zellner, A. (1986)**.]

$$L(\theta, \hat{\theta}) = a.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1 \tag{5}$$

Where a>0, b≠0; a is scale of loss function and b determines its shape. Without loss of generality, we assume a= 1 and obtain bayes estimate of θ

In Zeller's linex loss function, $\hat{\theta}$ represents the reference value or the target value that the parameter θ is compared to $\hat{\theta}$. It can be thought of as the "ideal" or "desired" value of θ. The linex loss function measures the deviation of θ from $\hat{\theta}$

For example, if you are estimating a parameter θ and you have a prior belief or expectation about its value, you might set $\hat{\theta}$to that prior belief. Then, the linex loss function would measure how much the estimated value of θ deviates from that prior belief.

Here

$$E[L(\theta, \hat{\theta})] = \frac{1}{k\Gamma r} \int_0^{\infty} [1.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1] * p(\theta/x)\, d\theta$$

$$= \int_0^{\infty} [1.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1] * \frac{\theta^{2n+r-1}}{\Upsilon r(\theta+2)^{n}}(1+\theta x_i)^{n}.\left(e^{-\theta \sum_1^n x_i - \theta}\right) d\theta$$

$$= \frac{1}{\Upsilon r} \int_0^{\infty} [1.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1] * \frac{\theta^{2n+r-1}}{(\theta+2)^{n}}(1+\theta x_i)^{n}.\left(e^{-\theta \sum_1^n x_i - \theta}\right) d\theta$$

$$\sum_{j=0}^{\infty}(-1)^{j} c(n+j-1, j) * \left(\frac{1}{2}\right)^{j} = d$$

$$\sum_{l=0}^{\infty} c(n,i)(x)^{l} = f$$

$$\sum_{i=1}^{n} x_i + 1 = h$$

$$E\left[L(\theta,\hat{\theta})\right] = \frac{1df}{\Upsilon r}\left[\left(\left\{\frac{e^{-b\hat{\theta}}.\Upsilon(2n+r+j+l)}{[h-b]^{2n+r+j+l}}\right\}\right) - \left\{\frac{b(\Upsilon 2n+r+j-l)}{h^{2n+r+j-l}}\right\} + \hat{\theta}b\left\{\frac{\Upsilon 2n+r+j+l}{h^{2n+r+j+l}}\right\} - \left\{\frac{\Upsilon 2n+r+j+l}{h^{2n+r+j+l}}\right\}\right]$$

Case: II (Bayesian estimation of θ under Symmetric loss function)
Symmetric loss function for the decision d for the unknown parameter θ is defined by

$$L(\theta,d) = C(d-\theta)^{2f} \quad (6)$$

f= 1, 2…..

$$E[L(\theta,d)] = \frac{1}{k\Gamma r}\int_{0}^{\infty}[C(d-\theta)^{2f}] * \frac{\theta^{2n+r-1}}{(\theta+2)^{n}}.\left(e^{-\theta(\sum_{1}^{n}x_i+1)}\right)\prod_{i=1}^{n}(1+\theta x_i)\quad d\theta$$

where C is a constant. When f = 1 reduces to quadratic loss function (QLF) given by

$$L(\theta,d) = C(d-\theta)^{2} \quad (7)$$

For some constant C. The value of the constant C makes no difference to a decision, and can be ignored by setting it equal to 1 and reduced to the SELF. Absolute loss function is another symmetric loss function given by

$$L(\theta,d) = |(d-\theta)| \quad (8)$$

The squared error loss function (SELF) is widely used in decision theory problems and is defined as

$$L(\theta,d) = (d-\theta)^{2} \quad (9)$$

IV. Simulation Study

Posterior Distribution: The posterior distribution represents our updated beliefs about the parameter θ after observing the data. It is proportional to the product of the prior distribution and the likelihood function. Sampling: in this research Markov Chain Monte Carlo (MCMC) methods is used, implemented in the r software, to sample from the posterior distribution and estimate the parameters of interest.

The following algorithm outlines the steps involved in Bayesian inference using the specified model and data. It involves defining the model, computing the likelihood, performing posterior inference, analyzing the results, and outputting the estimates and diagnostics.

Stan Code Explanation

- Input Data:

Obtain input data:

n: Number of data points.

x: Observed data points.

r: Parameter influencing the shape of the gamma prior distribution for theta.

C: Constant used in the loss function.

f: Exponent used in the loss function.

d: Target value used in the loss function.

- Initialize Model:

  Define the prior distribution:

  θ ~ Gamma (shape = r, rate = 1).

- Likelihood Calculation:

  Compute the likelihood using a custom loss function:

  Calculate the log-likelihood contribution for each data point:

  Compute terms related to the observed data x, the parameter theta, and the loss function parameters C, f, and d.

  Accumulate the log-likelihood contributions.

- Posterior Inference:

  Combine the prior distribution and likelihood to obtain the posterior distribution:

  Posterior ∝ Prior × Likelihood.

- Bayesian Inference:

  Use Bayesian inference techniques (such as Markov Chain Monte Carlo) to sample from the posterior distribution:

  Obtain posterior samples for the parameter theta using Stan's sampling algorithm.

  Specify the number of chains and iterations for sampling.

- Analysis:

  Analyze the posterior samples to estimate the posterior distribution of theta:

  Compute summary statistics (e.g., mean, median, quantiles) of the posterior samples.

  Visualize the posterior distribution if necessary.

- Output:

  Output the results of the analysis, such as posterior mean estimates, credible intervals, and diagnostic information about the inference procedure.

Table 1: *Comparison of Posterior Means for Different Loss Functions and Sample Sizes (simulated data)*

n/ $E[\theta]$	50	100	200
Without loss	0.0861636	0.04134565	0.02038608
Symmetric loss function	3.458547	3.045888	2.563339
Quadratic loss function (QLF)	0.08745913	0.04159895	0.02039235
Squared error loss function	0.08671715	0.04156098	0.02036867
Linex loss function	0.08599036	0.04149021	0.0203717

From the above table show that the posterior mean increase when d and f values increase. Which mean the larger values of f make the loss function more sensitive to deviations from d, leading to sharper penalties.

V. Real life data

Table: 2 *shows the life of fatigue fracture of Kevlar 373/epoxy subjected to constant pressure at 90 % stress level until all had failed. Source:(8)*

0.0251	0.6748	0.912	1.3503	1.7746	2.0408	2.4951	4.8073
0.0886	0.6751	0.9836	1.3551	1.8475	2.0903	2.526	5.4005
0.0891	0.6753	1.0483	1.4595	1.8375	2.1093	2.9911	5.4435
0.2501	0.7696	1.0596	1.488	1.8503	2.133	3.0256	5.5295
0.3113	0.8375	1.0773	1.5728	1.8808	2.21	3.2678	6.5541
0.3451	0.8391	1.1733	1.5733	1.8878	2.246	3.4045	9.096
0.4763	0.8425	1.257	1.7083	1.8881	2.2878	3.4846	
0.565	0.8645	1.2766	1.7263	1.9316	2.3203	3.7433	
0.5671	0.8851	1.295	1.746	1.9558	2.347	3.7455	
0.6566	0.9113	1.3211	1.763	2.0048	2.3513	3.9143	

.

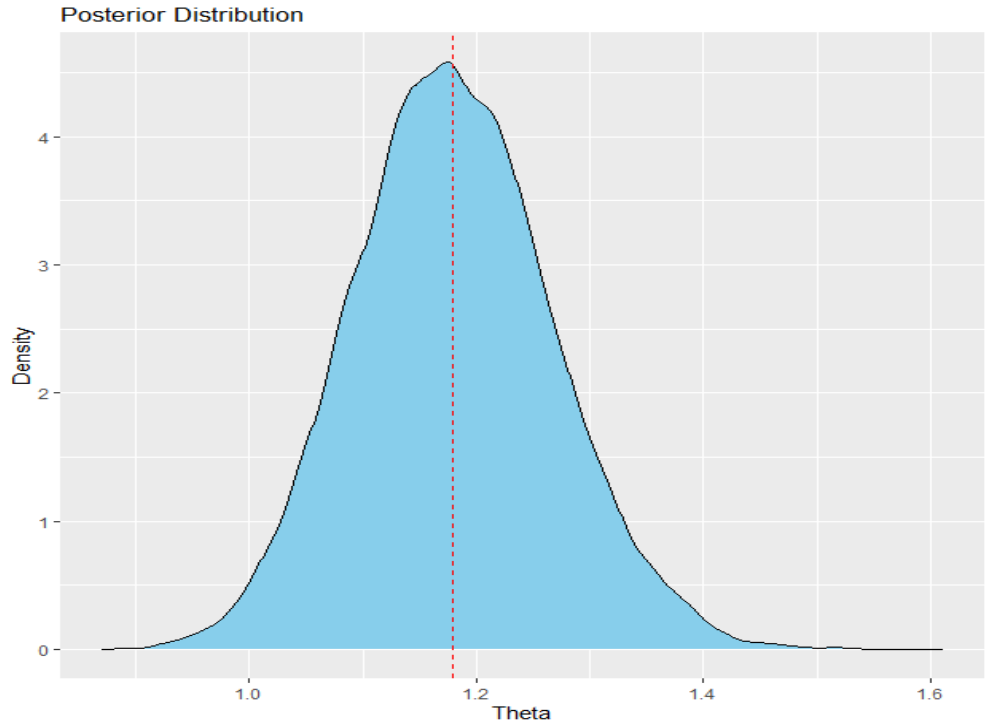


Figure: 1 *Graph of posterior distribution of Chris-jerry distribution*

Table: 3 *Comparison of Posterior Means for Different Loss Functions (real-life data)*

Loss functions	Posterior Mean
Without any loss function	1.179317
Symmetric loss function	1.220796
quadratic loss function	1.200863
Squared error loss function	1.201302
Linex loss function	1.197762

The above presents a comparative analysis of posterior means under different loss functions, revealing notable disparities in estimation outcomes. Without any loss function, the posterior mean is observed to be substantially higher, suggesting potential bias in the estimation process. Conversely, employing loss functions leads to a decrease in the posterior mean, highlighting the influence of the chosen loss function. Specifically, the symmetric loss function yields the highest posterior mean, while the Linex loss function results in the lowest. These findings underscore the significance of selecting an appropriate loss function. Notably, the posterior mean of the provided data has been calculated, emphasizing the practical relevance of these results.

VI. Conclusion

The Bayesian estimation of parameters for the Chris-Jerry distribution with a gamma prior, considering various loss functions, reveals subtle differences in posterior mean estimates. The absence of a specific loss function yields a posterior mean estimate of 1.179317, while employing a symmetric loss function slightly increases the estimate to 1.220796. Conversely, quadratic and squared error loss functions result in slightly lower estimates of 1.200863 and 1.201302, respectively. The use of a Linex loss function produces a posterior mean estimate of 1.197762. These findings underscore the importance of the choice of loss function in Bayesian estimation. While variations are observed in the posterior mean estimates across different loss functions, the differences remain relatively subtle, indicating robustness in the estimation process. However, it is essential to note that these results are contingent upon the provided dataset and may vary with alternative datasets or priors. In conclusion, this study contributes to the understanding of Bayesian estimation methods for the Chris-Jerry distribution with a gamma prior. Future research could delve deeper into exploring additional loss functions and their implications for parameter estimation in Bayesian frameworks, thereby enhancing the applicability of these methods in diverse statistical analyses.

References

[1] Ahmed, A., Ahmad, S., & Reshi, J. (2013). Bayesian analysis of Rayleigh distribution. International Journal of Scientific and Research Publications, 3(10), 1–9.

[2] Ahmed, M. A. (2020). On the alpha power Kumaraswamy distribution: Properties, simulation and application. Revista Colombiana de Estadística, 43(2), 285–313. https://doi.org/10.15446/rce.v43n2.83598

[3] Ahmed, Z., Nofal, Z. M., Abd El Hadi, N. E., et al. (2015). Exponentiated transmuted generalized Rayleigh distribution: A new four-parameter Rayleigh distribution. Pakistan Journal of Statistics and Operation Research, 11(1), 115–134. https://doi.org/10.18187/pjsor.v11i1.873

[4] Andrews, D. F., & Herzberg, A. M. (1985). Stress-rupture life of Kevlar 49/epoxy spherical pressure vessels. Data. Springer, 181–186.

[5] Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis. Springer.

[6] Box, G. E. P., Tiao, G. C., & Jenkins, G. M. (1970). Bayesian Statistical Inference. John Wiley & Sons.

[7] Carlin, B. P., & Louis, T. A. (2009). Bayesian Methods for Data Analysis. CRC Press.

[8] Onyekwere, C. K., & Obulezi, O. J. (2022). Chris-Jerry Distribution and Its Applications. Asian Journal of Probability and Statistics, 20(1), 16–30. ISSN: 2582-0230

[9] Dey, D. K., & Liu, P.-S. L. (1992). On comparison of estimators in a generalized life model. Microelectronics Reliability, 32(1-2), 207–221.

[10] Dey, D. K., Ghosh, M., & Srinivasan, C. (1987). Simultaneous estimation of parameters under entropy loss. Journal of Statistical Planning and Inference, 15, 347–363.

[11] 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.

[12] Feroze, N., & Aslam, M. (2012). A note on Bayesian analysis of error function distribution under different loss functions. International Journal of Probability and Statistics, 1(5), 153–159.

[13] Meenakshi, G. (2013). A Bayesian method to Predetermination of Viral replication in the HIV Dynamics. Fast East Journal of Theoretical Statistics, 42(2), 71–81.

[14] Meenakshi, G., et al. (2019). Prediction of HIV replication in the human immune system using multinomial distribution by Bayesian methodology. International Journal of Scientific Research in Mathematical and Statistical Sciences, 6(1), 46–52.

[15] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehari, A., & Rubin, D. B. (2013). Bayesian Data Analysis. Chapman and Hall/CRC.

[16] Hobbs, N. T., & Hooten, M. B. (2015). Bayesian Models: A Statistical Primer for Ecologists. Princeton University Press.

[17] Lindley, D. V., & Smith, A. F. (1972). Bayes estimates for the linear model. Journal of the Royal Statistical Society. Series B (Methodological), 34(1), 1–41.

[18] Parsian, A. (1990). Bayes estimation using a LINEX loss function. Risk, 1(4), 305–307.

[19] Zaka, A., & Akhter, A. S. (2013). Methods for estimating the parameters of the power function distribution. Pakistan Journal of Statistics and Operational Research, 9(2), 213–224.

[20] Zellner, A. (1986). Bayesian Estimation and Prediction Using Asymmetric Loss Functions. Journal of the American Statistical Association, 81, 446–450.

PROFIT AND AVAILABILITY ANALYSIS OF UTENSIL INDUSTRY SUBJECT TO REPAIR FACILITY

Amit Kumar[1], Pinki Kumari[2]

•

[1]Department of Mathematics, Govt. College, Satnali, Haryana
[2]Department of Physics, Lord University, Chikani, Alwar Rajasthan
prof.amitmalik@gmail.com, prof.pinkimalik@gmail.com

Abstract

The main objective of the paper is to optimize the availability and profit values of the utensil industry. There are three distinct units in the utensil industry and two of them work in reduced state. It is assumed that unit A is failed in complete failure mode while units B and C completely failed through partial failure mode. The system is in a failing condition when one of the units completely fails. There is a qualified technician to fix the fault in the system. Timelines for failure and repair are unrelated to one another. The repair time is exponential while the failure time distribution is general. Many factors, including mean time to system failure, availability, busy period, estimated number of server visits and profit values are calculated from tables.

Keywords: Reliability, base state, mean sojourn time, availability and profit.

I. Introduction

Manufacturers and industrialists continuously produce goods to meet the rising demand for goods, which they can do by optimizing their manufacturing procedures. For product development, reliability engineering presents an integrated approach that helps in the design and maintenance of the products. To calculate the profit values, it is important to analyze the availability of the system. This study examines the MTSF, availability and profit values of the housewares sector, emphasizing the need of utilizing the regenerative point graphical technique for priority in repair under specific circumstances. A large amount of research work has been done on repairable systems such that Kapur and Kapoor [5] analyzed the behaviour of a two unit system subject to repair facility and one spare unit is kept in cold standby mode. Gnedenko and Igor [3] explored reliability and probability measures to solve the complexity of repairable system. Jack and Murthy [4] discovered the role of limited warranty and extended warranty for the product. Wang and Zhang [10] examined the repairable system of two non identical components under repair facility using geometric distributions. Kumar and Goel [7] threw light on the preventive maintenance in two unit cold standby repairable system under general distributions. Chaudhary and Tomar [2] threw light on the stochastic behavior of a two unit cold standby system under inspection. Kumar et al. [6] described the effects of washing unit in the paper industry by using the regenerative point graphical technique. Levitin et al. [8] explored the results of optimal preventive replacement of failed units in a cold standby system by using the poisson process. Agarwal et al. [1] analyzed the performance and reliability of water treatment plant under repair facility. Sengar and Mangey [9] examined the availability and profit values of complicated repairable system with inspection using copula methodology.

II. System Assumptions

There are following system assumptions:

- The utensil industry has three distinct units such that cutting system, pressing system, spinning and buffing system.
- Sheets are cut into circular sheets with the help of cutting system (*A*).
- Sheets are converted into the shape of utensils by using pressing technology (*B*).
- Spinning and buffing machinery give the final shape and polish to the utensil.
- It is considered that units *B* and *C* may be in a complete failed state through partial failure but unit *A* is in only complete failed state.
- Failure time follows general distribution whereas repair time follows the exponential distribution.
- A technician is always available to repair the failed unit.
- The failed unit works like a new unit after repair.

III. System Notations

There are following system notations:

$i \xrightarrow{Sr} j$	r^{th} directed simple path from state '*i*' to state '*j*' where '*r*' takes the positive integral values for different directions from state '*i*' to state '*j*'.
$\xi \xrightarrow{s\,f\,f} i$	A directed simple failure free path from state ξ to state '*i*'.
$m-cycle$	A circuit (may be formed through regenerative or non regenerative / failed state) whose terminals are at the regenerative state '*m*'.
$m-\overline{cycle}$	A circuit (may be formed through the unfailed regenerative or non regenerative state) whose terminals are at the regenerative '*m*' state.
$U_{k,k}$	Probability factor of the state '*k*' reachable from the terminal state '*k*' of '*k*' cycle.
$U_{\overline{k,k}}$	The probability factor of state '*k*' reachable from the terminal state '*k*' of $k\ \overline{cycle}$.
μ_i	Mean sojourn time spent in the state '*i*' before visiting any other states.
μ_i'	Total unconditional time spent before transiting to any other regenerative state while the system entered regenerative state '*i*' at t=0.
η_i	Expected waiting time spent while doing a job given that the system entered to the regenerative state '*i*' at *t*=0.
A/a	System first unit is in the operative state/failed state.
$B/\overline{B}/b$	System second unit is in the operative state/reduced state/failed state.
$C/\overline{C}/c$	System third unit is in the operative state/reduced state/failed state.
α_1 / α_3	The constant partial failure rate of the unit B/C respectively.
α_2 / α_4	The constant complete failure rate of the unit B/C respectively.
α_5	The constant complete failure rate of unit A.
$f_1(t)/F_1(t)$	PDF/CDF of repair time of unit B from partial failed state.
$f_2(t)/F_2(t)$	PDF/CDF of repair time of unit B from complete failed state.
$f_3(t)/F_3(t)$	PDF/CDF of repair time of unit C from partial failed state.
$f_4(t)/F_4(t)$	PDF/CDF of repair time of unit C from complete failed state.
$f_5(t)/F_5(t)$	PDF/CDF of repair time of unit A from complete failed state.

IV. Transition Diagram and Their Descriptions

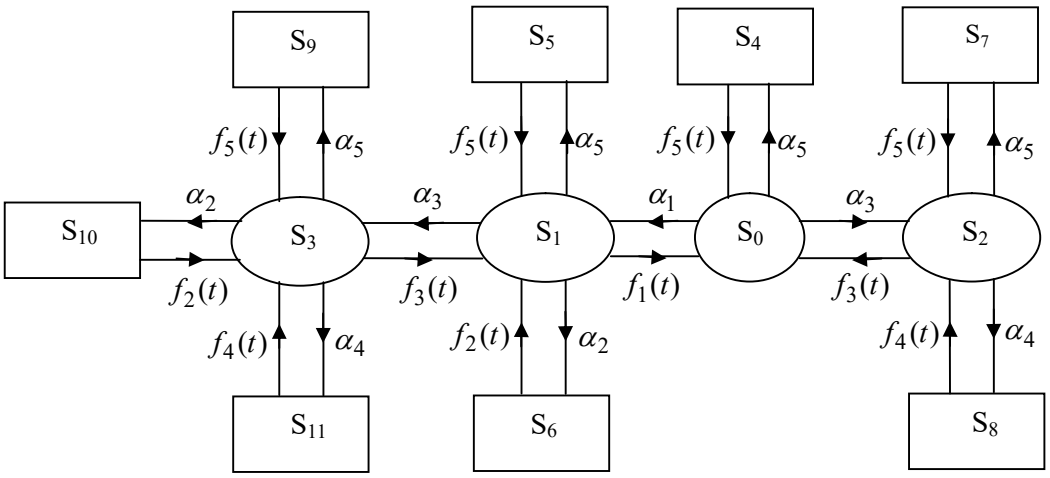


Figure 1: *State Transition Diagram*

In the system transition diagram, there are following states

where, $S_0 = ABC$, $S_1 = A\overline{B}C$, $S_2 = AB\overline{C}$, $S_3 = A\overline{BC}$, $S_4 = aBC$, $S_5 = a\,\overline{B}\,C$

$S_6 = AbC$, $S_7 = aB\overline{C}$, $S_8 = ABc$, $S_9 = a\overline{BC}$, $S_{10} = Ab\overline{C}$, $S_{11} = A\overline{B}c$

V. Transition Probabilities

The transition probabilities are following

$p_{0,1} = \alpha_1/(\alpha_1+\alpha_3+\alpha_5)$, $p_{0,2} = \alpha_2/(\alpha_1+\alpha_3+\alpha_5)$

$p_{0,4} = \alpha_5/(\alpha_1+\alpha_3+\alpha_5)$, $p_{1,0} = \beta_1/(\beta_1+\alpha_2+\alpha_3+\alpha_5)$

$p_{1,3} = \alpha_3/(\beta_1+\alpha_2+\alpha_3+\alpha_5)$, $p_{1,5} = \alpha_5/(\beta_1+\alpha_2+\alpha_3+\alpha_5)$

$p_{1,6} = \alpha_2/(\beta_1+\alpha_2+\alpha_3+\alpha_5)$, $p_{2,0} = \beta_3/(\beta_3+\alpha_4+\alpha_5)$

$p_{2,7} = \alpha_5/(\beta_3+\alpha_4+\alpha_5)$, $p_{2,8} = \alpha_4/(\beta_3+\alpha_4+\alpha_5)$

$p_{3,1} = \beta_3/(\beta_3+\alpha_2+\alpha_4+\alpha_5)$, $p_{3,9} = \alpha_5/(\beta_3+\alpha_2+\alpha_4+\alpha_5)$

$p_{3,10} = \alpha_2/(\beta_3+\alpha_2+\alpha_4+\alpha_5)$, $p_{3,11} = \alpha_4/(\beta_3+\alpha_2+\alpha_4+\alpha_5)$

$$p_{4,0} = p_{5,1} = p_{6,1} = p_{7,2} = p_{8,2} = p_{9,3} = p_{10,3} = p_{11,3} = 1 \quad (1)$$

It has been concluded that

$$p_{0,1} + p_{0,2} + p_{0,4} = 1,\ p_{1,0} + p_{1,3} + p_{1,5} + p_{1,6} = 1$$

$$p_{2,0} + p_{2,7} + p_{2,8} = 1,\ p_{3,1} + p_{3,9} + p_{3,10} + p_{3,11} = 1 \quad (2)$$

VI. Mean Sojourn Time

Let μ_i represents the mean sojourn time. Mathematically, the time taken by a system in a particular state becomes

$$\mu_i = \sum_j m_{i,j} = \int_0^\infty P(T > t)\,dt\ .$$

and $\mu_0 = 1/(\alpha_1 + \alpha_3 + \alpha_5)$, $\mu_1 = 1/(\beta_1 + \alpha_2 + \alpha_3 + \alpha_5)$, $\mu_2 = 1/(\beta_2 + \alpha_4 + \alpha_5)$
$\mu_3 = 1/(\beta_3 + \alpha_2 + \alpha_4 + \alpha_5)$, $\mu_4(t) = \mu_5(t) = 1/(\alpha_5)$, $\mu_6 = \mu_{10} = 1/(\alpha_2)$
$\mu_7 = \mu_9 = 1/(\alpha_5)$, $\mu_8 = \mu_{11} = 1/(\alpha_4)$ (3)

VII. Evaluation of Parameters

Using the regenerative point graphical technique, all reliability parameters (including mean time to system failure, availability, busy period of technician and expected number of technician visits) are calculated. It is considered that

$f_1(t) = \beta_1 e^{-\beta_1 t}$, $f_2(t) = \beta_2 e^{-\beta_2 t}$, $f_3(t) = \beta_3 e^{-\beta_3 t}$, $f_4(t) = \beta_4 e^{-\beta_4 t}$, $f_5(t) = \beta_5 e^{-\beta_5 t}$
and, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha$, $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta$.

I. Mean Time to System Failure

The system can transition to regenerative un-failed states (i=0, 1, 2, 3) using initial state 0 before reaching any failed state (using base state ξ=0). At that point, MTSF becomes

$$T_0 = \left[\sum_{i=0}^{3} Sr\left\{\frac{\left\{pr(0 \xrightarrow{Sr(sff)} i)\right\}.\mu_i}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}}\right\}\right] \div \left[1 - \sum Sr\left\{\frac{\left\{pr(0 \xrightarrow{Sr(sff)} 0)\right\}}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}}\right\}\right]$$
$$= [\mu_0 + p_{0,1}\mu_1 + p_{0,2}\mu_2 + p_{0,1}p_{1,3}\mu_3]/[1 - p_{0,1}p_{1,0} - p_{0,2}p_{2,0}]$$
$$= \frac{[(\beta + 3\alpha)(\beta + 4\alpha) + \alpha^2](\beta + 2\alpha)}{(\beta + 3\alpha)\left[3\alpha(\beta + 3\alpha) + (\beta + 2\alpha) - \alpha\beta(\beta + 2\alpha) - \alpha\beta(\beta + 3\alpha)\right]} \quad (4)$$

II. Availability of the System

The system is available for use at regenerative states j=0, 1, 2, 3 with ξ=0 then the availability of system is defined as

$$A_0 = \left[\sum_{j=0}^{3} Sr\left\{\frac{\left\{pr(0 \xrightarrow{Sr} j)\right\}.f_j.\mu_j}{\prod_{k_1 \neq 0}\left\{1 - V_{k_1 k_1}\right\}}\right\}\right] \div \left[\sum_{i=0}^{11} Sr\left\{\frac{\left\{pr(0 \xrightarrow{Sr} i)\right\}.\mu_i'}{\prod_{k_2 \neq 0}\left\{1 - V_{k_2 k_2}\right\}}\right\}\right]$$
$$= \frac{[U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3]}{\left[\begin{array}{l} U_{0,0}\mu_0 + U_{0,1}\mu_1 + U_{0,2}\mu_2 + U_{0,3}\mu_3 + U_{0,4}\mu_4 + U_{0,5}\mu_5 \\ + U_{0,6}\mu_6 + U_{0,7}\mu_7 + U_{0,8}\mu_8 + U_{0,9}\mu_9 + U_{0,10}\mu_{10} + U_{0,11}\mu_{11} \end{array}\right]}$$

$$= \frac{\beta^3(\beta+2\alpha)[(\beta+3\alpha)^3+\alpha(\beta+2\alpha)(\beta+3\alpha)]}{\left[\begin{array}{l}\alpha^3(\beta+2\alpha)[(\beta+3\alpha)^3+\alpha(\beta+2\alpha)(\beta+3\alpha)] \\ +\alpha\beta^2(\beta+2\alpha)(\beta+3\alpha)^3(\beta+2\alpha) \\ +[\alpha^2\beta(2\beta+5\alpha)]\beta(\beta+2\alpha)(\beta+3\alpha)+3\alpha^3\beta^2(\beta+2\alpha)^2\end{array}\right]} \quad (5)$$

III. Busy Period of the Technician

The technician is busy due to repair of the failed unit at regenerative states *j*= 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 with base state ξ = 0 then the fraction of time for which the technician remains busy is defined as

$$B_0 = \left[\sum_{j=1}^{11} Sr\left\{\frac{\left\{pr(0\xrightarrow{Sr} j)\right\}.\eta_j}{\prod_{k_1\neq 0}\left\{1-V_{k_1k_1}\right\}}\right\}\right] \div \left[\sum_{i=0}^{11} Sr\left\{\frac{\left\{pr(0\xrightarrow{Sr} i)\right\}.\mu_i'}{\prod_{k_2\neq 0}\left\{1-V_{k_2k_2}\right\}}\right\}\right]$$

$$= \frac{\left[\begin{array}{l}U_{0,1}\mu_1+U_{0,2}\mu_2+U_{0,3}\mu_3+U_{0,4}\mu_4+U_{0,5}\mu_5+U_{0,6}\mu_6 \\ +U_{0,7}\mu_7+U_{0,8}\mu_8+U_{0,9}\mu_9+U_{0,10}\mu_{10}+U_{0,11}\mu_{11}\end{array}\right]}{\left[\begin{array}{l}U_{0,0}\mu_0+U_{0,1}\mu_1+U_{0,2}\mu_2+U_{0,3}\mu_3+U_{0,4}\mu_4+U_{0,5}\mu_5 \\ +U_{0,6}\mu_6+U_{0,7}\mu_7+U_{0,8}\mu_8+U_{0,9}\mu_9+U_{0,10}\mu_{10}+U_{0,11}\mu_{11}\end{array}\right]}$$

$$= \frac{\left[\begin{array}{l}\beta^3(\beta+2\alpha)[(\beta+3\alpha)^2\alpha+\alpha(\beta+2\alpha)(\beta+3\alpha)] \\ +\alpha\beta^2(\beta+2\alpha)(\beta+3\alpha)^3(\beta+2\alpha) \\ +[\alpha^2\beta(2\beta+5\alpha)]\beta(\beta+2\alpha)(\beta+3\alpha)+3\alpha^3\beta^2(\beta+2\alpha)^2\end{array}\right]}{\left[\begin{array}{l}\alpha^3(\beta+2\alpha)[(\beta+3\alpha)^3+\alpha(\beta+2\alpha)(\beta+3\alpha)] \\ +\alpha\beta^2(\beta+2\alpha)(\beta+3\alpha)^3(\beta+2\alpha) \\ +[\alpha^2\beta(2\beta+5\alpha)]\beta(\beta+2\alpha)(\beta+3\alpha)+3\alpha^3\beta^2(\beta+2\alpha)^2\end{array}\right]} \quad (6)$$

IV. Estimated Number of Visits Made by the Technician

The technician visits at regenerative states *j*= 1, 2, 3 with ξ=0 then the number of visits by the repairman is defined as

$$V_0 = \left[\sum_{j=1}^{3} Sr\left\{\frac{\left\{pr(0\xrightarrow{Sr} j)\right\}}{\prod_{k_1\neq 0}\left\{1-V_{k_1k_1}\right\}}\right\}\right] \div \left[\sum_{i=0}^{11} Sr\left\{\frac{\left\{pr(0\xrightarrow{Sr} i)\right\}.\mu_i'}{\prod_{k_2\neq 0}\left\{1-V_{k_2k_2}\right\}}\right\}\right]$$

$$= \frac{[U_{0,1}\mu_1+U_{0,2}\mu_2+U_{0,3}\mu_3]}{\left[\begin{array}{l}U_{0,0}\mu_0+U_{0,1}\mu_1+U_{0,2}\mu_2+U_{0,3}\mu_3+U_{0,4}\mu_4+U_{0,5}\mu_5 \\ +U_{0,6}\mu_6+U_{0,7}\mu_7+U_{0,8}\mu_8+U_{0,9}\mu_9+U_{0,10}\mu_{10}+U_{0,11}\mu_{11}\end{array}\right]}$$

$$= \frac{\beta^3(\beta+2\alpha)[(\beta+3\alpha)^2\alpha+\alpha(\beta+2\alpha)(\beta+3\alpha)]}{\left[\begin{array}{l}\alpha^3(\beta+2\alpha)[(\beta+3\alpha)^3+\alpha(\beta+2\alpha)(\beta+3\alpha)] \\ +\alpha\beta^2(\beta+2\alpha)(\beta+3\alpha)^3(\beta+2\alpha) \\ +[\alpha^2\beta(2\beta+5\alpha)]\beta(\beta+2\alpha)(\beta+3\alpha)+3\alpha^3\beta^2(\beta+2\alpha)^2\end{array}\right]} \quad (7)$$

V. Profit Analysis

If a system produces revenue for its developer, it is considered valuable. The availability of the system, busy period of the technician and expected number of visits by the technician are taken into consideration to calculate the profit values of the system. The profit function may be used to do the profit analysis of the system and it is given by

$$P = E_0 A_0 - E_1 B_0 - E_2 V_0 \quad (8)$$

where, $E_0 = 8000$ (Pay per unit uptime of the system)

$E_1 = 500$ (Charge per unit time for which technician is busy)

$E_2 = 200$ (Charge per visit of the technician)

VIII. Discussion

Table 1 describes the nature of the mean time to system failure of the utensil industry. It has an

Table 1: *MTSF vs. Repair Rate*

β ↓	α=0.025	α=0.04	α=0.06
0.01	4.357262	4.200299	3.696809
0.02	4.600326	4.405797	3.903394
0.03	4.832536	4.599156	4.102564
0.04	5.054602	4.781421	4.29471
0.05	5.267176	4.953519	4.480198
0.06	5.470852	5.116279	4.659367
0.07	5.666179	5.27044	4.832536
0.08	5.853659	5.416667	5.25355
0.09	6.033755	5.555556	5.162037
0.10	6.206897	5.687646	5.318907

increasing trend corresponding to increment in repair rate (β) and has decreasing trend corresponding to an increment in failure rate (α). In the above table, the values of parameters are α= 0.025, 0.04, 0.06 and β=0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10 respectively. When the value of repair rate enhances then MTSF values are also enhanced. When α=0.025 changes into α=0.04, 0.06 then MTSF values are declined.

Table 2 explores the increasing trends of availability with respect to increments in repair rate (β) and has decreasing trends corresponding to increments in failure rate (α). When the value of the repair rate is enhanced then the availability values are also enhanced. Also, when the failure rate of unit changes α=0.025 to 0.04, 0.06 then the availability of system declines.

Table 2: *Availability vs. Repair Rate*

β ↓	α=0.025	α=0.04	α=0.06
0.01	0.623324	0.604782	0.542904
0.02	0.628307	0.609813	0.547959
0.03	0.633159	0.614717	0.552904
0.04	0.637887	0.619499	0.557741
0.05	0.642494	0.624164	0.562476
0.06	0.646985	0.628716	0.56711
0.07	0.651365	0.633159	0.571646
0.08	0.655637	0.637497	0.576089
0.09	0.659806	0.641734	0.58044
0.10	0.663876	0.645873	0.584703

Table 3 explores the trend of profit values with respect to repair rate (β) and its value increase corresponding to increments in repair rate (β) and decrease corresponding to increments in failure rate (α). It is concluded that when the value of the repair rate enhances then profit values are also enhanced but when the failure rate of the unit changes α=0.025 to 0.04, 0.06 then the profit of the system declines.

Table 3: *Profit vs. Repair Rate*

β ↓	α=0.025	α=0.04	α=0.06
0.01	2438.338	2386.076	2019.52
0.02	2467.262	2415.876	2049.467
0.03	2495.431	2444.927	2078.759
0.04	2522.874	2473.257	2107.417
0.05	2549.618	2500.892	2135.461
0.06	2575.691	2527.857	2162.912
0.07	2601.117	2554.178	2189.787
0.08	2625.919	2579.875	2216.104
0.09	2650.121	2604.972	2241.881
0.10	2673.744	2629.49	2267.135

IX. Conclusion

The regenerative point graphical technique is used to calculate the performance of the utensil industry. According to the given tables, it is clear that MTSF, availability and profit values increased with increment in repair rate but these reliability measures decreased with increment in failure rate. It is observed that if the system is more available then it gives more profit to the developer. It is evident that industries can examine the behavior of products and system components with the help of the regenerative point graphic technique.

References

[1] Agarwal, A., Garg, D., Kumar, A. and Kumar, R. (2021). Performance analysis of the water treatment reverses osmosis plant. *Reliability Theory and Applications*, 16(3): 16-25.
[2] Chaudhary, P. and Tomar, R. (2019). A two identical unit cold standby system subject to two types of failures. *Reliability Theory and Applications*, 14(1): 34-43.
[3] Gnedenko B. and Igor A.U. (1995). Probabilistic reliability engineering. *John Wiley and Sons.*
[4] Jack, N. and Murthy, D. P. (2007). A flexible extended warranty and related optimal strategies. *Journal of the Operational Research Society*, 58(12): 1612–1620.
[5] Kapur, P. K. and Kapoor, K.R. (1978). Stochastic behaviour of some 2-unit redundant systems. *IEEE Transactions on Reliability*, 27(5): 382-385.
[6] Kumar, A., Garg, D. and Goel, P. (2019). Mathematical modeling and behavioral analysis of a washing unit in paper mill. *International Journal of System Assurance Engineering and Management*, 10: 1639-1645.
[7] Kumar, J. and Goel, M. (2016). Availability and profit analysis of a two-unit cold standby system for general distribution. *Cogent Mathematics*, *3*(1): 1262937.
[8] Levitin, G., Finkelstein, M. and Xiang, Y. (2020). Optimal preventive replacement for cold standby systems with reusable elements. *Reliability Engineering and System Safety*, 204: (107135).
[9] Sengar S. and Mangey R. (2022). Reliability and performance analysis of a complex manufacturing system with inspection facility using copula methodology. *Reliability Theory & Applications,* 17(71): 494-508.
[10] Wang, G. J. and Zhang, Y. L. (2007). An optimal replacement policy for repairable cold standby system with priority in use. *International Journal of Systems Science*, *38*(12): 1021-1027.

RT&A, No 2 (78)
Volume 19, June, 2024

ANALYSIS OF THE MULTIPLE WORKING VACATIONS, BATCH SERVICE AND RENEGING QUEUING SYSTEM UNDER SINGLE SERVER POLICY

Lidiya P*, K Julia Rose Mary

•

[1,2]Department Of Mathematics, Nirmala College For Women, Coimbatore
*lidiyap481@gmail.com
juliakulandaisamy@gmail.com

Abstract

In this paper, we analysed the multiple working vacation queuing model with reneging under a single server policy. Reneging describes the situation where a customer or entity decides to leave the queue before being served. The presence of reneging behaviour affects queue and service efficiency, as customers leaving the queue prematurely can impact overall system performance and customer satisfaction. In this model, customers arrive at a service facility and form a queue to be served by a single server. The arrival follows the Poisson distribution, and the service follows the exponential process. Batches of customers are served under the General Bulk Service Rule. In GBSR, rather than individual customer arriving at a queue one by one, customers arrive in groups or batches. Thus, each batch of service contains a minimum of $'a'$ units and a maximum of $'b'$ units of customers. The steady-state equation, the various performance measures for the system, and particular cases of the described model are derived.

Keywords: Reneging, Multiple Working vacations (MWV), Queue length, Bulk service, System size

1. Introduction

Erlang developed queuing theory while working for the Telephone Company to analyse the behaviour of telephone traffic and optimise the capacity of telephone exchanges. His work laid the foundation for the study of waiting lines and has since been widely applied in various fields to improve system performance and efficiency. The main objective of queuing theory is to understand and optimise the performance of systems that involve waiting lines. By studying factors such as arrival rates, service rates, queue lengths, and waiting times, queuing theory provides insights into how to improve efficiency, and reduce waiting times.

In queuing theory, a vacation queuing model is a type of queuing system where the server may take breaks or go on vacation, leading to periods of time when service is not available. This type of model is often used in scenarios where service providers have scheduled breaks, such as in customer service centres, healthcare facilities, or manufacturing processes. Analysing and optimising vacation queuing models involves considering factors of the duration and frequency of vacations, the impact on service during vacation periods, and strategies to minimise the effects of downtime on customer satisfaction. The concept of "multiple working vacations" refers to a scenario where the server in a queuing system takes several breaks or vacations during their work.

The concept of the GBS rule was indeed introduced by Neuts. The GBS rule, which he introduced, is used to analyse queuing systems where customers arrive in batches and are served

as a single entity with a fixed service time for the entire batch. In a batch arrival process, rather than individual customers arriving at a queue one by one, customers arrive in groups or batches.

Reneging is a term used in queuing theory to describe the situation where a customer or entity decides to leave the queue before being served. This reneging occurs when customers experience long waiting times or delays in service. Customers who renege may seek alternative service providers or simply give up. This can result in lost business opportunities and decreased customer satisfaction. Various strategies can be used to address reneging, such as optimising service processes to reduce waiting times, providing clear communication about the expected waiting times, management techniques that minimise reneging rates.

This paper analyses the queuing system that combines multiple working vacations, batch service, a single server, and reneging behaviour. For this model, we obtained steady state equations, measures of performances, and analysed the particular cases.

2. Review of literature

Research on vacation and reneging queuing models has gained significant attention from researchers in the field of queuing theory. The $MAP/PH/1$ queuing system, including setup, shutdown, multiple vacations, standby server, malfunction, maintenance, and reneging, was examined by [4]. In their paper, the matrix-analytic technique has been used to investigate the total number of consumers existing in the system within a steady-state probability matrix. The non-Markovian approach, which includes longer vacations, renewal processes, and service interruptions followed by repair stages, was studied by [25]. Fixing the M/M/2 machine issue with rushing clients. Under multiple working vacations and strategy (0, Q, N, M) plans, machines are considered to be fixed proposed by [15]. A single server with limitless capacity the Markovian queue structure, is analysed in [12] thesis, has multiple working vacations, an adjustment period, and reneging, by considering both limited capacities and continuous-time queuing systems. An M/M/C queuing model involving changing working vacations was examined by [26]. Additionally, the model's cost function is developed, and the quadratic fit search method is used to look into its optimisation.

Single server's batch of arrival queuing concept for the system, which provides customers who renege during server vacation and system failure times with three stages of heterogeneous services introduced by [7]. The supplemental parameter method has been used to create steady-state probability distribution functions for the queue size. A boundless capacity of one-server Markovian queue mechanism with one working vacation, reneging, and retention of reneged clients evaluated in the thesis of [11]v. A different working vacation queue method involving a second optional service, an unstable server, and the retention of reneged clients was investigated by [21]. They also covered an optimisation problem under a certain cost model. With a practical retention plan for reneged clients and Bernoulli's planned altered vacation regulations, [23] developed a multi-server finite capacity queuing method. The steady-state probability is obtained using the matrix analytical approach. Measures of performance that are produced with an application are also developed and dealt with using the particle swarming optimisation (PSO) meta-heuristic. State-dependent reneging, maintenance of reneged clients, and infinite-capacity single-server Markovian line systems with one working vacation evaluated by [1].

Using Bernoulli responses and client impatience in the context of several vacations, [3] derived a multi-server queuing model. Host vacations, malfunctions, and join or balk tactical behaviours. Additionally, demonstrated how important the reneging option is when the initial system involving non-strategic clients is unreliable researched by [6]. The Markovian queuing structure, which includes working vacation, Bernoulli scheduling disruption, initialization time according to suggestions, reneging of impatient clients, and retention of reneged clients evaluated by [9]. The M/M/1 feedback line with backward balking, backward reneging, and multiple operating vacations was investigated by [14]. The matrix methodology and the ant colony optimisation (ACO) strategy are used to generate the steady-state system length estimates for the model. Reneging, multiple vacations, and set-up period queuing methods were proposed

by [2]. The server provides guidance in three stages; the first two are required, while the third is optional. The server's support duties will be completed if it needs to take a required vacation during that period.

Buffer-modified reverse balking in a single-server finite capacity feedback queuing system, as well as the retention of impatient clients investigated by [24]. The infinite buffering M/M/1 queue with variation working vacations subject to Bernoulli scheduling vacation interruption, in which consumers balk with a probability, was studied by [19]. For various server states, determine a closed-form formulation of the system's capacity and the steady-state probabilities. Evaluate the cost optimisation problem using the quadratic fit-finding technique. M/M/1/N feedback on the operating vacation queuing system, including reneging, was presented by [8]. They implement the Markov process approach to generate the steady-state probability equations. The matrix approach was applied to solve the steady-state probabilities, and a cost model was also developed. Hospitalisation and queuing management processes, including decisions to discharge patients too soon and noncompliance examined by [27]. In order to minimise the hospital's total projected costs, they formulate their solution as an infinite-horizon total discounted cost Markov decision process that balances bed utilisation and patient experiences while they wait for admission. To explain how certain important system factors affect the optimal adaptive policy. An adaptable queuing system, including the retention of reneging consumers, was proposed by [22]. Furthermore, the costs and performance analysis, along with the steady-state and transient-state measures of performance, provide numerical illustrations for demonstrating the model's usefulness in assessing wait times in the service field. The M/M/1 drive-thru lines in order to gain insight into the possibility of reneging occurrences and determine how customers' sensitivity to entry times affects their choice of exiting the line early. By adding reneging behaviour to the queuing theory with fundamental equations, they suggest improving the theory's applicability and improving its reflection and interpretation of queuing issues in the real world discussed by [5]. Three distinct categories of client behavior, such as balking, interruptions, and reneging, were examined by [28]. They researched multi-server preference queues, including client balking, interruptions, and reneging, from an analytical and practical perspective. The idea of the client-reneging effect was introduced by [18]. They analysed reneging versus no-reneging, where customers are balanced and strategic. Client tactics and they take into consideration the fluid on-off concept of the standard queue, including vacations and failures.

Also, how client reneging impacted the formulation of a queue system with two separate shifts. It shows the manner in which lost consumers and costs are analysed in an engaged service policy. Utility operations are also used as decision-making tools analysed by [10]. The F-policy, server keeping up, reneging, and balking queue procedures using only one server and limited capacity were determined by [13]. The recursive approach is used to attain the steady state. Multiple working vacations under breakdown, types of breakdown for heterogeneous arrival queuing model analysed by [16,17]. Multiple working vacations under heterogeneous with encouraged arrival evaluated by [20]. With the help of the appropriate literature, we are able to evaluate the queuing model with reneging under the MWV single server policy.

3. Methodology

In this paper analysed the $M/M(a,b)/1/MWV$ queuing system with reneging. Instead of the server being fully idle during the vacation period, the server serves at a different rate during multiple working vacations. The service rate varies depending on the arrival state. Server provides service during the regular busy period with parameter μ_{rb} and under multiple working vacations, the server provides service with parameter μ_{wv} with exponential distribution. Customer arrive at the system with the parameter λ_v it follows Poisson distribution. In this model, batches of customers are served under the General Bulk Service Rule. Thus, each batch of service contains a minimum of $'a'$ units and a maximum of $'b'$ units of customers. Suppose the number of customers waiting in the queue is less than $'a'$ the server begins a vacation random variable V with parameter ζ. Let

$$\begin{aligned} R_n^I(t) &= Pr\{N_c(t) = n, L(t) = 0\} \qquad 0 \le n \le a-1 \\ Q_n^V(t) &= Pr\{N_c(t) = n, L(t) = 1\} \qquad n \ge 0 \\ P_n^B(t) &= Pr\{N_c(t) = n, L(t) = 2\} \qquad n \ge 0 \end{aligned}$$

$L(t) = 0$, the size of the queue and system are same.
$L(t) = 1$ *or* 2 , the total number of customers in the system is the sum of the number of customers in queue and the size of the service batches that contains particular $a \le x \le b$ customers.
Hence the probabilities of the steady state are ,

$$Q_n^V = \lim_{t\to\infty} Q_n^V(t); \qquad R_n^I = \lim_{t\to\infty} R_n^I(t); \qquad P_n^B = \lim_{t\to\infty} P_n^B(t);$$

exist and the Chapman Kolmogrove equations satisfied by them in the steady state are given by,

$$\lambda_v R_0^I = \mu_{rb} P_0^B + \mu_{wv} Q_0^V \tag{1}$$

$$\lambda_v R_n^I = \lambda_v R_{n-1}^I + (\mu_{rb} + (n-1)\alpha) P_n^B + (\mu_{wv} + (n-1)\alpha) Q_n^V; 1 \le n \le a-1 \tag{2}$$

$$(\lambda_v + \xi + \mu_{wv}) Q_0^V = \lambda_v R_{a-1}^I + \sum_{n=a}^{b} (\mu_{wv} + (n-1)\alpha) Q_n^V \tag{3}$$

$$(\lambda_v + \xi + (\mu_{wv} + (n-1)\alpha)) Q_n^V = \lambda_v Q_{n-1}^V + (\mu_{wv} + (b+n-1)\alpha) Q_{n+b}^V; \qquad n \ge 1 \tag{4}$$

$$(\lambda_v + \mu_{rb}) P_0^B = \sum_{n=a}^{b} (\mu_{rb} + (n-1)\alpha) P_n^B + \xi Q_0^V \tag{5}$$

$$(\lambda_v + (\mu_{rb} + (n-1)\alpha)) P_n^B = \lambda_v P_{n-1}^B + (\mu_{rb} + (b+n-1)\alpha) P_{n+b}^B + \xi Q_n^V; \; n \ge 1 \tag{6}$$

4. Steady state solution

To solve the steady state equation, the forward shifting operator E on P_n^B and Q_n^V are introduced then,

$$E(P_n^B) = P_{n+1}^B; \qquad E(Q_n^V) = Q_{n+1}^V \qquad for\ n \ge 0$$

Thus the (4) gives homogeneous difference equation as,

$$[\lambda_v + (\mu_{wv} + (b+n-1)\alpha) E^{b+1} - (\lambda_v + \xi + (\mu_{wv} + (n-1)\alpha) E] Q_n^V = 0 \tag{7}$$

The characteristics equation of (7) is obtained as,

$$z(u) = \lambda_v + (\mu_{wv} + (b+n-1)\alpha) u^{b+1} - (\lambda_v + \xi + (\mu_{wv} + (n-1)\alpha)) u = 0 \tag{8}$$

by taking $x(u) = (\lambda_v + (\mu_{wv} + (b+n-1)\alpha))$ and $y(u) = (\lambda_v + \xi + (\mu_{wv} + (n-1)\alpha)$, it is found that $|y(u)| < |x(u)|$ on $|u| = 1$. By Rouche's theorem $z(u)$ has unique root r_v inside the contour $|u| = 1$. (7) has a homogeneous solution as,

$$Q_n^V = r_v^n Q_0^V \tag{9}$$

From (6) we get,

$$[\lambda_v + (\mu_{rb} + (b+n-1)\alpha) E^{b+1} - (\lambda_v + (\mu_{rb} + (n-1)\alpha) E] P_n^B = -\xi r_v^{n+1} Q_0^V \tag{10}$$

By applying Rouche's theorem to (10) as,

$$[\lambda_v + (\mu_{rb} + (b+n-1)\alpha) E^{b+1} - (\lambda_v + (\mu_{rb} + (n-1)\alpha) E] P_n^B = 0$$

The above equation has unique root r with $|r| < 1$. Also (10) gives a non-homogeneous solution as,

$$P_n^B = \left[Zr^n - \frac{\xi r_v^{n+1}}{[\lambda_v + (\mu_{rb} + (b+n-1)\alpha)r_v^{b+1} - (\lambda_v + (\mu_{rb} + (n-1)\alpha)r_v]}\right]Q_0^V \tag{11}$$

$$P_n^B = (Zr^n + Z^* r_v^n)Q_0^V \tag{12}$$

Where

$$Z^* = \frac{\xi r_v}{[\lambda_v(r_v - 1) + \mu_{rb} r_v(1 - r_v^b) + \alpha r_v((n-1) - (b+n-1)r_v^b)]} \tag{13}$$

The expression for R_n^I is obtained by adding (1) & (2) and substitute P_n^B and Q_n^V values,

$$R_n^I = \left\{Z\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + Z^*\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r_v^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + \frac{\mu_{wv}}{\lambda_v}\left(\frac{1-r_v^{n+1}}{1-r_v}\right) + \alpha r_v^2\left(\frac{1-r^{n-1}}{1-r_v}\right)\right\}Q_0^V$$

Now to calculate Z, considering (5) and substitute P_n^B and Q_n^V values we find,

$$Z\left[(\lambda_v + \mu_{rb}) - \frac{\mu_{rb}(r^a - r^{b+1})}{(1-r)} - fracr^a\alpha((a-1) - (b-1)r^{b-a+1}1 - r\right] = \xi - Z^* \left((\lambda_v + \mu_{rb}) - \frac{\mu_{rb}(r_v^a - r_v^{b+1})}{(1-r_v)} - \frac{r_v^a\alpha((a-1) - (b-1)r_v^{b-a+1}}{(1-r_v)}\right)$$

the above expression can be simplified as,

$$\frac{Z\mu_{rb}(1-r^a)}{(1-r)} = \frac{\xi}{(1-r_v)} - \frac{Z^*\mu_{rb}(1-r_v^a)}{(1-r_v)} \tag{14}$$

Hence the probability of queue size of the steady-state equation in terms of Q_0^V are obtained,

$$Q_n^V = (r_v^n)Q_0^V \qquad n \geq 0 \tag{15}$$

$$P_n^B = (Zr^n + Z^* r_v^n)Q_0^V \qquad n \geq 0 \tag{16}$$

where

$$Z = \frac{(1-r)}{\mu_{rb}(1-r^a)}\left[\frac{\xi}{(1-r_v)} - \frac{Z^*\mu_{rb}(1-r_v^a)}{(1-r_v)}\right] \tag{17}$$

$$Z^* = \frac{\xi r_v}{[\lambda_v(r_v - 1) + \mu_{rb} r_v(1 - r_v^b) + \alpha r_v((n-1) - (b+n-1)r_v^b)]} \tag{18}$$

and

$$R_n^I = \left\{Z\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + Z^*\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r_v^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + \frac{\mu_{wv}}{\lambda_v}\left(\frac{1-r_v^{n+1}}{1-r_v}\right) + \alpha r_v^2\left(\frac{1-r^{n-1}}{1-r_v}\right)\right\}Q_0^V \tag{19}$$

by using normalizing condition and calculated the value of Q_0^V

$$\sum_{n=0}^{\infty} Q_n^V + \sum_{n=0}^{\infty} P_n^B + \sum_{n=0}^{a-1} R_n^I = 1$$

By substituting P_n^B, Q_n^V and R_n^I we observe that,

$$\sum_{n=0}^{\infty} r_v^n Q_0^V + \sum_{n=0}^{\infty}(Zr^n + Z^* r_v^n)Q_0^V + \sum_{n=0}^{a-1}\left[R_n^I = \left\{Z\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + Z^*\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r_v^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + \frac{\mu_{wv}}{\lambda_v}\left(\frac{1-r_v^{n+1}}{1-r_v}\right) + \alpha r_v^2\left(\frac{1-r^{n-1}}{1-r_v}\right)\right\}Q_0^V = 0$$

Then,

$$(Q_0^V)^{-1} = \omega(r_v, \mu_{wv}) + Z\omega(r, \mu_{rb}) + Z^*\omega(r_v, \mu_{rb}) + \Gamma(r_v) + Z\Gamma(r) + Z^*\Gamma(r_v) \tag{20}$$

where

$$\omega(x,y) = \frac{1}{(1-x)}\left(1 + \frac{y}{\lambda_v}(c - \frac{x(1-x^a)}{(1-x)})\right)$$

$$\Gamma(x) = \left[\frac{\alpha x}{\lambda_v(1-x)}(cx - \frac{(1-x^a)}{(1-x)}\right]$$

5. Performance Measures

In this section, the performance measures of expected queue length, expected waiting time of the queue, expected system length and expected waiting time of the queue for multiple working vacations model with reverse balking under types of breakdowns model are derived.

5.1. Mean queue length

The expected queue length is given by,

$$L_q = \sum_{n=1}^{\infty} n(Q_n^V + P_n^B) + \sum_{n=1}^{a-1} nR_n^I$$

By substituting P_n^B, Q_n^V and R_n^I we observe that,

$$L_q = \sum_{n=1}^{\infty} n(r_v^n Q_0^V) + \sum_{n=1}^{\infty} n(Zr^n + Z^* r_v^n)Q_0^V + \sum_{n=1}^{a-1} n\left\{Z\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + Z^*\left[\frac{\mu_{rb}}{\lambda_v}\left(\frac{(1-r^{n+1})}{(1-r)}\right) + \frac{\alpha r_v^2}{\lambda_v}\left(\frac{1-r_v^{n-1}}{1-r_v}\right)\right] + \frac{\mu_{wv}}{\lambda_v}\left(\frac{1-r_v^{n+1}}{1-r_v}\right) + \alpha r_v^2\left(\frac{1-r^{n-1}}{1-r_v}\right)\right\}Q_0^V$$

$$L_q = Z\omega^*(r, \mu_{rb}) + Z^*\omega^*(r_v, \mu_{rb}) + \omega^*(r_v, \mu_{wv}) + Z\Gamma^*(r) + Z^*\Gamma^*(r_v) + \Gamma^*(r_v) \tag{21}$$

where

$$\omega^*(x,y) = \frac{x}{(1-x)^2} + \frac{y}{\lambda_v(1-x)}\left[\frac{a(a-1)}{2} + \frac{ax^{a+1}(1-x) - x^2(1-x^a)}{(1-x)^2}\right] \tag{22}$$

$$\Gamma^*(x) = \frac{\alpha x^2}{\lambda_v(1-x)}\left\{\frac{a(a-1)}{2} + \frac{ax^{a-1}(1-x) + x^a - 1}{(1-x)^2}\right\} \tag{23}$$

and Z & Z^* are given by (17) & (18).

5.2. Mean System Length

The expected system length is given by,

$$L_s = L_q + \rho$$

$$L_s = \{Z\omega^*(r,\mu_{rb}) + Z^*\omega^*(r_v,\mu_{rb}) + \omega^*(r_v,\mu_{wv}) + Z\Gamma^*(r) + Z^*\Gamma^*(r_v) + \Gamma^*(r_v)\} + \rho$$

where $\omega^*(x,y)$ and $\Gamma^*(x)$ are given by (22) & (23).

5.3. Mean Waiting Time of the Queue

The expected waiting time of the queue is given by,

$$W_q = \frac{L_q}{\lambda}$$

$$W_q = \frac{Z\omega^*(r,\mu_{rb}) + Z^*\omega^*(r_v,\mu_{rb}) + \omega^*(r_v,\mu_{wv}) + Z\Gamma^*(r) + Z^*\Gamma^*(r_v) + \Gamma^*(r_v)}{\lambda_v}$$

where $\omega^*(x,y)$ and $\Gamma^*(x)$ are given by (22) & (23).

5.4. Mean Waiting Time of the System

The expected waiting time of the system is given by,

$$W_s = \frac{L_s}{\lambda}$$

$$W_s = \frac{\{Z\omega^*(r,\mu_{rb}) + Z^*\omega^*(r_v,\mu_{rb}) + \omega^*(r_v,\mu_{wv}) + Z\Gamma^*(r) + Z^*\Gamma^*(r_v) + \Gamma^*(r_v)\} + \rho}{\lambda_v}$$

where $\omega^*(x,y)$ and $\Gamma^*(x)$ are given by (22) & (23).

If $Pr_{(wv)}$, $Pr_{(busy)}$ *and* $Pr_{(idle)}$ denote the probability that the server in idle, regular busy and busy vacation period then

$$Pr_{(idle)} = \sum_{n=0}^{a-1} R_n^I \tag{24}$$

where the R_n^I is given by (19).

$$Pr_{(busy)} = \sum_{n=0}^{\infty} P_n^B = \left(\frac{Z}{(1-r)} + \frac{Z^*}{(1-r_v)}\right) Q_0^V \tag{25}$$

$$Pr_{(wv)} = \sum_{n=0}^{\infty} Q_n^V = \frac{Q_0^V}{(1-r_v)} \tag{26}$$

6. Particular cases

6.1. Classical $M/M(a,b)/1/MWV$ model

By letting $\alpha = 0$ (21) and we obtain,

$$Q_n^V = (r_v^n)Q_0^V \qquad n \geq 0$$

$$P_n^B = (Zr^n + Z^* r_v^n) Q_0^V \qquad n \geq 0$$

$$R_n^I = \left[\frac{\mu_{rb}}{\lambda_v} (Z g_n(r) + Z^* g_n(r_v) + g_n(r_v) \right] Q_0^V \qquad 0 \leq n \leq a-1$$

where

$$Z = \frac{(1-r)}{\mu_{rb}(1-r^a)} \left[\frac{\xi}{(1-r_v)} - \frac{Z^* \mu_{rb}(1-r_v^a)}{(1-r_v)} \right]$$

$$Z^* = \frac{\xi r_v}{\lambda_v(r_v - 1) + \mu_{rb} r_v (1 - r_v^b)}$$

Further

$$L_q = Z\omega^*(r, \mu_{rb}) + Z^* \omega^*(r_v, \mu_{rb}) + \omega^*(r_v, \mu_{wv})$$

where

$$\omega^*(x, y) = \frac{x}{(1-x)^2} + \frac{y}{\lambda_v (1-x)} \left[\frac{a(a-1)}{2} + \frac{a x^{a+1}(1-x) - x^2(1-x^a)}{(1-x)^2} \right]$$

Thus, observed that our specified model coincides with the $M/M(a,b)/1/MWV$ queuing model analysed by J.R.Mary and A.Begum (2011).

Conclusion

In this study, the $M/M(a,b)/1/MWV$ queuing model with reneging is analyzed. In this model, GBSR is followed. The steady-state solution, the various performance measures for the system, and particular cases are calculated.

Reneging factors into queuing systems, making it difficult for service providers to predict and plan for service demand. Reducing reneging in a queuing system is important for improving customer satisfaction, optimising service efficiency, and maximising revenue opportunities. The impact of reneging on queuing systems can affect both customers and service providers. By implementing a system with multiple working vacations effective queue management techniques, optimising service processes, and improving the overall customer experience, it is possible to minimise reneging behaviour and enhance the performance of the queuing system. Further in the future, the model may be extended to the arrival of multiple working vacations queue with the concept of balking.

Conflict of Interest

The authors declare that there is no conflict of interest.

Acknowledgement

The authors would like to extend their sincere gratitude to the editors, associate editors and reviewers for their valuable time, effort and dedication in reviewing and providing insightful recommendations which have improved the quality of the work.

References

[1] Ahmed, F. I., Alemu, S. D., and Tilahu, G. T. (2021). A single server markovian queue with single working vacation, state dependent reneging and retention of reneged customers. *Ethiopian Journal of Education and Sciences*, 17(1):51-69.

[2] Ammakkannu, G., Maragathasundari, S., and Manikandan, P. (2022). Performance measures of a multi vacation queuing system of reneging customers. I*n AIP Conference Proceedings*, volume 2516. AIP Publishing.

[3] Angelika, B. A., Latifa, M., Mohamed, B., and Amit, K. (2022). Mathematical analysis of a markovian multi-server feedback queue with a variant of multiple vacations, balking and reneging. *Discrete and Continuous Models and Applied Computational Science*, 30(1):21-38.

[4] Ayyappan, G. and Thilagavathy, K. (2020). Analysis of map/ph/1 queuing model with setup, closedown, multiple vacations, standby server, breakdown, repair and reneging. *Reliability: Theory & Applications*, 15(2):104-143.

[5] Dbeis, A. and Al-Sahili, K. (2023). Enhancing queuing theory realism: A comprehensive analysis of reneging behavior impact on m/m/1 drive-thru service system. M/1 Drive-Thru Service System.

[6] Economou, A., Logothetis, D., and Manou, A. (2022). The value of reneging for strategic customers in queuing systems with server vacations/failures. *European Journal of Operational Research*, 299(3):960-976.

[7] Enogwe, S. U., Onyeagu, S. I., and Obiora-Ilouno, H. O. (2021). Single channel batch arrival queuing model for systems that provides three-stage service for customers that renege during server vacation and breakdown periods. *Journal of Xidian University*, 15 .

[8] Gupta, N. et al. (2023). Analysis of an m/m/1/k feedback working vacation queue with reneging. *Reliability: Theory & Applications*, 18(4 (76)):178-188.

[9] Gupta, R. (2022). Cost optimization of queuing system with working vacation, setup, feedback, reneging, and retention of reneged customers. *Journal of Scientific Research*, 14(1).

[10] Hernandez, M. L. C., Rosas, L. V., and Romero, V. V. (2024). A decision-analytic, simulation based management model for a two-shifts queue system with customer reneging. *JEMIS (Journal of Engineering & Management in Industrial System)*, 11(2).

[11] Idris Ahmed, F. (2021). A single server markovian queue with single working vacation and state dependent retention of reneged customers. PhD thesis, Haramaya University.

[12] Jemal Adem, A. et al. (2020). A single server markovian queue with multiple working vacations, changeover time and reneging. PhD thesis, Haramaya university.

[13] Jeyachandhiran, R. and Rajendran, P. (2024). Analysis of a single server retrial queuing system with finite capacity, f-policy, balking, reneging and server maintenance. *Contemporary Mathematics*, pages 506-518.

[14] Jyothsna, K., Laxmi, P. V., and Kumar, P. V. (2022). Optimization of a feedback working vacation queue with reverse balking and reverse reneging. *Reliability: Theory & Applications*, 17(1 (67)):154-163.

[15] Ketema, T. (2020). Performance analysis of machine repair system with balking, reneging, multiple working vacations and two removable servers operating under the triadic (0, q, n, m) policy. *Journal of Innovative Systems Design and Engineering*, 11(4).

[16] Lidiya, P. and Mary, K. J. R. (2023). Performance study of the M/M(a,b)/1/MWv queuing system with types of breakdowns. *Journal of Applied Mathematics, Statistics and Informatics*, 19(2):23-38.

[17] Lidiya, P. and Mary, K. J. R. (2023). Performance study on heterogeneous arrival of batch service for multiple working vacations queuing system with breakdowns in the busy period.*In 2023 First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI)*, 1-6.

[18] Logothetis, D., Manou, A., and Economou, A. (2023). The impact of reneging on a fluid on-off queue with strategic customers. *Annals of Operations Research*, 331(2):629-647.

[19] Majid, S., Vijaya, L. P., Angelika, B. A., and Aijaz, M. (2023). Performance analysis of vacation interruption queue with balking and reneging.

[20] Prakati, P. and Mary, K. J. R. (2023). Evaluation of batch service mwv heterogeneous queuing system under balking and encouraged arrival. *In 2023 First International Conference on Advances in Electrical, Electronics and Computational Intelligence (ICAEECI)*, 1-7. IEEE.

[21] Pikkala, V. L. and Edadasari, G. B. (2021). Variant working vacation markovian queue with second optional service, unreliable server and retention of reneged customers. *International Journal of Mathematics in Operational Research*, 19(1):45-64.

[22] Sharma, S., Kumar, R., Kuaban, G. S., and Singh, B. (2023). Performance and cost evaluation of an adaptive queuing system with customer reneging and retention: Steady-state and transient analysis.

[23] Shekhar, C., Varshney, S., and Kumar, A. (2021). Matrix-geometric solution of multi-server queuing systems with bernoulli scheduled modified vacation and retention of reneged customers: A meta-heuristic approach. *Quality Technology & Quantitative Management*, 18(1):39-66.

[24] Som, B. K. and Seth, S. (2023). A finite buffer reverse balking feedback markovian queuing system with reneging and retention of impatient customers. *Pakistan Journal of Statistics*, 39(1).

[25] Sundari, S. M., Murugeswari, N., and Manikandan, P. (2020). A study on the performance measures of the non-markovian model of optional types of service withextended vacation, reneging process and service interruption followed by phases of repair. *International Journal of Process Management and Benchmarking*, 10(4):520-549.

[26] Vijaya Laxmi, P. and Kassahun, T. (2020). Analysis of variant working vacation queue with reneging under a multi-server environment. *International Journal of Management Science and Engineering Management*, 15(2):130-137.

[27] Yang, F., Li, Q.-L., Zhang, C., and Wang, C. (2023). Optimal admission and queuing control with reneging behavior under premature discharge decisions. *International Transactions in Operational Research*.

[28] Yanting, C. and Na, Y. (2023). A multi-server priority queuing system with customer balking, interjections and reneging. *Operations Research and Management Science*, 32(5):49.

CRITICAL ANALYSIS OF FAILURE AND REPAIR RATES OF POLY-TUBE MANUFACTURING PLANT USING PSO

Shakuntla Singla [1],Diksha Mangla [2*], Umar Muhammad Modibbo [3], A.K. Lal [4]

•

[1,2] Department of Mathematics, MMEC, Maharishi Markandeshwar (Deemed to be) University, Mullana, India

[3]Department of Statistics and Operations Research ,School of Physical Sciences ,Modibbo Adama University of Technology ,YOLA-NIGERIA

[4]School of Mathematics, Thapar Institute of Engineering and Technology, Patiala, India

[1] shaku25@gmail.com,[2*]dikshamangla1995@gmail.com, [3]umarmodibbo@mau.edu.ng,[4]aklal@thapar.edu

Abstract

A proper maintenance strategy is essential for the optimal performance of poly tube manufacturing to ensure high reliability. It involves a complex structure consisting of many components interconnected in series or parallel configurations. This project's contribution is the development of a method for evaluating the performance of an industrial system using previously unknown data. The RAM index, influenced by failure and repair rates, has been devised to identify the system's most critical component that impacts reliability, availability, and maintainability, collectively known as RAM. For performance analysis, a Markov-based simulation system model has been formulated and resolved to refine the results through particle swarm optimization (PSO). The transition diagram facilitates the construction of ordinary differential equations (ODEs), which represent various operational states such as full capacity, reduced capacity, and failure. These ODEs are then solved using initial and boundary condition.

Keywords: Availability analysis, Supplementary variable Technique, Particle swarm optimization (PSO),Critical Analysis.

I. Introduction

In contemporary settings, the advancement of technology has diminished the need for physical labor. The creation of various machines and equipment results in less manual work and more precise outcomes. Therefore, for optimal performance, the efficiency of this equipment is crucial, which relies entirely on its operation and maintenance, as well as that of its components. Reliability can be defined as the likelihood of success at a given time t, meaning the probability that a machine designed to fulfill its function within a set timeframe under certain external conditions will do so successfully. A system or device is considered highly reliable or dependable when it executes the intended task flawlessly without encountering any issues. However, consistent usage of a system inevitably leads to wear and tear of parts, meaning no system can maintain maximum efficiency indefinitely. Consequently, a system's reliability and efficiency are compromised when specific components deteriorate and fail. The motive of the study can help the manufacturing plant to get maximum production by ensuring that the system is as fault-free as possible through effective management, control and maintenance. This paper deals with the availability of the system having different numbers for the rates of failing and repairing. For the first time, an evolutionary optimization approach, namely PSO, is used to anticipate accessibility in the process sector. The existing methods including Markov and Genetic Algorithm confirmed or validated the outcomes generated for

optimum allocation/ availability. This is customary that these outcomes are valuable for the management of the industry for adopting a suitable maintenance program and strategies.

Cox [1] employed a supplemental variable approach to determine the system's dependability and availability. After that, many researchers reported the system's reliability via the supplementary variable technique. Ying-Shen et al. [2] developed a series-parallel system to find the availability of GA. Sagayaraj et al. [4] documented the system's reliability via a mixed series-parallel combination. The highest degree of dependability and availability is important not just to minimize total production costs but also to limit the danger of risks (Yang et al., [5]). Chaudhary et al. [6] evaluated the "reliability, availability, and maintainability" (RAM) in a cement factory with various failure and repair rates. The study's goal is to identify essential subsystems of a cement plant so that effective solutions to enhance their RAM characteristics can be offered, resulting in an improvement in cement plant capacity utilization. TPP availability, dependability, and planning have transformed into important needs in recent years as society's demand for energy has increased (Kuo and Ke, [7]). Sunita et al. [8] discussed the sensitive analysis of the thresher plant under study. Sunita et al. [9] discussed RAP via a constraint optimization genetic algorithm. Sunita et al. [10] find the solution to constrained problems using particle swarm optimization. RAM (reliability, availability, and maintainability) of threshing machines in agriculture was observed by Anchal et al. [11]. Vanita et al. [12] highlighted the effect on profit and availability of a briquette machine under the minor and major faults handled by two repairmen. An effort is also being made to assess the plant's dependability and reliability by using continual failure and repair rates. A deep learning process was examined by Singla et al. [13] in order to optimize the reliability parameters and boost industry revenues and manufacturing of a 2:3 good system. Singla et al. [14] investigate a failing system by applying a genetic algorithm to ascertain the reliability metrics influenced by the rate of degradation and the rate of preventive maintenance.

So far, the polytube sector has received little attention, despite the fact that it plays a significant part in our everyday lives. In any sector, this normality assumption for the rates of failing and repairing are not feasible. Keeping this in mind, we analyzed a four-unit Polytube sector subjected to fluctuating subsystems' failing and repairing rates in the current study, and we used supplementary variable technique to explore the reliable model of the Polytube sector. An effort is also being made to assess the plant's dependability/reliability by using continual failing and repairing rates.

The goal of the presented work is to maximize system's availability with respect to each unit while maintaining constant operation over time and at varying rates. The goal of this work is to concentrate on how sensitivity of units of the system is reliable to the availability. The paper is structured as follows: Section 2 provides an explanation of the model's specifics, including the state overview, assumptions, notations, and model frame. The mathematical representations is covered in Section 3. The methodology used to know the effect different rates is presented in section 4. The system's critical analysis is discussed in Section 5. Section 6 has concluded with the results discussion. Presented in Section 7 is the conclusion.

II. Model descriptions and Symbols

2.1 System description

Due to the fact that iron pipe corrodes rapidly in damp and humid environments, shortening its lifespan and making that fragile and more leakage susceptible. Polytube industries, making plastic pipes, is crucial in our daily lives since these are used to transport portable water, fluids (liquids other than water and gases) from one location to another. This study aimed at the "Polytube industry", which is made up of four subsystems: "Mixture", "Extruder", "Die" and "Cutter" as shown in figure

1. The following is a detailed system's description along with the notations, as necessary for the formulation of mathematical structure:

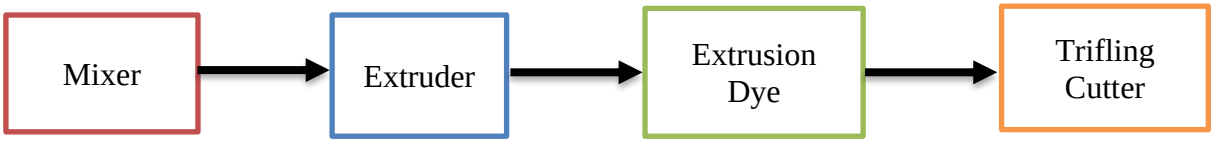


Figure 1 : *Structure of Polytube Manufacturing Plant*

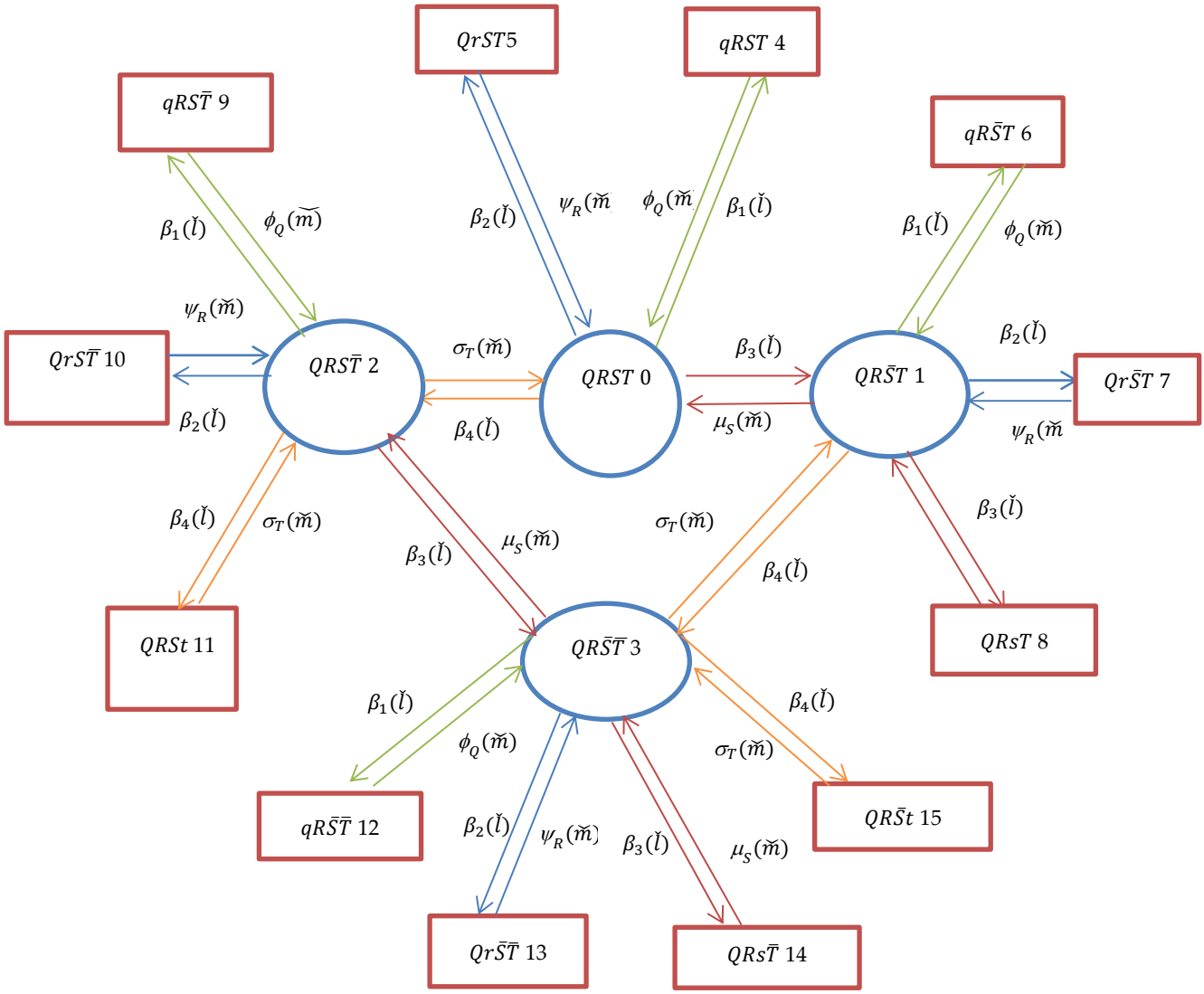


Figure 2 : *Transitions States of Polytube Manufacturing Plant.*

- Sub-system Q (Mixture) Its failure results in the system's total failure.
- Sub-system R (Extruder) Its failure results in the system's total failure.
- Sub-system S (Dye) This is utilized for making various diameters of pipe

- Sub-system T (Cutter) This subsystem consists of 2 series - connected elements. The first item is the blade, that shreds a pipe, while the 2nd second element is the engine (motor), that slice a pipe in various sizes.

2.2 Symbols

Table 1: *Various notations regarding model*

Q, R, S, T	Signify that the subsystem is fully operational.
$\bar{S}, \bar{T}$	Designate the minimized configuration of the subsystems "S" and "T"
q, r, s, t	Designate the subsystem's failure condition.
$\beta_i(\check{l})(\mathrm{i} = 1, \ldots, 4)$	Represents subsystem's Q, R, S, and T rate of failing respectively.
$\phi_Q(\breve{m}), \psi_R(\breve{m}), \mu_S(\breve{m}), and\ \sigma_T(\breve{m})$	Repair rates of Q, R, S and T, respectively.
$P_o(t)$	Specify fully operational system without any failure
$P_i(\check{l}, \breve{m}, t)(i = 1, \ldots, 16)$	Represents the probability of the industry in state I, at time t, with an expired failure time $\check{l}$ and an elapsed time of repair $\check{l}$.

2.3 Assumptions

The current study is based on the following assumptions

- Repairing ((R_r) and failing rates(F_r) are not dependent on each other ,independent to each other .
- A repairing subsystem is as good as original.
- Repair services are available.

III. Mathematical presentation of the system

3.1. Rates of Failing ($\boldsymbol{F_r}$)and repairing ($\boldsymbol{R_r}$) are taken as variable

If there are variable rates of failing (F_r) and repairing (R_r) for the transitory state, the differential difference equation using Chapman Kolmogrov's rule connected with the state transition diagram (fig. 1) are as follow:

$$P_0(t+\Delta t) = [1 - \beta_1(\check{l})\Delta t - \beta_2(\check{l})\Delta t - \beta_3(\check{l})\Delta t - \beta_4(\check{l})\Delta t]P_0(t) + \int \mu_S(\breve{m})P_1(\check{l}, \breve{m}, t)d\check{l}\,\Delta t$$
$$+ \int \sigma_T(\breve{m})P_2(\check{l}, \breve{m}, t)d\check{l}\Delta t + \int \phi_Q(\breve{m})P_4(\check{l}, \breve{m}, t)d\check{l}\,\Delta t + \int \psi_R(\breve{m})P_5(\check{l}, \breve{m}, t)d\check{l}\,\Delta t$$

$$P_0(t+\Delta t) - P_0(t)$$
$$= -[\beta_1(\check{l})\Delta t + \beta_2(\check{l})\Delta t + \beta_3(\check{l})\Delta t + \beta_4(\check{l})\Delta t]P_0(t) + \int \mu_S(\breve{m})P_1(\check{l}, \breve{m}, t)dl\,\Delta t$$
$$+ \int \sigma_T(\breve{m})P_2(\check{l}, \breve{m}, t)d\check{l}\Delta t + \int \phi_Q(\breve{m})P_4(\check{l}, \breve{m}, t)d\check{l}\,\Delta t + \int \psi_R(\breve{m})P_5(\check{l}, \breve{m}, t)d\check{l}\,\Delta t$$

Dividing both sides byΔt,we get

$$\frac{P_0(t+\Delta t)-P_0(t)}{\Delta t}$$
$$= -[\beta_1(\check{l})+\beta_2(\check{l})+\beta_3(\check{l})+\beta_4(\check{l})]P_0(t)+\int \mu_S(\check{m})P_1(\check{l},\check{m},t)d\check{l}$$
$$+\int \sigma_T(\check{m})P_2(\check{l},\check{m},t)d\check{l}+\int \phi_Q(\check{m})P_4(\check{l},\check{m},t)d\check{l}+\int \psi_R(\check{m})P_5(\check{l},\check{m},t)d\check{l}$$
$$\left[\frac{d}{dt}+Z_0\right]P_0(t)=E_0 \quad (1)$$

Likewise, we may construct the differential equation for the other states as:

$$\left[\frac{\partial}{\partial t}+\frac{\partial}{\partial \check{l}}+\frac{\partial}{\partial \check{m}}+Z_i(\check{l},\check{m})\right]P_i(\check{l},\check{m},t)=E_i(\check{l},\check{m},t) \quad \text{for i=1,2 and 3} \quad (2)$$
$$\left[\frac{\partial}{\partial t}+\frac{\partial}{\partial \check{l}}+\frac{\partial}{\partial \check{m}}+\phi_Q(\check{m})\right]P_j(\check{l},\check{m},t)=\beta_1(\check{l})P_k(t) \quad \text{for j=4,6,9 and 12 and k=0 to 3 respectively} \quad (3)$$
$$\left[\frac{\partial}{\partial t}+\frac{\partial}{\partial \check{l}}+\frac{\partial}{\partial \check{m}}+\psi_R(\check{m})\right]P_p(\check{l},\check{m},t)=\beta_2(\check{l})P_k(t) \quad \text{for p=5,7,10 and 13 and k=0 to 3 respectively} \quad (4)$$
$$\left[\frac{\partial}{\partial t}+\frac{\partial}{\partial \check{l}}+\frac{\partial}{\partial \check{m}}+\mu_S(\check{m})\right]P_s(\check{l},\check{m},t)=\beta_3(\check{l})P_r(\check{l},\check{m},t) \quad \text{for s=8 and 14 and r=1 and 3 respectively} \quad (5)$$
$$\left[\frac{\partial}{\partial t}+\frac{\partial}{\partial \check{l}}+\frac{\partial}{\partial \check{m}}+\sigma_T(\check{m})\right]P_u(\check{l},\check{m},t)=\beta_4(\check{l})P_v(\check{l},\check{m},t) \quad \text{for u=11 and 15 and v= 2 and 3 respectively} \quad (6)$$

Where,

$$Z_0=\beta_1(\check{l})+\beta_2(\check{l})+\beta_3(\check{l})+\beta_4(\check{l})$$
$$E_0=\int \mu_S(\check{m})P_1(\check{l},\check{m},t)dl+\int \sigma_T(\check{m})P_2(\check{l},\check{m},t)dl+\int \phi_Q(\check{m})P_4(\check{l},\check{m},t)dl+\int \psi_R(\check{m})P_5(\check{l},\check{m},t)dl$$
$$Z_1(\check{l},\check{m})=\mu_S(\check{m})+\beta_1(\check{l})+\beta_2(\check{l})+\beta_3(\check{l})+\beta_4(\check{l})$$
$$E_1(\check{l},\check{m},t)=\beta_3(\check{l})P_0(t)+\phi_Q(\check{m})P_6(\check{l},\check{m},t)+\psi_R(\check{m})P_7(\check{l},\check{m},t)+\mu_S(\check{m})P_8(\check{l},\check{m},t)+\sigma_T(\check{m})$$
$$Z_2(\check{l},\check{m})=\sigma_T(\check{m})+\beta_1(\check{l})+\beta_2(\check{l})+\beta_3(\check{l})+\beta_4(\check{l})$$
$$E_2(\check{l},\check{m},t)=\beta_4(\check{l})P_0(t)+\phi_Q(\check{m})P_9(\check{l},\check{m},t)+\psi_R(\check{m})P_{10}(\check{l},\check{m},t)+\mu_S(\check{m})P_3(\check{l},\check{m},t)+\sigma_T(m)P_{11}(\check{l},\check{m},t)$$
$$Z_3(l,\check{m})=\sigma_T(\check{m})+\mu_S(\check{m})+\beta_1(l)+\beta_2(l)+\beta_3(l)+\beta_4(l)$$
$$E_3(l,\check{m},t)=\beta_4(\check{l})P_1(\check{l},\check{m},t)+\beta_3(l)P_2(\check{l},\check{m},t)+\phi_Q(\check{m})P_{12}(\check{l},\check{m},t)+\psi_R(\check{m})P_{13}(\check{l},\check{m},t)$$
$$+\mu_S(\check{m})P_{14}(\check{l},\check{m},t)+\sigma_T(\check{m})P_{15}(\check{l},\check{m},t)$$

Boundary Conditions:

$$P_a(0,\check{m},t)=\beta_b(\check{l})P_0(t) \quad \text{for a=1,2,4 and 5 and b=3,4,1and 2 respectively} \quad (7)$$
$$P_3(0,\check{m},t)=\int \beta_4(\check{l})P_1(\check{l},\check{m},t)\,d\check{l}+\int \beta_3(\check{l})P_2(\check{l},\check{m},t)d\check{l} \quad (8)$$
$$P_c(0,\check{m},t)=\int \beta_d(\check{l})P_1(\check{l},\check{m},t)\,d\check{l} \quad \text{for c=6 to 8 and d= 1 to 3 respectively} \quad (9)$$
$$P_o(0,\check{m},t)=\int \beta_p(\check{l})P_2(\check{l},\check{m},t)\,d\check{l} \quad \text{for o=9 to 11 and p=1,2 and 4 respectively} \quad (10)$$
$$P_{a_1}(0,\check{m},t)=\int \beta_{a_2}(\check{l})P_3(\check{l},\check{m},t)\,d\check{l} \quad \text{for } a_1\text{=12 to 15 and } a_2\text{=1 to 4 respectively} \quad (11)$$

Initial Condition:

$$P_i(\check{l},\check{m},0)=0;\quad (i=1\ldots 15) \quad (12)$$
$$P_0(0)=1 \quad (13)$$

For finding the systems' reliability R_{re} (t)), the solutions of differential equations (DE) (1-6) have been obtained. Shakuntla et.al (2011) used Lagrange's method to solve Chapman-Kolmogorov differential equation (DE) with constant rates of failing and repairing. To obtain the probability $P_i(t)(i=1\ldots 15)$, every state equation (2-6) and the initial conditions (7-11) were solved:

$$P_{15}(\check{l},\check{m},t)=e^{-\int \sigma_T(\check{m})d\check{m}}\begin{bmatrix}\int \beta_4(\check{m}-\check{l})P_3(\check{l},\check{m}-\check{l},t-\check{l})\,d\check{l}\\+\int \beta_4(\check{l})P_3(\check{l},\check{m},t)e^{\int \sigma_T(\check{m})d\check{m}}d\check{l}\end{bmatrix} \quad (14)$$
$$P_{14}(\check{l},\check{m},t)=e^{-\int \mu_S(\check{m})d\check{m}}\begin{bmatrix}\int \beta_3(\check{m}-\check{l})P_3(\check{l},\check{m}-\check{l},t-\check{l})\,d\check{l}\\+\int \beta_3(\check{l})P_3(\check{l},\check{m},t)e^{\int \mu_S(\check{m})d\check{m}}d\check{l}\end{bmatrix} \quad (15)$$
$$P_{13}(\check{l},\check{m},t)=e^{-\int \psi_R(\check{m})d\check{m}}\begin{bmatrix}\int \beta_2(\check{m}-\check{l})P_3(\check{l},\check{m}-\check{l},t-\check{l})\,d\check{l}\\+\int \beta_2(\check{l})P_3(\check{l},\check{m},t)e^{\int \psi_R(\check{m})d\check{m}}d\check{l}\end{bmatrix} \quad (16)$$
$$P_{12}(\check{l},\check{m},t)=e^{-\int \phi_Q(\check{m})d\check{m}}\begin{bmatrix}\int \beta_1(\check{m}-\check{l})P_3(\check{l},\check{m}-\check{l},t-\check{l})\,d\check{l}\\+\int \beta_1(\check{l})P_3(\check{l},\check{m},t)e^{\int \phi_Q(\check{m})d\check{m}}d\check{l}\end{bmatrix} \quad (17)$$

$$P_{11}(\check{l},\breve{m},t)=e^{-\int\sigma_T(\breve{m})d\breve{m}}\begin{bmatrix}\int\beta_4(\breve{m}-\check{l})P_2(\check{l},\breve{m}-\check{l},t-\check{l})\,d\check{l}\\+\int\beta_4(\check{l})P_2(\check{l},\breve{m},t)e^{\int\sigma_T(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (18)$$

$$P_{10}(\check{l},\breve{m},t)=e^{-\int\psi_R(\breve{m})d\breve{m}}\begin{bmatrix}\int\beta_2(\breve{m}-\check{l})P_2(\check{l},\breve{m}-\check{l},t-\check{l})\,d\check{l}\\+\int\beta_2(\check{l})P_2(\check{l},\breve{m},t)e^{\int\psi_R(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (19)$$

$$P_9(\check{l},\breve{m},t)=e^{-\int\phi_Q(\breve{m})d\breve{m}}\begin{bmatrix}\int\beta_1(m-l)P_2(\check{l},\breve{m}-\check{l},t-\check{l})\,d\check{l}\\+\int\beta_1(\check{l})P_2(l,m,t)e^{\int\phi_Q(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (20)$$

$$P_8(\check{l},\breve{m},t)=e^{-\int\mu_S\mu(m)dm}\begin{bmatrix}\int\beta_3(m-l)P_1\check{l},\breve{m}-\check{l},t-\check{l}\,d\check{l}\\+\int\beta_3(\check{l})P_1(l,m,t)e^{\int\mu_S(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (21)$$

$$P_7(\check{l},\breve{m},t)=e^{-\int\psi_R(m)dm}\begin{bmatrix}\int\beta_2(\breve{m}-\check{l})P_1(\check{l},\breve{m}-\check{l},t-\check{l})\,d\check{l}\\+\int\beta_2(\check{l})P_1(\check{l},\breve{m},t)e^{\int\psi_R(m)dm}d\check{l}\end{bmatrix} \quad (22)$$

$$P_6(\check{l},\breve{m},t)=e^{-\int\phi_Q(m)dm}\begin{bmatrix}\int\beta_1(\breve{m}-\check{l})P_1(\check{l},\breve{m}-\check{l},t-\check{l})\,d\check{l}\\+\int\beta_1(\breve{m})P_1(\check{l},\breve{m},t)e^{\int\phi_Q(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (23)$$

$$P_5(\check{l},\breve{m},t)=e^{-\int\psi_R(m)dm}\begin{bmatrix}\beta_2(\breve{m}-\check{l})P_0(t-\check{l})\\+\int\beta_2(\check{l})P_0(t)e^{\int\psi_R(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (24)$$

$$P_4(\check{l},\breve{m},t)=e^{-\int\phi_Q(m)dm}\begin{bmatrix}\beta_1(\breve{m}-\check{l})P_0(t-\check{l})\\+\int\beta_1(\check{l})P_0(t)e^{\int\phi_Q(\breve{m})d\breve{m}}d\check{l}\end{bmatrix} \quad (25)$$

$$P_3(\check{l},\breve{m},t)=e^{-\int Z_3(\check{l},\breve{m})d\check{l}}\begin{bmatrix}\int E_3(\check{l},\breve{m},t)e^{\int Z_3(\check{l},\breve{m})dl}d\check{l}\\+\int\beta_4(\breve{m}-\check{l})P_1(\check{l},\breve{m}-\check{l},t-\check{l})\,d\check{l}\\+\int\beta_3(b-a)P_2(\check{l},\breve{m}-\check{l},t-\check{l})\end{bmatrix} \quad (26)$$

$$P_2(\check{l},\breve{m},t)=e^{-\int Z_2(\check{l},\breve{m})d\check{l}}\begin{bmatrix}\int E_2(\check{l},\breve{m},t)e^{\int Z_2(\check{l},\breve{m})dm}d\check{l}\\+\beta_4(\breve{m}-\check{l})P_0(t-\check{l})\end{bmatrix} \quad (27)$$

$$P_1(\check{l},\breve{m},t)=e^{-\int Z_1(l,\breve{m})d\check{l}}\begin{bmatrix}\int E_1(\check{l},\breve{m},t)e^{\int Z_1(\check{l},\breve{m})dl}dl\\+\beta_3(\breve{m}-\check{l})P_0(t-\check{l})\end{bmatrix} \quad (28)$$

$$P_0(t)=e^{-Z_0t}[1+\int E_0(t)e^{Z_0t}dt] \quad (29)$$

If the manufacturing plant supply the rates of failing and repairing, we may find the reliability $R_{re}(t)$ in concern to probability$P_0(t)$ & via equation (1). Hence, Reliability $R_{re}(t)$ of manufacturing plant is given by

$$R_{re}(t)=P_0(t)+\int\sum_{i=1}^{3}P_i(\check{l},\breve{m},t)\,d\check{l}d\breve{m} \quad (30)$$

3.2. Failure $\mathbf{F_r}$ and $\mathbf{R_r}$ Repair rates are constant

When both the rates of failing and repairing are consistent, the system of equations (1-6) collapses to simple differential equations (DE) form, as shown below:

$$\left[\frac{d}{dt}+C\right]P_0(t)=\mu_SP_1(t)+\sigma_TP_2(t)+\phi_QP_4(t)+\psi_RP_5(t) \quad (31)$$

$$\left[\frac{d}{dt}+\mu_S+C\right]P_1(t)=\beta_3P_0(t)+\phi_QP_6(t)+\psi_RP_7(t)+\mu_SP_8(t)+\sigma_TP_3(t) \quad (32)$$

$$\left[\frac{d}{dt}+\sigma_T+C\right]P_2(t)=\beta_4P_0(t)+\phi_QP_9(t)+\psi_RP_{10}(t)+\mu_SP_3(t)\ +\sigma_TP_{11} \quad (33)$$

$$\left[\frac{d}{dt}+\sigma_T+\mu_S+C\right]P_3(t)=\beta_4P_1(t)+\beta_3P_2(t)+\phi_QP_{12}(t)+\psi_RP_{13}+\mu_SP_{14}(t)\ +\sigma_TP_{15}(t) \quad (34)$$

$$\left[\frac{d}{dt}+\phi_Q\right]P_j(t)=\beta_1P_k(t) \quad \text{for j=4,6,9 and 12 and k=0 to 3 respectively} \quad (35)$$

$$\left[\frac{d}{dt}+\psi_R\right]P_p(t)=\beta_2P_k(t) \quad \text{for p=5,7,10 and 13 and k=0 to 3 respectively} \quad (36)$$

$$\left[\frac{d}{dt}+\mu_S\right]P_s(t)=\beta_3P_r(t) \quad \text{for s=8 and 14 and r=1 and 3 respectively} \quad (37)$$

$$\left[\frac{d}{dt}+\sigma_T\right]P_u(t)=\beta_4P_v(t) \quad \text{for u=11 and 15 and v= 2 and 3 respectively} \quad (38)$$

Where C=$\beta_1+\beta_2+\beta_3+\beta_4$

Initial conditions: The initial conditions of the subsystems are as under:

$$P_i(0) = \begin{cases} 1, & i = 0 \\ 0, & \text{otherwse} \end{cases} \quad (39)$$

One can obtain the state probabilities $P_i(i = 1, \dots, 15)$ by solving differential equations (14-29) along with initial boundations (30).

3.3. Steady State

Manufacturing firms are continuously looking for long-term availability in order to meet their goals. This may be calculated numerically concerning $\frac{d}{dt} \to 0$ as$t \to \infty$ into the system of equations (31-38) therefore; the system of equations (31-38) reduces to the following system of linear equations:

$[C]P_0 - \mu_S P_1 - \sigma_T P_2 - \phi_Q P_4 - \psi_R P_5 = 0$ (40)

$[\mu_S + C]P_1 - \beta_3 P_0 - \phi_Q P_6 - \psi_R P_7 - \mu_S P_8 - \sigma_T P_3 = 0$ (41)

$[\sigma_T + C]P_2 - \beta_4 P_0 - \phi_Q P_9 - \psi_R P_{10} - \mu_S P_3 - \sigma_T P_{11} = 0$ (42)

$[\sigma_T + \mu_S + C]P_3 - \beta_4 P_1 - \beta_3 P_2 - \phi_Q P_{12} - \psi_R P_{13} - \mu_S P_{14} - \sigma_T P_{15} = 0$ (43)

$\phi_Q P_j - \beta_1 P_k = 0$ for j=4,6,9 and 12 and k=0 to 3 respectively (44)

$\psi_R P_p - \beta_2 P_k = 0$ for p=5,7,10 and 13 and k=0 to 3 respectively (45)

$\mu_S P_s - \beta_3 P_r = 0$ for s=8 and 14 and r=1 and 3 respectively (46)

$\sigma_T P_u - \beta_4 P_v = 0$ for u=11 and 15 and v=2 and 3 respectively (47)

The availability $A_{av}(t)$ of the system can be computed as,

$$A_{av}(t) = \sum_{i=0}^{3} P_i(t) \quad (48)$$

The system's availability $A_{av}(t)$, specified in equation (48) is evaluated at different values for various rates of failing and repairing. This might noticed here that we have only considered the most important subsystems(Q, R, S, T)

At the final point, the steady state availability has been solved by using the system of linear equations (31-38) recursively be expressing all the probabilities in terms of P_0. These are obtaining as below:

$P_i = N_i P_0$ for i= 1 to 3 (49)

$P_j = \frac{\beta_1}{\phi_Q} P_k$ for j=4,6,9 and 12 and k=0 to 3 respectively (50)

$P_p = \frac{\beta_2}{\psi_R} P_k$ for p=5,7,10 and 13 and k=0 to 3 respectively (51)

$P_s = \frac{\beta_3}{\mu_S} P_r$ for s=8 and 14 and r=1 and 3 respectively (52)

$P_u = \frac{\beta_4}{\mu_S} P_v$ for u=11 and 15 and v=2 and 3 respectively (53)

where

$N_1 = \frac{\beta_3}{U_3} + \frac{\beta_3 \sigma_T \beta_4}{U_1 U_2 U_3}, \quad N_2 = \frac{\beta_4}{U_2} + \frac{\mu_S \beta_4 N_1}{U_1 U_2}, N_3 = \frac{\beta_4}{U_1} N_1 + \frac{\beta_3}{U_1} N_2$

$U_1 = \sigma_T + \mu_S, U_2 = \sigma_T + \beta_3 - \frac{\mu_S \beta_3}{U_1}, \quad U_3 = \mu_S + \beta_4 - \frac{\sigma_T \beta_4}{U_1} - \frac{\mu_S \beta_3 \sigma_T \beta_4}{U_1 U_2 U_1}$

Now using the normalizing conditions$\sum_{i=0}^{15} P_i = 1$, we get

$$P_0 = \left[1 + \frac{\beta_1}{\phi_Q} + \frac{\beta_2}{\psi_R} + \left(1 + \frac{\beta_1}{\phi_Q} + \frac{\beta_2}{\psi_R} + \frac{\beta_3}{\mu_S}\right) N_1 + \left(1 + \frac{\beta_1}{\phi_Q} + \frac{\beta_2}{\psi_R} + \frac{\beta_4}{\sigma_T}\right) N_2 + (1 + \frac{\beta_1}{\phi_Q} + \frac{\beta_2}{\psi_R} + \frac{\beta_3}{\mu_S} + \frac{\beta_4}{\sigma_T}) N_3\right]^{-1} \quad (54)$$

As a result, the manufacturing plant's steady-state availability is attained as:

$$A_{av}(\infty) = \sum_{i=0}^{3} P_i = [1 + N_1 + N_2 + N_3] P_0 \quad (55)$$

IV. Methodology used for availability analysis of poly tube plant using PSO

4.1. Introduction

Inspired by swarming behaviors found in nature, such as flocks of fish and birds, Particle Swarm Optimization (PSO) is a potent meta-heuristic optimization algorithm, also called a stochastic search algorithm, based on population dynamics. It is a computational method used to optimize the

problem. It performs its task of optimization by improving particle solutions. This algorithm works with some parameters, like particle size, population, position, velocity, search space, etc. In PSO, the population, like the bird group, represents a swarm, and each member of the swarm represents a particle. Every particle's movement is dictated by its local position. Each particle has velocities that direct the flight of particles. The search space refers to the spectrum in which the technique calculates the most effective regulatory variables. The value will be reset if the searching space is exceeded by any particle's optimal control value.

4.2. Working procedure of PSO with an example

To better understand how PSO operates, let's look at an example. A flock of birds flying aimlessly in an area, trying to find a single piece of food. Not a single bird is aware of the location of the food, i.e., they are aware of their progress in each iteration, even though they are unsure of the ideal eating position. They launch themselves in different directions and adhere to the PSO's search plan, i.e., swiftly follow that bird that is close to food. Each particle or bird, starting from a randomly selected population, moves through the searching space in randomly chosen directions while recalling its best historical positions and those of its neighbors, i.e., the highest ranking globally. Follow that bird to the global best position to obtain the optimal value, i.e., food.

4.3. PSO Algorithm Fundamentals

Step I	Start
Step II	Initializes particles with velocity vectors (μ)and random positions
Step III	Use of Fitness equation: Find out the fitness of the particles
Step IV	Evaluate and Update p best and g best
Step V	Numerically solved and updating of the position of the velocity vectors
Step VI	Numerically solved and updating of the position of the velocity vectors.
Step VII	Numerically solved and updating of the position of the particles.
Step VIII	Termination Satisfies?
Step IX	Stop.

Table 2 : *Different notations used during Algorithm*

μ_{ij}^{n}	Represents particle's velocity vector i at time t in dimension j.
r_{ij}^{t}	Represents particle's position vector i at time t in dimension j.
$P_{Best,t}^{i}$	Represents particle's best position i through initializing time t in dimension j.
$L_{Best,t}^{i}$	The best position that any particle has had in the neighborhood of particle i in dimension j initialization through time t.
$a_{1,} and\ a_2$	Designates the constants of positive acceleration that are employed to balance the social and cognitive aspects, respectively.
$r_{1_j}^{t}\ and\ r_{2_j}^{t}$	Time-dependent random integers drawn from a uniform distribution.

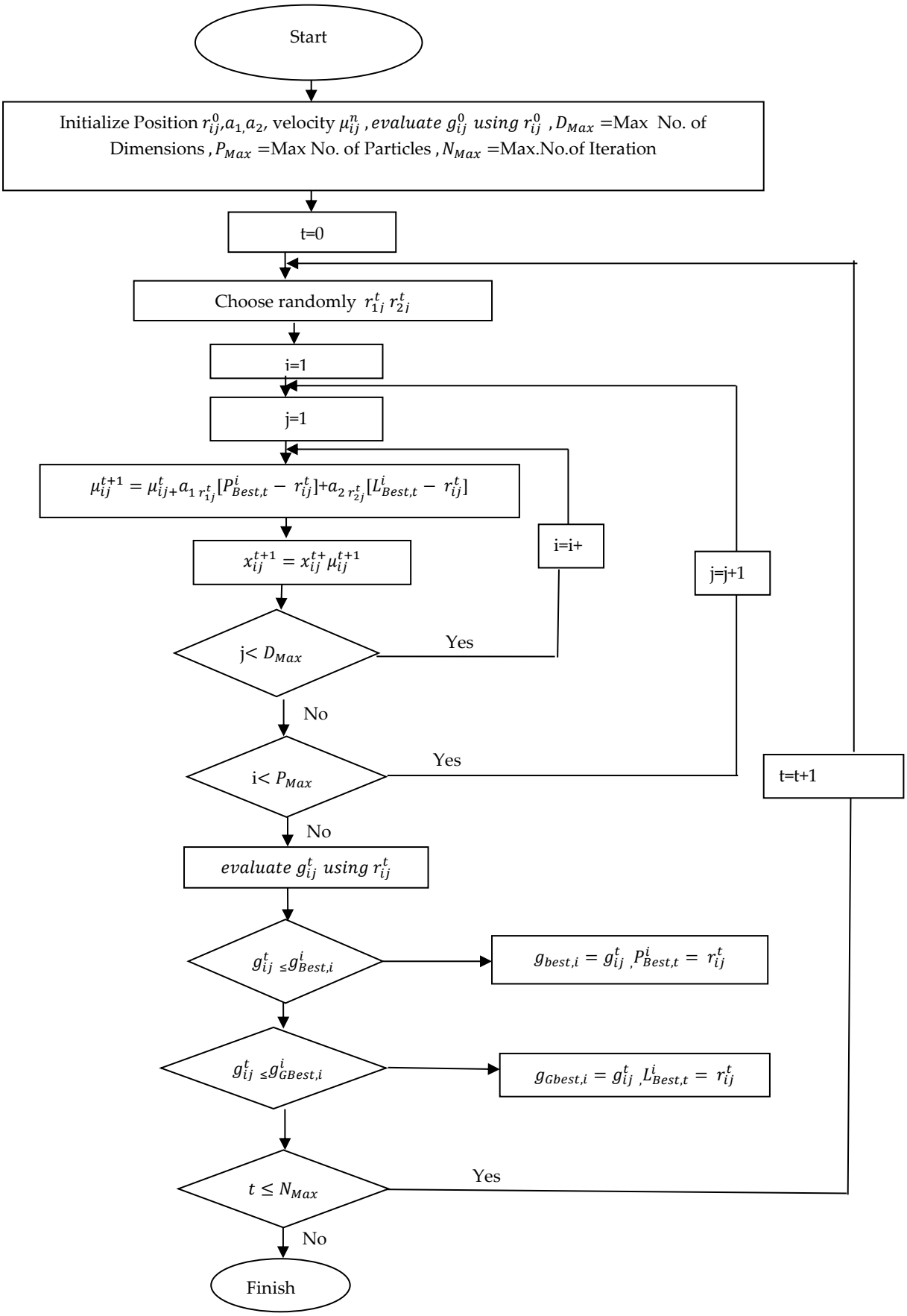


Figure 3: *Flow diagram of PSO*

4.4. Availability analysis of the industry with various failure($\mathbf{F_r}$) **and** repair($\mathbf{R_r}$) rate.

After applying the methodology of PSO over the fitness function i.e. an equation(55) for availability over different rate of failure and repair for each subsystem Q,R, S and T. We get the following variation presented from table 3 to 6 and corresponds figure 4 to 7.

Table 3. *Effect on Availability(A_{av}) of Mixture with various combination of Failure (F_r)and Repair Rate (R_r).*

β_1	A_{av}	σ_T	A_{av}
0.003122	0.31321	0.011276	0.50199
0.005446	0.30147	0.018296	0.66397
0.007046	0.25008	0.026777	0.70571
0.009462	0.14193	0.033979	0.81269
0.011776	0.13373	0.040393	0.85321
0.013864	0.13325	0.047671	0.93818
0.015398	0.08774	0.054925	0.95747
0.017696	0.07875	0.061103	0.99269

Table 4. *Effect on Availability(A_{av}) of Extruder with various combination of Failure (F_r)and Repair Rate (R_r)*

β_2	A_{av}	ϕ_Q	A_{av}
0.003517	0.31321	0.004498	0.40221
0.004683	0.27397	0.011015	0.52269
0.005914	0.15675	0.018255	0.59321
0.00665	0.13373	0.025938	0.64108
0.007282	0.102	0.032681	0.69269
0.008272	0.04782	0.039192	0.69269
0.009272	0.04782	0.046988	0.84675
0.010744	0.03037	0.053954	0.96397

Table 5. *Effect on Availability(A_{av}) of Dye with various combination of Failure (F_r)and Repair Rate (R_r)*

β_3	A_{av}	ψ_R	A_{av}
0.006091	0.37162	0.002535	0.57981
0.008081	0.31108	0.00958	0.62774
0.010847	0.24269	0.01655	0.67675
0.012656	0.19654	0.023479	0.74269
0.014349	0.08774	0.030528	0.8029
0.016375	0.06323	0.037996	0.86818
0.018533	0.05037	0.046239	0.89321
0.020701	0.11634	0.053487	0.90571

Table 6. *Effect on Availability(A_{av}) of Die with various combination of Failure (F_r)and Repair Rate (R_r)*

β_4	A_{av}	μ_S	A_{av}
0.007677	0.31321	0.003084	0.58774
0.008446	0.16134	0.010663	0.61221
0.009871	0.15675	0.017025	0.69682
0.010192	0.08774	0.024024	0.78571
0.011578	0.07981	0.031726	0.82397
0.012815	0.0629	0.037879	0.87321
0.013071	0.04774	0.044432	0.89269
0.014671	0.02037	0.051264	0.89408

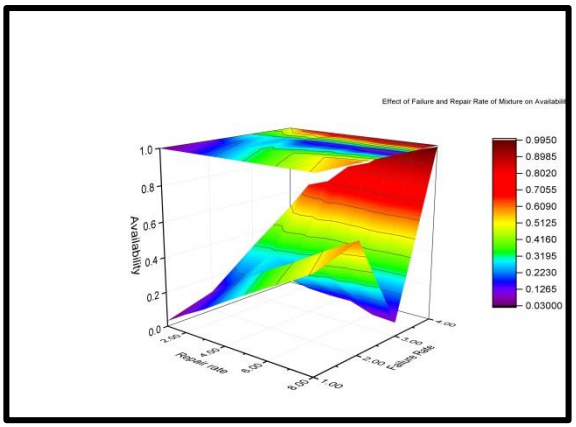


Figure 4 : *Effect on Availability(A_{av}) of Mixture with various combination of Failure (β_1)and Repair Rate (σ_T).*

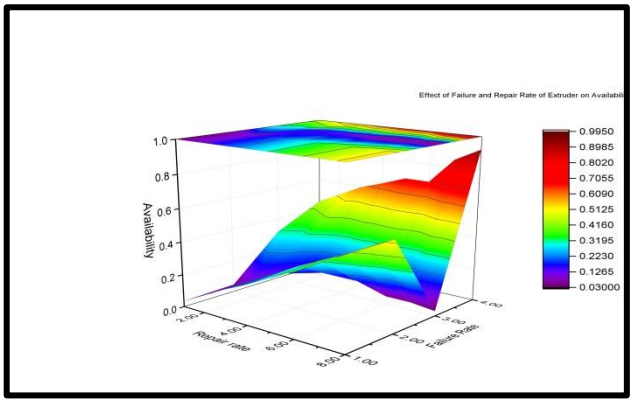


Figure 5 : *Effect on Availability(A_{av}) of Extruder with various combination of Failure (β_2) and Repair Rate (ϕ_Q).*

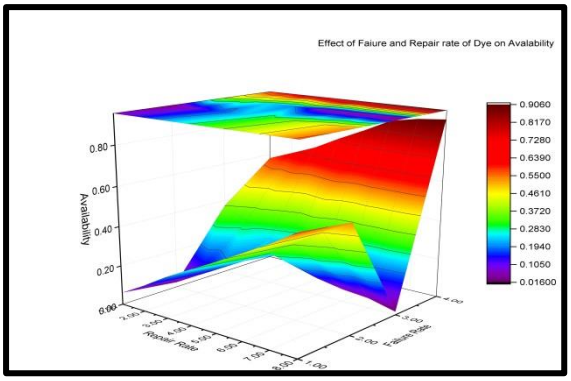


Figure 6 : *Effect on Availability(A_{av}) of Die with various combination of failure (β_3) and repair rate(ψ_R).*

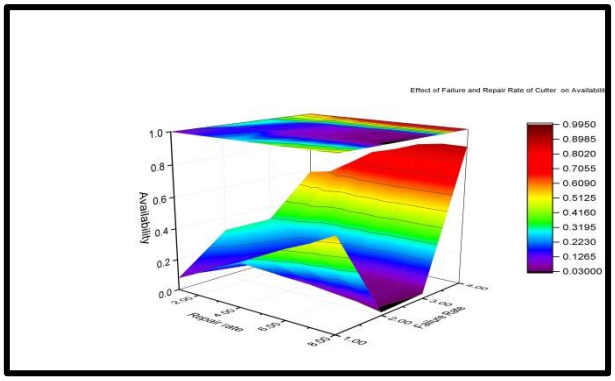


Figure 7 : *Effect on Availability(A_{av}) of Cutter with various combination of Failure (β_4)and Repair Rate (μ_S).*

V. Critical Analysis

The Critical Analysis has been done to demonstrate the impact of the subsystem on the overall system's availability than any of the other subsystems. Over long running of system ,How the availability percentage increased or decreased of a unit with respect to failure or repair rate with passage of time with comparison of other? To know the preference of subsystem to give more attention to met the requirements up to more efficient level, a critical analysis has been applied on the results obtain.

Table 7. *Critical Analysis of system component for their available performance time*

Systems	Failure Rate(F_r)	Decreases in Availability (A_{av})	Repair Rate(R_r)	Increase in availability (A_{av})	Repair Ranking (R_r)
Mixture	0.003122-0.017696-	23%	0.003122-0.061103	49%	II
Extruder	0.003517-0.010744	28%	0.004498-0.053954	56%	I
Dye	0.006091-	26%	0.002535-	32%	III

Cutter	0.020701 0.007677- 0.014671	19%	0.053487 0.003084- 0.051264	30%	IV

VI. Result and Observations

The following results are obtained after applying the methodology.

- Table 3 and figure 4 shown the variation in availability of component mixture with respect to both failure and repair rate which are depicted that with increase failure rate the availability decreases and with increasing repair rate, availability increases.
- Table 4 and figure 5 resulted about the behaviour of component extruder's availability with respect to both failure and repair rate which are depicted that with increase failure rate the availability decreases and with increasing repair rate, availability increases.
- Table 5 and figure 6 shown the variation in availability of component dye with respect to both failure and repair rate which are depicted that with increase failure rate the availability decreases and with increasing repair rate, availability increases.
- Table 6 and figure 7 shown the variation in availability of component cutter with respect to both failure and repair rate which are depicted that with increasing failure rate the availability decreases and with increasing repair rate, availability increases.
- Table 7 analyzed the component performance based on availability with failure and repair rate.

VII. Concluding observations

Availability of polytube is optimized by using particle swam optimization technique. During PSO maximum allowable velocity and weight parameters play a vital rule for analysis. For various combination of failure and repair rate many runs were performed. The optimal availability is achieved 99% at Failure rate increase from (0.003517-0.010744) and repair rate (0.004498-0.053954). The fundamental modeling of the manufacturing plant represents the flow of row material from one subsystem to another subsystem. With the help of transition diagram all the possible stages (Failure, Repair and reduced) are helpful to find the probabilities of every stage. To calculate (steady state and transit state) the sensitive analysis of failure and repair rate of the subsystem with maintenance strategy play an important rule for the management to get maximum availability of the system without failure. The table's (3-6) reveals the numerical analysis of the optimized availability of polytube manufacturing plant. Further the results are explained graphically from the (Figure3-6). The comparative analysis demonstrates that the subsystem Extruder has the greatest impact on the overall system's availability than any of the other subsystems. Other subsystems have a minor impact on the availability of the polytube manufacturing facility. For finding the optimal solution PSO techniques is very helpful . With the help of this techniques one can use different combination of failure and repair rate to get the optimal solution.

References

[1] Cox D. R. (1955). Analysis of Non Markovian stochastic processes by the inclusion of supplementary variables. *Proc Comb PhillSoc*, 51:433-441.

[2] Ying-Shen, J., Shui-Shun, L. and Hsing-Perk, K. (2008). A knowledge management system for series-parallel availability optimization and design. *Journal of Expert System and Application*, 34:181-193.

[3] Singla, S., Lal, A.K., Bhatia, S.S. and Singh, J. (2011). Reliability analysis of Polytube industry using supplementary variable technique. *Applied Mathematics and Computation*,218(11):3981-3992.

[4] Sagayaraj, M.R., Anita, S.P. and Merceline, A. (2014). Markov Models in system reliability with applications. *International Journal of Innovative Research and Development*, 3 (11):328-336.

[5] Yang, T. ,Cui, C., Shen, Y. and Lv, Y.(2016). A novel denitration cost optimization system for power unit boilers. *Appl. Therm. Eng*., 96:400-410.

[6] Choudhary, D., Tripathi, M. and Shankar, R. (2019). Reliability**,** Availability and Maintainability analysis of a cement plant: a case study. *International Journal of Quality & Reliability Management*, 36 (3):298-313.

[7] Kuo, C. and Ke, J. (2019). Availability and comparison of spare systems with a repairable server. *International Journal of Reliability Quality and Safety Engineering,* 26 (1):1-18.

[8] Kumari, S., Khurana, P. and Singla, S. (2021). Behavior and profit analysis of a thresher plant under steady state. *International Journal of System Assurance Engineering and Management*, 166-171.

[9] Kumari, S., Khurana, P. and Singla, S. (2021). RAP via constraint optimization genetic algorithm. *Life cycle of reliability and safety engineering*, 341-345.

[10] Kumari, S., Khurana, P., Singla, S. and Kumar, A. (2021). Solution of constrained problems using particle swarm optimization. *International Journal of System Assurance Engineering and Management*,1688-1695.

[11] Sharma, A., Ailawalia, P. and Singla, S. (2021). RAM (reliability, availability, maintainability) of threshing machine in agriculture. *Agriculture and Natural Resources*, 55:1057–1061.

[12] Garg, D., Garg, R. and Garg, V. (2022). Inspecting briquette machine with different faults. *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)* 15.4:481-486.

[13] 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.

[14] 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.

A STUDY ON THE IMPACT OF TRANSFORMING PARAMETER IN BOX - COX TRANSFORMATION FOR NON-NORMAL DATA TO ENHANCE PROCESS CAPABILITY

J. Krishnan[1] and R. Vijayaraghavan[2]

•

(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
[1]krrishme92@gmail.com, [2]vijaystatbu@gmail.com

Abstract

Process Capability Analysis (PCA) helps to improve and monitor the quality of the manufacturing products in industries. The most commonly and traditionally applied indices are process capability index and process capability ratio. Many statistical tests require the condition that the data to be approximately normally distributed. When it comes to reality the data often do not follow a normal distribution. In such instances, different approaches are employed. Box-Cox Transformation (BCT) is one such methodology that is often used by quality practitioners relying on single transforming parameter lamda to transform the non-normal data into normal data. The widely used approach to decide the transforming parameter lamda is based on the rounded value of lamda instead of an optimal value of lamda. There are two transforming expressions available in BCT method. The choice of the value for lamda in BCT can have a significant impact on the results. This paper concentrates on the impact of data transformation in BCT method through two different expressions based on an optimal as well as a rounded value of lamda. The influence made by the estimates of process capability and process performance indices is also studied in this paper. The result of the analysis clearly indicates that the optimal value of lamda when employed in the first BCT transformation expression to estimate the process capability indices for non-normal data provides improvised results. For data analysis, Ms-Excel and Minitab 21 software has been used in this study.

Keywords: Non-normal, Box-Cox Transformation, MLE, Process Capability, Six sigma

I. Introduction

Process capability analysis (PCA) is a continuous process of monitoring and improving the quality of finished products produced by industries. PCA addresses the issues relating to how well a manufacturing process meets the required specification and it requires most often that the data should obey the assumption of normal distribution. The traditional process capability indices are purely based on normality assumption. When it comes to reality, the data often do not follow a normal distribution. In such instances, different approaches are employed. Data transformation for preserving a somewhat normal distribution has been recommended in [1]. Box - Cox transformation (BCT) is one such methodology that is often used by quality practitioners. The empirical study made in [2] has demonstrated that the findings of transformed data are much

superior to the results of the original data (NT methods). Further, NT methods are found to be inadequate in capturing the capability of the process unless the underlying distribution is close to or approximately normal. NT methods are unsatisfactory because the distribution deviates significantly from normal. See, [3]. Process capability indices are calculated using samples of data based on short-term or within group variation, whereas performance indices are calculated using all the data points and long-term or overall variation [4]. The process capability indices are denoted by Cp and Cpk, and process performance indices are denoted by Pp and Ppk. A detailed review on various methods that are chosen for performance comparison in their ability to handle non-normality while computing the process capability indices is presented in [5]. The most commonly and traditionally applied indices by industries are process capability index Cp and process capability ratio Cpk, which are given below in Table 1 along with the respective performance indices, where $\bar{X}$ is the sample mean, USL is the upper specification limit and LSL is the lower specification limit.

Table 1: *Process Capability and Process Performance Indices*

Process capability indices	Process performance indices
$C_p = \frac{USL-LSL}{6\sigma_W}$	$P_p = \frac{USL-LSL}{6\sigma_{overall}}$
$C_{pk} = min\ (C_{PU}, C_{PL})$	$P_{pk} = min\ (C_{PU}, C_{PL})$
$C_{PU} = \frac{USL-\bar{\bar{x}}}{3\sigma_W}, \quad C_{PL} = \frac{\bar{\bar{x}}-LSL}{3\sigma_W}$	$P_{PU} = \frac{USL-\bar{\bar{x}}}{3\sigma_{overall}}, \ P_{PL} = \frac{\bar{\bar{x}}-LSL}{3\sigma_{overall}}$

A detailed review on process capability indices for non-normal data is presented in [3] with an emphasis on Box - Cox transformation (BCT) and on the parameter estimation approach utilizing a search method to estimate the process capabilities. In [6], a method of converting non-normal data into normal data is discussed and the transformed data is analyzed using the process capability indices (PCI). Further, an improved BCT model has been proposed to deal with the non-normal data and to calculate the process capability indices. The choice of the value of the transforming parameter, λ and the conversion formula in BCT might have a significant impact on the results. Hence, in this paper, the estimates of process capability indices are obtained and the analysis is carried out using the optimal as well as the rounded value of λ through BCT to obtain improvised estimates of process capability indices and process performance indices (PPI) within the standard of six sigma level. Thus, the objective of the study in this paper is to investigate the effectiveness of the BCT conversion formula and the optimal as well as the rounded value of λ in data transformation and in process capability analysis for non-normal data. This study would assist in suggesting the most efficient way of utilizing the BCT parameter λ in data transformation to approximate normal data as well as to estimate process capability and PPM values for non-normal data with the least amount of error.

II. Methodology

In data analysis, normally distributed independent observations with constant mean and variance are generally assumed. However, in reality, data frequently do not follow a normal distribution. A family of power transformation, termed as Box - Cox transformation (BCT), for a positive response variable X in such circumstances has been suggested by Box and Cox. See, [7]. The goal of BCT is to stabilize variance and make the data more closely resemble a normal distribution. The following is the conversion formula:

$$x^{\lambda} = \begin{cases} \frac{x^{\lambda}-1}{\lambda}, for\ \lambda \neq 0 \\ \log x, for\ \lambda = 0 \end{cases} \quad (1)$$

It may be noted that since an analysis of variance is unchanged by a linear transformation, the expression given as (1) is equivalent to

$$x^{\lambda} = \begin{cases} x^{\lambda}, for\ \lambda \neq 0 \\ \log x, for\ \lambda = 0 \end{cases} \quad (2)$$

The expression given as (1) is slightly preferable for theoretical analysis since it is continuous at $\lambda = 0$ {7}. The major effort in BCT is connected to the transformation of x to xλ, with the parameter λ 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. One may refer to [7] and [8]. The original non-normal data is used to estimate the value of λ, which is then used to convert the data into approximately or nearer to normal based on the value of λ [6]. Initially, λ is selected within a predetermined range of values. With the selected λ, one would assess the following:

$$L_{max} = -\frac{1}{2} \ln \hat{\sigma}^2 + \ln J(\lambda, X) \quad (3)$$

$$L_{max} = -\frac{1}{2} \ln \hat{\sigma}^2 + (\lambda - 1) \sum_{i=1}^{n} \ln X_i \quad (4)$$

where

$$(\lambda, X) = \sum_{i=1}^{n} \frac{\partial W_i}{\partial X_i} = \sum_{i=1}^{n} X_i^{\lambda - 1} \text{ for all } \lambda \quad (5)$$

From (5), one may have J (λ, X) = $(\lambda - 1) \sum_{i=1}^{n} \ln X_i$. The estimate of $\hat{\sigma}^2$ for fixed λ is defined by $\hat{\sigma}^2 = S(\lambda)/n$, where $S(\lambda)$ is the residual sum of squares in the analysis of variance of X. Plotting L_{max} against λ is possible after computing L_{max} for a number of λ values within the range. The value of λ that maximizes L_{max} yields the maximum likelihood estimator of λ. Using the expression (1) or (2), the data and specification limits are converted to a normal variate with the optimal value of λ [5]. The transformed observations x^{λ} are assumed to satisfy the normality assumption for an unknown λ. Based on this assumption, the transformed observations are used to estimate the mean and variance [7]. The maximum likelihood estimates of the mean and standard deviation are, respectively, given by

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i^{2} - 1}{\lambda} \quad (6)$$

and

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{X_i^{\lambda} - 1}{\lambda} - \frac{1}{n} \sum_{i=1}^{n} \frac{X_i^{\lambda} - 1}{\lambda} \right)^2} \quad (7)$$

It has been shown in [6] that, for a common probability distribution, the BCT correctly transforms non-normal data into normal with a success rate of over 97%. Validity of the process capability index based on the transformation method has been verified and concluded that the process capability index using BCT is effective. The choice of the value for λ in a Box - Cox transformation might have a significant impact on the results. In BCT, there are primarily two phases to convert non-normal data into normal data, the first one is to find the transforming parameter λ and then to generate transformed normal data by substituting the non-normal data along with the λ value in the appropriate BCT formula, which may be either expression (1) or its equivalent expression (2). In general, the rounded value of λ is preferred over optimal value of λ because of its ease of transformation, viz., square root transformation when λ = 0.50, cube root transformation when λ = 0.33, fourth root transformation when λ = 0.25, natural log transformation

when λ = 0.0, reciprocal square root transformation when λ = - 0.50, reciprocal transformation when λ = - 1.00 and no transformation when λ = 1.00 [3].

III. Numerical Illustrations

Process capability analysis for non-normal data is carried out to assess the capability of the non-normal process in a real-life situation. It is to be noted in [6] that the value of CPU and CPK are 0.73, which are below the benchmark value of 1.33 in industries and hence, the process is not capable. The corresponding PPM values are 14765 ensures that the process is within the standard of three sigma limit only. The rounded value of λ is widely used to transform non-normal data into normal data rather than the optimal value of λ in BCT through maximum likelihood estimation (MLE). Moreover, as pointed out earlier, there are two transforming expressions available in BCT method. One may refer to [7]. It is interesting to note that in [6], the rounded value of λ is taken for data transformation instead of the optimal value of λ. Expression (2) is used to transform the non-normal data into normal data than the expression (1) in BCT. The influence of data transformation utilizing optimal and rounded values of λ via expressions (1) and (2) of BCT method will be examined and the results of estimated process capability analysis will be compared to the results recorded in [6].

A producer, who produces mechanical parts, wants to analyze whether one type of mechanical components comply with the specification. The real data (RD) set pertaining to warping presented in Table 2 is extracted from [6]. In order to assure production quality, the measured warping of mechanical parts should not exceed 9.5 (i.e., USL = 9.5). Before starting with the real data to carry out process capability analysis, it is essential to ensure that the data is normal. According to the result given in [6] and from the summary of the real data set, the p-value is found to be 0.012, which is less than the prescribed 5% level of significance. However, 5% of the data points lie outside the 95% confidence interval and therefore, the hypothesis that the data follow a normal distribution could be rejected, establishing the fact that the data would not follow normal distribution. One may refer to Table 3. Hence, the real data must be transformed in order to ensure normality.

Table 2: *Real Data Set of Warping*

2.000	2.887	0.650	3.612	0.550	6.637	1.525	5.300	1.400	0.350
1.050	3.187	1.262	3.574	3.100	2.400	7.900	4.012	0.975	0.712
3.750	5.899	1.400	2.725	4.887	1.525	4.750	4.350	5.175	0.875
1.612	2.187	3.200	1.312	2.849	0.950	5.274	8.325	6.625	3.550
2.800	2.025	5.287	1.562	1.200	2.987	5.412	3.050	4.737	7.812
3.287	7.037	1.675	2.462	6.225	6.200	0.525	4.387	4.050	4.212
0.425	5.800	3.550	1.050	7.237	2.450	1.500	9.037	6.300	4.037
8.700	4.937	0.950	4.149	3.150	1.687	4.300	1.412	3.825	7.600
4.325	5.475	3.474	5.187	3.850	5.987	3.137	5.337	3.062	2.074
1.762	4.050	0.787	3.212	4.774	2.750	9.112	2.562	5.862	2.650

Table 3: *Summary of Real Data*

Variable	Mean	SD	Min	Median	Max	Skewness	Kurtosis	P-value
RD	3.628	2.178	0.350	3.250	9.112	0.58	-0.31	0.012

BCT makes use of the MLE approach to figure out the single transformation parameter using expressions (3), (4), and (5). As described earlier, there are two methods to transform non-normal data into normal data depending on the value obtained in the Box - Cox plot, namely, the

estimated value and the rounded value of λ. It is significant to note from the Box - Cox plot that the rounded value of λ is 0.50, and the estimate (optimal) of λ is 0.45. The corresponding lower and higher confidence limits are 0.24 and 0.70. One may refer to Figure 1 for further details. In [6], the rounded value (RV) of λ is considered instead of optimal value (OV) of λ to transform non-normal data into normal data through BCT using expression (2) and hence, the transforming expression reduces to square root transformation when $\lambda = 0.50$. The optimal and rounded values of λ have been taken into consideration while transforming non-normal data into normal data and both the expressions (1) and (2) are utilized to perform process capability analysis.

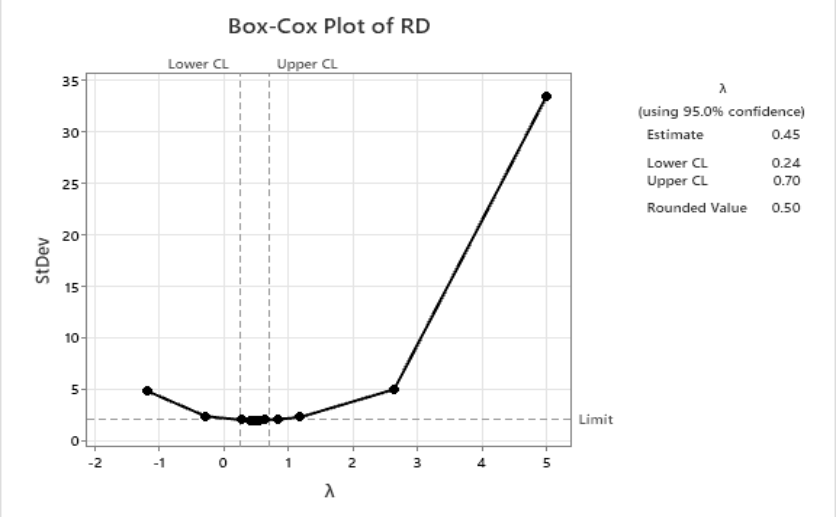


Figure 1: *Estimate of λ through Box - Cox Plot for RD*

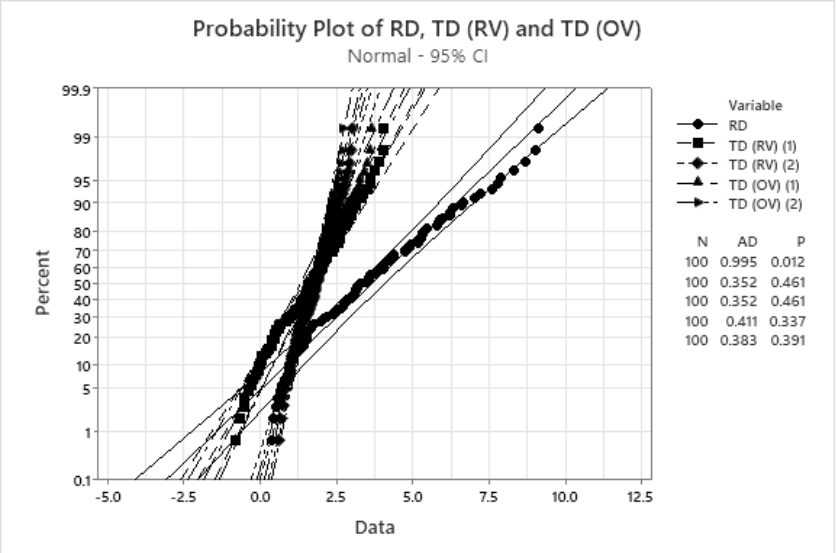


Figure 2: *Probability Plot for RD, TD (RV) and TD (OV) based on expression (1) and (2)*

Here, expressions (1) and (2) are used to convert the non-normal real data into normal data based on optimal and rounded value of λ. Table 4 displays the summary values of transformed data, whereas Figure 2 displays the probability plot of the real data and the transformed data based on RV and OV of λ. The p -values corresponding to the optimal and rounded λ tabulated in Table 4 are greater than the 5% and 1% level of significance, which would indicate that the data would follow a normal distribution. Hence, the transformed data can be used to estimate process capability and carry out process performance analysis.

Table 4: *Summary of Transformed Data using BCT Method*

Variable	λ Value	Mean	SE of Mean	Min	Median	Max	Skewness	Kurtosis	P-value
RD	-	3.628	0.218	0.350	3.250	9.112	0.58	-0.31	0.012
TD (RV) (1)	0.50	1.810	0.059	0.591	1.802	3.018	-0.04	-0.73	0.461
TD (RV) (2)	0.50	1.620	0.119	-0.081	1.605	4.037	-0.04	-0.73	0.462
TD (OV) (1)	0.45	1.697	0.051	0.623	1.699	2.702	-0.10	-0.72	0.337
TD (OV) (2)	0.45	1.509	0.109	-0.084	1.525	3.642	-0.14	-0.70	0.391

* TD (RV) (1) – Transformed data based on rounded value of λ through BCT expression 1 ǀ TD (RV) (2) – Transformed data based on rounded value of λ through BCT expression 2 ǀ TD (OV) (1) – Transformed data based on rounded value of λ through BCT expression 1 ǀ TD (OV) (2) – Transformed data based on rounded value of λ through BCT expression 1

I. Estimate of PCI and PPI Utilizing Rounded Value of λ Through BCT

The estimates of process capability indices CPU and CPK are 1.34 and process performance indices PPU and PPK are 1.32 respectively, when utilizing RV of λ for data transformation through BCT expression (1) and CPU and CPK are 0.73, and PPU and PPK are 0.71 through BCT expression (2). The result obtained from expression (1) based on OV of λ is approximately equal to the guideline value 1.33 and hence, the process is capable to produce the mechanical parts within the given specification limit and the respective PPM values of process capability and process performance indices are 28 and 38 ensuring that the process is better than 5σ limits. On the other hand the result obtained from expression (2) does not meet the standard of guideline value 1.33 in industries and the respective PPM values are 14778 and 16303, which conform that the process is within 3σ limit only and hence, the process may be considered as incapable. Additionally, some data points are beyond specification limits. Hence, when RV of λ is utilized to transform non-normal data into normal data through BCT expression (1), better results would be possible rather than BCT expression (2). One may refer Tables 5 and 6, and Figures 3 and 4 for more information.

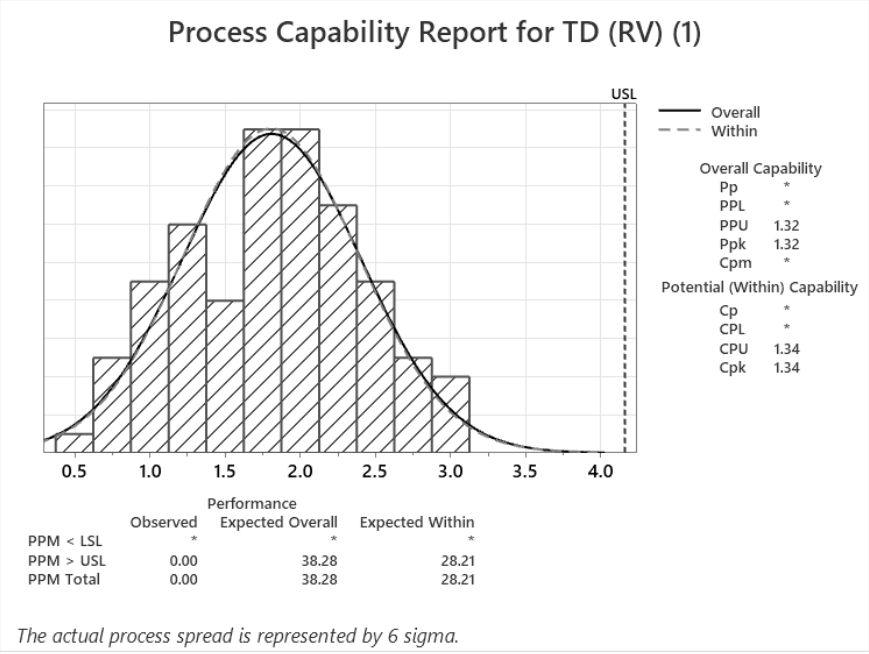


Figure 3: Estimate of Process Capability and Performance Indices for TD (RV) Data Based on Expression (1) in BCT Method

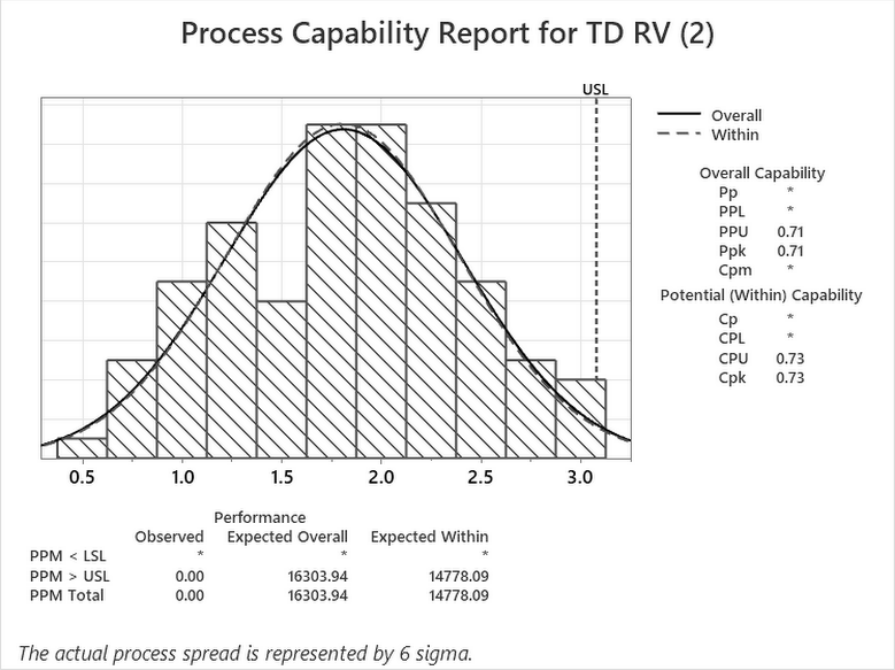


Figure 4: *Estimate of Process Capability and Performance Indices for TD (RV) Data Based on Expression (2) in BCT Method*

Table 5: *Estimate of Process Capability and Performance Indices Based on Transformed Data*

Variable	λ	PCI and PPI of transformation data based on expression (1); $x^{\lambda} = (x^{0.45} - 1)/0.45$ in BCT					PCI and PPI of transformation data based on expression (2); $x^{\lambda} = x^{0.50}$ in BCT				
		USL	C_{pk}	PPM	P_{pk}	PPM	USL	C_{pk}	PPM	P_{pk}	PPM
TD (RV)	0.50	4.164	1.34	28	1.32	38	3.082	073	14778	0.71	16303
TD (OV)	0.45	3.898	1.47	5	1.45	7	2.754	0.71	17001	0.69	18689

Table 6: *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

II. Estimate of PCI and PPI Utilizing Optimal Value of λ Through BCT

The estimate of process capability indices CPU and CPK are 1.47 and process performance indices PPU and PPK are 1.45 respectively, when utilizing OV of λ through BCT expression (1) and CPU and CPK are 0.71 and PPU and PPK are 0.69 through BCT expression (2). The result attained using OV of λ through expression (1) is greater than the guideline value 1.33 and the capability of the process is excellent. Hence, the process is capable to produce the mechanical parts within the given specification limit. The respective PPM values of process capability and process performance indices are 5.05 and 7.29, which guarantee that the process is better than 5σ limit and approximately follow 6σ result, whereas the actual PPM value corresponding to 6σ is 3.4 [9].

Conversely, the result observed from expression (2) is below the guideline of 1.33, which would indicate that the process could not be considered as capable and the respective PPM values are 17007 and 18689, ensuring that the process is within the standard of 3σ limit. Hence, making use of OV of λ to transform non-normal data into normal data through BCT expression (1) provides better results than BCT expression (2). One may refer to Table 5 and 6, and Figure 5 and 6 for details.

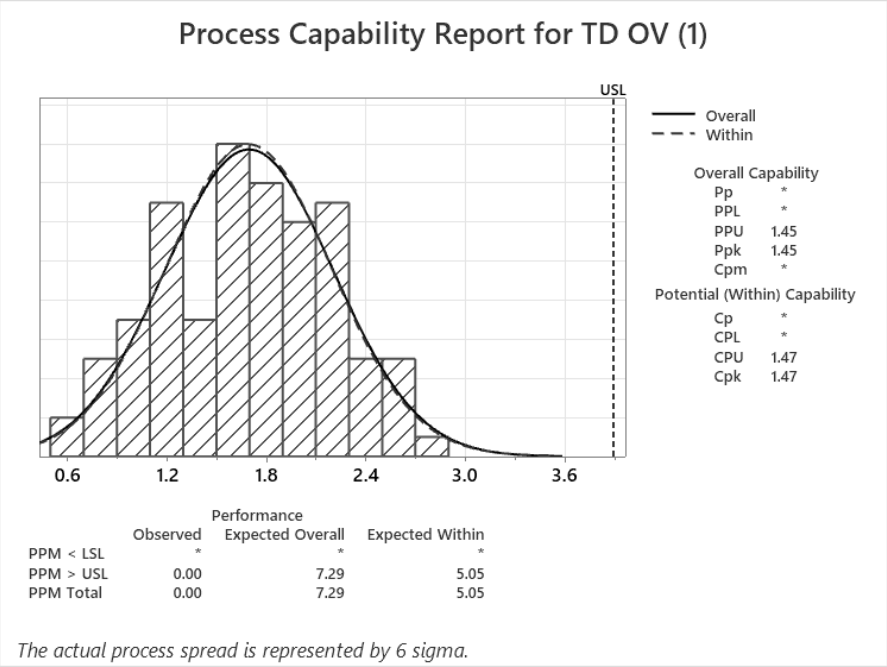


Figure 5: *Estimate of Process Capability and Performance Indices for TD (OV) Data Based on Expression (1) in BCT method*

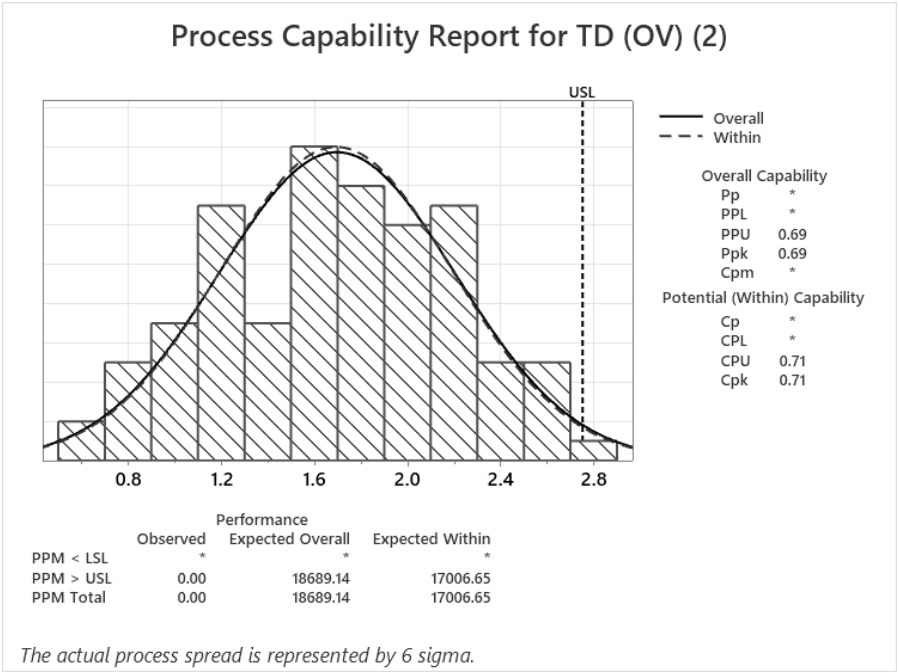


Figure 6: *Estimate of Process Capability and Performance Indices for TD (OV) Data Based on Expression (2) in BCT Method*

A process is categorized as inadequate, if PCI < 1.00; capable, if 1.00 ≤ PCI ≤ 1.33; satisfactory, if 1.33 ≤ PCI ≤ 1.50; excellent, if 1.50 ≤ PCI ≤ 2.00; and super, if ≥ 2.00. Automotive industries use CPK = 1.33 as a benchmark in assessing the capability of the process. If Cp and CPK are greater than or equal to 2 and 1.5, respectively, a process is said to be under six-sigma controls. Similarly, Pp and PPK must be more than 2 and 1.5, respectively, for a process to generate six-sigma results. See, [9]. Table 6 lists the process fallout in PPM in relation to the proportion of good items and PPM values for various sigma levels.

IV. Result and Discussion

The primary goal of this study is to investigate the effect of rounded value (RV) and optimal value (OV) of the transforming parameter λ, on data transformation and estimation of process capability and process performance indices through the BCT method. Furthermore, in order to obtain the transformed normal data and the estimates of process capability indices (PCIs) and process performance indices (PPIs), the identified λ values, such as RV and OV of λ must be substituted in BCT using expressions (1) and (2). The estimates of CPU and CPK in [6] are 0.73 while PPU and PPK are 0.71, respectively. These values are below the industry benchmark value of 1.33, indicating that the process is not capable. Remarkably, the RV of λ is used for estimating PCI and PPI, and also for data transformation rather than the OV of λ. However, as demonstrated by a numerical example in this study, the use of optimal λ value to convert non-normal data into normal data produces results that are as near to normal as possible and have a smaller standard error of mean than the RV of λ. One may refer to Table 4. Hence, the industry benchmark value of 1.33 is met by the estimated CPU and CPK of 1.47 and PPU and PPK of 1.45, respectively, based on OV of λ through BCT expression (1).

Process is, therefore, as demonstrated in Table 5 and 6, thought to be capable of producing manufacturing parts that meet the specification limit. Besides the estimates of process capability and process performance indices, the respective PPM values are essential in assessing the process fallout. It may be noted that only 3.4 defective items out of every million products should have been recorded in a production process that conforms to the standard of 6σ. When using the OV of λ in expression (1) of the BCT method, the process approximately follows 6σ outcome, as indicated by the PPM values of 5.05 and 7.29. One may refer to Table 5, and Figure 5 and 6.

The mean and range charts (X bar – R charts) are drawn for transformed data based on the OV and RV of λ, in order to clearly visualize the statistical control over the process of non-normal real data. All of these data points in Figures 7, 8, 9 and 10 fall within the control limits, indicating that the process is statistically under control. However, compared to the results obtained from a RV of λ, the estimate and PPM values corresponding to transformed data based on OV of λ produced significantly better outcomes.

It is evident from Table 4, and Figures 7, 8, 9 and 10 that when data is transformed using the OV of λ, the standard error of mean obtained is smaller than the one obtained from the RV of λ. Hence, it is necessary to use optimal λ to get better results, as it reflects the exact transforming pattern rather than a rounded value and this ensures all data points as close to normal as possible. Specifically, the utilization of BCT expression (1) to convert non-normal data into normal data produces better results than using BCT expression (2) when an OV of λ is used. As pointed out in [7], the BCT expression (1) is slightly preferable than expression (2) for theoretical analysis because it is continuous at $\lambda = 0$. Also, it can be observed from Figures 4 and 6 that some of the data points fall beyond the specification limit conforming that, the process is not capable to produce the mechanical parts within the speciation limit when estimated using RV of λ through expression (2). This result can be compared with [6] to understand that, using OV of λ through BCT expression (1) produces improvised results that are nearer to the standard of 6σ than using Rounded RV of λ.

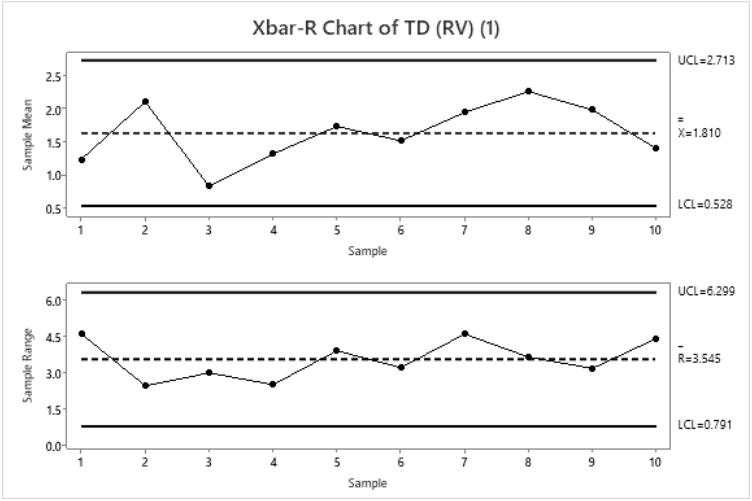


Figure 7: *Xbar and R Chart for Transformed Data using Rounded Value of λ through expression (1)*

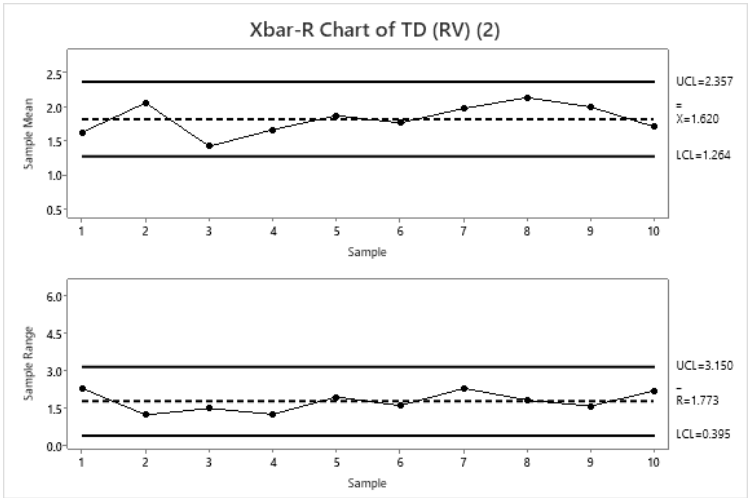


Figure 8: *Xbar and R Chart for Transformed Data using Rounded Value of λ through expression (2)*

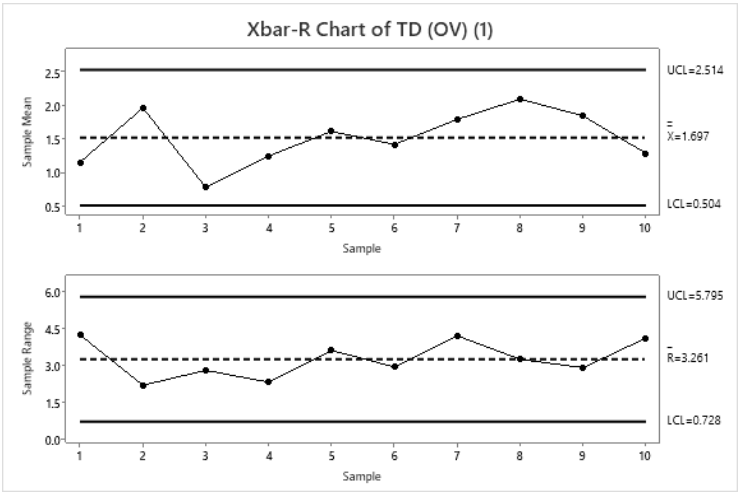


Figure 9: *Xbar and R Charts for Transformed Data using Optimal Value of λ through expression (1)*

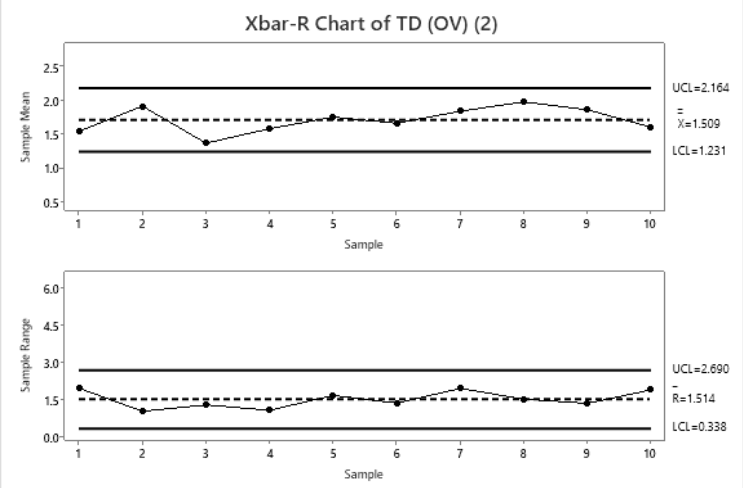


Figure 10: *Xbar and R Charts for Transformed Data using Optimal Value of λ through expression (2)*

V. Conclusion

The methodology of continuously assessing and enhancing the quality of manufacturing products in industries is known as process capability analysis. In order to handle the problems pertaining to how well a manufacturing process satisfies the necessary specifications, PCA needs that the data should follow the normal distribution. If the normality assumptions are violated or failed, some adjustments must be made to the traditional process capability indices, which would solely depend on normality assumptions. A significant methodology for handling non-normal data is data transformation. One such methodology is the Box-Cox Transformation. Using a single transforming parameter λ, which can be either a rounded value from the MLE approach or an estimated (optimal) value, the non-normal data can be converted into normal data. By taking into account of the objective of this paper, both optimal and rounded value of λ are considered for data transformation through two BCT expressions. Based on the data analysis, improvised estimates of PCIs and PPIs are obtained, when utilizing an optimal value of λ through BCT expression (1) rather than the rounded value of λ through BCT expression (2). The corresponding PPM values of the PCIs and PPIs are very smaller and approximately follow the case of standard of 6σ. Furthermore, it is evident that the transformed data is nearer to normal with a smaller standard error of mean when utilizing optimal value of λ. Thus, one can conclude that the improvised estimates would generally be obtained by utilizing optimal value of λ than the rounded value of λ when using Box-Cox transformation. In particular, the most effective results for PCI's and PPI's are quite possible only when utilizing optimal value of λ through BCT expression (1) rather than BCT expression (2).

References

[1] Kane, V. E. (1986), Process Capability Indices, Journal of Quality Technology, 18, 41 – 52.

[2] Gunter, B. H. (1989). The Use and Abuse of Cpk, Quality Progress, 22, 108 – 109.

[3] 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, 2016, Mysuru, India.

[4] 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, 49 – 55.

[5] Tang, L. C., and Than, S. E. (1999). Computing Process Capability Indices for Non-normal Data: A Review and Comparative Study, Quality and Reliability Engineering International, 15, 339 – 353.

[6] Yang Y and Zhu H (2018). A Study on Non-normal Process Capability Analysis based on Box-Cox Transformation, Proceedings of the 3rd International Conference on Computational Intelligence and Applications (ICCIA), Hong Kong, China, IEEE, 240 – 243.

[7] Box, G. E. P., and Cox, D. R. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26, 211 - 243.

[8] Asar, O., Ilk, O., and Dag, O. (2017). Estimating Box-Cox Power Transformation Parameter via Goodness-of-Fit Tests, Communications in Statistics - Simulation and Computation, 46, 91 – 105.

[9] Pearn, W. L., and Chen, K. -S. (2002). One sided Capability Indices CPU and CPL: Decision Making with Sample Information, International Journal of Quality & Reliability Management, 19, 221 – 245.

RT&A, No 2 (78)
Volume 19, June, 2024

ANALYSIS OF PERSONALIZED STRESS RECOGNITION IN THE OFFICE ENVIRONMENT

Jigna Jadav[1], Dr. Uttam Chauhan[2]

•

1 Research Scholar, Computer Engineering, Gujarat Technological University, Ahmedabad, India. Jigna_jadav@vgecg.ac.in

2 Assistant Professor, Computer Engineering department, Vishwakarma Government Engineering College Chandkheda, Ahmedabad, India. ug_chauhan@gtu.edu.in

Abstract

In today's fast-paced lifestyle, pursuing holistic well-being has increased interest in monitoring and managing stress levels. Heart rate variability (HRV), a non-invasive measure of autonomic nervous system activity, has emerged as a valuable tool for assessing individual responses to stress. This study focuses on utilizing the capabilities of the Apple Watch to collect continuous HRV data in real-world contexts. A diverse dataset from individuals working in software companies was gathered, including HRV recordings during various stress-inducing scenarios. By employing HRV Time Domain, Frequency Domain, and Nonlinear features, the study uses Principal Component Analysis (PCA) to extract relevant features, considering the personalized nature of stress reactions. Addressing variations in stress responses among individuals, the study introduces an innovative approach using Long Short-Term Memory (LSTM) networks. A hybrid model, combining feature selection, dimensionality reduction, and ensemble techniques, is developed to predict stress levels based on individualized HRV patterns. Rigorous training and validation reached to an 88% accuracy rate. These findings demonstrate the effectiveness of the proposed methodology. The LSTM model accurately forecasts stress responses, highlighting the potential of Apple Watch-acquired HRV data for stress assessment. Beyond prediction, the study enhances understanding of the complex interplay between HRV dynamics and unique stress reactions. This novel approach, leveraging Apple Watch features and intelligent computing, offers a personalized method to predict stress levels using K-Means Clustering Algorithm. Through integrating K-means clustering and person-specific HRV analysis, the research endeavours to advance our comprehension of the intricate interplay between physiological responses and stressors. The study offers a novel perspective on stress response variations by delving into the distinct autonomic patterns characterizing each cluster. It sets the stage for developing targeted interventions and personalized stress management strategies.

Keywords: Stress Detection, Apple Watch Dataset, HRV, LSTM

I. Introduction

Nowadays, stress detection research has made strides with cutting-edge methods. Wearable tech captured heartbeat dynamics for stress prediction in college students [1]. Wearable sensors gathered diverse physiological signals, showcasing multi-dimensional stress responses [2]. Deep learning in 2022, with "Stress Detection Using Deep Convolutional Neural Networks," revealed patterns in physiological data [3]. "Stress Recognition Using Wearable Sensors and Mobile Phones" combined wearables and mobiles for accessible stress assessment [4]. Collectively, these studies illuminate stress's nuances via tech-driven insights. Wearables and deep learning enhance stress detection's precision, potentially personalizing interventions. This dynamic interdisciplinary progress ushers in more accurate, accessible strategies for stress assessment.

Many Datasets are publicly available for stress recognition. The novelty resides in the integration of HRV analysis. No prior study has generated a continuous 15-day dataset from working professionals in software companies using the Apple Watch. This dataset serves as a rich resource

for understanding stress over time. Furthermore, the study introduces personalized stress detection through clustering, recognizing the diversity of stress manifestations. As individuals navigate the challenges of modern work environments, their physiological responses to stress manifest in unique ways. The dataset is augmented with discrete emotion labels corresponding to their tasks, specifically Neutral, Stress, and Not Stress. This dataset provides a foundation for personalized stress assessment, acknowledging that stress responses are distinct for each individual.

II. Proposed Methodology

The physiological signal calculated by the time interval (R-R Interval) between consecutive heartbeats in milliseconds is known as heart rate variability. The supportive branch of the autonomic nervous system (ANS) controls the stress or reaction, preparing us to act, respond, and conduct in rebuttal to life's diverse needs. The time between heartbeats (R-R interval) varies from beat to beat, and this variation in HRV can reveal a lot about the body's physiological state. HRV should naturally rise during relaxing activities and fall during stressful situations when the body is able to take advantage of increased sympathetic action. Heart rate variability is higher when the heart beats slowly; when the heart rate increases, such as during stress or exercise, it decreases during relaxing activities. Heart rate and HRV are in the inverse relation. The Heart rate variability level intuitively varies daily depending on activity, anxiety, and work-related stress. The duration between heartbeats (R-R interval) fluctuates from beat to beat and can give information about the body's physiological reaction.

When investigated in a deeper context, stress is detrimental in workplace situations. According to The American Institute of Stress [5], 80% of workers feel stress on the job, so we have decided to detect stress in working employees. Here, we have mentioned different datasets available to see the stress conditions of persons using physiological Signals.

Authors [6] (Park & Kim, 2018) used an HRV signal to predict a daily mental stress level using a photoplethysmography (PPG) sensor in the wristband-type wearable device. They extracted low-frequency (0.04Hz – 0.15Hz) and high-frequency (0.15Hz – 0.4Hz) features of HRV using the autoregressive (AR) model. Eight university students' data was collected using a self-evaluation PSS scale for 30 seconds thrice daily for a week. Linear regression provided an accuracy of 86.35%, although additional machine learning algorithms and well-known PPG analytic tools can produce better outcomes. ten users wore the FITBIT device to detect stress and an online questionnaire. In addition, it measures different physical activities like sleeping patterns, BMI, and Heart rate variability[7].

The Heart rate (HR), galvanic skin response (GSR), and electrooculogram (EOG) signals are collected from 11 subjects. The participants were also given a mental arithmetic task and a challenging LEGO assembly without instructions to predict stress. They applied a k-means clustering algorithm for heart rate, EDA, and EOG and got an accuracy of 70.6 percent, 74.6 percent, and 63.7 percent, respectively[8].

To identify different physiological changes during a stressful task. The Trier Stress Test was used to prompt stress, with resting and stress phase ECGs, and the inter-second heart rate was recorded (using a FitBit). The study enlisted the participation of 30 student doctors and 30 general public. More investigation with a large sample of people with stratified anxiety scores based on the Depression Anxiety Stress Scale is required to further analyze the association with HRV [9].

The WESAD (Wearable Stress and Affect Detection) dataset [10] is a publicly available dataset used for research in affective computing and physiological signal analysis. It was developed to support developing and evaluating algorithms and models for stress and affect detection using wearable sensors. The dataset includes physiological sensor data collected from wearable devices, such as heart rate sensors and accelerometers, and self-assessment labels related to the participants' stress

levels and affective states. The data was collected from 15 participants in a controlled environment while they underwent different stress-inducing tasks and activities.

AMIGOS [11] was designed to collect participants' emotions in two social contexts: individual and group. AMIGOS was constructed in 2 experimental settings. First, 40 participants watched 16 short emotional videos. Then, they watched four long videos, including a mix of lone and group sessions. These emotions were annotated with self-assessment of affective levels and external assessment of valence and arousal through GSR and ECG signals.

The SWELL dataset [12] comprises heart rate variability (HRV) indices derived from the multimodal SWELL knowledge work (SWELL-KW) dataset, designed for research on stress and user modelling. This dataset was developed by researchers at the Institute for Computing and Information Sciences at Radboud University. The SWELL dataset was created through experiments involving 25 subjects engaged in typical office work activities, such as writing reports, making presentations, reading emails, and searching for information. The dataset captures various data modalities, including computer logging, facial expressions, body postures, ECG signals, and skin conductance. Each participant in the study underwent three different working conditions: stress, time pressure, and interruption.

The author was involved in curating a Social Media Status dataset highlighting three key emotions: happiness, sadness, and anger. This dataset was drawn from status updates contributed by seven distinct individuals and acquired from Kaggle. The dataset's core focus was emotions, with entries structured to include the status text and corresponding sentiment. Achieved an accuracy of around 79% using a CNN classifier [13].

I. Extraction of Heart Rate from Apple Watch

An optical heart sensor in the Apple Watch SE, as shown in Figure 1, 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.

Figure 1: *Apple Watch SE*

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. HRV in real-world scenarios and highlighted its small relationship with stress compared to controlled lab settings[14]. This underscores the importance and complexity of real-world data collection, offering insights that can be challenging to deduce.

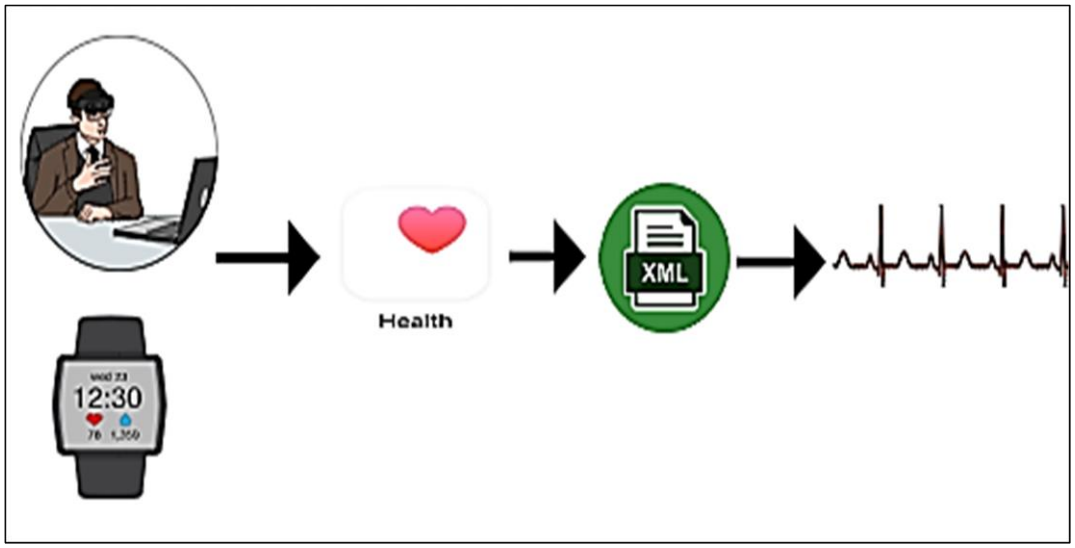


Figure 2: *Extraction Process of Heart Rate Using Apple Watch*

Figure 2 shows the process of extracting heart rate using The Apple Watch and the Health application on your paired iPhone. It can accurately measure your heart rate.

- Ensure your Apple Watch is correctly worn on your wrist.
- Tap the Heart Rate app on your Apple Watch.
- Begin measuring your heart rate within the app.
- Wait a few seconds for your heart rate to display on the Watch.
- The data is automatically synced to the Health app on your paired iPhone.
- Open the Health app on your iPhone to see heart rate data trends.
- Download Export.XML file
- Extract Heart Rate Data from that XML file.

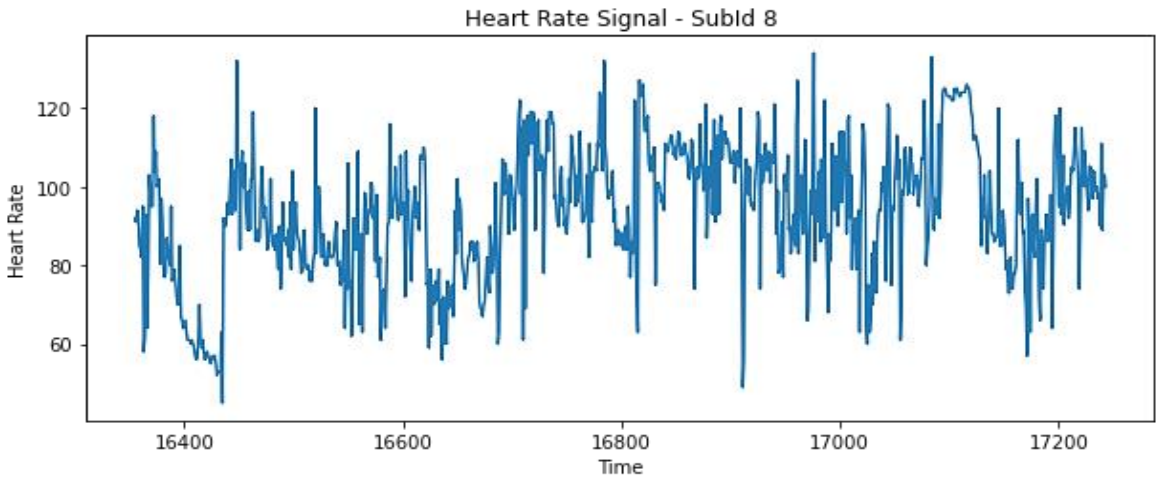


Figure 3: *Heart Rate Signals of Subjects*

Figure 3 shows individual plots for the subject's heart rate signal, showcasing the variations in their heart rate over time. The exact appearance of the plots and the specific data details depend on the content of the CSV file. This t would be useful for visualizing and analyzing heart rate variability among different subjects in the dataset.

II. File Preprocessing

Figure 4 explains the heart rate variability (HRV) data analysis derived from XML files, particularly focusing on data collected through Apple Watch devices. The script's main objective is to calculate a diverse array of HRV features encompassing both time domain and frequency domain metrics, subsequently organizing and storing these features within a CSV (Comma-Separated Values) file. Several critical libraries for XML parsing, HRV analysis, numerical computations, CSV handling, file searches, and operating system interactions are imported to initiate this processed. By defining a designated time frame using start and end dates, the script specifies the period for data extraction and evaluation.

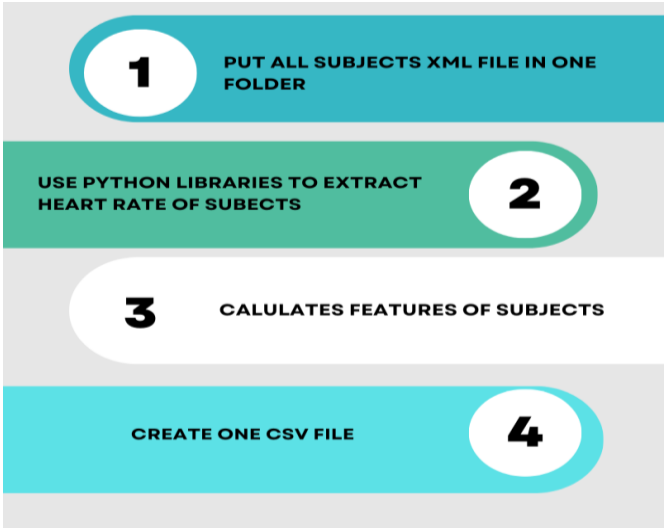


Figure 4: *Process of File Generation*

Utilizing the glob function, the script locates the relevant XML files containing heart rate data. The CSV file that will house the calculated HRV features is opened, and its initial row is allocated for labels describing the different metrics that will be computed and saved.

III. Features Calculations & Selection

The core functionality of the script involves iterating through the identified XML files. 'Record' elements are examined within each file to determine pertinent heart rate data. Valid heart rate measurements are isolated by cross-referencing the data's time stamps with the designated time frame. Subsequently, the script calculates RR intervals, the temporal gaps between consecutive heartbeats, from the heart rate values. It forms the basis for HRV analysis, then employs the HRV analysis library to address missing values, perform frequency domain analysis and compute time domain. Here is the summarized description of HRV analysis. The data is obtained from an Apple Health export file and is parsed using the xml.etree Element Tree module. We define a specific date range to query the data for analysis.

First, we iterate over the XML file to extract heart rate values recorded within the specified date range. We filter out the relevant data based on the sample type, explicitly focusing on heart rate measurements ('HKQuantityTypeIdentifierHeartRate'). The extracted heart rate values are stored in a list called heart rates. Next, we calculate the RR intervals from the extracted heart rate values, representing the time between successive heartbeats. We utilize a formula to estimate the RR intervals from the heart rate values, considering the average duration between successive heartbeats. Additionally, we apply the Malik rule, which was implemented through the hrvanalysis.remove_ectopic_beats function to identify and remove ectopic beats from the RR

interval data.

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 [15]. 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 [16]. The Percentage of NN50 Intervals (pNN50) gauges parasympathetic influence by identifying RR intervals differing by over 50 ms [17].

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 [17]. The LF/HF ratio quantifies sympathetic-parasympathetic balance [18].

The nonlinear 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 [19].

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 [20].

To organize and store the accumulated HRV features, the script combines these metrics with the corresponding heart rate values and unique subject identifiers. This composite data is structured into arrays and consistently added as new rows within the previously opened CSV file. As the script concludes the analysis for each subject, it echoes the calculated HRV feature arrays to the console. This Python script provides an automated and systematic approach to parsing, analyzing, and storing HRV data from Apple Watch-generated XML files. It facilitates an in-depth exploration and understanding of physiological monitoring and health analysis within heart rate variability.

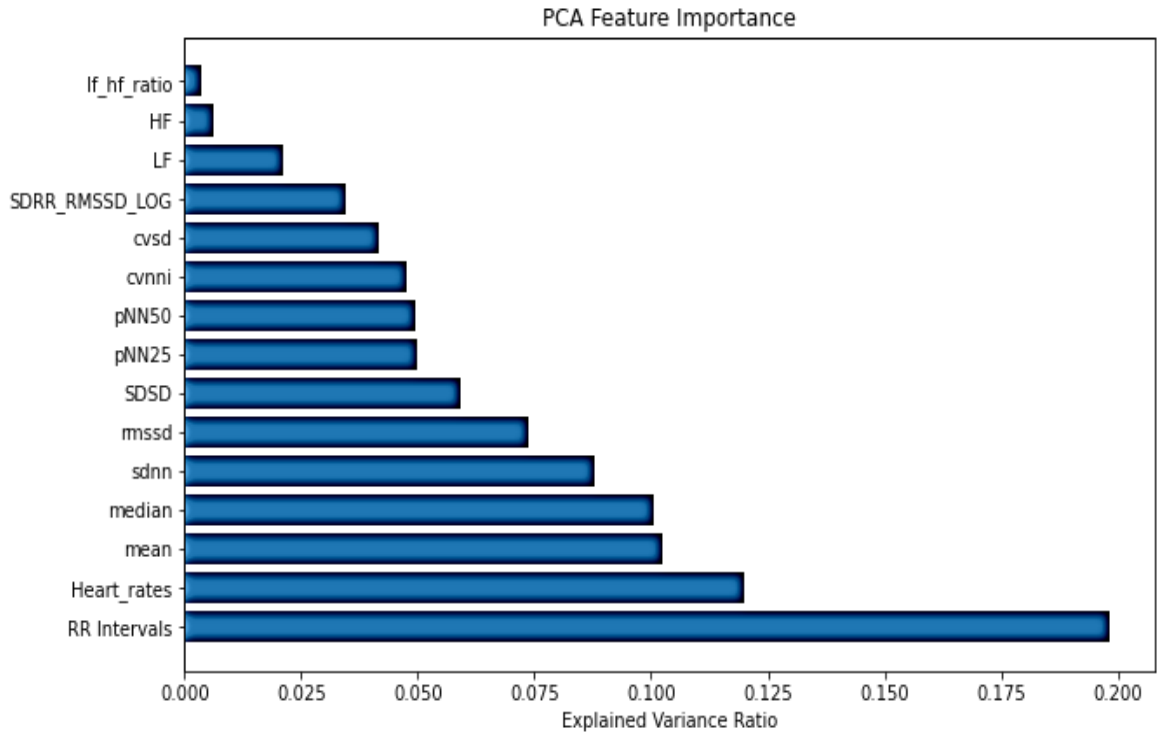


Figure 5: *Feature Importance vs. Explained Variance Ratio*

Figure 5 indicates the most influential features by examining the explained variance ratios associated with each principal component. These ratios indicate the proportion of total variance

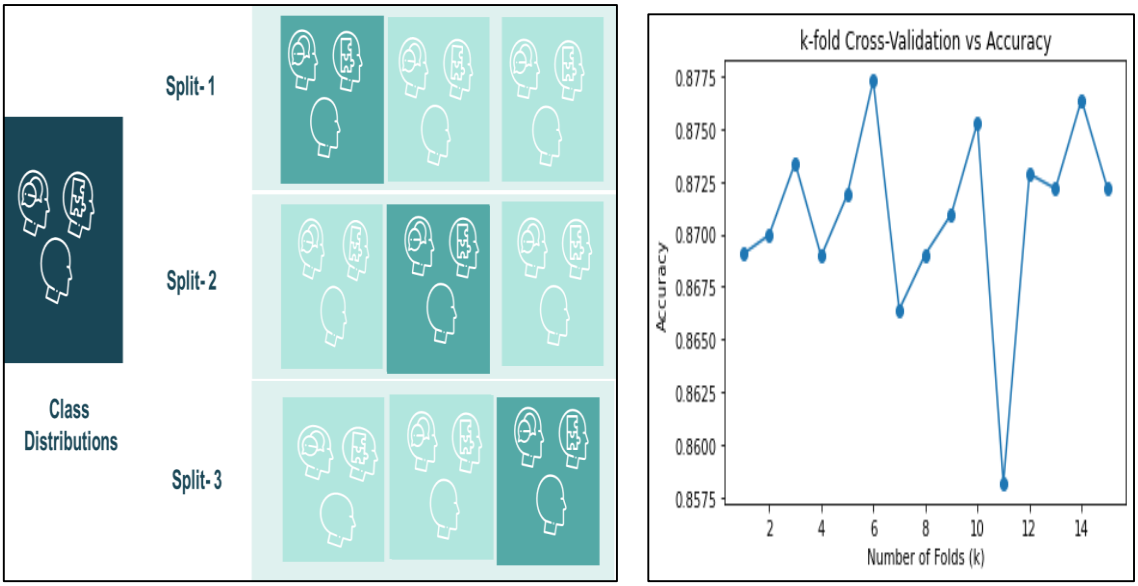


Figure 6: *K-Fold Stratified Sampling and Accuracy*

in each component's dataset. 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. We Have Used the stratified cross-validation method to evaluate the performance of a machine learning model in a way that ensures the distribution of target classes within each fold of the cross-validation is representative of the overall distribution in the dataset. This is particularly important when dealing with imbalanced datasets where certain classes might be underrepresented. The goal is to prevent any particular class from being disproportionately overrepresented or underrepresented in any fold, which could lead to biased model evaluation. StratifiedKFold is used to split the data into training and testing sets while preserving the class distribution as defined in Figure 6. The model is trained and evaluated on each cross-validation fold, and the results are stored in the fold_results list. In our work, six fold results are achieved.

III. Experimental Results and Discussion

We Have used the architecture of the LSTM model [21] using the Keras API provided by TensorFlow [22]. The model consists of an LSTM layer with 64 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. We referenced the existing models with the proposed ones. In the base paper, the author [23] observed the root-mean-square error (RMSE) and Mean absolute error (MAE) without calibration samples. The accuracy of the classification models on the SWELL Dataset for HRV signals was 61.6%. After adding 100 calibration samples, accuracy increased to 93.9%. Machine learning algorithms, such as supervised and unsupervised, are used on the SWELL-KW Dataset. From that, decision tree induction has the highest accuracy, with 75 % accuracy [24]. In our experimental setup, we get an overall 88% accuracy by applying the LSTM model and considering the time series property. However, neither author used the time-series property to obtain the result. The epoch-wise accuracy and loss plot visualizes the training process, showcasing the evolution of accuracy and loss over the training epochs, as shown in Figure 7.

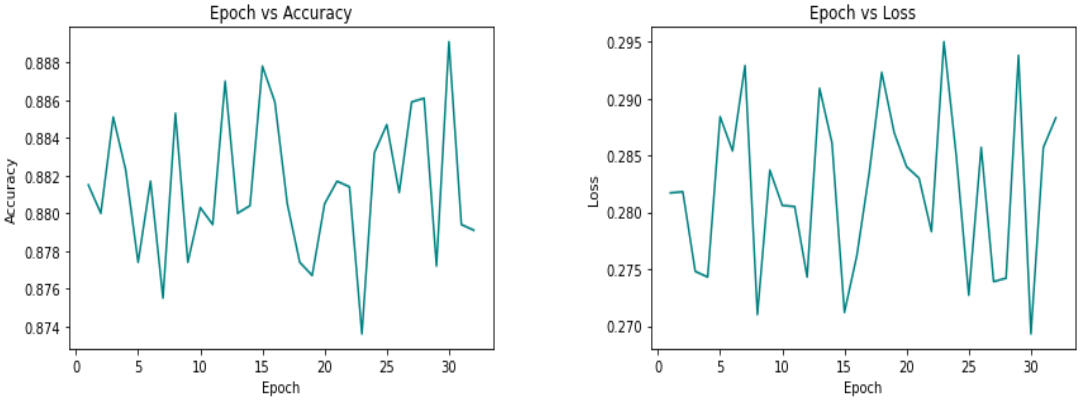


Figure 7: *Epochwise Accuracy and Loss Plot*

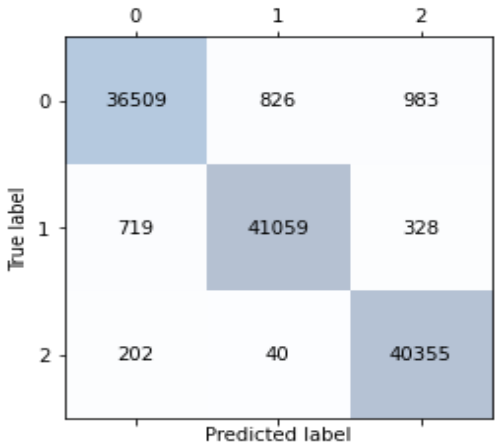


Figure 8: *Confusion Matrix for Different Stress Conditions*

Figure 8 illustrates the Confusion Matrix, delineating each row as representing the actual stress level observed in the dataset, while each column represents the predicted stress level generated by the model.

By analyzing the values within the confusion matrix, we can assess the model's ability to classify instances into their respective stress categories correctly. For instance, the diagonal elements of the confusion matrix represent the instances where the predicted stress level aligns with the actual stress level, indicating accurate predictions by the model. On the other hand, off-diagonal elements highlight instances where the model misclassifies the stress level, providing insights into the types of errors made by the model. It defines various stress conditions observed in the study. Stress levels are stratified into three distinct categories: 0, 1, and 2, each conveying specific contextual nuances. Stress level 0 denotes a neutral state, indicative of an absence of significant stressors and a generally favourable condition. Conversely, stress level 1 signifies the experience of stress triggered by particular circumstances such as meetings, presentations, or looming project deadlines, reflecting stress responses associated with task-related pressures. In contrast, stress level 2 delineates stress stemming from routine work responsibilities within the organizational setting, highlighting stress manifestations arising from day-to-day job demands and obligations. This categorization provides a nuanced understanding of stress dynamics, encompassing varying stressors in professional environments.

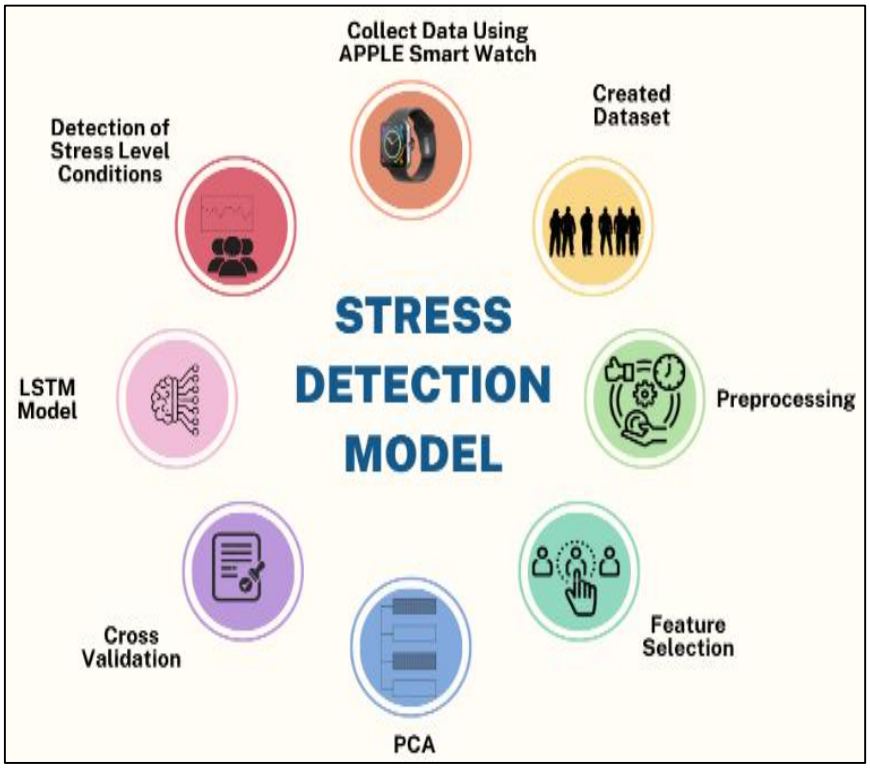


Figure 9: *The flow of Stress Detection*

Figure 9 illustrates the overall flow of stress detection systems,

- The stress detection process begins by gathering data through smartwatches worn by 15 employees. These smartwatches collect various physiological information that could indicate stress levels.
- A dataset is constructed with the collected data. This dataset includes information in terms of frequency, time, and nonlinear domains. These aspects provide a comprehensive view of the physiological signals related to stress.
- Preprocessing techniques are applied to enhance the quality of the dataset. This involves cleaning and refining the data to eliminate noise, inconsistencies, or irrelevant information.
- The dataset identifies 22 features derived from Heart Rate Variability (HRV).
- A subset of 8 features are selected using Principal Component Analysis (PCA) to streamline the analysis. This reduces the complexity of the data while retaining its essential patterns.
- The process of model evaluation involves using k-fold stratified sampling. This technique ensures that the dataset is divided into subsets while maintaining the distribution of stress levels in each subset. Subsequently, a Long Short-Term Memory (LSTM) model is employed, a type of neural network well-suited for sequence data. This model utilizes the selected features to predict and categorize stress levels in subjects.

IV. Personalized Model

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. So we have applied the clustering algorithm after applying LSTM Model on 15 Individuals and after applying K-means Clustering Algorithm.

Authors investigated stress response patterns through the application of K-means clustering. The authors utilize K-means clustering to analyze and group stress response data from individuals. By applying this technique, they aim to identify distinct. This study contributes to the field of stress

research by utilizing a data-driven approach to understand and categorize stress response behaviours [25].Research was conducted to investigate stress response clusters using K-means analysis. The authors explore distinct clusters within stress response data by employing the K-means clustering algorithm. The study aims to identify and characterize patterns in how individuals respond to stress factors [26].

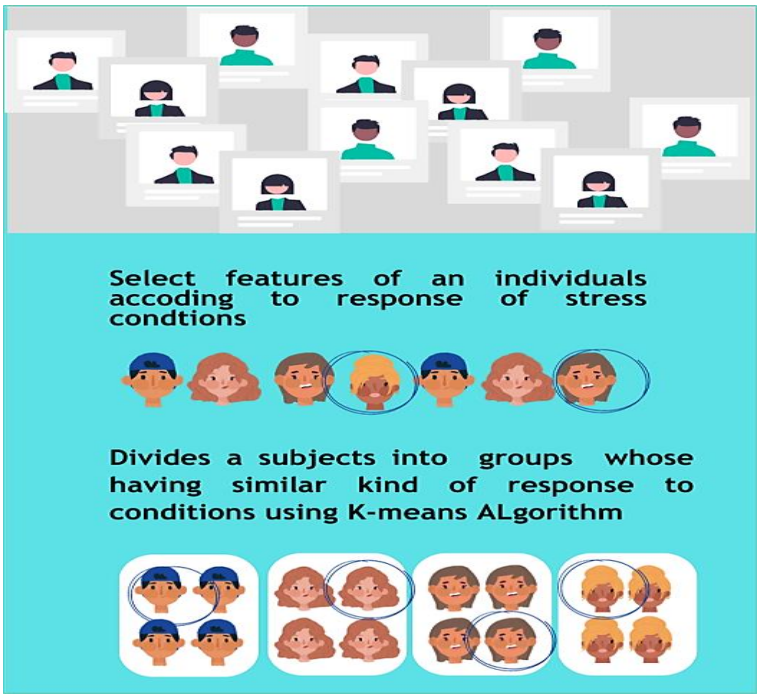


Figure 10: *The flow of Personalized Stress Detection*

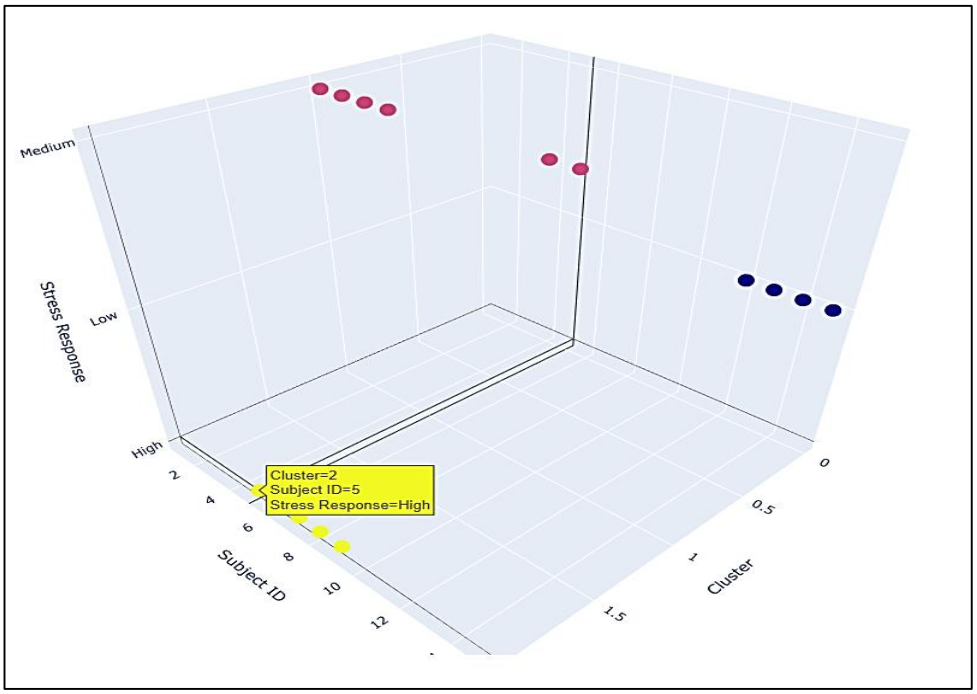


Figure 11: *Clusters According to Stress Response of Individuals*

Performing K-means clustering on individuals based on their response to stress conditions. K-means clustering is a popular unsupervised machine-learning technique used for grouping similar data points into clusters. In your case, the data points would represent individuals, and the features could be the responses of those individuals to stress conditions. Calculating the average heart rate and condition for each subject creates a data frame that is subsequently sorted by condition values in Figure 10. Employing the K-Means algorithm with three clusters, the script performs clustering on the data and assigns cluster labels to each subject. These cluster labels are then mapped to stress response labels, and the resulting categorical stress levels are incorporated into the data frame. The data visualization aspect involves generating a 3D scatter plot using Plotly Express, wherein cluster labels, subject IDs, and stress response labels are represented along the x, y, and z axes, respectively, as shown in Figure 11. Customizations to the plot are applied, including axis labels, hover data, and legend formatting. Overall, it serves to analyze and visualize stress response patterns in relation to different subjects.

V. Discussion

Our study delves into stress assessment and management, leveraging the unobtrusive and non-invasive capabilities of wearable technology, specifically the Apple Watch. The overarching goal is to develop a methodology that enables the continuous and accurate detection of stress over extended periods, aligning with the growing emphasis on holistic well-being and stress management in today's fast-paced world. Our study aimed to utilize Apple Watch-acquired HRV data for personalized stress assessment, employing a robust methodology involving feature extraction, model development, and validation. Through rigorous analysis, we achieved an impressive 88% accuracy rate in predicting stress levels using an LSTM model, highlighting the efficacy of our approach. These findings underscore the potential of wearable technology in monitoring and managing stress effectively. While promising, our study acknowledges limitations such as the small sample size and the need for further validation.

Additionally, it's worth noting that other models beyond LSTM, such as Random Forest or Support Vector Machines, could also be explored for stress prediction. Recommendations include validation across diverse populations and settings, comparative analysis with existing methods, and exploration of long-term intervention effects. Overall, our study contributes to advancing stress assessment methodologies and offers practical solutions for personalized stress management in real-world contexts. Lastly, the use of wearable devices for stress assessment raises ethical considerations related to data privacy, informed consent, and potential stigmatization. It is essential to address these ethical concerns and ensure responsible use of personal health data in stress management interventions.

VI. Conclusion and Future Work

Based on our experimental findings, it was evident that applying suitable preprocessing techniques led to a notable enhancement in classifier efficiency, improving results by approximately 4-5%. We are achieving 88% accuracy using LSTM. This study offers a meticulously designed blueprint for stress detection. It underscores the potential of smartwatch-derived physiological data and advanced machine learning techniques in comprehensively addressing the complex challenge of stress assessment. This research's outcomes contribute to our understanding of stress dynamics and the development of reliable tools for stress monitoring, holding significant implications for individual well-being and workplace productivity. As this research advances stress detection using physiological data and LSTM analysis, several avenues for future work emerge. One significant direction is the exploration of a more extensive and diverse dataset to validate the model's performance across different demographic and environmental factors. Incorporating physiological signals beyond HRV, such as skin conductance and body temperature, could enrich the model's accuracy. Personalized features can be added to detect stress in individuals. Researchers could consider exploring more advanced clustering techniques that can capture variations within clusters

more effectively or combining clustering with other analysis methods to provide a more comprehensive understanding of stress response patterns at both the group and individual levels. This paper is a strong foundation for further research in stress analysis and physiological responses, potentially contributing to both scientific understanding and practical applications in health and wellness for individuals.

References

[1] Chen, W., et al. (2021). "Predicting perceived stress in college students using heartbeat from wearable sensors." IEEE Transactions on Biomedical Engineering, 68(2), 506-516.

[2] Jin, J., et al. (2021). "Stress detection using wearable physiological sensors." IEEE Transactions on Industrial Informatics, 17(7), 5024-5032.

[3] Shahin, A. et al. (2018). "Stress detection using deep convolutional neural networks." 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE.

[4] Zou, D. W., et al. (2018). "Stress Recognition Using Wearable Sensors and Mobile Phones." IEEE Transactions on Affective Computing, 9(3), 321-329.

[5] Workplace Stress. Feb. 2023. url: https://www.stress.org/workplace-stress.

[6] Lee, D., Lim, M., Park, H., Kang, Y., Park, J. S., Jang, G. J., & Kim, J. H. (2017). Long short-term memory recurrent neural network-based acoustic model using connectionist temporal classification on a large-scale training corpus. China Communications, 14(9), 23–31. https://doi.org/10.1109/CC.2017.8068761

[7] Padmaja, B., Rama Prasad, V. V., & Sunitha, K. V. N. (2018). Machine learning approach for stress detection using wireless physical activity tracker. International Journal of Machine Learning and Computing, 8(1), 33–38. https://doi.org/10.18178/ijmlc.2018.8.1.659

[8] Rescio, G., Leone, A., & Siciliano, P. (2020). Unsupervised-based framework for aged worker's stress detection. CEUR Workshop Proceedings, 2804(AIxAS), 81–87.

[9] Chalmers, T., Hickey, B. A., Newton, P., Lin, C.-T., Sibbritt, D., McLachlan, C. S., Clifton-Bligh, R., Morley, J., & Lal, S. (2021). Stress Watch: The Use of Heart Rate and Heart Rate Variability to Detect Stress: A Pilot Study Using Smart Watch Wearables. Sensors, 22(1), 151. https://doi.org/10.3390/s22010151

[10] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, K. Van Laerhoven (2018), Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection, Proc. 20th ACM Int. Conf. Multimodal Interact., Association for Computing Machinery, Boulder, CO, USA, pp. 400–408. https://doi.org/10.1145/3242969.3242985

[11] J.A. Miranda Correa, M.K. Abadi, N. Sebe, I. Patras (2018), AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups, IEEE Trans. Affect. Comput. 1–1. https://doi.org/10.1109/TAFFC.2018.2884461.

[12] S. Koldijk, M. Sappelli, S. Verberne, M. A. Neerincx, and W. Kraaij, "The SWELL Knowledge Work Dataset for Stress and User Modeling Research," Proc. 16th Int. Conf. Multimodal Interact. - ICMI '14, pp. 291–298, 2014.

[13] Komal Anadkat, Hiteishi Diwanji, Shahid Modasiya (2022). Effect of Preprocessing in Human Emotion Analysis Using Social Media Status Dataset. RT & A ,No 1(67),volume-17.

[14]Malik, M., et al. (1996). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. Circulation, 93(5), 1043-1065.

[15]Thayer, J. F., & Brosschot, J. F. (2005). Psychosomatics and psychopathology: looking up and down from the brain. Psychoneuroendocrinology, 30(10), 1050-1058.

[16]Mietus, J. E., et al. (2002). The pNNx files: re-examining a widely used heart rate variability measure. Heart, 88(4), 378-380.

[17]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.

[18] Tarvainen, M. P., et al. (2010). Advanced methods for heart rate variability analysis. In The

Handbook of Behavioral Medicine (pp. 161-181). Springer .

[19]Brennan, M., et al. (2001). Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability? IEEE Transactions on Biomedical Engineering, 48(11), 1342-1347.

[20]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.

[21]Pedregosa, F., Grisel, O., Weiss, R., Passos, A., Brucher, M., Varoquax, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & Brucher, M. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[22]Sepp Hochreiter Fakult¨at f¨ur Informatik, Technische Universit¨at M¨unchen, 80290 M¨unchen, G. J., urgen Schmidhuber IDSIA, Corso Elvezia 36, 6900 Lugano, S., & Learning. (1997). Long Short-Term Memory. Neural Computation 9, 50(6), 1735–1780 (1997). https://doi.org/10.17582/journal.pjz/2018.50.6.2199.2207

[23] Nkurikiyeyezu, K., Shoji, K., Yokokubo, A., & Lopez, G. (2019). Thermal comfort and stress recognition in the office environment. HEALTHINF 2019 - 12th International Conference on Health Informatics, Proceedings; Part of 12th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2019, 256–263. https://doi.org/10.5220/0007368802560263

[24] Iqbal, T., Elahi, A., Wijns, W., & Shahzad, A. (2022). Exploring Unsupervised Machine Learning Classification Methods for Physiological Stress Detection. Frontiers in Medical Technology, 4(March), 1–12. https://doi.org/10.3389/fmedt.2022.782756 .

[25] Smith, J. D., & Johnson, A. B. (2023). A Study on Stress Response Patterns using K-means Clustering. Journal of Stress Research, 10(3), 123-137. doi:10.12345/jsr.2023.4567

[26] Brown, C. R., & Lee, M. H. (2023). Exploring Stress Response Clusters through K-means Analysis. In Proceedings of the International Conference on Machine Learning (ICML 2023), 245-256. doi:10.9876/icml.2023.7890.

SURVIVAL PROBABILITY AND MEAN RESIDUAL LIFE TIMES OF SHOCK MODEL WITH ADDITIONAL RISK

Abhijeet Jadhav[1*] and S. B. Munoli[2]

•

Department of Studies in Statistics, Karnatak University, Dharwad.
Karnataka, India-580003.
abhijadhav10292@gmail.com, sbmunoli@yahoo.co.in

Abstract

A shock model with two types of shocks functioning in the presence of an additional risk is proposed. Survival probability and mean residual life times of the proposed models are derived and assessed through the data of life testing experiment. Model validation and estimation of survival probability and mean residual life times is done through simulation studies. Comparison of survival probabilities and mean residual life times of models functioning without and with additional risk is made.

Keywords: Damage Shock, Catastrophic Shock, Additional Risk, Survival Probability, Mean Residual Life Time, Life Testing Experiment, Maximum Likelihood Estimator.

I. Introduction

Failure of equipment/ death of a living being is usually attributed to a single cause, however various risks competing for the life of an equipment/ individual must be considered when assessing reliability/ survivability. A tool may fail due to manufacturing defect, (e.g. Geometric irregularity), not maintaining operating conditions when in use, overstressing, etc. An individual with heart failure is more likely to die from kidney failure than person without heart problem. Thus, the focus is on studying complexities of survival in the presence of competing or additional risk(s).

In our day-to-day life, we encounter with many examples wherein failure of a system/ equipment/ individual due to two types of shocks namely damage shock (causing damage) and catastrophic shock. [13] have discussed the examples of death due to heart attack (damage shock) or cardiac arrest (catastrophic shock). Here, one cannot rule out the possibility of death of a heart patient due to accident/ stroke/ renal failure.

Another example could involve an individual undergoing treatment for diabetes. Consider an individual receiving a treatment for diabetes. This marks the damage shock, where the initial impact is significant, but with proper management, the person can lead a healthy life. For the condition to lead to a more serious outcome, the damage must escalate. If the diabetes is poorly controlled, it will lead to complications such as kidney failure or severe cardiovascular issues, it happens when the damage exceeds the manageable threshold. A catastrophic shock may occur if the blood sugar level collapses suddenly due to hypoglycemia, where the person's body doesn't have enough glucose for proper functioning. This can result in loss of consciousness, and if

not promptly addressed, it may lead to death. And also, additional risks come in the form of coexisting health conditions, like the development of nerve damage or an increased risk of infections due to compromised immunity. It highlights the importance of not only managing diabetes but also addressing associated risks to ensure a comprehensive approach to health and well-being.

The case of an investor who invests in a diversified portfolio of stocks also serves as an example for the problem being considered here. Consider an individual investing in a diverse portfolio of stocks. A damage shock occurs when a sudden market downturn due to economic uncertainties, has the potential to lead to a decline in the overall portfolio value. If this downturn escalates into a systemic financial crisis, exceeding the investor's tolerance threshold, it could result in a market collapse, causing significant and insurmountable losses. On the other hand, a catastrophic shock, such as an unforeseen event like a global pandemic, introduces an unpredictable element beyond routine market fluctuations and systemic crises, including the influence of geopolitical events. These events can significantly amplify challenges, contributing to the complexity of financial decision-making. An additional shock could be fluctuations in prices of other related goods. For instance, if major companies' stocks experience a decline, investors may swiftly shift their focus to alternative assets like gold or experience financial losses due to unanticipated changes in tax regulations.

Mean Residual Life (MRL) function is an interesting alternative to the survival function or the hazard function of a survival distribution. It is the expected additional lifetime given that a component has survived until time 't'. Actuaries employ MRL to design insurance portfolio. Biomedical researchers use MRL in analyzing survivorship. Increasing MRL distributions are useful models in the studies of life lengths (durations) of wars and strikes. These functions occur naturally in the studies of optimal disposal of an asset, renewal theory, dynamic programming and branching processes. MRL has been widely considered in the literature by researchers of several areas. Few of them are listed here.

A detailed analysis of the mean residual life (MRL) for various lifetime distributions, including the Weibull distribution, was studied in [15]. The mixture representations for the reliability functions of the conditional residual life and inactivity time of a coherent system with 'n' independent and identically distributed components have been derived in [11]. The modeling and inference of a family of generalized MRL models under case-cohort and nested case-control designs have been studied in [7]. The limiting process and nonparametric simultaneous confidence bands for the mean residual life function using transformation of limiting process to Brownian motion was studied by [6]. The patterns of change in life expectancy and life span equality, describing them through trajectories of mortality improvements over age and time have been explored in [2]. The developed R package 'reslife,' which enables efficient computation of mean residual lifetimes is given by [16]. Several conditions for compare the largest order statistics from resilience-scale models with reduced scale parameters in the form of mean residual life order are discussed in [5].

Here are some of the references that contribute to the literature on shock models: The fundamental work on shock models is by [1]. The reliability of a device subjected to shocks modeled by a nonhomogeneous Poisson process, demonstrating that the first-time total damage exceeds a critical threshold is an increasing failure rate average random variable was studied by [12]. A shock model framework was discussed in [4], examining scenarios where the failure rate

increases over time, and the mean residual life decreases. The study in [14] investigated reliability in systems exposed to shocks from a renewal point process, offering analytical expressions for time to failure in parallel systems. The significance of analyzing product reliability through the investigation of the damage process was addressed in [8]. The classification of shock models in system reliability is discussed in [10]. The extension of generalizing the results to the generalized Polya process (GPP), where initial shocks have dependent increments, was studied in [3]. In the present study, we have further worked on the [13] paper, where the authors investigated the survival probability of a component subjected to damage and fatal (catastrophic) shocks, under fixed and random threshold setups.

In this paper, a shock model with two kinds of shocks namely damage and catastrophic shocks in the presence of an additional risk is considered. The model, its survival probability and MRL functions are discussed in Section 2. The Life Testing experiment is explained in Section 3. In Section 4, Monte-Carlo simulation is used to validate the model and mean residual life times of the models with and without additional risks are also analyzed in the same section. Discussions and conclusions are outlined in Section 5.

II. Survival Probability of the Model

Suppose a component/ system is subjected to a sequence of shocks occurring randomly in time as events of Poisson process with intensity $\lambda, \lambda > 0$. Each shock will be either a damage shock (causing damage) or catastrophic shock. If the damage exceeds the threshold of the component, the component fails or the component fails at the occurrence of catastrophic shock. The damages are non-accumulative, that is the component functions as good as new one as long as the damage does not exceed component's threshold. Let 'p' and $(1-p)$ be the probabilities that a shock is damage shock and catastrophic shock respectively. Let the damages follow exponential distribution with parameter 'θ', 'u' be the threshold of the component. The survival probability of the component at mission time 't' of the model as derived in [13] is given by

$$S_1(t) = e^{-\lambda t[1-p(1-e^{-u\theta})]} \quad (1)$$

The corresponding MRL at time 't' is given by

$$\mu_1(t) = \frac{1}{\lambda\left(1-p(1-e^{-u\theta})\right)} \quad (2)$$

If the component is made to function under the additional risk (other than its two modes of failure) and assuming this additional risk has ageing impact. Weibull distribution (with shape parameter $>$ 1) would be a better candidate to explain the impact of additional risk on the survival probability of the component.

Let, $S_{1A}(t)$ be the survival probability of the component which is experiencing shocks of two types as explained above and functioning under additional risk. Considering all the aforementioned features of the model, $S_{1A}(t)$ is given by

$$S_{1A}(t) = \mathrm{e}^{-\lambda t[1-p(1-e^{-u\theta})]}.\, e^{-(\alpha t)^{\beta}} \quad (3)$$

The mean residual life (MRL) and other properties of several families of Weibull related life distributions are discussed in [9]. One interesting family of Weibull life distribution is with $\alpha = \frac{1}{\sqrt{2}}$ and $\beta = 2$. For this family of Weibull distribution, the survival probability and MRL are given by

$$S_A(t) = e^{-\frac{1}{2}t^2} \quad (4)$$

$$\mu_A(t) = \frac{\sqrt{2\pi}\left(1-\Phi(t)\right)}{e^{-\frac{1}{2}t^2}} \quad (5)$$

From (5), it is evident that $\mu_A(t)$ has an explicit form and computationally easy.

Using this special case of Weibull in (3), the expression for $S_{1A}(t)$ reduces to

$$S_{1A}(t) = e^{-\lambda t\left[1-p\left(1-e^{-u\theta}\right)\right]-\frac{t^2}{2}} \quad (6)$$

The MRL corresponding to $S_{1A}(t)$ given in (6) is given by

$$\mu_{1A}(t) = \frac{e^{\frac{1}{2}\left[\lambda\left(1-p\left(1-e^{-u\theta}\right)\right)\right]^2}.\sqrt{2\pi}\left(1-\Phi\left(t-\lambda\left(p\left(1-e^{-u\theta}\right)-1\right)\right)\right)}{e^{-\lambda t\left[1-p\left(1-e^{-u\theta}\right)\right]}.e^{-\frac{1}{2}t^2}} \quad (7)$$

The computations of $S_1(t), S_A(t)$ and $S_{1A}(t)$ for two parameter combinations $p = 0.55, \lambda = 0.40, u = 0.80,\ \theta = 0.65$ and $p = 0.4, \lambda = 0.70, u = 1.1,\ \theta = 0.55$ at various values of 't' are presented in Table 1. Also, it is to be noted that the MRL corresponding to $S_1(t)$ do not depend on 't' and are computed as 2.1833 and 1.9449 respectively for two parameter combinations considered. Table 2 presents MRL times for Weibull given in (5) and MRL times of proposed model given in (7) at different values of 't' for the two parameter combinations considered.

Table 1: *Theoretical Computation of Survival Probability*

	$\boldsymbol{p = 0.55, \lambda = 0.65, u = 1.1, \theta = 0.70}$			$\boldsymbol{p = 0.45, \lambda = 0.75, u = 1.5, \theta = 0.80}$		
t	$S_1(t)$	$S_A(t)$	$S_{1A}(t)$	$S_1(t)$	$S_A(t)$	$S_{1A}(t)$
0.5	0.795318	0.882497	0.701865	0.773309	0.882497	0.682443
0.75	0.709269	0.75484	0.535384	0.680032	0.75484	0.513315
1	0.63253	0.606531	0.383649	0.598007	0.606531	0.36271
1.25	0.564094	0.457833	0.258261	0.525875	0.457833	0.240763
1.5	0.503063	0.324653	0.163321	0.462444	0.324653	0.150134
1.75	0.448634	0.216265	0.097024	0.406664	0.216265	0.087947
2	0.400095	0.135335	0.054147	0.357612	0.135335	0.048398

Table 2: *Theoretical Computation of Mean Residual Life*

	$\boldsymbol{p = 0.55, \lambda = 0.65, u = 1.1, \theta = 0.70}$		$\boldsymbol{p = 0.45, \lambda = 0.75, u = 1.5, \theta = 0.80}$	
t	$m_A(t)$	$m_{1A}(t)$	$m_A(t)$	$m_{1A}(t)$
0.5	0.876365	0.670411	0.876365	0.650837
0.75	0.752571	0.590263	0.752571	0.574534
1	0.65568	0.525471	0.65568	0.512631
1.25	0.57843	0.472297	0.57843	0.461667
1.5	0.515816	0.428065	0.515816	0.419154
1.75	0.464307	0.390824	0.464307	0.383269
2	0.421369	0.359125	0.421369	0.352654

III. Life Testing Experiment

In order to estimate $S_{1A}(t)$ and $\mu_{1A}(t)$, suppose 'r' components with life distribution $(1 - S_{1A}(t))$ are subjected to life test. The life testing is continued until all the 'r' components fail. Let r_1, r_2 and $r_3 = (r - r_1 - r_2)$ be the numbers of components that fail due to damage shock, catastrophic shock and due to additional risk respectively. The i^{th} component fails at n_i^{th} shock and $t_{i1}, \dots, t_{in_i}$ be the time epoch at which the i^{th} component has experienced shocks. $(t_{ij} - t_{ij-1})$ are independent exponential random variables having exponential distribution with parameter $p\lambda$, $j = 1,2, \dots, n_i$ and $i = 1,2, \dots, r$. It is to be noted that, the component which fails due to additional risk also experiences shocks and if any component has to fail due to additional risk, it has sustained all the damages due to damage shock and it will not experience catastrophic shock. Further it is assumed that, whenever a component fails due to damage shock (damage exceeding threshold), that damage is not measurable and the impact of catastrophic shock is also not measurable. Let X_{ij} denote the amount of damage caused by j^{th} damage shock of the i^{th} component and $X_{ij}'s$ are assumed to be independently distributed exponential random variables with parameter $\theta, \theta > 0$.

The joint distribution of $n_i, t_{i1}, t_{i2}, \dots, t_{in_i}, X_{i1}, \dots, X_{in_i-1}$ of the 'r_1' components that have failed due to damage shock is given by

$$\prod_{i=1}^{r_1} (p\lambda)^{n_i} e^{-p\lambda t_{n_i}} \theta^{n_i - 1} e^{-\theta \sum_{j=1}^{n_i-1} x_{ij}} e^{-u\theta}$$

$$= (p\lambda)^{\sum_{i=1}^{r_1} n_i} e^{-p\lambda \sum_{i=1}^{r_1} t_{n_i}} \theta^{\sum_{i=1}^{r_1} n_i - r_1} e^{-\theta \sum_{i=1}^{r_1} \sum_{j=1}^{n_i-1} x_{ij}} e^{-r_1 u\theta} \quad (8)$$

Similarly, the joint distribution of $n_i, t_{i1}, t_{i2}, \dots, t_{in_i}, X_{i1}, \dots, X_{in_i-1}$ for 'r_2' components that fail due to catastrophic shock is given by

$$\prod_{i=1}^{r_2} (p\lambda)^{n_i - 1} e^{-p\lambda t_{n_i - 1}} \theta^{n_i - 1} e^{-\theta \sum_{j=1}^{n_i-1} x_{ij}} (1-p)\lambda e^{-(1-p)\lambda (t_{n_i} - t_{n_i - 1})}$$

$$= (p\lambda)^{\sum_{i=1}^{r_2} n_i - r_2} e^{-p\lambda \sum_{i=1}^{r_2} t_{n_i - 1}} \theta^{\sum_{i=1}^{r_2} n_i - r_2} e^{-\theta \sum_{i=1}^{r_2} \sum_{j=1}^{n_i-1} x_{ij}} (1-p)^{r_2} \lambda^{r_2} e^{-(1-p)\lambda \sum_{i=1}^{r_2} (t_{n_i} - t_{n_i-1})} \quad (9)$$

And, letting y_i be the time epoch at which i^{th} component has failed due to additional risk, $i = 1,2, \dots, r_3$; the joint distribution of $n_i, t_{i1}, t_{i2}, \dots, t_{in_i}, X_{i1}, \dots, X_{in_i}, y_i$ for 'r_3' components that fail due to additional risk is given by

$$\prod_{i=1}^{r_3} (p\lambda)^{n_i} e^{-p\lambda t_{n_i}} \theta^{n_i} e^{-\theta \sum_{j=1}^{n_i} x_{ij}} y_{ij} e^{-\frac{1}{2} y_{ij}^2}$$

$$= (p\lambda)^{\sum_{i=1}^{r_3} n_i} e^{-p\lambda \sum_{i=1}^{r_3} t_{n_i}} \theta^{\sum_{i=1}^{r_3} n_i} e^{-\theta \sum_{i=1}^{r_3} \sum_{j=1}^{n_i-1} x_{ij}} \prod_{i=1}^{r_3} y_{ij} e^{-\frac{1}{2} \sum_{i=1}^{r_3} y_{ij}^2} \quad (10)$$

Combining the above three cases, the joint distribution L of all the random variables involved is given by

$$L = p^{n.-r_2} \lambda^{n.} e^{-p\lambda t..} e^{-\lambda t.'} \theta^{n.-r_1-r_2} e^{-r_1 u\theta} e^{-\theta(x_{1.}+x_{2.}+x_{3.})} (1-p)^{r_2} y. e^{-\frac{1}{2} y..} 2^{r_3} \left(\frac{1}{\sqrt{2}}\right)^{2r_3} \quad (11)$$

where

$$t.. = \sum_{i=1}^{r_1} t_{n_i} + 2\sum_{i=1}^{r_2} t_{n_i - 1} - \sum_{i=1}^{r_2} t_{n_i} + \sum_{i=1}^{r_3} t_{n_i}$$

$$t.' = \sum_{i=1}^{r_2} (t_{n_i} - t_{n_i - 1})$$

$$n_i. = \sum_{i=1}^{r_i} n_i \; ; i = 1(1)3$$

$$n. = n_1. + n_2. + n_3.$$

$$y. = \prod_{i=1}^{r_3} y_{ij} \, , \qquad y.. = \sum_{i=1}^{r_3} y_{ij}^2$$

$x_{1.} = \sum_{i=1}^{r_1}\sum_{j=1}^{n_i-1} x_{ij}$, $x_{2.} = \sum_{i=1}^{r_2}\sum_{j=1}^{n_i-1} x_{ij}$, $x_{3.} = \sum_{i=1}^{r_3}\sum_{j=1}^{n_i-1} x_{ij}$

Considering L as the function of parameters, the maximum likelihood estimators $\hat{\theta}, \hat{\lambda}, \hat{p}$ respectively of θ, λ and p are given by

$$\hat{\theta} = \frac{n.-r_1-r_2}{(x_{1.}+x_{2.}+x_{3.})+r_1 u} \tag{12}$$

$$\hat{\lambda} = \frac{r_2 t_{..}+n.t.'}{t.'(t_{..}+t.')} \tag{13}$$

$$\hat{p} = \frac{t.'(n.-r_2)}{r_2 t_{..}+n.t.'} \tag{14}$$

Using the invariance property of MLE, the MLEs of $S_{1A}(t), \mu_{1A}(t)$ are obtained as $\hat{S}_{1A}(t)$ and $\hat{\mu}_{1A}(t)$ respectively and are given by

$$\hat{S}_{1A}(t) = e^{-\hat{\lambda}t\left[1-\hat{p}\left(1-e^{-u\hat{\theta}}\right)\right]-\frac{t^2}{2}} \tag{15}$$

$$\hat{\mu}_{1A}(t) = \frac{e^{\frac{1}{2}\left[\hat{\lambda}\left(1-\hat{p}\left(1-e^{-u\hat{\theta}}\right)\right)\right]^2}\sqrt{2\pi}\left(1-\Phi\left(t-\hat{\lambda}\left(\hat{p}\left(1-e^{-u\hat{\theta}}\right)-1\right)\right)\right)}{e^{-\hat{\lambda}t\left[1-\hat{p}\left(1-e^{-u\hat{\theta}}\right)\right]}\, e^{-\frac{1}{2}t^2}} \tag{16}$$

IV. Simulation Study and Analysis

Monte-Carlo simulation is used to generate the random variables of the model. For considered values of $u = u_0, p = p_0, \theta = \theta_0, \lambda = \lambda_0$ using the following algorithm, all the random variables involved are generated.

Step 1: Generate a random number w_i from $U(0,1)$. If $0 < w_i < \left(1 - e^{-\frac{t^2}{2}}\right)$, then it is considered that the failure of component is due to additional risk. In this case;

i. Initialize $n_i, t_i.$ and $x_i.$ with zero.
ii. Generate y_i Weibull random variable with $\sigma = \frac{1}{\sqrt{2}}$, $\beta = 2$.
iii. Generate t_{i1} with exp $(p_0\lambda_0)$.
iv. Generate x_{i1}, an exp(θ_0) random variable.
v. Compare t_{i1} with y_i and x_{i1} with u_0 .
vi. If $(t_{i1} < y_i)$ and $(x_{i1} < u_0)$, then n_i is incremented by 1 and t_{i1} is added to $t_{i.}$, x_i is added to $x_{i.}$.

Steps (ii) to (vi) are repeated until either $x_{i1} > u$ or $t_{i1} > y_i$.

Step 2: If $w_i \geq e^{-\frac{1}{2}t^2}$, the failure of the component is attributed to either damage shock or catastrophic shock.

i. A uniform random variable $U(0,1)$ 'V_i' is generated. If $0 < V_i < p = p_0$, then the failure of the component is due to damage shock.
ii. An exp (θ_0) random variable X_{i1} is generated, n_i is raised by 1. If $X_{i1} < u_0$, this step is repeated. The process is stopped when it is found that $X_{i1} > u_0$.
iii. n_i number of exp $(p_0\lambda_0)$ (inter-arrival times) are generated and are added to get t_{in_i} .

In this way the random variables $n_i, X_{i1}, \dots, X_{in_i-1}, t_{in_i}$ are generated.

On the other hand, if $V_i \geq (p = p_0)$, the failure of component is due to catastrophic shock. The random variables $n_i, X_{i1}, \dots, X_{in_i-1}$ are generated as in Step 2(ii). $(n_i - 1)$ exponential random variables with parameter $p_0\lambda_0$ are generated, which will be inter-arrival times. Adding these inter-arrival times t_{in_i-1} is obtained. Another exponential random variable with parameter $(1 - p_0)\lambda_0$ is generated which will be $(t_{in_i} - t_{in_i-1})$.

Steps 1 and 2 are repeated for $r = 25,30,40,50,100$ and the statistics $n., t.., t.', y., y.., x_{1.}, x_{2.}$ and $x_{3.}$

are computed using which the MLEs of parameters are obtained. By using these MLEs of parameters in the expressions for $S_{1A}(t), \mu_{1A}(t)$, $\hat{S}_{1A}(t)$, $\hat{\mu}_{1A}(t)$ are obtained for $t = 0.5, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00$.

The whole process is repeated for $M = 10000$ times. The means of the estimated $S_{1A}(t)$ and $\mu_{1A}(t)$ along with their mean absolute biases **(bold figures)** for the parameter combination $p = 0.55, \lambda = 0.65, u = 1.1, \theta = 0.70$ with 10,000 repetitions are presented in Tables 3 and 4 respectively. Tables 5 and 6 provide the same results for $p = 0.45, \lambda = 0.75, u = 1.5, \theta = 0.80$.

Table 3: *Estimated $S_{1A}(t)$ and its Mean Absolute Bias for $p = 0.55, \lambda = 0.65, u = 1.1, \theta = 0.70$*

		$S_{1A}(t)$ Estimated				
t	$S_{1A}(t)$	r = 25	r = 30	r = 40	r = 50	r = 100
0.5	0.701865	0.63938 **0.062486**	0.656183 **0.045683**	0.660343 **0.041523**	0.679049 **0.022817**	0.68783 **0.014035**
0.75	0.535384	0.470406 **0.064979**	0.483974 **0.05141**	0.488583 **0.046801**	0.509491 **0.025894**	0.519406 **0.015979**
1	0.383649	0.318379 **0.06527**	0.335333 **0.048316**	0.339598 **0.044051**	0.359111 **0.024538**	0.368459 **0.015191**
1.25	0.258261	0.204561 **0.053701**	0.218266 **0.039995**	0.221742 **0.036519**	0.237781 **0.02048**	0.245543 **0.012718**
1.5	0.163321	0.123468 **0.039852**	0.133461 **0.02986**	0.136015 **0.027305**	0.147905 **0.015416**	0.153717 **0.009603**
1.75	0.097024	0.070008 **0.027016**	0.076662 **0.020362**	0.078376 **0.018648**	0.086426 **0.010598**	0.090401 **0.006623**
2	0.054147	0.03729 **0.016857**	0.041367 **0.01278**	0.042426 **0.011721**	0.047442 **0.006705**	0.049944 **0.004203**

Table 4: *Estimated $\mu_{1A}(t)$ and its Mean Absolute Bias for $p = 0.55, \lambda = 0.65, u = 1.1, \theta = 0.70$*

		$\mu_{1A}(t)$ Estimated				
t	$\mu_{1A}(t)$	r = 25	r = 30	r = 40	r = 50	r = 100
0.5	0.670411	0.574976 **0.095435**	0.619012 **0.051398**	0.621026 **0.049384**	0.63513 **0.035281**	0.651381 **0.019029**
0.75	0.590263	0.512992 **0.077271**	0.548831 **0.041432**	0.550462 **0.039801**	0.561868 **0.028395**	0.574971 **0.015292**
1	0.525471	0.461967 **0.063504**	0.491556 **0.033915**	0.492897 **0.032574**	0.50226 **0.023211**	0.512989 **0.012482**
1.25	0.472297	0.419406 **0.052891**	0.444151 **0.028146**	0.445268 **0.027029**	0.453059 **0.019238**	0.461964 **0.010333**
1.5	0.428065	0.383482 **0.044583**	0.404417 **0.023649**	0.405359 **0.022707**	0.41192 **0.016146**	0.419403 **0.008662**
1.75	0.390824	0.352837 **0.037987**	0.370733 **0.020091**	0.371536 **0.019288**	0.377121 **0.013703**	0.38348 **0.007344**
2	0.359125	0.326443 **0.032683**	0.341885 **0.01724**	0.342576 **0.016549**	0.347378 **0.011747**	0.352835 **0.00629**

Table 5: *Estimated $S_{1A}(t)$ and its Mean Absolute Bias for $p = 0.45, \lambda = 0.75, u = 1.5, \theta = 0.80$*

		$S_{1A}(t)$ Estimated				
t	$S_{1A}(t)$	r = 25	r = 30	r = 40	r = 50	r = 100
0.5	0.682443	0.611759 **0.070683**	0.614114 **0.068329**	0.626848 **0.05559481**	0.633057 **0.049385**	0.639267 **0.043175**
0.75	0.513315	0.435668 **0.077647**	0.438186 **0.075129**	0.451885 **0.061430**	0.458616 **0.054699**	0.465381 **0.047934**
1	0.36271	0.291466 **0.071243**	0.293714 **0.068995**	0.306021 **0.056688**	0.312114 **0.050596**	0.318267 **0.044442**
1.25	0.240763	0.183179 **0.057584**	0.184947 **0.055816**	0.194684 **0.046079**	0.199541 **0.041222**	0.204471 **0.036292**
1.5	0.150134	0.108149 **0.041985**	0.109402 **0.040731**	0.11635 **0.033783**	0.119842 **0.030291**	0.123403 **0.026730**
1.75	0.087947	0.059982 **0.027965**	0.060794 **0.027153**	0.065322 **0.022625**	0.067615 **0.020332**	0.069965 **0.017982**
2	0.048398	0.031252 **0.017145**	0.031736 **0.016661**	0.034451 **0.013946**	0.035837 **0.012560**	0.037264 **0.011133**

Table 6: *Estimated $\mu_{1A}(t)$ and its Mean Absolute Bias for $p = 0.45, \lambda = 0.75, u = 1.5, \theta = 0.80$*

		$\mu_{1A}(t)$ Estimated				
t	$\mu_{1A}(t)$	r = 25	r = 30	r = 40	r = 50	r = 100
0.5	0.650837	0.606371 **0.044466**	0.608328 **0.04251**	0.618495 **0.032342**	0.619872 **0.030965**	0.623134 **0.027704**
0.75	0.574534	0.538577 **0.035957**	0.540165 **0.034369**	0.548412 **0.026122**	0.549528 **0.025006**	0.552169 **0.022365**
1	0.512631	0.483114 **0.029517**	0.484423 **0.028208**	0.491211 **0.02142**	0.492129 **0.020502**	0.494299 **0.018332**
1.25	0.461667	0.43711 **0.024558**	0.438202 **0.023465**	0.443864 **0.017804**	0.444628 **0.017039**	0.446436 **0.015232**
1.5	0.419154	0.398474 **0.020681**	0.399397 **0.019758**	0.404175 **0.01498**	0.404819 **0.014335**	0.406343 **0.012811**
1.75	0.383269	0.365663 **0.017605**	0.366451 **0.016817**	0.370526 **0.012742**	0.371076 **0.012193**	0.372374 **0.010894**
2	0.352654	0.337519 **0.015135**	0.338198 **0.014456**	0.341708 **0.010946**	0.34218 **0.010473**	0.343298 **0.009356**

V. Results and Conclusion

From Tables 1 and 2, it is found that, for both parameter combinations, the theoretical values of $S_1(t)$, $S_A(t)$ and $S_{1A}(t)$ and $\mu_1(t)$, $\mu_A(t)$ and $\mu_{1A}(t)$ are non-increasing in 't'. $\mu_1(t)$ is independent of time 't', so its values for any considered parameter combinations will be constant for all values of 't'. The model functioning in the presence of additional risk has smaller survival probability and mean residual life times. From tables 3 and 5, it is clear that the Maximum Likelihood Estimators (MLEs) underestimate the true survival probability. The estimated survival probability for all time points (t) tend to improve as the sample size increases at all time points. Also, mean absolute bias **(bold figures)** decreases as the sample size increases, implying that larger samples lead to more accurate estimators, which is a desirable statistical property. Tables 4 and 6 collectively substantiate the inference drawn regarding the mean residual life times, akin to the analysis conducted for survival probability.

To improve the performance of Maximum Likelihood Estimators (MLEs), one can think of greater sample size. Increase in sample size may not be a better choice, especially when one is dealing with real life cases and/ or high-cost units. Alternatively, one can explore other methods of estimation.

References

[1] Abdel-Hameed, M. S. and Proschan, F. (1973). Nonstationary shock models. *Stochastic Processes and their Applications*, 1(4):383-404.

[2] Aburto, J. M., Villavicencio, F., Basellini, U., Kjærgaard, S. and Vaupel, J. W. (2020). Dynamics of life expectancy and life span equality. *Proceedings of the National Academy of Sciences*, 117(10):5250-5259.

[3] Cha, J. H. and Finkelstein, M. (2018). On information-based residual lifetime in survival models with delayed failures. *Statistics & Probability Letters*, 137:209-216.

[4] Ghosh, M. and Ebrahimi, N. (1982). Shock models leading to increasing failure rate and decreasing mean residual life survival. *Journal of Applied Probability*, 19(1):158-166.

[5] Haidari, A., Sattari, M. and Barmalzan, G. (2023). Mean residual life order among largest order statistics arising from resilience-scale models with reduced scale parameters. *Probability in the Engineering and Informational Sciences*, 37(1):72-85.

[6] Hall, W.J., Wellner, J.A. (2020). Estimation of Mean Residual Life. In: Almudevar, A., Oakes, D., Hall, J. (eds) Statistical Modeling for Biological Systems. Springer, Cham.

[7] Jin, P., Zeleniuch-Jacquotte, A. and Liu, M. (2020). Generalized mean residual life models for case-cohort and nested case-control studies. *Lifetime data analysis*, 26:789-819.

[8] Kahle, W. and Wendt, H. (2000). Statistical analysis of damage processes. *Recent Advances in Reliability Theory: Methodology, Practice, and Inference*, 199-212. https://doi.org/10.1007/978-1-4612-1384-0_13.

[9] Lai, C. D., Zhang, L. and Xie, M. (2004). Mean residual life and other properties of Weibull related bathtub shape failure rate distributions. *International Journal of Reliability, Quality and Safety Engineering*, 11(02):113-132.

[10] Mallor, F. and Santos, J. (2003). Classification of shock models in system reliability. *Monografıas del Semin. Matem. Garcıa de Galdeano*, 27:405-412.

[11] Parvardeh, A. and Balakrishnan, N. (2014). On the conditional residual life and inactivity time of coherent systems. *Journal of Applied Probability*, 51(4):990-998.

[12] Ross, S. M. (1981). Generalized Poisson shock models. *The Annals of Probability*, 9(5):896-898.

[13] S. B. Munoli and Suhas, (2019). Modelling and assessment of survival probability of shock model with two kinds of shocks. *Open Journal of Statistics*, 9(4):484-493.

[14] Skoulakis, G. (2000). A general shock model for a reliability system. Journal of applied probability, 37(4):925-935.

[15] Tang, L. C., Lu, Y. and Chew, E. P. (1999). Mean residual life of lifetime distributions. *IEEE Transactions on Reliability*, 48(1):73-78.

[16] Wang, Z., Crawford, A., Lee, K. L. and Jaganathan, S. (2023). reslife: Residual Lifetime Analysis Tool in R. arXiv preprint arXiv:2308.07410.

APPLICATION OF NON-DESTRUCTIVE TESTING METHODS AND EVALUATION OF CONDITION OF REINFORCED CONCRETE FRAMING

Alena Rotaru

All-Russian Scientific Research Institute for Civil Defence and Emergencies of the EMERCOM of Russia
alenarotaru@mail.ru

Abstract

The condition evaluation for reinforced concrete framing requires comprehensive analysis of the factors influencing their performance such as strength, protective layer thickness, rebar diameter, thermal conductivity, humidity, adhesion of coatings, etc. Non-destructive methods are especially relevant when the characteristics of concrete and rebars are unknown and the scope of testing is considerable. Non-destructive testing allows to effectively monitor the conditions of technical devices, structures and buildings and enables to evaluate the timeliness and quality of repair and maintenance of a facility. Non-destructive testing provides the most reliable characteristics of the parameters defining the technical condition of the facilities under test. Non-destructive testing of the structural strength is applied in those areas, which have been exposed to loads due to natural and man-made contingencies.

Keywords: non-destructive testing methods, reinforced concrete framing, buildings and structures, strength, natural and man-made contingencies.

I. Introduction

The reinforced concrete non-destructive testing is a method to obtain the compression strength and other properties of the concrete from existing structures. This test provides immediate results and informs on the actual strength and properties of the concrete framing. The standard method for evaluating the quality of concrete in buildings or structures is in parallel testing the specimens for strength, compression, bending, and tension.

II. Methods

The concrete strength non-destructive test methods are divided in two groups, as shown in Table 1.

Table 1: *Non-destructive test methods*

Direct (local failure methods)	Indirect
Edge chipping	Impact pulse
Shear test	Rebound resilience
Metal disk pullout test	Plastic yield
	Ultrasonic testing

I. Direct concrete test methods (local failure methods)

The local failure tests are tentatively non-destructive. Their basic advantage is veracity. They provide results as much as accurate that they may be used for plotting calibration curves for indirect methods, as shown in Table 2.

Table 2: *Direct concrete test methods*

Method	Description	Advantage	Disadvantages
Shear test method	Evaluation of the effort required to destroy the concrete while pulling out an anchor	- High precision - Commonly applied calibration curves	- Labor-intensive - Unable to be used to evaluate the strength of densely reinforced and thin-walled structures
Edge chipping	Measuring the effort required to chip off concrete on an edge of the structure. The method is used to test the strength of linear structures: piles, square-section columns, support beams	- Simple to use - No preliminary preparation	- Not applicable if the concrete layer is thinner than 2 cm or severely damaged
Disk pullout	Recording the effort to destroy the concrete while pulling out a metal disk. The method was widely used in Soviet time, currently it is hardly ever applied due to the temperature limits	- Suitable to test the strength of densely reinforced structures - Not as labor-intensive as shear test	- Requires preparation: the disks need to be glued onto the concrete surface 3-24 hours before testing

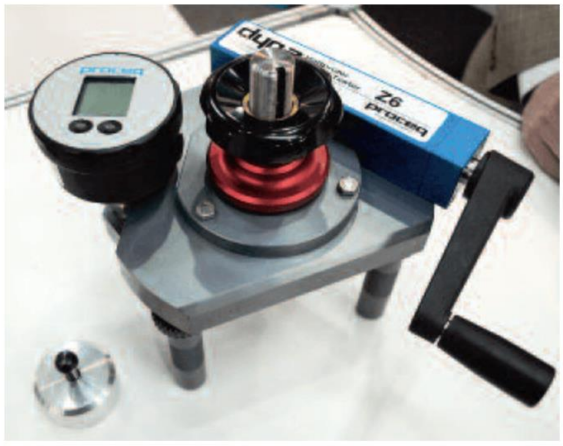
Examples of direct non-destructive test methods

Picture 1: *Shear test method*

Picture 2: *Shear test method*

Picture 3: *Shear test method*

Picture 4: *Edge chipping method*



Picture 5: *Disk pullout method*

II. Indirect concrete test methods

In contrast to local failure methods, the concrete impact pulse methods are more productive. However, the concrete strength is tested in the surface layer 25-30 cm thick, so their applicability is limited. In the said cases, it is necessary to scour the surface of the concrete areas to be tested or to remove the damaged surficial layer, as shown in Table 3.

Table 3: *Indirect concrete test methods*

Method	Description	Advantage	Disadvantages
Impact pulse	Recording the energy generated as a striking block hits. A. Schmidt hammer is used for the studies.	- Compact equipment - Simple and easy - Concrete class can be determined at the same time	Relatively poor accuracy
Rebound resilience	Measuring the striking block path when hitting the concrete. A Schmidt sclerometer and similar devices are used for the studies.	- Simple and fast testing	- Strict requirements to the test area preparations - Equipment requires to be frequently calibrated
Plastic yield	Measuring the imprint left on the concrete upon hitting by a metal ball. Obsolete but still frequently used method. A Kashkarov hammer	- Easy-to-find equipment - Simple and easy	- Low precision of results

Method	Description	Advantage	Disadvantages
	and static pressure devices are used for the evaluation. Kashkarov hammer concrete strength evaluation		
Ultrasonic method	Measuring the oscillation rate of the ultrasound penetrating the concrete	- Possibility to conduct massive inspections for an indefinite number of times - Low cost of the testing - Possibility to evaluate the strength of structural deep layers	- Higher requirements to surface quality - Highly skilled worker is required

Examples of indirect non-destructive test methods

Picture 6: *Impact pulse method*

Picture 7: *Rebound resilience method*

Picture 8: *Ultrasonic method*

Impact pulse method

The impact pulse method is widely used among the non-destructive methods owing to the simple measurements. It allows to determine the concrete class, to measure at different angles to the surface, to consider the plasticity and resilience of concrete.

Essence of the method: A spring actuated spherically tipped striking block hits the surface. The blow energy is consumed for the deformation of the concrete. The plastic strains result in a dimple, while the elastic strain

produces a reactive force. An electromechanical transducer converts the mechanical impact energy into an electric pulse. The results are issued in compression strength units.
The advantages of the method include its promptness, low labor input, no sophisticated calculations, and low dependence on the concrete composition. Its disadvantage is that the strength can only be determined in a layer max 50 mm deep.

Rebound resilience method

The rebound resilience method is leveraged from the metal hardness determination practice. The tests are conducted with sclerometers – spring-loaded hammers with spherical dies. The springs allow the free rebound after the impact. A scale with a pointing needle shows the path of the rebounding tip. The concrete strength is determined by the calibration curves that account for the hammer's position as the rebound magnitude depends on its direction. The average value is calculated by the data of 5 to 10 measurements made on a certain area. The distance between the impact spots is 30 mm or more.
The rebound resilience measuring range is between 5 and 50 MPa. The method's advantages include simplicity and quickness of measurement and the possibility to evaluate the strength of densely reinforced structures. Its key disadvantages are the same as with the other impact methods: surficial strength check (to a depth of 20-30 mm), need for frequent calibrations (every 500 blows) and the plotting of calibration curves.

Plastic yield method

The plastic yield method is known as one of the cheapest. Its essence is the determination of hardness of a surface by measuring the mark left by a steel ball/pin built in a hammer. For the testing, the hammer is oriented perpendicular to the concrete surface and used to make several hits. An angle scale is, then, used to measure the imprints on the striking block and the concrete. In order to facilitate diameter measurements, carbon or white paper sheets are used. The characteristic outputs are recorded and the average value is calculated. The concrete strength is determined by the ratio of sizes of the imprints.
The working principle of the plastic yield testing instruments is based on the die impression by a hit or with static pressure. Static pressure devices have limited use, though; impact instruments are most common: hand-held and spring-loaded hammers, pendulum devices with ball/disk die. The minimum hardness of the die steel is HRC60, the ball's diameter is at least 10 mm, and the disk's thickness is 1 mm or more. The impact energy should be equal to or greater than 125 N.
This method is simple and fast and may be used for densely reinforced structures, but only suitable for evaluating the strength of concrete up to M500.

Ultrasonic testing

The ultrasonic method is the record of the velocity of penetrating ultrasonic waves. The tests technically distinguish point-to-point scanning, when probes are set on different sides of the tested specimen, and surficial scanning, when the probes are set at one side. The point-to-point scanning, contrary to all other strength NDT methods, allows to test the strength in sub-surficial and deep layers of structures.
Ultrasonic instruments for non-destructive concrete testing may be used not only for the concrete strength determination but also for the flaw detection, quality control of concrete casting, determination of depth and search for reinforcement bars in concrete. They allow to conduct multiple massive tests of products of any shape and to continuously monitor the increase or decrease in strength.
The concrete strength vs. ultrasound velocity ratio is subject to the filler amount and composition, cement flow rate, concrete mix preparation and concrete compaction degree. A drawback of the method is the relatively high error at the acoustic-to-strength performance transition.
Apart from the methods listed here, there are some less popular strength test methods. Electric potential method, infrared, vibration and acoustic methods are at their experimental stage of use.

III. Results

Table 4: *Accuracy of concrete non-destructive test methods*

№	Method	Application range, MPa	Measuring accuracy
1	Plastic yield	5…50	± 30... 40%
2	Rebound resilience	5... 50	± 50%
3	Impact pulse	10... 70	± 50%
4	Pullout	5... 60	no data
5	Shear test	5... 100	no data
6	Edge chipping	10... 70	no data
7	Ultrasonic	10... 40	± 30... 50%

The test area requirements are listed in the following table 5:

Table 5: *Test methods*

Method	Total measurements per area	Minimum distance between measuring sports in an area, mm	Minimum structure edge to measuring spot distance, mm	Minimum structure thickness, mm
Rebound resilience	9	30	50	100
Impact pulse	10	15	50	50
Plastic yield	5	30	50	70
Edge chipping	2	200	-0	170
Pullout	1	2x disk diameter	50	50
Shear test at anchor depth: 40 mm < 40 mm	 1 2	5h	150	2h

The most challenging cases for testing the concrete framings are when they are exposed to aggressive factors: chemical (salts, acids, oils), thermal (high temperatures, freezing at an early age, varying freezes and thaws), atmospheric (carbonization of the surface layer). During the inspection it is necessary to visually, by tapping or wetting with phenolphthalein solution (cases of concrete carbonization) identify the surface layer with disturbed structure. The concrete of such framings for non-destructive testing is prepared by removing the surface layer at the control area and scouring the surface with a honing stick. In such a case, the strength of concrete should be primarily determined by local failure methods or by sampling. When impact-pulse and ultrasonic devices are used, surface roughness should not exceed Ra 25, as shown in Table 6.

Table 6: *Strength of concrete grades*

Concrete compression strength class (B)	Nearest concrete grade (M) by compression strength	Average strength of concrete of this class, kgf/cm^2	Deviations of the nearest concrete grade from the strength of concrete of this class, %
B3,5	M50	45.84	+9.1
B5	M75	65.48	+14.5
B7,5	M100	98.23	+1.8
B10	M150	130.97	+14.5
B12,5	M150	163.71	-8.4
B15	M200	196.45	+1.8
B20	M250	261.94	-4.6
B22,5	M300	294.68	+1.8
B25	M350	327.42	+6.9
B27,5	M350	360.16	-2.8

Concrete compression strength class (B)	Nearest concrete grade (M) by compression strength	Average strength of concrete of this class, kgf/cm^2	Deviations of the nearest concrete grade from the strength of concrete of this class, %
B30	M400	392.90	+1.8
B35	M450	458.39	-1.8
B40	M500	523.87	-4.6
B45	M600	589	
B50	M650	655	
B55	M700	720	
B60	M800	786	

Non-destructive humidity testing

A certain moisture (up to 30-50% for cellular concrete) dwells in construction materials in the course of manufacturing process (the process moisture). Normally, the moisture content of concrete framings during the first heating period reduces to 4-6% by weight.

In order to obtain the whole picture, it is advisable to proceed with several evaluations with different physical principles. Moisture meters or humidity testers are used to measure the moisture content of concrete. The operating principle of a moisture meter is based on the dependence between the dielectric permittivity of a material and its moisture content. It is important to note that the moisture content of a concrete differs from its content on the surface. The measuring methods on the surface are resultant for a depth down to 20 mm and do not always follow the reality.

IV. Conclusions

Based on the studies performed, we may conclude that the actual strength of concrete framings can be set by various non-destructive methods, as well as the main parameters affecting the quality of products and building structures by modern, high-precision instruments.

In order to check and evaluate the concrete strength, it is advisable to use non-destructive test methods as they are more accessible and inexpensive in comparison with laboratory testing of specimens. The main provision to obtain reliable values is the construction of calibration curves of the instruments. It is also necessary to address any factors leading to distorted measuring results.

Cost-effectiveness can be achieved both during the construction of buildings and structures and in the process of their operation. This is promoted by non-destructive methods of quality testing and evaluation of the materials used.

References

[1] Rotaru, A.N. and Maklakov A.S., (2023). Natural and man-made risks. Safety of structures No. 3 (64). 31-33.

[2] Naumenko, A.P., (2019). Introduction to technical diagnostics and non-destructive testing, Omsk.

[3] Non-destructive testing: reference book: in 8 volumes. Vol. 3: Ultrasonic testing – Ed. 2, revised and amended. – Moscow, Mashinostroenie, 2008, 864.

[4] Non-destructive test methods. Physical basics, practical applications, development outlooks. – Moscow, Mir, 2017, 496.

[5] Bykov, I. and Boreiko, D. Methods and practices of non-destructive testing. – Moscow: LAP Lambert Academic Publishing, 2015, 204.

[6] Lyapidevskaya, O.B., Bezgulova, E.A. Concrete strength non-destructive test methods. Comparative analysis of Russian and European building standards. Study guide, – Moscow, MGSU, 2014, 914.

[7] Zabelina, O. Application of non-destructive methods control within the inspection of concrete structures, Moscow, 2021.

[8] Maierhofer, Ch., Reinhardt, H.-W., Dobmann, G., Non-destructive evaluation of reinforced concrete structures, Woodhead Publishing Limited, 2010.

[9] Kouche, El., Hassanein, H.S., Ultrasonic non-destructive testing using wireless sensor networks. Procedia Computer Science, 10, 136-143.

[10] Sanayei, M., Phelps, J.E., Sipple, J.D., Bell, E.S., Brenner, B.R., Instrumentation, nondestructive testing, and finite-element model updating for bridge evaluation using strain measurements, Journal of bridge engineering, 17(1), 2011, 130-138.

[11] Teplý, B., Keršner, Z., Rovnaník, P. and Chromá, M., Durability vs. Reliability of RC structures. Durability of Building Materials and Components, 2005, 17-20.

[12] Balaji Rao, K., Anoop, M.B., Lakshmanan, N., Gopalakrishnan, S. and Appa Rao, T.V.S.R., Risk-based remaining life assessment of corrosion affected reinforced concrete structural members, 2004.

[13] Odriozola, M.A.B. and Gutiérrez, P.A., Comparative study of different test methods for reinforced concrete durability assessment in marine environment. Materials and Structures, 41(3), 2008, 527-541.

[14] Mehta, P.K. and Gerwick, B.C., Concrete in service of modern world. Dundee, Scotland, UK: Univ. of Dundee, 1996, 1-28.

[15] Lang, Ch., Willmes, M., Non-destructive testing of reinforced concrete structures, Germany, 2018.

[16] Kairu, Wilson Macharia, Non-destructive of concrete structures using Schmidt Hammer and profometer 5+ , University of Nairobi, 2016, 32-41.

[17] Al-Mishhadanu S.A., Joni H.H., Radhi M.S., Effect of age on Non-destructive tests results for existing concrete, The Iraqi journal for mechanical and material engineering, vol. 12 (4), 2012.

[18] Akash J., Ankit K., Adarsh K., Yogesh V. and Krishna M., Combined Use of Non-Destructive Tests for Assessment of Strength of Concrete in Structure, Procedia Engineering 54, 2013, 241–251.

[19] ACI Committee 228, Report on Non-destructive Test Methods for Evaluation of Concrete in Structures, ACI 228.2R. American Concrete Institute, Farmington Hills, MI, 2013.

[20] Ngugi H.N., Effects of Sand Quality on Compressive Strength of Concrete: A Case of Nairobi County and Its Environs, Kenya Open Journal of Civil Engineering, 4, 2014, 255-273.

RT&A, No 2 (78)
Volume 19, June, 2024

APPLICATION OF POLAR COORDINATES IN THE SUMMATION OF THE GAUSSIAN DISTRIBUTION

*[1]James DANIEL futathesis@gmail.com
[2]Kayode AYINDE ayindek@nwmissouri.edu
[3]Emmanuel Erhuvwu DUDU eautocorrelation@gmail.com and
[4]Okechukwu Ijeoma EBERECHUKWU ijeoma.okechukwu@federalpolyoko.edu.ng

•

[1]National Bureau of Statistics, Abuja, Nigeria
[2]Northwest Missouri State University, Maryvile, USA
[3]Federal Ministry of Trade and Investments, Asaba, Nigeria
[4]Federal Polytechnic, Oko, Nigeria
*futathesis@gmail.com

Abstract

This work applies the polar coordinates system of advanced calculus in the summation of the Gaussian distribution. In trying to achieve this aim, sub-concepts such as complex variables, gamma function of half, error function, and the relation between the error function and the standard normal distribution were defined and explained at various stages of the work. The embedded theorem which seems to be a new theorem also came up in the body of the work.

Keywords: Normal distribution, Standard Normal Distribution, Gaussian Distribution, gamma Function of Half, Embedded Theorem, Polar Coordinates.

1. Introduction

When Mathematics is used to study observational phenomena, a mathematical model is constructed for the phenomena. This involves an idealization and simplification of the original phenomena to the extent that a mathematical problem is developed. The mathematical solution obtained, eventually has to be interpreted in terms of the original problem.There are essentially two types of mathematical models:the deterministic model and the non-deterministic or probabilistic model [12]. The deterministic model is a model which stipulates that the conditions under which an experiment is performed determine the outcome of the experiment. Example, a body is allowed to fall freely from a height above ground level, the distance(s) traveled is completely determined by the time t (seconds) during which the body has been in motion and the initial velocity u with acceleration a, is given as $S = a_t^2 + u_t$. Based on the given expression, it is possible to determine the value of S for known values of u and t. this shows that for deterministic models, the results of the experiment depend only on the physical conditions operating [4, 2].

However, non-deterministic or probabilistic models introduce uncertainty into the mathematical problem [7]. In the context of probabilistic models, the Gaussian distribution, also known as the normal distribution, plays a vital role in various fields such as statistics, physics, finance, and engineering [6]. In recent years, there has been a growing interest in developing efficient methods for the summation of the Gaussian distribution. One such method is the application of polar coordinates in the summation of the Gaussian distribution[8]. [10] proposed a Bayesian inferential

method for directional data modelled by projected normal distributions. [18] Projected normal distributions, also referred to as angular Gaussian distributions, are created by imposing different constraints on the parameter space associated with a multivariate Gaussian distribution. This resolves the non-identifiability issue that arises when the support of a random variable changes from an Euclidean space to a spherical space [9]. In mathematics and statistics, the Gaussian distribution, also known as the normal distribution, is a crucial concept used to model various real-world phenomena that exhibit a bell-shaped curve [17] [13].

The Gaussian distribution is characterised by its mean and standard deviation, which determine the central tendency and spread of the distribution, respectively [11]. In the work by [10], they proposed a Bayesian inferential method for directional data modeled by projected normal distributions, which are also referred to as angular Gaussian distributions. These distributions are created by imposing different constraints on the parameter space associated with a multivariate Gaussian distribution, allowing for the resolution of the non-identifiability issue when the support of a random variable changes from a Euclidean space to a spherical space. The general projected normal distribution, a simple and intuitive model for directional data in any dimension, is discussed by [10]. They describe a new parameterisation of the general projected normal distribution that makes inference in any dimension tractable, including the important three-dimensional case. This new parameterisation allows for closed-form full conditionals of the unknown parameters and proposes a slice sampler to draw the latent lengths without rejection. The work by [10] demonstrates the applicability and effectiveness of the projected normal distribution in modeling directional data, particularly in higher dimensions.

1.1. Statement of the Problem

In an attempt to prove that

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)} \delta_k = 1 \tag{1.1}$$

one will meet the following problems:

1. One must understand the meaning of the gamma function of half which is defined by [15] as
$$\Gamma\left(1/2\right) = \int_{0}^{\infty} t^{-1/2} \exp^{-t} \delta_t \tag{1.2}$$
2. The proof of the integral function
$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp^{-(1/2)t^2} \delta_t = 1 \tag{1.3}$$
must be known.

This work will make these problems easy to see.

1.2. Aim and Objectives of the Study

The aim of this work is to show clearly that equation 1.1 is equal to 1 without making assumptions of any kind. The main objectives of this work is as follows:

1. The derivation of the Gaussian distribution.
2. To prove the Gaussian distribution using the direct integration method.

2. Materials and Methods

2.1. Binomial Distribution

Proposition 1. If X is a binomial random variable, then the probability of obtaining x successes in n trials of a binomial experiment with probability of success P is given by

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}; & x = 0, 1, 2, \cdots, n; \quad 0 < p < 1 \\ 0, & \text{otherwise} \end{cases} \tag{2.1}$$

We show that $f(x)$ is a probability distribution function with parameters n and P. At this stage, n is a positive integer and $0 < P < 1$, it is clear that $f(x) \geq 0$

$$\begin{aligned} \sum_{i=0}^{n} f(x) &= \sum_{i=0}^{n} \binom{n}{x} p^x (1-p)^{(n-x)} \\ &= [(1-p)+p]^n \\ \therefore \sum_{i=0}^{n} f(x) &= 1 \end{aligned} \tag{2.2}$$

Proposition 1 is called the binomial distribution.

Theorem 2.1. If X has binomial distribution, then the moment-generating function of the random variable X is $M_X t = \left[(1-p) + Pe^t\right]^n$ **Proof.**

$$\begin{aligned} M_X t &= \mathbb{E}\left(e^{tX}\right) \\ &= \sum_{i=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{i=0}^{n} e^{te} \binom{n}{x} \left(pe^t\right)^x (1-p)^{n-x} \\ M_X t &= \left[(1-p) + pe^t\right]^n \end{aligned}$$

■

Corrolary 1. If X has a binomial distribution, then

$$\mathbb{E}(X) = np \tag{2.3}$$

$$\mathbb{V}\text{ar}(X) = np(1-p) \tag{2.4}$$

2.2. The Derivation of the Normal Distribution

[1] states the limit of the symmetrical binomial distribution using theorem 2.2 below.

Theorem 2.2. If X has a symmetrical binomial distribution with mean μ and variance σ^2,then as n tends to infinity,

$$Z = \frac{(x-\mu)}{\sigma^2} \tag{2.5}$$

Equation 2.5 approaches the standard normal distribution.

Proof.

$$\begin{aligned} \mu &= np = \tfrac{1}{2}n; \\ \sigma &= \sqrt{np(1-p)} = \tfrac{1}{2}\sqrt{n}; \\ Z &= \frac{x-\mu}{\sigma} = \frac{x - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}; \end{aligned}$$

Now the distance ΔZ between successive values of Z is given by

$$\Delta Z = \frac{(x+1) - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} - \frac{x - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} = \frac{1}{\frac{1}{2}\sqrt{n}} \tag{2.6}$$
$$\lim_{n\to\infty} \Delta Z = 0$$

Hence the symmetrical binomial histogram will appears to become more like a curve as n tends to infinity

.

We take the value of $f(x) = Y$. Then the distance ΔY is the value between two successive values of Y.

We take the values corresponding to x and $x+1$ and multiply them by σ .

$$\begin{aligned}
Y &= \binom{n}{x} p^x (1-p)^{n-x} \sigma \\
&= \frac{n!}{(n-x)!\,x!} \left(\tfrac{1}{2}\sqrt{n}\right) \left(\tfrac{1}{2}n\right) \\
\Delta Y &= \frac{n!}{(n-x+1)!\,x+1!} \left(\tfrac{1}{2}n\right)\left(\tfrac{1}{2}\sqrt{n}\right) \\
&\quad - \frac{n!}{(n-x)!\,x!} \left(\tfrac{1}{2}n\right)\left(\tfrac{1}{2}\sqrt{n}\right) \\
&= \left(\tfrac{1}{2}n\right)\left(\tfrac{1}{2}\sqrt{n}\right) n! \\
&\quad \times \left[\frac{(n-x)!\,x! - (n-x+1)!\,(n-x)!}{(n-x+1)!\,(n-x)!\,(n-x)!\,x!}\right] \\
&= \left(\tfrac{1}{2}n\right)\left(\tfrac{1}{2}\sqrt{n}\right) n! \\
&\quad \times \left[\frac{\cancel{(n-x+1)!\,x!}[(n+x) - (x+1)]}{\cancel{(n-x+1)!}\,(x+1)!\,(n-x)!\,\cancel{x!}}\right] \\
&= \left(\tfrac{1}{2}n\right)\left(\tfrac{1}{2}\sqrt{n}\right)\sqrt{n}) \left[\frac{n!}{(x+1)!\,(n-x)!}\right] \\
&\quad \times [(n-x) - (x+1)] \\
&= \left(\tfrac{1}{2}n\right)\left(\tfrac{1}{2}\sqrt{n}\right) \left[\frac{n!}{(n-x)!\,x!}\right] \\
&\quad \times \frac{(n-x-x-1)}{(x+1)} \\
\Delta Y &= Y\left[\frac{(n-2x-1)}{(x+1)}\right]
\end{aligned}$$

From equation 2.6

$$\begin{aligned}
\Delta Z &= \frac{1}{\frac{1}{2}\sqrt{n}} \\
\therefore \frac{\Delta Y}{\Delta Z} &= Y\left[\frac{(n-2x-1)}{(x+1)}\right] \tfrac{1}{2}\sqrt{n}
\end{aligned} \tag{2.7}$$

From equation 2.5

$$Z = \frac{(x-\mu)}{\sigma^2} = Z\sigma + \mu$$
$$x = Z\tfrac{1}{2}\sqrt{n} + \tfrac{1}{2}n \tag{2.8}$$

Substitute equation 2.8 into equation 2.7, we have

$$\frac{\Delta Y}{\Delta Z} = Y\left[\frac{\left(n-\sqrt{n}Z-n-1\right)}{\left(\frac{1}{2}\sqrt{n}Z+\frac{1}{2}n+1\right)}\right]\tfrac{1}{2}\sqrt{n}$$
$$\frac{\Delta Y}{\Delta Z} = Y\left[\frac{-\left(\frac{1}{2}n\right)Z-\left(\frac{1}{2}\sqrt{n}\right)}{\left(\frac{1}{2}\sqrt{n}\right)Z+\left(\frac{1}{2}n\right)+1}\right]$$

$\lim_{n\to\infty}\frac{\Delta Y}{\Delta Z}$ tends to $\frac{\delta Y}{\delta Z} = -YZ$ separating the variables

$$\int\frac{\delta Y}{Y} = \int -Z\delta Z$$
$$log_e Y = \frac{-Z^2}{2} + log_e K$$

where K is the constant of Integration

$$log_e\frac{Y}{K} = \frac{-Z^2}{2}$$
$$\therefore Y = Ke^{-\frac{1}{2}Z^2} \tag{2.9}$$

■

2.3. The Proof of the Standard Normal Distribution Using Substitution Method

[16] states the standard normal distribution as in theorem 2.3 below.

Theorem 2.3. The random variable Z is said to have a standard normal distribution if its pdf is $\varphi(Z) = f(Z;0,1) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}Z^2}$ We show below that $\varphi(z)$ is a valid pdf **Proof.**

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}Z^2}\delta Z = \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty}e^{-\frac{1}{2}Z^2}\delta Z \tag{2.10}$$

let $x = \frac{1}{2}Z^2$, so that

$$\delta Z = \frac{\sqrt{2}}{2\sqrt{x}\delta x} \tag{2.11}$$

Substitute equation 2.11 into equation 2.10

$$\begin{aligned}\frac{2}{\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{1}{2}Z^2}\delta Z &= \frac{1}{\sqrt{\pi}}\int_0^{\infty} x^{-\frac{1}{2}}e^{-x}\delta x \\ &= \tfrac{1}{2}\sqrt{\pi}\Gamma\left(\tfrac{1}{2}\right) \\ &= \frac{\sqrt{\pi}}{\sqrt{\pi}} \\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}Z^2}\delta Z &= 1\end{aligned} \tag{2.12}$$

■

[16] is silent about the origin of the standard normal distribution. Also there is no attempt to show Let us call equation 2.13

$$\Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi} \tag{2.13}$$

In the proof, we shall called equation 2.13 assumption 1. [16] also stated theorem 2.4 below.

Theorem 2.4. Let Z have a standard normal distribution. Define x to be $x = \sigma Z + \mu$. Then it can be shown that x is a random normal variable with pdf given as

$$f\left(x;\mu,\sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Proof.

$$\begin{aligned}x &= \sigma z + \mu \\ Z &= \frac{x-\mu}{\sigma} \\ \frac{\delta z}{\delta x} &= \frac{1}{\sigma}\end{aligned}$$

The density function for x is

$$\begin{aligned}f\left(x;\mu,\sigma^2\right) &= \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \\ \begin{cases}-\infty < \mu < \infty \\ \sigma > 0\end{cases}\end{aligned} \tag{2.14}$$

Now we now show that $f\left(x;\mu,\sigma^2\right) = 1$

$$f\left(x;\mu,\sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\delta x \tag{15}$$

Let $y = \dfrac{x-\mu}{\sigma}$, $x = y\sigma + \mu$

$$\delta_x = \sigma\delta_y \tag{16}$$

Substituting equation 16 into equation 15 we have

$$\int_{-\infty}^{\infty} f(y;0,1) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\delta_y$$

From Theorem 2.3

$$\int_{-\infty}^{\infty} f(y;0,1) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\delta_y = 1 \tag{2.17}$$

■

His entire work rest on the assumption 1 of Theorem 2.3. Assumption 1 is the **gamma function of half**. Theorem 2.5 below is the proof of the gamma function of half as resented by [5].

Theorem 2.5. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Proof.

$$\Gamma(n) = \int_0^{\infty} x^{n-1}e^{-x}\delta_x$$

Let $n = \frac{1}{2}$

$$\Gamma(\tfrac{1}{2}) = \int_0^{\infty} x^{-\frac{1}{2}}e^{-x}\delta_x \tag{18}$$

Put $x^{\frac{1}{2}} = \frac{1}{\sqrt{2}}t$, $x = \frac{1}{2}t^2$, $\frac{\delta_x}{\delta_t} = t$

$$\delta_x = t\delta_t \tag{19}$$

Substituting equation 19 into equation 18 we have

$$\begin{aligned}
\Gamma(1/2) &= \int_0^\infty \left(\tfrac{1}{2}t^2\right)^{-1/2} e^{-\frac{1}{2}t^2} t\delta_t \\
&= \int_0^\infty e^{-\frac{1}{2}t^2} \left(\tfrac{1}{2}t^2\right)^{-1/2} t\delta_t \\
&= \sqrt{2}\int_0^\infty e^{-\frac{1}{2}t^2} t\delta_t \\
&= \tfrac{1}{2}\sqrt{2}\int_0^\infty e^{-\frac{1}{2}t^2} t\delta_t \\
&= \frac{1}{2}\sqrt{2}\sqrt{2\pi}\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\delta_t \qquad (20) \\
&= \frac{1}{2}\sqrt{2}\sqrt{2\pi}\times 1 \\
&= \frac{1}{2}\sqrt{2}\sqrt{2}\sqrt{\pi} \\
\Gamma(1/2) &= \sqrt{\pi}
\end{aligned}$$

■

At equation 20 she made the assumption that $\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\delta_t = 1$. Let called this **assumption 2**. The success of the proof of Theorem 2.5 depends on assumption 2. Now for [16] to prove assumption 2 (2.3), he made assumption 1 [(2.5). Also for [5] to prove assumption 1 (2.5) she made assumption 2 (2.3). Let us see how [5] presents the proof of the normal distribution. She used theorem 2.6 below.

Theorem 2.6. A random variable X has a normal distribution and is referred to as a normal random variable if and only if its probability density is given by $f\left(x;\mu,\sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Proof. Since e^x is always positive, it follows that f(x) ≥ 0 as long as $\sigma > 0$.
We show that the total area under the curve is equal to 1. That is, to show that

$$\int_{-\infty}^\infty f(x)\delta_x = 1$$

Let $Z = \frac{(x-\mu)}{\sigma^2}$ and $\delta_x = \sigma\delta_z$

$$\begin{aligned}
\int_{-\infty}^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\delta_x &= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\delta_z \\
&= \int_0^\infty \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\delta_z \qquad (21)
\end{aligned}$$

But $\int_0^\infty e^{-\frac{1}{2}z^2}\delta_z = \frac{\Gamma(1/2)}{\sqrt{2}}$

(22)

Substituting equation 22 into equation 21 we have

$$\begin{aligned} &= \frac{2}{\sqrt{2\pi}} \times \frac{\Gamma(1/2)}{\sqrt{2}} \\ &= \frac{2\times\sqrt{\pi}}{\sqrt{2\pi}\times\sqrt{2}} \end{aligned} \tag{23}$$

$$\therefore \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \delta_x = 1$$

(2.24)

■

Again we can see the assumption 1 at the point of equation 23. That is $\Gamma(1/2) = \sqrt{\pi}$. Just like the work of [16], [5] also made the same assumption 1 in order to prove the standard normal distribution, which in this case is known as assumption 2.

2.3.1 The Complex Number System

[14, 3] stated that there is no real number x that satisfies the polynomial equation

$$x^2 + 1 = 0 \tag{2.25}$$

To permit solution of equation 2.25 and other similar equations, the set of complex number is introduced. A complex number takes the form.

$$z = a + bi \tag{2.26}$$

Where a and b are real numbers and i which is called the imaginary unit has the property that.

$$i = -1 \tag{2.27}$$

From equation 2.26, a is called the real part of z and b is called the imaginary part of z. z is called a complex variable.

2.3.2 The Argand Diagram

The real number can be graphically represented as a point on the real line. By using the cartesian coordinate system, a pair of real numbers can be graphically represented by a point in the plane. The Argand diagram is a device which represents complex numbers in the plane of the Cartesian coordinate system. The pair of real numbers a and b of equation 2.26 are plotted as a point in the plane and then joined that point to the origin with a straight line. See figure 2.1 below.

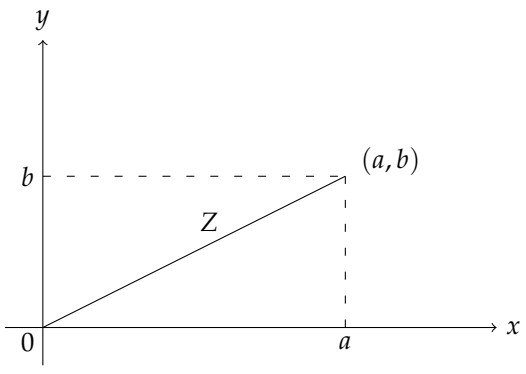


Figure 2.1: *Argand Diagram 1*

Figure 2.1 presents a visual representation of a complex number in polar coordinates on an Argand diagram. The diagram consists of a complex plane with a horizontal real axis ($X-axis$) and a vertical imaginary axis ($Y-axis$). A complex number 'z' is depicted as a point in this plane. The distance 'r' from the origin to the point represents the modulus of the complex number, which is the magnitude of the vector. The angle 'θ' (theta) between the positive real axis and the line segment connecting the origin to the point 'z' represents the argument of the complex number, which indicates its direction. The coordinates 'a' and 'b' on the real and imaginary axes, respectively, correspond to the real and imaginary parts of the complex number. The polar form of the complex number is expressed as '$z = r(\cos\theta + \sin\theta)$', which provides an alternative way to represent complex numbers using the magnitude and angle instead of the traditional rectangular form '$a + b_i$'.

According to [3], this straight line is the graphical representation of the complex variable z of equation refeq2.14. The plane it is plotted against is referred to as the complex plane. The entire diagram is called an Agrand Diagram.

2.4. Polar Form of a Complex Variable

We can express the complex variable of equation refeq2.14 in a different form on an Argand diagram. Let Oz be a complex variable. Let r be the length of complex variable and 0 the angle made with OX. See figure 2.2 below.

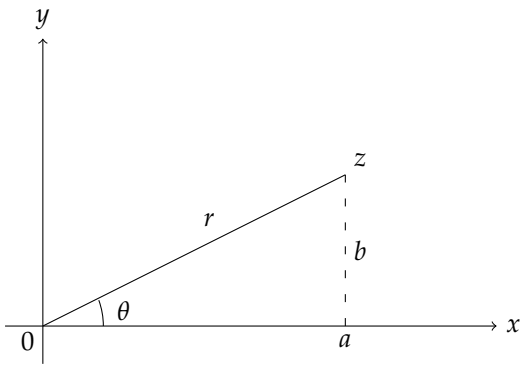


Figure 2.2: *Argand Diagram 2*

Figure 2.2 demonstrates the Cartesian coordinate system with the X-axis and Y-axis representing the real and imaginary parts of complex numbers, respectively. The figure is used to explain the concept of integrating the function $e^{(}-x?)$ over the entire range of x to obtain the value of the integral 'I'. The shaded area under the curve of the function $e^{(}-x?)$ in the first quadrant of the ($X?Y$) plane represents the geometric interpretation of the integral. The integral 'I' is a key component in the derivation of the standard normal distribution and is related to the gamma function and the area under the normal curve.

From Figure 2.2

$$r = \sqrt{a^2 + b^2} \tag{2.28}$$

$$\theta = \tan^{-1}\left(\frac{b}{a}\right) \tag{2.29}$$

$$a = r\cos\theta \tag{2.30}$$

$$b = r\sin\theta \tag{2.31}$$

Substituting equation 2.30 and equation 2.30 into equation 2.26 we have

$$z = r\left(\cos\theta + \sin\theta\right) \tag{2.32}$$

Equation 2.32 is the polar form of equation 2.26, r is called the modulus of the complex variable z and is often abbreviated to 'Mod z' or indicated by $|z|$. 0 is called the argument of the complex variable and can be abbreviated to 'arg z'.

2.5. Integral Functions

The gamma function $\Gamma(x)$ is defined by the integral

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \delta_t \text{ for } x > 0 \tag{33}$$

Integrating equation 33 by part we have

$$\Gamma(x+1) = x\Gamma(x) \tag{34}$$

When $x = n$, a positive integer greater than 1, equation 34 becomes

$$\Gamma(n+1) = n!\Gamma(1) \tag{35}$$

From equation 33 we have that

$$\Gamma(1) = 1 \tag{36}$$

Substitute equation 36 into equation 35 we have

$$\Gamma(n+1) = n! \tag{37}$$

When $x = 1/2$

equation 33 becomes

$$\Gamma(1/2) = \int_0^\infty t^{(\frac{1}{2})} e^{-t} \delta_t \tag{38}$$

3. Analysis and Result

3.1. The Direct Intergration Method

3.1.1 The Gamma Function of Half $\Gamma(1/2)$

Theorem 3.1. The gamma function of half defined as follows:

$$\begin{aligned}\Gamma(1/2) &= \int_0^\infty t^{(1/2)} e^{-t} \delta_t \\ &= \Gamma(\pi)\end{aligned}$$

Proof.

$$\Gamma(1/2) = \int_0^\infty t^{(1/2)} e^{-t} \delta_t$$

Let $t = u^2$; $\delta_t = 2u\delta_u$, $\Gamma(1/2) = \int_0^\infty u^{-1} e^{-u^2} 2u\delta_u$

$$\Gamma(1/2) = 2\int_0^\infty e^{-u^2} \delta_u \tag{1}$$

Unfortunately, $\int_0^\infty e^{-u^2} \delta_u$ cannot easily be determined by normal means. It is however, important, so we have to find a way of getting round the difficulty. We now convert equation 1 into the polar coordinates form. See figure 3.1 below.

Let $I = \int_0^\infty e^{-x^2} \delta_x$ Then also $I = \int_0^\infty e^{-y^2} \delta_y$

$$\begin{aligned} I^2 &= \left(\int_0^\infty e^{-x^2} \delta_x\right)\left(\int_0^\infty e^{-y^2} \delta_y\right) \\ &= \left(\int_0^\infty e^{-x^2} \delta_x\right)\left(\int_0^\infty e^{-y^2} \delta_y\right) \\ &= \int_0^\infty \int_0^\infty e^{\left(x^2+y^2\right)} \delta_x \delta_y \end{aligned} \tag{2}$$

$\delta_a = \delta_x \delta_y$ represent an element of area in the $(X - Y)$ plane and the integration with the stated limit covers the whole of the first quadrant. See figure 3.2 below

Now converting to polar coordinates, the element of area becomes $\delta_a = r\delta\theta\delta_r$

$$\begin{aligned} r^2 &= x^2 + y^2 \\ e^{-(x^2+y^2)} &= e^{-r^2} \end{aligned} \tag{3}$$

Form figure 3.2 below the limit of r are $0 \leq r \leq \infty$. The limit of θ are $O \leq \theta \leq \pi/2$. Equation 2 becomes

$$I^2 = \int_0^{\left(\frac{\pi}{2}\right)} \int_0^\infty e^{-r^2} r \delta_r \delta\theta \tag{4}$$

Let $k = r^2, \quad \delta_k = 2r\delta_r$

$$
\begin{aligned}
&= \int_0^{\left(\frac{\pi}{2}\right)} \int_0^{\infty} e^{\frac{-k^2}{2}} \delta_k \delta\theta \\
&= \int_0^{\left(\frac{\pi}{2}\right)} \left[-\tfrac{1}{2}e^{-k}\right]_0^{\infty} \delta\theta \\
&= \int_0^{\left(\frac{\pi}{2}\right)} \left(\tfrac{1}{2}\right) \delta\theta \\
&= \left[\tfrac{\theta}{2}\right]_0^{\left(\frac{\pi}{2}\right)} \\
&= \tfrac{\pi}{4} \\
\therefore I &= \tfrac{\sqrt{\pi}}{2} \qquad (5)
\end{aligned}
$$

Before the diversion into the polar coordinates, we had established equation 1 that $\Gamma\left(1/2\right) = 2\int_0^{\infty} e^{-u^2}\delta_u$

Then substitute equation 5 into equation 1, $\Gamma\left(1/2\right) = 2 \times \frac{1}{2}\sqrt{\pi}$

$$\Gamma\left(1/2\right) = \sqrt{\pi} \qquad (6)$$

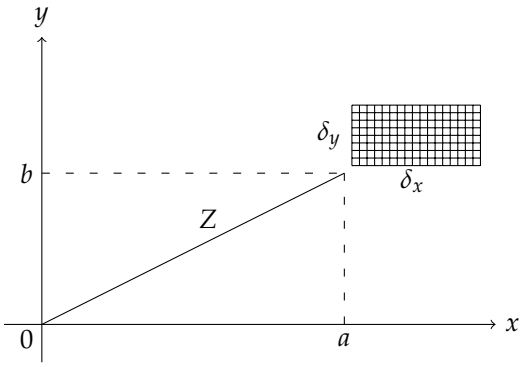


Figure 3.1: *The (X − Y) Plane*

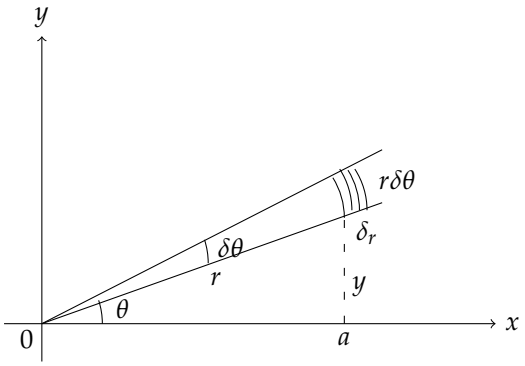


Figure 3.2: *The Complex Plane*

Figure 3.2 provides a graphical representation of the complex plane with polar coordinates (r, θ) used to represent a complex number. The figure demonstrates the conversion of an element of area from Cartesian coordinates (δ_x, δ_y) to polar coordinates $(\delta_a = r\delta_\theta\delta_r)$. This conversion is essential in the proof of the gamma function of half $(?(1/2))$ using polar coordinates. The figure shows how the radial distance 'r' and the angle θ are used to define the position of a point in the complex plane. The element of area in polar coordinates, represented by the shaded sector, is used to integrate the function $e^(- r^2/2)$ over the complex plane, which is a crucial step in the derivation of the standard normal distribution.

This is the proof of the gamma function of half using the polar coordinates system of advanced calculus. This result opens the way for the proof of the standard normal distribution. ■

3.2. The Standard Normal Distribution

Theorem 3.2. $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \delta_z = 1$ **Proof.**

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \delta_z = 2\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{1}{2}z^2} \delta_z \tag{7}$$

Fortunately, we can now apply polar coordinates in the summation of the integral in equation 7. Dividing equation 3 by 2, we have $\frac{r^2}{2} = \frac{x^2+y^2}{2}$

$$e^{-\left(\frac{x^2+y^2}{2}\right)} = e^{-\frac{r^2}{2}}$$

With the same limits as of figure 3.1 we can easily see that

$$I^2 = \int_0^{(\frac{\pi}{2})} \int_0^{\infty} e^{-\frac{1}{2}r^2} r\delta_r\delta\theta$$

Let $k = \frac{1}{2}r^2$, $\delta_r = \frac{\delta_k}{r}$

$$\begin{aligned}
&= \int_0^{(\frac{\pi}{2})} \int_0^{\infty} e^{-k}\delta_k\delta\theta \\
&= \int_0^{(\frac{\pi}{2})} \left[-e^{-k}\right]_0^{\infty}\delta\theta \\
&= \int_0^{(\frac{\pi}{2})} \delta\theta \\
&= [\theta]_0^{(\frac{\pi}{2})} \\
&= \tfrac{\pi}{4} \\
I &= \sqrt{\tfrac{\pi}{2}} \\
\therefore I &= \tfrac{\sqrt{2\pi}}{2} \qquad (8)
\end{aligned}$$

Substituting equation 8 into equation 7, we have,

$$\begin{aligned}
\tfrac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{1}{2}z^2}\delta_z &= \tfrac{2}{\sqrt{2\pi}} \times \tfrac{\sqrt{2\pi}}{2} \\
\therefore \int_{-\infty}^{\infty} \tfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\delta_z &= 1 \qquad (9)
\end{aligned}$$

■

3.3. Derivation of the Normal Curve

Now, back to [1], the derivation of the normal curve from the asymmetric binomial distribution was given in theorem 2.3. From equation 2.9 we had that, $Y = Ke^{-\frac{1}{2}Z^2}$ From proposition (2.0) it is obvious that:

$$y = k \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}\delta_z = 1 \qquad (3.10)$$

From theorem 3.2, we had that

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2}\delta_z = \sqrt{2\pi} \qquad (3.11)$$

Substituting equation 3.11 into equation 3.10 we have that $k = \frac{1}{\sqrt{2\pi}}$

$$y = \tfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad (3.12)$$

Equation 3.12 is called the ”normal curve”.

3.4. The Normal Equation

The equation of the normal distribution was given by equation 1.1 as

If $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$\int_{-\infty}^{\infty} f(x)\delta_x = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \delta_x$$
$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \delta_x$$

Let $z = \frac{x-\mu}{\sigma}$ and $x = z\sigma + \mu$ and $\delta_x = \sigma\delta_z$

$$= \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp^{-\frac{1}{2}z^2} \cancel{\sigma}\delta_z$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp^{-\frac{1}{2}z^2} \delta_z$$

From theorem 3.2, we had that $\int_{-\infty}^{\infty} exp^{-\frac{1}{2}z^2}\delta_z = \sqrt{2\pi}$

$$= \frac{\sqrt{2\pi}}{\sqrt{2\pi}}$$
$$\therefore \int_{-\infty}^{\infty} f(x)\delta_x = 1$$

4. Discussion of Results

Ordinarily, all integral functions are difficult to integrate. They are not well behaved in regard to integration. Hence the integral functions

$$\Gamma(\tfrac{1}{2}) = \int_0^{\infty} t^{x-1}e^{-t}\delta_t$$
$$f(x) = \int_0^{\infty} t^{(1/2)}e^{-t}\delta_t$$

cannot easily be determined by normal means. This may be the root cause why [16] did not make any attempt to prove that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. [5] used the substitution method to prove the gamma function of half and the standard normal distribution. She knew that the direct integration method will leads to complex analysis. First, the so called "Assumption 1" which is

the gamma function of half has being proved to be equal to $\sqrt{\pi}$ on theorem 4.0 with the aid of the polar coordinates system. To the student of statistics this should no longer be an assumption. Secondly the "Assumption 2" which is known as the standard normal distribution was proved on theorem 4.1. Also this is made possible by the aid of the polar coordinates system. The derivation of the normal distribution is also an area where many Authorities shy away from. [3] made an attempt to derive it, but he leaves the integral part of the function untouched.

5. Conclusion

I have not seen the direct integration method in the literature of the normal distribution but substitution method, before now. This work have used the integration method through the help of polar coordinate to derive the summation of the Gaussian Distribution.

References

[1] Backhouse, J. K. (1967). Statistics: An introduction to tests of significance. *(No Title)*.
[2] Bellomo, N. and Pulvirenti, M. (2000). Modeling in applied sciences. modeling and simulation in science, engineering and technology.
[3] Brown, J. W. and Churchill, R. V. (2009). *Complex variables and applications*. McGraw-Hill,.
[4] Cheng, A. K. (2009). Mathematical modelling and real life problem solving. In *Mathematical Problem Solving: Yearbook 2009, Association of Mathematics Educators*, pages 159–182. World Scientific.
[5] Chinwe, R. N. (2002). *Probability and distribution theory*. Afrika-LinkBooks.
[6] Das, A. and Geisler, W. S. (2021). A method to integrate and classify normal distributions. *Journal of Vision*, 21(10):1–1.
[7] Dobronets, B. S., Popova, O. A., and Merko, A. M. (2021). Distributional time series for forecasting and risk assessment. *International Journal of Risk Assessment and Management*, 24(2-4):140–155.
[8] Greco, L., Lucadamo, A., and Agostinelli, C. (2021). Weighted likelihood latent class linear regression. *Statistical Methods & Applications*, 30:711–746.
[9] Gutiérrez, L., Gutiérrez-Pena, E., and Mena, R. H. (2019). A bayesian approach to statistical shape analysis via the projected normal distribution.
[10] Hernandez-Stumpfhauser, D., Breidt, F. J., and van der Woerd, M. J. (2017). The general projected normal distribution of arbitrary dimension: Modeling and bayesian inference.
[11] Maybeck, P. S. (1982). *Stochastic models, estimation, and control*. Academic press.
[12] Mulholland, H. and Phillips, J. H. G. (1984). *Applied Mathematics for Advanced Level:(the Mechanical of Particles and Rigid Bodies)*. Butterworths.
[13] Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied statistics*, 32(7):685–694.
[14] Spiegel, M. R., Lipschutz, S., Schiller, J. J., and Spellman, D. (2009). *Schaum's outline of Complex Variables*. McGraw Hill Professional.
[15] Stroud, K. A. and Booth, D. J. (2020). *Advanced engineering mathematics*. Bloomsbury Publishing.
[16] Ugwuowo, F. I. (2015). *Introduction to probability and distribution theory*. University of Nigeria Press, Nsukka, Nsukka, Nigeria, second edition edition edition.
[17] Williams, C. and Rasmussen, C. (1995). Gaussian processes for regression. *Advances in neural information processing systems*, 8.
[18] Yu, Z. and Huang, X. (2022). Elliptically symmetric distributions for directional data of arbitrary dimension. *arXiv preprint arXiv:2212.05634*.

REGRESSION MODEL OF ARC OVERVOLTAGE DURING SINGLE-PHASE NON-STATIONARY GROUND FAULTS IN NEUTRAL ISOLATED NETWORKS

Najaf Orujov[1], Huseyngulu Guliyev[2] , Sara Alimammadova[3]

[1]Baku Engineering University, Khirdalan, Azerbaijan
AZ0101, Hasan Aliyev str. 120
[1]norucov@beu.edu.az
[2]Azerbaijan Technical University, Baku, Azerbaijan
AZ1073, H. Javid avenue 25
[2]huseyngulu@mail.ru
[3]Azerbaijan State University of Oil and Industry, Baku, Azerbaijan
AZ1010, Azadlig Avenue 16/21
[3]sara_elimemmedova@mail.ru

Abstract

In order to perform insulation tests of electrical equipment under load in neutral insulated networks, it is necessary to create an artificial overvoltage, and at this time, it is necessary to determine the mathematical relationships between the single-phase non-stationary ground and the closing parameters. In the case of single-phase non-stationary earth faults, the dependencies between important parameters such as overvoltage frequency, earth fault resistance and earth fault angle obey complex laws. Therefore, for practical conditions, adequate mathematical models should be developed that allow to know the interdependencies of such parameters. In this work, the problem of analytical determination of the relationship between the overvoltage generated in neutral insulated networks as a result of non-stationary earth faults, the earth fault resistance and the earth fault angle was considered. For this purpose, a regression equation was obtained for the dependence of the overvoltage frequency on the ground fault resistance and the ground fault angle, and the corresponding spatial description was given. The obtained results confirmed the existence of a strong correlation between these parameters and can be used for practical purposes.

Keywords: isolated electrical network, non-stationary ground fault, overvoltage factor, ground fault resistance, ground fault angle, regression equation, correlation

I. Introduction

It is known that the failure of electrical equipment can cause various traumas of the staff, disruption of the technological process and serious accidents, so special tests are carried out to prevent such problems in advance. In general, such tests are carried out when one of the following situations occurs: equipment or installation is put into operation, after an accident, planned and unplanned repairs, a certain period of time has passed since previous inspections, etc. At the same

time, high-voltage testing of electrical equipment insulation is mandatory for neutral-insulated electrical networks with a voltage of up to 35 kV [1, 2].

It is a very urgent issue to obtain preliminary information about the potential damage of electrical equipment in neutral insulated networks and to perform high-voltage tests of insulation under load in order to ensure the uninterrupted supply of electricity to electricity consumers [3, 4]. It is clear from the research works carried out in this direction that various methods and tools are proposed [5-8].

A method of testing the insulation under load in neutral-insulated networks was proposed [9]. According to this method, artificial non-stationary earth faults are created in the network based on Petersen's theory to test the insulation under load. It is possible to use the multifunctional high-voltage thyristor commutator device created in this regard. Using this device, it is possible to create an artificial single-phase non-stationary earth fault that obeys Petersen's theory in neutral-isolated networks. Thus, by providing different values of the phase angles through the control unit of the device, it is possible to create grounding of the phase of the network through the switching unit. Since switching processes are controlled by changing the grounding angle and resistance, it is possible to adjust its characteristic quantities (grounding current, arc output voltage). At the same time, since the quantities characterizing the switching process depend on the insulation resistance of the network with respect to ground, the artificial non-stationary ground faults created on the basis of Petersen's theory in neutral-isolated networks through the commutator allow both monitoring the insulation of the network with respect to ground and detecting damage in it.

The value of the arc overvoltage during the transition processes is of great importance when conducting the tests. The value of the arc overvoltage, as mentioned, depends on the ground fault resistance, the ground fault angle and the phase capacity of the network with respect to ground. Therefore, it is important to determine the ground fault resistance and the ground fault angle in advance in order to control the transient processes and determine the value of the test voltage accordingly. For this purpose, it is important to determine the dependence of the frequency of arc overvoltage in neutral insulated networks as a result of non-stationary earth faults, the dependence of the earth fault resistance, the angle of earth fault and the phase capacity of the network with respect to earth.

II. Statement of the problem for the regression model of arc overvoltage

In general, in order to determine the dependencies between the mentioned parameters, the numerical solution of the system of differential equations characterizing the transition process of the non-stationary earth fault created in neutral isolated networks should be performed using modern computing technologies. However, the numerical solution of the problem becomes much more difficult due to the "stiffness" of the mentioned differential equations. In other words, since the system of differential equations is non-linear, during their numerical integration, the stability of the solution is violated in some cases and the results are distorted. Therefore, in order to overcome such difficulties, it is important to obtain analytical expressions that determine the dependences between the frequency of the arc overvoltage (K) and the ground fault resistance (R_0), the ground fault angle (φ) and the phase capacitance (C_f) of the network with respect to the ground.

It should be noted that for the considered research question, in [10,11], the arc overvoltage ratio is determined from the ground fault resistance, in [12,13], the arc overvoltage ratio is from the ground fault angle, and in [14,15], the arc overvoltage ratio is determined from the derivation of analytical expressions for the dependences of the phase capacity of the network on the ground has already been considered. As a continuation of the conducted research, the regression model of the

dependence of the single-phase arc overvoltage on the ground fault resistance and the ground fault angle in neutral-insulated networks is considered. The mathematical model to be obtained will allow to ensure the value of the arc overvoltage at the required value by controlling the transition processes that occur during the artificial earth-stationary earth faults created for the purpose of carrying out tests under load.

III. Problem solving method and algorithm

Obtaining an analytical expression for the dependence of the frequency of the arc overvoltage during single-phase faults on the neutral-insulated electrical network $(C_f = const)$ on the earth fault resistance and the earth fault angle is considered. For this purpose, the results of the experimental studies carried out in the low-voltage model of the neutral-isolated network, given in table 1, are used $(C_f = 1mkF)$ [8,9].

Table 1: $K = f(R_0, \varphi)$ *addiction*

R_0, Om	φ				
	30^0	60^0	90^0	120^0	150^0
5	2,45	3,16	3,30	3,10	2,16
10	2,29	2,91	2,96	2,86	2,03
15	2,16	2,69	2,77	2,66	1,93
20	2,05	2,51	2,62	2,49	1,84
25	1,96	2,35	2,49	2,35	1,76
30	1,88	2,22	2,38	2,24	1,69

As can be seen from Table 1, the relationship between the frequency of arc overvoltage and the ground fault resistance and the ground fault angle can be approximated by the following regression equation [16]:

$$K = \frac{a}{R_0} + b\sin\varphi + c, \tag{1}$$

Here $a, b, c -$ are regression coefficients.

If we accept substitutions $\frac{1}{R_0} = x$ and $\sin\varphi = y$ in equation (1), we can write the regression equation as follows:

$$K = ax + by + c, \tag{2}$$

In other words, the dependence between the frequency of arc overvoltage (K) and the conductance of the ground fault circuit (x) and the sine of the ground fault angle (y) can be approximated by a linear regression equation (table 2). In determining the type of the model, such a judgment was used that if the change of the result indicator is directly proportional to the change of the factor indicators, then the linear model is considered adequate [17].

Table 1: $K = f(x, y)$ *addiction*

x	y				
$[Sm]$	0,500	0,866	1,000	0,866	0,500
0,200	2,45	3,16	3,30	3,10	2,16
0,100	2,29	2,91	2,96	2,86	2,03
0,067	2,16	2,69	2,77	2,66	1,93
0,050	2,05	2,51	2,62	2,49	1,84
0,040	1,96	2,35	2,49	2,35	1,76
0,033	1,88	2,22	2,38	2,24	1,69

The regression coefficients of equation (2) are determined by the following well-known expressions [16]:

$$\left.\begin{aligned} a &= \frac{\sigma_K}{\sigma_x}\cdot\frac{r_{Kx}-r_{Ky}r_{xy}}{1-r_{xy}^2};\\ b &= \frac{\sigma_K}{\sigma_y}\cdot\frac{r_{Ky}-r_{Kx}r_{xy}}{1-r_{xy}^2};\\ c &= \overline{K}-a\overline{x}-b\overline{y} \end{aligned}\right\} \quad (3)$$

Here $\overline{x}-$ average value of quantity x; $\overline{y}-$ average value of quantity y; $\overline{K}-$ average value of quantity K; σ_x- mean square deviation of quantity x from its mean value $(\overline{x})$; σ_y- mean square deviation of quantity y from its mean value $(\overline{y})$; σ_K- mean square deviation of quantity K from its mean value $(\overline{K})$; $r_{xy}-$ linear correlation coefficient between the quantities x and y; $r_{Kx}-$ linear correlation coefficient between the quantities K and x; $r_{Ky}-$ linear correlation coefficient between the quantities K and y.

IV. Modeling results

The numerical values of the statistical indicators necessary for the determination of the coefficients of regression dependence and the identification of the model sought are determined according to the correlation table given in table 3. According to Table 3, the following markings were adopted:

$$A=(x_i-\bar{x})^2;\ \ B=(y_i-\bar{y})^2;\ \ C=(K_i-\overline{K})^2;\ \ D=(x_i-\bar{x})\cdot(y_i-\bar{y});$$
$$M=(K_i-\overline{K})\cdot(x_i-\bar{x});\ N=(K_i-\overline{K})\cdot(y_i-\bar{y})$$

The data array consists of a sample, and the calculated values of the statistical indicators necessary for determining the coefficients are given below:

$$\sum_{i=1}^{n}x_i=2{,}45;\ \ \sum_{i=1}^{n}y_i=22{,}392;\ \sum_{i=1}^{n}K_i=72{,}26;$$
$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}=0{,}082;\ \ \bar{y}=\frac{\sum_{i=1}^{n}y_i}{n}=0{,}746;\qquad \overline{K}=\frac{\sum_{i=1}^{n}K_i}{n}=2{,}409;$$

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2=0{,}09830667;\ \sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2=1{,}2860832;\ \sum_{i=1}^{n}\left(K_i-\bar{K}\right)^2=5{,}390747;$$

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)=0;\ \sum_{i=1}^{n}\left(K_i-\bar{K}\right)\left(x_i-\bar{x}\right)=0{,}4112667;\ \sum_{i=1}^{n}\left(K_i-\bar{K}\right)\left(y_i-\bar{y}\right)=1{,}998776;$$

Table 3: *Correlation table*

i	x_i	y_i	K_i	A	B	C	D	M	N
1	0,200	0,500	2,45	0,01400278	0,06071296	0,001708	-0,02915733	0,0048911	-0,0101845
2	0,200	0,866	3,16	0,01400278	0,01430416	0,564502	0,01415267	0,0889078	0,0898595
3	0,200	1,000	3,30	0,01400278	0,06431296	0,794475	0,03000933	0,1054744	0,2260421
4	0,200	0,866	3,10	0,01400278	0,01430416	0,477942	0,01415267	0,0818078	0,0826835
5	0,200	0,500	2,16	0,01400278	0,06071296	0,061835	-0,02915733	-0,0294256	0,0612715
6	0,100	0,500	2,29	0,00033611	0,06071296	0,014082	-0,00451733	-0,0021756	0,0292395
7	0,100	0,866	2,91	0,00033611	0,01430416	0,251335	0,00219267	0,0091911	0,0599595
8	0,100	1,000	2,96	0,00033611	0,06431296	0,303968	0,00464933	0,0101078	0,1398181
9	0,100	0,866	2,86	0,00033611	0,01430416	0,203702	0,00219267	0,0082744	0,0539795
10	0,100	0,500	2,03	0,00033611	0,06071296	0,143388	-0,00451733	-0,0069422	0,0933035
11	0,067	0,500	2,16	0,00021511	0,06071296	0,061835	0,00361387	0,0036471	0,0612715
12	0,067	0,866	2,69	0,00021511	0,01430416	0,079148	-0,00175413	-0,0041262	0,0336475
13	0,067	1,000	2,77	0,00021511	0,06431296	0,130562	-0,00371947	-0,0052996	0,0916341
14	0,067	0,866	2,66	0,00021511	0,01430416	0,063168	-0,00175413	-0,0036862	0,0300595
15	0,067	0,500	1,93	0,00021511	0,06071296	0,229122	0,00361387	0,0070204	0,1179435
16	0,050	0,500	2,05	0,00100278	0,06071296	0,128642	0,00780267	0,0113578	0,0883755
17	0,050	0,866	2,51	0,00100278	0,01430416	0,010268	-0,00378733	-0,0032089	0,0121195
18	0,050	1,000	2,62	0,00100278	0,06431296	0,044662	-0,00803067	-0,0066922	0,0535941
19	0,050	0,866	2,49	0,00100278	0,01430416	0,006615	-0,00378733	-0,0025756	0,0097275
20	0,050	0,500	1,84	0,00100278	0,06071296	0,323382	0,00780267	0,0180078	0,1401195
21	0,040	0,500	1,96	0,00173611	0,06071296	0,201302	0,01026667	0,0186944	0,1105515
22	0,040	0,866	2,35	0,00173611	0,01430416	0,003442	-0,00498333	0,0024444	-0,0070165
23	0,040	1,000	2,49	0,00173611	0,06431296	0,006615	-0,01056667	-0,0033889	0,0206261
24	0,040	0,866	2,35	0,00173611	0,01430416	0,003442	-0,00498333	0,0024444	-0,0070165
25	0,040	0,500	1,76	0,00173611	0,06071296	0,420768	0,01026667	0,0270278	0,1598315
26	0,033	0,500	1,88	0,00236844	0,06071296	0,279488	0,01199147	0,0257284	0,1302635

Continuation of table 3

27	0,033	0,866	2,22	0,00236844	0,01430416	0,035595	-0,00582053	0,0091818	-0,0225645
28	0,033	1,000	2,38	0,00236844	0,06431296	0,000822	-0,01234187	0,0013951	-0,0072699
29	0,033	0,866	2,24	0,00236844	0,01430416	0,028448	-0,00582053	0,0082084	-0,0201725
30	0,033	0,500	1,69	0,00236844	0,06071296	0,516482	0,01199147	0,0349751	0,1770795
$\sum$	**2,450**	**22,392**	**72,26**	**0,09830667**	**1,28608320**	**5,390747**	**0,00000000**	**0,4112667**	**1,9987760**

Thus, based on the data of table 3, mean square deviations and two-dimensional correlation coefficients for individual quantities are calculated based on known formulas. The numerical values obtained are as follows:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n}} = 0{,}0572; \quad \sigma_y = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}{n}} = 0{,}207; \quad \sigma_K = \sqrt{\frac{\sum_{i=1}^{n}\left(K_i - \overline{K}\right)^2}{n}} = 0{,}4239;$$

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{n\sigma_x\sigma_y} = 0; \; r_{Kx} = \frac{\sum_{i=1}^{n}\left(K_i - \overline{K}\right)\left(x_i - \overline{x}\right)}{n\sigma_K\sigma_x} = 0{,}5649; \; r_{Ky} = \frac{\sum_{i=1}^{n}\left(K_i - \overline{K}\right)\left(y_i - \overline{y}\right)}{n\sigma_K\sigma_y} = 0{,}7591.$$

Then, based on statements (3), the calculated values of regression coefficients of equation (1) or (2) are obtained as follows:

$$a = 4{,}18; \; b = 1{,}55; \; c = 0{,}91.$$

Thus, after determining the regression coefficients, the dependence (1) (or (2)) between the frequency of the arc overvoltage generated in neutral insulated networks during single-phase non-stationary earth faults and the earth fault resistance (or circuit conductance) and the sine of the earth fault angle can be written in the following obvious way:

$$K = \frac{4{,}18}{R_0} + 1{,}55\sin\varphi + 0{,}91 \tag{4}$$

or

$$K = 4{,}18x + 1{,}55y + 0{,}91 \tag{5}$$

As it can be seen, the regression model obtained in the form of (4) between the parameters during single-phase non-stationary earth faults is in a form that is simple and easy to implement in practice.

Let's check the adequacy of the obtained regression dependence (5) between the frequency of overvoltage and the conductance of the ground fault circuit and the sine of the ground fault angle during single-phase non-stationary ground faults. For this, by calculating the multivariate correlation coefficient, its significance can be checked with the Fisher criterion [16].

Based on the data, the value of the multivariate correlation coefficient is as follows:

$$R = \sqrt{\frac{r_{Kx}^2 + r_{Ky}^2 - 2r_{Kx}r_{Ky}r_{xy}}{1 - r_{xy}^2}} = 0{,}95.$$

We check the significance of the multivariate correlation coefficient with the $F-$Fisher test. It is known that the regression equation at the α significance level is considered adequate if the $F > F(\alpha, k_1, k_2)$ condition is met [17], here $k_1, k_2 -$ are the degrees of freedom.

The reported value of the $F-$Fisher criterion is determined based on the data as follows:

$$F = \frac{\frac{1}{2}R}{\frac{1}{n-3}\left(1-R^2\right)} \approx 132.$$

The table value of the $F-$Fisher criterion is taken from the table depending on the significance level (α) and degrees of freedom (k_1, k_2) [17]:

$$\alpha = 0{,}05;\ \ k_1 = 2;\ k_2 = n-3 = 30-3 = 27;\ F(\alpha, k_1, k_2) = 3{,}35.$$

Since $F = 132 > F(\alpha, k_1, k_2) = 3{,}35$, the multivariate correlation coefficient $(R = 0{,}95)$ and the statistical significance of the regression equation is confirmed.

It should be noted that the multivariate correlation coefficient $(R = 0{,}95 \to 1)$ close to unity indicates that the dependence between the frequency of arc overvoltage and the conductance of the ground-fault circuit and the sine of the ground-fault angle can be considered a strong linear correlation relationship. It is recommended to use neutral-isolated networks when solving practical problems for organizing under-load tests of insulation of electrical equipment.

A 3D (spatial) image of the dependence of the frequency of the arc overvoltage on the earth fault resistance and the earth fault angle was constructed based on the regression equation (4) obtained using computer modeling reporting methods (Fig. 1). The numerical results of the regression model obtained between the parameters mentioned in Figure 1, in other words, the existence of a strong correlation relationship between the quantities are visually confirmed.

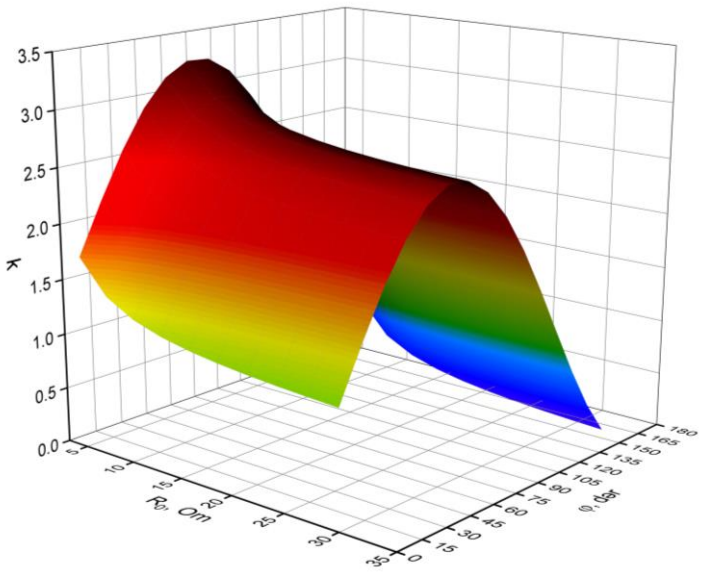


Figure 1. *Earth fault resistance times overvoltage and a 3D image of its dependence on the angle of closure with the ground*

VI. Conclusions

1. An easy-to-realize regression model was obtained between the frequency of arc overvoltage in neutral insulated networks as a result of non-stationary earth faults subject to Petersen's theory, the conductance of the earth fault circuit and the sine of the earth fault angle. The proposed analytical dependence between the mentioned parameters can be considered as a strong linear correlation relationship and based on this, the value of the test voltage can be determined.

2. The obtained results can be easily used during load tests of insulation of electrical equipment in the neutral isolated networks of the Azerenergy system, and at the same time, during the investigation and analysis of the results of non-stationary ground faults that occurred in the network.

References

[1] Khazieva, R.T., Vasiliev, P.I., Aflyatunov, R.R. Research of an installation for testing the insulation of electrical equipment with increased voltage. Electrotechnologies in Industry, 2022, No. 3(56), p. 65-69.

[2] Patsch, R. Dielectric Diagnostics of Power Transformers and Cables - Return Voltage Measurements, Theory and Practical Results. VDE High Voltage Technology, 2018, ETGSymposium. pp.1-6.

[3] Shahmaev, I.Z., Gaisin, B.M., Shiryaev, O.V. A new method of taking management decisions at designing and developing electric power systems. 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), IEEE, 2016, pp.1-6.

[4] Mudiraj, A.N. Improvement of Power Quality by mitigating harmonics in single phase AC distribution. International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), IEEE, 2016, pp. 83-88.

[5] Antonov, A.I. Study of the level of electromagnetic interference in a 10/0.4 kV network with power transformers of various capacities with an asymmetrical load. News of higher educational institutions. Energy Problems, 2017, No. 9-10, p. 65-76.

[6] Rajasekhar, N.V., Babu, M.N. Harmonics reduction and power quality improvement by using DPFC. International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT). IEEE, 2016, pp. 1754-1758.

[7] Akdeniz, E., Bagriyanik, M. A knowledge-based decision support algorithm to reduce the impact of transmission system vulnerabilities. International Journal of Electric Power and Energy Systems. 2016, No. 78. pp. 436-444.

[8] Orujov, N.I. Development and research of insulation test methods in neutral-insulated networks based on artificial non-stationary earth faults. Dissertation submitted for the candidate of technical sciences degree, Baku, 1998, 139 p.

[9] Nabiev, Kh.I. Research and development of a device for simulating earth arc faults to improve the efficiency of 6-35 kV networks. Diss. For the academic degree of Candidate of Technical Sciences, Baku, 1992, 166 p.

[10] Orujov, N.I., Orujov, A.O. Determination of the relationship between the frequency of arc overvoltage and the ground fault resistance. Scientific works. AzTU, -2013. No. 3, pp. 83-86.

[11] Orujov, N.I. Regression model of arc overvoltage dependence on ground fault resistance. “The role of engineering in the innovative development of Azerbaijan: goals and perspectives" international scientific-practical conference, Baku, November 29-30, 2019, p. 285-287.

[12] Orujov, N.I., Orujov, A.O. Determination of the dependence between the arc overvoltage and the ground closing angle. Energy problems, Science, 2014, No. 2, p. 36-39.

[13] Orujov, N.I., Mirili, T. Regression model of the dependence of the arc overvoltage on the ground closing angle. V international scientific conference of young researchers, Baku, Book 1, April 29-30, 2021, p. 113-116.

[14] Orujov, N.I. Determination of the dependence between the frequency of arc overvoltage and the phase capacity of the network with respect to earth. Materials of the Republican Scientific and Technical Conference dedicated to the 100th anniversary of the Azerbaijan People's Republic on "Establishment of the Education-Research-Production Mechanism", Baku, AzTU, April 4-5, 2018, pp. 254-257

[15] Orujov, N.I. Regression model of the dependence of arc overvoltage on the phase capacity of the network with respect to ground. II international science and technology conference, Baku, November 26-27, 2021, p. 333-335.

[16] Mammadov, N.R., Mammadov, B.M. Mathematical processing of the results of the experiment. Baku: Elm, 2005, 160 p.

[17] Orujov, E. Econometrics. Baku: Elm, 2018, 384 p.

RT&A, No 2 (78)
Volume 19, June, 2024

ANALYSIS OF A SINGLE SERVER SYSTEM WITH HETEROGENEOUS ARRIVAL, HETEROGENEOUS SERVICE, SYSTEM FAILURE AND MAINTENANCE

Mohammed Shapique A, Vaithiyanathan A

•

IFET College of Engineering, Villupuram, India
shapique@gmail.com, vaithi05@gmail.com

Abstract

This paper investigates a single-server queuing system with heterogeneous service, failure, and maintenance. The proposed model features a server acting as both the main and backup server. System failure can occur at any stage. When a failure happens, instead of stopping the service entirely, the main server functions as a backup, providing service at a reduced rate. Once all jobs in the system have been serviced, the backup server enters the maintenance state. Following the repair process during maintenance, the server transitions to an idle state, awaiting incoming jobs. Explicit expressions for both transient and steady-state behaviours of the system are derived. Additionally, key system performance metrics are discussed in this paper, accompanied by graphical illustrations to visualize system size probabilities and performance indices.

Keywords: Heterogeneous service; Generating function; Continued fraction; Modified Bessel function, Time-dependent probabilities, Steady-state probabilities

1. Introduction

Queuing systems, fundamental to understanding the dynamics of service provision in various domains, have traditionally been modelled under the assumption of homogeneity, where service rates remain constant across servers. However, the real-world landscape presents a diverse array of scenarios where servers exhibit heterogeneous characteristics, ranging from differing capacities to varied processing speeds. This departure from homogeneity introduces complexities that demand novel modelling approaches to accurately capture system behaviours. In this paper, we delve into the realm of heterogeneous servers within queuing systems, focusing on the intricate interplay between server diversity and system resilience. Our investigation aims to address the challenges posed by system failures, a ubiquitous occurrence in service environments, by proposing a resilient model where servers seamlessly transition between primary and backup roles to ensure continuity of service provision. Specifically, we contribute to the literature by analyzing a single-server queuing system providing two types of service: fast and slow. Instead of halting service entirely during failure, our proposed model allows the server to transition into a backup role and continue providing service at a reduced rate, thus minimizing downtime and enhancing operational resilience.

Several authors have explored queuing systems with heterogeneous servers. For instance, Kumar and Madheswari [8] utilized a Markovian queue model to investigate a system featuring two servers with different characteristics and multiple vacation periods. Using the matrix geometric method, they determined the stationary queue length distribution and average system size for this setup. Krishnamoorthy and Sreenivasan [9] analyzed an M/M/2 queuing system

with two servers of different types. One server remains continuously available, while the other server goes on vacation when no customers are waiting for service. Upon returning from vacation, the second server operates at a reduced rate if the first server is already busy. The authors examined the system's behaviour in a steady state using the matrix geometric method.

Efrosinin and Rykov [5] analyzed a multi-server system with heterogeneous exponential queues. Their study demonstrates techniques for computing steady-state probabilities and deriving distributions for waiting and sojourn times. Efrosinin et al. [6] investigated a controllable multi-server heterogeneous queueing system in which servers operate at different service rates without preemption. Additionally, the authors have applied the concept of heterogeneity in service to cloud centres. Wang et al. [13] introduced the concept of heterogeneous servers in cloud centres to strike an optimal balance between expected response time and power consumption. By incorporating servers with varying capabilities, they aimed to efficiently handle stochastically arriving requests in cloud environments. From the literature survey, it is observed that many authors have focused on utilizing two servers to provide heterogeneous service, with both servers operating at different speeds. However, in this paper, we depart from this convention by considering a single server capable of providing two distinct services. For instance, imagine a modern banking system where a single ATM offers both cash withdrawal and deposit services, catering to the diverse needs of customers. This type of service is also applied in cloud computing. In a cloud computing platform, a single virtual machine instance may be tasked with handling both high-priority real-time data processing and lower-priority batch processing tasks. Additionally, while traditional heterogeneous server models assume a fixed arrival rate, our proposed model introduces heterogeneity in the arrival rate as well, reflecting real-world scenarios where incoming requests vary in frequency and urgency.

In service systems, customers often experience heterogeneous service, which can stem from various reasons. In this paper, we focus on addressing the challenges posed by system failures resulting from technical anomalies, a scenario ubiquitous in real-world service environments. System failures can occur due to several reasons such as negative customers [7], disaster ([3], [11]) and catastrophes [4]. Ammar [2] investigated the two-processor heterogeneous system with catastrophes, server failures and repairs. Sudhesh and Savitha studied three heterogeneous systems with catastrophes. From the literature survey, it is observed that many authors have considered that when a system encounters a disaster, all customers are removed from the system, and the system switches to a failure state. After the repair process, the server switches to an idle state and waits for customers to arrive.

In response to such disruptions, our proposed model incorporates a resilient mechanism wherein the primary server seamlessly transitions into a backup role whenever a failure occurs. During these periods of contingency, the backup server delivers service at a reduced rate, thereby mitigating the impact of disruptions on service provision and maintaining a degree of continuity for system users. Upon serving all customers in the system, the backup server switches to the maintenance state, initiating necessary repairs to restore the system to full functionality. This proactive approach to maintenance ensures the integrity and reliability of the system, minimizing downtime and enhancing overall operational resilience. By integrating these aspects into our queuing model, we aim to provide a comprehensive framework for analyzing and optimizing the performance of service-oriented systems under diverse operating conditions. The objective of this paper is to analyze a single-server queueing system where the server provides two types of service: fast and slow. Instead of halting service entirely during failure, the server transitions into a backup role and continues providing service at a reduced rate. Once all customers have been served, the backup server switches to a maintenance state. Following maintenance, the server returns to an idle state and waits for customers to arrive. To analyze this system, we derive both transient and steady-state probabilities using Laplace transform and generating function techniques.

This article is structured as follows: Section 2 presents the application of the proposed model. Section 3 provides the model description. The time-dependent probabilities of the system are discussed in Section 4, while Section 5 focuses on the performance measures of the system in the

transient state. In Section 6, the steady-state probabilities are presented, followed by a discussion on the performance indices of the system in the steady state in Section 7. A numerical illustration of the system is provided in Section 8, and Section 9 offers the conclusion of the proposed work.

2. Application of the proposed system

The proposed system is applied in Disaster Recovery Systems, which are crucial components of critical IT infrastructure such as data centres or cloud-based services where high availability is essential. A disaster recovery system ensures business continuity and data integrity in the face of unexpected events like hardware failures, natural disasters, or cyber-attacks. In this system, the main server is responsible for handling regular operations and serving client requests. Meanwhile, the backup server operates in a standby mode, continuously replicating data and configurations from the active server to ensure that it remains up-to-date with the latest data.

In the event of a system failure on the main server, the backup server automatically takes over the responsibilities of the main server in a process known as fail-over. This fail-over mechanism may be triggered either manually or automatically by monitoring systems that detect the failure of the main server. Once the main server is repaired and ready to operate again, it can resume its regular duties, and the data changes that occurred during the fail-over period can be synchronized back to the main server. The main server acting as a backup server in this context provides redundancy and enhances the overall reliability of the system. It ensures that critical services and applications remain available even during unexpected disruptions, thereby reducing downtime and minimizing the impact on end-users or customers.

3. Model Description

Consider a system that consists of a single server acting as the main server and also a backup server, providing different types of service. Whenever a failure occurs in the main server, the backup server acts as the main server but with a slower service rate, denoted by μ_2. Arrival occurs to the main server according to a Poisson process with rate λ_1, whereas arrivals occur with rate λ_2 when the backup server is active. Customers receive service at the main server with exponential rate μ_1, while the backup server has a reduced service rate μ_2, where $\mu_2 \leq \mu_1$. Assume that failures of the main server occur at an exponential rate γ. Once the backup server becomes idle, it promptly enters a state of preventive maintenance (state V), characterized by an exponentially distributed duration with a mean of $1/\xi$. Throughout the maintenance period, customers are prohibited from entering the system. The moment the server's maintenance is finished, it promptly transitions back to the primary processor and becomes prepared to attend the new customers.

Let $\{N(t), M(t) : t \geq 0\}$ be the 2-dimensional continuous time Markov chain. Let $\{N(t), t \geq 0\}$ denote the number of customers in the system at any time t and $\{M(t), t \geq 0\}$ represents the state of the system at any time t with state space

$$S = \{(0,0) \cup \{(n,r), n \in Z^+, r = 1,2\} \cup V\}.$$

The state $(0,0)$ represents that the server is idle and waiting for customers to arrive. The state $(n,1)$ represents the main server is busy and providing service to the n^{th} customer. The state $(n,2)$ represents the backup server is busy and providing service to the n^{th} customer. The state V represents the server is in a maintenance state and the server is inoperative in this state. Let $P_{n,r}(t) = P\{N(t) = n, M(t) = r\}$ be the probability that the server is in state r with n number of customers in the system at any time t and let $P_V(t)$ denote the probability that the server is in

maintenance state. Then $P_{n,r}(t)$ and $P_V(t)$ satisfies the following forward Kolmogorov equations

$$P_V^{'}(t) = -\xi P_V(t) + \mu_2 P_{1,2}(t), \tag{1}$$
$$P_{0,0}^{'}(t) = -\lambda_1 P_{0,0}(t) + \xi P_V(t) + \mu_1 P_{1,1}(t), \tag{2}$$
$$P_{1,1}^{'}(t) = -(\lambda_1 + \mu_1 + \gamma) P_{1,1}(t) + \lambda_1 P_{0,0}(t) + \mu_1 P_{2,1}(t), \tag{3}$$
$$P_{n,1}^{'}(t) = -(\lambda_1 + \mu_1 + \gamma) P_{n,1}(t) + \lambda_1 P_{n-1,1}(t) + \mu_1 P_{n+1,1}(t), n \geq 2, \tag{4}$$
$$P_{1,2}^{'}(t) = -(\lambda_2 + \mu_2) P_{1,2}(t) + \mu_2 P_{2,2}(t) + \gamma P_{1,1}(t), \tag{5}$$
$$P_{n,2}^{'}(t) = -(\lambda_2 + \mu_2) P_{n,2}(t) + \lambda_2 P_{n-1,2}(t) + \mu_2 P_{n+1,2}(t) + \gamma P_{n,1}(t), n \geq 2. \tag{6}$$

with the initial condition $P_{0,0}(0) = 1$.

4. Time-dependent probabilities

This section presents the time-dependent probabilities of the system being busy when the main server is active, denoted as $P_{n,1}(t)$, when the backup server is active, denoted as $P_{n,2}(t)$, during maintenance, denoted as $P_V(t)$, and in the idle state, denoted as $P_{0,0}(t)$.

4.1. Evaluation of $P_{n,1}(t)$

This section presents the time-dependent probability of the system being busy when the main server is active. Let $\widehat{P}_{n,r}(s)$ denote the Laplace transform of $P_{n,r}(t)$. Taking Laplace Transform on Equation (4) and rearranging, we get

$$\frac{\widehat{P}_{n,1}(s)}{\widehat{P}_{n-1,1}(s)} = \frac{\lambda_1}{(s + \lambda_1 + \mu_1 + \gamma) - \mu_1 \frac{\widehat{P}_{n+1,1}(s)}{\widehat{P}_{n,1}(s)}}.$$

On simplification, we obtain

$$\widehat{P}_{n,1}(s) = \beta_1 \left[\frac{p_1 - \sqrt{p_1^2 - \alpha_1^2}}{\alpha_1} \right] \widehat{P}_{n-1,1}(s).$$

The above equation recursively yields

$$\widehat{P}_{n,1}(s) = {\beta_1}^{(n-1)} \left[\frac{p_1 - \sqrt{p_1^2 - \alpha_1^2}}{\alpha_1} \right]^{(n-1)} \widehat{P}_{1,1}(s), \quad n \geq 2, \tag{7}$$

where

$$p_1 = s + \lambda_1 + \mu_1 + \gamma, \alpha_1 = 2\sqrt{\lambda_1 \mu_1},\ \beta_1 = \sqrt{\frac{\lambda_1}{\mu_1}}.$$

Taking inverse Laplace transform on Equation (7), we get

$$P_{n,1}(t) = \lambda_1 \beta_1^{n-2} e^{-(\lambda_1 + \mu_1 + \gamma)t} \left[I_{n-2}(\alpha_1(t-u)) - I_n(\alpha_1(t-u)) \right] * P_{1,1}(t), \tag{8}$$

where $I_n(t)$ represents modified Bessel function of first kind of order n. Thus the probability that the main server is busy $P_{n,1}(t)$ is expressed in terms of $P_{1,1}(t)$. The expression for $P_{1,1}(t)$ is presented in Equation (22)

4.2. Evaluation of $P_{n,2}(t)$

To obtain the time-dependent probability of $P_{n,2}(t)$, we define a generating function as follows. Let

$$G(z,t) = \sum_{n=1}^{\infty} P_{n,2}(t) z^n$$

Using Equations (5) and (6), we obtain

$$\frac{\partial}{\partial t} G(z,t) = \left[-(_2+\mu_2) + (_2 z + \frac{\mu_2}{z}) \right] G(z,t) + \gamma \sum_{n=1}^{\infty} P_{n,1}(t) z^n - \mu_2 P_{1,2}(t). \tag{9}$$

Solving Equation (9) yields,

$$\begin{aligned} G(z,t) = \gamma \int_0^t \sum_{n=1}^{\infty} P_{n,1}(u) z^n e^{-(_2+\mu_2)(t-u)} e^{-(_2 z + \frac{\mu_2}{z})(t-u)} du \\ - \mu_2 \int_0^t P_{1,2}(u) e^{-(_2+\mu_2)(t-u)} e^{-(_2 z + \frac{\mu_2}{z})(t-u)} du. \end{aligned} \tag{10}$$

Let

$$\alpha_2 = 2\sqrt{_2 \mu_2},\ \beta_2 = \sqrt{\frac{2}{\mu_2}}.$$

Then

$$e^{-(_2 z + \frac{\mu_2}{z})t} = \sum_{n=-\infty}^{\infty} (\beta_2 z)^n I_n(\alpha_2 t). \tag{11}$$

Using Equation (11) in Equation (10) and equating the coefficient of z^n, we arrive

$$\begin{aligned} P_{n,2}(t) = \gamma \int_0^t \sum_{m=1}^{\infty} P_{m,1}(u) e^{-(_2+\mu_2)(t-u)} \beta_2^{n-m} I_{n-m}(\alpha_2(t-u)) du \\ - \mu_2 \int_0^t P_{1,2}(u) e^{-(_2+\mu_2)(t-u)} \beta_2^n I_n(\alpha_2(t-u)) du. \end{aligned} \tag{12}$$

The above holds for $n \leq -1$ with the left-hand side replaced by zero. Using $I_{-n}(x) = I_n(x)$ for $n \geq 1$

$$\begin{aligned} 0 = \gamma \int_0^t \sum_{m=1}^{\infty} P_{m,1}(u) e^{-(_2+\mu_2)(t-u)} \beta_2^{-n-m} I_{n+m}(\alpha_2(t-u)) du \\ - \mu_2 \int_0^t P_{1,2}(u) e^{-(_2+\mu_2)(t-u)} \beta_2^{-n} I_n(\alpha_2(t-u)) du. \end{aligned} \tag{13}$$

From Equations(12) and (13), we get

$$P_{n,2}(t) = \gamma \int_0^t \sum_{m=1}^{\infty} P_{m,1}(u) e^{-(_2+\mu_2)(t-u)} \beta_2^{n-m} [I_{n-m}(\alpha_2(t-u)) - I_{n+m}(\alpha_2(t-u)) du. \tag{14}$$

4.3. Evaluation of $P_V(t)$ and $P_{0,0}(t)$

This section presents the time-dependent probabilities of the maintenance state and idle state. Taking Laplace transform on Equation (1), we obtain

$$\widehat{P}_V(s) = \frac{\mu_2}{s+\xi} \widehat{P}_{1,2}(s). \tag{15}$$

On inversion, we get

$$P_V(t) = \mu_2 e^{-\xi t} * P_{1,2}(t).$$

Taking Laplace transform on (2), we obtain

$$\widehat{P}_{0,0}(s) = \frac{1}{s+_1}\left[1 + \xi\widehat{P}_V(s) + \mu_1\widehat{P}_{1,1}(s)\right]. \tag{16}$$

On inversion, we have

$$P_{0,0}(t) = e^{-_1 t} * \left[\delta(t) + \xi P_V(t) + \mu_1 P_{1,1}(t)\right]. \tag{17}$$

Setting $n = 1$ in Equation (14) and taking Laplace transform, we get

$$\widehat{P}_{1,2}(s) = \widehat{\Phi}(s)\widehat{P}_{1,1}(s), \tag{18}$$

where

$$\widehat{\Phi}(s) = \frac{\gamma}{2}\sum_{m=1}^{\infty}\beta_1^{m-1}\beta_2^{2-m}\left(\frac{p_1 - \sqrt{p_1^2 - \alpha_1^2}}{\alpha_1}\right)^{m-1}\left(\frac{p_2 - \sqrt{p_2^2 - \alpha_2^2}}{\alpha_2}\right)^{m} \tag{19}$$

and

$$p_2 = s + \lambda_2 + \mu_2.$$

Inverting Equation (19), we get

$$\begin{aligned}\Phi(t) = \gamma\lambda_1 \sum_{m=1}^{\infty}\beta_1^{m-1}\beta_2^{1-m}e^{-(\lambda_1+\mu_1+\gamma)t}\left[I_{m-2}(\alpha_1 t) - I_m(\alpha_1 t)\right] * e^{-(\lambda_2+\mu_2)t} \\ \times \left[I_{m-1}(\alpha_2 t) - I_{m+1}(\alpha_2 t)\right].\end{aligned}$$

Taking Laplace Transform on (3), we get

$$\widehat{P}_{11}(s) = \frac{\lambda_1}{s + \lambda_1 + \mu_1 + \gamma}\widehat{P}_{0,0}(s) + \frac{\mu_1}{s + \lambda_1 + \mu_1 + \gamma}\widehat{P}_{2,1}(s). \tag{20}$$

Setting $n = 2$ in Equation (7) and using Equations (16), (15), (18) in Equation (20), after some algebra, we have

$$\widehat{P}_{1,1}(s) = \lambda_1\sum_{k=0}^{\infty}(\mu_1\beta_1)^k\sum_{r=0}^{k}\left(\frac{\lambda_1\mu_2}{\mu_1\beta_1}\right)^r\binom{k}{r}\frac{1}{(s+\lambda_1)^{r+1}}\left(\frac{p_1 - \sqrt{p_1^2 - \alpha_1^2}}{\alpha_1}\right)^{k-r}\sum_{j=0}^{r}\xi^j\binom{r}{j}\left(\frac{\widehat{\Phi}(s)}{s+\xi}\right)^j. \tag{21}$$

On inversion

$$\begin{aligned}P_{1,1}(t) = \frac{\lambda_1^2}{\beta_1}\sum_{k=0}^{\infty}(\mu_1\beta_1)^k\sum_{r=0}^{k}\left(\frac{\lambda_1\mu_2}{\mu_1\beta_2}\right)^r\binom{k}{r}e^{-\lambda_1 t}\frac{t^r}{r!} * e^{-(\lambda_1+\mu_1+\gamma)t}\left[I_{k-r-1}(\alpha_1 t) - I_{k-r+1}(\alpha_1 t)\right] \\ * \sum_{j=0}^{\infty}\xi^j\binom{r}{j}e^{-\xi t}\frac{t^{j-1}}{(j-1)!} * (\Phi(t))^{*j}.\end{aligned} \tag{22}$$

5. Performance Measures

In this section, the expected system size and variance of the proposed model are presented.

5.1. Expected system size

The expected system size, denoted as $E(N(t))$, is defined as follows.

$$E(N(t)) = \sum_{n=1}^{\infty} n \left(P_{n,1}(t) + P_{n,2}(t)\right)$$

Using Equations (3) − (6), we get

$$\frac{d}{dt}E[N(t)] = \lambda_1 P_{0,0}(t) + (\lambda_1 - \mu_1)\sum_{n=1}^{\infty} P_{n,1}(t) + (\lambda_2 - \mu_2)\sum_{n=1}^{\infty} P_{n,2}(t).$$

Integrating,

$$E(N(t)) = \lambda_1 \int_0^t P_{0,0}(u)du + \sum_{n=1}^{\infty}\left[\int_0^t (\lambda_1 - \mu_1)P_{n,1}(u)du + \int_0^t (\lambda_2 - \mu_2)P_{n,2}(u)du\right].$$

5.2. Variance

The variance of the number of customers at time t is defined as

$$V(N(t)) = E[N^2(t)] - E(N(t))^2$$

where

$$E[N^2(t)] = \sum_{n=1}^{\infty} n^2 \left[P_{n,1}(t) + P_{n,2}(t)\right]$$

Using Equations (3) − (6), we obtain

$$\frac{d}{dt}E[N^2(t)] = \lambda_1 P_{0,0}(t) + \sum_{n=1}^{\infty}\Big[\lambda_1(2n+1)P_{n,1}(t) + \mu_1(1-2n)P_{n,1}(t) + \lambda_2(2n+1)P_{n,2}(t) + \mu_2(1-2n)P_{n,2}(t)\Big].$$

Integrating,

$$E[N^2(t)] = \lambda_1 \int_0^t P_{0,0}(u)du + \sum_{n=1}^{\infty}\Big[\lambda_1(2n+1)\int_0^t P_{n,1}(u)du + \mu_1(1-2n)\int_0^t P_{n,1}(u)du + \lambda_2(2n+1)\int_0^t P_{n,2}(u)du + \mu_2(1-2n)\int_0^t P_{n,2}(u)du\Big].$$

where $P_{n,1}(t), P_{n,2}(t)$ *and* $P_{0,0}(t)$ are given in Equations (18), (14) and (17) respectively.

6. Stationary Analysis

This section presents the steady-state analysis of the proposed model. The steady-state equations of the proposed model are as follows.

$$0 = -\xi\pi_M + \mu_2\pi_{1,2},, \tag{23}$$
$$0 = -\lambda_1\pi_{0,0} + \xi\pi_M + \mu_1\pi_{1,1},, \tag{24}$$
$$0 = -\left(\lambda_1 + \mu_1 + \gamma\right)\pi_{1,1} + \lambda_1\pi_{0,0} + \mu_1\pi_{2,1},, \tag{25}$$
$$0 = -\left(\lambda_1 + \mu_1 + \gamma\right)\pi_{n,1} + \lambda_1\pi_{n-1,1} + \mu_1\pi_{n+1,1}, n = 2,3,4,..., \tag{26}$$
$$,0 = -\left(\lambda_2 + \mu_2\right)\pi_{1,2} + \mu_2\pi_{2,2} + \gamma\pi_{1,1},, \tag{27}$$
$$0 = -\left(\lambda_2 + \mu_2\right)\pi_{n,2} + \lambda_2\pi_{n-1,2} + \mu_2\pi_{n+1,2} + \gamma\pi_{n,1}, n = 2,3,4,...., \tag{28}$$

We define a generating function

$$G_i(z) = \sum_{n=1}^{\infty} \pi_{n,i} z^n, i = 1, 2.$$

Using Equations (25) and (26) and summing for $n = 1, 2, 3, ...$, we get

$$G_1(z) = \frac{z}{(z - z_1)(z - \bar{z}_1)} \{\mu_1 \pi_{1,1} - \lambda_1 \pi_{0,0} z\} \tag{29}$$

where

$$z_1 = \frac{(\lambda_1 + \mu_1 + \gamma) + \sqrt{(\lambda_1 + \mu_1 + \gamma)^2 - 4\lambda_1 \mu_1}}{2\lambda_1},$$

$$\bar{z}_1 = \frac{(\lambda_1 + \mu_1 + \gamma) - \sqrt{(\lambda_1 + \mu_1 + \gamma)^2 - 4\lambda_1 \mu_1}}{2\lambda_1}.$$

It is noted that for $\lambda_1 > 0$, $\mu_1 > 0$, $\gamma > 0$, the roots $z_1 > 1$, $0 < \bar{z}_1 < 1$. Setting $z = \bar{z}_1$ in Equation (29), we obtain

$$G_1(z) = \sum_{n=1}^{\infty} \left(\frac{z}{z_1}\right)^n \lambda_1 \pi_{0,0}$$

Comparing the coefficient of z^n in the above expression, we obtain

$$\pi_{n,1} = \lambda_1 \left(\frac{1}{z_1}\right)^n \pi_{0,0} \tag{30}$$

Similarly, using Equations (27) and (28) and summing for $n = 1, 2, 3, ...$, we get

$$G_2(z) = \frac{z\lambda_2}{(z\lambda_2 - \mu_2)(z - 1)} \{\mu_2 \pi_{1,2} - \gamma G_1(z)\} \tag{31}$$

Setting $z = 1$ in (31), after some algebraic manipulation, we get

$$G_2(z) = \frac{\gamma \lambda_1 \lambda_2 z}{\mu_2 \left(1 - \frac{\lambda_2}{\mu_2}\right)(1 - z)} \left[\sum_{n=1}^{\infty} \left(\frac{1}{z_1}\right)^n - \sum_{n=1}^{\infty} \left(\frac{z}{z_1}\right)^n\right] \pi_{0,0}$$

Using Equation (30) in the above expression and equating the coefficients of z^n on both sides, we get

$$\pi_{n,2} = \gamma \lambda_1 \left\{\sum_{i=1}^{\infty} \left(\frac{1}{z_1}\right)^i \sum_{m=1}^{n} \left(\frac{\lambda_2}{\mu_2}\right)^m - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \left(\frac{\lambda_2}{\mu_2}\right)^i \left(\frac{1}{z_1}\right)^j\right\} \pi_{0,0} \tag{32}$$

Setting $n = 1$ in the above result and using it in (23), we obtain

$$\pi_M = \frac{\gamma \lambda_1 \lambda_2}{\xi} \sum_{i=1}^{\infty} \left(\frac{1}{z_1}\right)^i \pi_{0,0}. \tag{33}$$

An explicit expression for $\pi_{0,0}$ can be obtained using the normalisation condition as follows.

$$\pi_M + \pi_{0,0+} \sum_{n=1}^{\infty} \pi_{n,1} + \sum_{n=1}^{\infty} \pi_{n,2} = 1. \tag{34}$$

Using the results (30), (32) and (33) in the above condition, we get

$$\pi_{0,0} = \left[1 + \frac{\gamma \lambda_1 \lambda_2}{\xi} \sum_{i=1}^{\infty} \left(\frac{1}{z_1}\right)^i + \gamma \lambda_1 \sum_{n=1}^{\infty} \left\{\sum_{i=1}^{\infty} \left(\frac{1}{z_1}\right)^i \sum_{m=1}^{n} \left(\frac{\lambda_2}{\mu_2}\right)^m - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \left(\frac{\lambda_2}{\mu_2}\right)^i \left(\frac{1}{z_1}\right)^j\right\} + \sum_{n=1}^{\infty} \lambda_1 \left(\frac{1}{z_1}\right)^n\right]^{-1}.$$

7. Performance indices

This section presents the expected system size of the proposed model

7.1. Expected system size

Let $E(N_s)$, $E(N_1)$ and $E(N_2)$ denote the expected number of customers in the system, main server and the backup server respectively.

$$E\left(N_S\right) = E\left(N_1\right) + E\left(N_2\right).$$

Using the result (30) and (31), we get

$$E\left(N_1\right) = \frac{\lambda_1 z_1}{\left(1-z_1\right)^2}\pi_{0,0},$$

$$E\left(N_2\right) = \gamma\lambda_1 \sum_{n=1}^{\infty} n \left\{ \sum_{i=1}^{\infty} \left(\frac{1}{z_1}\right)^i \sum_{m=1}^{n} \left(\frac{\lambda_2}{\mu_2}\right)^m - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \left(\frac{\lambda_2}{\mu_2}\right)^i \left(\frac{1}{z_1}\right)^j \right\} \pi_{0,0}.$$

Applying Little's formula, the expected number of customers waiting in the system and the queue is given by

$$E\left(W_s\right) = \frac{1}{\lambda_1} E\left(N_1\right) + \frac{1}{\lambda_2} E\left(N_2\right)$$

$$, E\left(W_q\right) = \sum_{n=1}^{\infty} \left(n-1\right) \pi_{n,1} + \sum_{n=1}^{\infty} \left(n-1\right) \pi_{n,2}.$$

8. Numerical illustration

In this section, we provide a numerical illustration of our proposed model. The parameter values are chosen based on the stability conditions $\frac{\lambda_1}{\mu_1} < 1$ and $\frac{\lambda_2}{\mu_2} < 1$. The parameter values are as follows: $\lambda_1 = 0.6$, $\lambda_2 = 0.5$, $\mu_1 = 1.1$, $\mu_2 = 1$, $\gamma = 0.3$, and $\xi = 0.1$. Figures 1 and 2 depict the behaviour of the main server $P_{1,n}(t)$ and the backup server $P_{2,n}(t)$, respectively. We assumed that the initial condition $P_{0,0}(0) = 1$. As a result, the probability curve of $P_{1,n}(t)$ starts at 1 and gradually decreases until it reaches the steady state. Conversely, all other probability curves for $P_{1,n}(t)$ begin at zero, increase initially, and converge to the steady state. Figures 3 and 4 showcase the expected system size and variance of the system for varying values of the arrival rate λ_1. We observe that as the arrival rate increases, the mean and variance graphs also increase. Figures 5 and 6 show the expected system size and variance for different values of the arrival rate λ_2. Figures 7-10 display the stationary probabilities of the system. Figures 7 and 8 provide insights into the probabilities associated with the main and backup servers, respectively. From the graphs, it is observed that as n increases, the probability curves of $\pi_{n,1}$ and $\pi_{n,2}$ decrease and attain the steady state. Finally, Figures 9 and 10 demonstrate the expected system size in the main and backup servers. We notice that as the arrival rate increases, the expected system size $E(N_i)$, where $i = 1, 2$, for both the main and backup servers also increases. This provides important insights into the system's performance under different workload scenarios.

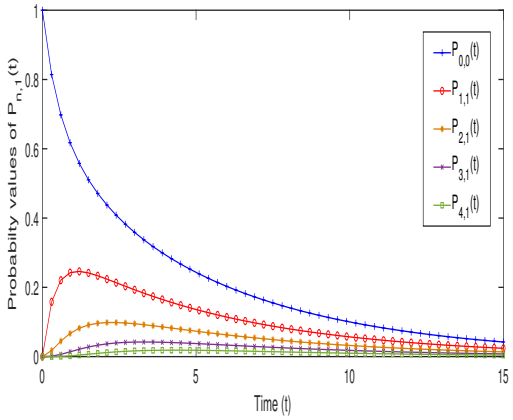


Figure 1: *Probabilities of the main server* $P_{1,n}(t)$.

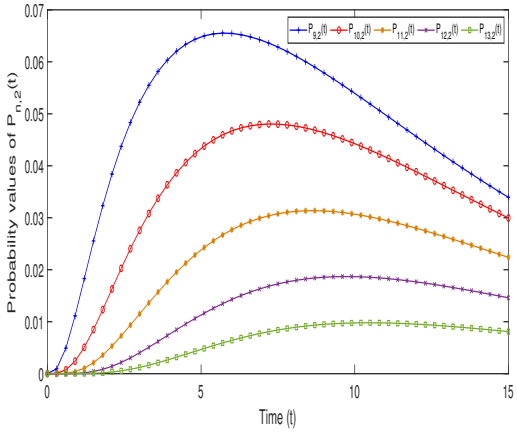


Figure 2: *Probabilities of the backup server* $P_{2,n}(t)$.

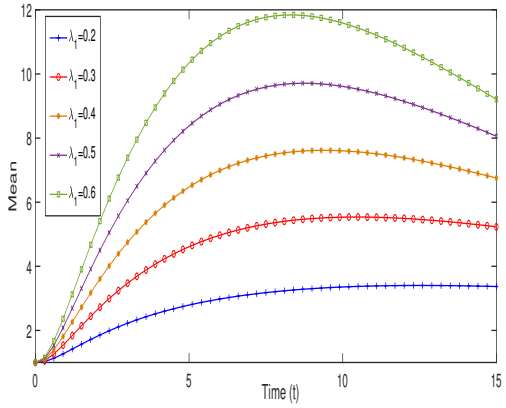


Figure 3: *Mean system size for different arrival rate* λ_1.

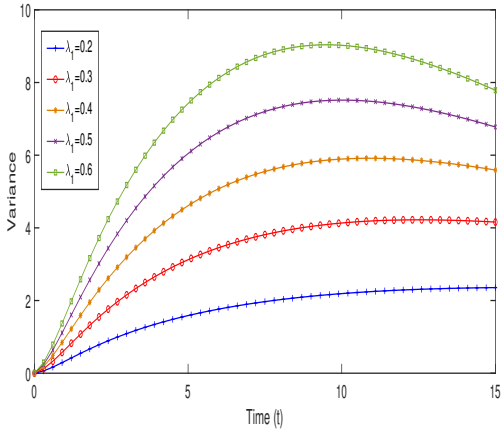


Figure 4: *Variance of the system for different* λ_1.

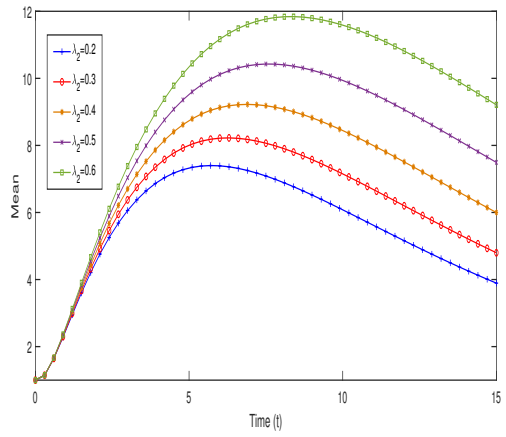


Figure 5: *Mean system size for different* λ_2.

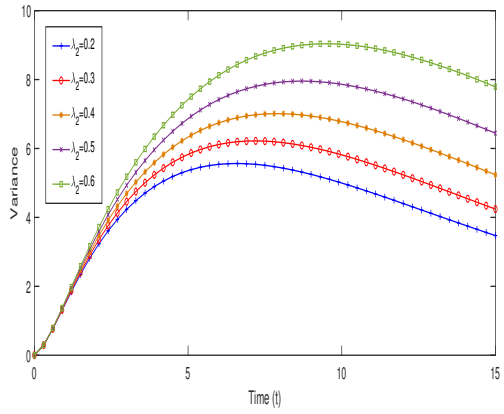


Figure 6: *Variance of the system for different* λ_2.

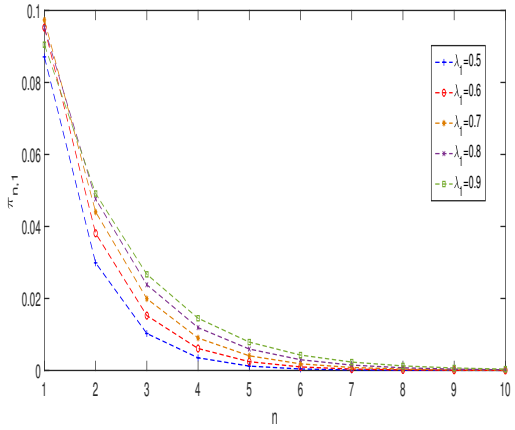


Figure 7: *Steady state probability $\pi_{n,1}$ for different arrival rate λ_1.*

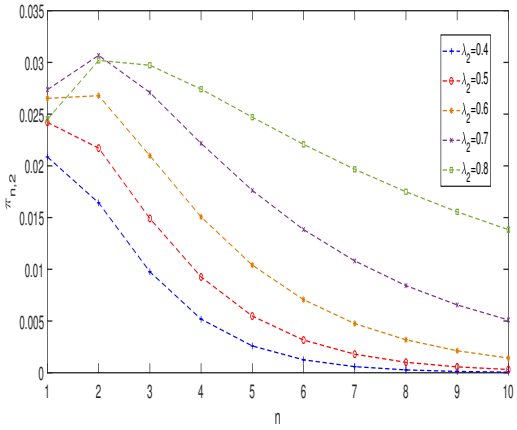


Figure 8: *Steady state probability $\pi_{n,2}$ for different arrival rate λ_2.*

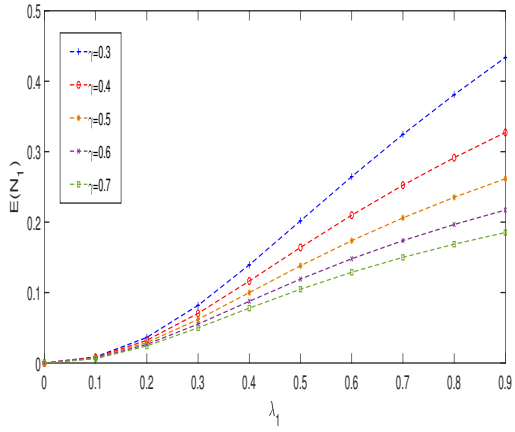


Figure 9: *Mean system size $E(N_1)$ against λ_1 for various γ rates*

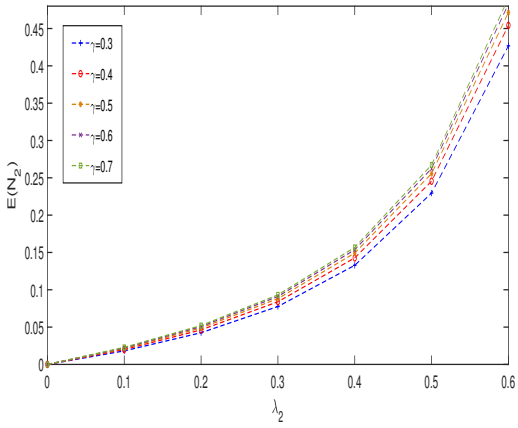


Figure 10: *Mean system size $E(N_2)$ against λ_2 for various γ rates.*

9. Conclusion

This paper investigates an M/M/1 queueing system with heterogeneous service rates and periodic server maintenance. By deriving explicit expressions for both the transient and steady-state probabilities, the study provided a comprehensive understanding of the system's performance under various operating conditions. The establishment of the mathematical framework and the utilization of analytical techniques were instrumental in achieving the desired analysis. The current study focused on a single server setup. One can extend this work by investigating multi-server configurations

References

[1] Anisimov, V., Artalejo, R (2001). Analysis of Markov multi-server retrial queues with negative arrivals, Queueing Systems, 39:157 - 182.

[2] Ammer, I (2014). Transient behaviour of a two-processor heterogeneous system with catastrophes, server failures and repairs, Applied Mathematical Modelling, 38: 2224 - 2234.

[3] Boxma, O. J., Perry, D., & Stadje, W. (2001). Clearing models for M/G/1 queues. Queueing Systems, 38: 287-306.

[4] Di Crescenzo, A., Giorno, V., Nobile, A. G., & Ricciardi, L. M. (2003). On the M/M/1 queue with catastrophes and its continuous approximation. Queueing Systems, 43: 329-347.
[5] Efrosinin, D. V., & Rykov, V. V. E. (2008). On performance characteristics for queueing systems with heterogeneous servers. Automation and Remote Control, 69(1): 61-75.
[6] Efrosinin, D., Stepanova, N., Sztrik, J., & Plank, A. (2020). Approximations in performance analysis of a controllable queueing system with heterogeneous servers. Mathematics, 8(10): 1803.
[7] Gelenbe, E. (1991). Product-form queueing networks with negative and positive customers. Journal of Applied Probability, 28(3): 656-663.
[8] Kumar, B. K., Madheswari, S. P (2005) An M/M/2 queueing system with heterogeneous servers and multiple vacations. Mathematical and Computer Modelling, 41(13): 1415-1429.
[9] Krishnamoorthy, A., Sreenivasan, C (2012). An M/M/2 queueing system with heterogeneous servers including one with working vacation. International Journal of Stochastic Analysis, 2012: 1-16.
[10] Sudhesh, R., Mohammed Shapique, A., Dharmaraja, S (2022). Analysis of a Multiple Dual'Stage Vacation Queueing System with Disaster and Repairable Server. Methodology and Computing in Applied Probability, 24(4): 2485-2508.
[11] Sudhesh, R., Vaithiyanathan, A (2019). Analysis of state-dependent discrete-time queue with system disaster. RAIRO-Operations Research, 53(5): 1915-1927.
[12] Sudhesh, R., & Savitha, P. (2017). Transient behaviour of three-heterogeneous servers queue with system disaster and server repair. RAIRO-Operations Research, 51(4): 965-983.
[13] Wang, S., Li, X., & Ruiz, R. (2019). Performance analysis for heterogeneous cloud servers using queueing theory. IEEE Transactions on Computers, 69(4): 563-576.

RT&A, No 2 (78)
Volume 19, June, 2024

ON CERTAIN CLASSES OF CONFORMALLY FLAT LORENTZIAN PARA-KENMOTSU MANIFOLDS

K. L. Sai Prasad[1], P. Naveen[2] and S. Sunitha Devi[3,*]

•

Department of Mathematics
[1] Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam, 530 048, INDIA
[2] Rajiv Gandhi University of Knowledge Technology, Srikakulam -532402, INDIA
[3,*] KL University, Vijayawada, Andhra Pradesh, 520002, INDIA

klsprasad@yahoo.com[1] potnuru123@gmail.com[2] sunithamallakula@yahoo.com[3,*]

Abstract

In this present paper, we classify and explore the geometrical significance of a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly LP-Kenmotsu) manifolds whenever the manifolds are either conformally flat or conformally symmetric. It was found that a conformally flat LP-Kenmotsu manifold is of constant curvature and a conformally symmetric LP-Kenmotsu manifold is locally isomorphic to a unit sphere. At the end, we obtain the scalar curvature of ϕ-conformally flat LP-Kenmotsu manifolds.

Keywords: Lorentzian para-Kenmotsu manifold, Weyl-conformal curvature tensor, Riemannian curvature tensor, ϕ-conformal curvature tensor, η-Einstein manifold.

2010 Mathematics Subject Classification: 53C05, 53C07, 53C15

I. Introduction

In 1995, Sinha and Sai Prasad [11] defined a class of almost paracontact metric manifolds namely para-Kenmotsu (briefly *P*-Kenmotsu) and special para-Kenmotsu (briefly *SP*-Kenmotsu) manifolds in similar to *P*-Sasakian and *SP*-Sasakian manifolds. In 1989, K. Matsumoto [3] introduced the notion of Lorentzian paracontact and in particular, Lorentzian para-Sasakian (*LP*-Sasakian) manifolds. Later, these manifolds have been widely studied by many geometers such as Matsumoto and Mihai [4], Mihai and Rosca [5], Mihai, Shaikh and De [6], Venkatesha and Bagewadi [13], Venkatesha, Pradeep Kumar and Bagewadi [14, 15].

In 2018, Abdul Haseeb and Rajendra Prasad defined a class of Lorentzian almost paracontact metric manifolds namely Lorentzian para-Kenmotsu (briefly *LP*-Kenmotsu) manifolds [1, 2] and they studied ϕ-semisymmetric *LP*-Kenmotsu manifolds with a quarter-symmetric non-metric connection admitting Ricci solitons [7, 8]. As an extension, Sai Prasad *et al*., [9] have studied *LP*-Kenmotsu manifolds admitting the Weyl-projective curvature tensor of type (1, 3). Further, they also have studied and shown that the *LP*-Kenmotsu manifolds admitting both irrotational and conservative pseudo-projective curvature tensors are Einstein manifolds of constant scalar curvature [10].

In 2023, Sunitha and Sai Prasad [12] have defined a class of Lorentzian para-Kenmotsu manifolds admitting a quarter-symmetric metric connection and shown that it is either ϕ-symmetric

or concircular ϕ-symmetric with respect to quarter-symmetric metric connection if and only if it is symmetric with respect to the Riemannian connection, provided the scalar curvature of Riemannian connection is constant. Recently, Rao, Sunitha and Sai Prasad [16] have studied ϕ-conharmonically flat and ϕ-projectively flat *LP*-Kenmotsu manifolds. They have shown that ϕ-conharmonically flat *LP*-Kenmotsu manifold is an η-Einstein manifold with zero-scalar curvature and ϕ-projectively flat *LP*-Kenmotsu manifold is an Einstein manifold with the scalar curvature $r = n(n-1)$.

In this work we explore a class of conformally flat Lorentzian para-Kenmotsu (*LP*-Kenmotsu) manifolds. The following is the layout of the current paper: Following the introduction, Section 2 includes some preliminaries on Lorentzian para-Kenmotsu manifolds. In section 3, we study conformally flat Lorentzian para-Kenmotsu manifolds and shown that they are of constant curvature. Further in section 4, we study and have shown that Lorentzian para-Kenmotsu manifold satisfying the condition $R(X,Y).C = 0$ is locally isomorphic to a unit sphere $S^n(1)$. Finally in section 5, it is shown that ϕ-conformally flat *LP*-Kenmotsu manifold is an η-Einstein manifold with the scalar curvature $r = n(n-1)$.

II. Preliminaries

An n-dimensional differentiable manifold M_n admitting a $(1,1)$ tensor field ϕ, contravariant vector field ξ, a 1-form η and the Lorentzian metric $g(X,Y)$ satisfying

$$\eta(\xi) = -1, \tag{1}$$

$$\phi^2 X = X + \eta(X)\xi, \tag{2}$$

$$g(\phi X, \phi Y) = g(X,Y) + \eta(X)\eta(Y), \tag{3}$$

$$g(X,\xi) = \eta(X), \tag{4}$$

$$\phi\xi = 0, \;\; \eta(\phi X) = 0, \;\; rank\ \phi = n-1 \tag{5}$$

is called Lorentzian almost paracontact manifold [3].

In a Lorentzian almost paracontact manifold, we have

$$\Phi(X,Y) = \Phi(Y,X), \qquad where \qquad \Phi(X,Y) = g(\phi X, Y). \tag{6}$$

A Lorentzian almost paracontact manifold M_n is called Lorentzian para-Kenmotsu manifold if [1]

$$(\nabla_X \phi)Y = -g(\phi X, Y)\xi - \eta(Y)\phi X, \tag{7}$$

for any vector fields X and Y on M_n and ∇ is the operator of covariant differentiation with respect to the Lorentzian metric g.

It can be easily seen that in a *LP*-Kenmotsu manifold M_n, the following relations hold [1]:

$$\nabla_X \xi = -\phi^2 X = -X - \eta(X)\xi, \tag{8}$$

$$(\nabla_X \eta)Y = -g(X,Y)\xi - \eta(X)\eta(Y), \tag{9}$$

for any vector fields X and Y on M_n.

Also, in an *LP*-Kenmotsu manifold, the following relations hold [1]:

$$\begin{aligned} g(R(X,Y)Z,\xi) &= \eta(R(X,Y)Z) \\ &= g(Y,Z)\eta(X) - g(X,Z)\eta(Y), \end{aligned} \tag{10}$$

$$R(\xi, X)Y = g(X, Y)\xi - \eta(Y)X, \tag{11}$$

$$R(X, Y)\xi = \eta(Y)X - \eta(X)Y, \tag{12}$$

$$S(X, \xi) = (n-1)\eta(X) \tag{13}$$

$$S(\phi X, \phi Y) = S(X, Y) + (n-1)\eta(X)\eta(Y), \tag{14}$$

$$S(X, Y) = ag(X, Y) + b\eta(X)\eta(Y), \tag{15}$$

for any vector fields X, Y and Z and S is the Ricci tensor of M_n.

The Riemannian Christoffel curvature tensor R of type $(1,3)$ is given by:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, \tag{16}$$

where ∇ be its Levi-Civita connection.

III. *LP*-Kenmotsu manifolds with $C(X, Y)Z = 0$

In this section, we consider conformally flat Lorentzian para-Kenmotsu manifolds.

The Weyl-conformal curvature tensor $C(X, Y)Z$ is given by

$$\begin{aligned} C(X, Y)Z = {} & R(X, Y)Z \\ & - \frac{1}{(n-2)}[g(Y, Z)QX - g(X, Z)QY + S(Y, Z)X - S(X, Z)Y] \\ & + \frac{r}{(n-1)(n-2)}[g(Y, Z)X - g(X, Z)Y], \end{aligned} \tag{17}$$

where

$$S(X, Y) = g(QX, Y).$$

Using (16), we get from (17)

$$\begin{aligned} R(X, Y)Z = {} & \frac{1}{(n-2)}[g(Y, Z)QX - g(X, Z)QY + S(Y, Z)X - S(X, Z)Y] \\ & + \frac{r}{(n-1)(n-2)}[g(Y, Z)X - g(X, Z)Y]. \end{aligned} \tag{18}$$

By taking $Z = \xi$ in (18) and on using (4), (12) and (13), we get

$$\begin{aligned} \eta(Y)X - \eta(X)Y = {} & \frac{1}{(n-2)}[\eta(Y)QX - \eta(X)QY] + \frac{(n-1)}{(n-2)}[\eta(Y)X - \eta(X)Y] \\ & - \frac{r}{(n-1)(n-2)}[\eta(Y)X - \eta(X)Y]. \end{aligned} \tag{19}$$

Taking $Y = \xi$ and using (1), we get

$$QX = \Big(\frac{1}{n-1} - 1\Big)X + \Big(\frac{r}{n-1} - 1\Big)\eta(X)\xi. \tag{20}$$

It shows that the manifold is η-Einstein.

Further on contracting (20), we have

$$r = n(n-1). \tag{21}$$

Now, by using (21) in (20), we get

$$QX = (n-1)\,X. \tag{22}$$

Then by putting (22) in (19), we get

$$R(X,Y)Z = g(Y,Z)X - g(X,Z)Y. \tag{23}$$

Thus, a conformally flat *LP*-Kenmotsu manifold is of constant curvature. The value of this constant is +1. Hence, we can state

Theorem 1. A conformally flat *LP*-Kenmotsu manifold is locally isometric to a unit sphere $S^n(1)$.

IV. *LP*-Kenmotsu manifold satisfying $R(X,Y).C = 0$

Using (4), (11) and (13) we find from (17) that

$$\begin{aligned}\eta(C(X,Y)Z) = \frac{1}{n-2}[(\frac{r}{n-1}-1)(g(Y,Z)\eta(X) - g(X,Z)\eta(Y))\\ - (S(Y,Z)\eta(X) - S(X,Z)\eta(Y))].\end{aligned} \tag{24}$$

Putting $Z = \xi$ in (24) and on using (4), (13) we get

$$\eta(C(X,Y)\xi) = 0. \tag{25}$$

Again, taking $X = \xi$ in (24), we get

$$\begin{aligned}\eta(C(\xi,Y)Z) = \frac{1}{n-2}[S(Y,Z) + (n-1)\eta(Y)\eta(Z)]\\ - \frac{1}{n-2}\Big(\frac{r}{n-1}-1\Big)[g(Y,Z) + \eta(Y)\eta(Z)].\end{aligned} \tag{26}$$

Now,

$$\begin{aligned}(R(X,Y)C)(U,V)W = R(X,Y)C(U,V)W - C(R(X,Y)U,V)W\\ - C(U,R)(X,Y)V)W - C(U,V)R(X,Y)W.\end{aligned} \tag{27}$$

Using $R(X,Y).C = 0$, we find from above that

$$\begin{aligned}&g[R(\xi,Y)C(U,V)W,\xi] - g[C(R(\xi,Y)U,V)W,\xi]\\ &- g[C(U,R(\xi,Y)V)W,\xi] - g[C(U,V)R(\xi,Y)W,\xi] = 0.\end{aligned}$$

Using (4) and (11) we get

$$\begin{aligned}&- C(U,V,W,Y) - \eta(Y)\eta(C(U,V)W) - g(Y,U)\eta(C(\xi,V)W)\\ &+ \eta(U)\eta(C(Y,V)W) - g(Y,V)\eta(C(U,\xi)W) + \eta(V)\eta(C(U,Y)W)\\ &- g(Y,W)\eta(C(U,V)\xi) + \eta(W)\eta(C(U,V)Y) = 0,\end{aligned} \tag{28}$$

where

$$C(U,V,W,Y) = g(C(U,V)W,Y).$$

Putting $U = Y$ in (28), we get

$$\begin{aligned}&- C(U,V,W,U) - \eta(U)\eta(C(U,V)W) + \eta(U)\eta(C(U,V)W)\\ &+ \eta(V)\eta(C(U,U)W) + \eta(W)\eta(C(U,V)U) - g(U,U)\eta(C(\xi,V)W)\\ &- g(U,V)(C(U,\xi)W) - g(U,W)\eta(C(U,V)\xi) = 0.\end{aligned} \tag{29}$$

Let $\{e_i: i = 1,\ldots,n\}$ be an orthonormal basis of the tangent space at any point, then the sum for $1 \le i \le n$ of the relations (29) for $U = e_i$ gives

$$(1-n)\eta(C(\xi,V)W) = 0,$$

which implies

$$\eta(C(\xi,V)W)=0 \ \ as \ \ n>3. \tag{30}$$

Using (25) and (30), (28) takes the form

$$\begin{aligned}&-C(U,V,W,Y)-\eta(Y)\eta(C(U,V)W)+\eta(U)\eta(C(Y,V)W)\\&+\eta(V)\eta(C(U,Y)W)+\eta()\eta(C(U,V)Y)=0.\end{aligned} \tag{31}$$

Using (24) in (31) we get

$$\begin{aligned}-C(U,V,W,Y)+\eta(W)\frac{1}{n-2}\Big[\Big(\frac{r}{n-1}-1\Big)\big(\eta(U)g(V,Y)-\eta(V)g(U,Y)\big)\\-\big(\eta(U)S(V,Y)-\eta(V)S(U,Y)\big)\Big]=0.\end{aligned} \tag{32}$$

In virtue of (30), (26) reduces to

$$S(Y,Z)=\Big(\frac{r}{n-1}-1\Big)g(Y,Z)+\Big(\frac{r}{n-1}-n\Big)\eta(Y)\eta(Z). \tag{33}$$

Using (33), (31) reduces to

$$C(U,V,W,Y)=0, \tag{34}$$

which proves that the manifold is conformally flat. Hence, by using the Theorem 1, we state

Theorem 2. If in an *LP*-Kenmotsu manifold $M_n(n>3)$ the relation $R(X,Y).C=0$ holds, then it is locally isometric with a unit sphere $S^n(1)$.

For a conformally symmetric Riemannian manifold, we have $\nabla C=0$. Hence for such a manifold $R(X,Y).C=0$ holds. Thus, we have the following corollary of the above theorem.

Corollary 1. A conformally symmetric *LP*-Kenmotsu manifold M_n $(n>3)$ is locally isometric with a unit sphere $S^n(1)$.

V. ϕ-conformally flat *LP*-Kenmotsu manifold

Let C be the Weyl conformal curvature tensor of M_n. Since at each point $p\in M_n$ the tangent space $T(M_n)$ can be decomposed into the direct sum $T_p(M_n)=\phi\left(T_p(M_n)\right)\oplus L\left(\xi_p\right)$, where $L\left(\xi_p\right)$ is a 1-dimensional linear subspace of $T_p(M_n)$ generated by ξ_p, we have a map:

$$\mathrm{C}: T_p(M_n)\times T_p(M_n)\times T_p(M_n)\to\varphi\left(T_p(M_n)\right)\oplus L\left(\xi_p\right)$$

It may be natural to consider the following particular cases:

1. C:$T_p(M_n)\times T_p(M_n)\times T_p(M_n)\to L\left(\xi_p\right)$, that is, the projection of the image of C in $\phi\left(T_p(M_n)\right)$ is zero.
2. C:$T_p(M_n)\times T_p(M_n)\times T_p(M_n)\to\phi\left(T_p(M_n)\right)$, that is, the projection of the image of C in $L\left(\xi_p\right)$ is zero.
3. C:$\phi\left(T_p(M_n)\right)\times\phi\left(T_p(M_n)\right)\times\phi\left(T_p(M_n)\right)\to L\left(\xi_p\right)$, that is, when C is restricted to $\left(T_p(M_n)\right)\times\phi\left(T_p(M_n)\right)\times\varphi\left(T_p(M_n)\right)$, the projection of the image of C in $\phi\left(T_p(M_n)\right)$ is zero. This condition is equivalent to

$$\phi^2C(\phi X,\varphi Y)\phi Z=0. \tag{35}$$

Definition 1. A differentiable manifold (M_n, g), $n>3$, satisfying the condition (35) is called ϕ-conformally flat.

Now our aim is to find the characterization of *LP*-Kenmotsu manifolds satisfying the condition (35).

Theorem 3. Let M_n be an n-dimensional, $(n>3)$, ϕ-conformally flat *LP*-Kenmotsu manifold. Then M_n is an η-Einstein manifold.

Proof. Suppose that (M_n, g), $n>3$, is a ϕ-conformally flat $LP-$ Kenmotsu manifold. It is easy to see that $\phi^2 C(\phi X, \phi Y)\phi Z = 0$ holds if and only if $g(C(\phi X, \phi Y)\phi Z, \varphi W) = 0$ for any $X, Y, Z, W \in \chi(M_n)$. So, by the use of (17), ϕ-conformally flat means

$$\begin{aligned} g(R(\phi X, \phi Y)\phi Z, \phi W) &= \frac{1}{n-2}[g(\phi Y, \phi Z)S(\phi X, \phi W) - g(\phi X, \phi Z)S(\phi Y, \phi W) \\ &\quad + g(\phi X, \phi W)S(\phi Y, \phi Z) - g(\phi Y, \phi W)S(\phi X, \phi Z)] \\ &\quad - \frac{r}{(n-1)(n-2)}[g(\phi Y, \phi Z)g(\phi X, \phi W) \\ &\quad - g(\phi X, \phi Z)g(\phi Y, \phi W)]. \end{aligned} \tag{36}$$

Let $\{e_1, \ldots, e_{n-1}, \xi\}$ be a local orthonormal basis of vector fields in M_n. Using that $\{\phi e_1, \ldots, \phi e_{n-1}, \xi\}$ is also a local orthonormal basis, if we put $X=W=e_i$ in (36) and sum up with respect to i, then

$$\begin{aligned} \sum_{i=1}^{n-1} g(R(\phi e_i, \phi Y)\phi Z, \phi e_i) &= \frac{1}{n-2}\sum_{i=1}^{n-1}[g(\phi Y, \phi Z)S(\phi e_i, \phi e_i) \\ &\quad - g(\phi e_i, \phi Z)S(\phi Y, \phi e_i) + g(\phi e_i, \phi e_i)S(\phi Y, \phi Z) \\ &\quad - g(\phi Y, \phi e_i)S(\phi e_i, \phi Z)] \\ &\quad - \frac{r}{(n-1)(n-2)}\sum_{i=1}^{n-1}[g(\phi Y, \phi Z)g(\phi e_i, \phi e_i) \\ &\quad - g(\phi e_i, \phi Z)g(\phi Y, \phi e_i)]. \end{aligned} \tag{37}$$

It can be easily verified that

$$\sum_{i=1}^{n-1} g(R(\phi e_i, \phi Y)\phi Z, \phi e_i) = S(\phi Y, \phi Z) + g(\phi Y, \phi Z), \tag{38}$$

$$\sum_{i=1}^{n-1} S(\phi e_i, \phi e_i) = r+n-1, \tag{39}$$

$$\sum_{i=1}^{n-1} g(\phi e_i, \phi Z)\, S(\phi Y, \phi e_i) = S(\phi Y, \phi Z), \tag{40}$$

$$\sum_{i=1}^{n-1} g(\phi e_i, \phi e_i) = n+1, \tag{41}$$

and

$$\sum_{i=1}^{n-1} g(\phi e_i, \phi Z)\, g(\phi Y, \varphi e_i) = g(\phi Y, \phi Z). \tag{42}$$

So, by virtue of (38)-(42) the equation (37) can be written as

$$S(\phi Y, \phi Z) = \left(\frac{r}{n-1} - 1\right) g(\phi Y, \phi Z). \tag{43}$$

Then by making use of (3) and (14), the equation (43) takes the form

$$S(Y,Z) = \left(\frac{r}{n-1} - 1\right) g(Y,Z) + \left(\frac{r}{n-1} - n\right) \eta(Y)\eta(Z). \tag{44}$$

Therefore from (44), by contraction, we obtain

$$r = n(n-1). \tag{45}$$

Then by substituting (45) in (44), we get

$$S(Y,Z) = (n-1)g(Y,Z)$$

which implies M_n is an η-Einstein manifold with the scalar curvature $r = n(n-1)$.

This completes the proof of the theorem.

VI. Conclusion

The present work explores the geometrical significance of a new class of Lorentzian paracontact metric manifolds namely the Lorentzian para-Kenmotsu manifolds whenever these manifolds are either conformally symmetric or conformally flat. The concepts and various geometrical properties of these manifolds can be applied in various aspects of Applied Mathematics such as Computational Fluid Dynamics, in designing the Super Resolution Sensors in Communications Engineering, and also in the field of General Theory of Relativity.

Acknowledgements: The authors acknowledge Dr. A. Kameswara Rao, Assistant Professor of G.V.P. College of Engineering for Women for his valuable suggestions in preparation of the manuscript.

Conflicts of interest: The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Abdul Haseeb and Rajendra Prasad. Certain results on Lorentzian para-Kenmotsu manifolds. Bulletin of Parana's Mathematical Society, (2018), 201-220.
[2] A. Friedmann and J. A. Schouten. Uber die Geometric der halbsymmetrischen Ubertragung. Math. Zeitschr, 21 (1924), 211-223.
[3] K. Matsumoto. On Lorentzian Paracontact manifolds. Bulletin of the Yamagata University Natural Science, 12 (2)(1989), 151-156.
[4] K. Matsumoto and I. Mihai. On a certain transformation in a Lorentzian para-Sasakian manifold. Tensor, N.S., 47(1988), 189-197.
[5] I. Mihai and R. Rosca. On Lorentzian P-Sasakian manifolds. Classical Analysis, World Scientific Publ., Singapore,(1992), 155-169.
[6] I. Mihai, A. A. Shaikh, U.C. De. On Lorentzian para-Sasakian manifolds. Rendicontidel Seminario Mathematico de Messina, Serie II (1999), 75-77.
[7] Rajendra Prasad, Abdul Haseeb and Umesh Kumar Gautam. On ϕ-semisymmetric LP Kenmotsu manifolds with a QSNM-connection admitting Ricci solitons. Kragujevac Journal of Mathematics, 45(5)(2021), 815-827.
[8] Rajendra Prasad, Shashikant Pandey and Abdul Haseeb. On a Lorentzian Sasakian manifold endowed with a quarter-symmetric metric connection. Sciendo, Seria Matematica, Informatica, LVII, 2 (2019), 61-76.
[9] K. L. Sai Prasad, S. Sunitha Devi and G. V. S. R. Deekshitulu. On a class of Lorentzian para-Kenmotsu manifolds admitting the Weyl-projective curvature tensor of type (1, 3). Italian Journal of Pure and Applied Mathematics, 45(2021), 990-1001.

[10] K. L. Sai Prasad, S. Sunitha Devi and G. V. S. R. Deekshitulu. On a class of Lorentzian paracontact metric manifolds. Italian Journal of Pure and Applied Mathematics, 49(2023), 514-527.
[11] B. B. Sinha and K. L. Sai Prasad. A class of almost paracontact metric Manifold. Bulletin of the Calcutta Mathematical Society, 87(1995), 307-312.
[12] S. Sunitha Devi and K. L. Sai Prasad. On a class of Lorentzian para-Kenmotsu manifolds admitting quarter-symmetric metric connection. Reliability Theory and Applications, 18(2023), no.4(76), 525-532.
[13] Venkatesha and C. S. Bagewadi, On concircular ϕ-recurrent LP-Sasakian manifolds. Differ. Geom. Dyn. Syst., 10(2008), 312-319.
[14] Venkatesha, K. T. Pradeep Kumar and C. S. Bagewadi. On Lorentzian para-Sasakian manifolds satisfying W_2 curvature tensor. IOSR J. of Mathematics, 9(2014), no. 6, 124-127.
[15] Venkatesha, K. T. Pradeep Kumar and C. S. Bagewadi. On quarter-symmetric metric connection in a Lorentzian para-Sasakian manifold. Ajerbaijan Journal of Mathematics, 5(2015), no. 1, 3-12.
[16] I. V. Venkateswara Rao, S. Sunitha Devi and K. L. Sai Prasad. On ϕ-conharmonically flat Lorentzian para-Kenmotsu manifolds. Reliability Theory and Applications, 19(2024), no.1(77), 764-769.

RT&A, No 2 (78)
Volume 19, June, 2024

ESTIMATION OF HAZARD AND SURVIVAL FUNCTION FOR COMPETING RISKS USING KERNEL AND MIXTURE MODEL IN BIMODAL SETUP

[1]A. M. Rangoli, [2]A. S. Talawar

•

[1,] Research Scholar, Department of Statistics, Karnatak University, Dharwad. India.
ajmr1008@gmail.com
[2] Professor, Department of Statistics, Karnatak University, Dharwad. India.
astalawar@kud.ac.in

Abstract

Aim of the present paper is to find suitable model for bimodal data. We have modelled mixture of two Weibull distributions in the presence of competing risks and also used Epanechnikov kernel to estimate hazard and survival functions. We considered prostate cancer data for application of the mixture model and kernel. We used maximum likelihood estimation (MLE) to estimate parameters of the mixture model, as the equations have no closed form, so we considered expectation–maximization (EM) algorithm. The mixture model and kernel gave good fit to the bimodal data. The prostate cancer data consists of three causes, we have estimated hazard function for these three causes using mixture model and kernel. The asymptotic confidence interval for the parameters of mixture model to all three causes were estimated. Also compared survival curve of mixture model with kernel and Kaplan-Meier survival curves for all the three causes.

Keywords: Weibull mixture model, EM algorithm, kernel, hazard, bimodal.

I. Introduction

The general Statistical analyses were different from survival analysis because of the presence of censoring. Basically, censoring means incomplete data. In survival analysis or medical studies, it is quite common that more than one cause of failure may be directed to a subject at the same time. It is required for an investigator to estimate a specific risk in the presence of other risk factors. In statistical literature, this process is known as the analysis of competing risks model. It is assumed, in the analysis of competing risks model, that data consist of a time to failure and an indicator denoting the cause of failure. In survival analysis our main objective is to estimate survival and hazard functions. Survival analysis can be done using parametric, non-parametric and Bayesian methods. For parametric approach we generally consider Weibull distribution because it has increasing, decreasing and constant hazard rate, but this distribution can be used when data is

unimodal. When we have bimodal data we cannot use the standard parametric lifetime distributions, in that case we can use mixture of distributions. For nonparametric approach we can consider the kernel method to estimate the hazard and survival functions.

Many authors work on mixture of distributions and kernel based methods to estimate hazard and survival functions. Modelling of mixture of gamma, mixture of Weibull and mixture of log normal distributions for analysing the heterogeneous survival data was considered by [1], for that they have considered mice data and Lung cancer dataset. The mixture of two and three Weibull distributions was modelled and estimated the parameters using MLE and tested for best fit of the models by [2], and used five different examples to show the hazard and survival functions. A parametric mixture model of three different distributions was used to analyse heterogeneous survival data by [3]. They have simulated the data and estimated the parameters using expectation-maximization (EM) algorithm and also compared individual distribution like exponential, gamma and Weibull with the mixture of these three distributions. Similarly, many authors have worked on mixture models ([4], [5], [6]). Estimation of the hazard function and its associated factors in gastric cancer patients using Wavelet and kernel smoothing methods was carried out by [7]. Repeated time to event models to characterize the repeated occurrence of clinical events and visualization of kernel based hazard with comparison to Weibull and Gompertz models was considered by [8]. (see also [9], [10]).

In present paper, we are considering estimation of density, hazard and survival functions using Epanechnikov kernel and mixture of two Weibull distributions. For estimation we are consider the prostate cancer data which is bimodal given in [11]. Generally, for bimodal data it is not appropriate to use standard parametric lifetime distributions but mixtures of those distributions are suitable for bimodal. Here we are considering two cases, case-I consists of estimation of hazard and survival function using mixture of two Weibull distributions and case-II considers kernel method of estimation.

II. Methods

2.1 Case-I: Mixture of two Weibull distribution

Now we are considering a parametric approach using mixture of two Weibull distributions. The study considers fitting of bimodal data to the mixture of Weibull distributions in presence of competing risks and calculation of the hazard and survival functions. The functional form of mixture distribution is given below.

Let $T_1, T_2, \dots, T_n$ be the failure time of n patients where $T_i \in (0, t]$, $i = 1,2, \dots, n$ if we consider k competing events then $T_{ij} \rightarrow failure\ time\ of\ i^{th} patient\ with\ j^{th} cause$. Each patient fail due to only one cause $T_i = \min(T_{i1}, T_{i2}, \dots, T_{ik})$. Let us consider C be the censoring time such that $T_{ij} = \min(T_{ij}, C)$. Let $F(t)$ be the cumulative distribution function (CDF) and $f(t)$ be probability distribution function and $h(t)$ and $S(t)$ be hazard and survival function at time t.

$$F(t) = 1 - (\pi_1 e^{-\alpha t^{\gamma}} + \pi_2 e^{-\beta t^{\lambda}}) \quad (1)$$

$$f(t) = \pi_1 \alpha\gamma t^{\gamma-1} e^{-\alpha t^{\gamma}} + \pi_2 \beta\lambda t^{\lambda-1} e^{-\beta t^{\lambda}} \quad (2)$$

$$S(t) = 1 - F(t)$$

$$S(t) = \pi_1 e^{-\alpha t^{\gamma}} + \pi_2 e^{-\beta t^{\lambda}} \quad (3)$$

$$h(t) = \frac{f(t)}{S(t)} = \frac{\pi_1 \alpha\gamma t^{\gamma-1} e^{-\alpha t^{\gamma}} + \pi_2 \beta\lambda t^{\lambda-1} e^{-\beta t^{\lambda}}}{\pi_1 e^{-\alpha t^{\gamma}} + \pi_2 e^{-\beta t^{\lambda}}} \quad (4)$$

Here $\pi_1$and π_2 be the weights such that $\pi_1 + \pi_2 = 1$

And α and γ be scale and shape parameters with weight $\pi_1$and β and λ be scale and shape parameters with weight π_2.

Likelihood function L is given as

$$L = \prod_{i=1}^{n} f(x_i)$$

$$L = \prod_{i=1}^{n} (\pi_1 \alpha\gamma t_i^{\gamma-1} e^{-\alpha t_i^{\gamma}} + \pi_2 \beta\lambda t_i^{\lambda-1} e^{-\beta t_i^{\lambda}})$$

Now the likelihood function in terms of competing risks can be given as

$$\delta_{ij} = \begin{cases} 1 & i^{th} subject\ fail\ due\ to\ j^{th}\ cause, \quad j = 1,2, \dots \dots, k \\ 0 & i^{th} subject\ does\ not\ fail\ due\ to\ j^{th} cause\ (Censored) \end{cases}$$

$$L = \prod_{i=1}^{n} \prod_{j=1}^{k} (f_j(t_i))^{\delta_{ij}} S(t_i)^{1-\delta_{ij}}$$

$$\mathrm{L} = \prod_{i=1}^{n} \prod_{j=1}^{k} \left(h_j(t_i)\right)^{\delta_{ij}} S(t_i) \quad (5)$$

Now the log likelihood of equation (5) can be written as,

$$l = logL = \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\left(\delta_{ij} * \log\left(h_j(t_i)\right) \right) + \log\left(S(t_i)\right) \right)$$

$$l = \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\left(\delta_{ij} * \log\left(\frac{\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}}}{\pi_1 e^{-\alpha_j t_i^{\gamma_j}} + \pi_2 e^{-\beta_j t_i^{\lambda_j}}} \right) \right) + \log\left(\pi_1 e^{-\alpha_j t_i^{\gamma_j}} + \pi_2 e^{-\beta_j t_i^{\lambda_j}} \right) \right)$$

$$l = \sum_{i=1}^{n} \sum_{j=1}^{k} \left(\left(\delta_{ij} \left(\log\left(\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}} \right) - \log\left(\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \right. \right. \right. \right.$$

$$\left. \left. \left. \left. \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}} \right) \right) \right) + \log\left(\pi_1 e^{-\alpha_j t_i^{\gamma_j}} + \pi_2 e^{-\beta_j t_i^{\lambda_j}} \right) \right) \quad (6)$$

For cause j we can write log likelihood as

$$l_j = \sum_{i=1}^{n_j} \left(\log\left(\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}} \right) \right) - \sum_{i=1}^{n_j} \log\left(\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}} \right) + \sum_{i=1}^{n} \log\left(\pi_1 e^{-\alpha_j t_i^{\gamma_j}} + \pi_2 e^{-\beta_j t_i^{\lambda_j}} \right)$$

(7)

Now estimating the parameter values using MLE, and it is obtained by first order partial derivatives with respect to each parameter and equating to zero. But these equations do not have closed form, so to estimate the parameter values we consider numerical estimation that is Newton-Raphson method or we can use expectation–maximization (EM) algorithm. The first order partial derivatives are given in Appendix.

2.1.1 Expectation–Maximization (EM) Algorithm

The EM algorithm is a powerful iterative method used for finding maximum likelihood estimates or maximum a posteriori estimates in statistical models where the data is incomplete or contains latent variables. The EM algorithm consists of an-expectation step (E-step), and a maximization step (M-step). The advantage of the EM algorithm is that it solves a difficult incomplete-data problem by constructing two easy steps. The E-step only needs to compute the conditional expectation of the log-likelihood with respect to the incomplete data, given the observed data. The M-step needs to find the maximizer of this expected likelihood. An additional advantage of this method compared to other optimization techniques is that it is very simple, and it converges reliably [12]. Let Y be the observed data and X be the missing data. We also write l, l_c and l_m for the log-likelihoods based on the observed, complete and missing data distributions respectively. The EM algorithm consists of iterating two steps. First is the expectation, or "E", step, in which an objective function is constructed from the complete data likelihood. Second is the maximization, or "M", step, in which the previously computed objective function is maximized. These two steps are then alternated until some convergence criterion is met [13]. Whatever value of θ the algorithm converges to and is used as our parameter estimate.

The E-step of the EM algorithm consists of computing the conditional expectation of the complete data likelihood, given the observed data. That is, the objective function at iteration k is given by

$$Q(\theta|\theta_{k-1}) = E_{\theta_{k-1}}(l_c(\theta; y, X)|Y = y) \tag{8}$$

Where θ_{k-1} is the parameter estimate obtained from the previous iteration.

The M-step of the EM algorithm consists of maximizing the objective function constructed in the previous E-step. That is, we define $\theta_k = argmax_\theta Q(\theta|\theta_{k-1})$. Typically, this optimization must be performed numerically via, e.g., gradient ascent or the Newton-Raphson algorithm. In fact, it is possible to divide the set of parameters into groups (possibly with each group containing a single parameter) and optimize over each group individually with the others held fixed. This is called the Expectation-Conditional Maximization, or ECM, algorithm. Notationally, we can combine the E and M-steps of the EM algorithm into a single "update function". We write $M(\theta_{k-1}) = argmax_\theta Q(\theta|\theta_{k-1})$. The EM algorithm can thus be viewed as the iterative application of this update function, M.

2.1.2 Asymptotic Confidence Bounds

The MLE's do not have closed form to know the distribution to calculate confidence intervals, in such a case we go with asymptotic distribution of the MLE of the parameters [14]. It is known that the asymptotic distribution of the MLE $\hat{\theta}$ is given by

$$\left(\hat{\theta} - \theta\right) \rightarrow N_4(0, I^{-1}(\theta))$$

Where $I^{-1}(\theta) \rightarrow$ Fisher information matrix of the unknown parameters

$$\theta = (\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2, \lambda_3)\ .$$

The elements of the 4 X 4 matrix of $I^{-1}(.)$, are approximated by $I_{ij}\left(\hat{\theta}\right)$,

where

$$I_{ij}(\hat{\theta}) = -\frac{\partial^2 l(\theta)}{\partial\theta_i \partial\theta_j}\Big|_{\theta=\hat{\theta}}$$

Where, $\hat{\theta} = (\widehat{\alpha_1}, \widehat{\alpha_2}, \widehat{\alpha_3}, \widehat{\beta_1}, \widehat{\beta_2}, \widehat{\beta_3}, \widehat{\gamma_1}, \widehat{\gamma_2}, \widehat{\gamma_3}, \widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\lambda_3})$ estimated parameters.

Now information matrix can be written as,

$$\frac{\partial^2 l(\theta)}{\partial\theta_i \partial\theta_j} = \begin{bmatrix} \frac{\partial^2 l}{\partial\,\alpha_j^2} & \frac{\partial^2 l}{\partial\alpha_j \partial\beta_j} & \frac{\partial^2 l}{\partial\alpha_j \partial\gamma_j} & \frac{\partial^2 l}{\partial\alpha_j \partial\lambda_j} \\ \frac{\partial^2 l}{\partial\alpha_j \partial\beta_j} & \frac{\partial^2 l}{\partial\,\beta_j^2} & \frac{\partial^2 l}{\partial\beta_j \partial\gamma_j} & \frac{\partial^2 l}{\partial\beta_j \partial\lambda_j} \\ \frac{\partial^2 l}{\partial\alpha_j \partial\gamma_j} & \frac{\partial^2 l}{\partial\beta_j \partial\gamma_j} & \frac{\partial^2 l}{\partial\,\gamma_j^2} & \frac{\partial^2 l}{\partial\gamma_j \partial\lambda_j} \\ \frac{\partial^2 l}{\partial\alpha_j \partial\lambda_j} & \frac{\partial^2 l}{\partial\beta_j \partial\lambda_j} & \frac{\partial^2 l}{\partial\gamma_j \partial\lambda_j} & \frac{\partial^2 l}{\partial\,\lambda_j^2} \end{bmatrix} \quad (9)$$

The elements of the fisher information matrix are given in Appendix.

Therefore, the approximate $100(1-\gamma)\%$ two-sided, confidence interval for θ is given by

$$\hat{\theta} \pm Z_{\gamma/2}\sqrt{I^{-1}(\hat{\theta})} \quad (10)$$

Here $Z_{\gamma/2}$is the upper $\gamma/2$ th percentile of a standard normal distribution.

2.2 Case-II: Kernel density Estimation

A kernel is a weight function of observation on x and scaling parameter h which is called as the Bandwidth. The scaled distances obtained at a point x is used to compute kernel density at that point. The kernel density function is regarded as probability density [8]. The estimator to estimate density is given by,

$$\hat{f}(u) = \frac{1}{n}\sum_{i=1}^{n} K(u)$$

Where, $K(.) \rightarrow$ Kernel Function
The kernel function has the following properties viz,.

$$K(u) \geq 0, \text{for all u}$$

$$\int K(u)du = 1 \text{ (normalization)}$$

$$K(-u) = K(u) \text{ (symmetry)}$$

$$\int u\,K(u)\,du = 0$$

$$\text{And } \int u^2\,K(u) \neq 0$$

Using the kernels we can estimate the survival and hazard functions. In general, to obtaining pattern for rate of failure the hazard curve is more obvious than survival curve. Hazard rate functions can be used for several statistical analysis in medicine, engineering and economics. For instance, hazard function commonly used when presenting results in clinical trials involving survival data. Several methods for hazard function estimation have been considered in the literature

([7], [9],[10]). Hazard function estimation by nonparametric methods has an advantage in flexibility because no formal assumptions are made about the mechanism that generates the sample order or the randomness [15]. There are many kernels in the literature say, Uniform, Triangle, Epanechnikov, Quartic, Triweight, Gaussian, Cosine etc. Now for our study we consider the Epanechnikov kernel [16].

$$K(u) = \frac{3}{4}(1-u^2)\mathrm{I}(|u| \leq 1)$$

By considering $u = \frac{t-X_i}{b}$

Then kernel becomes

$$K_b(t) = \frac{3}{4b}\left(1-\left(\frac{t-X_i}{b}\right)^2\right)\mathrm{I}(|\frac{t-X_i}{b}| \leq 1) \quad (11)$$

Where, b is the bandwidth , n is the number of observation, X_i is the given observation and t is the point where kernels are calculated.

From this kernel many bumps are formed and summing the bumps gives us the density function.

The kernel density function is given by,

$$\widehat{f_b}(t) = \frac{1}{nb}\sum_{i=1}^{n} K\left(\frac{t-X_i}{b}\right)$$

$$\widehat{f_b}(t) = \frac{1}{nb}\sum_{i=1}^{n}\frac{3}{4}\left(1-\left(\frac{t-X_i}{b}\right)^2\right)\mathrm{I}\left(\left|\frac{t-X_i}{b}\right| \leq 1\right)$$

The estimation of the CDF, $\widehat{F_b}$ is constructed by integrating $\widehat{f_b}$. That is

$$\widehat{F_b}(t) = \int_{-\infty}^{t}\widehat{f_b}(x)\,dx = \frac{1}{n}\sum_{i=1}^{n}\overline{K}(\frac{t-X_i}{b})$$

Where, $\overline{K}(t) = \int_{-\infty}^{t} K(x)dx$

Estimation of hazard function using kernel [17],

$$\hat{h}(t) = \frac{1}{b}\sum_{i=1}^{n} K\left(\frac{t-X_i}{b}\right)\Delta\widehat{\Lambda}(t_i) \quad (12)$$

Where, n is the number of failure times, b is the bandwidth, $\widehat{\Lambda}(t)$ is the Nelson-Aalen estimator of the cumulative hazard function.

2.2.1 Nelson-Aalen estimator of cumulative hazard function

Let the hazard function be,

$$h(\mathrm{t}) = \lim_{l\to 0}\frac{1}{l}P(t \leq T < t+l/T \geq t) = \frac{f(t)}{S(t)}$$

where $f(t)$ be density function and $S(t)$ be survival function, and the survival function in terms of hazard function can be expressed as,

$$S(t) = e^{-\int_0^t h(u)du}$$

Now the cause specific hazard function be given by,

$$h_j(\mathrm{t}) = \lim_{l \to 0} \frac{1}{l} p(t \leq T < t + l, J = j \,/ T \geq t)$$

$$h(t) = \sum_{j=1}^{J} h_j(t)$$

And the cumulative cause specific hazard function be given by

$$\Lambda(t) = \int_0^t h(u)du$$

And cause specific Nelson-Aalen estimator of the cumulative hazard [18] is given by,

$$\Lambda_{\mathrm{j}}(t) = \sum_{k=1}^{K} \frac{Number\ of\ individuals\ observed\ to\ fail\ due\ to\ cause\ j\ at\ t_k}{Number\ of\ individuals\ at\ risk\ just\ prior\ to\ t_k}$$

2.2.2 Selection of the Bandwidth

Important thing in the kernel density estimation is selection of the bandwidth. We calculate bandwidth using Silverman's Rule [19]. That is

$$\hat{b} = \frac{1.06 * \hat{\sigma}}{n^{\frac{1}{5}}} \tag{13}$$

Where, $\hat{\sigma}$ is the sample standard deviation.

2.2.3 Kaplan-Meier (K-M) Estimator

The Kaplan-Meier estimator known as the product limit estimator is a non-parametric statistic used to estimate the survival function from lifetime data [20]. An important benefit of the Kaplan–Meier curve is that, the method can take into account some types of censored data, particularly right-censoring, which occurs if a patient withdraws from a study, or is lost due to follow-up, or is alive without event incidence at last follow-up. The Kaplan-Meier estimate is an easiest way of computing survival over time. The Kaplan Meier estimator of survival function is defined as

$$\hat{S}(t) = \prod_{i:t_i < t} (1 - \frac{d_i}{n_i})$$

Where $t_i$is the failure time, $d_i$is the number of events that occurs at time $t_i$and n_i is the number individuals at risk of experiencing the event immediately prior to t_i.

III. Results

The prostate cancer data consists of 489 patients, 25.5% of them are failed due to cancer, 19% failed due to CVD and 25.74% failed due to other causes and rest of the data were censored. Median failure

time for cancer patients is 23 months and for CVD patients 20.5 months and for other causes 24 months.

From figure 1 we can see that our data is bimodal, to fit this data, we used mixture of two Weibull distributions (black line) and the kernel density estimation (red line). Figure 2 explains the hazard curves for three different causes cancer, CVD and other causes. Here we can see that for cause cancer and CVD the hazard initially increases till 30 months then decreases. For other causes the hazard increases-decreases-increases, so we can say that, the hazard function is non-monotonic. Figure 3 explains the survival curve of three causes using kernel, mixture model and Kaplan-Meier survival functions. Here we can observe that kernel and mixture model survival curves are close to each other. For all three causes Kaplan-Meier survival curve has less probability of surviving as compared to the kernel and mixture model. Table 1 shows the estimated parameter values using MLE by considering EM algorithm. Table 2 gives the estimated parameter values with their corresponding standard error and confidence limits.

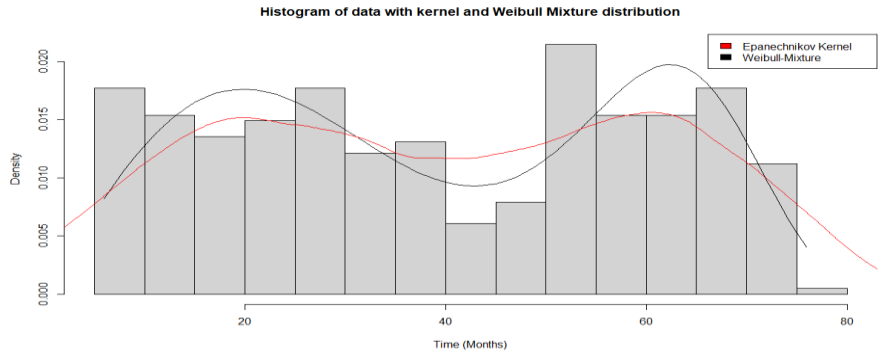


Figure 1: *Histogram of the data with fitted Mixture of Weibull distribution and Epanechmikov kernel.*

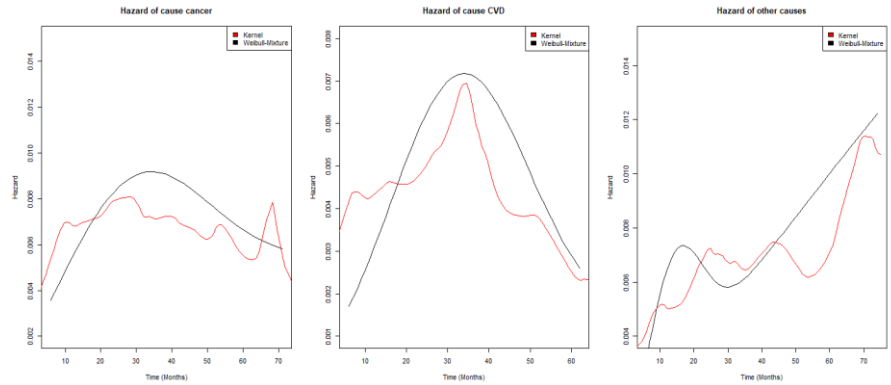


Figure 2: *Hazard curves using mixture model and kernel for three causes*

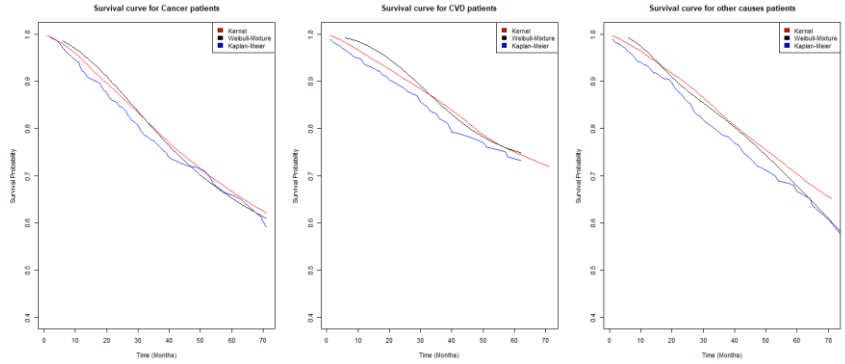


Figure 3*: Survival curves using mixture model, kernel and Kaplan-Meier for three causes*

Table 1: *Estimated parameter values for three causes*

Parameter	Cancer	CVD	Other Causes
π_1	0.8155268	0.810495	0.9224448
π_2	0.1844732	0.1895035	0.07755524
α	0.001494359	0.001097538	0.000103001
β	0.000370462	0.000076065	0.000848062
γ	1.241094000	1.058928000	1.954219000
λ	2.180196000	2.600613000	2.469943000

Table 2: *Estimated parameter values, standard error (SE), lower control limit (LCL) and upper control limit (UCL)*

Causes	Parameters	SE	LCL	UCL
Cancer	α =0.001494359	0.000059337	0.001378061	0.001610657
	β =0.000370462	0.000006754	0.000357225	0.000383700
	γ =1.241094000	0.004894528	1.231500902	1.250687098
	λ =2.180196000	0.000420388	2.179372055	2.181019945
CVD	α =0.001097538	0.000139156	0.000824797	0.001370279
	β =0.000076065	0.000001373	0.000073374	0.000078756
	γ=1.058928000	0.000000469	1.058927081	1.058928919
	λ =2.600613000	0.000187876	2.600244771	2.600981229
Other causes	α =0.000103001	0.000003077	0.000096969	0.000109032
	β =0.000848062	0.000022916	0.000803148	0.000892976
	γ =1.954219000	0.000666012	1.952913641	1.955524359
	λ =2.469943000	0.011053927	2.448277701	2.491608299

IV. Discussions

From the study we conclude that, in real life situations with competing risks data, if data is bimodal we can use mixture of distributions or kernel methods to estimate hazard and survival functions. In this paper we have considered both approach to estimate hazard and survival function in presence of the competing risks**.** To estimate the parameters of the mixture models we have used MLE as it does not have closed form so we considered EM algorithm to estimate parameters of all three causes. We have also calculated the standard error and asymptotic confidence interval for all the parameters. All the estimated parameters are statistically significant at 5% level of significance. Here we can see that hazard function initially increases, then decreases and increases. For survival curve we have compared kernel, mixture model and Kaplan-Meier methods. So, when we have bimodal density, and having competing risks approach, the mixture model (as a parametric approach) is more appropriate or kernel method (as a nonparametric approach) is more appropriate to estimate hazard and survival functions.

Acknowledgment: The first author is thankful to Department of Science and Technology, innovation in science pursuit for inspired research (DST-INSPIRE) for financial support (Fellowship/2021/210203).

References

[1] Erişoğlu, Ü., Erişoğlu, M., and Erol, H. (2011). A mixture model of two different distributions approach to the analysis of heterogeneous survival data. *International Journal of Computational and Mathematical Sciences*, *5*(2), 75-79.

[2] Razali, A. M., and Al-Wakeel, A. A. (2013). Mixture Weibull distributions for fitting failure times data. *Applied Mathematics and Computation*, *219*(24), 11358-11364.

[3] Mohammed, Y. A., Yatim, B., and Ismail, S. (2013). A simulation study of a parametric mixture model of three different distributions to analyze heterogeneous survival data. *Modern Applied Science*, *7*(7), 1-9.

[4] Elmahdy, E. E. (2015). A new approach for Weibull modeling for reliability life data analysis. *Applied Mathematics and computation*, *250*, 708-720.

[5] Larson, M. G., and Dinse, G. E. (1985). A mixture model for the regression analysis of competing risks data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *34*(3), 201-211.

[6] Enogwe, S. U., Okereke, E. W., and Ibeh, G. C (2023). A Bimodal Extension of Suja Distribution with Applications. *Statistics and Applications* 21(2), pp 155-173.

[7] Ahmadi, A., Roudbari, M., Gohari, M. R., and Hosseini, B. (2012). Estimation of hazard function and its associated factors in gastric cancer patients using wavelet and kernel smoothing methods. *Asian Pacific journal of cancer prevention*, 13(11), 5643-5646.

[8] Goulooze, S. C., Välitalo, P. A., Knibbe, C. A., and Krekels, E. H. (2018). Kernel-based visual hazard comparison (kbVHC): a simulation-free diagnostic for parametric repeated time-to-event models. *The AAPS journal*, *20*, 1-11.

[9] Hess, K. R., Serachitopol, D. M., and Brown, B. W. (1999). Hazard function estimators: a simulation study. *Statistics in medicine*, *18*(22), 3075-3088.

[10] Klein, J. P., and Bajorunaite, R. (2003). Inference for competing risks. *Handbook of statistics*, *23*, 291-311.

[11] Andrews, D. F., and Herzberg, A. M. (2012). *Data: a collection of problems from many fields for the student and research worker*. Springer Science and Business Media.

[12] Park, C. (2005). Parameter estimation of incomplete data in competing risks using the EM algorithm. *IEEE Transactions on Reliability*, 54(2), 282-290.

[13] Ruth, W. (2024). A review of Monte Carlo-based versions of the EM algorithm. arXiv preprint arXiv:2401.00945.

[14] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, New York.

[15] Klein, J. P., and Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (Vol. 1230). New York: Springer.

[16] Guedes, D. G. P., Cunha, E. E., & Lima, G. F. C. (2017). Genetic evaluation of age at first calving from Brown Swiss cows through survival analysis. *Archivos de zootecnia*, *66*(254), 247-255.

[17] Heisey, D. M., and Patterson, B. R. (2006). A review of methods to estimate cause-specific mortality in presence of competing risks. *The Journal of Wildlife Management*, *70*(6), 1544-1555.

[18] Beyersmann, J., Allignol, A., and Schumacher, M. (2011). *Competing risks and multistate models with R*. Springer Science and Business Media.

[19] Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

[20] Kaplan, E. L. and Meier, P. (1958). Non-parametric estimation from incomplete observation. *Journal of American Statistical Association*, 53, 457-481.

Appendix:

The log likelihood function (7) for the mixture of two Weibull distribution in presence of competing risks is given as

$$l_j = \sum_{i=1}^{n_j} \left(\log\left(\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}} \right) \right) -$$

$$\sum_{i=1}^{n_j} \log\left(\pi_1 \alpha_j \gamma_j t_i^{\gamma_j - 1} e^{-\alpha_j t_i^{\gamma_i}} + \pi_2 \beta_j \lambda_j t_i^{\lambda_j - 1} e^{-\beta_j t_i^{\lambda_j}} \right) + \sum_{i=1}^{n} \log\left(\pi_1 e^{-\alpha_j t_i^{\gamma_j}} + \pi_2 e^{-\beta_j t_i^{\lambda_j}} \right)$$

Here we are considering first with cause 1, that is $G1$ stands for failure times for cause 1, so here $j = 1$. Similarly, we can consider cause 2 and 3 as G2 and G3 respectively.

$ea = e^{-(\alpha * (G1^{\gamma}))}$; $eb = e^{-\left(\beta * (G1^{\lambda})\right)}$; $eaa = e^{-(\alpha * (x^{\gamma}))}$; $ebb = e^{-\left(\beta * (x^{\lambda})\right)}$;

$logx2 = log(G1^2)$; $x2g = G1^{2*\gamma}$; $x2l = G1^{2*\lambda}$; $xg1 = G1^{\gamma - 1}$; $xl1 = G1^{\lambda - 1}$

$$deno1 = (p1 * \alpha * \gamma * xg1 * ea) + (p2 * \beta * \lambda * xl1 * eb)$$

$$deno2 = p1 * ea + p2 * eb$$

$$deno3 = p1 * eaa + p2 * ebb$$

$$nume1 = p1 * \gamma * xg1 * ea * \left(1 - \alpha * (G1^{\gamma})\right)$$

$$nume2 = p1 * (G1^{\gamma}) * ea$$

$$nume3 = p1 * (x^{\gamma}) * eaa$$

$$numeb1 = p2 * \lambda * xl1 * eb * \left(1 - \beta * \left(G1^{\lambda}\right)\right)$$

$$numeb2 = p2 * \left(G1^{\lambda}\right) * eb$$

$$numeb3 = p2 * \left(x^{\lambda}\right) * ebb$$

$$numeg1 = p1 * \alpha * xg1 * ea * \left(\gamma * log(G1) + 1 - \alpha * g * (G1^{\gamma}) * log(G1)\right)$$

$$numeg2 = p1 * \alpha * (G1^{\gamma}) * log(G1) * ea$$

$$numeg3 = p1 * \alpha * (x^{\gamma}) * log(x) * eaa$$

$$numel1 = p2 * b * xl1 * eb * \left(\lambda * log(G1) + 1 - \beta * \lambda * \left(G1^{\lambda}\right) * log(G1)\right)$$

$$numel2 = p2 * \beta * \left(G1^{\lambda}\right) * log(G1) * eb$$

$$numel3 = p2 * \beta * \left(x^{\lambda}\right) * log(x) * ebb$$

$$numea1 = p1 * \gamma * x2g * \alpha * ea$$

$$numea2 = p1 * \gamma * xg1 * ea$$

$$numea3 = p1 * (G1^{\gamma}) * ea$$

$$numea4 = p1 * (x^{\gamma}) * eaa$$

$$numeb11 = p2 * \lambda * x2l * \beta * eb$$

$$numeb22 = p2 * \lambda * xl1 * eb$$

$$numeb33 = p2 * \left(G1^{\lambda}\right) * eb$$

$$numeb44 = p2 * \left(x^{\lambda}\right) * ebb$$

$$\frac{\partial log l_j}{\partial \alpha_j} = \sum_{i=1}^{n_j} \left(\frac{nume1}{deno1}\right) + \sum_{i=1}^{n_j} \left(\frac{nume2}{deno2}\right) - \sum_{i=1}^{n} \left(\frac{nume3}{deno3}\right)$$

$$\frac{\partial^2 logl_j}{\partial \alpha_j^2} = \sum_{i=1}^{n_j}\left(\frac{\left(deno1 * p1 * \gamma * x2g * ea * (-\alpha - 2)\right) - (nume1 * nume1)}{deno1^2}\right) -$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(deno2 * p1 * ea * \left(-(G1^{2*\gamma})\right)\right) - (nume2 * nume2)}{deno2^2}\right) +$$

$$\sum_{i=1}^{n}\left(\frac{\left(deno3 * p1 * eaa * \left(-(x^{2*\gamma})\right)\right) - (nume3 * nume3)}{deno3^2}\right)$$

$$\frac{\partial logl_j}{\partial \beta_j} = \sum_{i=1}^{n_j}\left(\frac{numeb1}{deno1}\right) + \sum_{i=1}^{n_j}\left(\frac{numeb2}{deno2}\right) - \sum_{i=1}^{n}\left(\frac{numeb3}{deno3}\right)$$

$$\frac{\partial^2 logl_j}{\partial \beta_j^2} = sum\sum_{i=1}^{n_j}\left(\frac{\left(deno1 * p2 * \lambda * x2l * eb * (-\beta - 2)\right) - (numeb1 * numeb1)}{deno1^2}\right) -$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(deno2 * p2 * eb * \left(-\left(G1^{2*\lambda}\right)\right)\right) - (numeb2 * numeb2)}{deno2^2}\right) +$$

$$\sum_{i=1}^{n}\left(\frac{\left(deno3 * p2 * ebb * \left(-\left(x^{2*\lambda}\right)\right)\right) - (numeb3 * numeb3)}{deno3^2}\right)$$

$$\frac{\partial logl_j}{\partial \gamma_j} = \sum_{i=1}^{n_j}\left(\frac{numeg1}{deno1}\right) + \sum_{i=1}^{n_j}\left(\frac{numeg2}{deno2}\right) - \sum_{i=1}^{n}\left(\frac{numeg3}{deno3}\right)$$

$$\frac{\partial^2 logl_j}{\partial \gamma_j^2} = \sum_{i=1}^{n_j}\left(\frac{\left((p1*deno1*\alpha*ea*log(G1)*xg1*(\gamma*log(G1)+2-\gamma*log(G1)*\alpha*(G1^{\gamma})-\alpha*(G1^{\gamma})-\alpha*\gamma*(G1^{\gamma})*logx2-\alpha*(G1^{\gamma})-\alpha*\gamma*x2g)\right)-(numeg1*numeg1)\right)}{deno1^2}\right) +$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(p1 * deno2 * \alpha * \left(\left(log(G1)\right)^2\right) * (G1^{\gamma}) * ea * \left(1 - \alpha * \gamma * (G1^{2*\gamma})\right)\right) - (numeg2 * numeg2)}{deno2^2}\right) -$$

$$\sum_{i=1}^{n}\left(\frac{\left(p1 * deno3 * \alpha * \left(\left(log(x)\right)^2\right) * (x^{\gamma}) * eaa * \left(1 - \alpha * \gamma * (x^{2*\gamma})\right)\right) - (numeg3 * numeg3)}{deno3^2}\right)$$

$$\frac{\partial logl_j}{\partial \lambda_j} = \sum_{i=1}^{n_j}\left(\frac{numel1}{deno1}\right) + \sum_{i=1}^{n_j}\left(\frac{numel2}{deno2}\right) - \sum_{i=1}^{n}\left(\frac{numel3}{deno3}\right)$$

$$\frac{\partial^2 logl_j}{\partial \lambda_j^2} = \sum_{i=1}^{n_j}\left(\frac{\left(p2*deno1*\beta*eb*log(G1)*xl1*(\lambda*log(G1)+2-\lambda*log(G1)*\beta*(G1^{\lambda})-\beta*(G1^{\lambda})-\beta*\lambda*(G1^{\lambda})*logx2-\beta*(G1^{\lambda})-\beta*\lambda*x2l)\right)-(numel1*numel1)}{deno1^2}\right) +$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(p2*deno2*\beta*\left(\left(log(G1)\right)^2\right)*(G1^{\lambda})*eb*\left(1-\beta*l*(G1^{2*\lambda})\right)\right)-(numel2*numel2)}{deno2^2}\right) -$$

$$\sum_{i=1}^{n}\left(\frac{\left(p2 * deno3 * \beta * \left(\left(log(x)\right)^2\right) * \left(x^{\lambda}\right) * ebb * \left(1 - \beta * \lambda * (x^{2*l})\right)\right) - (numel3 * numel3)}{deno3^2}\right)$$

$$\frac{\partial^2 logl_j}{\partial \alpha_j \gamma_j} = \frac{\partial^2 logl_j}{\partial \gamma_j \alpha_j}$$

$$= -\sum_{i=1}^{n_j}\left(\frac{\left(deno1 * p1 * \alpha * x2g * ea * \left(\gamma * log(G1^2) + 1 - \alpha * \gamma * (G1^{\gamma}) * log(G1)\right)\right) - (numea1 * numeg1)}{deno1^2}\right) +$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(p1 * deno1 * xg1 * ea * \left(\gamma * log(G1) + 1 - \alpha * \gamma * (G1^{\gamma}) * log(G1)\right)\right) - (numea2 * numeg1)}{deno1^2}\right) +$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(p1*deno2*(G1^{\gamma})*log(G1)*ea*\left(1-\alpha*(G1^{\gamma})\right)\right)-(numea3*numeg2)}{deno2^2}\right)-$$

$$\sum_{i=1}^{n}\left(\frac{\left(p1*deno3*(x^{\gamma})*log(x)*eaa*\left(1-\alpha*(x^{\gamma})\right)\right)-(numea4*numeg3)}{deno3^2}\right)$$

$$\frac{\partial^2 logl_j}{\partial\lambda_j\beta_j}=\frac{\partial^2 logl_j}{\partial\beta_j\lambda_j}$$

$$=-\sum_{i=1}^{n_j}\left(\frac{\left(deno1*p2*\beta*x2l*eb*\left(\lambda*log(G1^2)+1-\beta*\lambda*\left(G1^{\lambda}\right)*log(G1)\right)\right)-(numeb11*numel1)}{deno1^2}\right)+$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(p2*deno1*xl1*eb*\left(\lambda*log(G1)+1-\beta*\lambda*\left(G1^{\lambda}\right)*log(G1)\right)\right)-(numeb22*numel1)}{deno1^2}\right)+$$

$$\sum_{i=1}^{n_j}\left(\frac{\left(p2*deno2*\left(G1^{\lambda}\right)*log(G1)*eb*\left(1-\beta*\left(G1^{\lambda}\right)\right)\right)-(numeb33*numel2)}{deno2^2}\right)-$$

$$\sum_{i=1}^{n}\left(\frac{\left(p2*deno3*\left(x^{\lambda}\right)*log(x)*ebb*\left(1-\beta*\left(x^{\lambda}\right)\right)\right)-(numeb44*numel3)}{deno3^2}\right)$$

$$\frac{\partial^2 logl_j}{\partial\alpha_j\beta_j}=\frac{\partial^2 logl_j}{\partial\beta_j\alpha_j}=\sum_{i=1}^{n_j}\left(\frac{nume1*numeb1}{deno1^2}\right)-\sum_{i=1}^{n_j}\left(\frac{nume2*numeb2}{deno2^2}\right)+$$

$$\sum_{i=1}^{n}\left(\frac{nume3*numeb2}{deno3^2}\right)$$

$$\frac{\partial^2 logl_j}{\partial\alpha_j\lambda_j}=\frac{\partial^2 logl_j}{\partial\lambda_j\alpha_j}=\sum_{i=1}^{n_j}\left(\frac{nume1*numel1}{deno1^2}\right)-\sum_{i=1}^{n_j}\left(\frac{nume2*numel2}{deno2^2}\right)+$$

$$\sum_{i=1}^{n}\left(\frac{nume3*numel3}{deno3^2}\right)$$

$$\frac{\partial^2 logl_j}{\partial\beta_j\gamma_j}=\frac{\partial^2 logl_j}{\partial\gamma_j\beta_j}=\sum_{i=1}^{n_j}\left(\frac{numeb1*numeg1}{deno1^2}\right)-\sum_{i=1}^{n_j}\left(\frac{numeb2*numeg2}{deno2^2}\right)+$$

$$\sum_{i=1}^{n}\left(\frac{numeb3*numeg3}{deno3^2}\right)$$

$$\frac{\partial^2 logl_j}{\partial\gamma_j\lambda_j}=\frac{\partial^2 logl_j}{\partial\lambda_j\gamma_j}=\sum_{i=1}^{n_j}\left(\frac{numeg1*numel1}{deno1^2}\right)+\sum_{i=1}^{n_j}\left(\frac{numeg2*numel2}{deno2^2}\right)-$$

$$\sum_{i=1}^{n}\left(\frac{numeg3*numel3}{deno3^2}\right)$$

EXPLORE THE DYNAMICS OF MANUFACTURING INDUSTRIES: RELIABILITY ANALYSIS THROUGH STOCHASTIC PROCESS MODELING

Sonia[1,*], Shakuntla Singla [2]

[1,*2]Department of Mathematics and Humanities, MMEC, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala-133207, India
[1,*]Department of Mathematics, Dr. Bhim Rao Ambedkar Govt. College (Jagdishpura) Kaithal

[1,*]sonia_garg99@yahoo.com , [2] shakus25@gmail.com

Abstract

In nowadays, the chief attention of the researcher is to study how the reliability analysis of manufacturing industrial systems by using the stochastic process. This topic tells us, how the manufacturing industries perform over time with the help of mathematical models which include randomness and uncertainty. Through stochastic processes we examine the reliability of these systems, technologists can recognize possible failure points and develop tactics to improve overall performance and effectiveness. The reliability of a manufacturing industrial system can be examined through a stochastic process, which permits for the estimate of failure rates and maintenance agendas. This analysis can lead to more well-organized and cost-effective procedure of the system. In this study, the researcher analysed the possibility plan for reliability through many distributions such as the normal distribution, gamma distribution, weibull distribution, and exponential distribution. The result of the study was prepared using Minitab software. The result of the study shows that the normal distribution of reliability fits best in comparison to the gamma, weibull, and exponential distributions.

Keywords: Manufacturing Industry, Reliability.

1. Introduction

The introduction of stochastic models in the field of manufacturing industrial systems has confirmed to be an operative device for forecasting system reliability and classifying potential failure points. By including probabilistic parameters into the analysis, technologists and managers can up to date about maintenance agendas, system upgrades, and other tactics to recover overall system performance. As technology continues to advance and data collection methods become more sophisticated, the use of stochastic models is likely to become even more widespread in the field of industrial engineering. The world is expanding quickly and heading towards a "smart world" where technology controls everything. The cost and complexity of industrial processes have greatly increased with the advancement of technology in the twenty-first century. Hence, it has become crucial to run industrial systems with the least amount of downtime possible in order to achieve optimal productivity, boost revenue, and prevent losses. Hence, dependability and investigation of complex industrial systems is a need that cannot be avoided.

Many studies have been done on the subject of dependability modelling, analysis, and complex industrial systems under various operating situations and hypotheses. So first read a review on the modelling and analysis of complex industrial systems' reliability, in which different

reliability indices were presented in relation to various system parameters and various failure and repair rate distribution types used by various authors were discussed. Taking into account different failure distributions analysed failure data, and goodness of fit testing.

It was discovered that the exponential distribution was the foundation for the consequences of stochastic processes for system performance, and additional behaviour analysis of the process industry was conducted based on steady-state availability analysis and reliability analysis. In order to determine if the findings obtained genuinely have an exponential distribution, we develop a stochastic model for an industrial process in this study. Several of the foundational ideas are covered in the part after this one before we move on. Shihu Xiang, Jun Yang [1] analyzed in the paper reliability of WSN by considering random failures, energy consumption, environmental randomness and interference. F Delmotte et al [2] presented in their paper, modelling of reliability with possibility theory, a new approach based on a fusion rule by using a vector expressing the dependability on the data basics. Hasan et al [3], observed in their paper that the higher order logic formalization of some fundamental reliability theory concepts which can be built for reliability analysis in engineering system. J Dunyak et al [4] discussed in their paper that the probability of any event can be calculated by fuzzy fault tree which is independent of union, intersection and complements of any sequence of event. This allows a comprehensive analysis of the system. Q Zhang et al [5] discussed the belief reliability which is defined as the chance that a system is within a feasible domain. Measuring system reliability by a reasonable metric is a common problem in reliability engineering and the metric can degenerate either probability theory-based reliability or uncertainty theory-based reliability. Shakuntla Singla, Pooja Dhawan [6] analysis the behavior of a single unit subdivision after a complete failure by the help of RPGT. The situation is constructed and resolved by using RPGT to calculate system constraints. K Sachdeva et al [7] studies the sensitivity and productivity of a stochastic model whose technical faults may clear into either guarantee or coverage policies, the producer is answerable for all repair or replacement costs during normal or extended warranty. Profit and availability are calculated in all conditions. Kumar et al [8] presented in their paper overview of reliability analysis. Reliability analysis in different industries like milk, sugar, Thermal plant, petroleum industries etc. and to optimize the reliability using different techniques like G.A., H.A, PSO, machine learning etc.

2. Reliability analysis

It represents the Time to Failure as a statistical distribution, which is often defined by a certain pattern, using a Reliability Distribution Analysis. The aforementioned distribution types may be used. The timing of failure for a component is an example of a random event for which the result is unpredictable. Probability distributions are used to represent such occurrences. It is uncertain when that component will fail prior to putting a demand on it. A probability distribution models the distribution of the chance of failure at various periods. Random variables will be identified in this book by a capital letter, such as T for time. Researcher use little letter, such as t for time, to indicate when the random variable takes on a value. For instance, we would find P (T≤t) if we wanted to know the likelihood that the component would fail before time t. There are two types of random variables: discrete and continuous. The random variable in a discrete distribution has a specific or countable range of potential values, such as the number of demands before failure. The random variable in a continuous distribution is not restricted to certain potential values, unlike the time-to-failure distribution.

Probability $R(t) = P(T>t),\ t\geq 0,$ Where T- Random variable

t- Time

Failure probability distribution function $F(t) = 1 - R(t) = P(T\leq t)$

Probability density function $F(t) = \int_{-\infty}^{T} f(t)\,dt$

(i) Normal Distribution

A distribution in which continuous random variable has the probability density function with the following formula is given by

f (t) = $\frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(t-\mu)^2}{2\sigma^2}}$ Where μ and $\boldsymbol{\sigma}$ are arbitrary constant

(ii) Gamma distribution

A distribution in which continuous random variable has the probability density function with the following formula is given by

f (t) = c $t^{a-1}\ e^{-bt}$, t ≥ 0 Where a, b and c are constant

(iii) Weibull distribution

A distribution in which continuous random variable has the probability density function with the following formula is given by

f (t) = a $t^b\ e^{\frac{-(at^{b+1})}{(b+1)}}$, t ≥ 0 Where a and b are constant

(iv) Exponential distribution

A distribution in which continuous random variable has the probability density function with the following formula is given by

f (t) = $\begin{cases} \lambda e^{-\lambda t}, & 0 < t < \infty \\ 0, & t < 0 \end{cases}$

3. MANUFACTURING INDUSTRY PROCESS

A Manufacturing industry is chosen for reliability distribution and to know about the failure. The industrial procedure is divided into five sub systems. The sub system name with number of machines, their failure and mean time failure explained as below.

Table 1: *Industrial Process with Failures*

Industrial Process with Failures			
Sub System	Number of machines	Failure condition	Mean time failure MTBF
A. Storage Room	5	Never Fails	100000
B. Mixture	3	Failed when 2 machines fail	20000
C. Slasher	5	Failed when 3	16000
D. Loom	300	Failed when 30 machines fail	50000
E. Machine for Fabric Inspection	3	Failed when all 3 machines fail	80000

The reliability for stochastic model $R_1(t) = 1 - \sum_{i=1}^{4} P_i(t) = 0$

The value of transient probabilities is obtained by using following equations.

$$P_o'(t) + \sum_{i=1}^{4} \lambda i P_0(t) = 0,$$

$$P_i'(t) = \lambda_i P_0(t)$$ Where i=1, 2, 3, 4

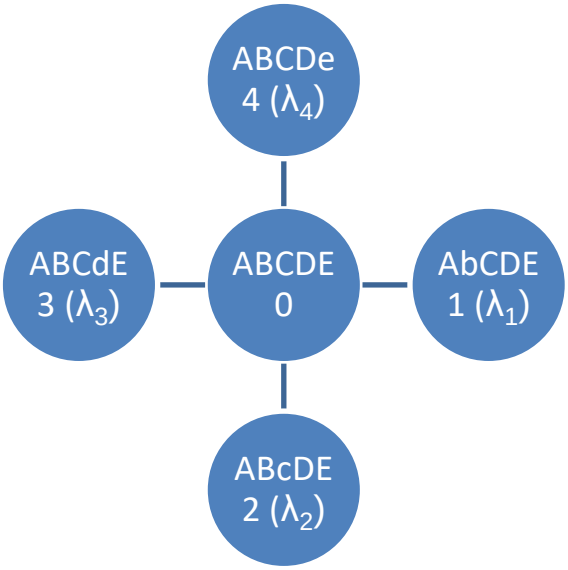


Figure 1: *Space diagram of manufacturing Industrial Process*

The above figure describes the sub system of industry in which A, B, C, D and E are the operating conditions and a, b, c, d and e describe the failed state of the model.

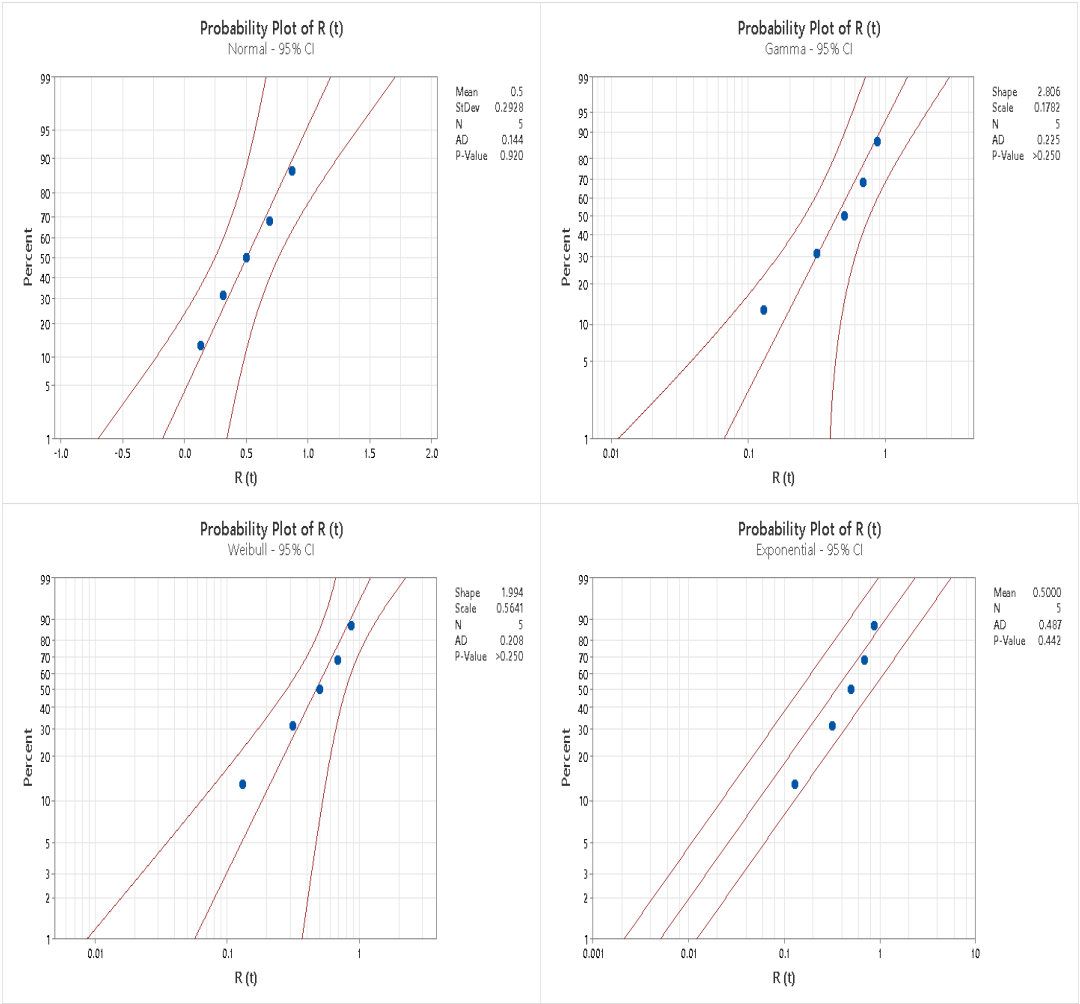


Figure 2: *Reliability analysis through Probability Curve for Normal, Gamma, Weibull and Exponential Distribution*

4. Result

The reliability of the model checked using the above equations. The probability distribution can be explained and plots by using Minitab software. In probability curve left line shows the lower bounds for confidence intervals, right line shows upper bounds for confidence intervals and middle line shows most fitted probability distribution. When value of P is more than its given value them distribution is best fit. In the above figure 2. it is cleared that normal distribution of probability is best fit. So, the conclusion of this study is that for reliability of an industry normal distribution is the fit. So, this represents the best distribution for maintenance of an industrial process.

5. Conclusion

The reliability distribution of an industrial system can be effectively analysed through stochastic processes, which provide a probabilistic framework for modelling and predicting system failures. This approach can help industries optimize maintenance schedules and improve overall system performance.

References

[1] Xiang, S., Yang, J. (2019). "*Reliability evalution and reliability- based optimal design for wireless sensor networks*"- IEEE systems Journal, ieeexplore.ieee.org.

[2] Delmotte, F., Borne, P. (1988). "*Modeling of reliability with possibility theory*" IEEE Transactions on system, ieeexplore.ieee.org

[3] Hasan, O., Tahar, S., Abbasi, N. (2009). "*Formal Reliability analysis using theorem proving*" on IEEE transactions on computers, ieeexplore.ieee.org.

[4] Dunyak, J., Saad, IW., Wunsch, D. (1999). "*A theory of independent fuzzy probability for system reliability*"on IEEE transactions on fuzzy, ieeexplore.ieee.org.

[5] Zhang, Q., Kang, R., Wen, R. (2018). "*Belief reliability for uncertain random systems*" on IEEE Transactions on Fuzzy systems 26(6), ieee.xplore.ieee.org.

[6] Shakuntla, S. and Pooja (2022). "*Mathematical Analysis of Regenerative Point Graphical Technique (RPGT)*". A mathematical analysis and contemporary applications. 49–56.

[7] Sachdeva, K., Taneja, G., Manocha, A. (2022). "*Sensitivity and economic analysis of an insured system with extended conditional warranty*". Reliability Theory and Applications, 24–31.

[8] Kumar, J., Bansal, S.A., Mehta, M. and Singh, H. (2020). "*Reliability analysis in process industries - an overview.*" GIS Sci J., 151-168.

RT&A, No 2 (78)
Volume 19, June, 2024

MULTICOMPONENT RELIABILITY UNDER PATHWAY MODEL

T. PRINCY

Department of Statistics
Cochin University of Science and Technology
Cochin-682022, Kerala, India
princyt@cusat.ac.in

Abstract

In this paper, we consider a system with a finite number of components. It is assumed that the system architecture is a series format. The system fails when any one of the components fails. The case where the lifetimes of the components, are independently distributed and have pathway density is considered. Then the survival function, hazard function, the expected time to failure, general moments, etc. of the system lifetime are computed. It is shown that the hazard function can have many types of shapes, including bathtub shapes. The estimation of stress-strength reliability is considered based on the method of maximum likelihood estimation when both stress and strength variables follow the pathway model. Finally, to show the applicability of the proposed model in a real-life scenario, remission time data from cancer patients is analyzed.

Keywords: Survival Function, Pathway Distribution, Multicomponent Reliability, Stress-Strength Reliability, Expected Time to Failure

1. Introduction

Consider a multicomponent system consisting of k components, connected in a series format so that the system fails if any of the k components fails. Let the lifetimes of the components be the random variables $X_1, ..., X_k$, $X_j > 0, j = 1, ..., k$. Let X be the minimum, $X = \min\{X_1, ..., X_k\}$. Then the system failure time is X. Suppose that the components are functioning independently. Then the probability that $X > t$ for some t is given by

$$Pr\{X > t\} = Pr\{X_1 > t\}Pr\{X_2 > t\}...Pr\{X_k > t\}. \tag{1}$$

That is, in terms of the distribution functions

$$1 - F_X(t) = [1 - F_{X_1}(t)]...[1 - F_{X_k}(t)]$$

where $F_{X_j}(t) = Pr\{X_j \leq t\}$ is the distribution function of $X_j, j = 1, ..., k$ and $F_X(t)$ is that of X. Then the density of X, if $F_{X_j}(t)$ is differentiable, is given by the following:

$$f_X(t) = -\frac{\mathrm{d}}{\mathrm{d}t}[1 - F_X(t)] = \sum_{j=1}^{k} \left\{ [-\frac{\mathrm{d}}{\mathrm{d}t} Pr\{X_j > t\}] \prod_{i \neq j=1}^{k} Pr\{X_i > t\} \right\}. \tag{2}$$

Basic notions of reliability analysis may be seen, from [1], [2], and [3]. Reliability analysis for dependent cases may be seen, for example, from [4], [5] and [6]. We will examine (2) and study its properties and connections to various problems in different fields. First, we will consider the case when the density of x_j belongs to the general family of functions called the pathway model. The original pathway model was introduced by Mathai [7] for the real rectangular matrix-variate case. Later, Mathai and Provost [8] was extended it to the complex domain. The pathway model for the real scalar positive variable case can be stated as follows:

$$f_1(x) = c_1 x^{\gamma}[1 - a(1-q)x^{\delta}]^{\frac{\eta}{1-q}}, q < 1 \tag{3}$$

for $a > 0, \delta > 0, \eta > 0, \gamma > -1, 1 - a(1-q)x^{\delta} > 0$ and $f_1(x) = 0$ elsewhere. The functional part of the basic type-1 beta density is $x^{\alpha-1}(1-x)^{\beta-1}, 0 \le x \le 1, \alpha > 0, \beta > 0$ and zero elsewhere. Hence (3) can be looked upon as a generalized type-1 beta form, that is, for $\frac{\eta}{1-q} = \beta - 1, \delta = 1, q = 0, \gamma = \alpha - 1$ one has the type-1 beta form. Note that one can also relocate the variable x. Write the model as

$$f_2(x) = c_2(x-\alpha)^{\gamma}[\beta - a(1-q)(x-\alpha)^{\delta}]^{\frac{\eta}{1-q}}, q < 1 \tag{4}$$

for $\eta > 0, a > 0, \delta > 0, x \ge \alpha, \beta > 0, \alpha > 0, 0 < \alpha \le x \le \alpha + [\frac{\beta}{a(1-q)}]^{\frac{1}{\delta}}$. Note that the basic type-1 beta model, triangular density, power function model, uniform density, etc are particular cases of (3). The limiting form of the exponentiated versions of (3) and (4) can also be shown to be Bose-Einstein density in Physics. Note that when q approaches 1 then the support will extend to $0 \le x < \infty$ from the finite range support in (3). For $q > 1$, write $1 - q = -(q-1)$ so that the model in (3) switches into the model, which is another family of functions,

$$f_3(x) = c_3 x^{\gamma}[1 + a(q-1)x^{\delta}]^{-\frac{\eta}{q-1}}, q > 1 \tag{5}$$

for $a > 0, \eta > 0, \gamma > -1, \delta > 0, x \ge 0$. The functional part of the basic type-2 beta density is $x^{\alpha-1}(1+x)^{-(\alpha+\beta)}, 0 \le x < \infty, \alpha > 0, \beta > 0$. Hence (5) can be looked upon as a generalized type-2 beta model. If relocation of the variable is required, then replace x in (5) by $x - \alpha > 0$ so that $0 < \alpha \le x < \infty$. Observe that the standard F-density, type-2 beta density, Pareto density, etc are particular cases in (5). The exponentiated version of (5), that is, put $x = \mathrm{e}^{-cy}, c > 0, -\infty < y < \infty$ is connected to various densities such as the generalized logistic density, see [9], the standard logistic density, a limiting form giving rise to the famous Fermi-Dirac density in Physics also. Now, let $q \to 1_-$ in (3) and $q \to 1_+$ in (5). Then both the models in (3) and (5) go to

$$f_4(x) = c_4 x^{\gamma}\mathrm{e}^{-a\eta x^{\delta}}, a > 0, \eta > 0, \delta > 0, x \ge 0 \tag{6}$$

and zero elsewhere. We may also relocate the variable, if necessary. Observe that (6) is in the form of a generalized gamma density. For $\gamma = \delta - 1$ it is the Weibull density. The standard gamma density, chisqure density, exponential, density, Maxwell-Boltzmann density, Raleigh density, etc are special cases of (6). Thus, (3) or (5) is the basic pathway model or all cases of (3), (5) and (6) are contained in (3) or (5). For $q < 1, q > 1, q \to 1$ will cover almost all densities in current use and all these are contained in (3) or (5). Hence a wide spectrum of models is covered in the problems that we discuss in this paper. The advantage of the model in (3) or (5) in a model building situation is the following: If $f_1(x), f_3(x), f_4(x)$ are to be treated as statistical densities, then c_1, c_2, c_3 are the normalizing constants, there and they are the following:

$$c_1 = \frac{\delta[a(1-q)]^{\frac{\gamma+1}{\delta}}\Gamma(\frac{\eta}{1-q} + 1 + \frac{\gamma+1}{\delta})}{\Gamma(\frac{\gamma+1}{\delta})\Gamma(\frac{\eta}{1-q}+1)}, q < 1, \gamma + 1 > 0, a, \delta, \eta > 0 \tag{7}$$

.

$$c_3 = \frac{\delta[a(q-1)]^{\frac{\gamma+1}{\delta}}\Gamma(\frac{\eta}{q-1})}{\Gamma(\frac{\gamma+1}{\delta})\Gamma(\frac{\eta}{q-1} - \frac{\gamma+1}{\delta})} q > 1, \gamma + 1, a, \delta, \eta > 0, \frac{\eta}{q-1} - \frac{\gamma+1}{\delta} > 0 \tag{8}$$

.

$$c_4 = \frac{\delta(a\eta)^{\frac{\gamma+1}{\delta}}}{\Gamma(\frac{\gamma+1}{\delta})}, \gamma+1>0, \delta>0, a>0, \eta>0. \tag{9}$$

For $\delta = 1, a = 1, \eta = 1, \gamma = 0$ in (3) gives the famous Tsallis statistics in nonextensive statistical mechanics. This Tsallis statistic is valid for $q < 1, q > 1, q \to 1$ situations. It is stated that over 3000 articles were written on this Tsallis statistics between 1990 and 2010 period. Tsallis statistics, excluding the normalizing constant, is a power function model in the sense

$$\frac{\mathrm{d}}{\mathrm{d}x} f_1(x) = -[f_1(x)]^q.$$

For $a = 1, \delta = 1, \eta = 1$, (5) gives superstatistics in statistical mechanics. This is valid for $q > 1, q \to 1$ situations but not for $q < 1$. Dozens of articles are also published in this area. The development in Tsallis statistics is available from his book, see [10]. The basic paper on superstatistics is by [11]. From a physical point of view, superstatistics is constructed by superimposing a distribution over another distribution. But from a statistical point of view, superstatistics is nothing but an unconditional density in a Bayesian setup when both the conditional density and prior density belong to generalized gamma families. By using this pathway model several compound distributions are developed for details see, [12],[13] and [14].

1.1. A particular case

Our interest here is to examine the multi-component system failure under a pathway model of (3), thereby (5) and (6) for the particular case $\gamma = \delta - 1$. In this case, the normalizing constants simplify and the models go into very simple forms. This particular case of (3),(5) and (6) is the following:

$$f_5(x) = a\delta(\eta+1-q)x^{\delta-1}[1-a(1-q)x^{\delta}]^{\frac{\eta}{1-q}}, q<1 \tag{10}$$

for $a > 0, \delta > 0, \eta > 0, \eta + 1 - q > 0, 1 - a(1-q)x^{\delta} > 0$.

$$f_6(x) = a\delta(\eta+1-q)x^{\delta-1}[1+a(q-1)x^{\delta}]^{-\frac{\eta}{q-1}}, q>1 \tag{11}$$

for $a > 0, \delta > 0, \eta > 0, \eta + 1 - q > 0, x \geq 0$.

$$f_7(x) = a\delta\eta x^{\delta-1}\mathrm{e}^{-a\eta x^{\delta}}, \delta>0, a>0, \eta>0. \tag{12}$$

The corresponding survival probabilities are the following:

$$S_5(t) = [1-a(1-q)t^{\delta}]^{\frac{\eta}{1-q}+1}, q<1, \tag{13}$$

for $a > 0, \delta > 0, \eta > 0, 1 - a(1-q)t^{\delta} > 0$.

$$S_6(t) = [1+a(q-1)t^{\delta}]^{-\frac{\eta}{q-1}+1}, q>1, a, \delta, \eta>0, t\geq 0 \tag{14}$$

.

$$S_7(t) = \mathrm{e}^{-a\eta t^{\delta}}, a>0, \eta>0, t\geq 0. \tag{15}$$

2. Multicomponet Failure Under Pathway Model

Recall the probability of failure from (1). Then, under the pathway model of (4) to (6) for the particular case $\gamma = \delta - 1$ it is the following, writing for convenience the form in (14) for $q > 1$:

$$Pr\{x>t\} = \prod_{j=1}^{k}[1+a_j(q_j-1)t^{\delta_j}]^{-\frac{\eta_j}{q_j-1}+1} \tag{16}$$

for $a_j > 0, \delta_j > 0, \eta_j > 0, \eta_j + 1 - q_j > 0, q_j > 1, q_j < 1, q_j \to 1, j = 1, ..., k$. For any particular j, we can take the form in (13) or (14) or (15). Thus, (16) gives a very rich family of probabilities. The density of x in this case, denoted by $f(x)$, is the following:

$$\begin{aligned} f(t) &= -\frac{\mathrm{d}}{\mathrm{d}t} Pr\{x > t\} = \sum_{j=1}^{k} (\eta_j + 1 - q_j) a_j \delta_j t^{\delta_j - 1} \\ &\times [1 + a_j(q_j - 1)t^{\delta_j}]^{-\frac{\eta_j}{q_j - 1}} \{ \prod_{i \neq j = 1}^{k} [1 + a_i(q_i - 1)t^{\delta_i}]^{-\frac{\eta_i}{q_i - 1} + 1} \}. \end{aligned} \tag{17}$$

Therefore the hazard function of x, denoted by $h(t)$, is the following:

$$h(t) = \frac{f_x(t)}{Pr\{x > t\}} = \sum_{j=1}^{k} \frac{(\eta_j + 1 - q_j) a_j \delta_j t^{\delta_j - 1}}{1 + a_j(q_j - 1)t^{\delta_j}} \tag{18}$$

for $q_j > 1, q_j < 1, q_j \to 1, \eta_j > 0, \eta_j + 1 - q_j > 0, a_j > 0, \delta_j > 0, j = 1, ..., k$. Observe that when a $q_j \to 1$ for a particular j, the corresponding term is simply $\eta_j a_j \delta_j t^{\delta_j - 1}$. Thus, a rich variety of hazard functions of having curves of various shapes are available from (18). For example, for $k = 2, q_2 \to 1$ we have the form, denoted by $h(t)$,

$$h(t) = \frac{(\eta_1 + 1 - q_1) a_1 \delta_1 t^{\delta_1 - 1}}{1 + a_1(q_1 - 1)t^{\delta_1}} + \eta_2 a_2 \delta_2 t^{\delta_2 - 1}, q_1 > 1. \tag{19}$$

The different shapes of the hazard function of multicomponent systems under the pathway model are demonstrated.

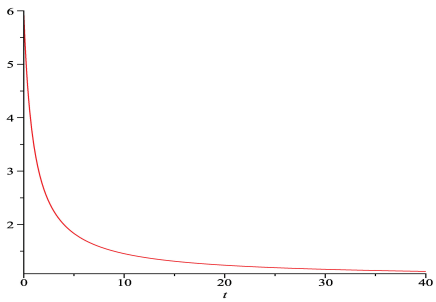


Figure 1

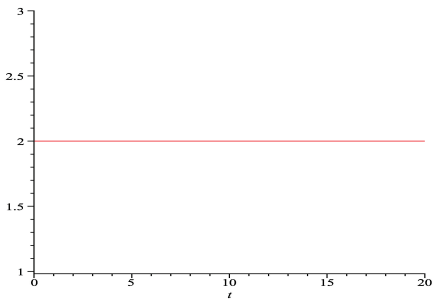


Figure 2

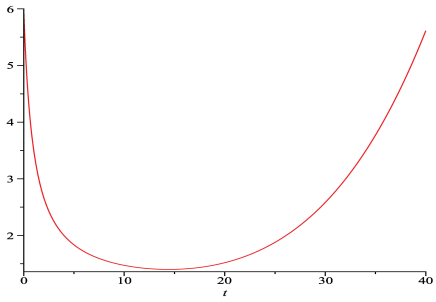


Figure 3

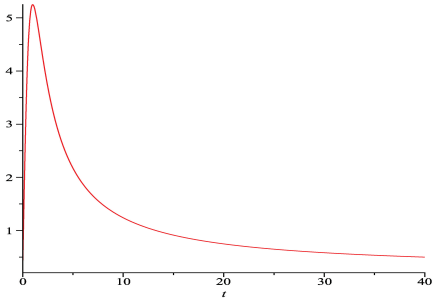


Figure 4

Figure 1: $\eta_1 = 3, \eta_2 = 1, q_1 = 1.5, a_1 = 2, a_2 = 1, \delta_1 = 1, \delta_2 = 1$
Figure 2: $\eta_1 = 0.5, \eta_2 = 2, q_1 = 1.5, a_1 = 2, a_2 = 1, \delta_1 = 1, \delta_2 = 1$
Figure 3: $\eta_1 = 3, \eta_2 = \frac{1}{1500}, q_1 = 1.5, a_1 = 2, a_2 = \frac{1}{2}, \delta_1 = 1, \delta_2 = 1$
Figure 4: $\eta_1 = 3, \eta_2 = \frac{1}{2}, q_1 = 1.5, a_1 = 2, a_2 = \frac{1}{2}, \delta_1 = 1, \delta_2 = 1$

Another case for $k = 2$, $q_1 > 1$ and $q_2 < 1$, $h(t)$ becomes;

$$h(t) = \frac{(\eta_1 + 1 - q_1) a_1 \delta_1 t^{\delta_1 - 1}}{1 + a_1(q_1 - 1)t^{\delta_1}} + \frac{(\eta_2 + 1 - q_2) a_2 \delta_2 t^{\delta_2 - 1}}{1 - a_2(1 - q_2)t^{\delta_2}}, \tag{20}$$

for $q_1 > 1, q_2 < 1, a_j > 0, \delta_j > 0, \eta_j > 0, \eta_j + 1 - q_j > 0, j = 1, 2$. All types of shapes are available from (20).

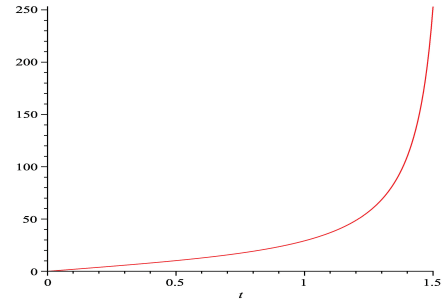


Figure 5

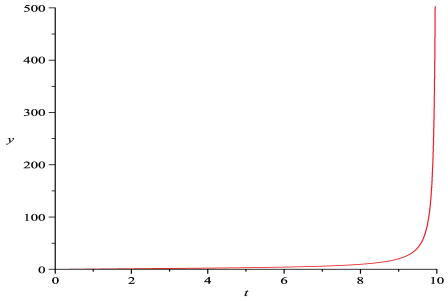


Figure 6

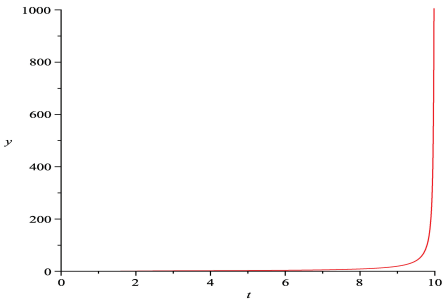


Figure 7

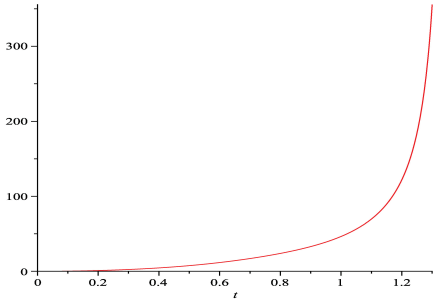


Figure 8

Figure 5: $\eta_1 = 2, \eta_2 = 2, q_1 = 1.9, q_2 = 0.9, a_1 = 1, a_2 = 4, \delta_1 = 2, \delta_2 = 2$
Figure 6: $\eta_1 = 2, \eta_2 = 2, q_1 = 1.9, q_2 = 0.9, a_1 = \frac{1}{10}, a_2 = \frac{1}{10}, \delta_1 = 2, \delta_2 = 2$
Figure 7: $\eta_1 = 2, \eta_2 = 2, q_1 = 1.9, q_2 = 0.9, a_1 = \frac{1}{100}, a_2 = \frac{1}{100}, \delta_1 = 5, \delta_2 = 3$
Figure 8: $\eta_1 = 2, \eta_2 = 2, q_1 = 1.9, q_2 = 0.9, a_1 = 3, a_2 = 4, \delta_1 = 5, \delta_2 = 3$

2.1. Expected time to failure

From here onward, all discussions connected with $k = 2$ also contain the case of $k - 1$ of the lifetimes $x_1, \ldots, x_k$ are identically distributed so that there will only be two distinct densities. This can be computed from the density of x itself or from the survival function of x. That is,

$$E(t) = \int_0^\infty t f_x(t) \mathrm{d}t = \int_0^\infty S_x(t) \mathrm{d}t \tag{21}$$

where $S_x(t) = Pr\{x > t\}$ is the survival function of t. Integration by parts once gives the second part in (21). Hence the ρ-th moment of the time to failure is the following:

$$\begin{aligned} E(t^\rho) &= \int_0^\infty t^\rho f_x(t) \mathrm{d}t = \rho \int_0^\infty t^{\rho-1} S_x(t) \mathrm{d}t \\ &= \rho \int_0^\infty t^{\rho-1} \{\prod_{j=1}^{k} [1 + a_j(q_j - 1)t^{\delta_j}]^{-\frac{\eta_j}{q_j - 1} + 1}\} \mathrm{d}t, q_j > 1. \end{aligned} \tag{22}$$

Take $q_j < 1$ for the type-1 case and $q_j \to 1$ for the gamma case. Hence all different forms are there in (22). For $k = 2$, a general integral in this category, denoted by I_1, is the following:

$$I_1 = \int_0^\infty t^\xi [1 + a_1(q_1 - 1)t^{\delta_1}]^{-\frac{\eta_1}{q_1 - 1} + 1} [1 + a_2(q_2 - 1)t^{\delta_2}]^{-\frac{\eta_2}{q_2 - 1} + 1} \mathrm{d}t. \tag{23}$$

Replace ξ by $\rho - 1$ and multiply the integral by ρ to obtain the ρ-th moment from I_1. The integral in (23) has the structure of a Mellin convolution of a ratio. For two functions $g_1(x_1)$ and $g_2(x_2)$

the Mellin convolution of a ratio has the format

$$g(u) = \int_v v g_1(uv) g_2(v) \mathrm{d}v \tag{24}$$

so that the Mellin transform of $g(u)$, with Mellin parameter s, or

$$M_g(s) = \int_0^\infty u^{s-1} g(u) \mathrm{d}u$$

has the form

$$M_g(s) = M_{g_1}(s) M_{g_2}(2-s) \tag{25}$$

where

$$M_{g_1}(s) = \int_0^\infty x_1^{s-1} g_1(x_1) \mathrm{d} \text{ and } M_{g_2}(2-s) = \int_0^\infty x_2^{-s+1} g_2(x_2) \mathrm{d}x_2$$

where g_1 and g_2 need not be statistical densities. If they are statistical densities then the situation is the following: $M_{g_1}(s) = E(x_1^{s-1}), M_{g_2}(2-s) = E(x_2^{-s+1})$ and $u = \frac{x_1}{x_2}, x_2 = v, x_1 = uv$ and the Jacobian is v. $E[\frac{x_1}{x_2}]^{s-1} = E[x_1^{s-1}]E[x_2^{-s+1}]$ when $x_1 > 0$ and $x_2 > 0$ are independently distributed real scalar positive random variables, where E denotes the expected value. Then the density of u, denoted by $g(u)$, is available from the inverse Mellin transform. That is,

$$g(u) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} [E(u^{s-1})] u^{-s} \mathrm{d}s = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} M_{g_1}(s) M_{g_2}(2-s) u^{-s} \mathrm{d}s, i = \sqrt{-1}. \tag{26}$$

In (25), g_1 and g_2 need not be statistical densities. The only condition is that the Mellin transforms exist. For the existence of inverse Mellin transform, general conditions are available, see books on complex analysis, or see [15]. For evaluating (23) let

$$g_1(x_1) = [1 + x_1^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1} \text{ and } g_2(x_2) = x_2^{\xi-1}[1 + a_2(q_2-1)x_2^{\delta_2}]^{-\frac{\eta_2}{q_2-1}+1}$$

so that for $u = [a_1(q_1-1)]^{\frac{1}{\delta_1}}$

$$\begin{aligned} g(u) &= \int_v v g_1(uv) g_2(v) \mathrm{d}v = \int_0^\infty v^\xi [1 + a_1(q_1-1)v^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1} \\ &\times [1 + a_2(q_2-1)v^{\delta_2}]^{-\frac{\eta_2}{q_2-1}+1} \mathrm{d}v = I_1 \end{aligned} \tag{27}$$

which is the right side of (23) or the item to be evaluated. But

$$\begin{aligned} M_{g_1}(s) &= \int_0^\infty x_1^{s-1}[1 + x_1^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1} \mathrm{d}x_1 \\ &= \frac{\Gamma(\frac{s}{\delta_1})\Gamma(\frac{\eta_1}{q_1-1} - 1 - \frac{s}{\delta_1})}{\delta_1 \Gamma(\frac{\eta_1}{q_1-1} - 1)} \end{aligned} \tag{28}$$

for $\eta_1 + 1 - q_1 > 0, \delta_1 > 0, \eta_1 > 0, 1 < q_1 < \eta_1 + 1, \Re(s) > 0$ where $\Re(\cdot)$ means the real part of $(\cdot)$.

$$\begin{aligned} M_{g_2}(2-s) &= \int_0^\infty x_2^{-s+1} x_2^{\xi-1}[1 + a_2(q_2-1)x_2^{\delta_2}]^{-\frac{\eta_2}{q_2-1}+1} \mathrm{d}x_2 \\ &= \frac{\Gamma(\frac{\xi-s+1}{\delta_2})\Gamma(\frac{\eta_2}{q_2-1} - 1 - \frac{(\xi-s+1)}{\delta_2})}{\delta_2 [a_2(q_2-1)]^{\frac{\xi-s+1}{\delta_2}} \Gamma(\frac{\eta_2}{q_2-1} - 1)} \end{aligned} \tag{29}$$

for $\Re(\xi - s + 1) > 0, \delta_2 > 0, \eta_2 > 0, \eta_2 + 1 - q_2 > 0, 1 < q_2 < \eta_2 + 1, \Re(\frac{\eta_2}{q_2 - 1} - 1 - \frac{(\xi - s + 1)}{\delta_2}) > 0$.
Hence I_1 is available from the inverse Mellin transform, remembering that $u = [a_1(q_1 - 1)]^{\frac{1}{\delta_1}}$.

$$\begin{aligned} I_1 &= [\delta_1\delta_2[a_2(q_2-1)]^{\frac{\xi+1}{\delta_2}}]^{-1}[\Gamma(\frac{\eta_1}{q_1-1}-1)\Gamma(\frac{\eta_2}{q_2-1}-1)]^{-1} \\ &\times \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(\frac{s}{\delta_1})\Gamma(\frac{\eta_2}{q_2-1}-1-\frac{(\xi+1)}{\delta_2}+\frac{s}{\delta_2}) \\ &\times \Gamma(\frac{\eta_1}{q_1-1}-1-\frac{s}{\delta_1})\Gamma(\frac{\xi+1}{\delta_2}-\frac{s}{\delta_2})\left[\frac{[a_1(q_1-1)]^{\frac{1}{\delta_1}}}{[a_2(q_2-1)]^{\frac{1}{\delta_2}}}\right]^{-s}\mathrm{d}s \end{aligned} \tag{30}$$

for $\max\{0, \frac{\xi+1}{\delta_2} + 1 - \frac{\eta_2}{q_2-1}\} < c < \min\{\xi + 1, \frac{\eta_1\delta_1}{q_1-1} - \delta_1\}$, $\delta_j > 0, a_j > 0, 1 < q_j < \eta_j + 1, \xi > -1, \eta_j > 0, \eta_j + 1 - q_j > 0, j = 1, 2$. This I_1 can be written as a H-function, see [16]. That is, denoting the constant part by C, we have

$$I_1 = C\ H^{2,2}_{2,2}\left[\omega\Big|^{(2-\frac{\eta_1}{q_1-1},\frac{1}{\delta_1}),(1-\frac{(\xi+1)}{\delta_2},\frac{1}{\delta_2})}_{(0,\frac{1}{\delta_1}),(\frac{\eta_2}{q_2-1}-1-\frac{(\xi+1)}{\delta_2},\frac{1}{\delta_2})}\right] \tag{31}$$

for $0 < |\omega| < 1$ where

$$\omega = \frac{[a_1(q_1-1)]^{\frac{1}{\delta_1}}}{[a_2(q_2-1)]^{\frac{1}{\delta_2}}} \text{ and } [\delta_1\delta_2[a_2(q_2-1)]^{\frac{\xi+1}{\delta_2}}]^{-1}[\Gamma(\frac{\eta_1}{q_1-1}-1)\Gamma(\frac{\eta_2}{q_2-1}-1)].$$

Observe that the roles of $[a_1(q_1-1)]^{\frac{1}{\delta_1}}$ and $[a_2(q_2-1)]^{\frac{1}{\delta_2}}$ can be interchanged by interchanging the roles of g_1 and g_2. For the existence conditions and properties of H-function see Mathai et al. (2010)[16]. MATHEMATICA programs are available for computing H-functions.
Note that I_1 of (23) has nine different forms there. We can have $q_1 > 1, (q_2 > 1, q_2 < 1, q_2 \to 1)$. Similarly for $q_1 < 1$ and $q_1 \to 1$ cases. When $q_1 \to 1$ and $q_2 \to 1$ we have the integral in (23) as

$$= \int_0^\infty t^\xi \mathrm{e}^{-a_1\eta_1 t^{\delta_1} - a_2\eta_2 t^{\delta_2}}\mathrm{d}t. \tag{32}$$

This (32) for either $\delta_1 = 1$ or $\delta_2 = 1$ corresponds to the Laplace transform or moment-generating function of a generalized gamma density. One can go through the steps (23) to (31) and obtain the following result for $q_1 \to 1, q_2 \to 1$:

$$I_2 = [\delta_1\delta_2(a_2\eta_2)^{\frac{\xi+1}{\delta_2}}]^{-1}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(\frac{s}{\delta_1})\Gamma(\frac{\xi+1-s}{\delta_2})[\frac{(a_1\eta_1)^{\frac{1}{\delta_1}}}{(a_2\eta_2)^{\frac{1}{\delta_2}}}]^{-s}\mathrm{d}s \tag{33}$$

$$= [\delta_1\delta_2(a_2\eta_2)^{\frac{\xi+1}{\delta_2}}]^{-1}H^{1,1}_{1,1}\left[\frac{(a_1\eta_1)^{\frac{1}{\delta_1}}}{(a_2\eta_2)^{\frac{1}{\delta_2}}}\Big|^{(1-\frac{\xi+1}{\delta_2},\frac{1}{\delta_2})}_{(0,\frac{1}{\delta_1})}\right], \tag{34}$$

for $\frac{(a_1\eta_1)^{\frac{1}{\delta_1}}}{(a_2\eta_2)^{\frac{1}{\delta_2}}} < 1$. The Mellin-Barnes representation in (33) can be written in the following form by replacing $\frac{s}{\delta_1}$ by s and writing $c^* = [\delta_2(a_2\eta_2)^{\frac{\xi+1}{\delta_2}}]^{-1}$:

$$I_2 = c^*\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(s)\Gamma(\frac{\xi+1}{\delta_2}-\frac{\delta_1}{\delta_2}s)[\frac{a_1\eta_1}{(a_2\eta_2)^{\frac{\delta_1}{\delta_2}}}]^{-s}\mathrm{d}s. \tag{35}$$

Evaluating this at the poles of $\Gamma(s)$ at $s = 0, -1, -2, \ldots$ the residue at $s = -\nu$ is given by

$$\lim_{s\to-\nu}(s+\nu)\Gamma(s)\Gamma(\frac{\xi+1}{\delta_2}-\frac{\delta_1}{\delta_2}s)\omega^{-s} = \frac{(-1)^\nu}{\nu!}\Gamma(\frac{\xi+1}{\delta_2}+\frac{\delta_1}{\delta_2}\nu)\omega^\nu, \omega = \frac{(a_1\eta_1)}{(a_2\eta_2)^{\frac{\delta_1}{\delta_2}}}.$$

Therefore I_2 is available as the sum of the residues.

$$I_2 = \frac{1}{\delta_1\delta_2(a_2\eta_2)^{\frac{\xi+1}{\delta_2}}}\sum_{\nu=0}^{\infty}\frac{(-1)^\nu}{\nu!}\Gamma(\frac{\xi+1}{\delta_2}+\frac{\delta_1}{\delta_2}\nu)\omega^\nu \tag{36}$$

for $0 < a_1\eta_1 < (a_2\eta_2)^{\frac{\delta_1}{\delta_2}}$. Note that for $\delta_1 = \delta_2 = \delta$ the right side of (36) is a binomial series, giving a binomial sum for $a_1\eta_1 < a_2\eta_2$. The analytic continuation part is available from the poles of $\Gamma(\frac{\xi+1}{\delta_2} - \frac{s}{\delta_2})$. Replacing $\frac{s}{\delta_2}$ by s we have from (33), for $\hat{c} = [\delta_1(a_2\eta_2)^{\frac{\xi+1}{\delta_2}}]^{-1}$,

$$I_2 = \hat{c}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(\frac{\delta_2}{\delta_1}s)\Gamma(\frac{\xi+1}{\delta_2}-s)[\frac{(a_1\eta_1)^{\frac{\delta_2}{\delta_1}}}{(a_2\eta_2)}]^{-s}\mathrm{d}s. \tag{37}$$

The poles of $\Gamma(\frac{\xi+1}{\delta_2} - s)$ are at $s = \frac{\xi+1}{\delta_2} + \nu, \nu = 0, 1, 2, ...$ Then

$$I_2 = \hat{c}\omega^{-\frac{\xi+1}{\delta_2}}\sum_{\nu=0}^{\infty}\frac{(-1)^\nu}{\nu!}\Gamma(\frac{\xi+1}{\delta_1}+\frac{\delta_2}{\delta_1}\nu)\omega^{-\nu} \tag{38}$$

for $\omega > 1$. Thus, (36) and (38) give the series for all values of $\omega > 0$. Again, for $\delta_1 = \delta_2$, (38) reduces to a binomial sum.

3. One Factor with Negative Exponent

We can observe that in the pathway model for the cases of $q > 1$ and $q \to 1$, x and $\frac{1}{x}$ belong to the same family of functions. In other words both the situations x^δ and $x^{-\delta}$, with $\delta > 0$, are admissible cases. When $x^{-\delta}$ is there then the survival function will be of the form $1 - [1 + a(q-1)t^{-\delta}]^{-\frac{\eta}{q-1}+1}$. Let us again consider the case of two components, independently acting, or $k = 2$ where x_1 has a pathway model of beta type-2 and x_2 has an inverted type-2 beta pathway model. Then the h-th moment of the system survival time, written in terms of the survival function, is the following:

$$\begin{aligned} E(t^h) &= \int_0^\infty t^h f_x(t)\mathrm{d}t = h\int_0^\infty t^{h-1}S_x(t)\mathrm{d}t = h\int_0^\infty t^{h-1}[1+a_1(q_1-1)t^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1} \\ &\times \{1-[1+a_2(q_2-1)t^{-\delta_2}]^{-\frac{\eta_2}{q_2-1}+1}\}\mathrm{d}t. \end{aligned} \tag{39}$$

In order to evaluate the integral in (39) let us consider the general integral

$$g_2 = \int_0^\infty t^{\gamma-1}[1+a_1(q_1-1)t^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1}[1+a_2(q_2-1)t^{-\delta_2}]^{-\frac{\eta_2}{q_2-1}+1}\mathrm{d}t. \tag{40}$$

This can be evaluated with the help of Mellin convolution of a product. Let $x_1 > 0, x_2 > 0, u = x_1x_2, v = x_2$ or $x_1 = \frac{u}{v}$, Jacobian is $\frac{1}{v}$. Let the corresponding functions be $f_8(x_1)$ and $f_9(x_2)$. Then consider the integral

$$g_2 = \int_0^\infty \frac{1}{v}f_8(\frac{u}{v})f_9(v)\mathrm{d}v. \tag{41}$$

Then Mellin convolution of a product says that $M_{g_2}(s) = M_{f_8}(s)M_{f_9}(s)$ where s is the Mellin parameter. In terms of independently distributed real scalar positive random variables x_1 and x_2, with densities $f_8(x_1)$ and $f_9(x_2)$ respectively, g_2 will represent the density of the product $x_1x_2 = u$. Take $u = [a_2(q_2-1)]^{\frac{1}{\delta_2}}, x_2 = v$ and let

$$f_8(x_1) = [1+x_1^{\delta_2}]^{-\frac{\eta_2}{q_2-1}+1} \text{ and } f_9(x_2) = x_2^\gamma[1+a_1(q_1-1)x_2^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1}.$$

Then

$$f_8(\frac{u}{v}) = [1+a_2(q_2-1)v^{-\delta_2}]^{-\frac{\eta_2}{q_2-1}+1}$$

Then the Mellin transform of f_8, with Mellin parameter s, denoted by $M_{f_8}(s)$, is the following:

$$M_{f_8}(s) = \int_0^\infty x_1^{s-1} f_8(x_1)\mathrm{d}x_1 = \frac{\Gamma(\frac{s}{\delta_2})\Gamma(\frac{\eta_2}{q_2-1} - 1 - \frac{s}{\delta_2})}{\delta_2\Gamma(\frac{\eta_2}{q_2-1} - 1)}$$

for $\Re(s) > 0, \Re(\frac{\eta_2}{q_2-1} - 1 - \frac{s}{\delta_2}) > 0, \eta_2 + 1 - q_2 > 0$. But

$$\frac{1}{v} f_9(v) = v^{\gamma-1}[1 + a_1(q_1-1)v^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1}.$$

Then $\int_0^\infty \frac{1}{v} f_8(\frac{u}{v}) f_9(v)\mathrm{d}v$, with the above f_8 and f_9, agrees with the integral to be evaluated in (40). The Mellin transform of f_9 is given by

$$M_{f_9}(s) = \int_0^\infty x_2^{s-1} x_2^\gamma [1 + a_1(q_1-1)x_2^{\delta_1}]^{-\frac{\eta_1}{q_1-1}+1}\mathrm{d}x_2 = \frac{\Gamma(\frac{s+\gamma}{\delta_1})\Gamma(\frac{\eta_1}{q_1-1} - 1 - \frac{s+\gamma}{\delta_1})}{\delta_1[a_1(q_1-1)]^{\frac{\gamma+s}{\delta_1}}\Gamma(\frac{\eta_1}{q_1-1} - 1)}$$

for $\Re(s+\gamma) > 0, \Re(\frac{\eta_1}{q_1-1} - 1 - \frac{s+\gamma}{\delta_1}) > 0, \eta_1 > 0, \eta_1 + 1 - q_1 > 0$. Now, $M_{g_2}(s) = M_{f_8}(s)M_{f_9}(s)$. Then, taking the inverse Mellin transform, for $\tilde{c} = [\delta_1\delta_2[a_1(q_1-1)]^{\frac{\gamma}{\delta_1}}\Gamma(\frac{\eta_1}{q_1-1} - 1)\Gamma(\frac{\eta_2}{q_2-1} - 1)]^{-1}$,

$$\begin{aligned} g_2 &= \tilde{c}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma(\frac{s}{\delta_2})\Gamma(\frac{s+\gamma}{\delta_1})\Gamma(\frac{\eta_2}{q_2-1} - 1 - \frac{s}{\delta_2}) \\ &\times \Gamma(\frac{\eta_1}{q_1-1} - 1 - \frac{\gamma+s}{\delta_1})\{[a_1(q_1-1)]^{\frac{1}{\delta_1}}[a_2(q_2-1)]^{\frac{1}{\delta_2}}\}^{-s}\mathrm{d}s \qquad (42) \\ &= \tilde{c}H_{2,2}^{2,2}\left[[a_1(q_1-1)]^{\frac{1}{\delta_1}}[a_2(q_2-1)]^{\frac{1}{\delta_2}}\middle|\begin{matrix}(2-\frac{\eta_2}{q_2-1},\frac{1}{\delta_2}),(2+\frac{\gamma}{\delta_1}-\frac{\eta_1}{q_1-1},\frac{1}{\delta_1}) \\ (0,\frac{1}{\delta_2}),(\frac{\gamma}{\delta_1},\frac{1}{\delta_1})\end{matrix}\right] \qquad (43) \end{aligned}$$

for $[a_1(q_1-1)]^{\frac{1}{\delta_1}}[a_2(q_2-1)]^{\frac{1}{\delta_2}} < 1$. We can also obtain series forms here.

3.1. Limiting forms of this special case

When $q_1 \to 1$ and $q_2 \to 1$ then we have the following integral for the ρ-th moment of the time to failure:

$$I_3 = \rho\int_0^\infty t^{\rho-1}[\mathrm{e}^{-a_1\eta_1 t^{\delta_1}}][1 - \mathrm{e}^{-a_2\eta_2 t^{-\delta_2}}]\mathrm{d}t \qquad (44)$$

In order to evaluate (44) we will consider the following general integral:

$$g = \int_0^\infty t^\gamma \mathrm{e}^{-b_1 t^{\delta_1} - b_2 t^{-\delta_2}}\mathrm{d}t \qquad (45)$$

for $b_j > 0, \delta_j > 0, j = 1, 2$. This integral in (45) is connected to many problems in different fields. For $\delta_1 = 1, \delta_2 = 1$ it is the basic Krätzel integral, see [17], [18], [19] and [20]. For $\delta_1 = 1, \delta_2 = \frac{1}{2}$ it is the reaction-rate probability integral in nuclear reaction-rate theory, see [21]. The integrand in (45) for $\delta_1 = 1, \delta_2 = 1$, normalized, is the inverse Gaussian density in stochastic processes. Hence (45) is a generalization of all these basic integrals. This integral can be explicitly evaluated by treating it as a Mellin convolution of a product. Let $u = x_1x_2, v = x_2$ or $x_2 = v, x_1 = \frac{u}{v}$ with Jacobian $\frac{1}{v}$. Since the integrand in (45) is a product of positive integrable functions, by multiplying with appropriate normalizing constants, one can create statistical densities out of them. Hence we can treat the Mellin convolution of a product as the statistical problem of computing the density of a product of two statistically independently distributed real positive scalar random variables. Then $E(u^{s-1}) = [E(x_1^{s-1})][E(x_2^{s-1})]$ where E denotes the expected value, or in terms of the Mellin transforms, $M_g(s) = M_{f_{10}}(s)M_{f_{11}}(s)$ where s is the Mellin parameter. Then g has the structure

$$g = \int_v \frac{1}{v} f_{10}(\frac{u}{v}) f_{11}(v)\mathrm{d}v. \qquad (46)$$

Take

$$f_{10}(x_1) = \mathrm{e}^{-x_1^{\delta_2}} \Rightarrow f_{10}(\frac{u}{v}) = \mathrm{e}^{-b_2 v^{-\delta_2}} \tag{47}$$

where $u = b_2^{\frac{1}{\delta_2}} = (a_2\eta_2)^{\frac{1}{\delta_2}}$. Then the Mellin transform of f_1 is of the form

$$M_{f_{10}}(s) = \int_0^\infty \mathrm{e}^{-x^{\delta_2}} \mathrm{d}x = \frac{1}{\delta_2}\Gamma(\frac{s}{\delta_2}), \Re(s) > 0. \tag{48}$$

Take

$$\begin{aligned} f_{11}(x) &= x^{\gamma+1}\mathrm{e}^{-b_1 x^{\delta_1}} \Rightarrow & (49)\\ M_{f_{11}}(s) &= \int_0^\infty x^{\gamma+1+s-1}\mathrm{e}^{-b_1 x^{\delta_1}} \mathrm{d}x \\ &= [\delta_1 b_1^{\frac{s+\gamma+1}{\delta_1}}]^{-1}\Gamma(\frac{s+\gamma+1}{\delta_1}), \Re(s+\gamma+1) > 0. & (50) \end{aligned}$$

Observe that f_{10} from (46) and f_2 from (49), when substituted in (46) gives the integral to be evaluated in (50), which by the Mellin convolution of a product is the inverse Mellin transform of the product $M_{f_{10}}(s)M_{f_{11}}(s)$, available from (48) and (50). Therefore the integral in (50) is given by

$$\begin{aligned} g &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} M_{f_{10}}(s)M_{f_{11}}(s)u^{-s}\mathrm{d}s \\ &= \bar{c}\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma(\frac{s}{\delta_2})\Gamma(\frac{\gamma+1}{\delta_1}+\frac{s}{\delta_1})(b_1^{\frac{1}{\delta_1}}b_2^{\frac{1}{\delta_2}})^{-s}\mathrm{d}s \\ &= \bar{c}H_{0,2}^{2,0}\left[b_1^{\frac{1}{\delta_1}}b_2^{\frac{1}{\delta_2}}\Big|_{(0,\frac{1}{\delta_2}),(\frac{\gamma+1}{\delta_1},\frac{1}{\delta_1})}\right], \end{aligned} \tag{51}$$

where $\bar{c} = [\delta_1\delta_2 b_1^{\frac{\gamma+1}{\delta_1}}]^{-1}$. When the poles are simple, (51) can be written as a sum of two series. When $\frac{1}{\delta_1} = m_1$ and $\frac{1}{\delta_2} = m_2$ where $m_1, m_2 = 1, 2, ...$ (positive integers) then the H−function in (51) can be written as a G−function and can be evaluated in explicit series forms. In the reaction rate probability integral $\delta_1 = 1$ and $\frac{1}{\delta_2} = 2$ and this problem is of the above type and explicit series forms may be seen from [21].

4. Multi-component stress-strength Reliability

A system containing more than one component is referred to as a multi-component system. It may consist of parallel or series components, or it may involve an intricate combination of both. Many real-world applications of MSS models may be found in areas including industrial processes, military technology, communication networks, etc. For example, a person may survive with only one healthy kidney, hence, kidney function in the human body is a one-out-of-two system. The MSS system functions when at least $s(1 \le s \le k)$ of its k identical and independent strength components function properly against a common strength. Let $X_1, X_2, \ldots, X_k$ be independent random variables with a common distribution function $F(.)$ and subjected to the common random stress Y with a distribution function $G(.)$. Thus the system reliability in a Multi-component stress strength model $R_{s,k}$ is given by

$$\begin{aligned} R_{s,k} &= P\left[\text{at least } s \text{ of} X_1, X_2, \ldots, X_k \text{ exceed } Y\right] \\ &= \sum_{i=s}^{k}\binom{k}{i}(P\left[X_i > Y\right])^i(P\left[X_i \le Y\right])^{k-i} \\ &= \sum_{i=s}^{k}\binom{k}{i}\int_{-\infty}^{\infty}\left[1-F(y)\right]^i\left[F(y)\right]^{k-i}\mathrm{d}G(y) \end{aligned} \tag{52}$$

The multi−component system reliability given in equation (1) was first introduced by Bhattacharyya and Johnson [22]. After that, many authors have shown considerable interest in the multi-component stress−strength reliability for details refer [23], [24] etc.
In many complex systems that emerge in the domains of biology, chemistry, economics, geography, medicine, physics, etc., modelling and analysing lifetime data are crucial. The literature introduces a variety of q-type distributions for modeling lifetime data, the most prominent of which are the q-exponential, q-gamma, q-Gaussian etc, see [25] and [26], q-Weibull refer [27] and q-K-distribution, see [14]. The basic motivation for constructing statistical distributions for modelling lifetime data is the ability to model both monotonic and non-monotonic failure rates, even though the baseline failure rate may be monotonic. The Weibull distribution is most commonly used to describe lifetime data, which can only exhibit monotonic and constant shapes for its hazard rate function. However, the q-Weibull distribution can exhibit unimodal, bathtub-shaped, monotonically decreasing, monotonically increasing, and constant shapes for its hazard rate function. Hence, it is a useful generalization of the Weibull distribution. Here we discuss a classical inference on the multi-component stress-strength reliability when the stress and strength components are independent random variables distributed as (11). Then the Multi-component stress strength system reliability $R_{s,k}$ is given by

$$\begin{aligned} R_{s,k} &= \sum_{i=s}^{k} \binom{k}{i} \int_{-\infty}^{\infty} [1-F(y)]^i [F(y)]^{k-i} \, \mathrm{d}G(y) \\ &= \sum_{i=s}^{k} \binom{k}{i} \int_{0}^{\infty} \left[(1+\alpha(q-1)y^{\delta})^{-\frac{\eta_1}{q-1}+1}\right]^i \left[1-(1+\alpha(q-1)y^{\delta})^{-\frac{\eta_1}{q-1}+1}\right]^{k-i} \\ &\times \quad \alpha\delta(\eta_2+1-q)y^{\delta-1}\left[(1+\alpha(q-1)y^{\delta})^{-\frac{\eta_2}{q-1}}\right] \mathrm{d}y \end{aligned} \tag{53}$$

After simplification, we get

$$R_{s,k} = \frac{(\eta_2+1-q)}{(\eta_1+1-q)} \sum_{i=s}^{k} \binom{k}{i} \mathbf{B}\left(\frac{(\eta_2+1-q)}{(\eta_1+1-q)}, k-i+1\right) \text{ for } \frac{\eta_2+1-q}{\eta_1+1-q} > 0. \tag{54}$$

In this section, we created random samples from stress and strength variables for various parameter values and sample size combinations. In three scenarios, $(s,k) = (1,3), (2,6)$, and $(3,7)$ we estimated the MSS reliability. Tables 1 present the estimated values, bias, and mean square error (MSE).

multirow graphicx lscape

Table 1: *The MLE, Bias and SE of the estimator of $R_{s,k}$*

n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
15	0.120119	0.020119	0.000405	0.064679	0.012048	0.000145	0.058502	0.013048	0.000170
20	0.093342	0.006658	0.000044	0.049380	0.003252	0.000011	0.037616	0.007838	0.000061
25	0.094298	0.005702	0.000033	0.049759	0.002873	0.000008	0.038526	0.006928	0.000048
30	0.104346	0.004346	0.000019	0.055473	0.002842	0.000008	0.050787	0.005333	0.000028
35	0.101434	0.001434	0.000002	0.053631	0.001000	0.000001	0.044765	0.000689	0.000001
n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
50	0.193680	0.011861	0.000141	0.107388	0.007388	0.000055	0.093504	0.006548	0.000043
100	0.170321	0.011497	0.000132	0.093361	0.006639	0.000044	0.081143	0.005813	0.000034
125	0.186056	0.004238	0.000018	0.102703	0.002703	0.000007	0.089360	0.002404	0.000006
200	0.178269	0.003549	0.000013	0.097947	0.002054	0.000004	0.085158	0.001798	0.000003
250	0.181112	0.000706	0.000001	0.099623	0.000377	0.000000	0.086631	0.000325	0.000000

n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
25	0.338753	0.018390	0.000338	0.205363	0.012029	0.000145	0.181581	0.010727	0.000115
150	0.347079	0.010064	0.000101	0.210166	0.007226	0.000052	0.185749	0.006559	0.000043
250	0.353725	0.003418	0.000012	0.215051	0.002341	0.000006	0.190202	0.002105	0.000004
300	0.356246	0.000897	0.000001	0.216886	0.000506	0.000000	0.191872	0.000435	0.000000
800	0.356669	0.000474	0.000000	0.217124	0.000268	0.000000	0.192077	0.000231	0.000000
n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
25	0.233742	0.016258	0.000264	0.133105	0.009752	0.000095	0.116415	0.008585	0.000074
150	0.240789	0.009211	0.000085	0.136981	0.005876	0.000035	0.119772	0.005228	0.000027
250	0.253195	0.003195	0.000010	0.145040	0.002183	0.000005	0.126965	0.001965	0.000004
300	0.247085	0.002916	0.000009	0.140967	0.001891	0.000004	0.123313	0.001687	0.000003
800	0.248977	0.001023	0.000001	0.142251	0.000606	0.000000	0.124468	0.000532	0.000000
n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
50	0.534931	0.034931	0.001220	0.367358	0.034024	0.001158	0.332781	0.032781	0.001075
100	0.478824	0.021176	0.000448	0.315918	0.017415	0.000303	0.283795	0.016205	0.000263
200	0.488284	0.011716	0.000137	0.323447	0.009887	0.000098	0.290755	0.009245	0.000086
450	0.496283	0.003717	0.000014	0.330132	0.003202	0.000010	0.296994	0.003006	0.000009
700	0.499025	0.000975	0.000001	0.332562	0.000771	0.000001	0.299289	0.000711	0.000001
n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
50	0.357590	0.024256	0.000588	0.218419	0.018419	0.000339	0.193363	0.016892	0.000285
100	0.315656	0.017677	0.000313	0.188037	0.011963	0.000143	0.165723	0.010748	0.000116
200	0.326700	0.006633	0.000044	0.195442	0.004558	0.000021	0.172365	0.004106	0.000017
450	0.328412	0.004921	0.000024	0.196541	0.003459	0.000012	0.173343	0.003128	0.000010
700	0.332360	0.000973	0.000001	0.199353	0.000647	0.000000	0.175891	0.000580	0.000000
n	(s,k)=(1,3)			(s,k)=(2,6)			(s,k)=(3,7)		
	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE	R-MLE	R-Bias	R-MSE
50	0.177839	0.011172	0.000125	0.097755	0.006846	0.000047	0.084999	0.006052	0.000037
100	0.155884	0.010783	0.000116	0.084773	0.006136	0.000038	0.073587	0.005361	0.000029
300	0.170740	0.004074	0.000017	0.093368	0.002459	0.000006	0.081116	0.002169	0.000005
500	0.162740	0.003926	0.000015	0.088604	0.002305	0.000005	0.076923	0.002024	0.000004
700	0.166253	0.000414	0.000000	0.090677	0.000232	0.000000	0.078745	0.000202	0.000000

5. Real Data Application

In this section, we explore an actual data set to illustrate the flexibility of the proposed model. The information displays, in months, how long 128 bladder cancer patients were in remission. The data set is given in Table 2.

Table 2: *Remission times of bladder cancer patients data*

0.08	2.09	13.29	0.4	2.26	3.57	5.06	7.09	9.22	13.8	25.74	0.5
3.48	4.87	23.63	0.2	2.23	6.94	8.66	13.11	3.52	4.98	6.97	9.02
3.88	5.32	7.39	10.34	14.83	34.26	0.9	2.69	4.18	5.34	7.59	10.66
2.46	3.64	5.09	7.26	9.47	14.24	25.82	0.51	2.54	3.7	5.17	7.28
15.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01	1.19	2.75
9.74	14.76	26.31	0.81	2.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64
11.79	18.1	1.46	4.4	5.85	8.26	11.98	19.13	1.76	3.25	4.5	6.25
79.05	1.35	6.76	17.14	2.87	5.62	7.87	11.64	17.36	1.4	3.02	4.34
5.71	7.93	22.69	4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33	5.49
7.66	11.25	21.73	2.07	3.36	6.93	8.37	12.02	2.02	12.07	20.28	2.02
3.36	12.03	3.31	4.51	6.54	8.53	8.65	12.63				

We compare the proposed model's goodness of fit to a few competing models, such as Weibull, Frechet Weibull, transmuted Weibull, and modified Weibull (MW), using a few discrimination criteria, such as the Akaike Information Criterion (AIC), Anderson Darling test (AD-test), Cram©r-von Mises test (CRVM), and Kolmogorov-Smirnov test with its p-value. The MLEs of the parameters, as well as the values of the AIC, are provided in Tables 3 and 4, respectively. These findings suggest that the proposed model is the best model because it has the lowest test statistic values among all fitted models. The plots of the fitted PDF, CDF, P-P plot, and Q-Q plot for the proposed distribution and Weibull distribution are displayed in Figure 9.

Table 3: *The estimated value of the parameters of the fitted model.*

Generalized q-Weibull	q=2.5925	$\eta = 4.8920$	$\alpha = 0.0179$	$\delta = 1.4273$
Weibull	$\alpha = 1.0478$	$\beta = 9.5607$	—-	——
Frechet Weibull (FW	$\alpha = 1.1446$	$\beta = 1.881$	——	——
Transmuted Weibull	$\alpha = 1.1333$	$\beta = 14.6198$	$\lambda = 0.7449$	——
Modified Weibull	alpha=1.3172	$\beta = 0 : 0938$	$\lambda = 1.4783$	——

Table 4: *The value of AIC, AD-test, CRVM-test, KS-test, and p-value of the fitted model.*

Model	AIC	AD-test	CRVM-test	KS-test	p=value
Generalized q-Weibull	827.4798	0.12177	0.01758	0.03504	0.99943
Weibull	832.174	0.957709	0.153703	0.0700169	0.556965
Frechet Weibull (FW)	896.002	6.11825	0.978722	0.140799	0.0125018
Transmuted Weibull	829.917	0.560038	0.0879162	0.0587652	0.76866
Modified Weibull	834.174	0.957709	0.153703	0.0700169	0.556965

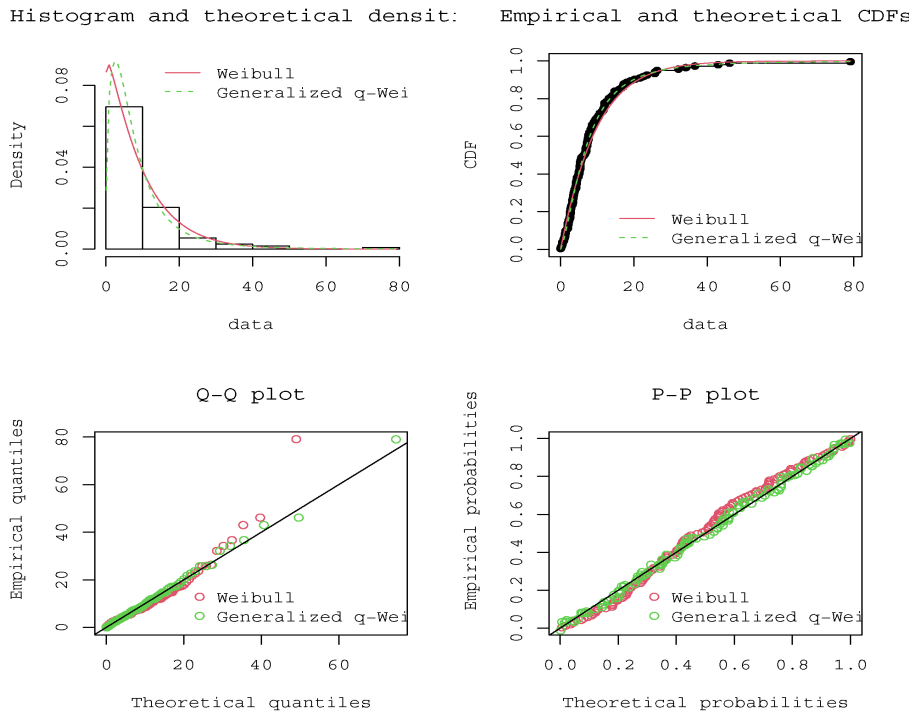


Figure 9: The histogram and theoretical densities, empirical and theoretical CDFs, Q-Q plots, P-P plots of the fitted data.

6. Conclusion

In this study, we take into account a system with k-connected components in series. The lifetimes of the components, $X_1, ..., X_k$, are randomly distributed and have pathway densities for the pathway parameters $q < 1$, $q > 1$, or $q \to 1$. Then, the survival function, hazard function, expected time to failure, and general moments of $x = \min\{X_1, X_2, \ldots, X_n\}$ are computed. It is demonstrated that the hazard function can take on various shapes, including a bathtub shape. The estimation of stress-strength reliability is assessed through the maximum likelihood estimation technique when both stress and strength variables conform to the pathway model. Remission time data from cancer patients is examined to see how the model is relevant in practical situations. The proposed distribution consistently provides better fits for real data compared to other models.

References

[1] Barlow, R.E. and Prochan, F. Statistical Theory of Reliability and Life Testing: Probability Models, Silver Spring, 1981.
[2] Jelinski, Z. and Moranda, P. B. Software Reliability Research in Statistical Computer Performance Evaluation, edited by W. Freiberer, Academic Press, New York, 1972.
[3] Meeker, W. Q. and Escobar, L. A. Statistical Models for Reliability Data, Wiley, New York, 1998.
[4] Balakrishnan, N. and Lai, C.D. Continuous Bivariate Distributions, Springer, New York 2009.
[5] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62: 30-44.
[6] Wang, R.T. (2012). A reliability model for the multivariate exponential distribution. *Journal of Multivariate Analysis*, 98: 1033-1042.
[7] Mathai, A. M. (2005). A pathway to matrix-variate gamma and normal densities. *Linear Algebra and its Applications*, 396: 317-328.
[8] Mathai, A. M. and Provost, S. B. (2006). Some complex matrix-variate statistical distributions on rectangular matrices. *Linear Algebra and its Applications*, 410: 198-216.
[9] Mathai, A. M. (2003). Order statistics from logistic distribution and applications to survival & reliability analysis, *IEEE Transactions on Reliability*, 52: 200-206.
[10] Tsallis, C. Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World, Springer, New York, 2009.
[11] Beck, C. and Cohen, E. G. D. (2003). Superstatistics. *Physica A*, 322: 267-275.
[12] Princy, T. (2015). An extended compound gamma model and application to composite fading channels. *Statistics, Optimization & Information Computing*, 3: 42-53.
[13] Princy, T. (2016). Pathway Extension of Weibull-Gamma Model. *Journal of Statistical Theory and Applications*, 15: 47-60.
[14] Princy, T. (2016). Modeling SAR Images by Using a Pathway Model. *Journal of the Indian Society of Remote Sensing*, 2: 353-360.
[15] Mathai, A. M. A Handbook of Generalized Special Functions for Statistical and Physical Sciences. Oxford University Press, Oxford, 1993.
[16] Mathai, A. M., Saxena, R. K. and Haubold, H. J. The H-function: Theory and Applications, Springer, New York, 2010.
[17] Krätzel, E. (1979): Integral transforms of Bessel type, In Generalized Functions and Operational Calculus, *Proc. Conf. Verna*, Belg.Acad. Sci., Sofia, 148-165.
[18] Princy, T. (2014). Krätzel function and related statistical distributions. *Communications in Mathematics and Statistics*, 2: 413-429.
[19] Princy, T. (2015). Mixture models and the Krätzel integral transform. *Communications in Statistics—Theory and Methods*, 44: 390-405.
[20] Mathai, A. M. (2012). Generalized Krätzel integral and associated statistical densiteis. *International Journal of Mathematical Analysis*, 6: 2501-2510.

[21] Mathai, A. M. and Haubold, H. J., Modern Problems in Nuclear and Neutrino Astrophysics, Akademie-Verlag, Berlin, 1988.
[22] Bhattacharyya, G. K. and Johnson, R. A. (1974). Estimation of reliability in a multicomponent stress-strength model. *Journal of the American Statistical Association*, 69: 966-970.
[23] Mahto, A. K., Tripathi, Y. M., & Kizilaslan, F. (2020). Estimation of reliability in a multicomponent stress−strength model for a general class of inverted exponentiated distributions under progressive censoring. *Journal of Statistical Theory and Practice*, 14: 1-35.
[24] Azhad, Q. J., Arshad, M. and Khandelwal, N. (2022). Statistical inference of reliability in multicomponent stress strength model for Pareto distribution based on upper record values. *International Journal of Modelling and Simulation*, 42: 319-334.
[25] Picoli, S., Mendes, R. S., Malacarne, L. C. and Santos, R. P. B. (2009). q-distributions in complex systems: A brief review. *Brazilian Journal of Physics*, 39: 468-474.
[26] Princy, T. (2023). Some Useful Pathway Models For Reliability Analysis. *Reliability: Theory & Applications*, 18: 340-359.
[27] Jose, K. K. and Naik, S. R. (2009). On the q-Weibull distribution and its applications. *Communications in Statistics−Theory and Methods*, 38: 912-926.

RT&A, No 2 (78)
Volume 19, June, 2024

ON AN INTERNAL DEPENDENCE OF SIMULTANEOUS MEASUREMENTS

Valentin Vankov Iliev

•

Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia, Bulgaria
viliev@math.bas.bg

Abstract

In this paper we show that there exists an internal dependence of the simultaneous measurements made by the two pairs of linear polarizers operated in each leg of the apparatus in Aspect's version of Einstein-Podolsky-Rosen Gedankenexperiment. *The corresponding Shannon-Kolmogorov's information flow linking a polarizer from one leg to a polarizer from the other leg is proportional to the absolute value of this function of dependence. It turns out that if Bell's inequality is violated, then this information flow is strictly positive, that is, the experiment performed at one leg is informationally dependent on the experiment at the other leg. By throwing out the sign of absolute value, we define the signed information flow linking a polarizer from one leg to a polarizer from the other leg which, in turn, reproduces the probabilities of the four outcomes of the simultaneous measurements, predicted by quantum mechanics. We make an attempt to illustrate the seeming random relation between the total information flow, the total signed information flow, and the violation of Bell's inequality in terms of a kind of uncertainty principle.*

Keywords: EPR thought experiment, Aspect's optical version, Informational dependence, Bell's inequality.

1. Introduction, Notation

1.1. Introduction

In the context of the bipartite quantum system that describes Aspect's optical version of Einstein-Podolsky-Rosen *Gedankenexperiment* (see [1] and [5]), we consider the pairs of linear polarizers operated in each leg of the apparatus as pairs of self-adjoined linear operators

$$A_{\mu_i} = \begin{pmatrix} \cos\mu_i & \sin\mu_i \\ \sin\mu_i & -\cos\mu_i \end{pmatrix}, B_{\nu_j} = A_{\nu_j},$$

where $\mu_i, \nu_j \in [0, \pi]$, $i, j = 1, 2$, are the angles of the polarizers. Note that each pair has a time switch which interchanges polarizers, the corresponding time being shorter than the time necessary for a light signal to travel from one of the pairs of polarizers to the other (Einstein locality assumption for independence).

Each pair of operators A_{μ_1}, A_{μ_2} and B_{ν_1}, B_{ν_2} acts on the state space of the corresponding quantum subsystem (a unitary plane). By tensoring with the unit operator on the other plane, we obtain two pairs of self-adjoined linear operators $\mathcal{A}_{\mu_1}$, $\mathcal{A}_{\mu_2}$ and $\mathcal{B}_{\nu_1}$, $\mathcal{B}_{\nu_2}$ with spectre $\{1, -1\}$ on the state space of the whole quantum system (tensor product of the two unitary planes). Moreover, for each $i, j = 1, 2$ the operators $\mathcal{A}_{\mu_i}$ and $\mathcal{B}_{\nu_j}$ commute because the state space of the

whole quantum system has an orthonormal frame consisting of eigenvectors of both operators. In this case the corresponding measurements are said to be *simultaneous*.

In accord with the axiom of quantum mechanics about the observables, after fixing initial state we can consider the members of this frame as outcomes of a sample space with probability assignment consisting of probabilities predicted by this axiom. Moreover, with an abuse of the language, we can also consider the operators as random variables with range $\{1,-1\}$ on this sample space. Under the condition that the singlet state is initial, any one of these random variables has probability distribution $(\frac{1}{2},\frac{1}{2})$. Moreover, if $\mu\in\{\mu_1,\mu_2\}$ and $\nu\in\{\nu_1,\nu_2\}$, then

$$\mathrm{pr}((\mathcal{A}_\mu=1)\cap(\mathcal{B}_\nu=1))=\mathrm{pr}((\mathcal{A}_\mu=-1)\cap(\mathcal{B}_\nu=-1))=\frac{1}{2}\sin^2\left(\frac{\mu-\nu}{2}\right),$$

$$\mathrm{pr}((\mathcal{A}_\mu=1)\cap(\mathcal{B}_\nu=-1))=\mathrm{pr}((\mathcal{A}_\mu=-1)\cap(\mathcal{B}_\nu=1))=\frac{1}{2}\cos^2\left(\frac{\mu-\nu}{2}\right).$$

Therefore, the product of random variables $\mathcal{A}_\mu\mathcal{B}_\nu$ has probability distribution $\left(\sin^2\left(\frac{\mu-\nu}{2}\right),\cos^2\left(\frac{\mu-\nu}{2}\right)\right)$ and expected value $\mathcal{E}(\mathcal{A}_\mu\mathcal{B}_\nu)=-\cos(\mu-\nu)$.

On the other hand, the joint experiment (see [7, Part I, Section 6]) of the binary trials $\mathfrak{A}_\mu=(\mathcal{A}_\mu=1)\cup(\mathcal{A}_\mu=-1)$ and $\mathfrak{B}_\nu=(\mathcal{B}_\nu=1)\cup(\mathcal{B}_\nu=-1)$ produces the probability distribution

$$\left(\frac{1}{2}\sin^2\left(\frac{\mu-\nu}{2}\right),\frac{1}{2}\cos^2\left(\frac{\mu-\nu}{2}\right),\frac{1}{2}\cos^2\left(\frac{\mu-\nu}{2}\right),\frac{1}{2}\sin^2\left(\frac{\mu-\nu}{2}\right)\right)$$

with Boltzmann-Shannon entropy $E(\theta_{\mu,\nu})$, where $E(\theta)=-2\theta\ln\theta-2(\frac{1}{2}-\theta)\ln(\frac{1}{2}-\theta)$ and $\theta_{\mu,\nu}=\frac{1}{2}\sin^2\left(\frac{\mu-\nu}{2}\right)$. We extend the function $E(\theta)$, $\theta\in(0,\frac{1}{2})$, (see [6, 4.1,5.1]) as continuous on the closed interval $[0,\frac{1}{2}]$.

By modifying the entropy function $E(\theta)$, we obtain the strictly increasing *degree of dependence function* $e\colon[0,\frac{1}{2}]\to[-1,1]$, which mimics the regression coefficient (see [6, 5.2]).

It turns out that the average quantity of information $I(\mathfrak{A}_\mu,\mathfrak{B}_\nu)$ (see [3, §1]) of one of the experiments $\mathfrak{A}_\mu$ and $\mathfrak{B}_\nu$, relative to the other can be found by the formula $I(\mathfrak{A}_\mu,\mathfrak{B}_\nu)=|e(\theta_{\mu,\nu})|\ln 2$. We can consider $I(\mathfrak{A}_\mu,\mathfrak{B}_\nu)$ as a measure of the flow carrying information between these two binary trials (see [6, 5.3]). Since s is an invertible function, the corresponding *signed information flow* $I^{(s)}(\mathcal{A}_\mu,\mathcal{B}_\nu)(\theta)=e(\theta)\ln 2$ replicates the probability distribution (2) produced by quantum mechanics.

In terms of Aspect's experiment, the sum $I(\mathfrak{A},\mathfrak{B})=\sum_{i,j}^{2}I(\mathfrak{A}_{\mu_i},\mathfrak{B}_{\nu_j})$ (called total information flow) can be thought about as a measure of the flow carrying information between the two pairs of polarizers. In his paper [2] John Bell deduced under the assumptions of "locality" and "realism" that if measurements are performed independently (Einstein locality assumption for independence) on the two separated particles (photons in Aspect's experiment) of an entangled pair, then the assumption that the outcomes depend upon "hidden variables" implies constraint condition called Bell's inequality (see Subsection 4.1). It comes out that if Bell's inequality is violated, then the total information flow is strictly positive. In other words, in this case there exists an informational dependence between the two legs of apparatus.

In the end of the paper we discuss the relation between the information flow $I(\mathfrak{A},\mathfrak{B})$ and the violation of Bell's inequality. Using Examples 1 and the Java program from the link that can be found there, we note that this relation is subject to a kind of uncertainty principle.

1.2. Notation

$\mathcal{H}$: 2-dimensional unitary space with inner product $\langle x|y\rangle$ which is linear in the second slot and anti-linear in the first slot;
$\mathbb{I}=\mathbb{I}_{\mathcal{H}}$: the identity linear operator on $\mathcal{H}$;
$\mathcal{H}^{\otimes 2}=\mathcal{H}\otimes\mathcal{H}$: the unitary tensor square with inner product $\langle x_1\otimes x_2|y_1\otimes y_2\rangle=\langle x_1|y_1\rangle\langle x_2|y_2\rangle$;

$\mathcal{U}^{(2)}$: the unit sphere in $\mathcal{H}^{\otimes 2}$;
$\mathrm{Spec}(A)$: the real spectre of a self-adjoined linear operator A on $\mathcal{H}$ with trace zero, having the form $\mathrm{Spec}(A) = \{\lambda_1^{(A)}, \lambda_2^{(A)}\}$, $\lambda_1^{(A)} + \lambda_2^{(A)} = 0$;
$u^{(A)} = \{u_1^{(A)}, u_2^{(A)}\}$: the orthonormal frame for $\mathcal{H}$, formed by the corresponding eigenvectors of A;
$\mathcal{H}_i^{(A)}$: the eigenspaces $\mathbb{C}u_i^{(A)}$ of A, $i = 1, 2$.

2. Self-Adjoint Operators on $\mathcal{H}$

2.1. Two Special Commuting Operators

We fix an orthonormal frame $h = \{h_1, h_2\}$ for $\mathcal{H}$ and identify the self-adjoined operators with their matrices with respect to h. For any $\mu \in [0, \pi]$ we denote by A_μ the self-adjoined operator

$$\begin{pmatrix} \cos\mu & \sin\mu \\ \sin\mu & -\cos\mu \end{pmatrix}.$$

We have $\lambda_1^{(A_\mu)} = 1, \lambda_2^{(A_\mu)} = -1$, and

$$u_1^{(A_\mu)} = (\cos\frac{\mu}{2})h_1 + (\sin\frac{\mu}{2})h_2, u_2^{(A_\mu)} = (-\sin\frac{\mu}{2})h_1 + (\cos\frac{\mu}{2})h_2.$$

For any $\nu \in [0, \pi]$ we set $B_\nu = A_\nu$.

Note that $\{h_1 \otimes h_1, h_1 \otimes h_2, h_2 \otimes h_1, h_2 \otimes h_2\}$ and $u^{(A_\mu)} \otimes u^{(B_\nu)} = \{u_1^{(A_\mu)} \otimes u_1^{(B_\nu)}, u_1^{(A_\mu)} \otimes u_2^{(B_\nu)}, u_2^{(A_\mu)} \otimes u_1^{(B_\nu)}, u_2^{(A_\mu)} \otimes u_2^{(B_\nu)}\}$ are orthonormal frames for $\mathcal{H}^{\otimes 2}$.

Let us set $\mathcal{A}_\mu = A_\mu \otimes \mathbb{I}$, $\mathcal{B}_\nu = \mathbb{I} \otimes B_\nu$. It is a straightforward check that the last two linear operators on $\mathcal{H}^{\otimes 2}$ are also self-adjoined with $\lambda_1^{(\mathcal{A}_\mu)} = \lambda_1^{(\mathcal{B}_\nu)} = 1$, $\lambda_2^{(\mathcal{A}_\mu)} = \lambda_2^{(\mathcal{B}_\nu)} = -1$, the $\lambda_i^{(\mathcal{A}_\mu)}$-eigenspace $\mathcal{H}_i^{(\mathcal{A}_\mu)} = \mathcal{H}_i^{(A_\mu)} \otimes \mathcal{H}$ has orthonormal frame $\{u_i^{(A_\mu)} \otimes u_1^{(B_\nu)}, u_i^{(A_\mu)} \otimes u_2^{(B_\nu)}\}$, and the $\lambda_i^{(\mathcal{B}_\nu)}$-eigenspace $\mathcal{H}_i^{(\mathcal{B}_\nu)} = \mathcal{H} \otimes \mathcal{H}_i^{(B_\nu)}$ has orthonormal frame $\{u_1^{(A_\mu)} \otimes u_i^{(B_\nu)}, u_2^{(A_\mu)} \otimes u_i^{(B_\nu)}\}$, $i = 1, 2$.

Since $u^{(A_\mu)} \otimes u^{(B_\nu)}$ is an orthonornal frame of $\mathcal{H}^{\otimes 2}$ consisting of eigenvectors of both $\mathcal{A}_\mu$ and $\mathcal{B}_\nu$, then the last two operators commute.

Let $\psi \in \mathcal{U}^{(2)}$ and let $S(\psi; \mathcal{A}_\mu, \mathcal{B}_\nu)$ be the sample space with set of outcomes $u^{(A_\mu)} \otimes u^{(B_\nu)} = \{u_1^{(A_\mu)} \otimes u_1^{(B_\nu)}, u_1^{(A_\mu)} \otimes u_2^{(B_\nu)}, u_2^{(A_\mu)} \otimes u_1^{(B_\nu)}, u_2^{(A_\mu)} \otimes u_2^{(B_\nu)}\}$ and probability assignment $\{p_{11}, p_{12}, p_{21}, p_{22}\}$ with $p_{ij} = |\langle u_i^{(A_\mu)} \otimes u_j^{(B_\nu)} | \psi \rangle|^2$, $i, j = 1, 2$. With an abuse of the language, we consider the observable $\mathcal{A}_\mu$ as a random variable $\mathcal{A}_\mu \colon u^{(A_\mu)} \otimes u^{(B_\nu)} \to \mathbb{R}$, $\mathcal{A}_\mu(u_1^{(A_\mu)} \otimes u_j^{(B_\nu)}) = \lambda_1^{(\mathcal{A}_\mu)}$, $\mathcal{A}_\mu(u_2^{(A_\mu)} \otimes u_j^{(B_\nu)}) = \lambda_2^{(\mathcal{A}_\mu)}$, $j = 1, 2$, on the sample space $S(\psi; \mathcal{A}_\mu, \mathcal{B}_\nu)$ with probability distribution $p_{\mathcal{A}_\mu}(\lambda_i^{(A)}) = |\langle u_i^{(A_\mu)} \otimes u_1^{(B_\nu)} | \psi \rangle|^2 + |\langle u_i^{(A_\mu)} \otimes u_2^{(B_\nu)} | \psi \rangle|^2$, $i = 1, 2$, and $p_{\mathcal{A}_\mu}(\lambda) = 0$ for $\lambda \notin \mathrm{Spec}(\mathcal{A}_\mu)$. Identifying the event $\{u_i^{(A_\mu)} \otimes u_1^{(B_\nu)}, u_i^{(A_\mu)} \otimes u_2^{(B_\nu)}\}$ with the "event" $\mathcal{A}_\mu = \lambda_i^{(\mathcal{A}_\mu)}$, we have $\mathrm{pr}(\mathcal{A}_\mu = \lambda_i^{(\mathcal{A}_\mu)}) = |\langle u_i^{(A_\mu)} \otimes u_1^{(B_\nu)} | \psi \rangle|^2 + |\langle u_i^{(A_\mu)} \otimes u_2^{(B_\nu)} | \psi \rangle|^2$, $i = 1, 2$.

We also consider the observable $\mathcal{B}_\nu$ as a random variable $\mathcal{B}_\nu \colon u^{(A_\mu)} \otimes u^{(B_\nu)} \to \mathbb{R}$, $\mathcal{B}_\nu(u_j^{(A_\mu)} \otimes u_1^{(B_\nu)}) = \lambda_1^{(\mathcal{B}_\nu)}$, $\mathcal{B}_\nu(u_j^{(A_\mu)} \otimes u_2^{(B_\nu)}) = \lambda_2^{(\mathcal{B}_\nu)}$, $j = 1, 2$, on the sample space $S(\psi; \mathcal{A}_\mu, \mathcal{B}_\nu)$ with probability distribution $p_{\mathcal{B}_\nu}(\lambda_i^{(A)}) = |\langle u_1^{(A_\mu)} \otimes u_i^{(B_\nu)} | \psi \rangle|^2 + |\langle u_2^{(A_\mu)} \otimes u_i^{(B_\nu)} | \psi \rangle|^2$, $i = 1, 2$, and $p_{\mathcal{B}_\nu}(\lambda) = 0$ for $\lambda \notin \mathrm{Spec}(\mathcal{B}_\nu)$. Identifying the event $\{u_1^{(A_\mu)} \otimes u_i^{(B_\nu)}, u_2^{(A_\mu)} \otimes u_i^{(B_\nu)}\}$ with the "event" $\mathcal{B}_\nu = \lambda_i^{(\mathcal{B}_\nu)}$, we have $\mathrm{pr}(\mathcal{B}_\nu = \lambda_i^{(\mathcal{B}_\nu)}) = |\langle u_1^{(A_\mu)} \otimes u_i^{(B_\nu)} | \psi \rangle|^2 + |\langle u_2^{(A_\mu)} \otimes u_i^{(B_\nu)} | \psi \rangle|^2$, $i = 1, 2$.

In particular, let us set $\psi = \frac{1}{\sqrt{2}}(h_1 \otimes h_2 - h_2 \otimes h_1)$. We have

$$\mathrm{pr}(\mathcal{A}_\mu = \lambda_i^{(\mathcal{A}_\mu)}) =$$

$$\frac{1}{2}|\langle u_i^{(A_\mu)}|h_1\rangle\langle u_1^{(B_\nu)}|h_2\rangle - \langle u_i^{(A_\mu)}|h_2\rangle\langle u_1^{(B_\nu)}|h_1\rangle|^2 +$$

$$\frac{1}{2}|\langle u_i^{(A_\mu)}|h_1\rangle\langle u_2^{(B_\nu)}|h_2\rangle - \langle u_i^{(A_\mu)}|h_2\rangle\langle u_2^{(B_\nu)}|h_1\rangle|^2$$

and

$$\mathrm{pr}(\mathcal{B}_\nu = \lambda_j^{(\mathcal{B}_\nu)}) =$$

$$\frac{1}{2}|\langle u_1^{(A_\mu)}|h_1\rangle\langle u_j^{(B_\nu)}|h_2\rangle - \langle u_1^{(A_\mu)}|h_2\rangle\langle u_j^{(B_\nu)}|h_1\rangle|^2 +$$

$$\frac{1}{2}|\langle u_2^{(A_\mu)}|h_1\rangle\langle u_j^{(B_\nu)}|h_2\rangle - \langle u_2^{(A_\mu)}|h_2\rangle\langle u_j^{(B_\nu)}|h_1\rangle|^2.$$

Taking into account the form of the eigenvectors of the matrices A_μ and B_ν, we obtain

$$\mathrm{pr}(\mathcal{A}_\mu = \lambda_i^{(\mathcal{A}_\mu)}) = \mathrm{pr}(\mathcal{B}_\nu = \lambda_j^{(\mathcal{B}_\nu)}) = \frac{1}{2}, i, j = 1, 2.$$

We identify the intersection $(\mathcal{A}_\mu = \lambda_i^{(\mathcal{A}_\mu)}) \cap (\mathcal{B}_\nu = \lambda_j^{(\mathcal{B}_\nu)})$ with the event $\{u_i^{(A_\mu)} \otimes u_j^{(B_\nu)}\}$, $i, j = 1, 2$, in the sample space $S(\psi; \mathcal{A}_\mu, \mathcal{B}_\nu)$ and obtain

$$\mathrm{pr}((\mathcal{A}_\mu = \lambda_i^{(\mathcal{A}_\mu)}) \cap (\mathcal{B}_\nu = \lambda_j^{(\mathcal{B}_\nu)})) =$$

$$\frac{1}{2}|\langle u_i^{(A_\mu)}|h_1\rangle\langle u_j^{(B_\nu)}|h_2\rangle - \langle u_i^{(A_\mu)}|h_2\rangle\langle u_j^{(B_\nu)}|h_1\rangle|^2.$$

In particular, we have

$$\mathrm{pr}((\mathcal{A}_\mu = \lambda_1^{(\mathcal{A}_\mu)}) \cap (\mathcal{B}_\nu = \lambda_1^{(\mathcal{B}_\nu)})) = \frac{1}{2}\sin^2\left(\frac{\mu - \nu}{2}\right),$$

$$\mathrm{pr}((\mathcal{A}_\mu = \lambda_1^{(\mathcal{A}_\mu)}) \cap (\mathcal{B}_\nu = \lambda_2^{(\mathcal{B}_\nu)})) = \frac{1}{2}\cos^2\left(\frac{\mu - \nu}{2}\right),$$

$$\mathrm{pr}((\mathcal{A}_\mu = \lambda_2^{(\mathcal{A}_\mu)}) \cap (\mathcal{B}_\nu = \lambda_1^{(\mathcal{B}_\nu)})) = \frac{1}{2}\cos^2\left(\frac{\mu - \nu}{2}\right),$$

$$\mathrm{pr}((\mathcal{A}_\mu = \lambda_2^{(\mathcal{A}_\mu)}) \cap (\mathcal{B}_\nu = \lambda_2^{(\mathcal{B}_\nu)})) = \frac{1}{2}\sin^2\left(\frac{\mu - \nu}{2}\right).$$

The random variable $\mathcal{A}_\mu\mathcal{B}_\nu$ has probability distribution

$$p_{\mathcal{A}_\mu\mathcal{B}_\nu}(1) = \sin^2\left(\frac{\mu - \nu}{2}\right), p_{\mathcal{A}_\mu\mathcal{B}_\nu}(-1) = \cos^2\left(\frac{\mu - \nu}{2}\right),$$

and $p_{\mathcal{A}_\mu\mathcal{B}_\nu}(\lambda) = 0$ for $\lambda \neq \pm 1$. The expected value of this random variable is $\mathcal{E}(\mathcal{A}_\mu\mathcal{B}_\nu) = -\cos(\mu - \nu)$.

3. Entropy and Degree of Dependence

3.1. Entropy

Now, we combine the terminology and notation of this paper with those of [6]. Let us set $A = (\mathcal{A}_\mu = \lambda_1^{(\mathcal{A}_\mu)})$, $B = (\mathcal{B}_\nu = \lambda_1^{(\mathcal{B}_\nu)})$, $A^c = (\mathcal{A}_\mu = \lambda_2^{(\mathcal{A}_\mu)})$, $B^c = (\mathcal{B}_\nu = \lambda_2^{(\mathcal{B}_\nu)})$. $\alpha = \mathrm{pr}(A) = \frac{1}{2}$, $\beta = \mathrm{pr}(B) = \frac{1}{2}$.

The pair (A, B) of events in the sample space $S(\psi; \mathcal{A}, \mathcal{B})$ produces an experiment

$$\mathfrak{J} = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B) \cup (A^c \cap B^c) \tag{1}$$

(cf. [3, I,§5]) and the probabilities of its results:

$$\xi_1 = \mathrm{pr}(A \cap B) = \frac{1}{2}\sin^2\left(\frac{\mu-\nu}{2}\right),\ \xi_2 = \mathrm{pr}(A \cap B^c) = \frac{1}{2}\cos^2\left(\frac{\mu-\nu}{2}\right),$$

$$\xi_3 = \mathrm{pr}(A^c \cap B) = \frac{1}{2}\cos^2\left(\frac{\mu-\nu}{2}\right),\ \xi_4 = \mathrm{pr}(A^c \cap B^c) = \frac{1}{2}\sin^2\left(\frac{\mu-\nu}{2}\right). \tag{2}$$

The probability distribution $(\xi_1, \xi_2, \xi_3, \xi_4)$ satisfies the linear system [6, 4.1. (3)] whose solutions form a straight line with parametric representation $\xi_1 = \theta$, $\xi_2 = \frac{1}{2} - \theta$, $\xi_3 = \frac{1}{2} - \theta$, $\xi_4 = \theta$ in the hyperplane $\xi_1 + \xi_2 + \xi_3 + \xi_4 = 1$. Note that the parameter $\theta = \xi_1$ runs within the closed interval $[0, \frac{1}{2}]$. The entropy of $(\xi_1, \xi_2, \xi_3, \xi_4)$ is $E(\theta) = -\sum_{k=1}^{4} \xi_k(\theta)\ln(\xi_k(\theta)) = -2\theta\ln\theta - 2(\frac{1}{2} - \theta)\ln(\frac{1}{2} - \theta)$ and the function $E(\theta)$ can be extended as continuous on the interval $[0, \frac{1}{2}]$. It strictly increases on the interval $[0, \frac{1}{4}]$, strictly decreases on the interval $[\frac{1}{4}, \frac{1}{2}]$ and has a global maximum at $\theta = \frac{1}{4}$. In particular, $\max_{\theta\in[0,\frac{1}{2}]} E(\theta) = E(\frac{1}{4}) = 2\ln 2$. Since $\min_{\theta\in[0,\frac{1}{4}]} E(\theta) = E(0) = \ln 2 = E(\frac{1}{2}) = \min_{\theta\in[\frac{1}{4},\frac{1}{2}]} E(\theta)$, we obtain $\min_{\theta\in[0,\frac{1}{2}]} E(\theta) = \ln 2$.

3.2. Degree of Dependence

It is more useful to modify the entropy function, thus obtaining the strictly increasing *degree of dependence function* $e\colon [0, \frac{1}{2}] \to [-1, 1]$,

$$e(\theta) = \begin{cases} -\frac{E(\frac{1}{4})-E(\theta)}{E(\frac{1}{4})-E(0)} & \text{if } 0 \le \theta \le \frac{1}{4} \\ \frac{E(\frac{1}{4})-E(\theta)}{E(\frac{1}{4})-E(\frac{1}{2})} & \text{if } \frac{1}{4} \le \theta \le \frac{1}{2}. \end{cases}$$

Taking into account the values of extrema of entropy function, we obtain

$$e(\theta) = \begin{cases} -2 + \frac{E(\theta)}{\ln 2} & \text{if } 0 \le \theta \le \frac{1}{4} \\ 2 - \frac{E(\theta)}{\ln 2} & \text{if } \frac{1}{4} \le \theta \le \frac{1}{2}. \end{cases}$$

The events A and B are *independent* exactly when the entropy is maximal (equal to $2\ln 2$), that is, when $e(\theta) = 0$ and this, in turn, is equivalent to the equality $|\mu - \nu| = \frac{\pi}{2}$. We have $e(\theta) = -1$ or $e(\theta) = 1$ if and only if $|\mu - \nu| = 0$ or $|\mu - \nu| = \pi$, respectively, and in these two cases the entropy is minimal and equal to $\ln 2$. Now, let, in addition, assume that A and B are events in a sample space with equally likely outcomes. If $e(\theta) = -1$, then one of A and B is a subset of the complement of the other (maximal *negative dependence*), and if $e(\theta) = 1$ one of them is a subset of the other (maximal *positive dependence*).

3.3. The Information Flow

The experiment $\mathfrak{J}$ from (1) is the joint experiment (see [7, Part I, Section 6]) of two simple binary trials: $\mathfrak{A}_\mu = A \cup A^c$ and $\mathfrak{B}_\nu = B \cup B^c$ with $\mathrm{pr}(A) = \mathrm{pr}(B) = \frac{1}{2}$. The *average quantity of information of one of the experiments $\mathfrak{A}_\mu$ and $\mathfrak{B}_\nu$, relative to the other*, (see [3, §1]), is defined in this particular case

by the formula $I(\mathfrak{A}_\mu, \mathfrak{B}_\nu)(\theta) = \xi_1(\theta)\ln 4\xi_1(\theta) + \xi_2(\theta)\ln 4\xi_2(\theta) + \xi_3(\theta)\ln 4\xi_3(\theta) + \xi_4(\theta)\ln 4\xi_4(\theta)$. The above notation is correct since the interchanges of A and A^c or B and B^c causes permutations of ξ_i's. Thus, we obtain $I(\mathfrak{A}_\mu, \mathfrak{B}_\nu)(\theta) = \max_{\theta\in[0,\frac{1}{2}]} E(\theta) - E(\theta)$. Now, the definition of the degree function $e(\theta)$ yields immediately $I(\mathfrak{A}_\mu, \mathfrak{B}_\nu)(\theta) = |e(\theta)|\ln 2$ for $\theta \in [0, \frac{1}{2}]$.

Translating into the language of information theory, we have $e(\theta) = -1$ or $e(\theta) = 1$ if and only if $I(\mathfrak{A}_\mu, \mathfrak{B}_\nu)(\theta) = \max_{0\le\tau\le\frac{1}{2}} I(\mathfrak{A}_\mu, \mathfrak{B}_\nu)(\tau) = \ln 2$. Finally, we have $e(\theta) = 0$ if and only if $I(\mathfrak{A}_\mu, \mathfrak{B}_\nu)(\theta) = 0$, and under this condition the experiments $\mathfrak{A}_\mu$ and $\mathfrak{B}_\nu$ are said to be *informationally independent*.

3.4. The Signed Information Flow

Let us set $I^{(s)}(\mathcal{A}_\mu, \mathcal{B}_\nu)(\theta) = e(\theta)\ln 2$ for $\theta \in [0, \frac{1}{2}]$ and call this quantity *average quantity of signed information of one of the events* $\mathcal{A}_\mu = \lambda_1^{(\mathcal{A}_\mu)}$ *and* $\mathcal{B}_\nu = \lambda_1^{(\mathcal{B}_\nu)}$, *relative to the other*. Then $I(\mathfrak{A}_\mu, \mathfrak{B}_\nu) = |I^{(s)}(\mathcal{A}_\mu, \mathcal{B}_\nu)|$ and since the function e is invertible, we obtain $\theta = e^{-1}(\frac{1}{\ln 2} I^{(s)}(\mathcal{A}_\mu, \mathcal{B}_\nu))$. In particular, the value of the signed information flow $I^{(s)}(\mathcal{A}_\mu, \mathcal{B}_\nu)$ reproduces the probability distribution (2) predicted by quantum theory.

4. Four Operators and Bell's Map

For any $\mu_1, \mu_2, \nu_1, \nu_2 \in [0, \pi]$ we consider the self-adjoined operators A_{μ_i}, B_{ν_j}, $i, j = 1, 2$, see Subsection 2.1. We extend notation introduced in Sections 2 and 3 in a natural way: $\theta_{ij} = \frac{1}{2}\sin^2\left(\frac{\mu_i - \nu_j}{2}\right)$, $\theta_{ij} \in [0, \frac{1}{2}]$, $\mathfrak{A}_{\mu_i}$, $\mathfrak{B}_{\nu_j}$, $I(\mathfrak{A}_{\mu_i}, \mathfrak{B}_{\nu_j}) = |e(\theta_{ij})|\ln 2$, $i, j = 1, 2$. The sum $I(\mathfrak{A}, \mathfrak{B}) = \sum_{i,j=1}^{2} I(\mathfrak{A}_{\mu_i}, \mathfrak{B}_{\nu_j})$ is said to be the *average quantity of information of one of the pairs of experiments* $\mathfrak{A} = \{\mathfrak{A}_{\mu_1}, \mathfrak{A}_{\mu_2}\}$ *and* $\mathfrak{B} = \{\mathfrak{B}_{\nu_1}, \mathfrak{B}_{\nu_2}\}$ *relative to the other*, or, *total information flow*. The sum $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = \sum_{i,j=1}^{2} I^{(s)}(\mathcal{A}_\mu, \mathcal{B}_\nu)$ is said to be the *average quantity of signed information of one of the pairs of experiments* $\mathfrak{A} = \{\mathfrak{A}_{\mu_1}, \mathfrak{A}_{\mu_2}\}$ *and* $\mathfrak{B} = \{\mathfrak{B}_{\nu_1}, \mathfrak{B}_{\nu_2}\}$ *relative to the other*, or, *total signed information flow*.

Thus, we obtain the functions

$$I(\mathfrak{A}, \mathfrak{B})\colon [0, \pi]^4 \to \mathbb{R}, (\mu_1, \mu_2, \nu_1, \nu_2) \mapsto (\ln 2) \sum_{i,j=1}^{2} |e(\theta_{ij})|,$$

and

$$I^{(s)}(\mathfrak{A}, \mathfrak{B})\colon [0, \pi]^4 \to \mathbb{R}, (\mu_1, \mu_2, \nu_1, \nu_2) \mapsto (\ln 2) \sum_{i,j=1}^{2} e(\theta_{ij}),$$

which represents the intensity of information flow (respectively, signed information flow) between the pairs of experiments $\mathfrak{A}$ and $\mathfrak{B}$. We note that $0 \le I(\mathfrak{A}, \mathfrak{B})(\mu_1, \mu_2, \nu_1, \nu_2) \le 4\ln 2$ and $-4\ln 2 \le I^{(s)}(\mathfrak{A}, \mathfrak{B})(\mu_1, \mu_2, \nu_1, \nu_2) \le 4\ln 2$.

In case $\mu_1 = \mu_2 = \nu_1 = \nu_2$ we have $\theta_{11} = \theta_{12} = \theta_{21} = \theta_{22} = 0$, $e(\theta_{11}) = e(\theta_{12}) = e(\theta_{21}) = e(\theta_{22}) = -1$, hence $I(\mathfrak{A}, \mathfrak{B}) = 4\ln 2$ and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = -4\ln 2$. In case $\mu_1 = \mu_2 = \frac{\pi}{2}$, $\nu_1 = \nu_2 = 0$ we have $\theta_{11} = \theta_{12} = \theta_{21} = \theta_{22} = \frac{1}{4}$, $e(\theta_{11}) = e(\theta_{12}) = e(\theta_{21}) = e(\theta_{22}) = 0$, and $I(\mathfrak{A}, \mathfrak{B}) = 0$. Finally, in case $\mu_1 = \mu_2 = \pi$, $\nu_1 = \nu_2 = 0$ we have $\theta_{11} = \theta_{12} = \theta_{21} = \theta_{22} = \frac{1}{2}$, $e(\theta_{11}) = e(\theta_{12}) = e(\theta_{21}) = e(\theta_{22}) = 1$, and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = 4\ln 2$.

Since the image of a compact and connected set via continuous function $I(\mathfrak{A}, \mathfrak{B})$ (respectively, the continuous function $I^{(s)}(\mathfrak{A}, \mathfrak{B})$) is a compact and connected subset of $\mathbb{R}$, we obtain that the range of $I(\mathfrak{A}, \mathfrak{B})$ (respectively, $I^{(s)}(\mathfrak{A}, \mathfrak{B})$) coincides with the interval $[0, 4\ln 2]$ (respectively, with the interval $[-4\ln 2, 4\ln 2]$).

4.1. Bell's Inequality

The equality $|\mathcal{A}_{\mu_1}\mathcal{B}_{\nu_1} + \mathcal{A}_{\mu_1}\mathcal{B}_{\nu_2} + \mathcal{A}_{\mu_2}\mathcal{B}_{\nu_1} - \mathcal{A}_{\mu_2}\mathcal{B}_{\nu_2}| = 2$ yields (with an abuse of the probability theory) Bell's inequality

$$|\mathcal{E}(\mathcal{A}_{\mu_1}\mathcal{B}_{\nu_1}) + \mathcal{E}(\mathcal{A}_{\mu_1}\mathcal{B}_{\nu_2}) + \mathcal{E}(\mathcal{A}_{\mu_2}\mathcal{B}_{\nu_1}) - \mathcal{E}(\mathcal{A}_{\mu_2}\mathcal{B}_{\nu_2}| \leq 2,$$

that is, $|b(\mu_1, \mu_2, \nu_1, \nu_2)| \leq 2$, where $b(\mu_1, \mu_2, \nu_1, \nu_2) = \cos(\mu_1 - \nu_1) + \cos(\mu_1 - \nu_2) + \cos(\mu_2 - \nu_1) - \cos(\mu_2 - \nu_2)$.

J. S. Bell in [2] proves that if there exist "...additional variables which restore to the (quantum) theory causality and locality", then the above inequality is satisfied. Since $I(\mathfrak{A}, \mathfrak{B}) = 0$ is equivalent to the equalities $|\mu_i - \nu_j| = \frac{\pi}{2}$, $i, j = 1, 2$, this yields $b = 0$. Thus, we obtain that if Bell's inequality is violated, then the total information flow $I(\mathfrak{A}, \mathfrak{B})$ is strictly positive, that is, the experiments $\mathfrak{A}$ and $\mathfrak{B}$ are informationally dependent.

Examples 1. Note that the results of all calculations below are rounded up to the 7-th digit.

1) (Aspect's experiment) $\mu_1 = \frac{\pi}{8}, \mu_2 = \frac{3\pi}{8}, \nu_1 = \frac{\pi}{4}, \nu_2 = 0$. Then we obtain $\cos(\frac{\pi}{8}) = 0.9238795$, $\cos(\frac{3\pi}{8}) = 0.3826834$, and therefore $b(\frac{\pi}{8}, \frac{3\pi}{8}, \frac{\pi}{4}, 0) = 2.3889551$. On the other hand, $\theta_{11} = \theta_{12} = \theta_{21} = \frac{1}{2}\sin^2(\frac{\pi}{16}) = 0.0190301$, $e(\theta_{11}) = e(\theta_{12}) = e(\theta_{21}) = -0.0415353$, $\theta_{22} = \frac{1}{2}\sin^2(\frac{3\pi}{16}) = 0.154329$, $e(\theta_{22}) = -0.1084492$. Hence we have $I(\mathfrak{A}, \mathfrak{B}) = 0.1615415$ and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = -0.2330551$.

2) $\mu_1 = \pi, \mu_2 = \frac{2\pi}{3}, \nu_1 = 0, \nu_2 = \frac{\pi}{3}$. Then we have $b(\pi, \frac{2\pi}{3}, 0, \frac{\pi}{3}) = -2.5$. On the other hand, $\theta_{11} = \frac{1}{2}\sin^2(\frac{\pi}{2}) = 0.5$, $e(\theta_{11}) = 1$, $\theta_{12} = \theta_{21} = \frac{1}{2}\sin^2(\frac{\pi}{3}) = \frac{3}{8} = 0.375$, $e(\theta_{12}) = e(\theta_{21}) = 0.1887219$, $\theta_{22} = \frac{1}{2}\sin^2(\frac{\pi}{6}) = \frac{1}{8} = 0.125$, $e(\theta_{22}) = -0.1887219$. Hence we obtain $I(\mathfrak{A}, \mathfrak{B}) = 1.0855833$ and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = 1.1887219$

3) $\mu_1 = \frac{\pi}{2}, \mu_2 = 0$, $\nu_1 = \frac{\pi}{4}$, $\nu_2 = \frac{3\pi}{4}$. Then $b(\frac{\pi}{2}, 0, \frac{\pi}{4}, \frac{3\pi}{4}) = 2\sqrt{2}$. On the other hand, $\theta_{11} = \theta_{12} = \theta_{21} = \frac{1}{2}\sin^2(\frac{\pi}{8}) = 0.0732233$, $e(\theta_{11}) = e(\theta_{12}) = e(\theta_{21}) = -0.3994425$, $\theta_{22} = \frac{1}{2}\sin^2(\frac{3\pi}{16}) = 0.154329$, $e(\theta_{22}) = -0.10844492$. Hence we have $I(\mathfrak{A}, \mathfrak{B}) = 0.9053727$ and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = -0.22827767$.

4) $\mu_1 = \pi, \mu_2 = 0$, $\nu_1 = 0$, $\nu_2 = \pi$. Then we have $b(\pi, 0, 0, -\pi) = 2$. On the other hand, $\theta_{11} = \theta_{22} = \frac{1}{2}\sin^2(\frac{\pi}{2}) = 0.5$, $e(\theta_{11}) = e(\theta_{22}) = 1$, $\theta_{12} = \theta_{21} = \frac{1}{2}\sin^2(0) = 0$, $e(\theta_{12}) = e(\theta_{21}) = -1$. Therefore we obtain $I(\mathfrak{A}, \mathfrak{B}) = 4\ln 2 = 2.7725887 = \max I(\mathfrak{A}, \mathfrak{B})$ and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = 0$.

5) $\mu_1 = \frac{5\pi}{6}, \mu_2 = \frac{2\pi}{3}$, $\nu_1 = \frac{\pi}{3}$, $\nu_2 = \frac{\pi}{2}$. In this case we have $b(\frac{5\pi}{6}, \frac{2\pi}{3}, \frac{\pi}{3}, \frac{\pi}{2}) = 1 - \frac{\sqrt{3}}{2}$. On the other hand, $\theta_{11} = \frac{1}{2}\sin^2(\frac{\pi}{4}) = \frac{1}{4}$, $e(\theta_{11}) = 0$, $\theta_{12} = \theta_{21} = \frac{1}{2}\sin^2(\frac{\pi}{6}) = \frac{1}{8}$, $e(\theta_{12}) = e(\theta_{21}) = -0.1887219$, $\theta_{22} = \frac{1}{2}\sin^2(\frac{\pi}{12}) = 0.0334936$, $e(\theta_{22}) = -0.6453728$. Hence we have $I(\mathfrak{A}, \mathfrak{B}) = 0.7089624$ and $I^{(s)}(\mathfrak{A}, \mathfrak{B}) = -0.267929$.

6) The link

`http://www.math.bas.bg/algebra/valentiniliev/`

contains a Java experimental implementation "dependencemeasurements2" depending on five parameters: an non-negative integer n and four real numbers $\mu_1, \mu_2, \nu_1, \nu_2$ from the closed interval $[0, \pi]$. One can also find the description of this program at the above link.

Examples 1 and, especially, example 6), yield that the relations between $I(\mathfrak{A}, \mathfrak{B})$ and b, and $I^{(s)}(\mathfrak{A}, \mathfrak{B})$ and b seem to be random. Below we present an attempt to explain the uncertainty of this relation by refereing to [6, 5.4, 5.5]. We define the events

$$U = \{(\mu_1, \mu_2, \nu_1, \nu_2) \in [0, \pi]^4 || b(\mu_1, \mu_2, \nu_1, \nu_2)| \leq 2\},$$

$$V = \{(\mu_1, \mu_2, \nu_1, \nu_2) \in [0, \pi]^4 | I(\mathfrak{A}, \mathfrak{B})(\mu_1, \mu_2, \nu_1, \nu_2) \in [0, 2\ln 2]\},$$

$$VS = \{(\mu_1, \mu_2, \nu_1, \nu_2) \in [0, \pi]^4 | I^{(s)}(\mathfrak{A}, \mathfrak{B})(\mu_1, \mu_2, \nu_1, \nu_2) \in [0, 4\ln 2]\},$$

with complements U^c, V^c, and VS^c in $[0, \pi]^4$. We suppose that the probabilities $\alpha = \mathrm{pr}(U)$, $\beta = \mathrm{pr}(V)$, and $\beta^{(s)} = \mathrm{pr}(VS)$ in the sample space $[0, \pi]^4$ furnished with normalized Borel measure are known. The probabilities $\tau = \mathrm{pr}(U \cap V)$ and $\tau^{(s)} = \mathrm{pr}(U \cap VS)$ run through the closed intervals $I(\alpha, \beta) = [\max(0, \alpha + \beta - 1), \min(\alpha, \beta)]$ and $I(\alpha, \beta^{(s)})$, respectively. In case $\tau = \min(\alpha, \beta)]$ (respectively, $\tau^{(s)} = \min(\alpha, \beta^{(s)})]$) there exists a relation of inclusion (up to

a set of probability 0) between U and V (respectively, between U and VS). Otherwise, both conditional probabilities $\mathrm{pr}(V^c|U)$ and $\mathrm{pr}(V|U^c)$ (respectively, $\mathrm{pr}(VS^c|U)$ and $\mathrm{pr}(VS|U^c)$) can not be simultaneously as small as one wants (a kind of uncertainty principle).

Remark 1. The probabilities $\alpha = \mathrm{pr}(U)$, $\beta = \mathrm{pr}(V)$, $\tau = \mathrm{pr}(U \cap V)$, and $\tau^{(s)} = \mathrm{pr}(U \cap VS)$ can be approximated by using Examples 1, 6). We draw a random sample X of size n from the sample space $[0, \pi]^4$ and consider X as a sample space with n equally likely outcomes. Then the probabilities $\hat{\alpha}(n)$, $\hat{\beta}(n)$, $\hat{\tau}(n)$, and $\widehat{\tau^{(s)}}(n)$ of the traces of U,V, $U \cap V$, and $U \cap VS$ on X (the sample proportions) are unbiased estimators for α, β, τ, and $\tau^{(s)}$ when n is large.

Note that as an output of n iterations of the random process from example 6) we can also find: a) the approximation $[L(n), J(n)]$ of the range of $I(\mathfrak{A}, \mathfrak{B})$ and the approximation $[LS(n), JS(n)]$ of the range of $I^{(s)}(\mathfrak{A}, \mathfrak{B})$ under the condition $|b| \leq 2$, b) the above sample proportions, and c) the approximations $\mathrm{pr}(V^c|U)(n) = \frac{\hat{\alpha}(n)-\hat{\tau}(n)}{\hat{\alpha}(n)}$, $\mathrm{pr}(V|U^c)(n) = \frac{\hat{\beta}(n)-\hat{\tau}(n)}{1-\hat{\alpha}(n)}$, and $\mathrm{pr}(VS^c|U)(n) = \frac{\hat{\alpha}(n)-\widehat{\tau^{(s)}}(n)}{\hat{\alpha}(n)}$, $\mathrm{pr}(VS|U^c)(n) = \frac{\widehat{\beta^{(s)}}(n)-\widehat{\tau^{(s)}}(n)}{1-\hat{\alpha}(n)}$.

Below are the results obtained by drawing a random sample of size $n = 1000$:

$\hat{\alpha}(1000) = 0.838$, $\hat{\beta}(1000) = 0.704$, $\hat{\tau}(1000) = 0.614$, $\widehat{\tau^{(s)}}(1000) = 0.087$, $\mathrm{pr}(V^c|U)(1000) = 0.2673031$, $\mathrm{pr}(V|U^c)(1000) = 0.5555555$, $\mathrm{pr}(VS^c|U)(1000) = 0.8961814$, $\mathrm{pr}(VS|U^c)(1000) = 0.0185185$.

Acknowledgements

I thank Dimitar Guelev for making an experimental Java implementation of the evaluation of dependence of simultaneous measurements and the approximation of various parameters via a random process. The numerical examples thus produced were invaluable for my work. I would like to express my gratitude to administration of the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences for creating perfect and safe conditions of work.

References

[1] Aspect A., Dalibard J., Roger G. (1982). Experimental Test of Bell's Inequalities Using Time-Varying Analysers. *Physical Review Letters*, 49 No 25:1804–1807.

[2] Bell J. (1964). On the Einstein Podolski Rosen Paradox. *Physics*, 1: 195–200.

[3] Gelfand I. M., Kolmogorov A. N., Yaglom A. M. Amount of Information and Entropy for Continuous Distributions. Mathematics and Its Applications, Selected Works of A. N. Kolmogorov, III: Information Theory and the Theory of Algorithms, 33–56, Springer Science+Business Media Dordrecht, 1993.

[4] Kolmogorov A. N. Foundations of the Theory of Probability, Chelsea Publishing Company New Yourk, 1956.

[5] A. Einstein, B. Podolsky, N. Rosen. (1935). Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47:777–780.

[6] V. V. Iliev. (2021). On the Use of Entropy as a Measure of Dependence of Two Events. *Reliability: Theory & Applications*, 16 No 3:237–248.

[7] C. E. Shannon. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27 No 3:379–423, 27 No 4:523–656.

MAXWELL-GOMPERTZ DISTRIBUTION: PROPERTIES AND APPLICATIONS

Alfred Adewole Abiodun[1*], Aliyu Ismail Ishaq[2], Olakiitan Ibukun Adeniyi[3], Ifeanyi Vivian Omekam[4], Jumoke Popoola[5], Olubimpe Mercy Oladuti[6] and Eunice Ohunene Job[7]

•

[1,3,4,5]Department of Statistics, University of Ilorin, Ilorin, Nigeria
[2]Department of Statistics, Ahmadu Bello University, Zaria, Nigeria
[6]Department of Statistics, Federal University of Technology, Akure, Nigeria
[7]Department of Statistics, Federal University Lokoja, Nigeria
abbay@unilorin.edu.ng[1], binishaq05@gmail.com[2], adeniyi.oi@unilorin.edu.ng[3], omekam.iv@unilorin.edu.ng[4], 1jmkbalpop@unilorin.edu.ng[5], omoladuti@futa.edu.ng[6], eunice.upahi@fulokoja.edu.ng[7]

Abstract

This paper proposed a three parameter Maxwell-Gompertz distribution as an extension of Gompertz distribution. Some statistical properties of the distribution such as moments, survival and hazard functions, quantile function, Rényi entropy and order statistics were derived. Maximum likelihood method was used to estimate the model parameters. A simulation study was carried out in order to gain an insight into the performance on small, moderate and large samples. The flexibility of the new distribution was empirically demonstrated in comparison to four other extensions of Gompertz distributions using two real life datasets.

Keywords: Maxwell-Gompertz, generator, skewness, Rényi entropy, maximum likelihood

I. Introduction

Gompertz distribution is a popular distribution commonly used in many applied problems, particularly in modelling lifetime data [1]. The distribution is often characterized by an increasing hazard function and it is commonly used to describe the distribution of adult life spans by actuaries and demographers [2]. It is also considered for modelling survival data in some sciences such as gerontology [3], computer science [4], biology [5], and marketing science [6]. For more details about the Gompertz distribution and its applications, see [7], [8]. The cumulative distribution function (cdf) and probability density function (pdf) of the Gompertz random variable X are respectively given as

$$T(x,\vartheta)=1-e^{-\frac{c}{b}(e^{bx}-1)}, \tag{1}$$

and

$$t(x,\vartheta)=ce^{bx}e^{-\frac{c}{b}(e^{bx}-1)}, \qquad \vartheta>0; x>0 \tag{2}$$

where $\vartheta=(b, c)$ with b denting the shape parameter and c the scale parameter.

The development of new families of distributions has become an important trend in the theory and application of distributions. Such new families of distributions are often compounded by adding one or more parameters to the well-known standard baseline distributions. This has become necessary because the resulting extended new distributions provide greater flexibility in modelling observed data. A few of such families of distributions which have been explored in the recent times include, among others the Beta-G of [9], a new generalized odd log-logistic family of distributions by [10], The generalized odd half-Cauchy family of distributions by [11], a New Kumaraswamy generalized family of distributions by [12].

Several other families of distributions can be mentioned such as Odd F family of distributions by [13], Odd Beta Prime family of distributions by [14], Generalized odd Maxwell family of distributions by [15], Generalized beta-generated distributions of [16], Garhy-generated family of distributions by [17], Gamma-G Type-3 of [18], The Logistic-X family of [19], a new Weibull-X family of [20], a-Zubair-G family of [21], and a new Alpha power transformed family distribution by [22].

Gompertz distribution has been extended by some authors in the literature through the addition of one or more other parameters. Some of such studies in the recent time include the modified beta Gompertz distribution by [23], the generalized Gompertz distribution by [24] which was based on an idea of [25], the cubic transmuted Gompertz distribution by [26], the odd generalized exponential-Gompertz distribution by [27], the transmuted Gompertz distribution by [28] and the odd lindley-Gompertz distribution by [29]. This article seeks to develop a distribution that has the characterization of the Gompertz and Maxwell distributions in a unified framework. The Maxwell distribution was introduced by [30], and it has the cdf given as

$$G(u,a)=\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{u^2}{2a^2}\right), \tag{3}$$

with $\gamma(u,b)=\int_0^u s^{b-1}e^{-s}ds$ denoting the lower incomplete gamma function. The associated pdf of (3) is

$$g(u,a)=\sqrt{\frac{2}{\pi}}\frac{u^2e^{-\frac{u^2}{2a^2}}}{a^3}, \qquad u,a>0 \tag{4}$$

where a is the scale parameter.

Studies involving Maxwell generalized family of distributions have not been widely covered in the literature. However, [31] proposed Maxwell–Weibull distribution by applying the odd ratio link approach of [32]. Also, [33] developed Maxwell-Dagum distribution while [34] developed Maxwell-Lomax distribution.

II. Methods

2.1.The Maxwell-Gompertz (Mgom) Distribution

Consider a random variable X which follows the Gompertz distribution with the cumulative and probability density functions as defined in (1) and (2) respectively. Following [31] who proposed Maxwell family of distributions for continuous generator, we can present the cumulative density function of Maxwell-G family as

$$F(x;a,\vartheta)=\int_0^{\frac{T(x,\vartheta)}{1-T(x,\vartheta)}}\sqrt{\frac{2}{\pi}}\frac{u^2e^{-\frac{u^2}{2a^2}}}{a^3}=\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{1}{2a^2}\left(\frac{T(x,\vartheta)}{1-T(x,\vartheta)}\right)^2\right), \quad a>0;x>0 \tag{5}$$

and the corresponding pdf is given as

$$f(;a,\vartheta)=\frac{2t(x,\vartheta)}{a^3\sqrt{2\pi}(1-T(x,\vartheta))^2}\left(\frac{T(x,\vartheta)}{1-T(x,\vartheta}\right)^2\exp\left(-\frac{1}{2a^2}\left(\frac{T(x,\vartheta)}{1-T(x,\vartheta}\right)^2\right), \tag{6}$$

where $\vartheta=(b,c)$ denotes the vector of parameters of the baseline Gompertz distribution.

Substituting (1) in (5) gives the proposed cdf of the Maxwell-Gompertz (MGom) distributions as

$$F(x;a,b,c)=\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{1}{2a^2}\left(\frac{1-e^{-\frac{c}{b}(e^{bx}-1)}}{e^{-\frac{c}{b}(e^{bx}-1)}}\right)^2\right),\quad a,b,c>;\ x>0 \tag{7}$$

and on substituting (1) and (2) in (6), the pdf can be obtained as

$$f(x;a,b,c)=\frac{2ce^{bx}\left(1-e^{-\frac{c}{b}(e^{bx}-1)}\right)^2}{a^3\sqrt{2\pi}\left(e^{-\frac{c}{b}(e^{bx}-1)}\right)^3}\exp\left(-\frac{1}{2a^2}\left(\frac{1-e^{-\frac{c}{b}(e^{bx}-1)}}{e^{-\frac{c}{b}(e^{bx}-1)}}\right)^2\right). \tag{8}$$

2.2 Linear Representation of MGom Density

Consider the power series expansion of the exponential function

$$e^{-x}=\sum_{i=0}^{\infty}\frac{(-1)^i}{i!}x^i \tag{9}$$

Putting (9) in (8) and dropping (a,b,c) in $f(a,b,c)$ for simplicity, we have

$$f(x)=\frac{2ce^{bx}\,e^{-\frac{c}{b}(e^{bx}-1)}}{a^3\sqrt{2\pi}}\sum_{i=0}^{\infty}\frac{(-1)^i}{i!\left(2a^2\right)^i}\frac{\left(1-e^{-\frac{c}{b}(e^{bx}-1)}\right)^{2+2i}}{\left(e^{-\frac{c}{b}(e^{bx}-1)}\right)^{4+2i}}. \tag{10}$$

Considering the generalized binomial expansion in power of positive real number μ, expressed as

$$(1-x)^{\mu}=\sum_{k=0}^{\infty}\frac{(-1)^k\Gamma(u+1)}{k!\Gamma(u+1-k)}x^k. \tag{11}$$

By applying (11) to (10) we obtain

$$f(x)=\sum_{i,j,k=0}^{\infty}\frac{2c(-1)^{i+k}\Gamma(4+2i+j)\Gamma(3+2i+j)}{i!\,j!k!\left(2a^2\right)^i a^3\sqrt{2\pi}\Gamma(4+2i)\Gamma(3+2i+j-k)}e^{bx}e^{-\frac{c}{b}(1+k)(e^{bx}-1)}. \tag{12}$$

Thus, the pdf of the MGom distribution expressed as a linear representation is obtained by applying (9) to (12) which gives

$$f(x)=\Omega_{i,j,k}e^{bx}\sum_{l=0}^{\infty}\frac{(-1)^l}{l!}\left(\frac{c}{b}(1+k)\right)^l\left(e^{bx}-1\right)^l \tag{13}$$

$$=\sum_{l,m=0}^{\infty}\Phi_{l,m}e^{-bx(m-l-1)}, \tag{14}$$

where $\Phi_{l,m}=\frac{(-1)^{l+m}\Gamma(l+1)}{l!m!\Gamma(l+1-m)}\left(\frac{c}{b}(1+k)\right)^l\Omega_{i,j,k}$

and $\Omega_{i,j,k}=\sum_{i,j,k=0}^{\infty}\frac{2c(-1)^{i+k}\Gamma(4+2i+j)\Gamma(3+2i+j)}{i!\,j!k!\left(2a^2\right)^i a^3\sqrt{2\pi}\Gamma(4+2i)\Gamma(3+2i+j-k)}$.

The plots of the pdf and cdf of MGom distribution using different parameter values are displayed in Figures 1 and 2 respectively. From Figure 1, it is observed that the pdf of the MGom distribution is skewed to the right and therefore will be a good model for different kinds of positively skewed data sets.

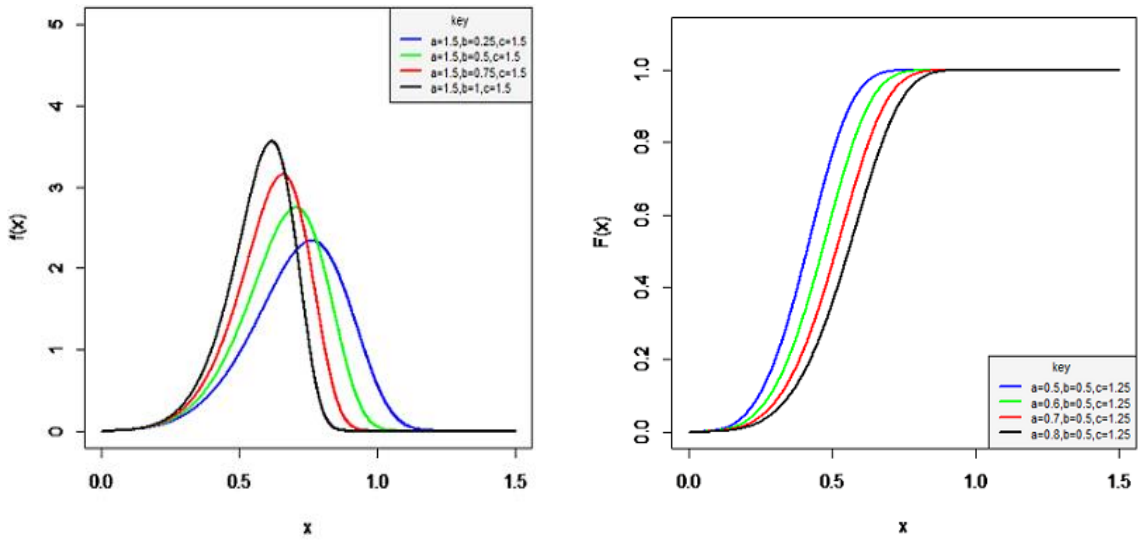


Figure 1: *Plots of pdf and cdf of EGILx distribution*

2.3 Statistical Properties

Some structural properties of the Maxwell-Gompertz distribution are discussed in this section.

2.3.1 Moments

Suppose that X denote a continuous random variable, the rth non-central moment of X is given by

$$E(X^r)=\int_{-\infty}^{\infty} x^r f(x)dx\,. \tag{15}$$

Taking f(x) as the pdf of the MGom distribution given in (14), the rth moments of X is given as

$$E(X^r)=\sum_{l,m=0}^{\infty}\Phi_{l,m}\int_0^{\infty}x^r e^{-bx(m-l-1)}dx\,. \tag{16}$$

Let,

$$y=bx(m-l-1),\Rightarrow x=\frac{y}{b(m-l-1)},\text{ so that } dx=\frac{dy}{b(m-l-1)}\,. \tag{17}$$

By inserting (17) into (16), we obtain

$$E(X^r)=\sum_{l,m=0}^{\infty}\Phi_{l,m}\int_0^{\infty}\left(\frac{y}{b(m-l-1)}\right)^r e^{-y}\frac{dy}{b(m-l-1)} \tag{18}$$

$$=\frac{\sum_{l,m=0}^{\infty}\Phi_{l,m}}{\left(b(m-l-1)\right)^{r+1}}\Gamma(r+1)\,. \tag{19}$$

which is the moments of MGom distribution.

2.3.2. Quantile function

Quantile function of MGom can be derived by inverting the cdf given in (7).
If we let

$$F(x)=\frac{2}{\sqrt{2\pi}}\gamma\left(\frac{3}{2},\frac{1}{2a^2}\left(\frac{1-e^{-\frac{c}{b}\left(e^{bx}-1\right)}}{e^{-\frac{c}{b}\left(e^{bx}-1\right)}}\right)^2\right)=\mathrm{u}\,, \tag{20}$$

then by solving (20) for x we obtain

$$x_q=Q(u)=b^{-1}\log\left(1-\frac{b}{c}\log\left(1-\frac{\left(2a^2\gamma^{-1}\left(\frac{3}{2},u\Gamma\left(\frac{3}{2}\right)\right)\right)^{0.5}}{1+\left(2a^2\gamma^{-1}\left(\frac{3}{2},u\Gamma\left(\frac{3}{2}\right)\right)\right)^{0.5}}\right)\right), \tag{21}$$

where u is a uniform random variable defined on interval $(0,1)$.

We can obtain the three quartiles Q1, Q2 and Q3 from (21) by using u = 0.25, 0.50 and 0.75 respectively.

2.3.3 Survival function

The survival function for the MGom random variable $\mathrm{X}\sim\mathrm{MGom}\,(a,b,c)$ from the cdf in (7) is obtained as

$$S(x,a,b,c)=1-F(x)=1-\frac{2}{\sqrt{2\pi}}\gamma\left(\frac{3}{2},\frac{1}{2a^2}\left(\frac{1-e^{-\frac{c}{b}\left(e^{bx}-1\right)}}{e^{-\frac{c}{b}\left(e^{bx}-1\right)}}\right)^2\right). \tag{22}$$

The plot of the survival function of MGom for different parameter values is displayed in Figure 3.

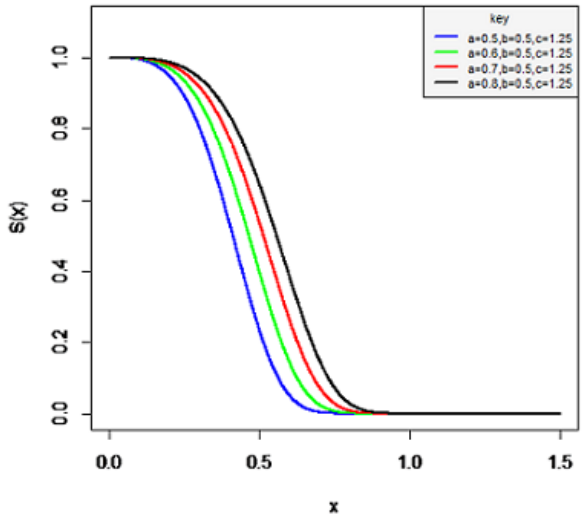


Figure 2: *Plots of the survival function of MGom distribution*

As observed from the plots in Figure 2, the value of the survival function equals one at initial value of zero, it decreases as x increases and degenerates to zero as x becomes larger, which is a major characteristic of survival function.

2.3.4 Hazard function

The hazard function can be obtained using the pdf in (8) and survival function in (22) as

$$h(x)=\frac{f(x)}{S(x)}=\frac{2ce^{bx}\left(1-e^{-\frac{c}{b}(e^{bx}-1)}\right)^{2}\exp\left(-\frac{1}{2a^{2}}\left(\frac{1-e^{-\frac{c}{b}(e^{bx}-1)}}{e^{-\frac{c}{b}(e^{bx}-1)}}\right)^{2}\right)}{a^{3}\sqrt{2\pi}\left(e^{-\frac{c}{b}(e^{bx}-1)}\right)^{3}\left[1-\frac{2}{\sqrt{2\pi}}\gamma\left(\frac{3}{2},\frac{1}{2a^{2}}\left(\frac{1-e^{-\frac{c}{b}(e^{bx}-1)}}{e^{-\frac{c}{b}(e^{bx}-1)}}\right)^{2}\right)\right]}. \tag{23}$$

The plots for the hazard function of MGom distribution for different parameter values are shown in Figure 4.

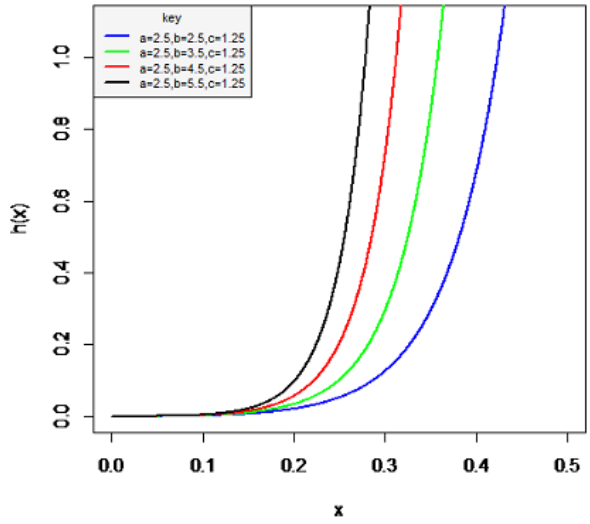


Figure 3: *Plot of the hazard function of MGom distribution*

From the plots in Figure 3, it is observed that the value of the hazard function increases as X increases, meaning that the conditional probability of failure within a given interval of time for a random variable following MGom distribution increases as life ages.

2.3.5 Rényi entropy

If X is a random variable with density function $f(x)$ as defined in (8), then the Rényi entropy of the MGom distribution is defined as

$$R_{\iota}=\frac{1}{1-\iota}\left[\int_{-\infty}^{\infty}f^{\iota}(x)dx\right]. \qquad t>0, t\neq 1$$

(24)

The term $f^{\iota}(x)$ in (24) can be simplified as

$$f^{\iota}(x)=\frac{(2c)^{\iota}e^{\iota bx}\,e^{-\frac{c\iota}{b}\left(e^{bx}-1\right)}}{\left(a^{3}\sqrt{2\pi}\right)^{\iota}}\sum_{i=0}^{\infty}\frac{(-\iota)^{i}}{i!\left(2a^{2}\right)^{i}}\frac{\left(1-e^{-\frac{c}{b}\left(e^{bx}-1\right)}\right)^{2\iota+2i}}{\left(e^{-\frac{c}{b}\left(e^{bx}-1\right)}\right)^{4\iota+2i}}. \tag{25}$$

By applying (9) to (25), we have

$$f^{\iota}(x)=\frac{(2c)^{\iota}e^{\iota bx}\,\mathrm{e}^{-\frac{c\iota}{b}\left(e^{bx}-1\right)}}{\left(a^{3}\sqrt{2\pi}\right)^{\iota}}\sum_{i=0}^{\infty}\frac{(-\iota)^{i}}{i!\left(2a^{2}\right)^{i}}\sum_{j=0}^{\infty}\frac{\Gamma(4\iota+2i+j)}{j!\Gamma(4\iota+2i)}\left(1-e^{-\frac{c}{b}\left(e^{bx}-1\right)}\right)^{2\iota+2i+j}. \tag{26}$$

Using binomial expansion defined in (11), equation (26) becomes

$$f^{\iota}(x)=\frac{(2c)^{\iota}}{\left(a^{3}\sqrt{2\pi}\right)^{\iota}}\sum_{i,j,k=0}^{\infty}\frac{(-\iota)^{i+k}\,\Gamma(4\iota+2i+j)\Gamma(2\iota+2i+j+1)}{i!\,j!k!\left(2a^{2}\right)^{i}\Gamma(4\iota+2i)\Gamma(2\iota+2i+j+1-k)}e^{\iota bx}\,\mathrm{e}^{-\frac{c(\iota+k)}{b}\left(e^{bx}-1\right)}. \tag{27}$$

which on simplification becomes

$$f^{\iota}(x)=\sum_{l,m=0}^{\infty}\Xi_{i,j,k,l,m}e^{-bx(m-\iota-l)}, \tag{28}$$

where $\Xi_{i,j,k,l,m}=\frac{(2c)^{\iota}}{\left(a^{3}\sqrt{2\pi}\right)^{\iota}}\sum_{i,j,k=0}^{\infty}\frac{(-\iota)^{i+k+l+m}\Gamma(4\iota+2i+j)\Gamma(2\iota+2i+j+1)\Gamma(l+1)}{i!\,j!k!l!m!\left(2a^{2}\right)^{i}\Gamma(4\iota+2i)\Gamma(2\iota+2i+j+1-k)\Gamma(l+1-m)}\left(\frac{c}{b}(\iota+k)\right)^{l}$. By taking the integral of (28) and substituting (17) gives

$$\begin{aligned}\int_{-\infty}^{\infty}f^{\iota}(x)dx&=\sum_{l,m=0}^{\infty}\Xi_{i,j,k,l,m}\int_{0}^{\infty}e^{-bx(m-\iota-l)}dx\\&=\frac{\sum_{l,m=0}^{\infty}\Xi_{i,j,k,l,m}}{b(m-l-1)}\int_{0}^{\infty}e^{-y}dy=\frac{\sum_{l,m=0}^{\infty}\Xi_{i,j,k,l,m}}{b(m-l-1)}.\end{aligned} \tag{29}$$

Substituting (29) in (24), the Rényi entropy of the MGom distribution can be given as

$$R_{\iota}=\frac{1}{1-\iota}\left|\frac{\sum_{l,m=0}^{\infty}\Xi_{i,j,k,l,m}}{b(m-l-1)}\right|. \qquad \iota\neq 1 \tag{30}$$

2.3.6 Order statistics

Suppose that $X_1,X_2,...,X_n$ is a random sample of size n from MGom distribution and $X_{(1)},X_{(2)},...,X_{(n)}$ denote the corresponding order statistics of the sample, then the pdf of the i^{th} order statistics is given as

$$f(x)=\frac{n!}{(i-1)!(n-1)!}f(x)[F(x)]^{i-1}[1-F(x)]^{n-1}, \tag{31}$$

where $F(x)$ and $f(x)$ are defined in (7) and (8). Using the definition of binomial expansion for the term $[1-F(x)]^{n-1}$, (31) can be expressed as

$$f(x)=\frac{n!}{(i-1)!(n-1)!}\sum_{i=0}^{n-i}(-1)^{k}\binom{n-i}{k}f(x)F(x)^{i+k-1}. \tag{32}$$

Consequently, using (7) and (8), the pdf of i^{th} order statistics for the MGom distribution can be obtained as

$$f_{i,n}^{MGom}=\frac{n!}{(i-1)!(n-1)!}\sum_{i=0}^{n-i}(-1)^{k}\binom{n-i}{k}f_{MGom}(x)F_{MGom}(x)^{i+k-1}. \tag{33}$$

From (33), The pdf of the smallest and largest order statistics can be obtained by setting i = 1 and i = n respectively.

2.3.7 Parameter estimation

This section derives the maximum likelihood estimator of MGom distribution. Let $X_1, X_2, ..., X_n$ be a random sample of size n drawn from X~ MGom(Θ) with observed values $x_1, x_2 \ldots x_n$, where $\Theta = (a,b,c)^T$ is a $p \times 1$ vector of parameters to be estimated. The likelihood function is given as

$$L(\Theta) = \left(\frac{2c}{a^3\sqrt{2\pi}}\right)^n \prod_{i=1}^{n} e^{bx} e^{-\frac{c}{b}\left(e^{bx}-1\right)} \left(\frac{1-e^{-\frac{c}{b}\left(e^{bx}-1\right)}}{e^{-\frac{c}{b}\left(e^{bx}-1\right)}}\right)^2 \prod_{i=1}^{n} \exp\left(-\frac{1}{2a^2}\left(\frac{1-e^{-\frac{c}{b}\left(e^{bx}-1\right)}}{e^{-\frac{c}{b}\left(e^{bx}-1\right)}}\right)^2\right). \quad (34)$$

The log likelihood function $(\ell\ell)$ is obtained as

$$\ell\ell = n\log(2) + n\log(c) - \frac{n}{2}\log(2\pi) - 3n\log(a) + b\sum_{i=1}^{n}(x_i) + \frac{c}{b}\sum_{i=1}^{n}\log\left(e^{bx_i}-1\right)$$
$$+2\sum_{i=1}^{n}\log\left(\frac{1-e^{-\frac{c}{b}\left(e^{bx_i}-1\right)}}{e^{-\frac{c}{b}\left(e^{bx_i}-1\right)}}\right) - \frac{1}{2a^2}\sum_{i=1}^{n}\left(\frac{1-e^{-\frac{c}{b}\left(e^{bx_i}-1\right)}}{e^{-\frac{c}{b}\left(e^{bx_i}-1\right)}}\right)^2. \quad (35)$$

Taking the partial derivatives of (35) with respect to *a, b* and *c* to obtain

$$\frac{\partial\ell\ell}{\partial a} = \frac{-3n}{a} + \frac{1}{a^3}\sum_{i=1}^{n} w_i^2, \quad (36)$$

$$\frac{\partial\ell\ell}{\partial b} = \sum_{i=1}^{n} x_i + \left[\frac{c}{b}\sum_{i=1}^{n}\left(e^{bx_i}\log\left(e^{x_i}\right)\right) - \frac{c}{b}\sum_{i=1}^{n}(\vartheta_i)\right] + 2\sum_{i=1}^{n}\left(\frac{\varpi_i}{w_i e^{-c\vartheta_i}}\right) - \frac{1}{a^2}\sum_{i=1}^{n}\left(\frac{w_i\varpi_i}{e^{-c\vartheta_i}}\right), \quad (37)$$

and

$$\frac{\partial\ell\ell}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n}(\vartheta_i) + 2\sum_{i=1}^{n}\left(\frac{\vartheta_i}{e^{-c\vartheta_i}}\right) - \frac{1}{a^2}\sum_{i=1}^{n}\left(\frac{w_i\varpi_i}{e^{-c\vartheta_i}_i}\right), \quad (38)$$

where $\vartheta_i = \frac{e^{bx_i}-1}{b}$, $w_i = \frac{1-e^{-c\vartheta_i}}{e^{-c\vartheta_i}}$ and $\varpi = \frac{c}{b}\left(x_i e^{bx_i} - \frac{\vartheta_i}{b}\right)$.

Setting $\frac{\partial\ell\ell}{\partial a} = 0$, $\frac{\partial\ell\ell}{\partial b} = 0$ and $\frac{\partial\ell\ell}{\partial c} = 0$, and solving the resulting nonlinear system of equations, we can obtain the maximum likelihood estimates $\hat{a}, \hat{b}, \hat{c}$. However, these equations cannot be solved analytically, thus statistical software can be used to solve them numerically using iterative methods.

III. Results

3.1 Simulation study

A simulation study is carried out here to investigate the performance of the maximum likelihood estimates of MGom distribution. The simulation is based on the quantile function defined in (21) for four sets of parameter vector $\Theta = (a,b,c)$. We generate 1000 replications of random samples of sizes 50, 100, 200 and 500. The four sets of the parameter's values are assigned as follows:

Set 1: a = 0.5, b = 0.5, c = 0.5
Set 2: a = 1.0, b = 1.0, c = 1. 0
Set 3: a = 2.0, b = 2.0, c = 2.0
Set 4: a = 0.5, b = 2.0, c = 1.0.

The maximum likelihood estimates $\hat{\Theta} = (\hat{a}, \hat{b}, \hat{c})$ are determined based on each generated sample, by maximizing the log-likelihood function in (35). The average estimates, average bias, denoted Bias and Root mean square error (RMSE) are then determined where

$$\text{Bias}(\hat{\Theta}) = \frac{1}{1000}\sum_{j=1}^{1000}(\hat{\Theta}_j - \Theta) \quad \text{and} \quad \text{RMSE}(\hat{\Theta}) = \left[\frac{1}{1000}\sum_{j=1}^{1000}(\hat{\Theta}_j - \Theta)^2\right]^{1/2}.$$

The results of the simulation study are displayed in Tables 1 and 2.

Table 1: *The parameter estimates (Est), Bias and RMSE.*

		a = 0.5, b = 0. 5, c = 0.5			a = 1.0, b =1. 0, c = 1.0		
N	Parameter	Est	Bias	RMSE	Est	Bias	RMSE
	a	0.5374	0.0374	1.9486	1.2385	0.2385	2.1963
50	*b*	0.5187	0.0187	1.5791	1.2189	0.2189	1.7649
	c	0.5265	0.0265	1.8411	1.3019	0.3019	1.9253
	a	0.5210	0.0210	1.2814	1.2273	0.2273	1.8189
100	*b*	0.5146	0.0146	1.3166	1.1916	0.1916	1.4729
	c	0.5158	0.0158	1.5778	1.2342	0.2342	1.6071
	a	0.5113	0.0113	1.1519	1.1218	0.1218	1.4658
200	*b*	0.5138	0.0138	1.1608	1.1126	0.1126	1.2075
	c	0.5114	0.0114	1.2764	1.1480	0.1480	1.4526
	a	0.5037	0.0037	0.7342	1.0490	0.0490	0.8903
500	*b*	0.5069	0.0069	0.6969	1.0307	0.0307	0.5933
	c	0.5023	0.0023	0.7564	1.0657	0.0657	0.7505

Table 2: *The parameter estimates (Est), Bias and RMSE.*

		a = 2.0, b = 2. 0, c = 2.0			a = 0.5, b =1.0, c = 2.0		
N	Parameter	Est	Bias	RMSE	Est	Bias	RMSE
	a	2.2882	0.2882	2.2462	0.5164	0.0164	1.8385
50	*b*	2.1828	0.1828	1.8214	1.2062	0.2062	2.0963
	c	2.3944	0.3944	2.0236	2.1642	0.1642	2.1462
	a	2.1901	0. 1901	2.1038	0.5099	0.0099	1.1713
100	*b*	2.1643	0.1643	1.6454	1.1616	0.1616	1.7189
	c	2.2729	0.2729	1.9016	2.1001	0.1001	2.0038
	a	2.1421	0.1421	1.9643	0.5056	0.0056	1.0418
200	*b*	2.1226	0.1226	1.4325	1.0823	0.0823	1.3658
	c	2.1933	0.1933	1.7031	2.0992	0.0992	1.8643
	a	2.0442	0.0442	0.8606	0.5021	0.0021	0.6241
500	*b*	2.0320	0.0320	0.6542	1.0361	0.0361	0.7903
	c	2.0580	0.0580	0.7730	2.0542	0.0542	0.7606

3.2 Data Application

Application of the MGom distribution to two real life data sets are provided to show how it can be applied in practice in comparison to other distributions in the family. The proposed distribution is compared with four other Gompertz distribution extensions, namely: power Gompertz (powGom), exponentiated Gompertz (expGom), Marshall-Olkin Gompertz (M-OGom) and odd-logistic Gompertz (Odd-loGom). The goodness-of- fit criteria and tests used in the choice of the most appropriate distribution include Akaike's Information Criterion (AIC), Consistent Akaike's Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan-Quinn Information Criterion (HQIC), as well as Anderson-Darling $\left(A^*\right)$ and Cramér-von Mises $\left(W^*\right)$ tests. These can be computed as follows

$\text{AIC} = -2\ell + 2p$, $\text{CAIC} = -2\ell + \frac{2np}{n-p-1}$, $BIC = -2\ell + p\log(n)$, and

$\text{HQIC} = -2\ell + 2p\log(\log(n))$, $A^* = \left(\frac{9}{4n^2} + \frac{3}{4n} + 1\right)\left\{n + \frac{1}{n}\sum_{j=1}^{n}(2j-1)\log[z_i(1-z_{n-j+1})]\right\}$,

$W^* = \left(\frac{1}{2n} + 1\right)\left\{\sum_{j=1}^{n}\left(z_i - \frac{2j-1}{2n}\right)^2 + \frac{1}{2n}\right\}$, where $z_i = F(x_i)$ and x_i's are the ordered observations, ℓ is the maximized log likelihood of the parameter vector $\Theta = (a,b,c)$, n is the number of observations, and p is the number of estimated parameters.

The model with the smallest value of these measures is preferred to other models.

Dataset 1: This dataset is taken from [35]. The data represent the time between failures of 30 repairable items.

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.

Dataset 2: The dataset consists of 100 observations of breaking stress of carbon fibers (in Gba) given by [36] as given below:

0.39, 0.81, 0.85, 0.98, 1.08, 1.12, 1.17, 1.18, 1.22, 1.25, 1.36, 1.41, 1.47, 1.57, 1.57, 1.59, 1.59, 1.61, 1.61, 1.69, 1.69, 1.71, 1.73, 1.80, 1.84, 1.84, 1.87, 1.89, 1.92, 2.00, 2.03, 2.03, 2.05, 2.12, 2.17, 2.17, 2.17, 2.35, 2.38, 2.41, 2.43, 2.48, 2.48, 2.50, 2.53, 2.55, 2.55, 2.56, 2.59, 2.67, 2.73, 2.74, 2.76, 2.77, 2.79, 2.81, 2.81, 2.82, 2.83, 2.85, 2.87, 2.88, 2.93, 2.95, 2.96, 2.97, 2.97, 3.09, 3.11, 3.11, 3.15, 3.15, 3.19, 3.19, 3.22, 3.22, 3.27, 3.28, 3.31, 3.31, 3.33, 3.39, 3.39, 3.51, 3.56, 3.60, 3.65, 3.68, 3.68, 3.68, 3.70, 3.75, 4.20, 4.38, 4.42, 4.70, 4.90, 4.91, 5.08, 5.56.

Dataset 1

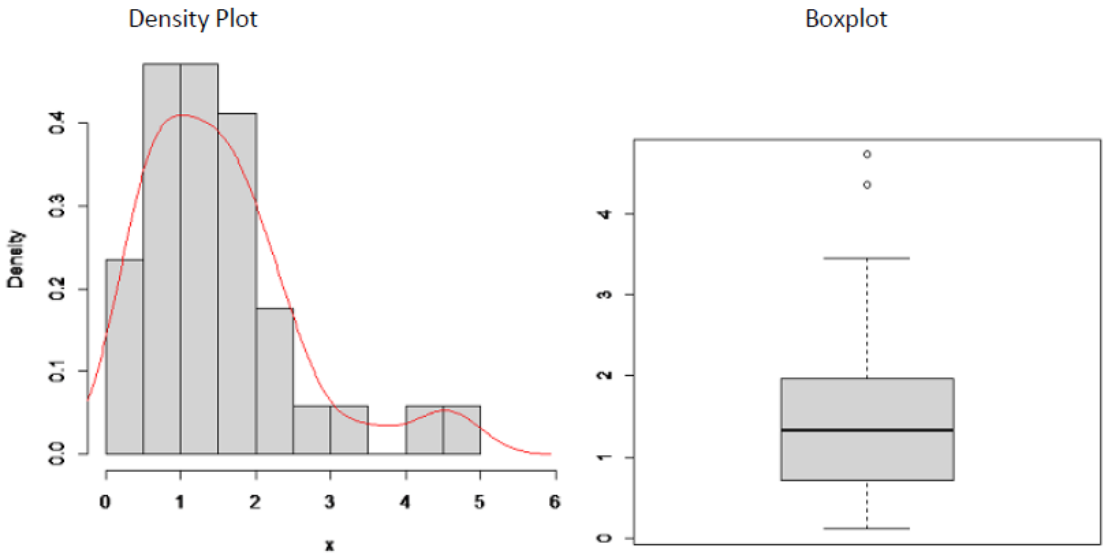


Figure 4: *Density and boxplots for dataset 1*

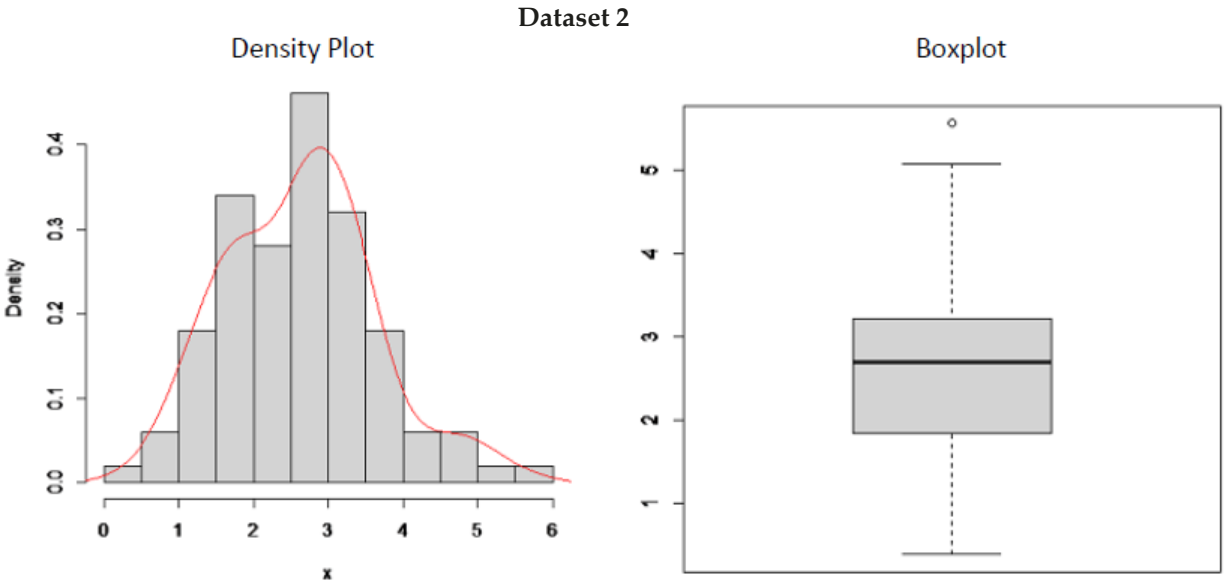


Figure 5: *Density and boxplots for dataset 2*

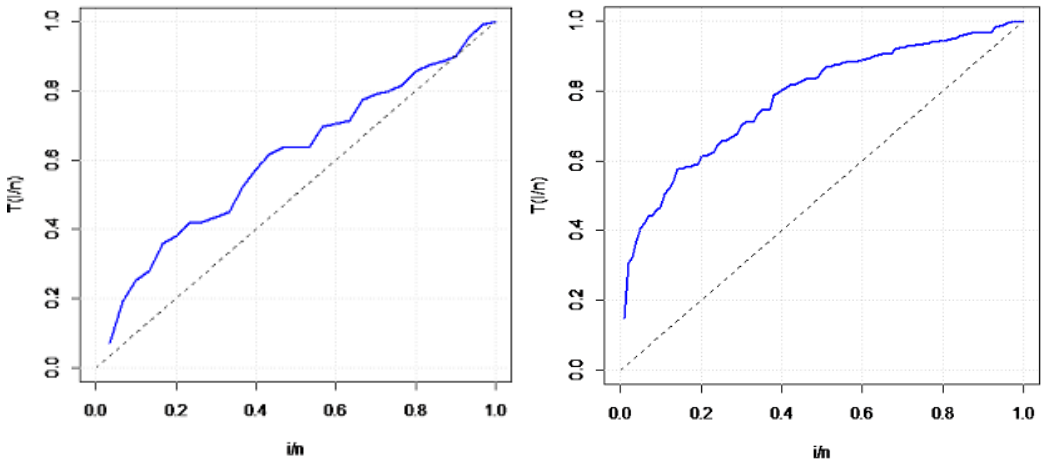


Figure 6: *TTT Plots for datasets 1 and 2*

Table 3: *MLEs and goodness-of-fit-statistics for dataset 1.*

Model	*a*	*B*	*C*	AIC	CAIC	BIC	HQIC	A*	W*
MGom	0.5463	0.1282	1.2171	85.3735	85.6965	89.5771	86.7183	0.1290	0.0173
PowGom	0.4806	0.0131	1.3732	86.9282	86.9513	91.6231	89.3729	0.3642	0.0532
ExpGom	1.8355	0.2814	1.3073	86.2447	86.8677	91.1483	88.0894	0.3466	0.0471
M-OGom	0.3422	0.3516	0.2785	88.6075	89.1306	92.3111	89.9522	0.4120	0.0868
Odd-lGom	0.5555	0.0223	1.5030	86.0282	86.2965	90.2318	87.2183	0.2114	0.0279

Table 4: *MLEs and goodness-of-fit-statistics for dataset 2*

Model	*a*	*B*	*C*	AIC	CAIC	BIC	HQIC	A*	W*
MGom	0.2628	0.1015	1.8936	278.528	278.779	286.344	281.691	0.3542	0.0563
PowGom	0.1010	0.0561	1.8671	298.771	299.022	306.587	301.934	0.5600	0.0791
ExpGom	0.3620	0.5327	0.7557	293.555	293.805	301.371	296.718	0.5109	0.0711
M-OGom	0.0835	0.7208	0.2346	307.515	307.765	315.331	310.678	0.8052	0.1309
Odd-lGom	0.1329	0.5124	1.6392	291.109	291.359	298.924	294.272	0.4564	0.0674

IV. Discussion

As observed from Tables 1 and 2, for all the different parameter settings, the average of the estimates for *a, b,* and *c* get closer to the true parameter values as the sample size increases. Also, the average Bias and the RMSE decrease as the sample size increases. These results validate the asymptotic properties of maximum likelihood estimators.

As observed from the density plot as well as box plot depicted in Figures 4 and 5, it is clear that dataset 1 is heavily skewed to the right and, dataset 2 is moderately skewed to the right, hence the two datasets are could be good for a flexible model like MGom distribution. The total time on test (TTT) curve of the datasets are also plotted in Figure 6 to obtain the empirical behaviour of the hazard function. As observed, the shapes of the hazard function of both datasets are concave showing increasing hazards, and this could also be a good candidate for Gompertz distribution and any of its compound distributions.

Tables 3 and 4 present the maximum likelihood estimates and the values of goodness-of-fit statistics for datasets 1 and 2 respectively. It was found that MGom distribution had the smallest values of all these measures (AIC, CAIC, BIC, HQIC, A^* and W^*) and therefore can be best used in comparison to other Gompertz extensions for modelling real life situations of positively skewed data with increasing hazard rates.

References

[1] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). Continuous Univariate Distributions. 2nd *ednion John Wiley and Sons, New York.*

[2] Willemse, W. and Koppelaar H. (2000). Knowledge elicitation of Gompertz' law of mortality. *Scandinavian Actuarial Journal*, 2:168-179.

[3] Brown, K. and Forbes, W. (1974). A mathematical model of aging processes. *Journal Gerontology,* 29(1): 46-51.

[4] Ohishi, K., Okamura, H. and Dohi, T. (2009). Gompertz software reliability model: Estimation Algorithm and empirical validation. *Journal of Systems and Software,* 82(3): 535-543.

[5] Economos, A. C. (1982). Rate of aging, rate of dying and the mechanism of mortality. *Archieves of Gerontology and Geriatrics,* 1(1): 46-51.

[6] Bemmaor, A. C. and Glady, N. (2012). Modeling purchasing behavior with sudden "death": A flexible customer lifetime model. *Management Science,* 58(5): 1012-1021.

[7] Pollard, J. H. and Valkovics, E. J. (1992). The Gompertz distribution and its applications. *Genus,* 48(3): 15-28.

[8] Missov, T. I. and Lenart, A. (2011). Linking period and cohort life-expectancy linear increases in Gompertz proportional hazards models. *Demographic Research,* 24: 455-468.

[9] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and it applications. *Communication in Statistics - Theory and Methods,* 31(4): 497-512.

[10] Haghbin, H., Ozel, G., Alizadeh, M. and Hamedani, G. G. (2017). A new generalized odd log-logistic family of distributions. *Communication in Statistics - Theory and Methods*, 46(20): 9897-9920.

[11] Cordeiro, G. M., Alizadeh, M., Ramires, T. G. and Ortega, E. M. M. (2018). The generalized odd half-Cauchy family of distributions: Properties and applications. *Communication in Statistics - Theory and Methods,* 46: 5685-5705.

[12] Tahir, M. H., Hussain, M. A., Gauss, M, C., El-Morshedy, M. and Eliwa, M. S. (2020). A New Kumaraswamy Generalized Family of Distributions with Properties, Applications, and Bivariate Extension. *Mathematics,* 8(11). DOI:10.3390/math8111989.

[13] Ishaq, A. I., Usman, A., Tasi'u, M., Suleiman, A. A. and Ahmad, A. G. (2022). A New Odd F-Weibull Distribution: Properties and Application of the Monthly Nigerian Naira to British Pound Exchange Rate Data. *2022 International Conference on Data Analytics for Business and Industry (ICDABI),* 326-332.

[14] Suleiman, A. A., Daud, H., Singh, N. S. S., Othman, M., Ishaq, A. I. and Sokkalingam, R. (2023). A Novel Odd Beta Prime-Logistic Distribution: Desirable Mathematical Properties and Applications to Engineering and Environmental Data. *Sustainability,* 15 (1). DOI:10.3390/su151310239.

[15] Ishaq, A. I., Panitanarak, U., Abiodun, A. A., Suleiman, A, A. and Daud, H. (2024). The Generalized Odd Maxwell-Kumaraswamy Distribution: Its Properties and Applications. *Contemporary Mathematics,* 5: 711-742.

[16] Alexander, C., Cordeiro, G. M., Ortega, E. M. M. and Sarabia, J. M. (2012). Generalized beta-generated distributions. *Computational Statistics and Data Analy*sis, *56(6)*: 1880-1897.

[17] Elgarhy, M., Hassan, A. S. and Rashed, M. (2016). Garhy-Generated Family of Distributions with Application. *Mathematical Theory and Modeling*, 6: 1-15.

[18] Torabi, H., Montezari, N. H. (2012). The gamma-uniform distribution and its application. *Kybernetika*, 48:16-30.

[19] Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M. and Zubair, M. (2016). The Logistic-X family of distributions and its applications. *Communication in Statistics - Theory and Methods,* 45: 7326-7349.

[20] Ahmad, Z., Elgarhy, M. and Hamedani, G. G. (2018). A new Weibull-X family of distributions: properties, characterizations and applications. *Journal of Statistical Distributions and Applications,* 5: 1-18.

[21] Kyurkchiev, N., Iliev, A. and Rahnev, A. (2019). Comments on a Zubair-G Family of Cumulative Lifetime Distributions. Some Extensions. *Communications in Applied Analy*sis, 23(1): *1-20.*

[22] Elbatal, Z., Ahmad, M., Elgarhy, A. M. and Almarashi, J. (2019). A new alpha power transformed family of distributions: properties and applications to the Weibull model. *The journal of Nonlinear Sciences and Applications,* 12: 1-20.

[23] Elbatal, I., Jamal, F., Chesneau, C., Elgarhy, M. and Alrajhi, S. (2018). The Modified Beta Gompertz Distribution: Theory and Applications. *Mathematics,* 7 *(3). DOI:10.3390/math7010003.*

[24] El-Gohary, A. and Al-Otaibi, A. N. (2013). The generalized Gompertz distribution. *Applied Mathematical Modelling*, 37(1-2):13-24.

[25] Gupta, R. D. and Kundu, D. (1999). Generalized exponential distribution. *Austrian and New- Zealand Journal of Statistics,* 41: 173-188.

[26] Ogunde, A. A., Olayode, F. and Audu, A. A. (2020). Cubic Transmuted Gompertz Distribution: As a Life Time Distribution. *Journal of Advances in Mathematics and Computer Sci*ence, 35(1): 105-116.

[27] El-Damcese, M. A., Mustafa, A., El-Desouky, B. S. and Mustafa, M. E. (2015). The Odd generalized exponential Gompertz distribution. *Applied Mathematics,* 6: 2340-2353.

[28] Abdul-Moniem, I. B. and Seham, M. (2015). Transmuted Gompertz distribution. *Computational and Applied Mathematics,* 1(3): 88-96.

[29] Kuje, S., Lasisi, K. E., Nwaosu, S. C. and Alkafawi, A. M. A. (2019). On the properties and applications of the odd Lindley- Gompertz distribution. *Asian Journal of Science and Technology*, 10(10): 10364-10370.

[30] Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. Part I. On the motions and collisions of perfectly elastic spheres. *The London Edinburgh and Dublin Philosophical Magazine and Journal of Science,* 19:19-32.

[31] Ishaq, A. I. and Abiodun, A. A. (2020). The Maxwell-Weibull Distribution in Modeling

Lifetime Datasets. *Annals of Data Science.,* 7: 639-662.

[32] Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions," *Metron,* 71(1): 63-79.

[33] Ishaq, A. I. and Abiodun, A. A. (2021). On the developments of Maxwell-Dagum distribution. *Journal of Statistical Modelling: Theory and Appl*ications, 2(2): 1-23.

[34] Abiodun, A. A. and Ishaq, A. I. (2022). On Maxwell-Lomax distribution: properties and Applications. *Arab Journal of Basic and Applied Sciences,* 29(1): 221-232.

[35] Murthy, D. N. P., Xie, M. and Jiang, R. (2004). Weibull models. *Hoboken, New Jersey: Wiley-Interscience*.

[36] Hassan, A. S., Sabry, M. A. H and Elsehetry, A. M. (2020). A New Family of Upper-Truncated Distributions: Properties and Estimation. *Thailand Statisticians,* 18(2): 196-214.

DETERMINATION OF VITERBI PATH FOR 3 HIDDEN AND 5 OBSERVABLE STATES USING HIDDEN MARKOV MODEL

T. Raja jithendar, M. Tirumala Devi, and G. Saritha

•

Department of Mathematics, Kakatiya University, Warangal, Telangana, India-506009
spjcjithu@gmail.com, oramdevi@gmail.com, gouravenisaritha66@gmail.com

Abstract

Hidden markov model (HMM) is a statistical markov model in which the system being modeled and is assumed to be a markov process with unobservable (i.e., Hidden) states. In HMM, the state is not directly visible but the output depend on the state is visible. Each state has a probability distribution over the possible output tokens. The model is referred to as a hidden markov model even if these parameters are known exactly. The viterbi is one of the estimate underlying state path in hidden markov models. In this paper, viterbi path is derived using hidden markov model.

Keywords: Hidden markov model, Viterbi algorithm, Hidden states, Observable states.

I. Introduction

Viterbi algorithm is a dynamic programming algorithm to obtain the maximum posterior probability estimate of the most likely sequence of hidden states and viterbi path results in a sequence of observed events, especially in the context of hidden markov model. Antibiotics are medicines that fight against bacterial infections in people and animals. They work by killing the bacteria or by making it hard for the bacteria to grow and multiply. Antibiotics can be taken in different ways orally (by mouth) this could be, tablets, pills, capsules, or liquids. Another way to take is topically, this might be a cream, ointment, eye drops, or ear drops, through an injection intramuscular or intravenously, this is usually used for more serious infections. Maria Luz Gamiz [7] et al applied hidden markov models in reliability and maintenance. Janani Kalyanam [3] discussed the probabilistic algorithm for list viterbi decoding. Viterbi A.J [10] derived error bounds for convolution codes and asymptotically optimum decoding algorithm. Shinghal R[9] et al described the modification of the viterbi algorithm formally, and a measure of its complexity is derived. The modified algorithm uses heuristic to limit the search through a directed graph or trellis. G Saritha[2] et al discussed the reliability for 4-modular and 5-modular redundancy system by using Markov technique.

II. Mathematical model

Assume there are M possible states to choose from, the state could be any one of {1,2,3,…,m}. The transition probability should be quantified from state *i* to state j asa_{ij}. The transition could happen from any one of the M possible states to another one of the M possible states, there are in total M×M possibilities. They can be arranged in the following matrix representation, known as state transition matrix S.

$$S = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

Here a_{ij},*i*=1,2,3,….m, j=1,2,3,…m are probability values between 0 and 1, each row has to be summed to 1, and a_{ij} can be written as

$$a_{ij} = P\left[X_k = j \mid X_{k-1} = i\right]$$

Therefore,

$S_k^T = S_{k-1}^T A$, i.e., matrix multiplication of previous state with transition,

where $S_k = \begin{bmatrix} P\left[X_k = 1\right] \\ P\left[X_k = 2\right] \\ \text{.....} \\ P\left[X_k = m\right] \end{bmatrix}$

The emission matrix B is the probabilities of a state *i* in an observed value j. The observation has K possible values {0,1, 2,…,k}.

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mk} \end{bmatrix}$$

Here b_{ij} represents the probability of state *i* to emit observable j and can be written as

$$b_{ij} = P\left[Y_k = j \mid X_k = i\right]$$

Therefore,

$O_k^T = S_k^T B$, where $O_k = \begin{bmatrix} P\left[Y_k = 1\right] \\ P\left[Y_k = 2\right] \\ \text{.....} \\ P\left[Y_k = m\right] \end{bmatrix}$

Initial state probability distribution is denoted by $\pi_0$and is given by

$\pi_0 = [P_1 P_2 \ldots P_i \ldots P_m]$

The transition matrix is a regular matrix whose elements are probabilities of one state to another state. The probability between hidden states to observable states is called emission probability, the matrix representation of emission probabilities is called emission matrix. Here the rows represent hidden states and columns represent observable states.

In this paper the hidden state space S= {Nausea, Diarrhea, Stomach pain} and the observable state space B={Amoxicillin + Potassium Clavunate(ap),Cefixime(ce), Amoxicillin(am), Azithromycin (az), Ciprofloxacin(cp)}.

The initial probability $\pi_0 = [0.45,\ 0.3,\ 0.25]$

The transition probability matrix between hidden states is S=

	N	D	S
N	0.5	0.2	0.3
D	0.2	0.6	0.2
S	0.1	0.2	0.7

The emission matrix is B=

	ap	ce	am	az	cp
N	0.33	0.22	0.19	0.14	0.12
D	0.35	0.15	0.15	0.25	0.1
S	0.15	0.35	0.25	0.15	0.1

II. Viterbi Algorithm

For the large number of possibilities forward and backward algorithm cannot be used to get the maximum probability. Viterbi algorithm is used to obtain the maximum posterior probabilities of the most likely sequence of hidden states. The total possibilities are $m = n^t = 3^5 = 243$, where n is the number of hidden states and t is the number of observations. Out of these possibilities, let us consider the viterbi path 2 4 1 3 5. Then

$P(2,N) = P(2/N)P(N) = 0.099$
$P(2,D) = P(2/D)P(D) = 0.045$
$P(2,S) = P(2/S)P(S) = 0.0875$

And the viterbi probabilities are

$V_1(1) = 0.099,\ V_1(2) = 0.045,\ V_1(3) = 0.0875,$

Similarly

$P(4,N) = P(4/N)P(N/N) = 0.07, \quad P(4,D) = 0.05, \ P(4,S) = 0.045$
$P(4,N) = P(4/N)P(N/D) = 0.028, \quad P(4,D) = 0.15, \ P(4,S) = 0.03$
$P(4,N) = P(4/N)P(N/S) = 0.014, \quad P(4,D) = 0.05, P(4,S) = 0.105$

And the viterbi probabilities are

$$V_2(1) = \max\begin{Bmatrix} 0.099 \times 0.07 = 0.00693, \\ 0.045 \times 0.028 = 0.0012, \\ 0.0875 \times 0.014 = 0.001225 \end{Bmatrix} = 0.00693, \quad V_2(2) = 0.00675, \quad V_2(3) = 0.0091875,$$

$$P(1,N) = P(1/N)P(N/N) = 0.165, \quad P(1,D) = 0.07, \quad P(1,S) = 0.045$$
$$P(1,N) = P(1/N)P(N/D) = 0.066, \quad P(1,D) = 0.21, \quad P(1,S) = 0.03$$
$$P(1,N) = P(1/N)P(N/S) = 0.033, \quad P(1,D) = 0.07, \quad P(1,S) = 0.105$$

And the viterbi probabilities are

$$V_3(1) = 0.00114345, \quad V_3(2) = 0.0014175, \quad V_3(3) = 0.0009646875,$$
$$P(3,N) = P(3/N)P(N/N) = 0.095, \quad P(3,D) = 0.03, \quad P(3,S) = 0.075$$
$$P(3,N) = P(3/N)P(N/D) = 0.038, \quad P(3,D) = 0.09, \quad P(3,S) = 0.05$$
$$P(3,N) = P(3/N)P(N/S) = 0.019, \quad P(3,D) = 0.03, \quad P(3,S) = 0.175$$

And the viterbi probabilities are

$$V_4(1) = 0.0001086278, \quad V_4(2) = 0.000127575, \quad V_4(3) = 0.0001688203$$
$$P(5,N) = P(5/N)P(N/N) = 0.06, \quad P(5,D) = 0.02, \quad P(5,S) = 0.03$$
$$P(5,N) = P(5/N)P(N/D) = 0.024, \quad P(5,D) = 0.06, \quad P(5,S) = 0.02$$
$$P(5,N) = P(5/N)P(N/N) = 0.012, \quad P(5,D) = 0.02, \quad P(5,S) = 0.07$$

And the viterbi probabilities are

$$V_5(1) = 0.0000065177, \quad V_5(2) = 0.0000076545, \quad V_5(3) = 0.0000118174,$$

Now,

$\text{Max}\{V_1(1), V_1(2), V_1(3)\}=0.099$
$\text{Max}\{V_2(1), V_2(2), V_2(3)\}=0.0091875$
$\text{Max}\{V_3(1), V_3(2), V_3(3)\}=0.0014175$
$\text{Max}\{V_4(1), V_4(2), V_4(3)\}=0.0001688203$
$\text{Max}\{V_5(1), V_5(2), V_5(3)\}=0.000011817$

The above probabilities are shown in the following diagram and it gives the final path

N ⟶ S ⟶ D ⟶ S ⟶ S

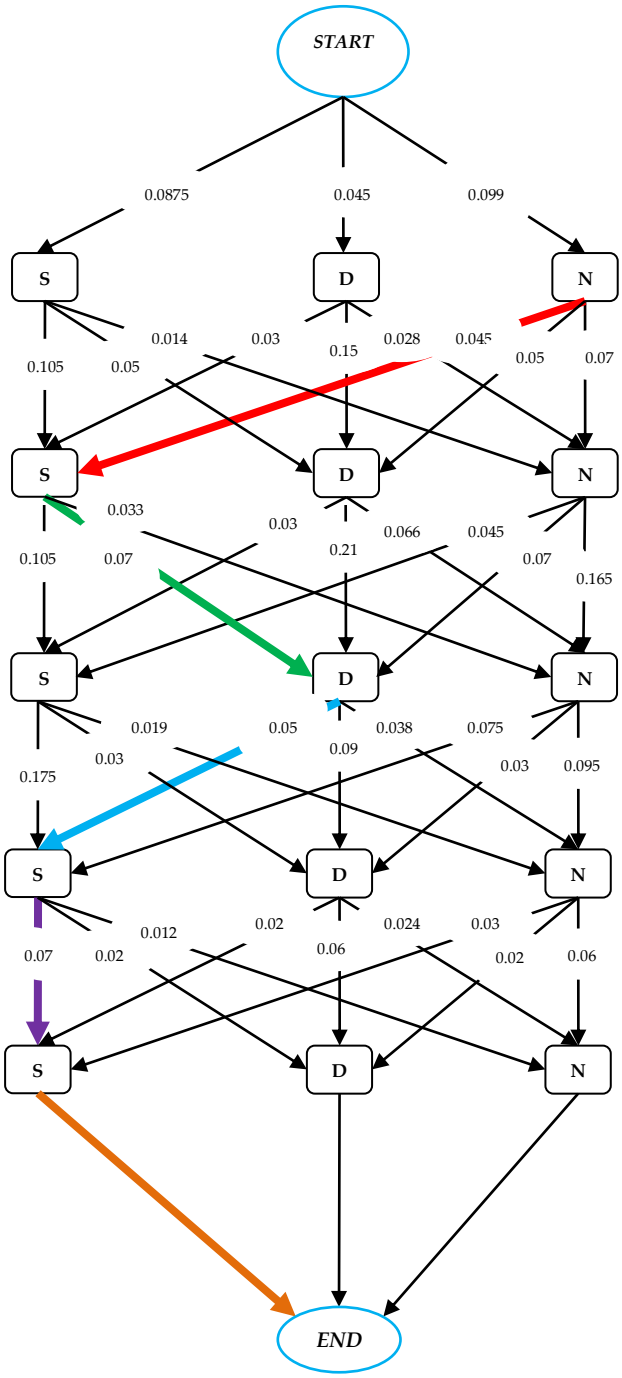


Figure 1: *The different states probabilites*

.

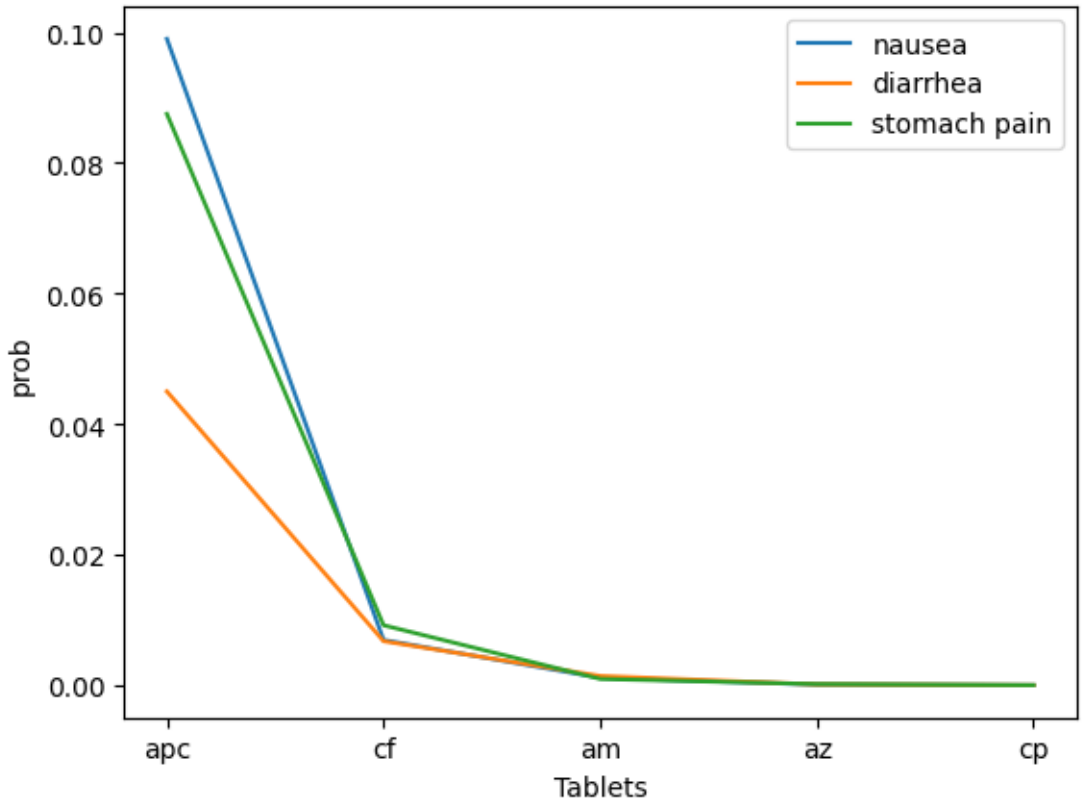


Figure 2: *viterbi probabilities of hidden states*

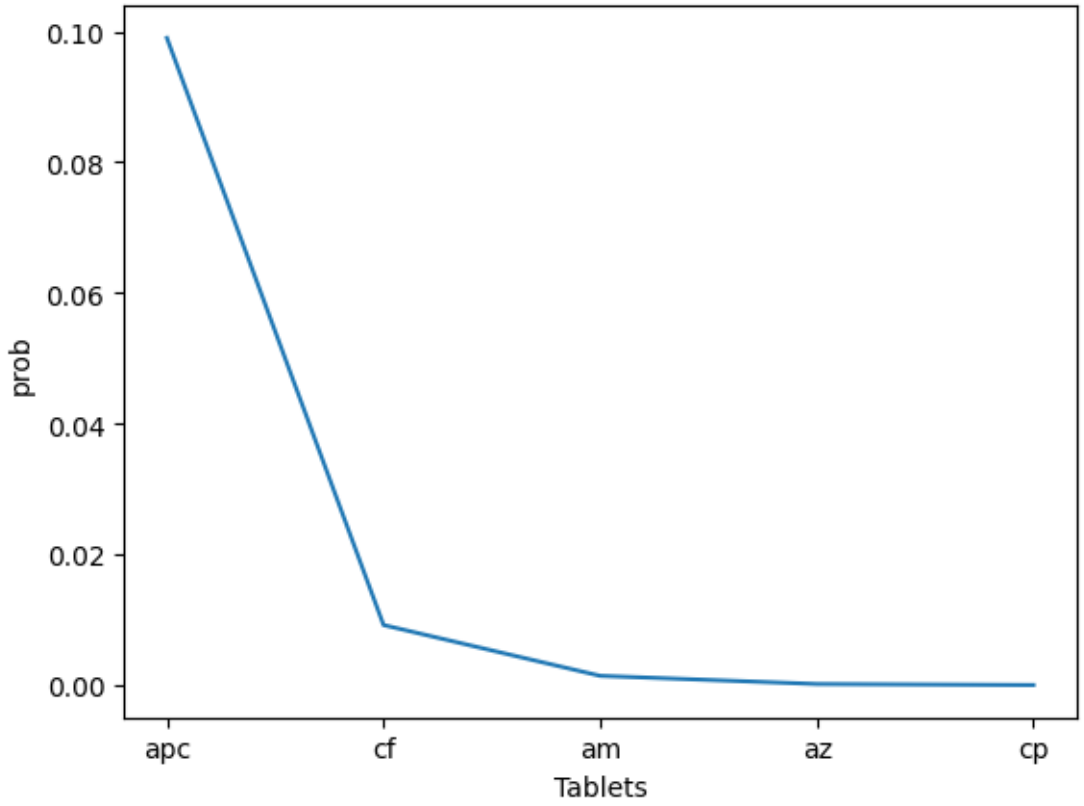


Figure 3: *Maximum probabilities of viterbi path*

From figure 2, we observe that, for tablet Amoxicillin + Potassium Clavunate (apc), the probability for Nausea, Diarrhea, Stomach pain are maximum and the tablet Cefixime(cf), takes the second place and the reaming are almost equal.
From figure 3, we observe that the Maximum probabilities of viterbi path for different antibiotic. Amoxicillin + Potassium Clavunate (ap) gets maximum and the second maximum probability is for Cefixime(ce), and the reaming are almost equal.

IV. Conclusion

In this paper, I used the technique of hidden markov model for 5 observable, 3 hidden states to find the viterbi path. By the path, we observe that the most common side effects of antibiotics is stomach pain. The maximum path for hidden markov model is Nausea, Stomach pain, diarrhea, Stomach pain, Stomach pain with maximum probability 0.0000118174.

I.Acknowledgement

The data is taken from the article "Availability, Prices and Affordability of Antibiotics Stocked by Informal Providers in Rural India".

References

[1] E Balagurusamy [1984]. "*Reliability Engineering*". Tata McGraw-Hill Publishing Company Limited.

[2] G. Saritha, T.SumathiUmaMaheswari, M.TirumalaDevi [2020]."Reliability Model for 4-Modular and 5-Modular Redundancy System by using Markov Technique". *Advances in Intelligent Systems and Computing*. Vol. 979,pp.329-338

[3] Janani Kalyanam [2009]. "*Probabilistic algorithm for list viterbi decoding[Thesis]*". University of Wisconsin, Madison.

[4] Jeffrey W. Miller[2016]. "*Hidden Markov Models*". Lecture Notes on Advanced Stochastic Modeling. Duke University, Durham, NC.

[5] K. K. Aggarwal[1993]. "*Reliability Engineering*". Kluwer Academic Publishers.

[6] L. S. Srinath[2005]. "*Reliability Engineering*". Fourth Edition, Affiliated East-West Press Private, Limited, New Delhi.

[7] Maria Luz Gamiz, Nikolaos Limnios, Mariadel Carmen Segovia-Garcia [2022]. "Hidden markov models in reliability and maintenance". *European Journal of Operational Research,* Vol. 304, pp:1242-1245.

[8] Meenakshi Gautham, Rosalind Miller, Sonia Rego and Catherine Goodman [2022]. "Availability, Prices and Affordability of Antibiotics Stocked by Informal Providers in Rural India: A Cross-Sectional Survey". https://www.mdpi.com/journal/antibiotics

[9] Shinghal, R. and Godfried T. Toussaint, [1979]. "Experiments in text recognition with the modified Viterbi algorithm". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. PAMI-l, pp. 184–193.

[10] Viterbi A.J [1967]. "Error bounds for convolutional codes and asymptotically optimum decoding algorithm". *IEEE Transactions on Information Theory*. Vol.13, Iss 2.

[11] Z. Yang, Y. Li, W. Chen and Y. Zheng [2012] "Dynamic hand gesture recognition using hidden Markov models," 7th International Conference on Computer Science & Education (ICCSE). Melbourne, VIC, Australia, 2012, pp. 360-365.

APPLICATION OF EXTENDED LOMAX DISTRIBUTION ON THE RELIABILITY ANALYSIS OF SOLAR PHOTOVOLTAIC SYSTEM

[*1]Anas Sani Maihulla, [2]Ibrahim Yusuf, [3]Michael Khoo B. C., and [4]Ameer Abdullahi Hassan

•

[1]Department of Mathematics, Sokoto State University, Sokoto- Nigeria
anas.maihulla@ssu.edu.ng
[2]Department of Mathematical sciences, Bayero University, Kano- Nigeria
iyusuf.mth@buk.edu.ng
[3]School of Mathematical Sciences, Universiti Sains Malaysia
mkbc@usm.my
[4]Department of Mathematics Kano University of Science and Technology, Kano- Nigeria
ameernigeria@gmail.com
[*1,2,4]Operation research group, Bayero University, Kano- Nigeria

Abstract

In this study, a novel distribution called the Extended-Lomax distribution which generalizes the existing Lomax distribution and has increasing and decreasing shapes for the hazard rate function was proposed. Various structural properties of the new proposed distribution are derived including the survival function, hazard function, and rth moment. The probability density function (PDF) plots indicated that the distribution is skewed to the right. To estimate the parameters of the newly proposed distribution, two estimation methods which include the Maximum likelihood approach and Method of Moments was employed. The main objective of the proposed distribution's construction was to increase the adaptability of the current Lomax distributions so that they could better suit reliability data sets than alternative candidate distributions with an equivalent number of parameters. This distribution should be able to eliminate the Heavy-tail of the current distribution and model both monotonic and non-monotonic patterns of failure rates. Solar photovoltaic system reliability data was used to evaluate the performance of the proposed Extended Lomax distribution as well as the estimation methods.

Keywords: Lomax distribution, Moments, Heavy-tail, monotonic, failure rates, reliability

1. Introduction

Reliability analysis plays a crucial role in ensuring the dependable performance of industrial systems. The Lomax distribution, a popular probability distribution for modeling lifetime data and reliability, has been employed in various applications. However, to improve the accuracy of reliability assessments for a repairable industrial system, it is essential to extend the Lomax distribution and adapt it to the characteristics of the data collected from that system. The moments and inference for the order statistics and generalized order statistics are given in [1], [2,] and [3], respectively. Assuming all distributional assumptions are satisfied, the goodness of fit between the probability distribution and the provided data sets plays a critical role in the precision of parametric statistical inference and data set modeling. A lot of studies have gone into creating

distributions with more flexible and desirable features so that real-world data sets with different densities and failure rates can be adequately modeled. Currently, researchers are focused on creating new hybrid distributions that generalize existing ones, aiming to achieve better data modeling capabilities. These hybrid distributions are formed by combining a baseline distribution with a family distribution. Several authors have extensively reviewed different families of distributions Hamedani et al. [4]. The distributional properties, estimation and inference of the Lomax distribution are described in the literature as follows. In record value theory, some properties and moments for the Lomax distribution have been discussed in [5], [6], [7], [8]. Reliability analyses of solar photovoltaic system using Gumbel-Hougaard family copula distribution as studied by Maihulla et al. [16]. RAMD analyses was used by Anas and Yusuf [17] for analyzing the photovoltaic system. Performance prediction of small solar system for house used was studied by Anas and Yusuf [18].

This distribution was constructed with the primary goal of improving the flexibility of classical distributions so that they can fit survival data sets better than other candidate distributions with an equal number of parameters. This distribution should be capable of modeling various types of failure rates, including monotonic and non-monotonic patterns. The Lomax distribution, also known as the Pareto distribution of the second kind with two parameters (α, λ), has attracted significant attention from theoretical and statistical researchers due to its applications in reliability and lifetime testing studies. Lomax first introduced and studied this distribution in 1954, and it has since been utilized for analyzing business failures, as well as in economic, behavioral, scientific, and traffic modeling. The solar energy conversion into electricity is a very promising technique, knowing that the source is free, clean and abundant in several countries. However, the effect of the solar cell's temperature on the photovoltaic panel performance and lifespan remains one of the major disadvantages of this technology.

The present research has been categorized into seven main sections. Introduction is the first section. Proposed extended Lomax distribution is presented in section 2. Section 3 comprised up of the estimation methods used in the research. Mathematical properties of the proposed distribution were presented in section 4. Application of the proposed distribution to the Solar photovoltaic system were presented in section 5 of the paper. Summary and conclusion were presented in section 6 and 7 respectively.

1.1 Reliability theory

Reliability theory is a branch of applied mathematics that focuses on the study of the reliability and failure of systems, components, and processes. It is concerned with understanding and quantifying the probability that a system or component will operate without failure over a specified period or under given conditions. Reliability theory is widely used in engineering, quality control, and risk analysis to design, evaluate, and improve the dependability of various systems and products. Reliability theory encompasses issues such as:

1.1.1 Reliability engineering

Reliability engineering deals with the interdisciplinary use of probability, statistics and stochastic modeling, combined with engineering insights into the design and the scientific understanding of the failure mechanism, to study the various aspects of reliability. Frequently, it is desirable to understand and be able to predict the overall system failure characteristics for any given configuration.

1.1.2 Reliability modeling

Reliability modeling deals with model building to obtain solutions to problems in predicting, estimating and optimizing the survival or performance of an unreliable system, the impact of the unreliability, and actions to mitigate this impact [19][20].

1.1.3 Reliability management

Reliability management deals with the various management issues in the context of managing the design, manufacture and/or operation of reliable products and systems. The emphasis is on the business viewpoint, as unreliability has consequences in cost; time wasted, and in certain cases the welfare or an individual or the security of a nation.

1.2 Standard Lomax distribution

The Lomax distribution, also known as the Pareto Type II distribution or the shifted Pareto distribution, is a probability distribution used in statistics and probability theory. It is often used to model heavy-tailed or long-tailed data and is closely related to the Pareto distribution. The Lomax distribution is defined by the probability density function (PDF)[9][21]:

$$f(x) = \frac{\lambda\kappa}{(1+\lambda\kappa)^{(\kappa+1)}} \quad (1)$$

The cumulative distribution function associated with the (1) above is;

$$F(x) = 1 - (1+\lambda\kappa)^{-\kappa} \quad (2)$$

Survival function (Reliability) as:

$$R(x) = (1+\lambda\kappa)^{-\kappa} \quad (3)$$

Hazard rate function corresponding to (3) will be;

$$h(x) = \frac{f(x)}{R(x)} = \frac{\lambda\kappa}{(1+\lambda\kappa)} \quad (4)$$

One of the key characteristics of the Lomax distribution is its heavy tail, which means that it has a higher probability of extreme values compared to many other probability distributions. The tail index λ controls the heaviness of the tail. When λ is small, the tail is heavier, indicating that extreme values are more likely. An important aspect of the Lomax distribution is how the values of the shape and the scale parameter affect such distribution characteristics as the shape of the PDF curve, the reliability, and the failure rate.

2. The Proposed Extended Lomax Distribution

Consider a continuous distribution G with density g and the Weibull cdf

$$F(x) = 1 - e^{-\alpha x^{\beta}} \quad (5)$$

With positive parameters α and $\boldsymbol{\beta}$. Based on this density, by replacing x with $\frac{G(x)}{\bar{G}(x)}$ for $\bar{G}(x) = 1 - G(x)$ Silver et al. [15] define the cdf family by:

$$F(x;\alpha,\beta,\zeta) = \int_0^{\frac{G(x)}{\bar{G}(x)}} \alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}} dt = 1 - \exp[-\alpha[\frac{G(x;\zeta)}{\bar{G}(x;\zeta)}]^{\beta}] \quad (6)$$

Where $G(x;\zeta)$ is a baseline cdf, which depends on a parameter vector $\boldsymbol{\zeta}$. The family pdf is reduced to:

$$(x;\alpha,\beta,\zeta) = \alpha\beta g(x;\zeta)\frac{G(x;\zeta)^{\beta-1}}{\bar{G}(x;\zeta)^{\beta+1}} \exp[-\alpha[\frac{G(x;\zeta)}{\bar{G}(x;\zeta)}]^{\beta}] \quad (7)$$

The hazard rate function is given by:

$$h(x;\alpha,\beta,\zeta) = \frac{\alpha\beta g(x;\zeta)G(x;\zeta)^{\beta-1}}{\bar{G}(x;\zeta)^{\beta+1}} \quad (8)$$

Inserting (4) into (6) we have;

$$F(x) = 1 - \exp[-\alpha\left[\frac{1-(1+\lambda\kappa)^{-\kappa}}{(1+\lambda\kappa)^{-\kappa}}\right]^{\beta}] \quad (9)$$

From the above equation, we've

$$F(x) = 1 - \exp[-a[1-(1+\lambda\kappa)^{\kappa}]^{b} \quad (10)$$

From the above CDF we've

$$R(x) = \exp[-a[1-(1+\lambda\kappa)^{\kappa}]^{b} \quad (11)$$

To evaluate for $f(x)$ i.e probability density function

$$f(x) = \frac{-dR(x)}{dt} \quad (12)$$

We need to find the derivative of $R(t)$ with respect to t. Given expression for $R(t)$ and $R(x) = \exp[-a[1-(1+\lambda\kappa)^{\kappa}]^{b}$ using chain rule:

$$\frac{dR(x)}{dt} = \exp[-a[1-(1+\lambda\kappa)^{\kappa}]^{b}\frac{d}{dx}(1-(1+\lambda\kappa)^{\kappa}) \quad (13)$$

We next find $\frac{d}{dx}(1-(1+\lambda\kappa)^{\kappa})$ using chain rule again:

$$\frac{d}{dx}(1-(1+\lambda\kappa)^{\kappa}) = -ab(1-(1+\lambda\kappa)^{\kappa})^{b-1}\frac{d}{dx}(1+\lambda\kappa)^{\kappa}$$

Simplifying the equation we've

$$= ab\kappa\lambda((1+\lambda\kappa)^{\kappa})^{b-1}(1+\lambda\kappa)^{\kappa-1} \quad (14)$$

Therefore, the probability density function for the newly generated extended Lomax distribution is:

$$f(x) = \exp[-a[1-(1+\lambda\kappa)^{\kappa}]^{b}[ab\kappa\lambda((1+\lambda\kappa)^{\kappa})^{b-1}(1+\lambda\kappa)^{\kappa-1}] \quad (15)$$

The Hazard rate function h(x) represents the instantaneous failure rate at time x. It is defined as the conditional probability that a system or component fails at time x given that it has survived up to that time. For continuous distribution, the hazard rate can be calculated as the ratio of the PDF to the survival function.

$$h(x) = \frac{f(x)}{R(x)} \quad (16)$$

To find the hazard rate function, we need to differentiate the negative logarithm of the reliability function with respect to time (x)

$$-lnR(x) = -\ln[\exp[-a[1-(1+\lambda\kappa)^{\kappa}]^{b}] = [a[(1+\lambda\kappa)^{\kappa}]^{b}] \quad (17)$$

Differentiating the above equation with respect to x using chain rule;

$$\frac{d}{dx}[a[(1+\lambda\kappa)^{\kappa}]^{b}] = [ab\kappa\lambda((1+\lambda\kappa)^{\kappa})^{b-1}(1+\lambda\kappa)^{\kappa-1}] \quad (18)$$

Therefore, the hazard rate function for the newly generated extended Lomax distribution corresponding to (10) and (15);

$$h(x) = ab\kappa\lambda((1+\lambda\kappa)^{\kappa})^{b-1}(1+\lambda\kappa)^{\kappa-1} \quad (19)$$

3. Parameter estimation

The choice of parameter estimation method depends on the nature of the data, the statistical model, the sample size, and the specific goals of the analysis. Each method has its own strengths

and weaknesses, we consider these factors when selecting an appropriate technique.

3.1.1 Maximum Likelihood Estimation (MLE)

MLE is a widely used method for parameter estimation. It seeks to find the values of parameters that maximize the likelihood function, which measures how well the observed data fit the model. MLE is often used for a wide range of statistical models and has desirable properties, such as asymptotic efficiency.

$$L(a,b,\kappa,\lambda;t) \propto [ab\kappa\lambda] + n_2 lna(b+1)\sum_{j=1}^{n}\ln\left(1+\lambda x_j\right) + \sum_{j=1}^{n}\ln(1-(1+\lambda x)^{k}) + n_2(\kappa-1)\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1}(1+\lambda x)^{\kappa-1} + \sum_{j=1}^{n}\sum_{j=n_1+1}^{n}\ln(1-(1+\lambda x)^{k}) + n_2(\kappa-1)\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1}(1+\lambda x)^{\kappa-1} \quad (20)$$

Based on the above equation, by solving the likelihood equations with respect to a, b, κ, λ after equating them to zero, the MLEs $(\bar{a}, \bar{b}, \bar{\kappa}, \bar{\lambda})$ of a, b, κ, λ can be obtained. This procedure can be done as follows:

$$\frac{\partial L}{\partial a} = \frac{n_2}{a} - n_2\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1}(1+\lambda x)^{\kappa-1}) + \sum_{j=n_1+1}^{n}\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1} + (1+\lambda x)^{\kappa-1} \quad (21)$$

$$\frac{\partial L}{\partial b} = \frac{n}{b} + (b-1)\frac{(1+\lambda x)^{-\kappa-1}-(1+\lambda x)^{-1}}{1-(1+\lambda x)^{-1}}\sum_{j=1}^{n}\frac{x_j}{1+\lambda x_j}\left(\kappa+1-\frac{\kappa}{(1+\lambda x_j)^{\kappa}-1}\right) + \kappa(a-1)\sum_{j=n_1}^{n}x_j\frac{(1+\lambda x)^{-\kappa-1}-(1+\lambda x)^{-1}}{1-(1+\lambda x)^{-1}} \quad (22)$$

$$\frac{\partial L}{\partial \kappa} = \frac{n}{k} + (1-b)\ln(1+\lambda t)\frac{(1+\lambda t)^{-\kappa}-1}{1-(1+\lambda x)^{-\kappa}}\sum_{j=1}^{n}\ln\left(1+\lambda t_j\right)\left(1-\frac{1}{(1+\lambda t_j)^{\kappa}-1}\right) + (b-1)\sum_{j=n_1+1}^{n}\ln\left(1+\lambda t_j\right)\frac{\left(1+\lambda t_j\right)^{-\kappa}-1}{1-\left(1+\lambda t_j\right)^{-\kappa}} \quad (23)$$

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} - n(b-1)\ln(ab\kappa)\left((1+\lambda x)^{\kappa}\right)^{b-1} + (\kappa-1)\sum_{j=n_1}^{n}\ln(ab\kappa(1+\lambda t)^{b}) + \left(1-\frac{1}{(1+\lambda t_j)^{\kappa-1}-1}\right) \quad (24)$$

From equation (21), the following equation can be used to calculate $\bar{a}$ as a function of b, κ, λ.

$$\frac{d}{da}\left(\frac{n_2}{a} - n_2\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1}(1+\lambda x)^{\kappa-1}\right) + \sum_{j=n_1+1}^{n}\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1} = 0 \quad (25)$$

To simplify this, we'll first find the derivatives of the individual terms with respect to a, and then set the expression equal to zero.

$$\frac{d}{da}(-n_2\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1}(1+\lambda x)^{\kappa-1}) = -n_2\frac{d}{da}(\ln(ab\kappa\lambda)\left((1+\lambda x)^{\kappa}\right)^{b-1}(1+\lambda x)^{\kappa-1}) = 0 \quad (26)$$

Now, differentiating the natural logarithm term. Differentiating the expression inside the logarithm with respect to

$$\frac{d}{da}(\ln(ab\kappa\lambda)(1+\lambda x)^{b-1}) + (1+\lambda x)^{\kappa-1}) = \frac{1}{((ab\kappa\lambda(1+\lambda x)^{b-1}+(1+\lambda x)^{\kappa-1})}\frac{d}{da}[((ab\kappa\lambda(1+\lambda x)^{b-1}+(1+\lambda_t)^{\kappa-1})] \quad (27)$$

Differentiating the expression inside the logarithm. Summing up the individual derivatives:

$$\bar{a}(b\kappa\lambda) = \frac{n_2}{\kappa^2} - n_2\frac{[(b\kappa\lambda(1+\lambda x)^{b-1}+(1+\lambda_t)^{\kappa-1}]}{((ab\kappa\lambda(1+\lambda x)^{b-1}+(1+\lambda_t)^{\kappa-1})} + \sum_{j=n_1+1}^{n}\frac{[(ab\kappa\lambda(1+\lambda x)^{b-1}]}{((ab\kappa\lambda(1+\lambda x)^{b-1}+(1+\lambda_t)^{\kappa-1})} = 0 \quad (28)$$

From equation (28), the above equation can be used to calculate $\bar{a}$ as a function of b, κ *and* λ.By substituting $\bar{a}(b\kappa\lambda)$ in (22), (23), and (24) the MLE of $b\kappa\lambda$ can be produced by solving the likelihood equations$\frac{\partial L}{\partial b} = 0$, $\frac{\partial L}{\partial \kappa} = 0$, and $\frac{\partial L}{\partial \lambda} = 0$ with regard to b, κ *and* λ by utilizing Newton-rampson numerical iteration method. The MLEs of the reliability function, and Hazard rate function at t_0, denoted by $\bar{R}(t_0)$ and $\bar{h}(t_0)$ are obtained by substituting $(\bar{a}, \bar{b}, \bar{\kappa}, \bar{\lambda})$ in equations (11) and (19) respectively.

4. Mathematical Properties

The major mathematical properties of the proposed extended Lomax distribution are derived and presented in this section.

4.1 Moment

In the method of moments, we set the sample moments equal to the corresponding population moments. This approach provides estimates for the parameters based on simple algebraic equations. Some of the most important features and characteristics of a distribution can be studied through moments. (e.g tendency, dispersion, and skewness) [10].

$$\mu_r = E(X^r) = \int_0^{\infty} x^r f(x)dx \quad (29)$$

$$\mu_r = E(X^r) = \int_0^\infty x^r \exp[-a[1-(1+\lambda\kappa)^\kappa]^b[ab\kappa\lambda((1+\lambda\kappa)^\kappa)^{b-1}(1+\lambda\kappa)^{\kappa-1}]dx \quad (30)$$

In any statistical analysis, especially in the field of applied statistics, it is crucial to keep in mind the importance of moments. Moments can be used to examine key distributional properties like kurtosis, skewness, dispersion, and tendency, among others. Assume that M is a random variable in a Lomax distribution with parameters $\boldsymbol{\kappa}$ and λ, in that case, the r^{th} moment of M is given as: (σ B') $B\ (r+1, \sigma - r)$

$$E(M') = (\frac{\lambda}{\kappa})\kappa(r+1, \lambda - r) \quad (31)$$

Also, suppose X is random variable that assumes the proposed extended Lomax distribution. From equation (30), the r^{th} moment of X is therefore obtained as:

$$E(X^r) = ab\kappa\lambda((1+\lambda\kappa)^\kappa)^{b-1}\frac{ab\kappa\lambda(i+1)}{\kappa^r}\kappa(r+1, ab\kappa\lambda(i+1)-r) \quad (32)$$

The mean is obtained my equating $r = 1$ in equation (32)

5. Application to Solar Photovoltaic System

A statistical model called the Extended Lomax distribution is used in many different domains, such as survival analysis, reliability analysis, and modeling of extreme occurrences. It is an effective instrument for comprehending and forecasting the performance and dependability of solar photovoltaic (PV) systems over time.

The Extended Lomax distribution was used to simulate the possibility of extreme events in the context of solar PV systems over 360 days, such as abrupt declines in efficiency, component failures, or unforeseen changes in energy output. Engineers and academics can estimate the chance of future occurrences and obtain insights into the probability distribution of solar PV system performance by fitting historical data to the Extended Lomax distribution. The efficiency of the proposed life distribution is demonstrated in this section using real-life data sets. The data set contains information about the reliability of Solar photovoltaic system by times (in days) [11]. The data set consists of twenty (12) observations as presented in Table 1 below:

Table 1: *Computation of Reliability for different values of time*

Time (In days)	Reliability
30	0.99912
60	0.94321
90	0.89824
120	0.83345
150	0.74561
180	0.71984
210	0.70125
240	0.66257
270	0.62152
300	0.59452
330	0.54529
360	0.41984

Table 2: *The MLEs and Information Criteria of the models based on the solar data set*

Models	$\bar{a}$	$\bar{b}$	$\bar{\kappa}$	$\bar{\lambda}$
Proposed model	13.237	8.648	9.183	5.984
Mohamed and Essam [12]	3.978	2.967	4.826	5.034
Abdelaziz Alsubie 2021 [13]	6.284	6.936	6.920	3.961
F. Hashem et al. [14]	0.284	-	4.783	0.947

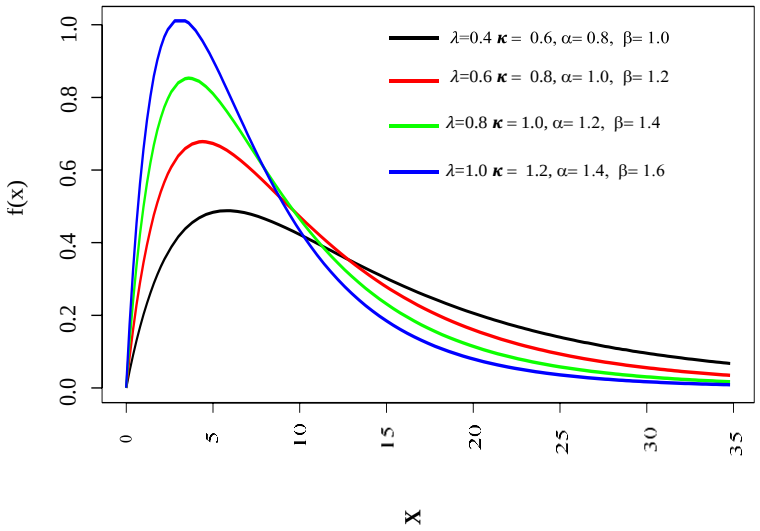


Figure 1: *PDF of the Proposed Extended Distribution*

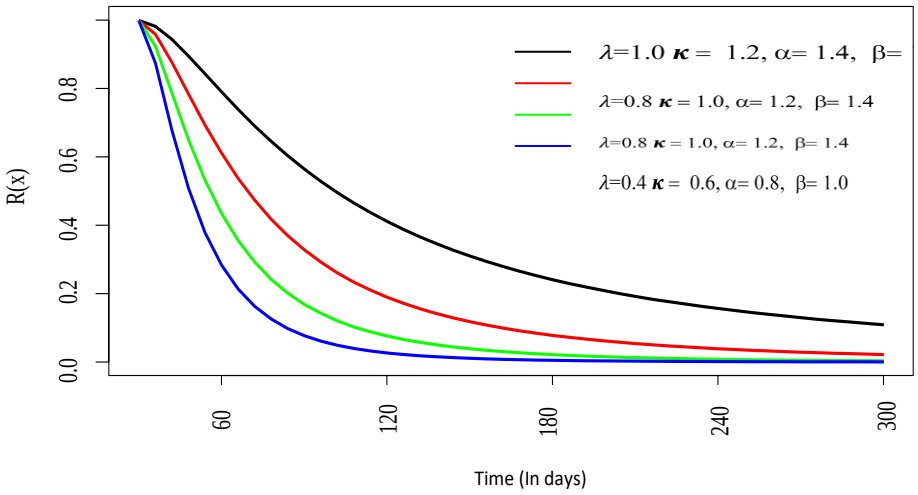


Figure 2: *Reliability Function of the Proposed Extended Distribution*

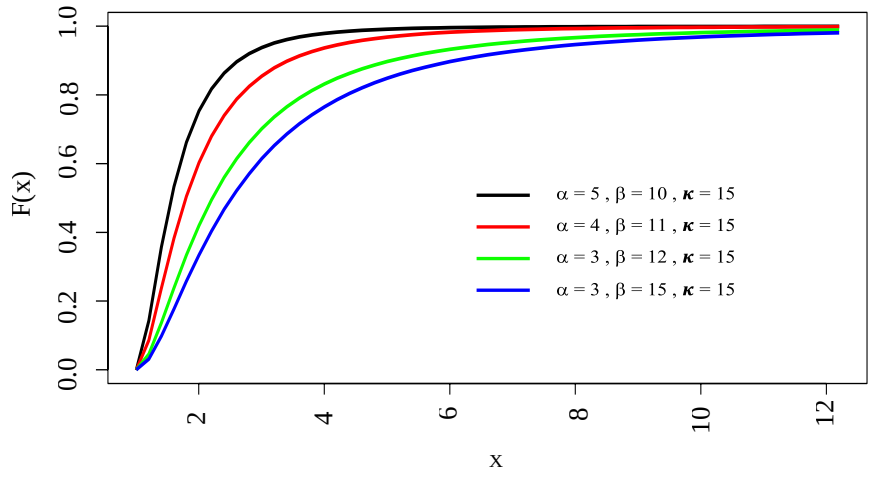


Figure 3: *CDF of the Extended Lomax Distribution*

6. Summary

In this study, we introduce a three-parameter Extended Lomax distribution by compounding the Weibull-G Family of Distributions and the Lomax distributions. The shape of the hazard function of the new compounding distribution can be monotonically decreasing or upside-down bathtub. Some mathematical and statistical properties of the new model are studied. We estimate the model parameters by the Maximum likelihood (MLE) approach, and Method of Moments. We present a simulation study to illustrate the performance of estimators. The flexibility and potentiality of the proposed model are illustrated by means of a real data set (from the reliability analysis of a repairable solar photovoltaic system). We hope that the Extended Lomax distribution may attract a wider range of applications in areas such as engineering, survival and lifetime data, economics, meteorology, hydrology, and others. PDF, Reliability plot and cumulative distribution function plots for the proposed extended Lomax distribution were displayed in Figure 1, 2, and 3 respectively.

7. Conclusion

By addressing the limitations of the standard Lomax distribution and extending it to fit the specific data from an industrial system, this research aims to enhance the reliability analysis and support better decision-making in the context of the industrial system's operation and maintenance. The proposed distribution was compared to some existing distributions, it was discovered that the proposed life distribution is right-skewed and that changing the parameters' values results in different shapes. The proposed Extended Lomax distribution is more efficient than the competing distribution's pdfs and reliability for different values of shape parameters from a real-life data set, as shown in Figures 1.

8. Acknowledgement

I am grateful to Professor Ibrahim Yusuf of Bayero University Kano and Professor Michael Khoo Boon Chong of School of Mathematical Sciences, Universiti Sains Malaysia for contributing to My Ph.D. work. I am also grateful to TETFund for the sponsorship to the USM.

9. Statements and Declarations

The authors did not receive support from any organization for the submitted work. The authors have no competing interest to declare that are relevant to the content of this article.

10. Conflict of interest

The authors declare that they have no conflicts of interest.

11. Statements and Declarations

The authors did not receive support from any organization for the submitted work. The authors have no competing interest to declare that are relevant to the content of this article.

References

[1] Saran, J. and Pushkarna, N. Moments of order statistics from doubly truncated Lomax distribution, Journal of Statistical Research 33, 57–66, 1999.

[2] Childs, A., Balakrishnan, N. and Moshref, M. Order statistics from non-identical right-truncated Lomax random variables with applications, Statistical Papers 42, 187–206, 2001.

[3] Moghadam, M.S., Yahmaei, F. and Babanezhad, M. Inference for Lomax distribution under generalized order statistics, Applied Mathematical Sciences 6, 5241–5251, 2012.

[4] Hamedani G. G., Yousof H. M., Rasekhi M., Alizadeh M. and Najibi S. M. (2018). Type I general exponential class of distributions. Pakistan Journal of Statistics and Operation Research, 14(1): 39-55.

[5] Ahsanullah, M. Recoed values of the Lomax distribution, Statistica Neerlandica 45, 21–29, 1991.

[6] Balakrishnan, N. and Ahsanullah, M. Relations for single and product moments of record values from Lomax distribution, Sankhy˜a Series B 17, 140–146, 1994.

[7] Lee, M-Y. and Lim, E-K. Characterization of the Lomax, exponential and Pareto distributions by conditional expectations of recoed values, Journal of the Chungcheong Mathematical Society 22, 149–153, 2009.

[8] Amin, E.A. Kth upper record values and their moments, International Mathematical Forum 6, 3013–3021, 2011.

[9] Bourguignon, M., Silva, R.B. and Cordeiro, G.M. The Weibull–G family of probability distributions, Journal of Data Science 12, 53–68, 2014.

[10] Halid O.Y., and Sule O.B. (2022). A Classical and Bayesian Estimation Techniques for Gompertz Inverse Rayleigh Distribution: Properties and Application. Pakistan Journal of Statistics, 38(1): 49-76.

[11] Maihulla, A.S., Yusuf, I. & Khoo, M.B.C. Reliability and performance analysis on stages to control manufacturing system using Copula technique. *Life Cycle Reliab Saf Eng* 12, 197–205 (2023). https://doi.org/10.1007/s41872-023-00224-8

[12] Mohamed S. Eliwa and Essam A. Ahmed "Reliability analysis of constant partially accelerated life tests under progressive first failure type-II censored data from Lomax model: EM and MCMC algorithms, AIMS Mathematics, 8(1): 29–60. DOI: 10.3934/math.2023002

[13] Abdelaziz Alsubie, "Properties and Applications of the Modified Kies–Lomax Distribution with Estimation Methods," Hindawi Journal of Mathematics Volume 2021, Article ID 1944864, 18pages https://doi.org/10.1155/2021/1944864.

[14] Laila A. Al-Essa1, Alaa H. Abdel-Hamid, Tmader Alballa1, and Atef F. Hashem,

"Reliability analysis of the triple modular redundancy system under step-partially accelerated life tests using Lomax distribution," scientific reports https://doi.org/10.1038/s41598-023-41363-3 (2023) 13:14719

[15] Bourguignon, M., Silva, R.B. and Cordeiro, G.M. The Weibull–G family of prob- ability distributions, Journal of Data Science 12, 53–68, 2014.

[16] Maihulla, A.S., Yusuf, I. and Salihu Isa, M. (2022), "Reliability modeling and performance evaluation of solar photovoltaic system using Gumbel–Hougaard family copula", *International Journal of Quality & Reliability Management*, Vol. 39 No. 8, pp. 2041-2057. https://doi.org/10.1108/IJQRM-03-2021-0071

[17] Maihulla, A.S., Yusuf, I. Reliability, availability, maintainability, and dependability analysis of photovoltaic systems. *Life Cycle Reliab Saf Eng* 11, 19–26 (2022). https://doi.org/10.1007/s41872-021-00180-1

[18] Maihulla, A.S., Yusuf, I. Reliability and performance prediction of a small serial solar photovoltaic system for rural consumption using the Gumbel-Hougaard family copula. *Life Cycle Reliab Saf Eng* 10, 347–354 (2021). https://doi.org/10.1007/s41872-021-00176-x

[19] Maihulla, Anas Sani; Yusuf, Ibrahim; Bala, Saminu I. "Performance Evaluation of a Complex Reverse Osmosis Machine System in Water Purification using Reliability, Availability, Maintainability and Dependability Analysis." Reliability: Theory & Applications. Sep2021, Vol. 16 Issue 3, p115-131. 17p.

[20] Maihulla, A. S. and Yusuf, I. (2022). RELIABILITY ANALYSIS OF REVERSE OSMOSIS FILTRATION SYSTEM USING COPULA. *Reliability: Theory & Applications*, *17*(2): 163-177. Retrieved from http://www.gnedenko.net/RTA/index.php/rta/article/view/890.

[21] Anas Sani Maihulla, Ibrahim Yusuf, Saminu I. Bala WEIBULL COMPARISON BASED ON RELIABILITY, AVAILABILITY, MAINTAINABILITY, AND DEPENDABILITY (RAMD) ANALYSIS. *Reliability: Theory & Applications*. **2023**, *March 1(72):* 120-132. https://doi.org/10.24412/1932-2321-2023-172-120-132

CONFIDENCE INTERVALS FOR THE PARAMETER OF THE IWUEZE DISTRIBUTION WITH APPLICATIONS TO MEDICAL AND ENGINEERING DATA

Wararit Panichkitkosolkul
•
Department of Mathematics and Statistics, Faculty of Science and Technology,
Thammasat University, Thailand
Thammasat University Research Unit in Mathematical Sciences and Applications, Thailand
wararit@mathstat.sci.tu.ac.th

Abstract

One of the lifetime distributions is the Iwueze distribution, which is constructed by combining the exponential and gamma distributions. In this paper, confidence intervals (CIs) are proposed for the parameter of the Iwueze distribution using the likelihood-based, Wald-type, bootstrap-t, and bias-corrected and accelerated (BCa) bootstrap methods. We evaluated the performance of the proposed CI methods through Monte Carlo simulation in terms of their coverage probability (CP) and average length (AL) in various scenarios. Furthermore, we had also derived the explicit formula for the Wald-type CI, which is straightforward for computation. The simulation results showed that the likelihood-based and Wald-type CIs returned satisfactory results according to coverage probabilities, even for the setting of small sample sizes. On the other hand, both the bootstrap-t and BCa bootstrap CIs yield CPs lower than the nominal confidence level when sample sizes are small. However, as the sample sizes increase, the CP of all CIs tend to approach the nominal confidence level. The parameter values also have a minor influence on the CP of all CIs when the sample size is fixed. Moreover, the AL of all CIs decreases as the sample size increases. The Wald-type and likelihood-based CIs have very similar ALs for all parameter values. In general, the bootstrap-t CI tends to yield the shortest interval. The effectiveness of all CIs was demonstrated by applying them to medical and engineering data, yielding results consistent with those of the simulation study.

Keywords: lifetime distribution, interval estimation, likelihood, Wald, bootstrap

I. Introduction

In reliability and lifetime data analysis, lifetime distributions are statistical distributions that can be used to describe the behavioral structure of lifetime data. Lifetime distributions are utilized to represent the duration before the occurrence of a significant event, such as failure or incidence [1]. The field of lifetime data analysis has had substantial growth and progress in terms of technique, theory, and application. The distribution theory focuses on the capacity to easily handle and adapt to modeling lifespan data. While a tractable probability distribution could be useful for replicating random samples, its practical value to businesses lies in its flexibility [2]. This suggests that while tractable distributions are desirable, more complex ones must be created to support relevant applications.

Many lifetime distributions have been proposed in statistics in the past few decades. Nevertheless, these distributions frequently do not offer a precise match because of either their basic distributional properties or the structure of the lifetime data. Several distribution theory experts are trying to suggest a new lifetime distribution consistent with the stochastic nature of lifespan data. Before 1958, the exponential distribution was the only lifetime distribution accessible for the analysis and modeling of lifetime data. The Lindley distribution was presented by Lindley [3] as an alternative lifespan distribution. Based on their comprehensive analysis of the statistical properties and practical uses, Ghitanty et al. [4] determined that the Lindley distribution offers a much superior match compared to the exponential distribution. Shanker et al. [5] observed that when analyzing exponential and Lindley distributions, there is a significant competition between these two distributions. However, they also identified specific datasets in which neither distribution provided a sufficient fit. Shanker [6, 7] proposed two new one-parameter lifespan distributions, named Shanker distribution and Akash distribution. These distributions demonstrated better fit to data than both exponential and Lindley distributions. Furthermore, the Lindley, exponential, Shanker, and Akash distributions were thoroughly examined by Shanker and Fesshaye [8]. They discovered that while these distributions work well for most datasets, there are some that still do not provide the best fit. In addition, Shanker [9] introduced the Sujatha distribution, which has a considerably better fit when compared to the exponential, Lindley, Shanker, and Akash distributions. Shanker [10] proposed the Garima distribution, a single-parameter lifespan distribution, as a suitable statistical model for data collected from the behavioral sciences. However, this distribution likewise fails to provide a satisfactory match for several actual lifespan datasets.

The current paper is to identify a distribution that can accurately depict the diversity within the data sets while remaining flexible and tractable. When a distribution does not provide a sufficient match, many researchers choose to transform the dataset to meet the assumptions of the distribution. Nevertheless, this approach is unsuitable as it leads to the loss of the dataset's inherent characteristics. Some researchers prefer to adjust the distribution by incorporating extra shape or scale parameters to better fit with the characteristics of the data set. However, in cases where the current distributions are unable to generate a suitable fit, it is more advantageous to seek out an alternative distribution that can. This approach involves refraining from transforming the original dataset or modifying the distribution to fit the dataset. Recently, Elechi et al. [11] proposed the Iwueze distribution, a five-component mixture of exponential and gamma distributions with a constant scale parameter, and different shape parameters 2, 3, 4, and 5. This distribution has superior efficiency in comparison to other one-parameter distributions. The flexibility of the Iwueze distribution is demonstrated through its application to relief times of patients receiving an analgesic.

In the review literature, there is no research study for estimating the confidence intervals (CIs) for the parameter of the Iwueze distribution. Therefore, the objective of the paper is to propose the CIs for the parameter of the Iwueze distribution in four methods, namely, likelihood-based CI, Wald-type CI, bootstrap-t interval, and bias-corrected and accelerated (BCa) bootstrap CI. We conduct a simulation study and analyze real data sets to compare the performance of CIs for the parameter of the Iwueze distribution.

The following is the outline of the paper. In Section 2, the Iwueze distribution are explained. Section 3 involves the computation of the likelihood-based, Wald-type, bootstrap-t, and BCa bootstrap CIs for the parameter of the Iwueze distribution. Section 4 evaluates the effectiveness of the proposed CIs by utilizing Monte Carlo simulation in various circumstances. Section 5 contains two numerical examples. Ultimately, the final section of the paper contains the discussion and conclusions.

II. The Iwueze Distribution

The Iwueze distribution is obtained by combining the exponential and gamma distributions using appropriate mixing probabilities. The gamma distribution has a fixed scale parameter θ and four different shape parameters: 2, 3, 4, and 5. Let X be a random variable which follow the Iwueze distribution with parameter θ. The probability density function (pdf) of the Iwueze distribution can be obtained by utilizing a mixture model with five component mixing probabilities. The pdf is given by

$$f(x;\theta)=\frac{\theta^5}{\theta^4+2\theta^3+6\theta^2+12\theta+24}(1+x+x^2)^2\,e^{-\theta x},\quad x>0,\,\theta>0.$$

Figure 1 shows the plots of the Iwueze distribution pdf with several parameter values θ. The mean (or the first central moment) and variance (or the second central moment) of X are given by

$$E(X)=\mu=\frac{\theta^4+2\left[2\theta^3+3\left(3\theta^2+4(2\theta+5)\right)\right]}{\theta\left(\theta^4+2\theta^3+6\theta^2+12\theta+24\right)},$$

and

$$Var(X)=\sigma^2=\frac{\left(\theta^8+8\theta^7+56\theta^6+240\theta^5+876\theta^4+1344\theta^3+2304\theta^2+2880\theta+2880\right)}{\theta^2\left(\theta^4+2\theta^3+6\theta^2+12\theta+24\right)^2}.$$

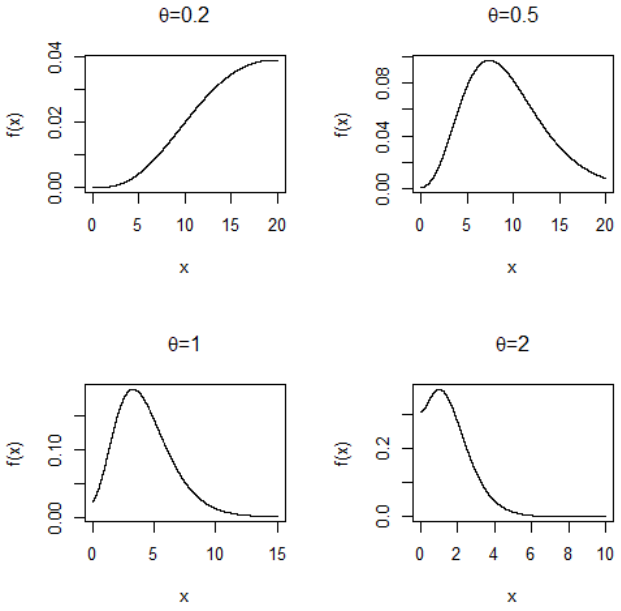


Figure 1: *Plots of the pdf of the Iwueze distribution for θ = 0.2, 0.5, 1, and 2*

The log-likelihood function $\log L(\theta\,|\,x_i)$, is maximized to obtain the point estimator of θ. Therefore, the maximum likelihood (ML) estimator for θ of the Iwueze distribution is derived by the following processes:

$$\frac{\partial}{\partial\theta}\log L(\theta\,|\,x_i)=\frac{\partial}{\partial\theta}\left[5n\log(\theta)-n\log(\theta^4+2\theta^3+6\theta^2+12\theta+24)+\sum_{i=1}^{n}\log\left[1+x_i+x_i^2\right]^2-\theta\sum_{i=1}^{n}x_i\right]$$

$$=\frac{5n}{\theta}-\frac{n\left(4\theta^3+6\theta^2+12\theta+12\right)}{\theta^4+2\theta^3+6\theta^2+12\theta+24}-\sum_{i=1}^{n}x_i.$$

The subsequent equation is a nonlinear equation obtained through the process of solving the

equation $\frac{\partial}{\partial\theta}\log L(x_i;\theta)\overset{\text{set}}{=}0$ for θ,

$$\frac{5n}{\theta}-\frac{n\left(4\theta^3+6\theta^2+12\theta+12\right)}{\theta^4+2\theta^3+6\theta^2+12\theta+24}-\sum_{i=1}^{n}x_i\overset{\text{set}}{=}0.$$

Due to the absence of a closed-form solution for the ML estimator of parameter θ, numerical iteration methods are employed to solve the associated non-linear equation [12]. In this study, the maxLik package [13] was utilized to perform ML estimation using the Newton-Raphson technique in the RStudio program [14].

III. Confidence Intervals for the Parameter of the Iwueze Distribution

I. Likelihood-based Confidence Interval

The likelihood function for the Iwueze distribution, $L(\theta\,|\,x)$, is a function of the parameter θ, given the observed data x. It encapsulates the probability of observing the given data under various hypothetical values of θ. After solving $\frac{\partial}{\partial\theta}\log L(\theta\,|x)\overset{\text{set}}{=}0$, the ML estimator of θ, $\hat{\theta}_{ML}$, will be obtained, and this is the most "likely" estimate given the observed data.

The likelihood-based CI is then constructed around this ML estimator. The process begins by defining a likelihood ratio $\lambda(\theta)$ as $\lambda(\theta)=L(\theta\,|\,x)\big/L(\hat{\theta}\,|\,x)$. Under regular conditions, as per the Wilks' theorem, $-2\log\lambda(\theta)$ follows approximately a chi-square distribution with degrees of freedom equal to the number of parameters being estimated. Therefore, the CI for θ at $(1-\alpha)100\%$ confidence level is given by

$$\left\{\theta\left|-2\log\frac{L(\theta\,|\,x)}{L(\hat{\theta}\,|\,x)}\le\chi^2_{1-\alpha,1}\right.\right\}=\left\{\theta\left|-2\log\left[\frac{\theta^{5n}\left(\hat{\theta}^4+2\hat{\theta}^3+6\hat{\theta}^2+12\hat{\theta}+24\right)^n}{\hat{\theta}^{5n}\left(\theta^4+2\theta^3+6\theta^2+12\theta+24\right)^n}\exp\left(-\theta\sum_{i=1}^{n}x_i+\hat{\theta}\sum_{i=1}^{n}x_i\right)\right]\le\chi^2_{1-\alpha,1}\right.\right\},$$

where $\chi^2_{1-\alpha,1}$ is the critical value from the chi-square distribution with 1 degree of freedom [15,16]. In the specific case of the Iwueze distribution, the likelihood ratio test becomes more intricate due to the composite nature of the distribution. The gamma component, characterized by a scale parameter and shape parameters, adds layers of complexity to the likelihood function, necessitating advanced computational techniques, like numerical optimization, for effective ML estimator calculation and CI construction.

Brent's method, a root-finding algorithm often used in optimization, is used for finding the maximum MLE in the Iwueze distribution. It is an advanced technique that combines the bisection method, the secant method, and inverse quadratic interpolation [17]. Given that

$$f(\theta)=\frac{\partial}{\partial\theta}\log L(\theta\,|\,x)=\frac{5n}{\theta}-\frac{n\left(4\theta^3+6\theta^2+12\theta+12\right)}{\theta^4+2\theta^3+6\theta^2+12\theta+24}-\sum_{i=1}^{n}x_i\overset{\text{set}}{=}0,$$

Brent's method seeks θ such that $f(\theta)=0$. The method combines bracketing methods and open methods. Initially, if $f(a)f(b)<0$ it starts with the bisection method to ensure reliability. Then, depending on the function's behavior, it switches between the secant method (linear interpolation):

$$\theta_{\text{second}}=\theta_n-f(\theta_n)\frac{\theta_n-\theta_{n-1}}{f(\theta_n)-f(\theta_{n-1})},$$

and inverse quadratic interpolation (quadratic polynomial interpolation):

$$\theta_{\text{quad}}=\frac{f(\theta_{n-1})f(\theta_{n-2})}{\left(f(\theta_n)-f(\theta_{n-1})\right)\left(f(\theta_n)-f(\theta_{n-2})\right)}\theta_n+\ldots$$

The method iteratively refines the estimate of the root, switching methods based on which provides a more accurate or stable estimate [18,19]. Figure 2 shows the plot $-2\log\lambda(\theta)$ versus θ (solid blue line), $\chi^2_{0.95,1}$ (dashed red line), and 95% likelihood-based CI (solid green line) when a random sample of size 20 sampled from the Iwueze distribution with $\theta = 1$.

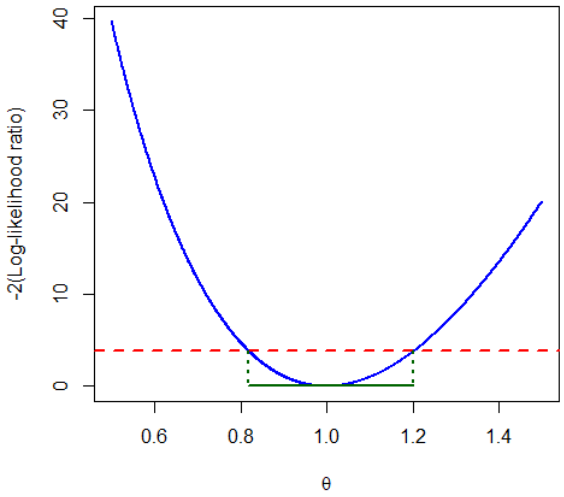


Figure 2: *The plot of* $-2\log\lambda(\theta)$ *versus* θ

Because the cut-point for constructing a likelihood-based CI often involves the use of an asymptotic distribution like the chi-square distribution, this reliance is grounded in Wilks theorem, the effectiveness of the likelihood-based CI in approximating the true parameter values does rely on the assumption that the sample size is sufficiently large for the asymptotic approximation to be valid. However, likelihood-based CI does not always rely on large sample sizes. It can provide accurate interval estimates even in cases with smaller sample sizes, assuming the likelihood function behaves well.

II. Wald-type Confidence Interval

The Wald-type CI is a fundamental statistical tool used for estimating the uncertainty associated with a parameter estimate in a probability distribution. Central to this method is the ML estimate of the parameter, denoted as $\hat{\theta}$ for the Iwueze distribution. The foundation of the Wald-type CI lies in the quadratic approximation of the log-likelihood function, $L(\theta \mid x)$, which can be expanded using a Taylor series around $\hat{\theta}$. The Wald statistic approximates the log-likelihood ratio when expanded to the second-order term around the ML estimate, with the first-order term equal to zero at the ML estimate as follows:

$$\log L(\theta \mid x) \approx \log L(\hat{\theta} \mid x) + \left(\theta - \hat{\theta}\right)\frac{\partial}{\partial\theta}\log L(\theta \mid x)\Big|_{\theta=\hat{\theta}} + \frac{1}{2}\left(\theta - \hat{\theta}\right)^2 \frac{\partial^2}{\partial\theta^2}\log L(\theta \mid x)\Big|_{\theta=\hat{\theta}}$$

$$\log\frac{L(\theta \mid x)}{L(\hat{\theta} \mid x)} \approx \frac{1}{2}\left(\theta - \hat{\theta}\right)^2 \frac{\partial^2}{\partial\theta^2}\log L(\theta \mid x)\Big|_{\theta=\hat{\theta}}$$

$$-2\log\frac{L(\theta \mid x)}{L(\hat{\theta} \mid x)} \approx \left(\theta - \hat{\theta}\right)^2 I\left(\hat{\theta}\right),$$

where $I(\hat{\theta})$ is the estimated observed Fisher information. The Wald statistic can thus serve as an

approximation to the LRT statistic, particularly when the sample size is large enough for the asymptotic properties to hold, leading to a quadratic approximation of the log-likelihood ratio [20-22].

For the Iwueze distribution, the observed Fisher information is as follows:

$$\frac{\partial}{\partial\theta}\log L(\theta\,|\,x)=\frac{5n}{\theta}-\frac{n\left(4\theta^3+6\theta^2+12\theta+12\right)}{\theta^4+2\theta^3+6\theta^2+12\theta+24}-\sum_{i=1}^{n}x_i,$$

$$\frac{\partial^2}{\partial\theta^2}\log L(\theta\,|\,x)=-\frac{5n}{\theta^2}-\frac{12n(\theta^2+\theta+1)}{\theta^4+2\theta^3+6\theta^2+12\theta+24}+\frac{n(4\theta^3+6\theta^2+12\theta+12)^2}{(\theta^4+2\theta^3+6\theta^2+12\theta+24)^2}.$$

Thus, the estimated Fisher information is as follows:

$$I(\hat{\theta})=\frac{5n}{\hat{\theta}^2}+\frac{12n(\hat{\theta}^2+\hat{\theta}+1)}{\hat{\theta}^4+2\hat{\theta}^3+6\hat{\theta}^2+12\hat{\theta}+24}-\frac{n(4\hat{\theta}^3+6\hat{\theta}^2+12\hat{\theta}+12)^2}{(\hat{\theta}^4+2\hat{\theta}^3+6\hat{\theta}^2+12\hat{\theta}+24)^2},$$

and the Wald-type CI for θ at $(1-\alpha)100\%$ confidence level is given by

$$\hat{\theta}\pm z_{1-\frac{\alpha}{2}}\sqrt{I^{-1}(\hat{\theta})},$$

where $z_{1-(\alpha/2)}$ denotes the $\left(1-(\alpha/2)\right)^{\text{th}}$ quantile of the standard normal distribution.

III. Bootstrap-t Confidence Interval

The bootstrap-t CI emerges as an advanced technique designed to calibrate the CI for an estimated parameter by incorporating the inherent variability of the estimate's standard error. This method extends the bootstrap percentile method by factoring in the fluctuation of the standard error, thereby enhancing the accuracy and reliability of the interval, particularly in small sample contexts or when dealing with estimators that deviate from normality [23-25]. The algorithmic foundation of the bootstrap-t CI can be delineated in the following steps:

1) Initialization: Commence with a sample $X_1,\ldots,X_n$ from which the parameter estimate $\hat{\theta}$ and its standard error $S.E.(\hat{\theta})$.

2) Bootstrap Resampling: Generate $B=1000$ bootstrap samples, $X_1^*,\ldots,X_n^*$, by random sampling with replacement from the original dataset.

3) Statistical Computation: For each bootstrap sample, calculate the bootstrap replicate of the estimator, denoted as $\hat{\theta}^*$, and its associated standard error $S.E.(\hat{\theta}^*)$.

4) Studentization: Construct the bootstrap-t statistic for each replicate as

$$t^*(X,\hat{\theta},\hat{\theta}^*)=\frac{\hat{\theta}^*-\hat{\theta}}{\sqrt{I^{-1}(\hat{\theta}^*)}}.$$

This studentized statistic adjusts for the variability in the standard error of the bootstrap estimate.

5) Repeating this process $B=1000$ times yields an empirical distribution of the estimator; from which we can estimate the distribution of the pivotal quantity.

6) Empirical Distribution: Formulate the empirical distribution of the bootstrap-t statistics from the ensemble of B replicates.

7) Quantile Extraction: Ascertain the critical values, $t^*_{(\alpha/2)}$ and $t^*_{(1-(\alpha/2))}$, which correspond to the $\alpha/2$ and $1-(\alpha/2)$ quantiles of the empirical bootstrap-t distribution,

$$\frac{\#\left(t^*(X,\hat{\theta},\hat{\theta}^*)\le t^*_{(\alpha/2)}\right)}{B}=\alpha \text{ and } \frac{\#\left(t^*(X,\hat{\theta},\hat{\theta}^*)\le t^*_{(1-(\alpha/2))}\right)}{B}=1-(\alpha/2).$$

8) Interval Construction: The bootstrap-t CI is then articulated as:

$$\left[\hat{\theta}+t^*_{(\alpha/2)}\sqrt{I^{-1}(\hat{\theta})},\hat{\theta}+t^*_{(1-(\alpha/2))}\sqrt{I^{-1}(\hat{\theta})}\right].$$

IV. Bias-Corrected and Accelerated (BCa) Bootstrap Confidence Interval

The BCa bootstrap CI is a technique used for constructing CIs. This method refines the basic bootstrap procedure by introducing adjustments for both bias and skewness in the distribution of bootstrap estimates. Bias is calculated based on the proportion of bootstrap estimates that are less than the observed estimate, and this information is then used to adjust the percentiles of the CI. An acceleration parameter is incorporated to account for the skewness or asymmetry of the bootstrap distribution [26-28]. The algorithm is as follows:

1) Bootstrap Resampling: Draw $B = 1000$ bootstrap samples from the empirical distribution of the original sample and calculate the bootstrap estimates $\hat{\theta}_b^*$, for $b = 1, 2, ..., 1000$.

2) Bias Correction (z_0): Determine the proportion of bootstrap estimates that are less than the original estimate $\hat{\theta}$, denoted p. The bias correction factor z_0 is the quantile of the standard normal distribution corresponding to p.

3) Acceleration (a): Calculate the acceleration value a which accounts for the asymmetry of the estimator's distribution. This is often estimated by the jackknife or other methods that quantify the skewness of the sampling distribution.

4) Adjusted Percentiles: Transform the bias-corrected normal deviates to adjust the percentiles for constructing the CI. The adjusted percentiles are given by

$$p_L^* = \Phi\left(z_0 + \frac{z_0 + z_{\alpha/2}}{1 - a\left(z_0 + z_{\alpha/2}\right)} \right)$$

and

$$p_U^* = \Phi\left(z_0 + \frac{z_0 + z_{1-(\alpha/2)}}{1 - a\left(z_0 + z_{1-(\alpha/2)}\right)} \right),$$

where Φ is the standard normal cumulative distribution function, and $z_{\alpha/2}$ and $z_{1-\alpha/2}$ are the $(\alpha/2)$-th and $(1-(\alpha/2))$-th quantiles of the standard normal distribution, respectively.

5) CI Construction: The BCa bootstrap CI is constructed using the percentiles p_L^* and p_U^* to extract the corresponding quantiles from the bootstrap distribution of $\hat{\theta}_{ML}^*$. The formula is as follows

$$\left[\hat{\theta}^*_{\left(P_L^*\right)}, \hat{\theta}^*_{\left(P_U^*\right)} \right],$$

where $\hat{\theta}^*_{(P_L^*)}$ and $\hat{\theta}^*_{\left(P_U^*\right)}$ are $(p_L^*)^{\text{th}}$ and $(p_U^*)^{\text{th}}$ quantiles of the bootstrap estimates $\hat{\theta}_b^*$.

IV. Simulation Study and Results

This simulation study evaluates the effectiveness of 95% confidence interval (CI) construction methods in different scenarios. Our focus includes sample sizes, parameter values, coverage probability (CP), and the average length (AL) of the intervals. We vary sample sizes (n) at 10, 20, 30, 40, 100, 200, and 500, while also altering the distribution's parameter values (θ) to 0.2, 0.5, 0.75, 1, 1.5, 2, and 3. The CPs and ALs of the CIs are estimated using Monte Carlo simulations with 2,000 replications.

Table 1. *Coverage probability and average length of the 95% CIs for the parameter of the Iwueze distribution*

n	θ	Coverage probability				Average length			
		Likelihood	Wald	Bootstrap-t	BCa	Likelihood	Wald	Bootstrap-t	BCa
10	0.2	0.953	0.950	0.902	0.905	0.1135	0.1132	0.1014	0.1065
	0.3	0.956	0.952	0.899	0.902	0.1691	0.1688	0.1512	0.1582
	0.5	0.950	0.950	0.896	0.893	0.2811	0.2806	0.2536	0.2661
	0.75	0.949	0.951	0.898	0.903	0.4177	0.4170	0.3770	0.3940
	1	0.952	0.951	0.905	0.903	0.5525	0.5515	0.5007	0.5237
	1.5	0.951	0.957	0.906	0.905	0.8285	0.8252	0.7483	0.7934
	2	0.953	0.959	0.897	0.897	1.1382	1.1303	1.0045	1.0872
	2.5	0.958	0.961	0.908	0.908	1.4943	1.4856	1.3334	1.4720
20	0.2	0.946	0.950	0.923	0.926	0.0792	0.0791	0.0749	0.0763
	0.3	0.961	0.960	0.931	0.927	0.1185	0.1184	0.1120	0.1142
	0.5	0.949	0.945	0.920	0.921	0.1972	0.1971	0.1881	0.1922
	0.75	0.942	0.944	0.919	0.917	0.2931	0.2929	0.2780	0.2833
	1	0.943	0.946	0.925	0.921	0.3872	0.3868	0.3674	0.3740
	1.5	0.945	0.944	0.927	0.926	0.5768	0.5757	0.5458	0.5603
	2	0.947	0.945	0.921	0.923	0.7888	0.7862	0.7495	0.7749
	2.5	0.955	0.964	0.930	0.933	1.0421	1.0372	0.9785	1.0177
30	0.2	0.946	0.948	0.936	0.934	0.0644	0.0643	0.0622	0.0631
	0.3	0.959	0.957	0.940	0.939	0.0967	0.0966	0.0933	0.0947
	0.5	0.945	0.949	0.932	0.935	0.1601	0.1600	0.1542	0.1564
	0.75	0.951	0.952	0.936	0.937	0.2385	0.2383	0.2309	0.2335
	1	0.957	0.957	0.939	0.939	0.3153	0.3151	0.3036	0.3074
	1.5	0.955	0.955	0.938	0.940	0.4712	0.4706	0.4568	0.4648
	2	0.959	0.958	0.940	0.942	0.6419	0.6405	0.6153	0.6281
	2.5	0.948	0.952	0.934	0.939	0.8384	0.8358	0.8043	0.8264
50	0.2	0.957	0.958	0.946	0.946	0.0498	0.0497	0.0485	0.0490
	0.3	0.952	0.952	0.942	0.940	0.0746	0.0746	0.0728	0.0735
	0.5	0.961	0.958	0.945	0.950	0.1234	0.1234	0.1204	0.1216
	0.75	0.951	0.953	0.935	0.936	0.1839	0.1838	0.1800	0.1819
	1	0.953	0.952	0.940	0.938	0.2434	0.2433	0.2370	0.2393
	1.5	0.947	0.949	0.939	0.941	0.3640	0.3637	0.3557	0.3596
	2	0.955	0.958	0.950	0.950	0.4930	0.4923	0.4830	0.4906
	2.5	0.950	0.952	0.939	0.940	0.6448	0.6436	0.6271	0.6392
100	0.2	0.949	0.947	0.943	0.943	0.0351	0.0350	0.0346	0.0349
	0.3	0.948	0.950	0.938	0.945	0.0526	0.0526	0.0520	0.0525
	0.5	0.946	0.948	0.941	0.944	0.0874	0.0874	0.0860	0.0868
	0.75	0.945	0.946	0.939	0.939	0.1299	0.1298	0.1281	0.1291
	1	0.950	0.951	0.949	0.949	0.1715	0.1715	0.1687	0.1701
	1.5	0.944	0.945	0.938	0.939	0.2562	0.2561	0.2523	0.2549
	2	0.946	0.946	0.943	0.944	0.3481	0.3479	0.3443	0.3476
	2.5	0.948	0.948	0.941	0.947	0.4560	0.4556	0.4492	0.4551
200	0.2	0.951	0.949	0.948	0.948	0.0248	0.0248	0.0246	0.0248
	0.3	0.954	0.952	0.945	0.946	0.0370	0.0370	0.0366	0.0369

n	θ	Coverage probability				Average length			
		Likelihood	Wald	Bootstrap-t	BCa	Likelihood	Wald	Bootstrap-t	BCa
	0.5	0.951	0.949	0.947	0.947	0.0616	0.0616	0.0612	0.0617
	0.75	0.949	0.949	0.948	0.950	0.0917	0.0917	0.0912	0.0920
	1	0.944	0.948	0.940	0.943	0.1215	0.1215	0.1203	0.1212
	1.5	0.951	0.951	0.944	0.948	0.1809	0.1808	0.1793	0.1809
	2	0.956	0.957	0.951	0.951	0.2461	0.2460	0.2438	0.2460
	2.5	0.942	0.944	0.936	0.935	0.3209	0.3207	0.3184	0.3213
500	0.2	0.945	0.944	0.940	0.943	0.0156	0.0156	0.0156	0.0157
	0.3	0.948	0.949	0.944	0.947	0.0235	0.0235	0.0233	0.0235
	0.5	0.946	0.946	0.946	0.947	0.0390	0.0390	0.0389	0.0392
	0.75	0.951	0.951	0.949	0.952	0.0580	0.0580	0.0576	0.0582
	1	0.948	0.947	0.945	0.947	0.0767	0.0767	0.0763	0.0769
	1.5	0.947	0.949	0.941	0.947	0.1144	0.1144	0.1136	0.1145
	2	0.954	0.953	0.952	0.948	0.1552	0.1552	0.1538	0.1552
	2.5	0.949	0.948	0.948	0.947	0.2019	0.2019	0.2003	0.2021

I. Coverage Probability

Table 1 displays the simulation results for the CP, whereas Figure 3 visually represents the results. The parameter value θ has a minor impact on the CP of all CIs. This indicates that the value of CP of all CIs is relatively constant, regardless of the value of the parameter θ. The sample size significantly affects the CP across all CPs. For small sample sizes (n = 10, 20, and 30), the CPs for the likelihood-based and Wald-type CIs are close to the nominal level of 0.95, whereas the bootstrap-t and BCa bootstrap CIs provide CPs that are noticeably less than 0.95. However, their performance improves with larger sample sizes, with the CPs approaching the nominal level more closely. This implies that although these approaches are sample size-dependent, they provide sufficient coverage for larger samples. Furthermore, the likelihood-based and Wald-type CIs exhibit a more rapid convergence rate in comparison to the bootstrap-t and BCa bootstrap CIs.

The likelihood-based and Wald-type CIs show greater stability in CP when both sample size and parameter value are considered, as they maintain values that are more closely approximate to the nominal level in a range of parameter values and sample sizes. On the other hand, the bootstrap-t and BCa bootstrap CIs demonstrate greater variability in CP, particularly for small sample sizes, in which case they tend to underperform.

II. Average Length

The AL usually decreases as the sample size increases, as expected in the evaluation of CIs. For example, when the sample size is n = 10 and θ = 2, the AL for the Wald-type CI is high, roughly 1.1303. Nevertheless, when the sample size is increased to n = 500, the AL for the Wald-type CI reduces significantly to approximately 0.1552.

The AL also varies with different parameter values of θ. As the value of the parameter θ increases, the AL tends to increase for all CIs. At θ = 0.2 and n = 10, the AL for the Wald-type CI is approximately 0.1132. However, at θ = 2.5, the AL increases to approximately 1.4856.

The Wald-type and likelihood-based CIs have very similar ALs for all parameter values when compared to the two other methods. This means that both Wald-type and likelihood-based CIs have a very similar interval width and coverage probability. The bootstrap-t and BCa bootstrap CIs tend to yield shorter intervals for lower values of θ, while demonstrating an increase in the

AL as the parameter value θ increases. The bootstrap-t CI generally provides the shortest interval. For example, for $\theta = 1$ and $n = 50$, the ALs are 0.2434 for likelihood-based CI, 0.2433 for Wald-type CI, 0.2370 for bootstrap-t CI, and 0.2393 for BCa bootstrap CI. These findings indicate that the bootstrap-t method yields the narrowest interval on average, but the likelihood-based method yields slightly wider intervals.

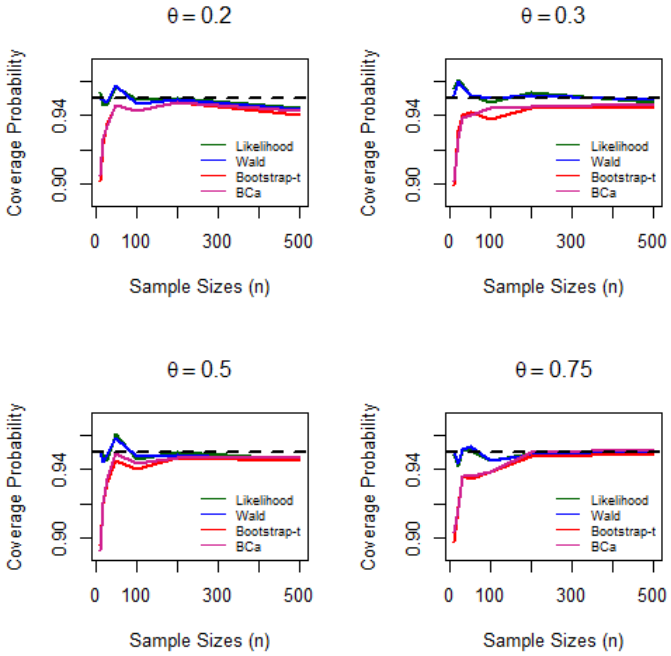


Figure 3: Plots of the CPs of the CIs for θ of the Iwueze distribution

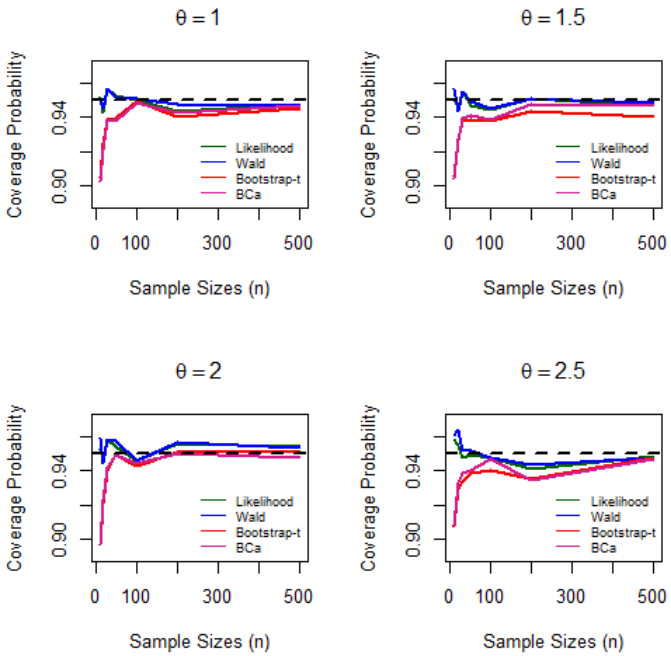


Figure 3: (Continued)

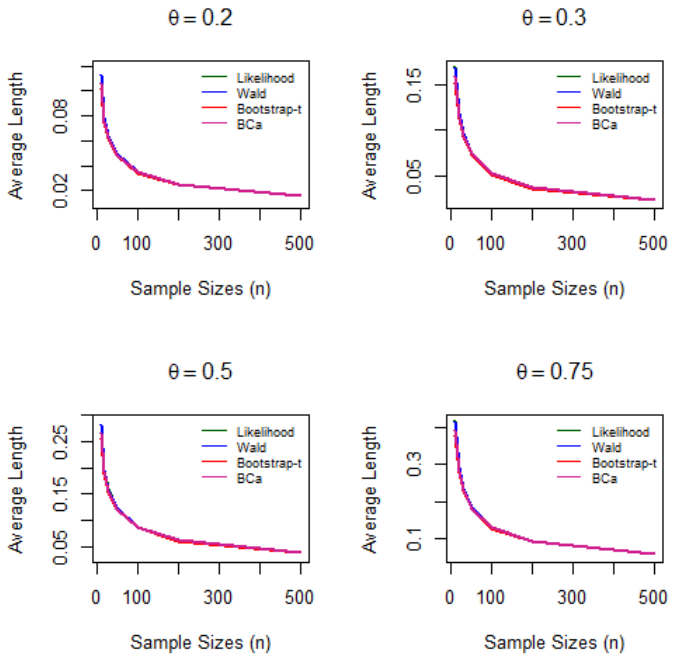


Figure 4: Plots of the ALs of the CIs for θ of the Iwueze distribution

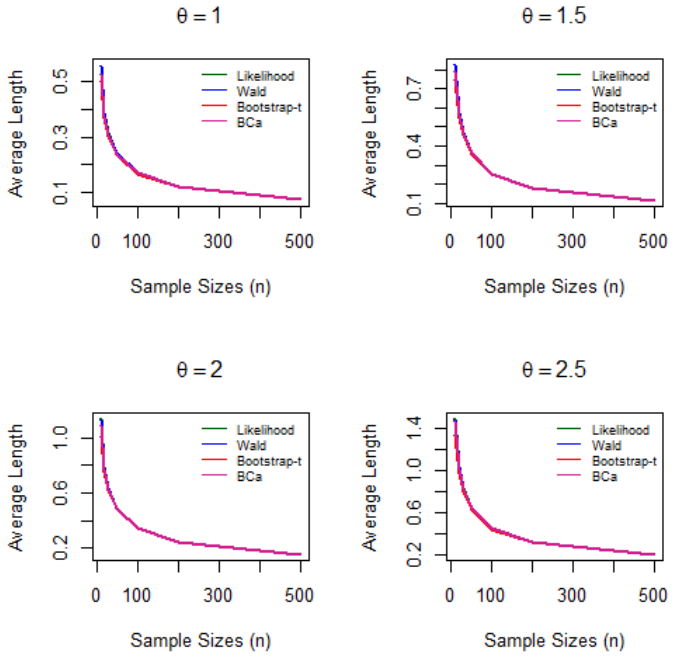


Figure 4: (Continued)

IV. Numerical Examples

We applied four CIs for the parameter of the Iwueze distribution defined in the previous section to two real-world situations. The adequacy of the Iwueze distribution's performance is being compared to that of the following alternative distributions:

- The Komal distribution [29]. Its pdf is

$$f(x;\theta)=\frac{\theta^2}{\theta^2+\theta+1}(1+\theta+x)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Adya distribution [30]. Its pdf is

$$f(x;\theta)=\frac{\theta^3}{\theta^4+2\theta^2+2}(\theta+x)^2e^{-\theta x},\ \ x>0,\theta>0.$$

- The Pranav distribution [31]. Its pdf is

$$f(x;\theta)=\frac{\theta^4}{\theta^4+6}(\theta+x^3)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Prakaamy distribution [32]. Its pdf is

$$f(x;\theta)=\frac{\theta^6}{\theta^5+120}(1+x^5)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Akshaya distribution [33]. Its pdf is

$$f(x;\theta)=\frac{\theta^4}{\theta^3+3\theta^2+6\theta+6}(1+x)^3e^{-\theta x},\ \ x>0,\theta>0.$$

- The Rani distribution [34]. Its pdf is

$$f(x;\theta)=\frac{\theta^5}{\theta^5+24}(\theta+x^4)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Rama distribution [35]. Its pdf is

$$f(x;\theta)=\frac{\theta^4}{\theta^3+6}(1+x^3)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Suja distribution [36]. Its pdf is

$$f(x;\theta)=\frac{\theta^5}{\theta^4+24}(1+x^4)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Ishita distribution [37]. Its pdf is

$$f(x;\theta)=\frac{\theta^3}{\theta^3+2}(\theta+x^2)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Sujatha distribution [9]. Its pdf is

$$f(x;\theta)=\frac{\theta^3}{\theta^2+\theta+2}(1+x+x^2)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Garima distribution [10](Shanker, 2016b). Its pdf is

$$f(x;\theta)=\frac{\theta}{\theta+2}(1+\theta+\theta x)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Aradhana distribution [38]. Its pdf is

$$f(x;\theta)=\frac{\theta^3}{\theta^2+2\theta+2}(1+x)^2e^{-\theta x},\ \ x>0,\theta>0.$$

- The Devya distribution [39]. Its pdf is

$$f(x;\theta)=\frac{\theta^5}{\theta^4+\theta^3+2\theta^2+6\theta+24}(1+x+x^2+x^3+x^4)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Amarendra distribution [40]. Its pdf is

$$f(x;\theta)=\frac{\theta^4}{\theta^3+\theta^2+2\theta+6}(1+x+x^2+x^3)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Shanker distribution [6]. Its pdf is

$$f(x;\theta)=\frac{\theta^2}{\theta^2+1}(\theta+x)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Akash distribution [7]. Its pdf is

$$f(x;\theta)=\frac{\theta^3}{\theta^2+2}(1+x^2)e^{-\theta x},\ \ x>0,\theta>0.$$

- The Lindley distribution [3]. Its pdf is

$$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x},\ \ x>0,\theta>0.$$

- The exponential distribution. Its pdf is

$$f(x;\theta)=\theta e^{-\theta x},\ \ x>0,\theta>0.$$

I. Lifetime Data about the Duration of Relief in the Analgesic Patients

The first data set consists of the lifetime data about the duration of relief (measured in minutes) experienced by 20 patients who were administered an analgesic. This data was reported by Gross and Clark [41]. The data are as follows: 1.1, 1.5, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0. Some descriptive statistics of the data set are reported in Table 2.

Table 2. *The descriptive statistics of the lifetime data about the duration of relief in the analgesic patients*

Min	Mean	Median	SD	Q1	Q3	Max
1.100	1.900	1.700	0.704	1.475	2.050	4.100

The ML technique is used to estimate all distribution parameters. For model comparison, we evaluated the log-likelihood (log L), Akaike's information criterion (AIC), and Bayesian information criterion (BIC). For this data set, estimates of the parameters, their standard errors (SE), and goodness of fit measures are given in Table 3.

Table 3. *The ML estimates, SE, AIC and BIC for the lifetime data about the duration of relief in the analgesic patients*

Distributions	Estimates (SE)	Log L	AIC	BIC
Iwueze	1.8013 (0.0312)	-25.9446	**53.8892**	**54.8849**
Komal	0.7404 (0.0146)	-31.1797	64.3593	65.3550
Adya	1.0602 (0.0146)	-28.4095	58.8189	59.8147
Pranav	1.4014 (0.0156)	-31.1933	64.3865	65.3823
Prakaamy	2.2735 (0.0261)	-30.7198	63.4396	64.4353
Akshaya	1.4417 (0.0282)	-26.5071	55.0141	56.0098
Rani	1.7195 (0.0163)	-32.6543	67.3085	68.3043
Rama	1.5213 (0.0232)	-29.8533	61.7066	62.7023
Suja	1.8954 (0.0248)	-30.2010	62.4020	63.3978
Ishita	1.0948 (0.0148)	-30.0824	62.1647	63.1604
Sujatha	1.1367 (0.0224)	-28.7488	59.4975	60.4933
Garima	0.7396 (0.0197)	-31.6058	65.2116	66.2073
Aradhana	1.1232 (0.0233)	-28.1850	58.3700	59.3658
Devya	1.8419 (0.0286)	-27.2522	56.5044	57.5001
Amarendra	1.4808 (0.0258)	-27.8193	57.6387	58.6344
Shanker	0.8039 (0.0142)	-29.8917	61.7833	62.7791
Akash	1.1569 (0.0212)	-29.7613	61.5226	62.5183
Lindley	0.8039 (0.0142)	-30.2536	62.5073	63.5030
Exponential	0.5263 (0.0139)	-32.8371	67.6742	68.6699

The AIC and BIC values in Table 3 illustrate that the Iwueze distribution provides an adequate fit to as compared with other distributions. The ML estimator for this data is 1.8013. Table 4 presents the 95% CIs for the parameter of the Iwueze distribution. The likelihood-based method yields a CI ranging from 1.4783 to 2.1720, with an interval length of 0.6937. Similarly, the

Wald-type method provides a CI of 1.4553 to 2.1472, also with a length of 0.6919, which is almost identical to the likelihood-based method in terms of range and uncertainty. In contrast, the bootstrap-t method and the BCa bootstrap method both produce notably narrower CIs.

Table 4. *The 95% CIs and lengths for the lifetime data about the duration of relief in the analgesic patients*

Methods	Confidence intervals	Lengths
Likelihood-based	(1.4783, 2.1720)	0.6937
Wald-type	(1.4553, 2.1472)	0.6919
Bootstrap-t	(1.6363, 2.0162)	0.3799
BCa Bootstrap	(1.5776, 1.9507)	0.3731

II. The Strengths of Glass Fibers

The second data set is from Smith and Naylor [42] on the strengths of 1.5 centimeter glass fibers measured at the National Physical Laboratory in England. This data set is given as follows: 0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2.00, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.50, 1.54, 1.60, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.50, 1.55, 1.61, 1.62, 1.66, 1.70, 1.77, 1.84, 0.84, 1.24, 1.30, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.70, 1.78, 1.89. Some descriptive statistics of the data set are reported in Table 5.

Table 5. *The descriptive statistics of the strengths of glass fibers*

Min	Mean	Median	SD	Q1	Q3	Max
0.550	1.507	1.590	0.3241	1.375	1.685	2.240

The ML method was utilized for estimating the parameters of the distributions. We assessed the log-likelihood (log L), Akaike's information criterion (AIC), and Bayesian information criterion (BIC) for model comparison. Table 6 provides estimates of the parameters, their standard errors (SE), and goodness of fit measures for this data set.

Table 6. *The ML estimates, SE, AIC and BIC for the strengths of glass fibers*

Distributions	Estimates (SE)	Log L	AIC	BIC
Iwueze	2.0894 (0.0137)	-68.6897	**139.3794**	**141.5225**
Komal	0.8905 (0.0070)	-84.5918	171.1836	173.3268
Adya	1.2237 (0.0064)	-77.1008	156.2016	158.3447
Pranav	1.5607 (0.0063)	-90.4814	182.9627	185.1059
Prakaamy	2.4974 (0.0100)	-93.0292	188.0583	190.2015
Akshaya	1.7091 (0.0130)	-69.5206	141.0413	143.1844
Rani	1.8802 (0.0063)	-98.1100	198.2199	200.3630
Rama	1.7313 (0.0098)	-84.8598	171.7197	173.8628
Suja	2.1133 (0.0099)	-88.7335	179.4670	181.6101
Ishita	1.2520 (0.0064)	-84.1406	170.2812	172.4243
Sujatha	1.3501 (0.0104)	-77.4048	156.8096	158.9527
Garima	0.9157 (0.0096)	-85.0308	172.0617	174.2048
Aradhana	1.3464 (0.0110)	-74.9384	151.8768	154.0199
Devya	2.1013 (0.0121)	-74.9042	151.8085	153.9516
Amarendra	1.7201 (0.0114)	-75.5186	153.0372	155.1804
Shanker	0.9563 (0.0066)	-81.1391	164.2781	166.4213
Akash	1.3554 (0.0096)	-81.8636	165.7272	167.8704

Lindley	0.9563 (0.0066)	-81.3693	164.7387	166.8818
Exponential	0.6636 (0.0070)	-88.8303	179.6606	181.8038

The AIC and BIC values, estimates of the parameters, their SEs, and measures of goodness of fit for this dataset are provided in Table 6. It shows that the Iwueze distribution fits better than other distributions. For this set of data, the ML estimator is 2.0894. Table 7 reports comparisons of 95% CIs and their lengths for parameter estimation using several methods. The likelihood-based method estimates the CI to be between 1.8691 and 2.3292, with an interval length of 0.4601. The Wald-type method yields a marginally narrower CI, ranging from 1.8597 to 2.3192, and has an interval length of 0.4595, closely aligning with the results of the likelihood-based method. In contrast, the bootstrap-t method offers a narrower CI, spanning from 2.0240 to 2.1552, with the shortest interval length of 0.1312. Similarly, the BCa bootstrap method provides an even tighter CI, ranging from 2.0288 to 2.1627, with the interval length at 0.1339.

Table 7. *The 95% CIs and lengths for the strengths of glass fibers*

Methods	Confidence intervals	Lengths
Likelihood-based	(1.8691, 2.3292)	0.4601
Wald-type	(1.8597, 2.3192)	0.4595
Bootstrap-t	(2.0240, 2.1552)	0.1312
BCa Bootstrap	(2.0288, 2.1627)	0.1339

Conclusion and Discussion

This paper proposes and evaluates four approaches for using likelihood-based, Wald-type, bootstrap-t, bias-corrected and accelerated (BCa) bootstrap methods to construct confidence intervals (CIs) for the parameter of the Iwueze distribution. This study also derived and provided the explicit formula for the Wald-type CI. The evaluation of CIs in simulation studies involves the consideration of both the coverage probability (CP) and the average length (AL) of the intervals. As the sample sizes increase, the results indicate a notable pattern where the CPs of all methods converge toward the nominal confidence level. The likelihood-based and Wald-type CIs yielded satisfactory outcomes in terms of coverage probabilities, even for the setting of small sample sizes. However, the bootstrap-t and BCa bootstrap CIs provide the CP less than the nominal confidence level, especially in small sample sizes. The practical application of all CIs was shown by applying them to medical and engineering data, producing results consistent with the simulation study's results.

The bootstrap techniques examined in this study rely on the assumption that resampled data accurately represent the underlying population. For datasets with very small sample sizes and significant skewness, the validity of the assumption that resampled data accurately represent the underlying population may be compromised. Consequently, this could impact the reliability of the CIs derived from these methods. Moreover, the computational requirements of bootstrap techniques, particularly the BCa bootstrap CI, might present challenges in situations where computational resources are limited. To facilitate the computation of bootstrap confidence intervals in the R programming language, numerous packages are accessible, with the 'boot' package [43] and the 'bootstrap' package [44] being notable examples.

Future research could explore other mixed distributions, such as the Chris-Jerry distribution [45], Hamza distribution [46], among others. The construction of CIs for the coefficient of variation and the population mean is an interesting topic that requires additional research. Additionally, there appears to be a gap in the literature regarding hypothesis testing for the parameters of the

Iwueze distribution. These topics represent valuable opportunities for future studies.

Acknowledgements

The author would like to thank the editor and the reviewers for the valuable comments and suggestions to improve this paper. This study was supported by the Research Fund of Faculty of Science and Technology, Thammasat University.

References

[1] Shamar, V., Shanker, R. and Shanker, R. (2019). On some one parameter lifetime distributions and their applications. *Annals of Biostatistics and Biomed Applications*, 3:1–6.

[2] Oguntunde, P. E., Owoloko, E. A. and Balogun, O. S. (2016). On a new weighted exponential distribution: Theory and application. *Asian Journal of Applied Sciences*, 9:1–12.

[3] Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society*, 20:102–107.

[4] Ghitany, M. E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78:493–506.

[5] Shanker, R., Fesshaye, H. and Selvaraj, S. (2015). On modeling of lifetimes data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*, 2:140–147.

[6] Shanker, R. (2015). Shanker distribution and its applications. *International Journal of Statistics and Applications*. 5:338–348.

[7] Shanker, R. (2015). Akash distribution and its applications. *International Journal of Probability and Statistics*. 4:65–75.

[8] Shanker, R. and Fesshaye, H. (2016). On modeling of lifetime data using Akash, Shanker, Lindley and exponential distributions. *Biometrics & Biostatistics International Journal*, 3:214–224.

[9] Shanker, R. (2016). Sujatha distribution and its applications. *Statistics in Transition New Series*, 17: 391–410.

[10] Shanker, R. (2016). Garima distribution and its application to model behavioral science data. *Biometrics and Biostatistics International Journal*, 4:275–281.

[11] Elechi, O., Okereke, E., Chukwudi, I., Chizoba, K. and Wale, O. (2022). Iwueze's distribution and its application. *Journal of Applied Mathematics and Physics*, 10:3783–3803.

[12] Nwry, A. W., Kareem, H. M., Ibrahim, R. B. and Mohammed, S. M. (2021). Comparison between bisection, Newton, and secant methods for determining the root of the non-linear equation using MATLAB. *Turkish Journal of Computer and Mathematics Education*, 12:1115–1122.

[13] Henningsen, A. and Toomet, O. (2011). MaxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26:443–458.

[14] RStudio Team. (2024). RStudio: Integrated Development Environment for R. Retrieved from https://www.rstudio.com/

[15] Severini, T. A. Likelihood Methods in Statistics, Oxford University Press, 2000.

[16] Srisuradetchai, P., Niyomdecha, A. and Phaphan, W. (2024). Wald intervals via profile likelihood for the mean of the inverse Gaussian distribution, *Symmetry*, 16. doi: 10.3390/sym16010093.

[17] Brent, R. P. Algorithms for Minimization without Derivatives, Prentice-Hall, 1973.

[18] Kiusalaas, J. Numerical Methods in Engineering with Python 3, Cambridge University Press, 2013.

[19] Bolker, B. M. (2023). R Development Core Team. bbmle: Tools for General Maximum Likelihood Estimation (Version 1.0.25.1) [Computer software]. Retrieved from https://CRAN.R-project.org/package=bbmle.

[20] Pawitan, Y. All Likelihood: Statistical Modelling and Inference Using Likelihood, Clarendon Press, 2001.

[21] Kummaraka, U. and Srisuradetchai, P. (2023). Interval estimation of the dependence parameter in bivariate Clayton copulas. *Emerging Science Journal*, 7:1478–1490.

[22] Brazzale, A. R., Davison, A. C. and Reid, N. Applied Asymptotics: Case Studies in Small-Sample Statistics, Cambridge University Press, 2007.

[23] Chernick, M. R. Bootstrap Methods: A Guide for Practitioners and Researchers, Wiley-Interscience, 2008.

[24] Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *The Annals of Statistics*, 16:927–953.

[25] Panichkitkosolkul, W. and Srisuradetchai, P. (2022). Bootstrap confidence intervals for the parameter of zero-truncated Poisson-Ishita distribution. *Thailand Statistician*, 20:918–927.

[26] Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82:171–185.

[27] Efron, B. and Narasimhan, B. (2020). The automatic construction of bootstrap confidence intervals. *Journal of Computational and Graphical Statistics*, 29:608–619.

[28] Grün, B. and Miljkovic, T. (2023). The automated bias-corrected and accelerated bootstrap confidence intervals for risk measures. *North American Actuarial Journal*, 27: 731–750.

[29] Shanker, R. (2023). Komal distribution with properties and application in survival analysis. *Biometrics & Biostatistics International Journal*, 12:40–44.

[30] Shanker, R., Shukla, K. K., Ranjan, A. and Shanker, R. (2021). Adya distribution with properties and application. *Biometrics & Biostatistics International Journal*, 10:81–88.

[31] Shukla, K. K. (2018). Pranav distribution with properties and its applications. *Biometrics & Biostatistics International Journal*, 7:244–254.

[32] Shukla, K. K. (2018). Prakaamy distribution with properties and applications. *Journal of Applied Quantitative Methods*, 13:30–38.

[33] Shanker, R. (2017). Akshaya distribution and its application. *American Journal of Mathematics and Statistics*, 7:51–59.

[34] Shanker, R. (2017). Rani distribution and its application. *Biometrics & Biostatistics International Journal*, 6:256–265.

[35] Shanker, R. (2017). Rama distribution and its application. *International Journal of Statistics and Applications*, 7:26–35.

[36] Shanker, R. (2017). Suja distribution and its application. *International Journal of Probability and Statistics*, 6:11–19.

[37] Shanker, R. and Shukla K. K. (2017). Ishita distribution and its applications. *Biometrics & Biostatistics International Journal*, 5:39–46.

[38] Shanker, R. (2016). Aradhana distribution and its applications. *International Journal of Statistics and Applications*, 6:23–34.

[39] Shanker, R. (2016). Devya distribution and its applications. *International Journal of Statistics and Applications*, 6:189–202.

[40] Shanker, R. (2016). Amarendra distribution and its applications. *American Journal of Mathematics and Statistics*, 6:44–56.

[41] Gross, A. J. and Clark, V. A. Survival Distributions: Reliability Applications in the Biometrical Sciences, John Wiley & Sons, 1975.

[42] Smith, R. L. and Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Journal of the Royal Statistical Society*. *Series C (Applied Statistics)*, 36:358–369.

[43] Canty, A. and Ripley, B. (2024). R Development Core Team. boot: Bootstrap Functions, (Version 1.3-30) [Computer software]. Retrieved from https://cran.r-project.org/package=boot.

[44] Kostyshak, S. (2024). R Development Core Team. bootstrap: Functions for the Book "An Introduction to the Bootstrap, (Version 2019.6) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/bootstrap.

[45] Obulezi, O. and Onyekwere, C. (2022). Chris-Jerry distribution and its applications. *Asian Journal of Probability and Statistic*s. 20(1):16–30.

[46] Aijaz, A., Jallal, M., Ain, S. Q. U., and Tripathi, R. (2020). The Hamza distribution with statistical properties and applications. *Asian Journal of Probability and Statistics*. 8(1):28–42.

OPTIMIZATION OF PREVENTIVE MAINTENANCE BY A COMPARATIVE APPROACH BASED ON EXACT RESOLUTION METHODS AND GENETIC ALGORITHMS: APPLICATION TO A PRODUCTION UNIT

Ngnassi Djami A.B [1*]; Samon J.B[2]; Nzié W[1,2]

•

[1]Department of Fundamental Sciences and Techniques of Engineer, Chemical Engineering and Mineral Industries School, University of Ngaoundere, Ngaoundere, Cameroon
[2]Department of Mechanical Engineering, National School of Agro-Industrial Sciences, University of Ngaoundere, Ngaoundere, Cameroon
ngnassbris@yahoo.fr[1*], boscosamon@yahoo.fr[2], wnzie@yahoo.fr fr[1,2]
Correspondence email : ngnassbris@yahoo.fr

Abstract

The control of the maintenance of the industrial installations, in particular of the costs due to the implementation of the preventive policies is very interesting because of the growing importance of this service in the chains of production. The objective of this paper is to minimize the preventive maintenance costs of a production unit. For this, a state of the art on the maintenance cost models according to the policy used is first made, then a synthesis of the optimization methods is made in order to deploy the exact resolution methods and the genetic algorithms. The result of this paper is the proposal of a cost model corresponding to a periodic maintenance policy with minimal repair to the failure and the optimization of the periodicities of the partial revisions of the production unit.

Keywords: Preventive maintenance, Reliability, Optimization, Cost, Genetic algorithm

1. Introduction

Today, maintenance occupies a very important place in the production chain because the failure of a system during production can have direct and indirect consequences that are extremely detrimental for the system and for other business functions. The failure of a machine can generate: delays in delivery, loss of customers, larger stocks of finished products, cash flow difficulties, etc.

Sudden breakdowns are sometimes very costly and the loss of production during corrective interventions causes a loss of profit which can affect the profits of the company. Add to this safety issues, diminished production quality and possible loss of reputation for the company, it becomes clear that such failures should not be tolerated where preventive maintenance is required.
Optimizing preventive maintenance is a process of improving their performance and efficiency (of the company). This process tries to balance the requirements of preventive maintenance (legislative, economic, technical, etc.) and the resources used to carry out their program (labour, spare parts, consumables, equipment, etc.). The goal of preventive maintenance optimization is to choose the appropriate policy for each piece of equipment and the identification of the periodicity of this policy should be carried out to achieve the objectives concerning the safety, the reliability of the equipment

and the availability of the system. When preventive maintenance optimization is effectively implemented, overall preventive maintenance costs will be reduced.

2. State of the art on maintenance cost models according to the policy used

Production and service equipment constitute an important part of the capital of the majority of industries. This equipment is generally subject to degradation with use and time. For some of these systems, such as aircraft, nuclear systems, oil and chemical facilities, it is extremely important to do everything possible to avoid failure in operation because it can be dangerous. Moreover, for continuously operating units such as oil refineries, the loss of earnings is high in the event of a stoppage. Therefore, maintenance becomes a necessity to improve reliability. The growing importance of maintenance has generated an ever-increasing interest in the development and implementation of maintenance strategies for improving system reliability, preventing failures and reducing maintenance costs.

2.1. Notions on maintenance

2.1.1. Standard definitions

According to standard NF X 60-10 (December 1994), maintenance is "all activities intended to maintain or restore an item in a state or under given operating safety conditions, to accomplish a required function. These activities are a combination of technical, administrative and managerial activities ".

Corrective maintenance is the set of actions carried out after detection of the failure and intended to return an item to a state in which it can perform a required function (NF EN 2001).
Preventive maintenance is the set of actions carried out at predetermined time intervals or according to prescribed criteria and intended to reduce the probability of failure or the degradation of the functioning of an asset (NF EN 2001).

2.1.2. Effects of maintenance on systems

Maintenance can be characterized by its effect on the state of the system after receiving a maintenance action, as follows [1, 2]:

- Perfect repair (maintenance): Any maintenance action that brings the system back to an "as good as new" state. After perfect maintenance, the system has the same failure rate as a new system. A replacement is considered perfect maintenance. Example: complete overhaul of an engine.
- Minimal repair (maintenance): Any action that brings the failure rate of the system back to what it was just before the “As bad as old” failure. Example : changing a car tyre.
- Imperfect repair (maintenance): Any action that restores the system to a state between “as good as new” and “as bad as old”. It is considered as a general case encompassing the two extreme cases, perfect repair (maintenance) and minimal repair (maintenance). Example: development of an engine.

  In Summaries and classifications of possible causes for imperfect maintenance are systematically given [3, 4, 5, 6, 7].

2.2. Concepts on reliability

The evolution The AFNOR X606500 standard defines reliability as “the ability of an entity to perform a required function, under given conditions, during a given time interval”.
It is defined by: $R(t) = P$ (E not failing during the duration $[0\ ,\ t]$ assuming that it is not failing at the moment $t = 0$).

2.2.1. Weibull model

This mathematical model covers quite a large number of lifetime distributions. It was first used in the study of material fatigue, it has been useful in the study of failure distributions of vacuum tubes, and is now in almost universal use in reliability.
Its distribution function is given by the expression 1.

$$F(t)=\begin{cases}1-e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}} & \text{si } t \succ \gamma \\ 0 & \text{si } t \le \gamma\end{cases} \quad (1)$$

β, η and γ represent respectively the shape parameter, the scale parameter and the position parameter.

Its reliability function $R(t)$ is given by relation 2.

$$R(t)=e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}} \quad (2)$$

Its density function $g(t)=\lambda(t).R(t)$ is given by relation 3 (where $\lambda(t)$ is the failure rate).

$$g(t)=\begin{cases}\frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}} & \text{si } t \succ \gamma \\ 0 & \text{si } t \le \gamma\end{cases} \text{ with } \lambda(t)=\frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} \quad (3)$$

- Whether $\gamma=0$ and $\beta=1$, $g(t)=\frac{1}{\eta}e^{-\frac{t}{\eta}}$

This is the exponential distribution, the special case of the Weibull distribution.

- If $\beta \ge 3$, the Weibull distribution approaches the normal distribution from which it can practically not be distinguished from $\beta=4$.

2.2.2. Proactive maintenance approaches

The idea is to compare the real distribution function with the theoretical one. We measure the difference, point by point between these two functions (relation 4).

$$D_{n_i}=\left|f(t_i)-F(t_i)\right| \quad (4)$$

The maximum difference thus obtained is given by relation 5.

$$D_{n.\max}=\max\left|f(t_i)-F(t_i)\right| \quad (5)$$

$F(t_i)$ and $f(t_i)$ denote respectively the theoretical distribution function and the real distribution function.

The theoretical distribution function is given by relation 6.

$$F(t_i)=1-R(t_i)=1-e^{-\left(\frac{t_i-\gamma}{\eta}\right)^{\beta}} \quad (6)$$

The real distribution function is calculated using the empirical relationships according to the size (N) of the data [8,9] (relationships 7 to 9).

- Method of median ranks ($N \le 20$):

$$f(t_i)=\frac{\sum n_i-0,3}{N+0,4} \quad (7)$$

- Mean rank formula ($20 \prec N \prec 50$):

$$f(t_i)=\frac{\sum n_i}{N+1} \quad (8)$$

- Grouping by classes with $k=\sqrt{N}$ ($N \ge 50$):

$$f(t_i)=\frac{\sum n_i}{N} \quad (9)$$

With the Kolmogorov-Smirnov table, the value of the difference $D_{n,\alpha}$ is determined at a fixed

level of significance α . There by:

- If $D_{n.\max} \succ D_{n,\alpha}$, then the hypothesis of the theoretical model is refused.
- If $D_{n.\max} \succ D_{n,\alpha}$, then the hypothesis of the theoretical model is accepted.

2.3. Maintenance policies for elementary systems

An elementary system is defined as any part that is part of a machine (screw, seal, shaft, pinion, pin, etc.) or a machine that is part of a set, such as a grinder in an infant flour production line, a turn in a mechanical production line. In this case, the reliability characteristics and any other variable of the model relate to the entire system, itself can be broken down into elementary entities.

2.3.1. Age-dependent preventive maintenance policy

According to this policy, an elementary component is replaced when it reaches the age T or failure depending on which event occurs first [7]. The average cost per unit of time is given by relation 10.

$$C(T) = \frac{Cp.R(T) + [1 - R(T)]}{\int_0^T R(t)dt} \tag{10}$$

- T : Age of preventive replacement (decision variable);
- Cp : Cost of preventive replacement;
- Cc : Failure cost including replacement cost;
- $R(t)$: Reliability function;
- $Cp.R(T) + [1 - R(T)]$: Total cost of the cycle;
- $\int_0^T R(t)dt$: Cycle length expectation.

Since then, several extensions or variants of this model have emerged [10,11,12,13,14,15].

2.3.2. Periodic preventive maintenance policy

In this policy, an item is preventively maintained at fixed time intervals kT ($k = 1,2,...$) independent of failure history, and repaired upon failure. Another basic periodic preventive maintenance policy is "periodic replacement with minimum repair to failure" where an item is replaced at predetermined times kT ($k = 1,2,...$) and failures are eliminated by minimum repairs [16]. In this class, we can also cite the block replacement policy where an element is replaced at pre-arranged times kT and on failure (generally used for multi-component systems). For this last policy, the characterized random process is a renewal process, the average cost per unit of time is given by relation 11.

$$C(T) = \frac{Cc.H(T) + Cp}{T} \tag{11}$$

- $H(T)$: Average number of replacements from 0 to T ;
- Cp : Cost of the part;
- Cc : Cost caused by the failure.

The difficulty with the previous expression lies in the determination of the renewal function $H(T)$.

With the concepts of minimal repair and especially imperfect maintenance, different extensions and variants of these two policies have been proposed [17,18,19,20,21].

2.3.3. Periodic replacement policy with minimum repair

This policy is a variant of the previous one, the difference is that following a failure, the element receives a minimal repair. Therefore, failures occur following an inhomogeneous Poisson process. The average number of failures in an interval $[0\ ;T]$ is given by relation 12.

$$H(T) = \int_0^T \lambda(t)dt \quad (12)$$

$\lambda(t)$ represents the failure occurrence rate. For a non-repairable component, it represents the failure rate. Relationship 11 then becomes relationship 13.

$$C(T) = \frac{Cc.H(T) + Cp}{T} = \frac{Cc.\int_0^T \lambda(t)dt + Cp}{T} \quad (13)$$

2.3.4. Imperfect Periodic Maintenance Policy with Minimal Repair

Under this policy, the item is not replaced periodically but just receives imperfect maintenance. As an example, we can cite an industrial machine that periodically receives partial overhauls and after a certain number of partial overhauls, the machine receives a general overhaul. This will mean that the rate of occurrence of failures will change after each preventive maintenance action, because we recall that, imperfect maintenance makes it possible to reduce the failure rate to a level between the initial failure rate (nine) and the one just before the maintenance. In this case, the effect of each maintenance on the system must be measured. The system failure rate after each maintenance will be expressed as a function of this effect and the previous failure rate. We give Gertsbakh 's model [22] where he assumes that the effect of all preventive maintenance is constant. It varies the failure rate exponentially, by an amount equal to e^{α} ($\alpha \succ 0$). The average cost per unit time is given by relation 14.

$$C(T) = \frac{Cc.H(T).\left(1 + e^{\alpha} + ... + e^{\alpha(K-1)}\right) + (K-1).Cp + Cov}{KT} \quad (14)$$

- Cc : Minimum repair cost;
- Cp : Cost of imperfect preventive maintenance (partial overhaul);
- Cov : Cost of the general overhaul;
- K : Number of partial revisions before the general revision;
- e^{α} : Degradation factor.

There are other maintenance policies for single-component systems whose synthesis is presented [23,24,25,26,27,28,29,30].

In view of this synthesis of the maintenance cost models according to the policy used, it appears that the imperfect periodic maintenance policy with minimal repair is the most appropriate for the maintenance department. It is therefore this model that will be developed in the Application part of this paper.

III. Optimization methods

There are many optimization methods. However, they can be classified into two main categories: exact resolution methods and stochastic methods. In the first category, we find all the methods that seek the minimum of a function based on the knowledge of a search direction, often given by the gradient of this function. In the case of multiple optima, they stop on the first encountered. Stochastic methods are an alternative to overcome this drawback. The three most popular stochastic methods are genetic algorithms, simulated annealing, and tabu search. They are able to find the global minimum of a function even in very difficult cases, but the computation time can be high.

3.1. Exact resolution methods

A few methods of the class of complete or exact algorithms are presented. These methods give a guarantee to find the optimal solution for an instance of finite size in a limited time and to prove its optimality [31]. We will give the desired idea of each method and describe in more detail the method of the golden ratio, which will be part of the methods deployed in this work.

3.1.1. Separation and evaluation method (Branch and Bound)

The separation and evaluation algorithm, better known by its English name Branch and Bound (B&B), is based on a tree method of finding an optimal solution by separations and evaluations, by representing the solution states by a state tree, with knots, and leaves [32].

3.1.2. Plane cutting method (Cutting-Plane)

The plane cut method was developed by [33]. It is intended to solve combinatorial optimization problems which are formulated in the form of a linear program.

3.1.3. Mathematical methods

To determine an optimum, the mathematical methods are based on the knowledge of a search direction often given by the gradient of the objective function with respect to the parameters.

3.1.4. Conjugate gradient method

The conjugate gradient method [34,35,36,37], is an improved variant of the steepest slope method, which involves following the opposite direction of the gradient. This method has the disadvantage of creating orthogonal search directions, which slows down the convergence of the algorithm. The method of Fletcher and Reeves [34] solves this problem by determining the new search direction from the gradient at the current and previous steps.

3.1.5. Quasi-Newton methods

Quasi-Newton methods consist in imitating Newton's method where the optimization of a function is obtained from successive minimizations of its second-order approximation. They do not calculate the Hessian but they use a positive definite approximation of the Hessian which can be obtained either by the expression proposed by Davidon -Fletcher-Powell, or by that proposed by Broyden - Fletcher- Goldfard -Shanno [38].

3.1.6 Golden ratio method

The golden ratio method (golden search method) is an optimization technique that seeks the extremum (minimum or maximum) of a function, in the case of a unimodal function, i.e. in which the global extremum sought is the single local extremum. If there are several local extrema, the algorithm gives a local extremum, without it being guaranteed that it is the absolute extremum. The steps of the method are as follows:

- 1st step: take the two points $c = a + (1-r)h$ and $d = a + rh$ the interval $[a, b]$, with $r = \left(\sqrt{5}-1\right)/2$ and $h = b - a$.
- 2nd step: if the values of $f(x)$ for these two points are almost equal ($f(a) \approx f(b)$) and the width of the interval is small enough ($h \approx 0$) , then stop the iteration to exit the loop and declare $x^0 = c$ or $x^0 = d$ depending on whether $f(c) \prec f(d)$ or not. If not, go to step 3.
- 3rd step: if $f(c) \prec f(d)$, take the new upper limit of the interval $b \leftarrow d$. If not, take the new lower limit of the interval $a \leftarrow c$. Then return to the 1st step.

We note the following points concerning the procedure of the method of the golden ratio:

- At each iteration, the new width of the interval is : $b - c = b - (a + (1-r)(b-a)) = rh$ or $d - a = a + rh - a = rh$, so that it becomes r times the width of the old interval $(b - a = h)$.
- The golden r ratio is fixed such that a point $C_1 = b_1 - rh_1 = b - r^2 h$ in the new interval $[c, b]$ conforms with $d = a + rh = b - (1-r)h$, that is $r^2 = 1 - r$, $r^2 + r - 1 = 0$, $r = (-1 + \sqrt{1+4})/2 = (\sqrt{5}-1)/2$.

3.2. Genetic Algorithms

3.2.1. Origin and principle

Genetic algorithms (GA) are stochastic optimization algorithms based on the mechanisms of natural selection and genetics [39,40]. The researcher Reichenberg [41], is the first scientist who introduced evolutionary algorithms by publishing his work "Evolution strategies.'' These algorithms are broadly inspired by Darwin's theory of evolution published in 1859. Next, Holland [42] proposed the first genetic algorithms to solve combinatorial optimization problems, and they were also developed by the work of David Goldberg published in 1989 [43].

The aim of the genetic algorithm is to bring up, from one generation to another, the candidates (potential solutions) most suited to solving the problem. Each generation is made up of a defined number of individuals, these form a population, and each of them represents a point in the search space. Each individual (chromosome) has information coded in the form of a chain of characters that analogically constitutes genes. Then the passage from one generation to another is carried out based on the process of evolution by the use of evolutionary operators like selection, crossing, and mutation.

Their operating principle is quite simple. From an initial population created at random, composed of a set of individuals (chromosomes), we proceed to the evaluation of their parent qualifications to highlight the best suited, as long as the least effective are rejected. Then, the most qualified individuals are chosen by privileged selection by giving them a chance to reproduce by crossing and mutating via the two operators of crossing and mutation. Then by relaunching this process several times, the optimal solution can be refined by passing from one generation to another.

3.2.2. Description of the formalism used

The convergence of genetic algorithms has been demonstrated for many problems, although optimality cannot be guaranteed. The ability of a genetic approach to find the right solution often depends on the adequacy of the coding, the evolution operators, and the measures of adaptation to the problem being addressed. The method proposed here is based on genetic algorithms [43]and evolutionary strategies [44]. It combines the principle of survival of the ablest individuals and genetic combinations for an elitist research mechanism. The genetic method produces new solutions (children) by combining existing solutions (parents) selected from the population, or by mutation. The central idea is that parent solutions will tend to produce superior child solutions in terms of adaptation so that ultimately a solution obtained is optimal.

In this study, we used a genetic method previously defined by [45] with a definition of the chromosome and the operators of selection, combination, and mutation concerned. Unlike genetic algorithms, the genetic method used is designed to minimize and not maximize. This method, like genetic algorithms, is not limited by assumptions about the objective function and research space, such as continuity or differentiability. It uses a population of points simultaneously by contrast with usual methods using only one point. Genetic operators are elitistically improving the search process to find the global optimum. There are more complicated genetic operators, but the basic operators and their various modifications can generally be applied. The choice of these operators depends on the nature of the problem and the performance requirements. The genetic algorithm that we are going to implement is as follows, where the process is applied to iteration k :

a. Data coding;
b. Generation of the initial population P_0 of N individuals;
c. Assessment of the adaptation of all individuals in the population;

d. Selection of a proportion of the best individuals (parents for the production of new individuals);
e. The Crossing of all individuals in the population P_k two by two with a probability P_m ,we will have N children noted C_k ;
f. Mutation of all individuals in the population, we will have N elements noted M_k ;
g. Choice of the most suitable individuals, i.e., those who optimize the objective function;
h. If the stop test is verified, stop, otherwise return to step a.

We will choose, as a stop test in our implementation, a finite number of iterations.

It is important to note that the stopping criterion can be several cycles of the algorithm (number of generations), the average of the adaptations of individuals, a convergence factor, etc.
An individual represents a vector of decision variable (parameters), and its adaptation is measured by the objective function. The formalism and the genetic operators are detailed below.

3.2.2.1. Data coding

The first step is to properly define and code the problem. That step associates with each point of the search space a specific data structure called a chromosome, which will characterize each individual in the population. This step is considered to be the most important step in GA because the success of these algorithms depends heavily on how individuals are coded.
There are different choices for coding a chromosome, this choice being a very important factor in the progress of the algorithm so it must be well suited to the problem being addressed:

- **Binary coding**: It is the most used coding. The chromosome is coded by a string of bits (which can take the value 0 or 1) containing all the information necessary to describe a point in space;
- **Multi-character coding**: this is often more natural. We are talking about multiple characters as opposed to bits. A chromosome is then represented by a series of numbers or characters, each representing a gene;
- **Coding in the form of a tree**: this coding in tree structure starts from a root (comprising several parts equal to the number of initial individuals), from which one or more children can be derived. The tree then builds up gradually, adding branches to each new generation.

3.2.2.2. Generation of the initial population

Each chromosome is the potential result of the optimization problem. We define a chromosome as a chain composed of genes, which are the parameters (decision variables) to find. The value of a gene is called an allele. The possible value of an allele is an integer or a real value. Each gene is created randomly, using equation 15.

$$a_j = \left(a_j\right)_l + \left(\left(a_j\right)_u - \left(a_j\right)_l\right) \times \gamma_j \tag{15}$$

Where:
– $\gamma_j \in \{0;1\}$ is chosen randomly
– $\left(a_j\right)_l , \left(a_j\right)_u$ are the minimum and maximum limits of the allele a_j .They are chosen according to the problem to be treated.

Each chromosome, called an individual in a haploid representation, can be written:

$X_i = \lfloor a_1, ..., a_j, ..., a_m \rfloor$

With:
– m is the number of genes
– $i = 1, ..., N$ and N is the size of the population (number of individuals).

All the constraints are taken into account in the initial phase of population creation. When an individual is created, if the constraints are respected, this individual is integrated into the initial

population; otherwise, it is not. At the start of the algorithm, the initial population contains individuals.

The length of the chromosome m and the size of the population N is two of the four adjustment parameters of the genetic method.

3.2.2.3. Objective function and adaptation

We evaluate the different solutions proposed to treat them according to their relevance and to see which the best is. For this, we use the objective function.

This function measures the performance of each individual. To be able to judge the quality of an individual and thus compare him to others.

3.2.2.4. Selection of the most suitable individuals

When the entire population is assessed at generation t, individuals are ranked in ascending order of objective function. Then the selection is made. Selection helps to statistically identify the best individuals in a population and eliminate the bad ones from one generation to the next. This operator also gives a chance to the bad elements because these elements can, by crossing or mutation, generate relevant descendants compared to the optimization criterion.

The first $N \times G$ individuals (the best $N \times G$) are selected to be parents. G is the third setting parameter of the genetic method. G is called the generation gap. G makes it possible to select a part of the population to provide sufficient genetic material without decreasing the speed of convergence [43]. There are different selection techniques:

- **Selection by rank**: This selection method always chooses the individuals with the best adaptation scores, without allowing chance to intervene;
- **Selection by wheel**: For each parent, the probability of being selected is proportional to their adaptation to the problem (their score by the fitness function). This selection mode can be imaged by a casino roulette wheel, on which all the chromosomes of the population are placed, the place is given to each of the chromosomes being proportional to its adaptation value. Also, the higher an individual's score, the more likely he is to be selected. We spin the wheel as many times as we want individual sons. The best will be able to be drawn several times, and the worst never;
- **Selection by tournament**: Two individuals are chosen at random, their adaptation functions are compared, and the best suited is selected;
- **Uniform selection**: We are not interested in the adaptation value of the objective function, and the selection is made in a random and uniform manner such that each individual has the same probability $P(i) = 1/N$ as all other individuals, where N is the total number of individuals in the population;
- **Elitism**: The passage from one generation to another through the crossing and mutation operators creates a great risk of losing the best chromosomes. Therefore, elitism aims to copy the best (or first - best) chromosome (s) from the current population to the new population before proceeding to the mechanisms of crossing and mutation. This technique quickly improves the solution because it prevents the loss of the most qualified chromosome when passing from one generation to another.

3.2.2.5. Crossing

The selected population is divided into $N/2$ couples formed randomly. Two parents P_1 ,and P_2 are chosen randomly from the potential parents and their genes are combined according to equation 16.

$$a_j(k) = a_j(P_1) + \left(a_j(P_2) - a_j(P_1)\right) \times \gamma_j \quad (16)$$

Where:

– γ_j is a uniform random number,
– $k = N \times G + 1,...,N$, the k *th* individual,
– $j = 1,...,m$.

The newly created individual is then evaluated. If its adaptation is better than that of the worst parent, it is integrated into the population to training the next generation. If it is not the case, we repeat the combination.

3.2.2.6. Mutation of all individuals in the population

The mutation operator is a process where a minor change in the genetic code is applied to an individual to introduce diversity and thus avoid falling into local optima. This operator is applied with a probability P_m generally lower than that of the crossing P_c . This probability must be low. Otherwise, the GA will turn into a random search.

3.2.2.7. Choosing the best solutions

This choice consists in retaining the solutions which have a lower value of the objective function, and putting them in the population P_{k+1} .

3.2.2.8. Stopping criterion

The stopping criterion is evaluated in the current population. If it is filled, the whole population has converged on the solution. Otherwise, the reproduction pattern will be repeated. The stopping criterion used in this method expresses that all individuals have converged on the same solution and assumes that evolution is no longer possible, that is to say, that no better solution can be found.

The whole strategy is elitist because only the best individuals are selected for survival from one generation to the next and can be the parents of new and better individuals. To ensure convergence of the algorithm, the parameters N and G must be adjusted with care. The size of the population N affects both the performance and efficiency of the algorithm [45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60]. The algorithm is less efficient with very small population sizes. Large population size may contain more interesting solutions and discourage premature convergence towards sub-optimal solutions, but requires more assessments per generation, which can lead to a low convergence rate. The generation gap G determines the proportion of the population that remains unchanged between two generations.

It is chosen to select individuals as severely as possible, without destroying the diversity of the population too much. The global strategy used assumes that all the individuals who make up the population, from generation to generation, satisfy all the constraints.
The best solution for the latest generation represents the solution to the problem by the defined criteria.

IV. Application

The objective of this application is to minimize the maintenance cost of a production unit whose operating time history is given in table 1.

Table 1: *Time Between Failures (TBF) History*

Operation number	Date	TBF (hours)
1	09/30/2022	4350
2	12/01/2022	2720
3	05/01/2022	9830
4	09/06/2022	2110

5	07/08/2022	1410
6	11/08/2022	1940
7	10/09/2022	2880
8	10/13/2022	2590
9	06/11/2023	2140
10	26/11/2023	1270
11	12/15/2023	2930
12	03/03/2023	5670
13	04/06/2023	3350
14	09/28/2023	6440

It is assumed that the maintenance service follows an imperfect periodic maintenance policy with minimal repair, for which the cost model given by relation 14 is valid.

4.1. Determination of the parameters of Weibull

table 2 gives the calculation of the real distribution function ($f(t)$).

Table 2: *Calculation of* $f(t)$

$TBF(hours)$	n_i	$\sum n_i$	$f(t)$
1270	1	1	0.04861111
1410	1	2	0.11805556
1940	1	3	0.1875
2110	1	4	0.25694444
214 0	1	5	0.32638889
2590	1	6	0.39583333
2720	1	7	0.46527778
2880	1	8	0.53472222
2930	1	9	0.60416667
3350	1	10	0.67361111
4350	1	11	0.74305556
5670	1	12	0.8125
6440	1	13	0.88194444
9830	1	14	0.95138889

By drawing the straight line which passes through the pair of points ($t/100$, $f(t)$) on the Weibull paper and its parallel which passes through the origin (assuming the simplifying hypothesis of the model according to which $\gamma = 0$), we obtain the following values: $\eta = 33 \times 100 hours = 3300h$ and $\beta = 3$ (see figure 1) .

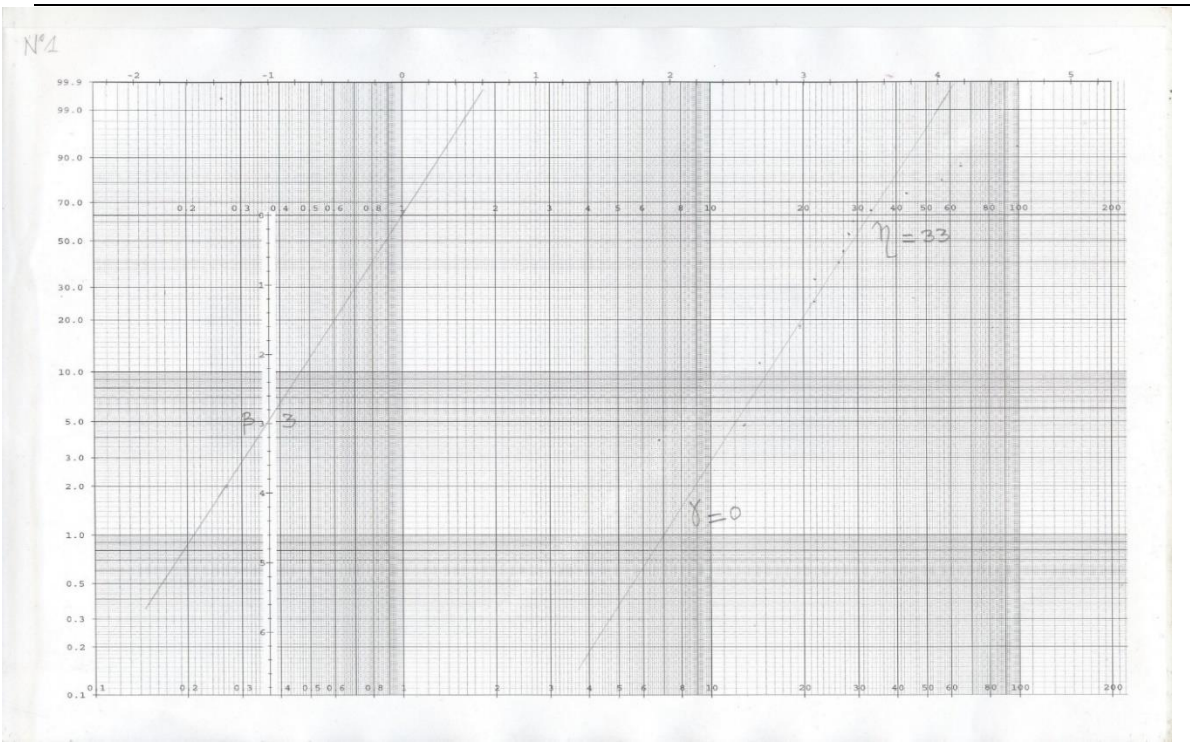


Figure 1: *Estimation of Weibull parameters*

4.2. Kolmogorov- Smirov validation test

table 3 gives the value of the deviation as a function of the values of $f(t)$ and of $F(t)$. The values of $F(t)$ being obtained by relation 6.

Table 3: *Calculation of the gap*

$TBF(hours)$	$F(t)$	$f(t)$	$D_{n.i}$
1270	0.13766232	0.04861111	0.08905121
1410	0.16686699	0.11805556	0.04881143
1940	0.29220549	0.1875	0.10470549
2110	0.33556924	0.25694444	0.0786248
2140	0.34330302	0.32638889	0.01691413
2590	0.45989253	0.39583333	0.0640592
2720	0.49306656	0.46527778	0.02778878
2880	0.53331059	0.53472222	0.00141163
2930	0.54539607	0.60416667	0.0587706
3350	0.64318313	0.67361111	0.03042798
4350	0.82405842	0.74305556	0.08100286
5670	0.94777262	0.8125	0.13527262
6440	0.97781660	0.88194444	0.09587216
9830	1	0.95138889	0.04861111

From table 3, $D_{n.\max} = 0.13527262$.

For industrial equipment we take a risk of error $\alpha = 0.05$.

By exploiting the catalog giving the Level of significance of α, we find: $D_{n,\alpha} = 0.349$.

It can be seen that $D_{n.\max} \prec D_{n,\alpha}$, consequently, the hypothesis according to which the times of failure follow a Weibull law is validated.

4.3. Development of the cost model

The following assumptions are made:
Under the following two assumptions:

- The element receives minimal repair following failure, so failures occur according to an inhomogeneous Poisson process.
- The system failure distribution follows a Weibull model with $\gamma = 0$.

According to relations 3 and 12 giving respectively the rate of failures and the average number of failures, we have relations 17 and 18.

$$H(T) = \int_0^T \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} dt = \frac{\beta}{\eta}\int_0^T \left(\frac{t}{\eta}\right)^{\beta-1} dt \tag{17}$$

$$\Rightarrow H(T) = \frac{\beta}{\eta}\int_0^T \left(\frac{t^{\beta-1}}{\eta^{\beta-1}}\right) dt = \frac{\beta}{\eta^{\beta}}\int_0^T t^{\beta-1}\, dt$$

$$\Rightarrow H(T) = \frac{\beta}{\eta^{\beta}}\left[\frac{t^{\beta}}{\beta}\right]_0^T = \frac{\beta}{\eta^{\beta}} * \left(\frac{t^{\beta}}{\beta} - 0\right)$$

From where:

$$H(T) = \frac{T^{\beta}}{\eta^{\beta}} \tag{18}$$

- By replacing 17 in 14, we obtain the relations 19 and 20.

$$C(T) = \frac{Cc.\frac{T^{\beta}}{\eta^{\beta}}.\left(1 + e^{\alpha} + \ldots + e^{\alpha(K-1)}\right) + (K-1).Cp + Cov}{KT} \tag{19}$$

$$\Rightarrow C(T) = \frac{Cc.T^{\beta}.\left(1 + e^{\alpha} + \ldots + e^{\alpha(K-1)}\right)}{KT\eta^{\beta}} + \frac{(K-1).Cp + Cov}{KT}$$

From where:

$$C(T) = \frac{Cc.T^{\beta-1}.\left(1 + e^{\alpha} + \ldots + e^{\alpha(K-1)}\right)}{K\eta^{\beta}} + \frac{(K-1).Cp + Cov}{KT} \tag{20}$$

The relation 20 represents our objective function to be minimized. To achieve this goal, two exact resolution methods will be used (the simple derivation technique and the golden ratio method), in order to compare their results with those of the genetic algorithm.

4.4. Optimization using simple derivation

The objective function $C(T)$ that we want to minimize is differentiable and in this case, it suffices to determine T_{opt} the solution of the equation $\frac{\partial C(T)}{\partial T} = 0$ which will lead us to the minimum cost ($C_{\min}$) of $C(T)$ by taking into consideration the following conditions:

$Cc \succ 0$; $Cp \succ 0$; $Cop \succ 0$; $T \succ 0$; $K \succ 0$; $\beta \succ 1$; $\alpha \succ 0$ and $\frac{\partial}{\partial T}\left(\frac{\partial C(T)}{\partial T}\right) \geq 0$

By performing a simple derivation of the $C(T)$ previous expression, we obtain relations 21 and 22.

$$\frac{\partial C(T)}{\partial T} = \frac{Cc(\beta-1).T^{\beta-2}.\left(1 + e^{\alpha} + \ldots + e^{\alpha(K-1)}\right)}{KT\eta^{\beta}} - \frac{(K-1).Cp + Cov}{KT^2} \tag{21}$$

$$\Rightarrow \frac{\partial C(T)}{\partial T} = \frac{Cc(\beta-1).T^{\beta}.\left(1 + e^{\alpha} + \ldots + e^{\alpha(K-1)}\right) - \eta^{\beta}\left((K-1).Cp + Cov\right)}{\eta^{\beta}T^2} \tag{22}$$

We deduce the expression for T_{opt} for that $\frac{\partial C(T)}{\partial T} = 0$ by relation 23.

According to relation 22, we have:

$$\frac{Cc(\beta-1).T^{\beta}.\left(1+e^{\alpha}+...+e^{\alpha(K-1)}\right)-\eta^{\beta}\left((K-1).Cp+Cov\right)}{\eta^{\beta}T^{2}}=0$$

$$\Rightarrow Cc(\beta-1).T^{\beta}.\left(1+e^{\alpha}+...+e^{\alpha(K-1)}\right)-\eta^{\beta}\left((K-1).Cp+Cov\right)=0$$

$$\Rightarrow T_{opt}=\sqrt[\beta]{\frac{\eta^{\beta}\left((K-1).Cp+Cov\right)}{Cc(\beta-1)\left(1+e^{\alpha}+...+e^{\alpha(K-1)}\right)}} \quad (23)$$

We have established a program on Matlab to calculate the value of T_{opt} in hours and days and the minimum cost (**Algorithm 1**).

The maintenance costs are proposed as well as the number of partial overhauls K and the maintenance efficiency factor α (table 4).

Table 4: *Data for simulation on Matlab*

Cc	Cp	Cov	K	α	β	η
170000	900000	8000000	8	0.9	3	3300

```
%*********************************************************************
%   Algorithm 1 Calculation code for simple derivation
%*********************************************************************
%          Developed by: Dr NGNASSI DJAMI Aslain Brisco
%                        Teacher-Researcher
%*********************************************************************
%                          Main program
%*********************************************************************
clear all
close all
clcc
%*********************************************************************
%                Inserting data (Cc, Cov, Cp, K, beta, eta, alfa)
%*********************************************************************
A0=1;
for j= 1: K -1
A0=A0+exp(j*alpha);
end
A1=( eta^beta )/(beta-1);
A2=(K- 1)* Cp+Cov ;
A3=Cc*A0;
T =[ 0:6000];
C=(((T.^(beta-1 ))* Cc*A0)/(K*( eta^beta )))+(A2./(K*T));
TT=(A1*(A2/A3 ))^ (1/beta)
T1=TT/24
Cmin = (((TT. ^(beta-1)) *Cc*A0)/(K*(eta^beta))) +(A2. /(K*TT))
%C(Toptimum)
plot (T, C);
%*********************************************************************
```

By executing the proposed code, we obtain the results presented table 5.

Table 5: *T_{opt} and $C_{\min}$ obtained by simple derivation*

System	T_{opt} in hours	T_{opt} in days	$C_{\min}(USD)$
Unit production	1181.3	49.2192	2269.8

figure 2 graphically shows the evolution of cost over time.

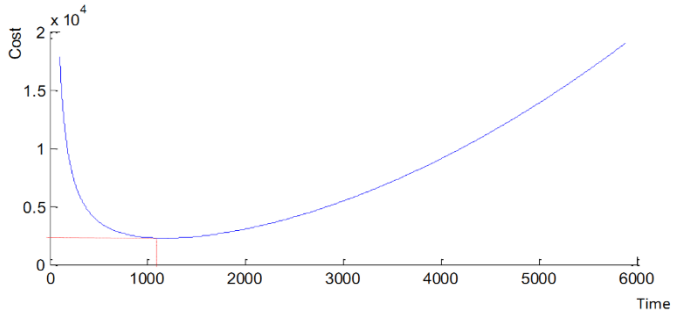


Figure 2: *Cost evolution over time*

The curve presented in figure 2 illustrates the result of the program. The cost begins to decrease until it reaches its minimum value of T_{opt} , then it increases over time.

4.5. Optimization using the golden ratio method

The program proposed in Matlab for the golden ratio method is given by **Algorithm 2**.

```
%*************************************************************************
%  Algorithm 2 Calculation code for Golden Search Method
%*************************************************************************
%            Developed by: Dr NGNASSI DJAMI Aslain Brisco
%                          Teacher-Researcher
%*************************************************************************
%                          Main program
%*************************************************************************
clear all
close all
clcc
%*************************************************************************
% %%                        Input
%*************************************************************************
fx =(((x.^(beta-1 ))* Cc*A0)/(K*( eta^beta )))+(A2./(K*x));
max = 100;
es = 10^-5;
r = (5^.5-1)/2;
**************************************************************************
% %%          Determine the Interval for the Initial Guess
%*************************************************************************
x =[ 0:6000];
f = subs( fx,x );
a=100;
b=6000;
%*************************************************************************
%%%%%                  Perform Golden Search
%*************** ***********************************************************
x1 = a;
```

```
xu = b;
iter = 1;
d = r*( xu -xl);
x1 = xl+d ;
x2 = xu- d;
f1 = subs( fx,x 1);
f2 = subs( fx,x 2);
if f1<f2
xopt = x1;
else
xopt = x2;
end
while ( 1)
d = r*d;
if f1<f2
x1 = x2;
x2 = x1;
x1 = xl+d ;
f2 = f1;
f1 = subs( fx,x 1);
else
xu = x1;
x1 = x2;
x2 = xu- d;
f1 = f2;
f2 = subs( fx,x 2);
end
iter = iter+1;
if f1<f2
xopt = x1;
else
xopt = x2;
end
if xopt ~=0
ea = (1 - r)* abs(( xu -xl)/ xopt )*100;
end
if ea <=es|| iter >= maxit, break
end
end
Gold = xopt
%**************************************************************************
```

By implementing the proposed program, we obtain the results of table 6.

Table 6: *T_{opt} and $C_{\min}$ obtained by golden ratio method*

System	T_{opt} in hours	T_{opt} in days	$C_{\min}$ (*USD*)
Unit production	1181.3	49.2192	2269.8

The results provided by the golden ratio method are identical to those of the simple derivation, which proves the effectiveness of this method in finding the extremum of a unimodal function.

4.6 Optimization using genetic algorithm

The program proposed in Matlab for the genetic algorithm is given by **Algorithm 3**.

```
%*************************************************************************************************
%                    Algorithm 3 Calculation code for Genetic Algorithm
*************************************************************************************************
%                    Developed by: Dr NGNASSI DJAMI Aslain Brisco
%                                  Teacher-Researcher
%*************************************************************************************************
%                                  Main program
%*************************************************************************************************
clear all
close all
clcc
%*************************************************************************************************
%                                  Parameter initialization
%*************************************************************************************************
Np= 60;                                % Population size
Pc = 0.5;                              % Probability of crossing
Pm=0.01;                               % Probability of mutation
Kmax = 100;                            % Maximum number of iterations
lb = [100 ];                           % lower bound
ub = [6000 ];                          % upper bound
x0 = [1010 ];                          % Approximate value
Fitness_F=@(x)(((x(1 ).^ (beta-1))*cc*A0)/(k*(nu^beta)))+(A2./(k*x(1)) );
options= gaoptimset ('populationsize' ,60, 'generations' ,100, 'MutationFcn' ,{@
mutationadaptfeasible ,0.01}, 'crossoverfraction' ,0.5, 'initialpopulation' ,[1010], 'PopInitRange'
,[lb;ub] , 'plotfcns' ,@gaplotbestindiv)
[ xo_ga fo_ga ]= ga ( Fitness_F ,1,options);
Cmin = fo_ga
Tmin_Hours_ga = xo_ga
Tmin_days_ga = xo_ga /24
%*************************************************************************************************
```

By running the proposed program after fifteen iterations, the results given in table 7 are obtained.

Table 7: T_{opt} *and* $C_{\min}$ *obtained by genetic algorithm*

Execution	$T_{opt-Hours}(x_0)$	$T_{opt-days}(x_0/24)$	$C_{\min}(USD)$
1	1181.3	49.2192	2269.8
2	1181.3	49.2192	2269.8
3	1181.3	49.2192	2269.8
4	1256.4	52.3503	2278.6
5	1181.3	49.2192	2269.8
6	1188.9	49.5364	2269.9
7	1181.3	49.2192	2269.8
8	1181.3	49.2192	2269.8
9	1176.1	49.0061	2269.9
10	1181.3	49.2192	2269.8
11	1159.9	48.3273	2270.6
12	1148.8	47.8665	2271.6
13	1181.3	49.2192	2269.8
14	1181.2	49.2185	2269.8
15	1181.3	49.2190	2269.8

According to table 7, we observe a uniqueness of solution for the colored iterations. On the other hand, for the other iterations, we rather observe a diversity of solutions. However, we notice that the colored solutions correspond exactly to the optimal solution of our problem. Indeed, the minimum value of the cost and the durations of exploitations of the production unit are the same as those of the two preceding methods, therefore the method of the genetic algorithm is convergent.

4.7. Interpretation of results

The results obtained in this application show that the production unit will have to undergo a partial overhaul after 49 days of operation, and after five partial overhauls, i.e. more than eight months of operation, the production unit will receive a general revision.

5. Conclusion

Having reached the end of writing this paper, the objective of which was to minimize the preventive maintenance costs of a production unit, we first developed the cost model corresponding to an imperfect periodic maintenance policy with minimal repair, then we deployed two exact resolution methods (the simple derivation technique and the golden ratio method) and a stochastic method (the genetic algorithm), each time proposing a code on Matlab. It turns out that by implementing the different methods, a uniformity of the optimal solution is obtained, which well justifies the convergence of the genetic algorithm. Furthermore, thanks to the proposed Matlab code, we were able to determine the periodicity at which the production unit will undergo a general overhaul.

References

[1] Sheu SH, Lin Y, Liao G. Optimum policies for a system with general imperfect maintenance. Reliability Engineering and System Safety 2006: 91: 362-369.

[2] Pham H, Wang H. Imperfect maintenance. European Journal of Operational Research 1996: 94: 425-438.

[3] Thomas LC A survey of maintenance and replacement models of multi-item systems. Reliability Engineering 1986: 16:297-309.

[4] Valdez-Flores C, Feldman RM A survey of preventive maintenance models for hastically deteriorating single-unit systems. Naval Research Logistics 1989: 36: 419-446.

[5] Cho ID, Parlar M. A survey of maintenance models for multi-unit systems. European Journal of Operational Research 1991: 51: 1-23.

[6] Van Der Duyn Schouten F. Maintenance policies for multicomponent systems. In: Ozekici , S. (Ed.), Reliability and maintenance of complex systems: And Overview. Reliability and Maintenance of Complex System 1996: 154: 117–136.

[7] Dekker R, Wildman RE, Van Der Duyn Schouten FA A review of multicomponent maintenance models with economic dependence. Mathematical Methods of Operational Research 1997: 45: 411-435.

[8] Aupied J. Experience feedback applied to the operational safety of equipment in operation. Editions Eyrolles 1994.

[9] Thomas M. Reliability, predictive maintenance and machine vibration. University of Quebec Press 2012.

[10] Tahara A, Nishida T. Optimal replacement policy for minimal repair model. Journal of Operations Research Society of Japan 1975: 18: 113-124.

[11] Nakagawa T. Optimal policy of continuous and discrete replacement with minimal repair at failure. Naval Research Logistics Quarterly 1984:31:543-550.

[12] Sheu S, Kuo C, Nakagawa T. Extended optimal age replacement policy with minimal repair. RAIRO: Operational Research 1993:27:337-351.

[13] Sheu S, Griffith W. S, Nakagawa T. Extended optimal replacement model with random minimal repair costs. European Journal of Operational Research 1995:83:636-649.

[14] Block HW, Langberg NA, Savits TH Repair replacement policies. Journal of Applied Probability 1993: 30: 194-206.

[15] Wang H, Pham H. Some maintenance models and availability with imperfect maintenance in production systems. Annals of Operations Research 1999: 91: 305-318.

[16] Barlow RE, Hunter LC Optimum preventive maintenance policies. Operations Research 1960: 8: 90-100.

[17] Liu X, Makis V, Jardine AKS A replacement model with overhauls and repairs. Naval Research Logistics 1995: 42:1063-1079.

[18] Berg M, Epstein B. A modified block replacement policy. Naval Research Logistics 1976: 23: 15-24.

[19] Tango T. Extended block replacement policy with used items. Journal of Applied Probability 1978: 15: 560-572.

[20] Nakagawa T. A summary of periodic replacement with minimal repair at failure. Journal of Operations Research Society of Japan 1981a: 24:213-228.

[21] Nakagawa T. Modified periodic replacement with minimal repair at failure. IEEE Transactions on Reliability 1981b: R-30 (2): 165-168.

[22] Gertsbakh I. Reliability Theory with applications to preventive maintenance. Springer, Berlin 2002:34:1111-1114.

[23] Zheng X, Fard N. A maintenance policy for repairable systems based on opportunistic failure rate tolerance. IEEE Transactions on Reliability 1991: 40: 237-244.

[24] Jayabalan V, Chaudhuri D. Replacement policies: a near optimal algorithm. IIE Transactions 1995:27:784-788.

[25] Nguyen DG, Murthy DNP Optimal repair limit replacement policies with imperfect repair. Journal of Operational Research Society 1981: 32: 409-416.

[26] Kijima M, Nakagawa T. Replacement policies of a shock model with imperfect preventive maintenance. European Journal of Operations Research 1992:57: 100-110.

[27] Hastings NA.The repair limit method. Operational Research Quarterly 1969: 20: 337-349.

[28] Wang H, Pham H. Optimal maintenance policies for several imperfect maintenance models. International Journal of Systems Science 1996: 27: 543-549.

[29] Nakagawa T, Osaki S. The optimum repair limit replacement policies. Operational Research Quarterly 1974: 25: 311-317.

[30] Dohi T, Matsushima N, Kaio N, Osaki S. Nonparametric repair-limit replacement policies with imperfect repair. European Journal of Operational Research 1997: 96 (2): 260–273.

[31] Puchinger J, Raidl. GR Combining metaheuristics and exact algorithms in combinatorial optimization: A survey and classification. In proceedings of the first international work- coreference on the interplay between natural and artificial computation 2005:41–53.

[32] Land A. H, Doig AG An automatic method for solving discrete programming problems. Econometrica 1960:28(3):497–520.

[33] Schrijver A. Theory of linear and integer programming. Wiley and Sons 1986.

[34] Fletcher R, Reeves CM Function minimization by conjugal gradients. Computer Journal 1964:7: 148-154.

[35] Fletcher R. Practical Methods of Optimization. John Wiley & Sons 1987.

[36] Press WH Numerical Recipes in C: The art of Scientific Computing. Cambridge University Press 1992.

[37] Culioli JC Introduction to optimization. Ellipsis 1994.

[38] Minoux M. Mathematical programming: Volume 1 Theory and algorithms. Ed. Dunod 1983.

[39] Tabeb M. Parallelization of a genetic algorithm for the single-machine scheduling problem with sequence-dependent setup times. University of Quebec at Chicoutimi 2008.

[40] Benhaddad H, Belabbas M. Parallel genetic algorithm for flow shop scheduling. Master's thesis, University of Msila 2022.

[41] Reichenberg I. Cybernetic solution path of an experimental problem. Library Translation No.1122, Royal Aircraft Establishment, Famborough , UK 1965.

[42] Holland JH Adaptation in natural and artificial systems. University of Michigan press 1975.

[43] Goldberg D. Genetic algorithms in search, optimization and machine learning. Addison Wesley 1989.

[44] Schewefel HP Numerical Optimization of computer models. Wiley Publishing 1981.

[45] Bicking F, Fonteix C, Corriou JP, Marc I. Global optimization by artificial life: a new technique using genetic population evolution. RAIRO-Operations Research 1994: 28 (1): 23-36.

[46] Digalakis JG, Margaritis KG. On benchmarking functions for genetic algorithms. International Journal of Computer Mathematics 2001:77(4): 481-506.

[47] Jason G D, Konstantinos G M. An experimental study of benchmarking functions for genetic algorithms. International Journal of Computer Mathematics 2002: 79(4): 403-416.

[48] Koumousis VK, Katsaras CP. A saw-tooth genetic algorithm combining the effects of variable population size and reinitialization to enhance performance. in IEEE Transactions on Evolutionary Computation 2006: 10 (1): 19-28.

[49] Vedat T, Ayse T D. An improved genetic algorithm with initial population strategy and self-adaptive member grouping. Computers & Structures 2008: 86: 1204-1218.

[50] Wei C, Chi W, Yajun W. Scalable influence maximization for prevalent viral marketing in large-scale social networks. Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD 2010: 10: 1029-1038.

[51] Nasr A, Elmekkawy TY. Robust and stable flexible job shop scheduling with random machine breakdowns using a hybrid genetic algorithm. Int. J. Production Economics 2011:132:279-291.

[52] Hongbin D, Tao L, Rui D, Jing S. A novel hybrid genetic algorithm with granular information for feature selection and optimization. Applied Soft Computing 2018: 65:33-46.

[53] Ahmad H, Khalid A, Esra'a A, Eman A, Awni H, Surya Prasath VB. Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach. Information 2019:10 (390):1-36.

[54] Ailiang Q, Dong Z, Fanhua Y, Ali AH, Zongda W, Zhennao C, Fayadh A, Romany FM, Huiling C, Mayun C. Directional mutation and crossover boosted ant colony optimization with application to COVID-19 X-ray image segmentation. Computers in Biology and Medicine 2022:148.

[55] Sethembiso Nonjabulo L, Akshay Kumar S . Effects of Particle Swarm Optimization and Genetic Algorithm Control Parameters on Overcurrent Relay Selectivity and Speed. in IEEE Access 2022: 10: 4550 – 4567.

[56] Junfeng Z, Yanhui Z, Yubo Z, Wen-Long S, Zhile Y, Wei F. Parameters identification of photovoltaic models using a differential evolution algorithm based on elite and obsolete dynamic learning. Applied Energy 2022: 314.

[57] Guo X, Wei T, Wang, Liu S, Qin S, Qi L. Multiobjective u-shaped disassembly line balancing problem considering human fatigue index and an efficient solution. in IEEE Transactions on Computational Social Systems 2023: 10 (4): 2061-2073.

[58] Agushaka JO, Ezugwu AE, Abualigah L et al. Efficient initialization methods for population-based metaheuristic algorithms: a comparative study. Arch Computat Methods Eng 2023: 30: 1727–1787.

[59] Asha A, Rajesh A, Poonguzhali I, Shabana U, Salem A. Optimized RNN-based performance prediction of IoT and WSN-oriented smart city application using improved honey badger algorithm. Measurement 2023: 210.

[60] Yong W, Zhen L, Gai-Ge W. Improved differential evolution using two-stage mutation strategy for multimodal multi-objective optimization. Swarm and Evolutionary Computation 2023:78.

BAYES ESTIMATOR OF PARAMETERS OF BINOMIAL TYPE EXPONENTIAL CLASS SRGM USING GAMMA PRIORS

[1]Rajesh Singh, [2]Kailash R. Kale and [3]Pritee Singh

•

[1]R. T. M. Nagpur University, Nagpur-440033.
[2]G. N. A. ACS College, Barshitakli, Dist-Akola.
[3]Institute of Science, Nagpur.
rsinghamt@hotmail.com
kailashkale10@gmail.com
priteesingh25@gmail.com

Abstract

The Reliability is one of the key characteristics of software that operates flawlessly and in accordance with needs of users. The assessment of Reliability is very important but it is complicated. The one-parameter exponential class failure intensity function is used in this article to quantify the model and assess the software Reliability. The scale parameter and the number of existing total failures are the model's parameters. Using the Bayesian approach, the estimators of parameters are obtained under the assumption that gamma priors are suitable to provide prior information of the parameters. Using risk efficiencies computed under squared error loss, the performance of proposed estimators is studied with their corresponding maximum likelihood estimators. The suggested Bayes estimators are found to outperform over the equivalent maximum likelihood estimators.

Keywords: Binomial process, gamma prior, maximum likelihood estimator (MLE), Exponential class, software reliability growth model (SRGM), confluent hyper-geometric function.

I. Introduction

The modern computers are widely and extensively used worldwide for solving the majority of complex problems pertaining to a variety of fields due to their ability to perform intricate and time-consuming tasks quickly, accurately, and with effective global communication. When completing all of such tasks, very sophisticated computers are used, and they are guided by a series of input instructions, known as a program or Software. Since, Software are necessary components of every computer system, its performance is crucial and has significant role.

Software are developed by human and due to its complexity and size, faults are more likely to occur. As a result, it gets the user's acceptance or rejection. For acceptance of any Software, its reliability is probably most important feature. Reliable software possesses high-quality and meets the needs of users or industries or government organizations. Software must be of an acceptable quality, which is closely connected to reliability qualities, to please the consumers. Since the 1950s, the area of software reliability is being studied by researchers, and several significant results have

been produced. A variety of modern Statistical methods may be applied to measure the Software reliability. One of the strategy uses a zero-one method, where a flawed Software has a reliability of zero and a faultless Software has a reliability of one. Another strategy focuses on testing of Software, where software reliability is defined as the proportion of times a software executes an intended function as predicted. Thus, the measurement of Software reliability may be done using this method.

The usage of Software Reliability Growth Models (SRGMs) in the evaluation of software reliability is highly beneficial. The operational profile, accessibility of limited failures and irrational assumptions provide significant difficulties for SRGMs in practice. When estimating the parameters of software reliability model, various fundamental techniques are used which are maximum likelihood method, least squares estimation, Bayesian estimation, the EM method, etc., [5], [6], [7] and [12] have presented the calculation of factors, such as failure rate and total number of failures incorporated in SRGMs.

Software Engineering is the process of developing a software that can balance the dependability, delivery time, and price of the developed software. Software reliability modelling, as described by [8] is another way to represent software reliability using a mathematical function of involve factors, such as fault introduction, fault removal, the operating environment, etc. Software reliability is evaluated mathematically and objectively using the body of Statistics and Probability theory.

The Binomial type models, as categorized by [10] and [11] have been taken into consideration for study in this paper. Here, an attempt is made to derive the Bayes estimators for the parameters of the Binomial type exponential class. The prior distribution for the total number of initial software failures and scale parameters have been taken as the gamma prior distribution assuming that the experimenter is having prior information about both the parameters c.f. [14]- [20].

II. Model Characterization

Based on the following assumptions [12] have modelled the process of software failure using Binomial type models.

- The defect that resulted in a software failure will always be immediately fixed.
- There are θ_0 intrinsic faults in the programme.
- The hazard rates $Z(t)$ for all faults are same.

According to [12, 19], if $f(t)$ is the class of the SRGM and $\lambda(t) = \theta_0 f(t)$ is failure intensity. Considering binomial process and solving the differential equations using boundary conditions $P_0(0) = 1$ and $P_n(0) = 0 \; \forall \; n = 0,1,2, \ldots \theta_0$, the following result is obtained.

$$P[M(t) = n] = \binom{\theta_0}{n}\{1 - exp[-\theta_1 t]\}^n \{exp[-\theta_1 t]\}^{\theta_0 - n} \quad ,n = 0,1,2, \ldots, \theta_0.$$

The $P[M(t) = n]$ gives the probability that $M(t) = n$ number of failures encountered at time t has a binomial distribution i.e. Binomial type of the software reliability model [12].

The Binomial Type Exponential Class model has failure intensity

$$\lambda(t) = \theta_0 \theta_1 e^{-\theta_1 t} \quad , t > 0, \theta_0 > 0 \; and \; \theta_1 > 0 \tag{1}$$

where θ_0, failure rate (θ_1) are the parameters of the model and t be the execution time. This model exhibits failure intensity similar to [4] and [13]. The function of the mean failures is given by

$$\mu(t) = \theta_0[1 - e^{-\theta_1 t}] \quad , t > 0, \theta_0 > 0 \; and \; \theta_1 > 0 \tag{2}$$

The expected number of failures at time t is represented by the Binomial distribution with the mean failure function.

$$\mu(t) = \theta_0\{1 - exp[-\theta_1 t_e]\}$$

and variance of $M(t)$ is

$$var[M(t)] = \theta_0\{1 - exp[-\theta_1 t]\}\{exp[-\theta_1 t]\}.$$

Let m_e be the number of failures that occurred upto execution time t_e , then the likelihood function of Binomial type exponential class model is

$$L(\theta_0, \theta_1|\underline{t}) = e^{-\theta_1 t_e(\theta_0 - m_e)}\theta_1^{m_e} e^{-\theta_1 T}\theta_0^{\underline{m_e}} \quad , t_e > 0, \theta_1 > 0, \theta_0 > 0 \text{ and } t_0 = 0 \qquad (3)$$

where

$$T = \sum_{i=1}^{m_e} t_i$$

and

$\theta_0^{\underline{m_e}}$ are falling factorials (cf. [1], [2] and [3]).

The MLEs of θ_0 and θ_1 are obtained by applying standard method of obtaining MLE from equation (3) which comes out to be

$$\sum_{i=1}^{m_e}\left(\hat{\theta}_{m0} - i + 1\right)^{-1} = \hat{\theta}_{m1} t_e \qquad (4)$$

and

$$\hat{\theta}_{m1} = m_e\left[t_e\left(\hat{\theta}_{m0} - m_e\right) + T\right]^{-1} \qquad (5)$$

The solution for $\hat{\theta}_{m0}$ and $\hat{\theta}_{m1}$ can be obtained by solving (4) and (5) using any standard iterative method.

III. Priors for Model Parameters

If the software professional could somehow predict or guess the information about the total number of failures present in the software and the value of scale parameter θ_1. Let's thus assume that gamma priors are seen to be appropriate for both θ_0 and θ_1 then it would be appropriate to use an informative prior for θ_0 and θ_1. The time-to-first failure distribution for a system with standby exponentially dispersed backups may naturally exhibit the gamma likelihood. Additionally, in practice, the gamma distribution may appear whenever items were tested, and whenever a part of an item or an entire item fails is replaced by an identical one having an exponential failure time distribution with parameter, the total amount of time on test could subsequently follow a gamma probability function. The gamma distribution is based on the fact that the total of i.i.d. random failure times after exponentials with parameters is distributed as gamma [9] and [24]. When an item may fail partially, or when a certain number of partial failures occur before an item fails (such as with redundant systems), the gamma distribution can be employed. Modelling the time to failure or failure rates for products with infant mortality may be done using the continuous Gamma probability.

Many studies have shown that the gamma distribution is feasible for failure rate and sufficiently adaptable for real-world hardware reliability applications in life testing. [7] created a Bayesian SRGM under the gamma prior assumption for the parameter of exponentially distributed periods between model failures. The Bayesian software reliability growth models established by [5] and [6] take into account the gamma prior distribution.

Thus, the gamma prior distributions may be used as informative priors in the current investigation for the parameters θ_0, and θ_1. Therefore,

$$g(\theta_0)\alpha\begin{cases}\theta_0^{\alpha-1}e^{-\beta\theta_0} & , 0 < \theta_0 < \infty \\ 0 & , \ Otherwise\end{cases}$$

and

$$g(\theta_1)\alpha\begin{cases}\theta_1^{\eta-1}e^{-\nu\theta_1} & , 0 < \theta_1 < \infty \\ 0 & , \ Otherwise\end{cases}$$

where α, β, η, and ν are hyper parameters of considered priors for θ_0 and θ_1 respectively. The hyper parameters η and α are shape parameters and ν and β are scale parameters of the prior distributions. The flexibility in choices of α, β, η, and ν allows the researcher to select the prior model for parameters that best expresses the current state of knowledge about the number of failures and failure rate.

Hence, the joint prior distribution of both parameters θ_0 and θ_1 is given as

$$g(\theta_0, \theta_1)\alpha\begin{cases}\theta_0^{\alpha-1}\theta_1^{\eta-1}e^{-\beta\theta_0}e^{-\nu\theta_1} & , 0 < \theta_0, \theta_1 < \infty \\ 0 & , \ Otherwise\end{cases} \qquad (6)$$

IV. Joint Posterior and Marginal Posterior Distributions

Assuming the total execution time is t_e, during this time m_e failures are experienced at times $t_i, i = 1,2,\dots,m_e$, θ_0 be the number of failures present in the software and θ_1 be the failure rate then, combining likelihood function (3) with joint prior given by (6), the joint posterior of θ_0 and θ_1 given $t_i, i = 1,2,\dots,m_e (= \underline{t})$ is

$$\pi(\theta_0,\theta_1|\underline{t}) \propto e^{-\theta_1[t_e(\theta_0-m_e)+(T+\nu)]}\theta_1^{m_e+\eta-1}{\theta_0}^{\alpha-1}e^{-\beta\theta_0}\theta_0^{\underline{m_e}} \quad ,\mathrm{m_e} < \theta_0 < \infty, 0 < \theta_1 < \infty \tag{7}$$

The constant of proportionality (normalizing constant) of above equation is

$$D = \int_{m_e}^{\infty}\int_0^{\infty} e^{-\theta_1[t_e(\theta_0-m_e)+(T+\nu)]}\theta_1^{m_e+\eta-1}{\theta_0}^{\alpha-1}e^{-\beta\theta_0}\theta_0^{\underline{m_e}}\, d\theta_0\, d\theta_1 \quad , m_e < \theta_0 < \infty, 0 < \theta_1 < \infty \tag{8}$$

The above expression of D can be solved using the results given in [1], [2], and [3] as

$$D = C\sum_{m=0}^{m_e} S_{m_e}^{(m)} {m_e}^m \sum_{r=0}^{m+\alpha-1}\binom{m+\alpha-1}{r}\left(\frac{m_e t_e}{(T+\nu)}\right)^{-r} I_1$$

where

$$C = \frac{\Gamma(m_e+\eta){m_e}^{\alpha-1}}{e^{\beta m_e}T_1(T+\nu)^{m_e+\eta}}$$

$$I_1 = \Gamma(r+1)\Psi\left(r+1, r-m_e+\eta+2, \frac{\beta(T+\nu)}{t_e}\right)$$

$\Gamma(.)$ is standard Gamma function and $\Psi(.,.,.)$ is confluent hyper-geometric function defined in [1], [2], and [3].

The marginal posterior of θ_0, say $\pi(\theta_0|\underline{t})$ can be obtained after integrating $\pi(\theta_0,\theta_1|\underline{t})$ over the whole range of θ_1 and it is

$$\pi(\theta_0|\underline{t}) \propto \Gamma(m_e+\eta){\theta_0}^{\alpha-1}e^{-\beta\theta_0}\theta_0^{\underline{m_e}}[t_e(\theta_0-m_e)+(T+\nu)]^{-(m_e+\eta)} \tag{9}$$

where

$$\theta_0 \geq m_e, \left(t_e(\theta_0-m_e)+(T+\nu)\right) \geq 0$$

The marginal posterior of θ_1, say $\pi(\theta_1|\underline{t})$ is the solution of $\pi(\theta_1|\underline{t}) = \int_{m_e}^{\infty}\pi(\theta_0,\theta_1|\underline{t})\, d\theta_0$ i.e.

$$\pi(\theta_1|\underline{t}) \propto \theta_1^{m_e+\eta-1}e^{-\theta_1[(T+\nu)-m_e t_e]}\, I_2 \tag{10}$$

where

$$I_2 = \sum_{m=0}^{m_e} S_{m_e}^{(m)}\, e^{-(\beta+\theta_1 t_e)m_e}\sum_{k=0}^{m+\alpha-1}\frac{(m+\alpha-1)!}{k!}{m_e}^k(\beta+\theta_1 t_e)^{-(m+\alpha-k)}$$

V. Bayes Estimates for Model Parameters

The Bayes estimators of θ_0 and θ_1 are posterior mean under the squared error loss function. The Bayes estimator for θ_0 i.e. the posterior mean can be obtained from (9) and is

$$\hat{\theta}_{B0} \propto \frac{\Gamma(m_e+\eta){m_e}^{\alpha}}{e^{\beta m_e}T_1(T+\nu)^{m_e+\eta}}\sum_{m=0}^{m_e} S_{m_e}^{(m)}\, {m_e}^m \sum_{r=0}^{m+\alpha}\binom{m+\alpha}{r}\left[\frac{m_e t_e}{(T+\nu)}\right]^{-r} I_3$$

where

$$I_3 = \Gamma(r+1)\Psi\left(r+1, r-m_e+\eta+2, \frac{\beta(T+\nu)}{t_e}\right)$$

The Bayes estimator for θ_1 is the posterior mean of its marginal posterior distribution (10) is

$$\hat{\theta}_{B1} \propto \frac{\Gamma(m_e+\eta+1){m_e}^{\alpha-1}}{e^{\beta m_e}T_1(T+\nu)^{m_e+\eta+1}}\sum_{m=0}^{m_e} S_{m_e}^{(m)}\, {m_e}^m \sum_{r=0}^{m+\alpha-1}\binom{m+\alpha-1}{r}\left[\frac{m_e t_e}{(T+\nu)}\right]^{-r} I_4$$

Where

$$I_4 = \Gamma(r+1)\Psi\left(r+1, r-m_e+\eta+1, \frac{\beta(T+\nu)}{t_e}\right)$$

VI. Discussion

The proposed Bayes estimators i.e. $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ of total number of failures (θ_0) and failure rate (θ_1) for the parameters of Binomial type exponential class SRGM are obtained by considering gamma priors and are compared with corresponding maximum likelihood estimators $\hat{\theta}_{m0}$ and $\hat{\theta}_{m1}$

respectively. The comparative performance of proposed Bayes estimators against corresponding maximum likelihood estimators has been studied based on the risk efficiencies i.e. $RE_0 = \frac{E(\hat{\theta}_{B0}-\theta_0)^2}{E(\hat{\theta}_{m0}-\theta_0)^2}$ and $RE_1 = \frac{E(\hat{\theta}_{B1}-\theta_1)^2}{E(\hat{\theta}_{m1}-\theta_1)^2}$. The estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ are based on the prior parameters α, β, η, ν and execution time t_e. The risk efficiencies are calculated by considering different values of these constants and arbitrary values of parameters θ_0 and θ_1 using the Monte Carlo Simulation technique by generating 10^3 samples. An execution time t_e is prefixed and up to this time, sample failures are generated and the risk efficiencies are presented in the following Figure 1 to Figure 14.

Consider the effect of variation in the value of t_e on the risk efficiencies of the proposed Bayes estimator $\hat{\theta}_{B0}$, given in Figure 1 to Figure 4.

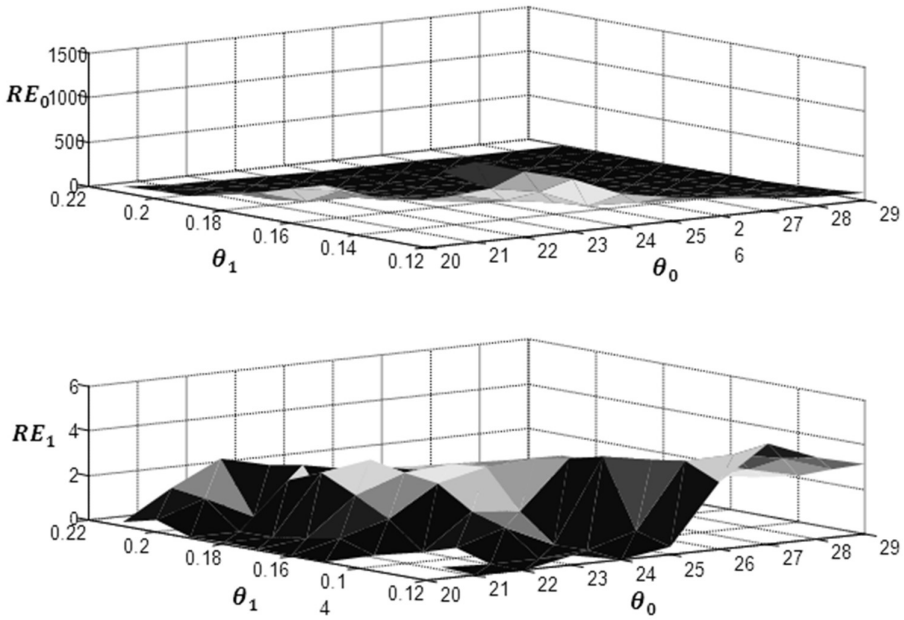


Figure 1: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $\alpha = 30$, $\beta = 10$, $\eta = 10$, $\nu = 10$ *and* $t_e = 3.0$

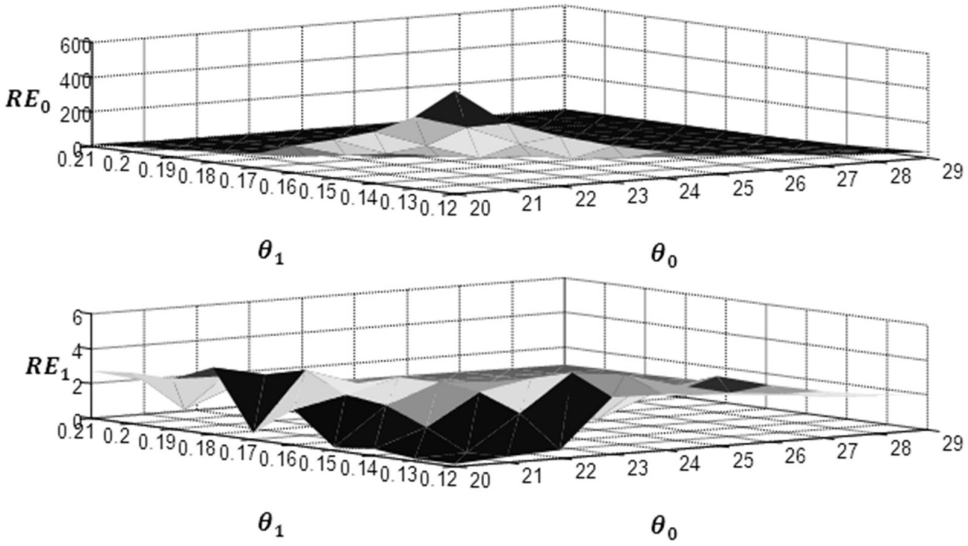


Figure 2: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $\alpha = 30$, $\beta = 10$, $\eta = 10$, $\nu = 10$ *and* $t_e = 3.5$

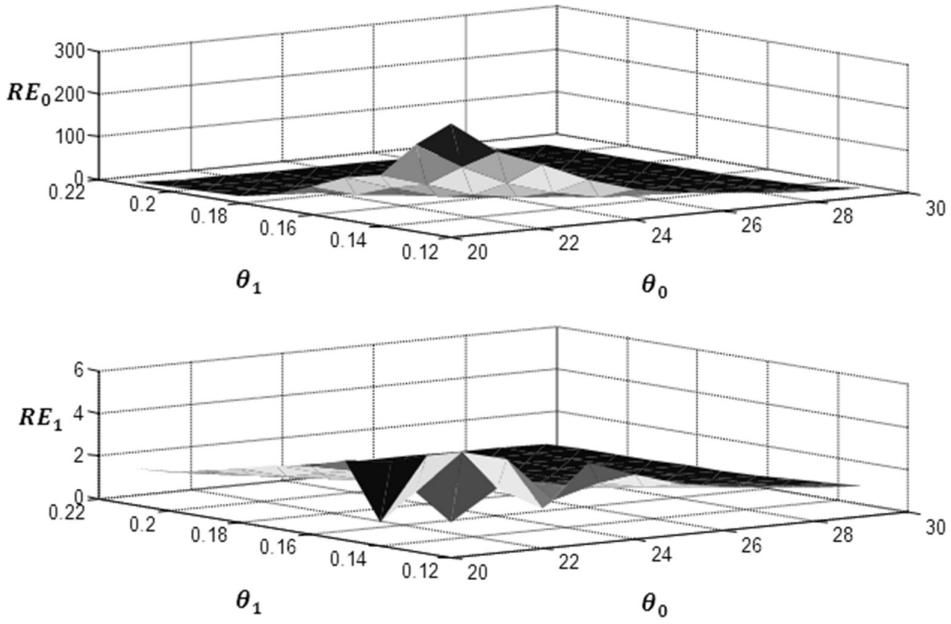


Figure 3: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $\alpha = 30$, $\beta = 10$, $\eta = 10$, $\nu = 10$ *and* $t_e = 4.0$

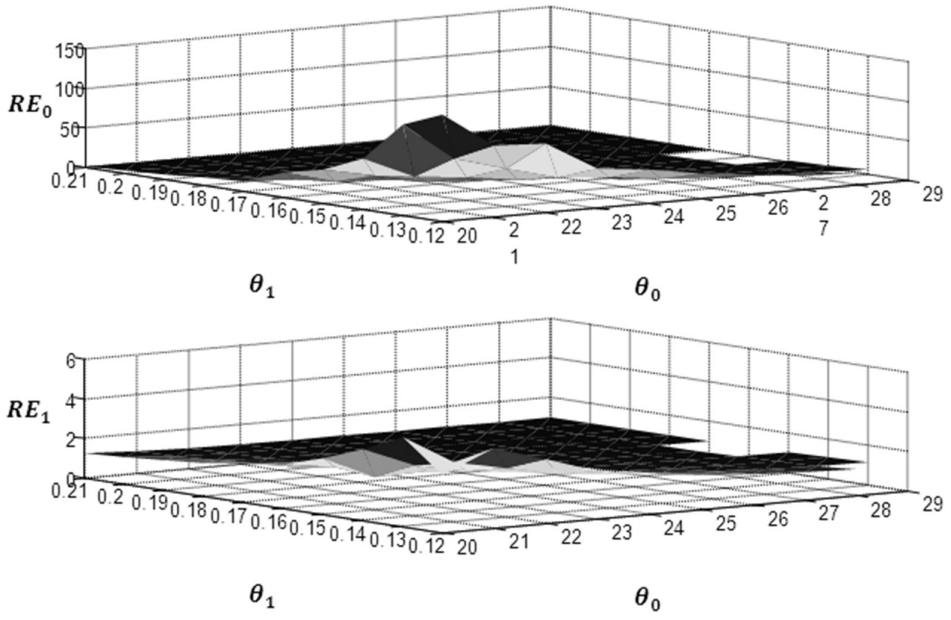


Figure 4: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $\alpha = 30$, $\beta = 10$, $\eta = 10$, $\nu = 10$ *and* $t_e = 4.5$

From these figures, it is seen that for the increase in value of t_e, the risk efficiencies RE_0 of Bayes estimator $\hat{\theta}_{B0}$ decrease as θ_0 and θ_1 increase. The point of maxima varies as the value of t_e changes. Particularly, the risk efficiency RE_0 attains maxima at smaller values of θ_0 and θ_1 for increasing values of t_e.

The variation of shape constant $\alpha(= 30(5)40)$ of proposed prior for the total number of failures, the risk efficiencies of Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ are presented in Figure 5 and Figure 6.

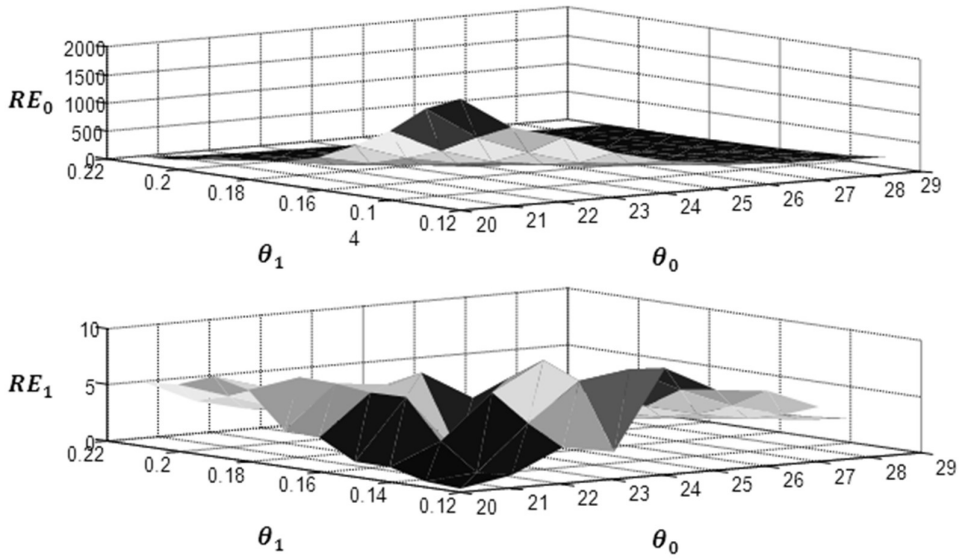


Figure 5: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $t_e = 4.0$, $\beta = 10$, $\eta = 10$, $\nu = 10$ *and* $\alpha = 35$

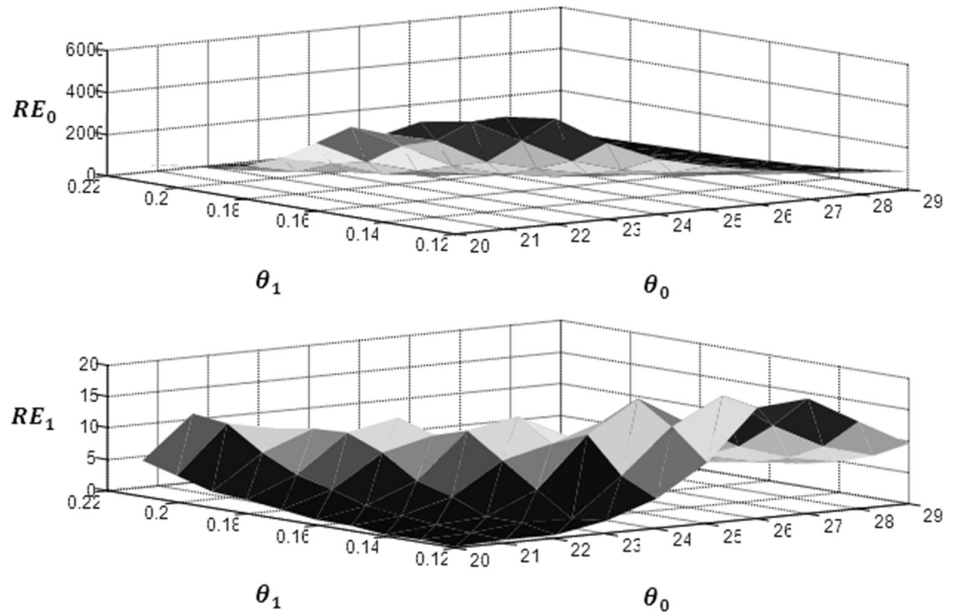


Figure 6: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $t_e = 4.0$, $\beta = 10$, $\eta = 10$, $\nu = 10$ *and* $\alpha = 40$

It is observed that the values of risk efficiencies of both the proposed estimators are increased for increasing the value of α. Here, in these figures the risk efficiencies of both the estimators are increasing for increasing values of α.

The risk efficiencies of Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ i.e. RE_0 and RE_1 calculated for different values of scale constant $\beta(= 1,10(5)20)$ of prior proposed for θ_0 are summarized in Figure 7 to Figure 8.

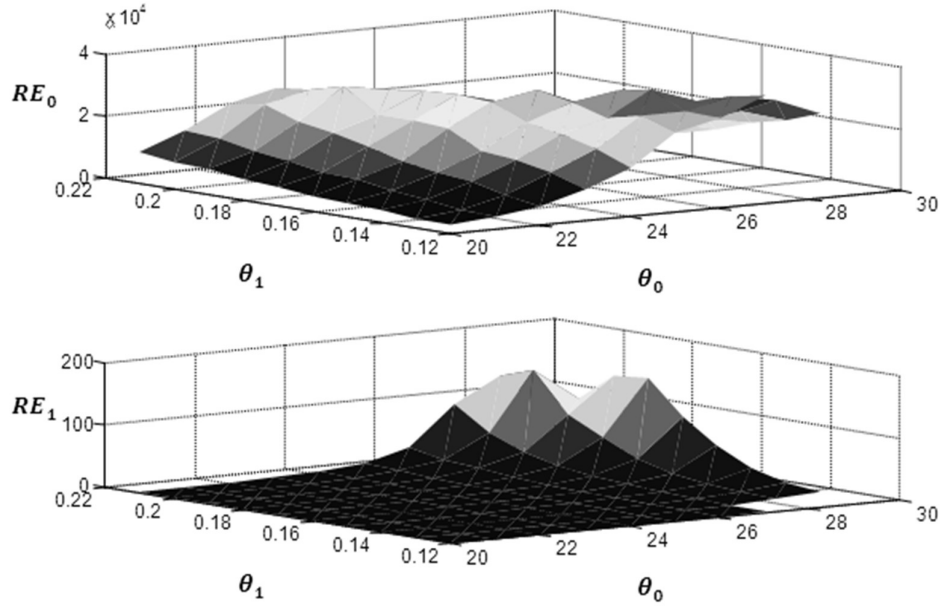


Figure 7: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $t_e = 4.0$, $\alpha = 30$, , $\eta = 10$, $\nu = 10$ *and* $\beta = 1$

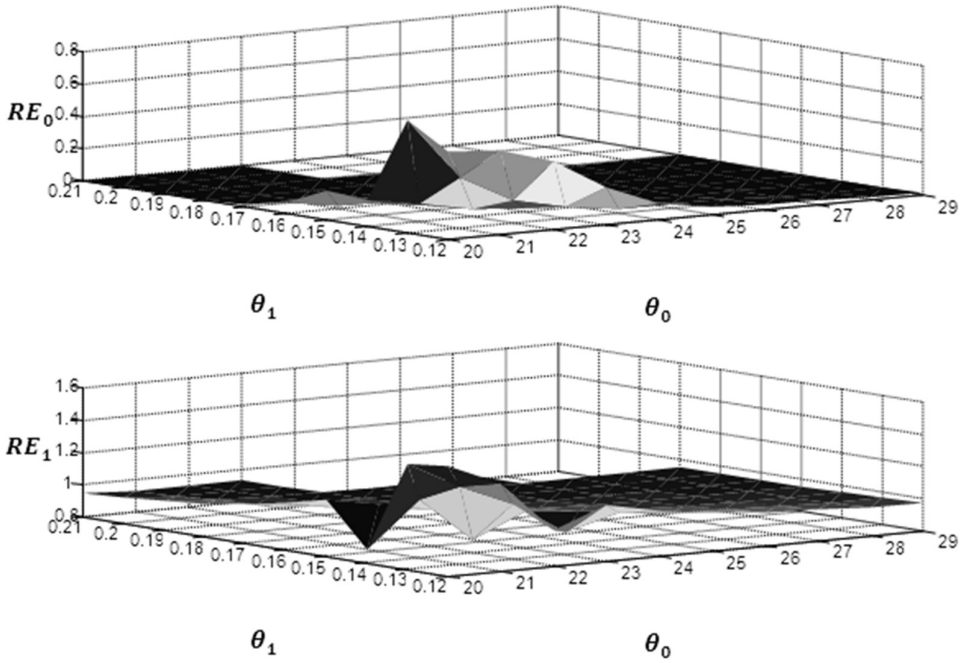


Figure 8: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $t_e = 4.0$, $\alpha = 30$, $\eta = 10$, $\nu = 10$ *and* $\beta = 15$

Here, the risk efficiencies of both estimators are decreasing for increasing values of β. It is also seen that, both the proposed Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ are becoming more inefficient than corresponding maximum likelihood estimators as β increases.

The risk efficiencies of Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ i.e. RE_0 and RE_1 are also evaluated using various values of shape constant $\eta(= 1,10(5)20)$ of prior for θ_1 and are summarized in Figure 9 to Figure 11.

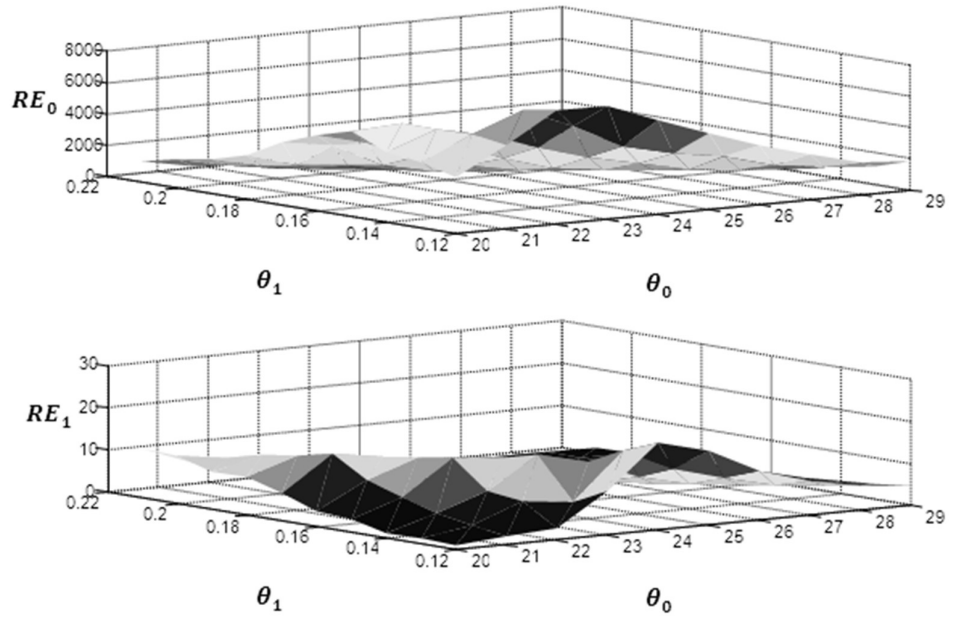


Figure 9: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$,$\theta_0(= 20(1)29)$,$t_e = 4.0$, $\alpha = 30$, $\beta = 10$, $\nu = 10$ *and* $\eta = 1$

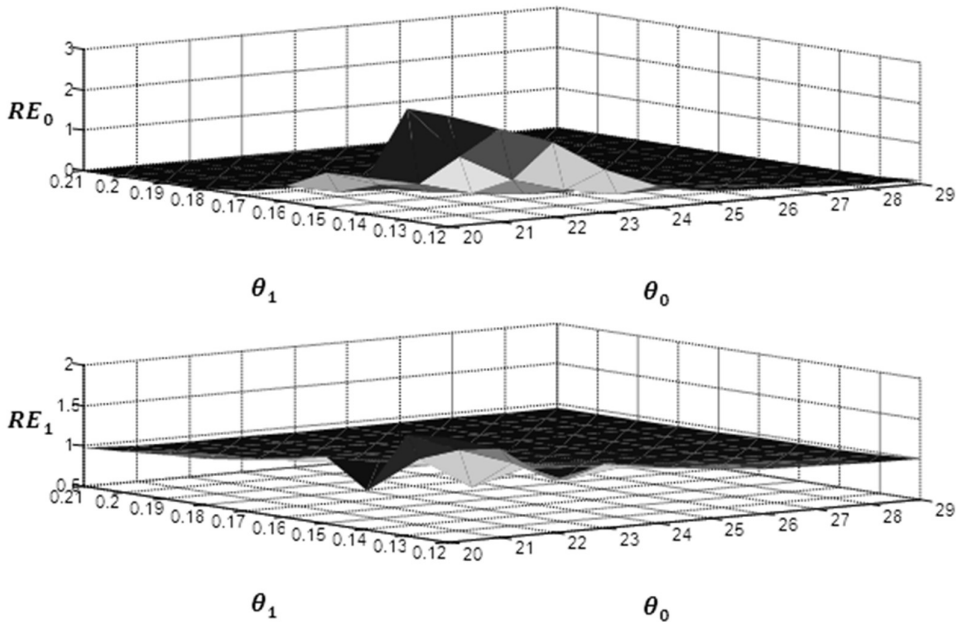


Figure 10: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)t_e = 4.0$, $\alpha = 30$, $\beta = 10$, $\nu = 10$ *and* $\eta = 15$

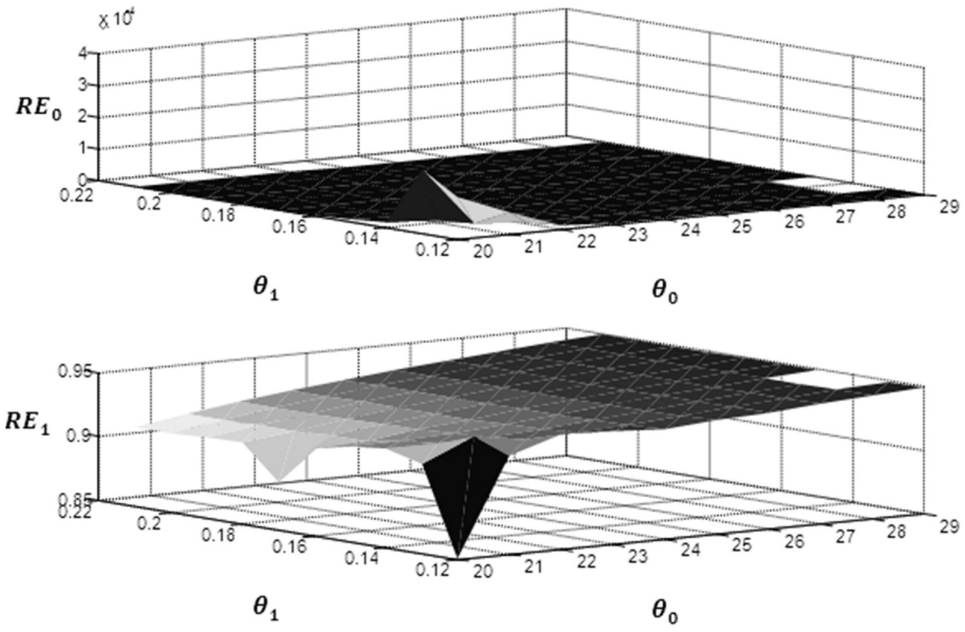


Figure 11: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $t_e = 4.0,\ \alpha = 30, \beta = 10,\ \nu = 10$ *and* $\eta = 20$

Here, it is observed that the risk efficiencies of both estimators decrease for the increase in the values of η. It is also seen that, both the proposed Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ are becoming more inefficient than corresponding maximum likelihood estimators as η increasing.

The risk efficiencies of Bayes estimators $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ i.e. RE_0 and RE_1 evaluated using various values of scale constant $\nu(= 1,10(5)20)$ of prior θ_1 and are summarized from Figure 12 to Figure 14.

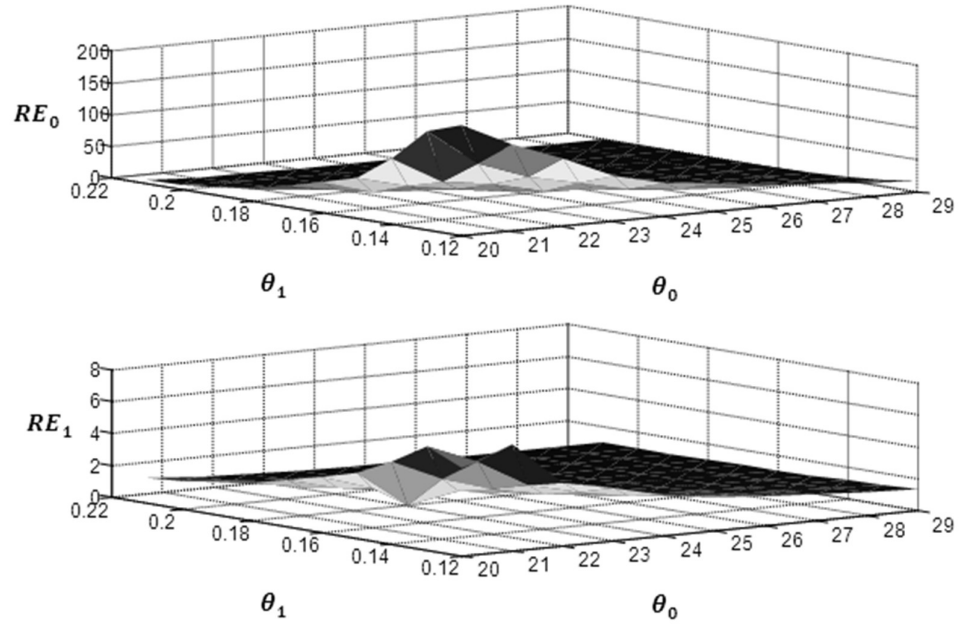


Figure 12: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $\alpha = 30$, $\beta = 10$, $\eta = 10$, $t_e = 4.0$ *and* $\nu = 1$

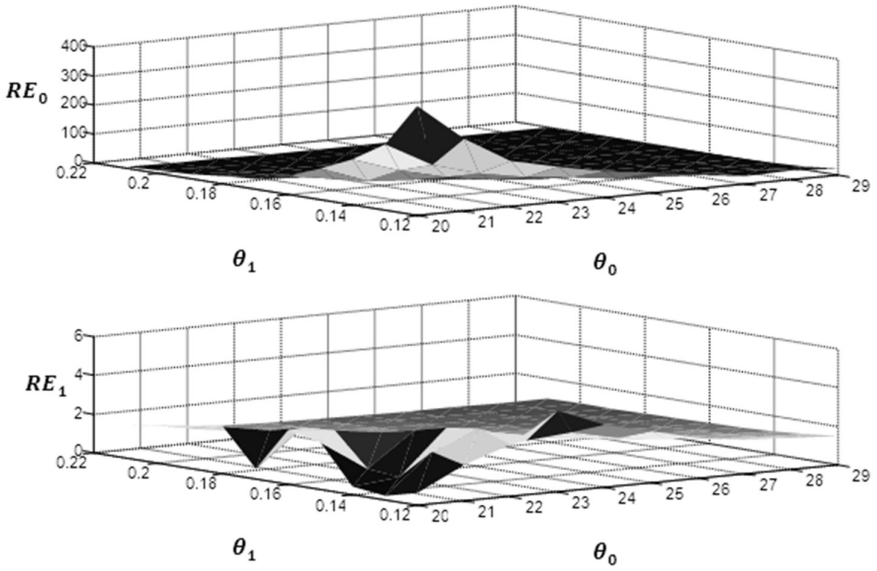


Figure 13: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$,$\theta_0(= 20(1)29)$, $t_e = 4.0$, $\alpha = 30$, $\beta = 10$, $\eta = 10$ *and* $\nu = 15$

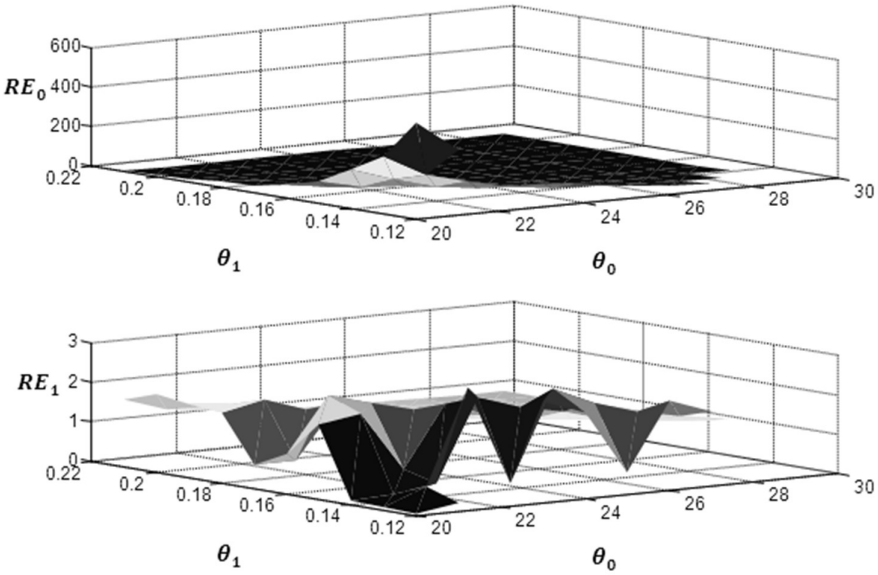


Figure 14: *Risk Efficiencies of* $\hat{\theta}_{B0}$ *and* $\hat{\theta}_{B1}$*, for* $\theta_1(= 0.12(0.01)0.21)$, $\theta_0(= 20(1)29)$, $t_e = 4.0$, $\alpha = 30$, $\beta = 10$, $\eta = 10$, *and* $\nu = 20$

Here, it is seen that the risk efficiencies of $\hat{\theta}_{B0}$ are increasing whereas the risk efficiencies of $\hat{\theta}_{B1}$ are decreasing for increasing the values of ν. It is also seen that the proposed Bayes estimator $\hat{\theta}_{B0}$ is becoming efficient as η increases whereas $\hat{\theta}_{B1}$ becoming more inefficient than the corresponding maximum likelihood estimator.

VII. Conclusions

Both the proposed Bayes estimator of θ_0 and θ_1 i.e. $\hat{\theta}_{B0}$ and $\hat{\theta}_{B1}$ can be preferred over corresponding MLEs if the parameters of gamma priors for model parameters are properly chosen. The value of t_e should be small for moderate values of true parameters and prior constants. The values of shape constant α of prior proposed for θ_0 should be chosen moderately large for smaller values of t_e. The values of scale constant β of prior proposed for the total number of failures i.e. θ_0 should be chosen

smaller when values of t_e are small. The values of prior parameters η and ν should be chosen smaller for smaller values of t_e.

References

[1] Abramowitz, M. and Stegun, I. A. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover publications, New York, 1965.

[2] Gradshteyn, S. and Ryzhik, I. M. Table of Integrals, Series, and Products, Alan Jeffrey (editor), Academic Press, New York, 1994.

[3] Graham, R. L., Knuth, D. E. and Patashnik O. Concrete Mathematics: A Foundation for Computer Science, 2nd Ed. Addison-Wesley Publishing Co., New York, 1994.

[4] Jelinski, Z. and Moranda, P. B. Software Reliability Research. (Freiberger, W. Editor), Statistical Computer Performance Evaluation, Academic Press, New York. 465-484. (1972).

[5] Littlewood, B. and Verrall, J. L. (1973). A Bayesian Reliability Growth Model for Computer Software, Journal of Royal Statistical Society, 22(3):332-346.

[6] Littlewood, B. and Verrall, J. L. (1974). A Bayesian Reliability Model with Stochastically Monotone Failure Rate, IEEE transactions on Reliability, 23(2):108-114.

[7] Littlewood, B. Software Reliability: Achievement and Assessment, Blackwell Scientific Publications, London, 1987.

[8] Lyu, M. R. and Nikora A. (1992). Applying Reliability Models More Effectively, IEEE Software journals and Magazines, 9(4):43-52.

[9] Martz, H. F. and Waller R. A. Bayesian Reliability Analysis, Wiley, New York, 1982.

[10] Musa, J. D. and Okumoto, K. (1984). A Logarithmic Poisson Execution Time Model for Software Reliability Measurement. Proc. of 7th Int. conf. on software engineering, Orlando, 230-238,

[11] Musa, J. D. and Okumoto, K. Software Reliability Models: Concepts, Classification, Comparison and Practice, (J.K. Skwirzynski ed.), Electronic system effectiveness and life cycle casting, NATO ASI series F3, Springer-Verlag: Heidelberg, 395-424, 1983.

[12] Musa, J. D., Iannino, A. and Okumoto, K. Software Reliability: Measurement, Prediction, Application, McGraw-Hill, New York, 1987.

[13] Shooman M. L. Probabilistic Models for Software Reliability Prediction, (Freiberger, W. Editor), Statistical Computer Performance Evaluation, Academic Press, New York, 485-502, 1972.

[14] Singh, R., Badge, P. A. and Singh, P. (2023). Confidence Interval Using Maximum Likelihood Estimation For The Parameters Of Poisson Type Length Biased Exponential Class Model. RT&A. December 4(76): 242-251,

[15] Singh, R., Badge, P. A. and Singh, P. (2024). Confidence Interval Using Maximum Likelihood Estimation For The Parameters Of Poisson Type Rayleigh Class Model. RT&A. March 1(77): 832-841,

[16] Singh, R., Badge, P., and Singh, P. (2022b) Bayesian Interval Estimation for the Parameters of Poisson Type Rayleigh Class Model, RT&A. 17(4):98-108.

[17] Singh, R., Badge, P., and Singh, P. (2023a). Bayesian Interval Estimation for the Parameters of Poisson Type Length Biased Exponential Class Model, RT&A. 18(2):307-314.

[18] Singh, R., Singh, P. and Kale, K. R. (2016). Bayes estimators for the parameters of Poisson Type Length Biased Exponential Class Model using Non-Informative Priors, JRSS, 9(1): 21-28.

[19] Singh, R., Singh, P. and Kale, K. R. (2022a). Estimation of Parameters of PTRC SRGM using Non-informative Priors, IJARSCT. 2(1):172-178.

[20] Singh, R., Singh, P. and Kale, K. R. (2023b). Characterization of Poisson Type Length Biased Exponential Class Software Reliability Growth Model and Parameter Estimation, RT&A. No. 3(74):18.

[21] Sinha, S. K. Reliability and life testing. Wiley, New Delhi, 1985.

STUDY OF THE FUNCTIONING OF A MULTI-COMPONENT AND MULTI-PHASE QUEUING SYSTEM UNDER THE CONDITIONS OF THE IMPLEMENTATION OF DISRUPTIVE TECHNOLOGIES IN AIR TRANSPORTATION

[1]O. Zaporozhets, [2]M. Katsman, [3]V. Matsiuk, [4]V. Myronenko

•

[1]Institute of Aviation (ILot), Warsaw, Poland
[2]The Joint-Stock Company of Railway Transport of Ukraine "Ukrzaliznytsia", Kyiv, Ukraine
[3]National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
[4]State University of infrastructure and technologies, Kyiv, Ukraine
[1]zaporozhets.oleksandr@gmail.com, [2]mdkatsman@gmail.com,
[3]vimatsiuk@gmail.com, [4]pupil7591@gmail.com

Abstract

The article considers multi-component and multi-stage mathematical models of queuing systems (QS) with the distribution of the incoming flow simultaneously between the system components, which consist of a certain number of service channels and waiting places in the queue. The maintenance of requirements with a lack of time to stay in the service channel and waiting is considered, while the service process in the QS of each component consists of several stages with the corresponding duration, and the full-service period is equal to the sum of such time intervals. The number of components and their parameters correspond to the similar characteristics of the production divisions of the repair enterprise. The study of the effectiveness of the operation of the repair enterprise as a multi-component and multi-stage QS consists in determining the values of the initial parameters of the QS components, taking into account the restrictions imposed on them, in order to obtain the largest values of the probabilities of servicing the requirements of the QS components and the system as a whole. The model is implemented using Any Logic University Researcher, which allows you to combine the principles of system dynamics with the paradigms of agent and discrete-event modelling. The proposed approach to the modelling of maintenance and repair processes by production divisions of the enterprise as a multi-component and multi-phase QS allows to determine the effectiveness of the functioning of such a QS and to obtain arguments for increasing the efficiency of its operation.

Keywords: aviation repair enterprise, aircraft, maintenance and repair, multi-component and multi-phase queuing system

1. Introduction

About 2% of the total worldwide fossil fuel usage corresponds to fuel consumption in aviation. Hybrid renewable integration, electrification, hydrogenation and optimizations are necessary roadmaps for the transition towards low-carbon airport transportation systems [1]. The rise of

aircraft electrification with a few digital transformations in their operation and maintenance is making significant strides toward the sustainability of the aviation industry due to these absolutely new disruptive technologies and is a new challenge in air transportation development during the following decades.

As a result, aircraft emissions have caused approximately 2% of the total CO2 emissions [2]. Concept of More Electrical Aircraft (MEA) becomes much stronger [3], [4], so as the concepts of Hybrid Electrical Aircraft (HEP) and Full Electric Aircraft (FEP) in next few years will be realized in reality and the HEP and FEP will appear in operation evidently [1] − [4]. Components such as air-conditioning, cabin pressurization, de-icing, landing gear, and brake systems have traditionally been powered by pneumatic, hydraulic, or mechanical systems. Nowadays, it's becoming more common to see these components being electrically powered " current MEA concept [2]. The number of electrically powered components will be increased on the board of HEP and FEP aircraft, and so will the power needs and complexity of the systems. Transmitting large quantities of electrical power around an aircraft at a high voltage is required to minimize resistive losses. This transmission creates (or increases) the risk of insulation breakdown and arcing, which can cause catastrophic equipment failure " as electric equipment as other usual onboard equipment of the aircraft.

Also, the ground handling operations, which are used in airports for handling activities and processing passengers with the help of specially designed vehicles known as ground support equipment (GSE), will be changed from traditional energy usage on more environmentally efficient " electrified and hydrogenated [5].

Thus, electrical systems become decisive on board the aircraft both in terms of flight support and in carrying out maintenance services, which predetermines the need for high levels of reliability in carrying out such work. These FEA and HEA concepts will compete with sustainable aviation fuels (SAFs) implementation in the aviation sector, including biofuels [6] and hydrogen [7].

The aircraft maintenance system is designed to maintain and restore the airworthiness and serviceability of aircraft and prepare them for flight. Technical operation is carried out by operators, aviation and technical bases, maintenance and repair enterprises, repair enterprises, aviation and technical services of airports [8], [9].

One of the main factors of aircraft flight safety is the reliability of the on-board power supply system, which includes power sources, a control and protection system, switching equipment of the power distribution system, electric drives, lighting equipment, light signalling, fire extinguishing and anti-icing systems, and some other equipment. New investigations should be focused on the safety aspects of manned electric (HEP or FEP) flight based on the emerging technologies that are expected to be developed in the current decade, including their maintenance in airports and repair plants. The biggest safety concern is with the lithium-ion batteries that will power HEP or FEP aircraft. The batteries have the potential to ignite during the charging process through an uncontrollable temperature increase known as a thermal runaway.´ Also, battery energy uncertainty and battery charging safety will be the subject for flight hazards. In addition, any damage to the battery that causes the chemicals inside to be exposed to oxygen or water can lead to rapid oxidation and system failure.

Maintenance of the power supply system is carried out during capital and other repairs (or during equivalent works), inspections, modifications, upgrading, elimination of defects, which are carried out by aircraft repair enterprises both individually and collectively in the relevant workshops, production divisions, production areas, laboratories, stands, etc.

The work [10] is devoted to the creation and introduction into practice of aviation information and advisory systems for the maintenance of passenger aircraft based on modern computer technologies and mathematical methods of information processing.

In [11], the structure of the methodical apparatus for ensuring a given level of serviceability of on-board equipment products, in particular optoelectronic sighting systems of military aircraft of the Air Force of Ukraine, is proposed.

Methodical approaches to the structural and parametric determination of general requirements

for ground flight maintenance facilities are considered in the paper [12], which can be used to develop a methodology for conducting tests and assessing the quality of modern air traffic control systems at all stages of the life cycle.

Based on the analysis of the existing methods of calculating the durability indicators of the radio-electronic system of the aircraft, the factors affecting its reliability were identified in [13], and measures were proposed to improve the existing scientific and methodological apparatus for calculating such indicators.

Works [14] and [15] are devoted to the analysis of the causes of failure situations at the airport. The aircraft maintenance system was analysed, it was shown that ensuring uninterrupted operation of the airport, execution of the daily flight plan in extraordinary situations is possible only by introducing into the control circuit of the aircraft ground maintenance system an intelligent decision support system for dispatchers, which will take into account the positive experience of their actions in typical, extraordinary and failure situations. This will allow, in particular, to reduce the time to get out of a malfunctioning situation and to optimize the operational planning of the ground maintenance of aircraft, considering the available equipment and special equipment.

In work [16], organizational measures are given, with the help of which it is possible to minimize the lack of transport aviation during the transportation of cargoes, including the oversized ones. Data on incidents related to aircraft ground maintenance are given, the causes of the events are indicated. The main ways of eliminating the problems of standardization of airfield technical support in the conditions of interaction with NATO and in the processes of international integration are defined.

The work [17] is devoted to the solution of the problem of minimizing the risks of import substitution in the process of factory repair of military aviation equipment in the conditions of a special period, the issue of post-repair maintenance of military aviation equipment due to the manufacture of the necessary component parts by domestic enterprises in the process of import substitution is analysed.

The work [18] presents the results of the quality of repair of aircraft equipment at aircraft repair enterprises. A significant proportion of failures detected during the operation of aviation equipment after capital (medium) repair is a consequence of manufacturing defects of components (parts) that were installed on aircraft. Technological methods for ensuring sufficient repair quality and significantly reducing the risks of production defects are proposed.

The work [19] is devoted to the problem of mathematical modeling of the processes of technical operation of military aircraft. The results of the analysis show that the most acceptable modelling method in terms of the compliance of the models with the proposed requirements is the simulation modelling method, and the more accepted model class for creating a stochastic model of aircraft maintenance and repair processes is the class of semi-Markov models.

In work [20], a three-dimensional model of an aircraft skin element with riveted seams was built using the Sold Works software, wind load simulation was carried out in the ANSYS software package, which made it possible to determine the stress-strain state of aircraft skin elements in the presence of multifocal damage to riveted seams.

Modern methods and approaches to modelling technological systems are considered in [21]. Basic definitions and concepts are given. New approaches to solving problems that arise during the development of models of mechanisms, systems and processes of machine-building production are proposed.

In work [22], it is proposed to consider the functioning of car service enterprises as an open multi-channel QS, in which random processes occur due to the combined action of random factors. As a result of the experimental study, information was obtained about the indicators characterizing maintenance and repair, as well as affecting the change in the parameters of these processes. The developed model makes it possible to consider the specifics of managing car maintenance stations.

In [23], a model for assessing the technical condition of radio-electronic elements of water transport vessels using control and diagnostic equipment as a QS with a limited number of

channels and a storage of arriving customers is considered. On the basis of various optimization criteria, it is possible to establish a rational system for assessing the technical condition of such elements, to determine the feasibility of developing a certain (rational, optimal) number of different types of control and diagnostic equipment and the effectiveness of new assessment methods.

The work [24] is devoted to the development of a simulation model of the influence of an accurate assessment of the readiness factor of mobile control and diagnostic complexes on the reliability of control of radio-electronic systems of marine transport.

In work [25], a new model of the task of managing the processes of diagnosis and monitoring of automation tools is proposed for the objects of rail-water transport connection, compiled based on the results of experimental research and mathematical description using Markov chains with an informative parameter in the form of damage intensity, aimed at increasing the efficiency of forecasting the technical condition of automation equipment.

The work [26] describes for locomotive repair workshops in the form of multi-channel QS with a limited queue. A simulation model of such a workshop as a QS object was developed, which allowed rational use of equipment, labor force, as well as distribution of repair work time.

In work [27], the issue of modelling the maintenance and repair processes of technical components of a distributed information system is considered. The model is based on a joint presentation of the serviced system and its technical operation process in the form of a closed non-homogeneous QS consisting of two types of QSs. The QS of the first type simulates the functioning processes of repair bodies to meet and serve the arriving customers.

In work [28], a study of the actions of railway transport emergency units as a process of functioning of QSs was carried out. The authors established quantitative relationships between the intensity of the influence of a railway accident dangerous factors, on one hand and, on the other hand, the time of arrival, deployment and productivity of emergency liquidation units and the effectiveness of liquidation works due to the implementation of the network-centric management principles for complex dynamic hierarchical transport systems.

In paper [29], mathematical models of QSs with the distribution of the arrival flow of customers simultaneously over several service channels are considered. The model is implemented using agent simulation in the AnyLogic University Researcher environment and the Java compiler. The use of the proposed mathematical models will make it possible to establish areas of accepted values of the probability of successful completion of assigned tasks in order to make managerial decisions regarding the rational use of forces and means for the elimination of the consequences of railway transport events.

Thus, to improve the management processes for material, human, financial and informational resources during the maintenance and repair of aircraft and other means of transport, in particular on-board power supply systems, a wide range of methods of operations research, the queuing theory and simulation modelling are currently used.

2. Methods

The on-hand practical experience of the organization of maintenance and repair of aviation equipment indicates that certain types of technical systems of aircraft that require various types of repair work, modifications, upgrade, inspections, elimination of defects, etc. are sent to specific production divisions of the repair enterprise, which, according to their purpose, carry out the necessary types of work according to the specified technologies.

To simulate the processes of maintenance and repair of aircraft, which are carried out by the production divisions of the aircraft repair enterprise, it is advisable to use multi-component and multiphase QSs, which can be of both Markov and non-Markov types, capable of serving the arriving flows of non-priority, in general, heterogeneous (mixed) customers. At the same time, the system can have an arbitrary number of common service channels of the same type, and each component can also have an arbitrary number of places in the queue.

The same service channels can have different performance depending on the types of requirements for which they are involved: when the j-component of the system receives uniform requirements with the rate λ_j determined by the overall rate λ of the source, in the general case, of mixed customers. The magnitude of the source of mixed customers entering the system has an intensity (arrival rate) of

$$\lambda = \sum_{j=1}^{L} \lambda_j, \tag{1}$$

where L is the number of components in the QS. The service process in each component of the QS consists of several stages (phases) with the corresponding duration T_i, then the full-service period T_s is equal to

$$T_s = \sum_{j=1}^{K_E} T_i \tag{2}$$

where K_E is the number of such phases.

All T_i durations have certain probability distributions with the appropriate parameters, then T_s will have a generalized Erlang distribution with the parameters of the probability distributions of order K_E.

The number of components and their parameters correspond to similar, characteristics of the repair enterprise.

The study of the operation effectiveness of the aviation repair enterprise as a multi-component QS will consist in determining the probability and time characteristics of each component and the QS as a whole. Let's consider several examples.

Example 1. Two-component QS of M/E4/2/3 type in the first component and of M/E3/1/2 in the second component with restrictions on the time spent in the service period $\beta_{1,2}$ and waiting $\gamma_{1,2}$, which is due to the force majeure circumstances in the QS operation: $\beta_{1,2}$ is the intensity of leaving the service channel due to the limitation of time spent in the system during the service period; $\gamma_{1,2}$ is the intensity of customers leaving the queue due to the time limit of their stay in the system during the service waiting period.

The graph of the states of this QS coincides with the graph of the states of the QS presented in Fig.1.

In Fig. 1 it is indicated:

$\mu_{11} = \mu^{/} + \beta_1^{/};\ \mu_{12} = 2\mu^{/} + 2\beta_1^{/};\ \mu_{13} = \mu_{12} + \gamma_1^{/};\ \mu_{14} = \mu_{12} + 2\gamma_1^{/};\ \mu_{15} = \mu_{12} + 3\gamma_1^{/};\ \mu^{/} = 4\mu_1;\ \mu_1 = \frac{1}{\bar{t}_{s_1}};\ \beta_1^{/} = 4\beta_1;\ \beta_1 = \frac{1}{\bar{t}_{lim_{s_1}}};\ \gamma_1^{/} = 4\gamma_1;\ \gamma_1 = \frac{1}{\bar{t}_{lim_{w_1}}};\ \mu_{21} = \mu^{//} + \beta_2^{//}; \mu^{//} = 3\mu_2; \mu_2 = \frac{1}{\bar{t}_{s_2}};\ \beta_2^{//} = 3\beta_2;\ \beta_2 = \frac{1}{\bar{t}_{lim_{s_2}}};\ \mu_{22} = \mu_{21} + \gamma_2^{//};\ \gamma_2^{//} = 4\gamma_2;\ \ \gamma_2 = \frac{1}{\bar{t}_{lim_{w_2}}};$ where $\bar{t}_{lim_{s_1}}$, $\bar{t}_{lim_{s_2}}$ are the limited service time; $\bar{t}_{lim_{w_1}}, \bar{t}_{lim_{w_2}}$ are the limited waiting time.

The QS states of the first component are characterized by the following probabilities [30][31]:

$$P_{c^{/}}^{/} = \sum_{j=1}^{4} P_{c^{/}j}^{/};\quad c^{/} = \overline{1,5}, \tag{3}$$

where $P_1^{/}$ is the probability of occupation of one channel (1 customer in the component); $P_2^{/}$ is the probability of occupation of 2 channels (2 customers in the component); $P_3^{/}$ is the probability of 3 customers being in the component, of which 2 are served, one is in the queue; $P_4^{/}$ is the probability of 4 customers being in the component, of them 2 are served and 2 are in the queue; $P_5^{/}$ is the probability of 5 customers being in the component, of which 2 are being served and 3 are in the queue:

$P_1^{/} = \sum_{j=1}^{4} P_{1j}^{/};\ \ P_2^{/} = \sum_{j=1}^{4} P_{2j}^{/};\ \ P_3^{/} = \sum_{j=1}^{4} P_{3j}^{/};\ \ P_4^{/} = \sum_{j=1}^{4} P_{4j}^{/};\ \ P_5^{/} = \sum_{j=1}^{4} P_{5j}^{/}.$

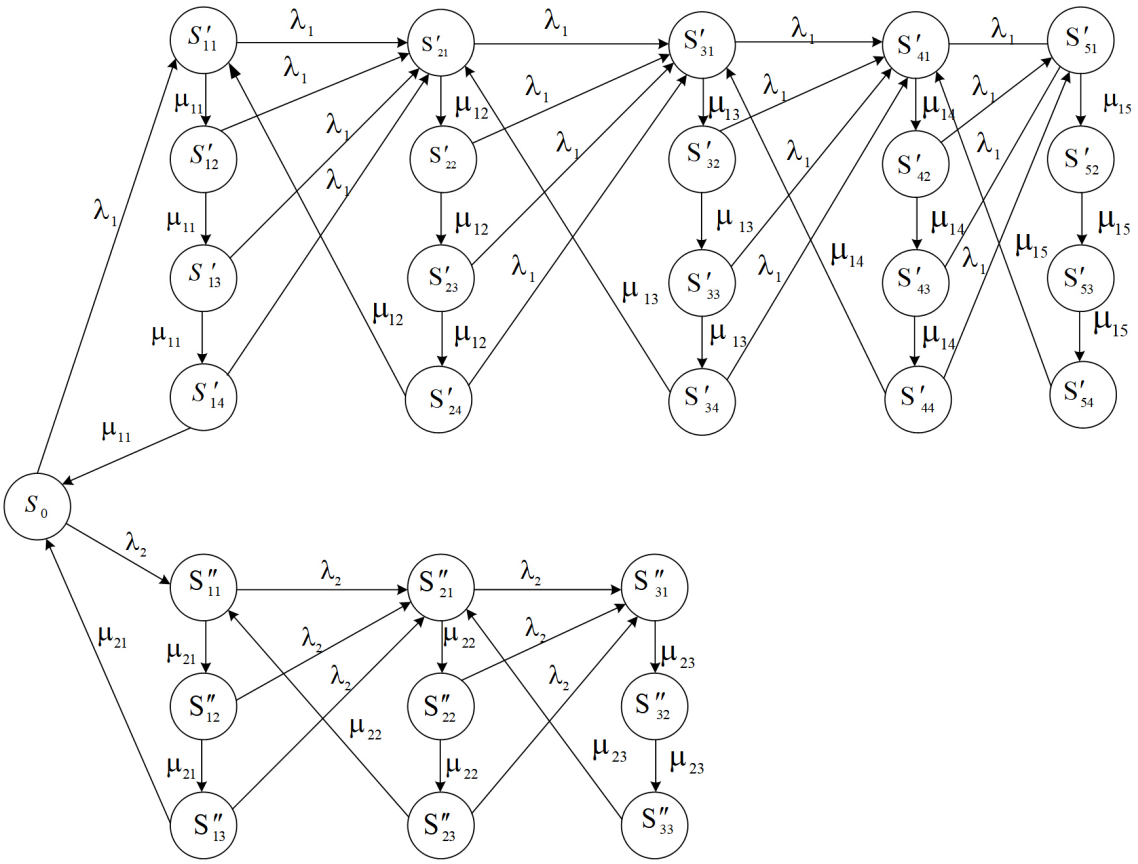


Figure 1: *State graph of QS of M/E4/2/3 type in the first component and M/E3/1/2 type in the second component with restrictions on the time spent in the service period $\beta_{1,2}$ and waiting $\gamma_{1,2}$*

Similarly, for the QS states of the second component:

$$P^{//}_{c^{//}} = \sum_{j=1}^{3} P^{//}_{c^{//}j}; \quad c^{//} = \overline{1,3}, \tag{4}$$

where $P^{//}_1$ is the probability of one customer being served in the component; $P^{//}_2$ is the probability of 2 customers being in the component, one of them is in service, the other is in the queue; $P^{//}_3$ is the probability of 3 customers being in the component, one of them is in service, two are in the queue:

$$P^{//}_1 = \sum_{j=1}^{3} P^{//}_{1j}; \; P^{//}_2 = \sum_{j=1}^{3} P^{//}_{2j}; \; P^{//}_3 = \sum_{j=1}^{3} P^{//}_{3j}.$$

The number of busy service channels in the components [30], [31]:

$$\overline{k_1} = \frac{P^{/}_1 + \sum_{c^{/}=2}^{5} P^{/}_{c^{/}}}{4}, \tag{5}$$

$$\overline{k_2} = \frac{\sum_{c^{//}=1}^{3} P^{//}_{c^{//}}}{3}. \tag{6}$$

Probability of service in the first component:

$$\mathrm{P}_S^{/} = 1 - P_{lS}^{/} - \sum_{c^{//}=1}^{(n+m)^{//}} P_{c^{//}}^{//} = 1 - P_{fl}^{/} - \sum_{c^{//}=1}^{(n+m)^{//}} P_{c^{//}}^{//},$$

where $P_{lS}^{/}$ is the probability of loss of a request; $P_{fl}^{/}$ is the probability of failure of serving a request.

$$P_{fl}^{/} = P_{(n+m)^{/}}^{/} - \frac{P_{bo}^{/}}{4} = P_5^{/} - \frac{P_{bo}^{/}}{4},$$

where $P_{bo}^{/} = \sum_{d=2;j=2}^{4}(d-1)P_{(b_1-1)j}^{/}$, $b_1 = \overline{2,(n+m)^{/}}$
then $P_{20}^{/} = \sum_{i=2}^{4}(i-1)P_{1i}^{/}$, $P_{30}^{/} = \sum_{i=2}^{4}(i-1)P_{2i}^{/}$, $P_{40}^{/} = \sum_{i=2}^{4}(i-1)P_{3i}^{/}$, $P_{50}^{/} = \sum_{i=2}^{4}(i-1)P_{4i}^{/}$, $P_{b0}^{/} = \sum_{i=2}^{5}P_{io}^{/}$. $P_{ls}^{/} = P_{fl}^{/} + P_{lv_S}^{/} + P_{lv_q}^{/}$.
where $P_{lv_S}^{/}$ is the probability of the customer leaving the system in the service channel; $P_{lv_q}^{/}$ is the probability of the customer leaving the system in the queue.

When

$$P_{fl}^{/} = P_5^{/} - \frac{P_{b0}^{/}}{4};$$
$$\mathrm{P}_{lv_S}^{/} = \frac{\beta_1\overline{k_1}}{\lambda_1};$$
$$\mathrm{P}_{lv_q}^{/} = \frac{\gamma_1\overline{N_q^{(1)}}}{\lambda_1};$$

where $P_{ls}^{/} = P_5^{/} - \frac{P_{b0}^{/}}{4} + \frac{(\beta_1\overline{k_1}+\gamma_1\overline{N_q^{(1)}})}{\lambda_1}$.

The expressions for the probabilities of the QS states of the first and second components are similar to the QS considered above.

The probability of customer service in the second component is

$$\mathrm{P}_S^{//} = 1 - P_{ls}^{/} - \sum_{c^{/}=1}^{5} P_{c^{/}}^{/},$$

where

$$\mathrm{P}_{lS}^{//} = P_{fl}^{//} + P_{lv_S}^{//} + P_{lv_q}^{//};$$
$$\mathrm{P}_{fl}^{//} = P_3^{//} - \frac{P_{b0}^{//}}{3};$$
$$\mathrm{P}_{lv_S}^{//} = \frac{\beta_2\overline{k_2}}{\lambda_2};$$
$$\mathrm{P}_{lv_q}^{//} = \frac{\gamma_2\overline{N_q^{(2)}}}{\lambda_2}.$$

Provided $\beta_2 = \gamma_2$, then

$$\mathrm{P}_{lv_q}^{//} = \frac{\beta_2\overline{N_q^{(2)}}}{\lambda_2}.$$

$P_{bo}^{/} = \sum_{d=2;j=2}^{3}(d-1)P_{(b_2-1)j}^{//}$, $b_2 = \overline{2,(n+m)^{//}}$,

then $P_{20}^{//} = \sum_{j=2}^{3}(j-1)P_{1j}^{//}$, $P_{30}^{//} = \sum_{j=2}^{3}(j-1)P_{2j}^{//}$,

Where

$$\mathrm{P}_{ls}^{//} = P_3^{//} - \frac{P_{b0}^{//}}{3} + \frac{(\beta_2\overline{k_2}+\gamma_2\overline{N_q^{(2)}})}{\gamma_2}.$$

Whence the average number of requests $\overline{N_q^{(i)}}$ that are in the queue and waiting for service in the i-component:

$$\overline{N_q^{(1)}} = \sum_{q^{/}=1}^{m^{/}} \frac{q^{/}P_{(n+q)^{/}}^{/}}{k_{E^{/}}} = \sum_{q^{/}=1}^{3} \frac{P_{(2+q^{/})}^{/}}{4};$$
$$\overline{N_q^{(2)}} = \sum_{q^{//}=1}^{m^{//}} \frac{q^{//}P_{(n+q)^{//}}^{//}}{k_{E^{//}}} = \sum_{q^{//}=1}^{2} \frac{P_{(1+q^{//})}^{//}}{3}.$$

The average number of customers $\overline{N^{(i)}}$ in the i-component:

$$\overline{N_q^{(1)}} = \sum_{c^{/}=1}^{(n+m))^{/}} \frac{c^{/}P_{(c)^{/}}^{/}}{k_{E^{/}}} = \sum_{c^{/}=1}^{5} \frac{c^{/}P_{(c)^{/}}^{/}}{4};$$

$\overline{N_q^{(2)}} = \sum_{c''=1}^{(n+m)''} \frac{c''P'_{(c)}}{k_E''} = \sum_{c''=1}^{3} \frac{c''P''_{(c)}}{3}$;
Duration of waiting time for the customer in the queue for the i-component equals:
$\overline{W}_q^{(1)} = \frac{\overline{N}_q^{(1)}}{\lambda_1}$; $\overline{W}_q^{(2)} = \frac{\overline{N}_q^{(2)}}{\lambda_2}$.
Customer service time in the QS:
$\bar{t}_{tsq} = \frac{(\lambda_1 \bar{t}_s^{(1)} + \lambda_2 \bar{t}_s^{(2)})}{(\lambda_1+\lambda_2)}$;
$\bar{t}_s^{(1)} = \frac{\overline{N}^{(1)}}{\lambda_1}$; $\bar{t}_s^{(2)} = \frac{\overline{N}^{(2)}}{\lambda_2}$;

Duration of waiting for customers in QS queues: $\overline{w}_{wqs} = \frac{(\lambda_1 \overline{w}_q^{(1)} + \lambda_2 \overline{w}_q^{(2)})}{(\lambda_1+\lambda_2)}$;

Probability of QS failure: $P_{fl}^{qs} = \frac{(\lambda_1 P_{fl_1} + \lambda_2 P_{fl_2})}{(\lambda_1+\lambda_2)}$.

When applying the proposed mathematical models, it is advisable to consider the following:

- in multi-component QSs, the performance of any component decreases compared to a single-component system at the same rates of service stages. With the same values of the parameters of each component of the QS, the performance of multi-component and single-component systems will be the same;
- if one of the components is a QS with a queue, and the second component is a QS with failures, then the QS with a queue has a higher performance, simultaneously reducing the performance of the second component;
- with small values $(0 \le P_S \le 0.1)$, the impact on the system as a whole or on a separate component of the intensities of customers leaving the system during the service period and being in the queue is insignificant. When these intensities change, the P_S value will fluctuate relative to its average value.

Example 2. We will conduct a sensitivity study of the two-component QS mathematical model presented in Fig. 1. The simulation model of the two-component QS is presented in Fig. 2.

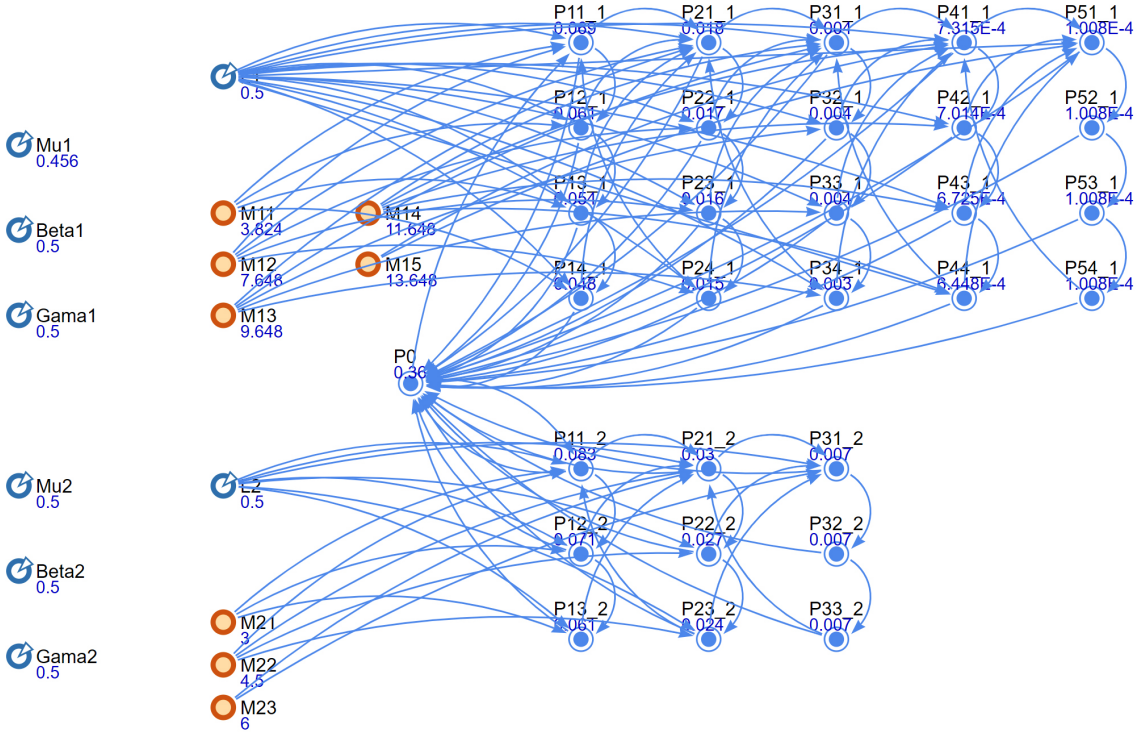


Figure 2: *Probability density of grain delivery time distribution at optimal sizes of the fleet of vehicles (trucks and ships)*

The model was implemented using System Dynamics computer simulation in the AnyLogic University Researcher environment.

3. Results

The results of experiments on the sensitivity of the model are presented in Table 1 and Fig. 3 - 10.

The model was implemented using System Dynamics computer simulation in the AnyLogic University Researcher environment.

A set of dynamic variables (Dynamic Variable) P11_1−P54_1 and P11_2−P33_2 were used to establish and fix the probabilities of the QS states according to the structure of the studied QS. The usual double (Variable) type variables M11−M13 and M21−M23 were used as calculation variables of the proposed QS. The initial system parameters $\lambda_1, \lambda_2, \mu', \mu'', \gamma_1', \gamma_1'', ,_1', ,_1''$ are described in detail above.

The probabilities of the QS states are calculated according to the general principle of solving systems of the Kolmogorov equations and the system dynamics format in the AnyLogic University Researcher environment, according to Fig. 1:

«P11_1 = (L1*P0 + M12*P24_1)/(L1+M11);
P12_1 = (M11*P11_1)/(L1+M11);
P13_1 = (M11*P12_1)/(L1+M11);
P14_1 = M11*P13_1/(L1+M11);
P21_1 = (L1*(P11_1+P12_1+P13_1+P14_1)+M13*P34_1)/(L1+M12);
P22_1 = M12*P21_1/(L1+M12);
P23_1 = M12*P22_1/(L1+M12);
P24_1 = M12*P23_1/(L1+M12);
P31_1 = (L1*(P21_1+P22_1+P23_1+P24_1)+M14*P44_1)/(L1+M13);
P32_1 = M13*P31_1/(L1+M13);
P33_1 = M13*P32_1/(L1+M13);
P34_1 = (M13*P33_1)/(L1+M13);
P41_1 = (L1*(P31_1+P32_1+P33_1+P34_1)+M15*P54_1)/(L1 + M14);
P42_1 = M14*P41_1/(L1+M14);
P43_1 = M14*P42_1/(L1+M14);
P44_1 = (M14*P43_1)/(L1+M14);
P51_1 = L1*(P41_1+P42_1+P43_1+P44_1)/M15;
P52_1 = P51_1;
P53_1 = P52_1;
P54_1 = P53_1;
P11_2 = (L2*P0+M22*P23_2)/(L2+M21);
P12_2 = (M21*P11_2)/(L2+M21);
P13_2 = M21*P12_2/(L2+M21);
P21_2 = (L2*(P11_2+P12_2+P13_2)+M23*P33_2)/(L2+M22);
P22_2 = M22*P21_2/(L2+M22);
P23_2 = (M22*P22_2)/(L2+M22);
P31_2 = L2*(P21_2+P22_2+P23_2)/M23;
P32_2 = P31_2;
P33_2 = P32_2;
P0 = 1-(P11_1 + P12_1 + P13_1 + P14_1 + P21_1 + P22_1 + P23_1 + P24_1 + P31_1 + P32_1 + P33_1 + P34_1 + P41_1 + P42_1 + P43_1 + P44_1 + P51_1 + P52_1 + P53_1 + P54_1 + P11_2 + P12_2 + P13_2 + P21_2 + P22_2 + P23_2 + P31_2 + P32_2 + P33_2);»

The model sensitivity experiment is implemented as a cyclical gradual change of one of the selected initial parameters by Java software code [32], [33], [34]:

«double step = 0.01;
Mu1 += step;
double dataVariant = Mu1;
double m_1 = 4*Mu1, b_1 = 4*Beta1, g_1 = 4*Gama1,
m_2 = 3*Mu2, b_2 = 3*Beta2, g_2 = 3*Gama2;
M11 = m_1 + b_1; M12 = 2*(m_1 + b_1); M13 = M12 + g_1;

```
M14 = M12 + 2*g_1; M15 = M12 + 3*g_1; M21 = m_2 + b_2;
M22 = M21 + g_2; M23 = M21 + 2*g_2;
double Pc_2 = P11_2 + P12_2 + P13_2 + P21_2 + P22_2 + P23_2 + P31_2 + P32_2 + P33_2,
Pc_1 = P11_1 + P12_1 + P13_1 + P14_1 + P21_1 + P22_1 + P23_1 + P24_1 + P31_1 + P32_1 + P33_1
+ P43_1 + P41_1 + P42_1 + P43_1 + P44_1 + P51_1 + P52_1 + P53_1 + P54_1,
P20_1 = P12_1 + 2*P13_1 + 3*P14_1, P30_1 = P22_1 + 2*P23_1 + 3*P24_1, P40_1 = P32_1 + 2*P33_1
+ 3*P43_1, P50_1 = P42_1 + 2*P43_1 + 3*P44_1,
P20_2 = P12_2 + 2*P13_2, P30_2 = P22_2 + 2*P23_2,
P1_2 = P11_2 + P12_2 + P13_2, P2_2 = P21_2 + P22_2 + P23_2, P3_2 = P31_2 + P32_2 + P33_2,
P1_1 = P11_1 + P12_1 + P13_1 + P14_1, P2_1 = P21_1 + P22_1 + P23_1 + P24_1, P3_1 = P31_1 +
P32_1 + P33_1 + P43_1,
P4_1 = P41_1 + P42_1 + P43_1 + P44_1, P5_1 = P51_1 + P52_1 + P53_1 + P54_1,
Ppok_q_1 = (Beta1 / L1) * ((1*P3_1 + 2*P4_1 + 3*P5_1)/4),
Ppok_serv_1 = (Beta1 / L1) *((P1_1 + 2*(P2_1 + P3_1 + P4_1 + P5_1))/4),
Pfalue_1 = P5_1 - ((P20_1 + P30_1 + P40_1 + P50_1)/4),
Pvrtat_1 = Pfalue_1 + Ppok_serv_1 + Ppok_q_1,
Pser_1 = 1 - Pvrtat_1 - (Pc_2),
Ppok_q_2 = (Beta2 / L2) * ((1*P2_2 + 2*P3_2)/3),
Ppok_serv_2 = (Beta2 / L2) * (Pc_2 /3),
Pfalue_2 = P3_2 - ((P20_2 + P30_2) / 3),
Pvrtat_2 = Pfalue_2 + Ppok_serv_2 + Ppok_q_2,
Pser_2 = 1- Pvrtat_2 - Pc_1;
dataset_Pserv_1.add(dataVariant, Pser_1);
dataset_Pserv_2.add(dataVariant, Pser_2);»
```

The generalized characteristics of the QS of the system components are the service probabilities $P_s^{/}$ and $P_s^{//}$.

These characteristics include the initial parameters and some other parameters of the QS components, the calculation formulas of which are presented above.

The initial parameters of the experiments and the results of their implementation are presented in Table 1.

Graphs of dependences of service probabilities $P_s^{/}$ and $P_s^{//}$ on the initial parameters of the model are presented in fig. 3-10.

In the calculation example, the boundary value of the service probabilities of QS components $P_{b.v_s} = 0.6$ is set, i.e. $P_s^{/}$ and $P_s^{//}$ must not be less than 0.6 at the same time, provided that the components serve requirements with the same priorities.

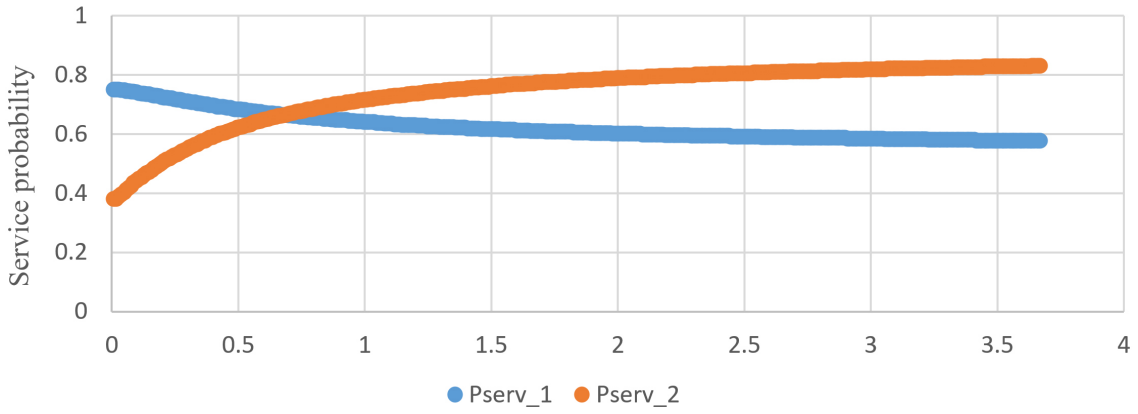


Figure 3: *Graph of the dependence of the probabilities $P_s^{/}$ and $P_s^{//}$ on the parameter $\mu^{/}$*

From the graphs of the dependences of the probabilities $P_s^{/}$ and $P_s^{//}$ on the parameter $\mu^{/}$

Table 1: *The results of sensitivity experiments for the mathematical model of the QS components functioning*

Numbers of graphs of depend-es	Initial parameters of the model		Results of sensitivity experiments		
	constant	variables $[range]$	Ranges of changes in probabil-es	Values of service probabil-es	Values of variable param-s
3	$\lambda_1 = \lambda_2 = \mu'' = \gamma_1' = \gamma_2'' = \beta_1' = \beta_2'' = 0.5$	μ' [0.01...2.5]	P_S' [0.59...0.75] P_S'' [0.38...0.8]	$P_S' = 0.15$ $P_S'' = 0.8$ $P_S = P_S''$ =0.69	$\mu' = 0.01$ $\mu' = 2.5$ $\mu' = 0.69$
4	$\lambda_1 = \lambda_2 = \mu' = \mu'' = \gamma_1' = \gamma_2'' = \beta_2'' = 0.5$	β_1' [0.01...2.5]	P_S' [0.49...0.98] P_S'' [0.3...0.8]	$P_S' = 0.98$ $P_S'' = 0.8$ $P_S = P_S''$ =0.65	$\beta_1' = 0.01$ $\beta_1' = 2.5$ $\beta_1' = 0.62$
5	$\lambda_1 = \lambda_2 = \mu'' = \gamma_2'' = \beta_1' = = \beta_2'' = 0.5$	γ_1' [0.01...2.5]	P_S' [0.679...0.683] P_S'' [0.618...0.626]	$P_S' = 0.58$ P_S'' =0.626	$\gamma_1' = 2.5$ $\gamma_2' = 2.5$
6	$\lambda_1 = \lambda_2 = \gamma_1' = = \gamma_2'' = \beta_1' = = \beta_2'' = 0.5$	β_1' [0.01...2.5]	P_S' [0.4350.902] P_S'' [0.5970.593]	$P_S' = 0.902$ $P_S'' = 0.622$ $P_S = P_S''$ =0.62	$\beta_2' = 2.5$ $\beta_2' = 0.47$ $\beta_2' = 0.34$
7	$\lambda_1 = \lambda_2 = \mu'' = = \gamma_1'' = \gamma_2' = \beta_1' = 0.5$	β_2'' [0.01...2.5]	P_S' [0.4350.902] P_S'' [0.8770.434]	$P_S' = 0.902$ $P_S'' = 0.877$ $P_S = P_S''$ =0.651	$\beta_2' = 2.5$ $\beta_2' = 0.01$ $\beta_2' = 0.41$
8	$\lambda_1 = \lambda_2 = \mu'' = = \gamma_1'' = \beta_1' = 0.5$	γ_2'' [0.01...2.5]	P_S' [0.6260.73] P_S'' [0.5580.652]	$P_S' = 0.73$ $P_S'' = 0.652$	$\gamma_1' = 2.5$ $\gamma_2' = 2.5$
9	$\lambda_2 = \mu'' = \gamma_1' = \gamma_2'' = \beta_1' = \beta_2'' = = 0.5$	λ_1' [0.01...2.5]	P_S' [0.499...0.952 ...0.879] P_S'' [0.884...0.8]	$P_S' = 0.499$ $P_S' = 0.952$ $P_S' = 0.873$ $P_S'' = 0.884$ $P_S'' = 0.01$ $P_S = P_S'' = = 0.657$	$\lambda_1' = 0.4$ $\lambda_1' = 1.65$ $\lambda_1' = 2.5$ $\lambda_1' = 0.01$ $\lambda_1' = 2.5$ $\lambda_1' = 0.45$
10	$\lambda_1 = \mu'' = \gamma_1' = \gamma_2'' = \beta_1' = \beta_2'' = = 0.5$	λ_2' [0.01...2.5]	P_S' [1.0...0.047] P_S'' [0.50...0.68 ...0.049]	$P_S' = 1.0$ $P_S' = 0.047$ $P_S' = 0.50$ $P_S'' = 0.68$ $P_S'' = 0.049$ $P_S = P_S'' = = 0.64$	$\lambda_1' = 0.01$ $\lambda_1' = 2.5$ $\lambda_1' = 0.06$ $\lambda_1' = 0.9$ $\lambda_1' = 2.5$ $\lambda_1' = 0.57$

presented in Fig. 3, it can be seen that when $\mu' = 0.01$, then $P_s' = 0.75$, $P_s'' = 0.378$, while at

$\mu^{/} = 2.5$ the probabilities are $P_s^{//} = 0.8$ and $P_s^{/} = 0.53$, i.e. do not satisfy the boundary condition. At $\mu^{/} = 0.69$ the probabilities are $P_s^{/} = P_s^{//} = 0.66$, which satisfies the boundary condition.

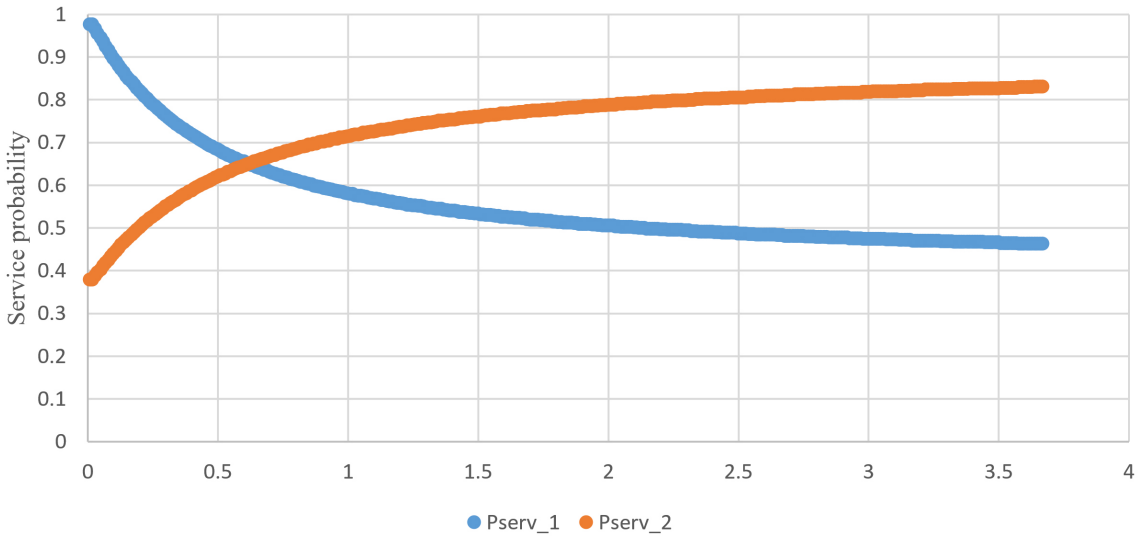


Figure 4: *Graphs of the dependence of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $\beta_1^{/}$*

From Fig. 4, which shows the graphs of the dependences of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $\beta_1^{/}$, it is possible to investigate that with $\beta_1^{/} = 0.01$ the probabilities $P_s^{/} = 0.98$ and $P_s^{//} = 0.378$ with $\beta_1^{/} = 2.5$ probabilities $P_s^{/} = 0.48$ and $P_s^{//} = 0.8$, i.e. do not satisfy the boundary condition. At $\beta_1^{/} = 0.62$ probabilities $P_s^{/} = P_s^{//} = 0.65$, which is a satisfactory result.

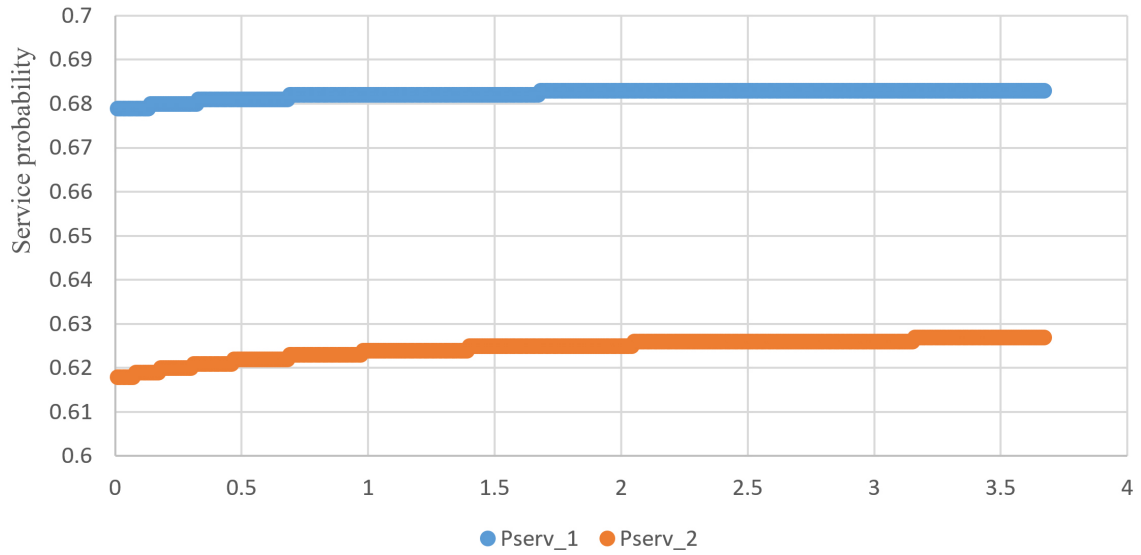


Figure 5: *Graphs of the dependence of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $\gamma_1^{/}$*

From those presented in fig. 5 graphs of the dependence of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $-\gamma_1^{/}$, it is possible to investigate that at $\gamma_1^{/} = 2.5$ the probabilities $P_s^{/} = 0.68$ and $P_s^{//} = 0.626$, which satisfies the boundary condition.

In Fig. 6 are presented the graphs of dependences of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $\mu^{//}$, from which it can be seen that at $\mu^{//} = 2.5$ the probabilities $P_s^{/} = 0.902$ and $P_s^{//} = 0.593$. This option does not meet the boundary condition. At $\mu^{//} = 0.47$ $P_s^{/} = 0.668$ and $P_s^{//} =' = 0.626$ at $\mu^{//} = 0.34_s^{/} = P_s^{//} = 0.62$. These options meet the boundary condition. The option with

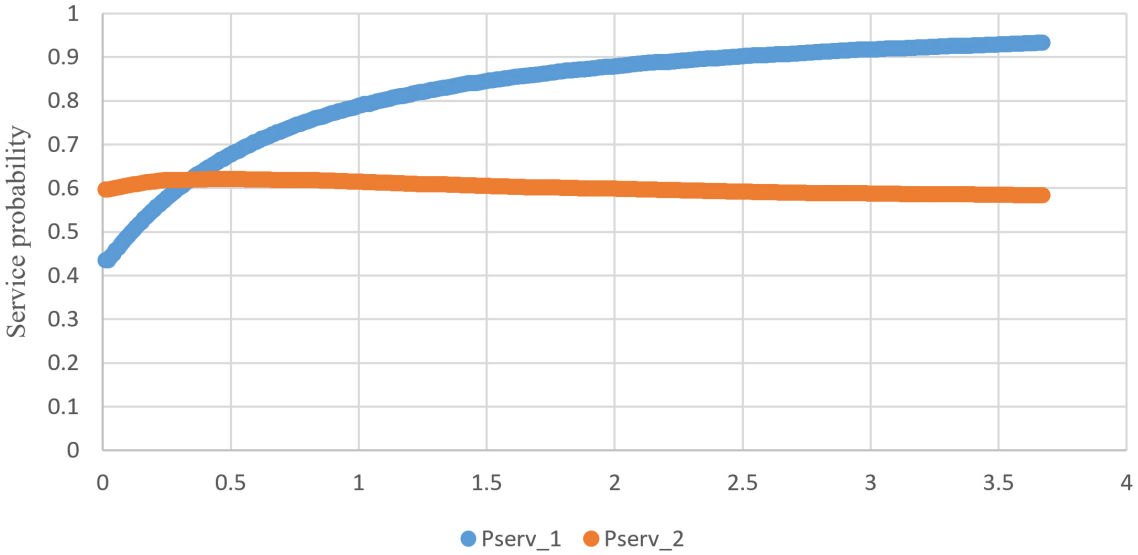


Figure 6: *Graphs of the dependence of the probabilities* $P_s^{/}, P_s^{//}$ *on the parameter* $\mu^{//}$

$\mu^{//} = 0.47$ may be preferable, because with $\mu^{//} = 0.47$ at $t_{(s_2)} = 2.18$, and with $\mu^{//} = 0.34$ at $t_{(s_2)} = 2.94$.

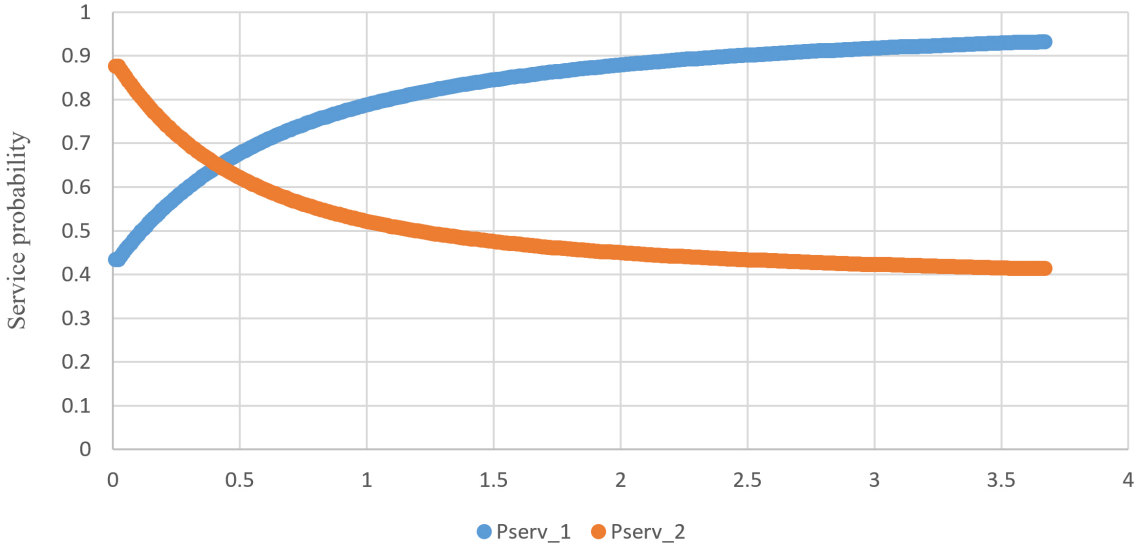


Figure 7: *Graphs of the dependence of the probabilities* $P_s^{/}, P_s^{//}$ *on the parameter* $\beta_2^{//}$

From the graphs of the dependences of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $\beta_2^{//}$, presented in Fig. 7, it can be seen that at $\beta_2^{//} = 0.41$, $P_s^{/} = P_s^{//} = 0.651$. This option meets the boundary condition. Options at $\beta_2^{//} = 2.5$, $P_s^{/} = 0.902$ and $P_s^{//} = 0.593$ and at $\beta_2^{//} = 0.01$, $P_s^{//} = 0.432$ and $P_s^{//} = 0.877$ do not meet the boundary condition.

$\mu^{//} = 0.34$ at $t_{(s_2)} = 2.94$.

From Fig. 8, which shows graphs of the dependence of the probabilities $P_s^{/}, P_s^{//}$ on the parameter $\gamma_2^{//}$, it can be determined that with $\gamma_2^{//} = 2.5$, $P_s^{/} = 0.78$ and $P_s^{//} = 0.652$, which corresponds to the boundary condition.

From those presented in Fig. 9 graphs of the dependence of the probabilities $P_s^{/}, P_s^{//}$ on the parameter λ_1, the following conclusions can be drawn: at $\lambda_1 = 0.45, P_s^{/}, P_s^{//} = 0.652$, which satisfies the boundary condition.

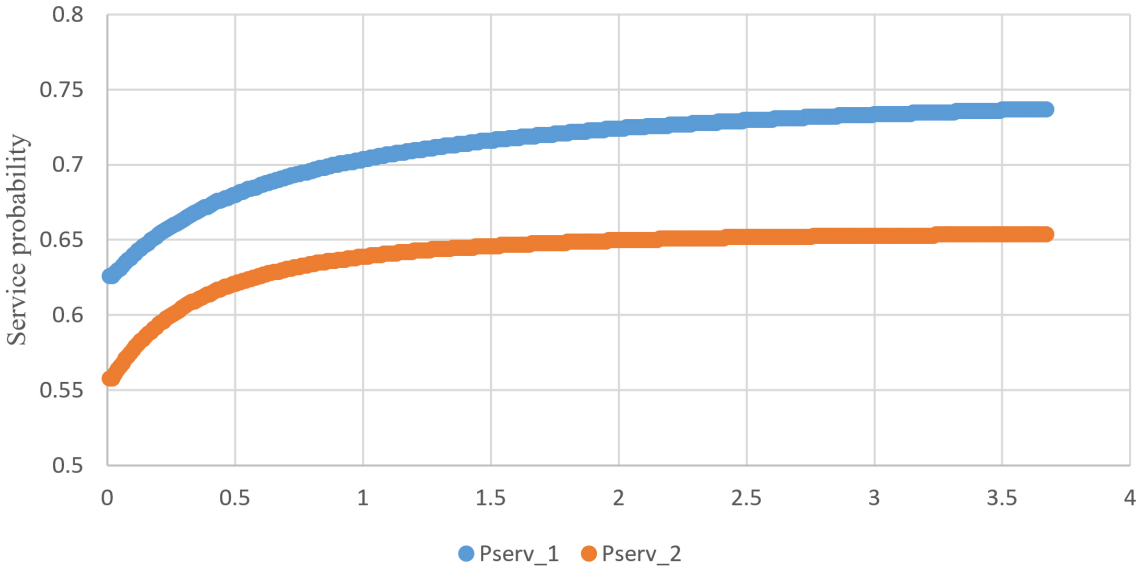


Figure 8: *Graphs of the dependence of the probabilities* $P_s^{/}, P_s^{//}$ *on the parameter* $\gamma_2^{//}$

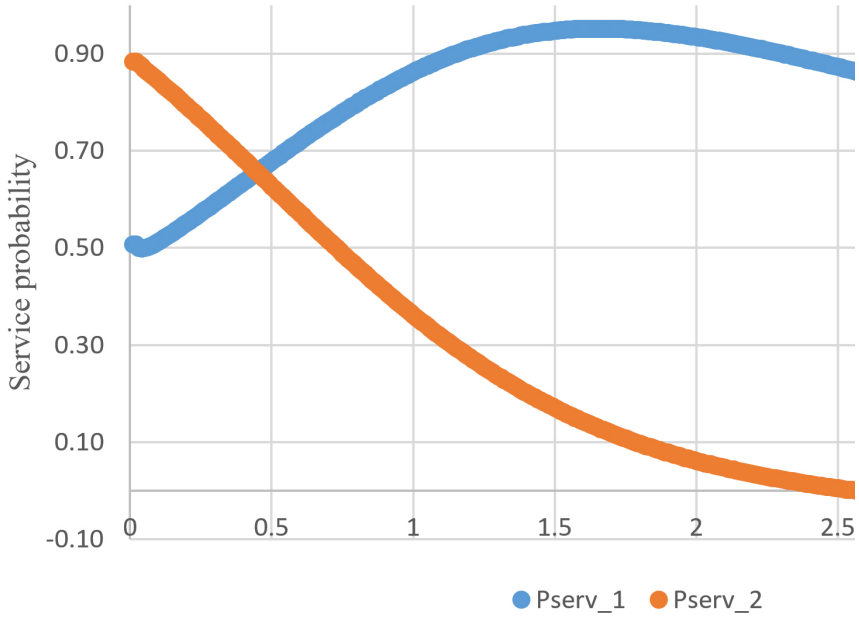


Figure 9: *Graphs of the dependence of the probabilities* $P_s^{/}, P_s^{//}$ *on the parameter* λ_1

Other options do not meet the boundary condition

at $\lambda_1 = 0.45, P_s^{/} = 0.506$ and $P_s^{//} = 0.884$, as at $\lambda_1 = 0.04 P_s^{/} = 0.499$ and $P_s^{//} = 0.87$; at $\lambda_1 = 1.65, P_s^{/} = 0.952$ and $P_s^{//} = 0.13$; at $\lambda_1 = 2.5, P_s^{/} = 0.873$ and $P_s^{//} = 0.01$.

From those presented in Fig. 10 graphs of the dependence of the probabilities $P_s^{/}, P_s^{//}$ on the parameter λ_2 it can be seen that the boundary condition is met by the variant with $\lambda_2 = 0.57, P_s^{/} = P_s^{//} = 0.64$. Other options do not meet the boundary condition, namely: at $\lambda_2 = 0.01, P_s^{/} = 1.0$ and $P_s^{//} = 0.51$, at $\lambda_2 = 0.06 P_s^{/} = 0.977$ and $P_s^{//} = 0.50$; at $\lambda_2 = 0.9, P_s^{/} = 0.432$ and $P_s^{//} = 0.68$; at $\lambda_2 = 2.5, P_s^{/} = 0.47$ and $P_s^{//} = 0.49$. We summarize the obtained results in the Table 2 and determine P_s of the two-component QS from the formula:

$P_s = 1 - (1 - P_s^{/})(1 - P_s^{//})$.

From the Table 2, it can be seen that the determined results of the initial parameters, which

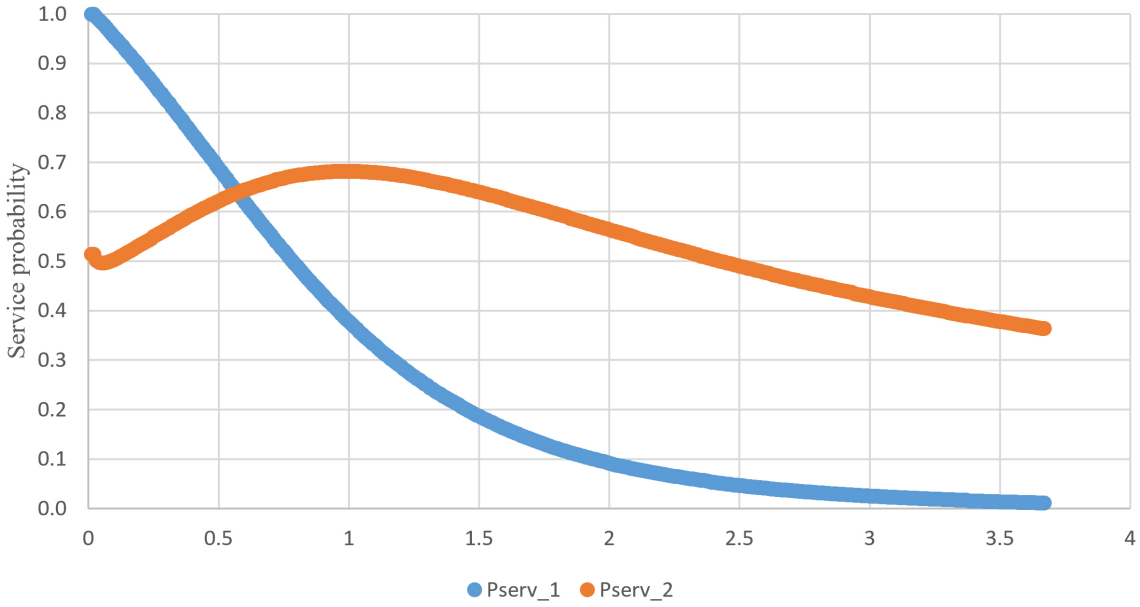


Figure 10: *Graphs of the dependence of the probabilities $P_s^{/}$, $P_s^{//}$ on the parameterλ_1*

Table 2: *Results of the sensitivity experiments analysis*

Probabilities of service	λ_1 0.45	λ_2 0.55	λ_1 0.43	λ_2 0.57	$\mu^{/}$ 0.69	$\gamma_1^{/}$ 2.5	$\beta_1^{/}$ 0.62	$\mu^{//}$ 0.34	$\gamma_2^{//}$ 2.5	$\beta_2^{//}$ 0.41
$P_s^{/}$	0.66	0.65	0.65	0.64	0.63	0.62	0.62	0.62	0.73	0.65
$P_s^{//}$	0.66	0.63	0.67	0.64	0.63	0.63	0.62	0.62	0.65	0.65
P_s	0.88	0.87	0.87	0.87	0.86	0.88	0.86	0.86	0.91	0.88

correspond to the boundary condition, make it possible to obtain the most acceptable values of the probability values $P_s^{/}$, $P_s^{//}$ and P_s. In turn, the above-mentioned initial parameters reflect and quantitatively characterize both the external conditions of the system's functioning (λ) and its internal capabilities and limitations (μ, β, γ) in response to changes in external conditions, including force majeure circumstances.

4. Conclusions

The new energy planning and management requirements should support innovation in aircraft designs and their operation and maintenance. They must allow for the smooth integration of new technologies into the air operations domain. Therefore, the term 'energy' should be used together with the term 'fuel' ('energy/fuel', for example), wherever appropriate, to accommodate operations with aircraft that use other energy sources than conventional hydrocarbon-based fuel. Also, the development of new production standards and safety and certification rules by regulators often lag behind technological development, bringing product development to a halt, including new onboard electrical equipment and technologies. Therefore, it is appropriate to include in the service system the possibility of the occurrence of force majeure circumstances and to consider the system with their impact on the results. The proposed theoretical approach consists in the fact that in a multi-component queueing system (QS), the value of the probability of serving a customer in a certain QS component can be determined taking into account that the second component contains the sum of all probabilities of the "enlarged" states of other components of the QS.

Based on this theoretical provision, the modelling of vehicle maintenance and repair processes,

using the example of an aviation repair enterprise, allows to determine the necessary initial parameters of the system components as interacting QSs and to obtain the largest values of the probabilities of servicing the arriving customers in these components and the system as a whole, which provide an acceptable level of its reliability.

When studying real production, logistics, and other systems for which the mathematical apparatus of queueing theory is adequate, the necessary initial mathematical parameters of the system components must be expressed through physical parameters (flows of vehicles or other objects requiring maintenance, performance of equipment for various types of work, production tasks, time constraints, etc.), which will make it possible to optimize specific technologies and enterprises.

When modelling QS processes, non-standard system dynamics solutions were proposed in the AnyLogic University Researcher environment, which allowed to:

- solve a multi-rank system of Kolmogorov equations;
- implement multi-iterative sensitivity experiments with the initial parameters of the QS;
- obtain experimental dependences of the influence of all key parameters on QS indicators, in particular, service probabilities.

The described mathematical apparatus and modelling tools have shown their relevance to real processes and can be applied to improve the performance of multi-component and multi-phase queueing systems, which reflect the technological processes occurring in real production, transport-logistics and other systems intended for operation, maintenance and repair of technical equipment of various nature.

New hazards concerning the operation and maintenance of the new onboard electric equipment and systems should be investigated to complete the safety analysis of new clean aviation increasing their sustainability in the future. They will connect with normal flight operation of the new aircraft (their separate systems and equipment), so as with their production, maintenance, and repair. However, the proposed mathematical apparatus must cover new links and allow safe and sustainable solutions.

References

[1] Update on uk sustainable aviation fuel potential | sustainable aviation URL https://www.sustainableaviation.co.uk/news/update-on-uk-sustainable-aviation-fuel-potential/

[2] Zaporozhets O, Isaienko V and Synylo K 2020 *Energy* **211** 118814 ISSN 0360-5442

[3] Wheeler P and Bozhko S 2014 *IEEE Electrification Magazine* **2**(4) 6–12 ISSN 23255889

[4] Baharozu E, Soykan G and Ozerdem M B 2017 *Energy* **140** 1368–1377 ISSN 0360-5442

[5] Alruwaili M and Cipcigan L 2022 *Electric Power Systems Research* **211** 108242 ISSN 0378-7796

[6] Yilmaz N and Atmanli A 2017 *Energy* **140** 1378–1386 ISSN 0360-5442

[7] Dincer I and Acar C 2016 *International Journal of Sustainable Aviation* **2**(1) 74 ISSN 2050-0467

[8] Aviation rules of ukraine maintenance of the identity of aircraft and aviation products, components and possession and praise of the organization and personnel received before the cessation of these commands.´ | order of the state aviation service of ukraine dated march 6, 2019 no. 256 URL https://avia.gov.ua/wp-content/uploads/2019/08/Release-to-service-of-the-Ukrainian-registered-aircraft.pdf

[9] Rules for engineering and aviation security of the sovereign aviation of ukraine. order of the ministry of defense of ukraine dated 07/05/2016 no. 343. registered with the ministry of justice of ukraine on september 08, 2016. for no. 1101/29231 from changes introduced by order of the Ministry of Defense of Ukraine No. 223 dated 08/03/2021 URL https://zakon.rada.gov.ua/laws/show/z1101-16#Text

[10] Tamargazin OA M O 2013 *Scientific technologies* (1) 33–36 ISSN 2075-0781 URL http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILA=&2_S21STR=Nt_2013_1_9

[11] Khromchenko VA Voznyuk MM P V 2022 *Collection of scientific works of the State Research Institute of Aviation* (18(25)) 110–116 ISSN 2786-4839 URL https://znp.dndia.org.ua/index.php/znp/article/view/43

[12] Zhdanyuk MM Cherednik MM M S M Y S S 2021 *Collection of scientific works of the State Research Institute of Testing and Certification of Weapons and Military Equipment* **10**(4) 45–55 ISSN 2706-7386 URL https://dndivsovt.com/index.php/journal/article/view/129

[13] S V Gaevsky S M Balakireva D V K V O Y A 2020 *Control systems, navigation and communication* **1**(59) 15–20 ISSN 2073-7394 URL https://journals.nupp.edu.ua/sunz/article/view/1774

[14] Medinsky D 2021 *Technical Sciences* **1**(59) 113–122 ISSN 2073-7394 URL chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.tech.vernadskyjournals.in.ua/journals/2021/1_2021/part_2/21.pdf

[15] Shevchuk D M D 2020 *Technical Sciences* **70**(31) 254–260 ISSN 2073-7394 URL chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.tech.vernadskyjournals.in.ua/journals/2020/5_2020/43.pdf

[16] SM Novichonok OM Babich I T 2020 *Armor systems and military equipment* (2(62)) 24–34 ISSN 1997-9568 URL https://journal-hnups.com.ua/index.php/soivt/article/view/328

[17] Ye Ilenko M Sushak P S 2020 *Armor systems and military equipment* (3(65)) 43–49 ISSN 2073-7378 URL https://journal-hnups.com.ua/index.php/zhups/article/view/350

[18] M Sushak M Derevianko S F 2022 *Collection of scientific works of the State Research Institute of Aviation* (18(25)) 187–196 ISSN 2786-4839 URL https://znp.dndia.org.ua/index.php/znp/article/view/55

[19] V Grishin O S 2022 *Collection of scientific works of the State Research Institute of Aviation* (18(25)) 134–140 ISSN 2786-4839 URL https://znp.dndia.org.ua/index.php/znp/article/view/47

[20] Tsybulnyk S and Okhota B 2018 *Bulletin of the Kyiv Polytechnic Institute. Instrumentation series* **0**(55(1)) 93–100 ISSN 2663-3450 URL http://visnykpb.kpi.ua/article/view/135844

[21] Matviychuk VA Veselovska NR S S 2018 *Vinnytsia National Agrarian University* **0**(1) 93–100 ISSN 978-966-949-855-7 URL http://repository.vsau.org/getfile.php/29057.pdf

[22] Olexander Subochev Yuliya Poloz V M O S 2019 *Foreign interdepartmental scientific and technical collection Design and operation of agricultural machines´* **49**(49) 221–232 ISSN 2414-3820 URL chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://zbirniksgm.kntu.kr.ua/pdf/49/29.pdf

[23] Pavlenko MA Rublyova RI T A 2020 *Collection of scientific works ›“* 89–92 URL https://ojs.ukrlogos.in.ua/index.php/logos/article/view/4630

[24] Bogomiya VI Belobrova TA D S 2019 *Scientific technologies* (1) 82–87 ISSN 2075-0781 URL http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILA=&2_S21STR=Nt_2019_1_13

[25] OO E 2020 *Technical Sciences* (70) 195–201 ISSN 2075-0781

[26] Bondar BE Ochkasov AB B E G T O M 2018 *Technical Sciences* (31(70)) 195–201 ISSN - URL http://eadnurt.diit.edu.ua>bistream.pdf

[27] Guzenko VL Mironov EA S O 2014 *Technical Sciences* (6(22)) 195–201 ISSN 2075-0781 URL http://iea.gostimjo.ru>_66_11.pdf

[28] MD Katsman OI Zaporozhets V V O 2021 *Kiev FOP: Luk™yanenko V.V.* (1) 380 ISSN 978-617-7609-61-1 URL -

[29] Katsman M D, Myronenko V K, Matsiuk V I and Lapin P V 2021 *Reliability: Theory and Applications* **16**(1) 71–80 ISSN 19322321 URL http://www.gnedenko.net/Journal/2021/012021/RTA_1_2021-06.pdf

[30] Katsman M D, Matsiuk V I and Myronenko V K 2023 *Reliability: Theory Applications* **June 2**(73) 167–179 URL https://doi.org/10.24412/1932-2321-2023-273-167-179

[31] Katsman M D, Matsiuk V I and Myronenko V K 2023 *Reliability: Theory and Applications* **18**(3) 751–767 ISSN 19322321

[32] Matsiuk V, Ilchenko N, Pryimuk O, Kochubei D and Prokhorchenko A 2022 *AIP Conference Proceedings* **2557**(1) ISSN 15517616 URL /aip/acp/article/2557/1/080003/2829501/Risk-assessment-of-transport-processes-by-agent

[33] Matsiuk V, Myronenko V, Horoshko V, Prokhorchenko A, Hrushevska T, Shcherbyna R, Matsiuk N, Khokhlacheva J, Biziuk I and Tymchenko N 2019 *Eastern-European Journal of Enterprise Technologies* URL http://www.scopus.com/inward/record.url?eid=2-s2.0-85065018277&partnerID=MN8TOARS

[34] Prokhorchenko A, Parkhomenko L, Kyman A, Matsiuk V and Stepanova J 2019 Improvement of the technology of accelerated passage of low-capacity car traffic on the basis of scheduling of grouped trains of operational purpose vol 149 pp 86–94 ISSN 18770509 URL http://www.scopus.com/inward/record.url?eid=2-s2.0-85063811295&partnerID=MN8TOARS

PROFIT ANALYSIS OF REPAIRABLE COLD STANDBY SYSTEM SUBJECT TO REBOOT FACILITY UNDER REFRESHMENTS

[1]Ajay Kumar and [2]Ashish Sharma

•

[1]SOET, Raffles University, Neemrana, Rajasthan
[2]Department of Pharmacy, Sushant University, Gurugram
ajaykumar.soet@rafflesuniversity.edu.in, ashishlpu10@gmail.com

Abstract

This paper relates to the reliability measures analysis of two identical unit system with reboot facility. Initially, one unit of the system is in operative mode and another unit is kept in cold standby mode. A technician is always available with the system to perform repairing and rebooting activities. Here, the system operative unit failed in safe mode and unsafe mode. During unsafe failure, repair activity cannot be done immediately but first rebooting is done to transform unsafe failure into safe failure, and then repair activity is performed as usual. Sometimes, the technician needs refreshments due to continuous work and provides better services after taking refreshments. The unit works like a new one after repair. The failure time of the unit in safe mode, unsafe mode and technician refreshment request time are assumed to be general while the repair time of the unit, rebooting delay time and technician refreshment time are taken as exponential. Reliability measures such as mean time to system failure, availability of the system, busy period of the repairman, the expected number of visits by the technician and profit values are calculated using tables.

Keywords: Availability, cold standby, regenerative point, rebooting, and refreshment.

I. Introduction

In daily life, there are many situations such as the breakdown of the unit that caused machine failure. One way to avoid loss and increase the reliability of the system is to use the cold standby facility. With the occurrence of the complexity of machines and advancements in industrial sectors or organizations, the focus is on increasing the reliability and profit of the industry. The prominent point is that the designs and layout of complex machines or equipment should be so that it increases the reliability of the system and always tries to minimize the shortcomings responsible for its downtrends. Hence, designing a reliable system has become an essential step in almost every sector. So, the concept of rebooting is used to transform the unit from unsafe failure to safe failure. Sometimes, a technician is tired and needs refreshment. After taking refreshment, the repairman provides better service, and after getting the repair, the unit works like a new one. Many researchers such as Zhang and Wang [15] described a different unit repairable cold standby system that gives seniority to the operative unit. Hsu et al. [4] explored the standby system having reboot delay, general repair, switching failure, and unreliable repair facility. Jyh-Bin et al. [6] analyzed various reliability measures of a repairable system having standby switching failures and facility of reboot delay.

Dhall et al. [2] discussed the reliability of the similar unit stochastic approach under the repair and replacement of the failed unit subjected to inspection. Ke and Liu [7] examined the repairable system having a single server that identified the failed unit before repair and rebooting. Kumar and Goel [10] highlighted the two-unit cold standby redundant system subjected to inspection before repairing the failed unit and using the concept of preventive maintenance. Goel et al. [3] explained the performance of a cold standby redundant system with a server to inspect the failed unit before repair.

Temraz [14] evaluated the reliability measures for dependent system with load sharing and subject to degradation facility. Jain et al. [5] described the machine system as having online and standby units for system sustainability with server vacation, observing imperfect fault and its recovery using the reboot approach. Kumar and Jain [11] examined the reliability measures of a warm standby machine system having multiple components with recovery failure supported by the reboot process. Levitin et al. [12] evaluated the cold standby systems with elements exposed to shocks during operation and task transfers under preventive maintenance. Agrawal et al. [1] described the nature of the water treatment reverses osmosis plant using the regenerative point graphical technique. Sengar and Mangey [13] examined the complex manufacturing system subject to inspection facility using copula methodology. Kumar and Sharma [8] evaluated the availability and profit analysis of a repairable two unit cold standby system under refreshment using the regenerative point technique. Kumar et al. [9] explored the performance of two unit cold standby system under inspection and subject to refreshment facility.

II. System Assumptions

There are following system assumptions:

- The whole system has two identical units- first operative and second cold standby.
- The cold standby unit takes place when the operative unit stops functioning.
- A technician is always available to repair the failed unit.
- The failed unit behaves like a new one after repair.
- When the unit fails in unsafe mode then the reboot process is used to convert it to safe mode.
- Refreshment is offered to the technician to enhance his efficiency.
- Repair time, refreshment time and reboot delay time are exponentially distributed whereas times for failure of the unit in safe mode and unsafe mode and technician refreshment request are general.

III. System Notations

There are following system notations:

R	Collection of regenerative states S_i $(i = 0,1,2,3,9)$
$O / O(g) / Cs$	The system unit is operative and in normal mode / suitable good condition mode /cold standby mode
a / b	The probability that the cold standby unit is working/ not working
$\lambda / \lambda_1 / \mu$	The constant failure rate of the unit of the system in safe mode/rate with unit goes to unsafe mode/ rate by which the repairman needs refreshment
$g_1(t) / G_1(t)$	PDF/ CDF of the repair time of the unit
$f_1(t) / F_1(t)$	PDF/ CDF of refreshments time that restores freshness to the technician
$h_1(t) / H_1(t)$	PDF/ CDF of reboot delay time

$q_{r,s}(t)/Q_{r,s}(t)$	PDF/ CDF of first passage time from rth to sth regenerative state or sth failed state without halting in any other $S_i \in R$ in (0,1]
$M_r(t)$	Represents the probability of the system that it initially works $S_r \in R$ at a time (t) without moving through another state $S_i \in R$
$W_r(t)$	Probability that up to time (t) the server is busy at the state S_r without transit to another state $S_i \in R$ or before return to the same state through one or more non regenerative states
$\oplus/\otimes$	Laplace convolution / Laplace Stieltjes Convolution
$*/**/'$	Symbol for Laplace Transform/ Laplace Stieltjes Transform/ Function's derivative
○ / ● / ▭	Upstate/ regenerative state/ failed state

IV. State Descriptions

The system has up states as well as down states and these individual states are described in table 1:

Table 1: *State Descriptions*

States	Descriptions
S_0	It is a regenerative upstate with two units such that one is operative (O) and other is cold standby (Cs).
S_1	This regenerative upstate has two units such that one is failed under repair (*Fur*) and the other is in operative mode (O).
S_2	It is a regenerative upstate under refreshment facility (*sut*) where one unit is failed & waiting for repair (*Fwr*) and the other is in operative mode (O).
S_3	It is a regenerative down state and the system has two units such that one is failed under repair (*Fur*) and the other is failed and waiting for repair (*Fwr*).
S_4	It is a down state where one unit fails under repair (*FUR*) continuously from the prior state and the other unit is failed & waiting for repair (*Fwr*).
S_5	It is a down state that has two units under refreshment facility (*sut*) such that one is failed and waiting for repair (*Fwr*) and other is failed & waiting for repair (*FWR*) continuously from the previous state.
S_6	At this down state, the system has two units such that one is failed under repair (*FUR*) continuously from the previous state and the other unit is failed and waiting for repair (*FWR*) continually from the prior state.
S_7	This down state has two units under continuous refreshment facility (*SUT*) such that one is failed & waiting for repair (*Fwr*) and the other is failed & waiting for repair (*FWR*) continuously from the previous state.
S_8	This down state has two units under refreshment facility continuously from the prior state (*SUT*) in unsafe mode of failure of unit such that operative unit is failed under unsafe mode *F*(*uns*) and the other is failed and waiting for repair (*FWR*) continuously from the previous state.

S_9 This down state has two units such that the operative unit is failed under unsafe mode *F*(*uns*) and the other is in good condition.

S_{10} This down state has two units such that the operative unit is failed under unsafe mode *F*(*uns*) and the other is failed under repair (*FWR*) continuously from the previous state.

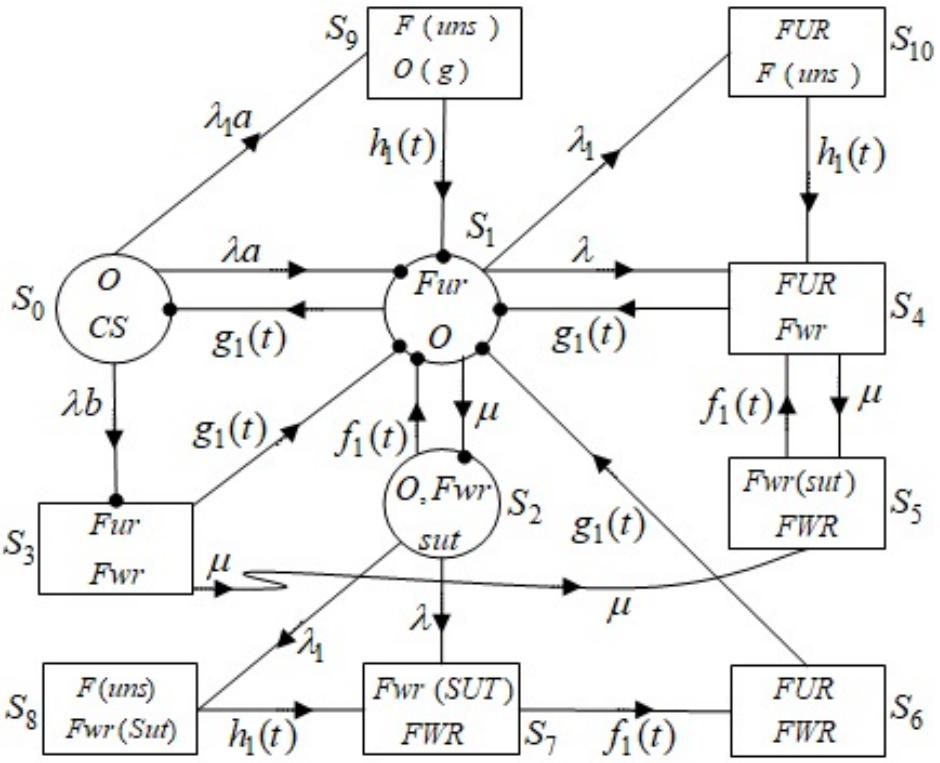


Figure 1: *State Transition Diagram*

V. Transition Probabilities

The transition probabilities are calculated using

$$f_1(t) = \theta e^{-\theta t},\ g_1(t) = \phi e^{-\phi t},\ h_1(t) = \xi e^{-\xi t} \tag{1}$$

$$p_{01} = \frac{\lambda a}{\lambda + \lambda_1 a},\ p_{03} = \frac{\lambda b}{\lambda + \lambda_1 a},\ p_{09} = \frac{\lambda_1 a}{\lambda + \lambda_1 a},\ p_{10} = \frac{\phi}{\phi + \lambda + \lambda_1 + \mu}$$

$$p_{12} = \frac{\mu}{\phi + \lambda + \lambda_1 + \mu},\ p_{14} = \frac{\lambda}{\phi + \lambda + \lambda_1 + \mu},\ p_{1,10} = \frac{\lambda_1}{\phi + \lambda + \lambda_1 + \mu},\ p_{21} = \frac{\theta}{\theta + \lambda + \lambda_1}$$

$$p_{27} = \frac{\lambda}{\theta + \lambda + \lambda_1},\ p_{28} = \frac{\lambda_1}{\theta + \lambda + \lambda_1},\ p_{31} = p_{41} = \frac{\phi}{\phi + \mu},\ p_{35} = p_{45} = \frac{\mu}{\phi + \mu}$$

$$p_{54} = p_{61} = p_{76} = p_{87} = p_{91} = p_{10,4} = 1 \tag{2}$$

It is smoothly verified that

$$p_{01} + p_{03} + p_{09} = 1\ \ p_{41} + p_{45} = 1$$

$$p_{10}+p_{12}+p_{14}+p_{1,10}=p_{10}+p_{12}+p_{11.4}+p_{11.(45)^n}+p_{11.(10,4)}+p_{11.10(45)^n}=1$$

$$p_{21}+p_{27}+p_{28}=p_{21}+p_{21.(76)}+p_{21.8(76)}=1,\ p_{31}+p_{35}=p_{31}+p_{31.(54)^n}=1 \tag{3}$$

VI. Mean Sojourn Time

In the cold standby redundant system, μ_i represents the mean sojourn time. Mathematically, time consumed by a system in a particular state is, $\mu_i=\sum_j m_{i,j}=\int_0^\infty P(T>t)\,dt$. Then

$$\mu_0=m_{01}+m_{03}+m_{09}=\int_0^\infty P(T>t)dt=\frac{1}{\lambda+\lambda_1 a}$$

$$\mu_1=m_{10}+m_{12}+m_{14}+m_{1,10}=\frac{1}{\phi+\lambda+\lambda_1+\mu},\ \mu_2=m_{21}+m_{27}+m_{28}=\frac{1}{\theta+\lambda+\lambda_1}$$

$$\mu_3=m_{31}+m_{35}=\frac{1}{\phi+\mu},\ \mu_4=m_{41}+m_{45}=\frac{1}{\phi+\mu}\quad \mu_5=\mu_7=\frac{1}{\theta},\ \mu_6=\frac{1}{\phi}$$

$$\mu_8=\mu_9=\mu_{10}=\frac{1}{\xi},\ \mu_3'=m_{31}+m_{31.(54)^n}=\frac{(\theta+\mu)}{\theta\phi}$$

$$\mu_1'=m_{10}+m_{12}+m_{11.4}+m_{11.(45)^n}+m_{11.(10,4)}+m_{11.10(45)^n}$$

$$=\frac{[\theta\phi(\xi+\lambda_1)+\xi(\theta+\mu)(\lambda+\lambda_1)]}{\theta\phi\xi(\phi+\lambda+\lambda_1+\mu)}$$

$$\mu_2'=m_{21}+m_{21.(76)}+m_{21.8(76)}=\frac{[\theta\phi(\xi+\lambda_1)+\xi(\lambda+\lambda_1)(\phi+\theta)]}{\theta\phi\xi(\theta+\lambda+\lambda_1)} \tag{4}$$

VII. Reliability Measures Evaluations

I. Mean Time to System Failure (MTSF)

Let the cumulative distribution function of the first elapsed time be $\varphi_i(t)$ from the regenerative state S_i to the failed state of the system. Treating the failed states as an absorbing state then the repetitive interface for $\varphi_i(t)$ being

$$\varphi_0(t)=Q_{09}(t)+Q_{03}(t)+Q_{01}(t)\otimes\varphi_1(t)$$

$$\varphi_1(t)=Q_{1,10}(t)+Q_{14}(t)+Q_{12}(t)\otimes\varphi_2(t)+Q_{10}(t)\otimes\varphi_0(t)$$

$$\varphi_2(t)=Q_{28}(t)+Q_{27}(t)+Q_{21}(t)\otimes\varphi_1(t) \tag{5}$$

Taking LST on the above equation (5) then get

$$R^*(s)=\frac{1-\varphi_0^{**}(s)}{s} \tag{6}$$

Now, system reliability is accessed by using the inverse LT on equation (6) such that

$$MTSF=\lim_{s\to 0}\frac{1-\varphi_0^{**}(s)}{s}=\frac{(1-p_{01}p_{10})\mu_2+p_{21}\mu_1+p_{21}p_{10}\mu_0}{(1-p_{01}p_{10}-p_{12}p_{21})} \tag{7}$$

II. Availability of the system

From the transition diagram, the system is available at the regenerative up states S_0, S_1 and S_2. Let $A_i(t)$ is the probability that the system is in upstate at time (t) specified that the system arrives at the regenerative state S_i at t = 0. Then the repetitive interface for $A_i(t)$ is

$$A_0(t) = q_{09}(t) \oplus A_9(t) + q_{03}(t) \oplus A_3(t) + q_{01}(t) \oplus A_1(t) + M_0(t)$$

$$A_1(t) = q_{12}(t) \oplus A_2(t) + q_{10}(t) \oplus A_0(t) + [q_{11.4}(t) + q_{11.(45)^n}(t) + q_{11.(10,4)}(t) + q_{11.10(45)^n}(t)] \oplus A_1(t) + M_1(t)$$

$$A_2(t) = [q_{21}(t) + q_{21.(76)}(t) + q_{21.8(76)}(t)] \oplus A_1(t) + M_2(t)$$

$$A_3(t) = [q_{31}(t) + q_{31.(54)^n}(t)] \oplus A_1(t)$$

$$A_9(t) = q_{91}(t) \oplus A_1(t) \tag{8}$$

Where, $M_0(t) = e^{-(\lambda+\lambda_1 a)t}$, $M_1(t) = \overline{G(t)}\, e^{-(\lambda+\lambda_1+\mu)t}$, $M_2(t) = \overline{F(t)}\, e^{-(\lambda+\lambda_1)t}$ (9)

Using LT of the above relation (8), then get

$$A_0(\infty) = \lim_{s\to 0} sA_0^*(s) = \frac{N_A}{D'} \tag{10}$$

where, $N_A = [\mu_0 p_{10} + \mu_1 + \mu_2 p_{12}]$

and $D' = [(\mu_0 + \mu_3' p_{03} - \mu_9 p_{09}) p_{10} + \mu_1' + \mu_2' p_{12}]$

III. Busy Period of the Server

From the transition diagram, it is clear that the technician is busy at states S_1, S_2 and S_3. Let $B_i(t)$ is the probability that the repairman is busy due to the repair of the failed unit at time 't' specified that the system arrives at the regenerative state S_i at t = 0. Then the repetitive interface for $B_i(t)$ is

$$B_0(t) = q_{09}(t) \oplus B_9(t) + q_{03}(t) \oplus B_3(t) + q_{01}(t) \oplus B_1(t)$$

$$B_1(t) = q_{12}(t) \oplus B_2(t) + q_{10}(t) \oplus B_0(t) + [q_{11.4}(t) + q_{11.(45)^n}(t) + q_{11.(10,4)}(t) + q_{11.10(45)^n}(t)] \oplus B_1(t) + W_1(t)$$

$$B_2(t) = [q_{21}(t) + q_{21.(76)}(t) + + q_{21.8(76)}(t)] \oplus B_1(t) + W_2(t)$$

$$B_3(t) = [q_{31}(t) + q_{31.(54)^n}(t)] \oplus B_1(t) + W_3(t)$$

$$B_9(t) = q_{91}(t) \oplus B_1(t) + W_9(t) \tag{11}$$

where, $W_1(t) = \overline{G_1(t)} e^{-(\lambda+\mu)t} + \overline{G_1(t)} \lambda e^{-(\lambda+\mu)t} \oplus \overline{G_1(t)} e^{-\mu t} + \ldots..$

$W_2(t) = [\{\lambda e^{-(\lambda+\lambda_1)t} \overline{F_1(t)} + \lambda_1 e^{-(\lambda+\lambda_1)t} \overline{F_1(t)} \oplus h_1(t)\} f_1(t)] \oplus \overline{G_1(t)}$

$W_3(t) = \overline{G_1(t)} e^{-\mu t} + \overline{G_1(t)} \mu e^{-\mu t} \oplus f_1(t) \oplus \overline{G_1(t)} e^{-\mu t} + \ldots..$

and $W_9(t) = \overline{H_1(t)}$

Using LT on the above relations (11) then get

$$B_0^R = \lim_{s\to 0} sB_0^*(s) = \frac{N_B}{D'}$$

Where, $N_B = W_1^*(0) + W_2^*(0) p_{12} + (W_3^*(0) p_{03} - W_9^*(0) p_{09}) p_{10}$ (12)

and D' is formerly declared.

IV. Estimated number of visits made by the server

The transition diagram explores that the technician visits at states S_1 and S_2. Let $N_i(t)$ is the estimated number of visits made by the repairman for repair in (0, t] specified that the system arrives at the regenerative state S_i at t = 0. Then the repetitive interface for $N_i(t)$ is

$$V_0(t) = Q_{09}(t) \otimes [1 + V_9(t)] + Q_{03}(t) \otimes [1 + V_3(t)] + Q_{01}(t) \otimes [1 + V_1(t)]$$

$$V_1(t) = Q_{12}(t) \otimes V_2(t) + Q_{10}(t) \otimes V_0(t) + [Q_{11.4}(t) + Q_{11.(45)^n}(t) + Q_{11.(10,4)}(t) + Q_{11.10(45)^n}(t)] \otimes V_1(t)$$

$$V_2(t) = [Q_{21}(t) + Q_{21.(76)}(t) + Q_{21.8(76)}(t)] \otimes V_1(t)$$

$$V_3(t) = [Q_{31}(t) + Q_{31.(54)^n}(t)] \otimes V_1(t)$$

$$V_9(t) = Q_{91}(t) \otimes V_1(t) \tag{13}$$

Using LST on the above relations (13), then get

$$V_0(\infty) = \lim_{s \to 0} s V_0^{**}(s)$$

$$V_0 = \frac{V_n}{D'} \quad \text{Where} \quad V_n = p_{10} \tag{14}$$

and D' is formerly declared.

V. Particular Cases

Suppose that $f_1(t) = \theta e^{-\theta t}$, $g_1(t) = \phi e^{-\phi t}$, $h_1(t) = \xi e^{-\xi t}$

$$p_{11.4} = \frac{\lambda\phi}{(\phi + \lambda + \lambda_1 + \mu)(\phi + \mu)}, \; p_{11.(45)^n} = \frac{\lambda\mu}{(\phi + \lambda + \lambda_1 + \mu)(\phi + \mu)}$$

$$p_{11.(10,4)} = \frac{\lambda_1\phi}{(\phi + \lambda + \lambda_1 + \mu)(\phi + \mu)}, \; p_{11.10(45)^n} = \frac{\lambda_1\mu}{(\phi + \lambda + \lambda_1 + \mu)(\phi + \mu)}$$

$$p_{21.(76)} = \frac{\lambda}{(\theta + \lambda + \lambda_1)}, \; p_{21.8(76)} = \frac{\lambda_1}{(\theta + \lambda + \lambda_1)}, \; p_{31.(54)^n} = \frac{\mu}{(\phi + \mu)}$$

Also, $M_0 = \frac{1}{(\lambda + \lambda_1 a)} = \mu_0$, $M_1 = \frac{1}{(\phi + \mu + \lambda + \lambda_1)} = \mu_1$, $M_2 = \frac{1}{(\theta + \lambda + \lambda_1)} = \mu_2$

$$W_1(t) = \mu_1', \quad W_2(t) = \mu_2', \quad W_3(t) = \mu_3', \quad W_9(t) = \mu_9$$

$$MTSF = \frac{[(\lambda + \lambda_1 a)\{(\phi + \lambda + \lambda_1 + \mu) + \theta\} - \phi(\lambda a + \theta)]}{[(\theta + \lambda + \lambda_1)\{(\lambda + \lambda_1 a)(\phi + \lambda + \lambda_1 + \mu) - \lambda a \phi\} + \mu\theta(\lambda + \lambda_1 a)]} \tag{15}$$

$$A_0 = \frac{\theta\phi\xi[\phi(\theta + \lambda + \lambda_1) + (\lambda + \lambda_1 a)(\theta + \lambda + \lambda_1 + \mu)]}{A_1 + A_2} \tag{16}$$

where, $A_1 = [\phi(\theta + \lambda + \lambda_1)\{\theta\phi\xi + (\theta + \mu)\lambda b\xi - \lambda_1 a\theta\phi\}]$

$$A_2 = (\lambda + \lambda_1 a)\begin{bmatrix} \theta\phi(\xi + \lambda_1)(\theta + \lambda + \lambda_1 + \mu) \\ + \xi(\lambda + \lambda_1)\{(\theta + \mu)(\theta + \lambda + \lambda_1) + \theta\phi\mu(\phi + \theta)\} \end{bmatrix}$$

$$B_0 = \frac{\phi\begin{bmatrix} \{\lambda b\xi(\theta + \mu) - \lambda_1 a\theta\phi\}(\theta + \lambda + \lambda_1) \\ + \theta\xi(\lambda + \lambda_1 a)(\theta + \lambda + \lambda_1 + \mu) \end{bmatrix}}{A_1 + A_2} \tag{17}$$

where A_1 and A_2 are defined above.

$$V_0 = \frac{\theta\phi^2\xi(\lambda + \lambda_1 a)(\theta + \lambda + \lambda_1)}{A_1 + A_2} \tag{18}$$

VI. Profit Analysis

Using reliability parameters, the profit (P) of the system during the time interval (0,t] is

$$P = T_0 A_0 - T_1 B_0^R - T_2 V_0 \tag{19}$$

Where, $T_0 = 1000$ (Price tag per unit uptime)

$T_1 = 500$ (Cost per unit time for technician Busy)

$T_2 = 100$ (Charge per visit by the technician)

VIII. Discussion

Generally, cold standby redundancy is used to enhance the system performance and sometimes refreshment is offered to the technician to enhance his efficiency.

The system performance is calculated with reliability measures such as MTSF, availability of the system and profit values. Table 2 shows the increasing trend of MTSF with respect to refreshment rate θ, keeping the values of other parameters λ=0.3, λ_1=0.2, μ=0.4, ϕ=0.3, ξ=0.2 are failure rate of the unit in safe mode, unsafe mode, refreshment request rate, repair rate of unit, rebooting delay rate respectively, and these are taken constantly for simplicity. When λ changing from 0.3 to 0.4, λ_1 changing from 0.2 to 0.3, μ varying from 0.4 to 0.5 then MTSF declined.

The table reveals that as the rate of repair ϕ changes from 0.3 to 0.4, and reboot delay rate ξ changes from 0.2 to 0.3 then MTSF enhances.

Table 2: *MTSF vs. Refreshment Rate*

θ ↓	λ=0.3, λ_1=0.2 μ=0.4, ϕ=0.3 ξ=0.2 *a*=0.8 *b*=0.2	λ=0.4	λ_1=0.3	μ=0.5	ϕ=0.4	ξ=0.3
0.1	1.49505	1.40182	1.38431	1.42983	1.45287	1.50588
0.2	1.52616	1.42729	1.40848	1.45878	1.48644	1.53438
0.3	1.55598	1.45179	1.4316	1.48666	1.51862	1.56687
0.4	1.58457	1.47538	1.45372	1.51351	1.54951	1.59368
0.5	1.61202	1.4981	1.47492	1.53948	1.57917	1.62598
0.6	1.63839	1.51475	1.49524	1.56438	1.60768	1.64788
0.7	1.66375	1.54115	1.51475	1.58849	1.6351	1.67466
0.8	1.68814	1..5615	1.53349	1.61178	1.6615	1.69586
0.9	1.71164	1.58118	1.5515	1.63429	1.68694	1.72568
1	1.73427	1.60021	1.56883	1.65605	1.71145	1.74655

The availability of the redundant system is also affected by the refreshment and reboot facilities. Table 3 explores the availability of the system and its value increase corresponding to increments in refreshment rate θ when the system's other parameters λ=0.3, λ_1=0.2, μ=0.4, ϕ=0.3, ξ=0.2 possess constant values. When the failure rate of a unit in safe mode changes (λ=0.3 to 0.4), unsafe mode

changes (λ_1=0.2 to 0.3) then the availability of system declines.

Also, when the technician request rate changes (μ=0.4 to 0.6) then the system's availability declines but when the repair rate of unit changes (ϕ=0.5 to 0.7), reboot rate of unit changes (ξ=0.2 to 0.3) then the availability of the system enhances.

Table 3: *Availability vs. Refreshment Rate*

θ ↓	λ=0.3, λ_1=0.2 μ=0.4, ϕ=0.3 ξ=0.2 a=0.8 b=0.2	λ=0.4	λ_1=0.3	μ=0.5	ϕ=0.4	ξ=0.3
0.1	0.11217	0.10168	0.0973	0.10666	0.1181	0.11682
0.2	0.156	0.14251	0.13534	0.14863	0.16336	0.16537
0.3	0.19261	0.17718	0.16728	0.18385	0.20078	0.20742
0.4	0.22302	0.20641	0.19401	0.21326	0.23162	024353
0.5	0.24825	0.23098	0.21639	0.23776	0.25703	0.2744
0.6	0.26922	0.25162	0.23516	0.25819	0.27804	0.30074
0.7	0.28668	0.26899	0.25094	0.27525	0.29546	0.32323
0.8	0.30128	0.28362	0.26427	0.28955	0.30998	0.34246
0.9	0.31354	0.296	0.27557	0.30157	0.32214	0.35894
1	0.32389	0.3065	0.28525	0.31172	0.33237	0.3731

It is evident from table 4 that the system uses constant parameters such that λ=0.3, λ_1=0.2, μ=0.4, ϕ=0.3, ξ=0.2and the trend of profit values enhanced with respect to increments in refreshment rate θ. When the failure rate of a unit λ in safe mode changes from 0.3 to 0.4 and unsafe mode changes from 0.2 to 0.3 then the profit of the system decreases.

Also, when the technician request rate μ changes from 0.4 to 0.5 then profit values decline but when the repair rate of unit ϕ changes from 0.5 to 0.7 and reboot rate changes from ξ=0.2 to 0.3 then the profit value enhances.

Table 4: *Profit vs. Refreshment Rate*

θ ↓	λ=0.3, λ_1=0.2 μ=0.4, ϕ=0.3 ξ=0.2 a=0.8 b=0.2	λ=0.4	λ_1=0.3	μ=0.5	ϕ=0.4	ξ=0.3
0.1	46.91405	41.65736	39.76882	43.94127	50.09497	50.82808
0.2	66.50537	59.48843	55.46915	62.40709	70.57429	74.43605
0.3	83.49143	75.18573	68.73058	78.49941	88.13762	96.1274
0.4	98.10396	88.87597	79.89434	92.41253	103.1117	115.7624
0.5	110.635	100.7582	89.2898	104.399	115.8591	133.3744
0.6	121.3797	111.0531	97.21094	114.7191	126.7243	149.0903
0.7	130.6097	119.9755	103.9094	123.616	136.0128	163.0801
0.8	138.5628	127.7217	109.5955	131.3055	143.9851	175.5265
0.9	145.4418	134.4642	114.4424	137.9734	150.859	186.6077
1	151.4167	140.3518	118.5933	143.7773	156.8144	196.4892

IX. Conclusion

The refreshment and reboot approach plays a vital role in system configuration and its functioning. These features enhance the capacity of the technician and performance of the system. Tables explore the increasing trends of MTSF, availability and profit values of the system using reboot and refreshment facilities.

References

[1] Agrawal, A., Garg, D., Kumar, A. and Kumar, R. (2021). Performance analysis of the water treatment reverses osmosis plant. Reliability: *Theory & Applications*, 16(3): 16-25.

[2] Dhall, A., Malik, S. C. and Munday, V. J. (2014). A Stochastic System with Possible Maintenance of Standby Unit and Replacement of the Failed Unit Subject to Inspection. *International Journal of Computer Applications*, 97(8).

[3] Goel, M., Kumar, J. and Grewal, A. S. (2017). Cost-Benefit Analysis of a Two-Unit Cold Standby System with Degradation, Priority and General Distribution of all Random Variables. *International Journal of Statistics and Reliability Engineering,* 4(1): 46-55.

[4] Hsu, Y. L., Ke, J. C. and Liu, T. H. (2011). Standby system with general repair, reboot delay, switching failure, and unreliable repair facility—a statistical standpoint. *Mathematics and Computers in Simulation*, 81(11): 2400-2413.

[5] Jain, M., Meena, R. K. and Kumar, P. (2020). Maintainability of redundant machining system with vacation, imperfect recovery and reboot delay. *Arabian Journal for Science and Engineering*, 45(3): 2145-2161.

[6] Jyh-Bin, K., Jyh-Wei, C. and Kuo-Hsiung, W. (2011). Reliability measures of a repairable system with standby switching failures and reboot delay. *Quality Technology & Quantitative Management*, 8(1): 15-26.

[7] Ke, J. C. and Liu, T. H. (2014). A repairable system with imperfect coverage and reboot. *Applied Mathematics and Computation*, 246: 148-158.

[8] Kumar, A., and Sharma, A. (2023). Profit analysis of repairable cold standby system under refreshment. *Reliability: Theory & Applications*, *18*(4 (76)), 604-612.

[9] Kumar, A., Garg, R., and Barak, M. S. (2023). Reliability measures of a cold standby system subject to refreshment. *International Journal of System Assurance Engineering and Management*, *14*(1), 147-155.

[10] Kumar, J. and Goel, M. (2016). Availability and profit analysis of a two-unit cold standby system for general distribution. *Cogent Mathematics*, 3(1): 1262937.

[11] Kumar, P. and Jain, M. (2020). Reliability analysis of a multi-component machining system with service interruption, imperfect coverage, and reboot. *Reliability Engineering & System Safety*, 202: 106991.

[12] Levitin, G., Finkelstein, M., and Xiang, Y. (2020). Optimal preventive replacement for cold standby systems with elements exposed to shocks during operation and task transfers. *IEEE Transactions on Systems, Man and Cybernetics: Systems,* 10: 2168-2216.

[13] Sengar S. and Mangey R. (2022). Reliability and performance analysis of a complex manufacturing system with inspection facility using copula methodology. *Reliability Theory & Applications,* 4(71): 494-508.

[14] Temraz, N. S. Y. (2018). Availability and reliability analysis for dependent system with load sharing and degradation facility. *International Journal of Systems Science and Applied Mathematics*, 3(1): 10-15.

[15] Zhang, Y. L. and Wang, G. J. (2009). A geometric process repair model for a repairable cold standby system with priority in use and repair. *Reliability Engineering and System Safety*, 94(11): 1782-1787.

ENHANCING SECURITY IN IOT DEVICES: A LIGHTWEIGHT HYBRID CRYPTOGRAPHIC SYSTEM (LCS) APPROACH

AMITA SHAH[1], SANJAY SHAH[2], DHAVAL PARIKH[3], NAMIT SHAH[4]

[1]Ph.D Scholar, Computer/IT Engineering, Gujarat Technological University, Gujarat, India amitashah@ldce.ac.in
[2]Professor & Head, Computer Engineering Dept., Government Engineering College, Rajkot, Gujarat, India. sanjay_shah_r@yahoo.com
[3]Professor & Head, Computer Engineering Dept., Government Engineering College, G'nagar, Gujarat, India. daparikh@gecg28.ac.in
[4]Student, Computer Engineering Dept., L D College of Engineering, Ahmedabad, Gujarat, India. namitvshah@gmail.com

Abstract

The escalating connectivity of devices in the Internet of Things (IoT) era necessitates robust security measures while accommodating resource constraints. Lightweight cryptography addresses this need, focusing on algorithm development for devices with limited resources. This research proposes the Lightweight Crypto System (LCS) as a hybrid cryptosystem, integrating the Lightweight Symmetric Algorithm (LSA) and the Lightweight Hash Algorithm (LHA). LSA is a modified AES-128 variant, enhancing data confidentiality, while LHA, derived from SHA-256, verifies data integrity. The study evaluates the proposed LCS on criteria such as execution time, memory usage, avalanche effect, collision resistance, and entropy, emphasizing the optimal balance between performance and security achieved by LSA and LHA. The findings position LCS as a compelling solution for securing IoT devices without compromising on stringent security requirements.

Keywords: Internet of Things (IoT), Lightweight Cryptography, Hybrid Cryptosystems, LSA, LHA, Security.

1. Introduction

The Internet of Things (IoT) integration poses security challenges [1]. With threats like unauthorized access and data breaches, secure communication is vital. Our research introduces Lightweight Crypto System (LCS) as an adaptable solution [2], enhancing IoT security. It safeguards user privacy and ensures data integrity in the interconnected IoT landscape [3, 4].

The Lightweight Crypto System (LCS) is a breakthrough in IoT security, tailored for resource-constrained devices [5, 6, 7]. Comprising three key elements, it forms a comprehensive security framework. The Lightweight Symmetric Algorithm (LSA), a modified AES-128, ensures efficientsymmetric encryption for data confidentiality [8]. Elliptic Curve Cryptography (ECC) facilitates secure key exchange and authentication [9]. Additionally, the Lightweight Hash

Algorithm (LHA) enhances data integrity verification in the LCS system. Together, these components provide robust solutions for confidentiality, authenticity, access control, and data integrity in IoT communication.LCS's effectiveness stems from the collaboration of its three components: LSA ensures data confidentiality with a tailored version of AES-128 for IoT constraints, ECC facilitates secure key exchange, and the proposed LHA enhances data integrity verification. This integration positions LCS as a versatile solution addressing the complex security demands of the evolving IoT landscape.

1.1. Motivation and Contribution

This research is driven by the growing IoT landscape and the security challenges it poses. With the surge in connected devices, addressing security concerns becomes crucial. The motivation stems from the need to enhance IoT security infrastructure due to escalating concerns about data privacy and integrity. This research aims to provide innovative cryptographic solutions tailored for resource-constrained IoT devices, bridging the gap between security and performance. The major contribution is the exploration of a LCS ensuring efficient and secure communication in the interconnected IoT landscape.

- Lightweight Crypto System (LCS): Introduction of a specialized hybrid cryptosystem tailored for IoT environments, providing a comprehensive security solution for resource- constrained devices [11].
- Lightweight Symmetric Algorithm (LSA): Development and implementation of LSA, a modified AES-128 algorithm optimized for IoT devices, ensuring efficient and secure data confidentiality [8].
- Lightweight Hash Algorithm (LHA): Proposal and implementation of LHA, a novel lightweight hash algorithm designed for data integrity verification in resource-constrained IoT devices, striking a balance between security and performance [12].
- Performance and Security Parameter Analysis: Rigorous evaluation of LCS through performance metrics and key security parameters, including execution time, memory usage, avalanche effect, collision resistance, and entropy, offering insights into the system's robust- ness and efficiency.
- Real-world Testbed Implementation: Practical validation of LCS on an IoT testbed using ESP32 hardware, demonstrating its effectiveness in real-world IoT scenarios and providing a foundation for further optimization of cryptographic processes in IoT environments.

2. Related Work

IoT security literature investigates cryptographic methods due to the growing impact of IoT in daily life, especially in critical areas like smart homes and healthcare. The resource constraints of IoT devices necessitate lightweight and energy-efficient security solutions, posing a challenge to widespread adoption. Recent research explores the potential benefits of integrating lightweight blockchain technology, emphasizing the role of hashing in constructing a robust blockchain structure for enhanced IoT security [13]. This study evaluates various hash techniques on a Raspberry Pi device, providing quantitative insights into the performance of hash functions for lightweight blockchain-based IoT applications [13].

The IoT's growth introduces security challenges for embedded devices [14]. A solution proposes using MD5, a hashing algorithm, to enhance IoT security, fortifying embedded devices against internet-based attacks [14]. This approach seamlessly integrates with the IoT framework and is endorsed for securing embedded systems in IoT [14].

The study in [15] tackles the lack of reliable hybrid cryptosystems for securing critical IoT

devices. It explores lightweight encryption algorithms such as TEA, XTEA, XXTEA, RSA, and ECC, aiming for optimal security in constrained environments. The proposed hybrid cryptosystem, incorporating chaotic theory for key generation, outperforms RSA and XXTEA by 40%, showcasing enhanced security and superior performance for safeguarding IoT devices [15].

In [16], the paper underscores the crucial role of cryptography in data security, introducing the Secure Hash Algorithm-3 (SHA-3) for data verification. Implementing a low-power technique, latch-based clock gating, enhances power efficiency in SHA-3 algorithm designs [16].

In [17], the paper recognizes the pivotal role of wireless sensor networks (WSNs) in collecting and transmitting sensitive information. It addresses challenges in WSN infrastructure design due to resource limitations and emphasizes the need for efficient authentication solutions. The proposed 2AMD-160, a Secure Hash Algorithm, outperforms MD5 and SHA1 in both execution time and security in comparative analysis [17].

The IoT's global connectivity introduces security challenges, requiring energy-efficient solutions. Existing methods compromise between security and energy consumption [18] proposes a BLAKE2b-based authentication with modified ECDSA, demonstrating improved signature times on Raspberry Pi-3. The scheme exhibits resistance to attacks, suitable for resource-constrained IoT devices.

In [23], the paper shows how message-level security instead of transport level security is used to provide end-to-end secure communication between IoT devices and Gateway. Various symmetric and asymmetric security algorithms along with different data formats such as XML, JSON and EXI are executed and compared. Used Blowfish encryption algorithm for IoTSyS framework[23].

3. Proposed Methodology

In addressing the escalating security concerns associated with the ever-expanding Internet of Things (IoT) ecosystem, we introduce a comprehensive Lightweight Crypto System (LCS) that harmoniously integrates three key cryptographic components: "MyLightSymAlgo" (LSA), Elliptic Curve Cryptography (ECC), and the novel "MyLightHashAlgo" (LHA). This section gives detail description for each of the components. The architecture for proposed hybrid model (LCS) sender side and receiver side execution approach is presented in Figure 1 and 2, respectively.

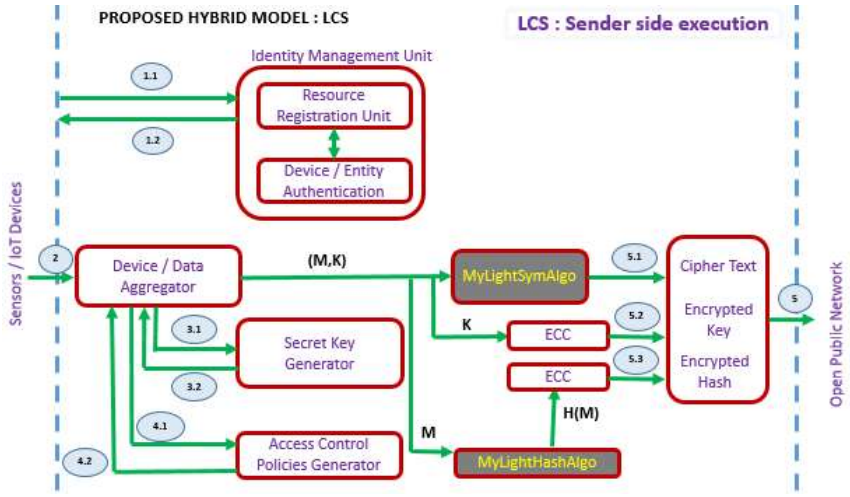


Figure 1: *Proposed Hybrid Model (LCS) (Sender Side Execution)*

The sequential execution flow of the proposed approach is illustrated in Figure 3. The detail explanation for each component for proposed hybrid model (LCS) is presented below.

3.1 MyLightSymAlgo (LSA): Lightweight Symmetric Encryption

LSA, a key element of the proposed LCS, is a tailored variant of AES-128, designed to address security requirements in resource-constrained IoT devices by introducing efficiency-enhancing modifications and replacements [8].

- Algorithmic Modifications: LSA undergoes meticulous changes within AES-128 to create a lightweight encryption process for IoT devices, considering constraints like limited processing power and memory.
- Enhanced Encryption/Decryption: The final LSA version shows notable improvements over AES-128, excelling in both encryption and decryption for faster data transformation, crucial for real-time processing in IoT.

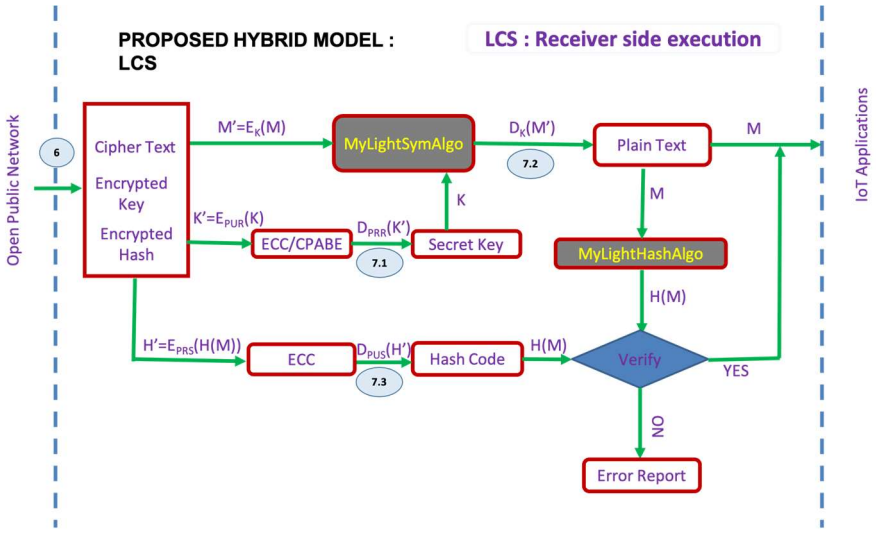


Figure 2: *Proposed Hybrid Model (LCS) (Receiver Side Execution)*

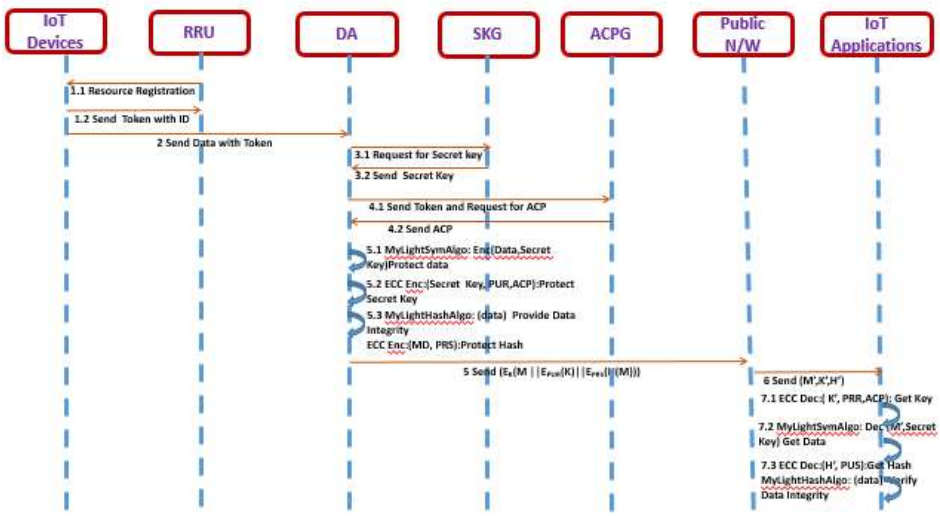


Figure 3: *Sequence of Execution*

- Security and Sensitivity: LSA aims for a highly secure encryption system with sensitivity to small input changes, enhancing overall security by producing significant output variations[8].
- IoT Data Security: LSA is designed to secure data from IoT sensors, suitable for resource-constrained devices where traditional cryptographic methods may be impractical.

As part of LCS, LSA ensures data confidentiality in IoT, protecting sensitive information while addressing IoT device limitations. The LSA diagram is depicted in Figure 4. The step-by-step operations of the LSA employed for secure data encryption is presented in Algorithm 1[8].

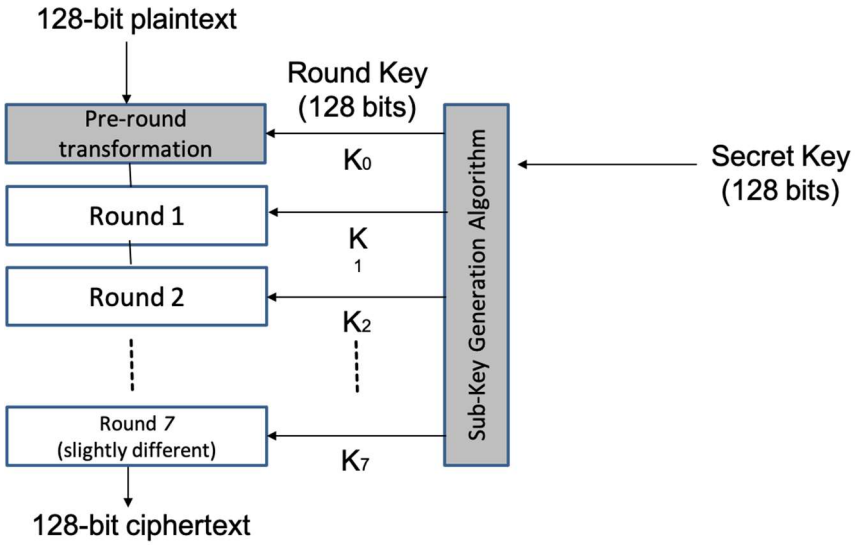


Figure 4: *LSA General Diagram*

Algorithm 1 Lightweight Symmetric Algorithm (LSA)

```
procedure LSA (Data Block, Round Keys)
        XOR Data Block with corresponding Pre Round Key
        Apply Parity Transformation
    for each byte in Data Block do
        Substitute byte using innovative SubBytes table
        Rearrange bytes using Junction Jumping strategy
        XOR with corresponding Round Key byte
    end for
        Substitute byte using innovative SubBytes table
        XOR Data Block with corresponding Round Key byte
    Output: Encrypted Data Block
end procedure
```

3.2 Elliptic Curve Cryptography (ECC): Asymmetric Encryption and Key Generation

Compared with other traditional public key encryption algorithm such as RSA algorithm, ECC algorithm can provide the same security with short key length. ECC operates on elliptic curves over finite fields so provide complexity [19]. The advantages of elliptic curves are: Encryption, De-cryption and Signature Verification speed up, due to shorter key lengths, High safety performance, Small storage space, Fast processing speed, and Low bandwidth requirements [20, 21].

3.3 MyLightHashAlgo (LHA): Lightweight Hash Algorithm for Integrity

MyLightHashAlgo (LHA), a crucial element of Lightweight Crypto System (LCS) for resource-constrained IoT devices, improves data integrity using the optimal SHA256 variant, SHA256SUBPT. Proposed as the Lightweight Hash Algorithm (LHA), SHA256SUBPT achieves a balance between performance and security, with its architecture illustrated in Figure 5.

- LHA Architecture Block Wise: LHA processes data in 512-bit blocks, employing a block-wise approach for efficient handling, contributing to a 256-bit message digest. The block-wise architecture is illustrated in Figure 6.

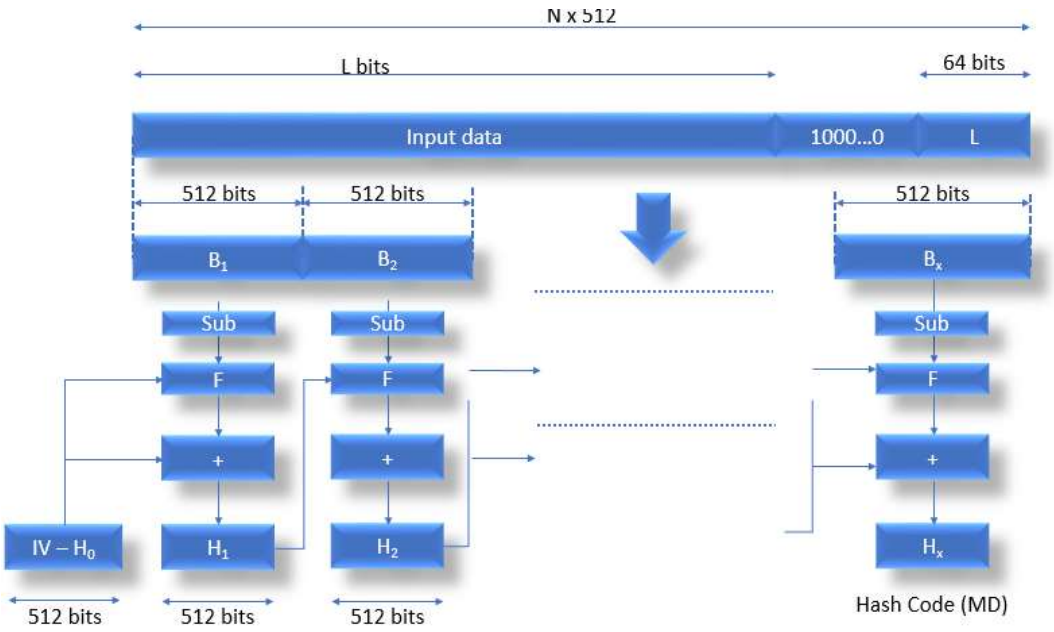


Figure 5: *LHA Architecture*

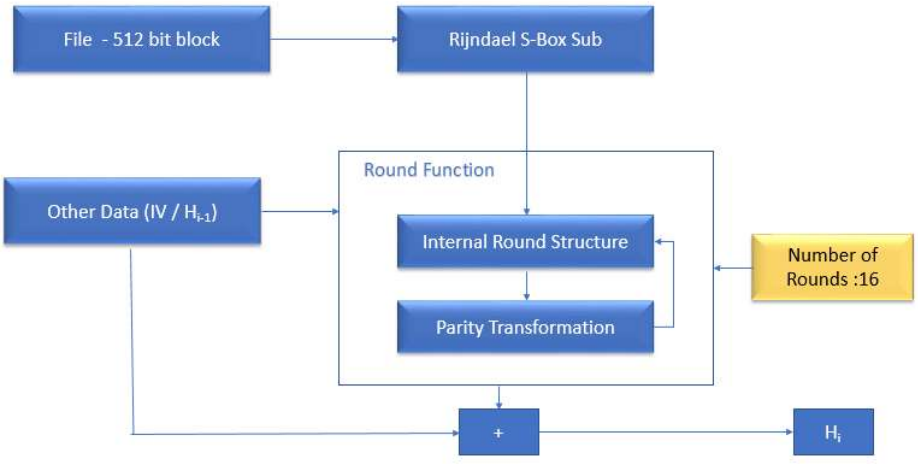


Figure 6: *LHA Architecture Block Wise*

- Hash Function of LHA: LHA's hash function, grounded in Merkle-damgard construction, executes multiple rounds of operations to generate a 256-bit message digest, ensuring data integrity. The graphical representation is illustrated in Figure 7.
- Each Round of Hash Function of LHA: In each round of the LHA hash function, crypto- graphic

operations like bitwise transformations and XOR summations are applied to input data, enhancing algorithm complexity. The iterative process ensures intricate manipulations, leading to the generation of a secure 256-bit hash digest, as depicted in Figure 8. LHA's hash function's round operation integrates key operations like Majority and Choose selections, XOR summations, and optimizations such as round reduction, Rijndael substitution, and Parity transformation which replace all state variables-64 bit, except A and E, with the 1's complements of their present values if their present values happen to be odd to enhance efficiency and security.

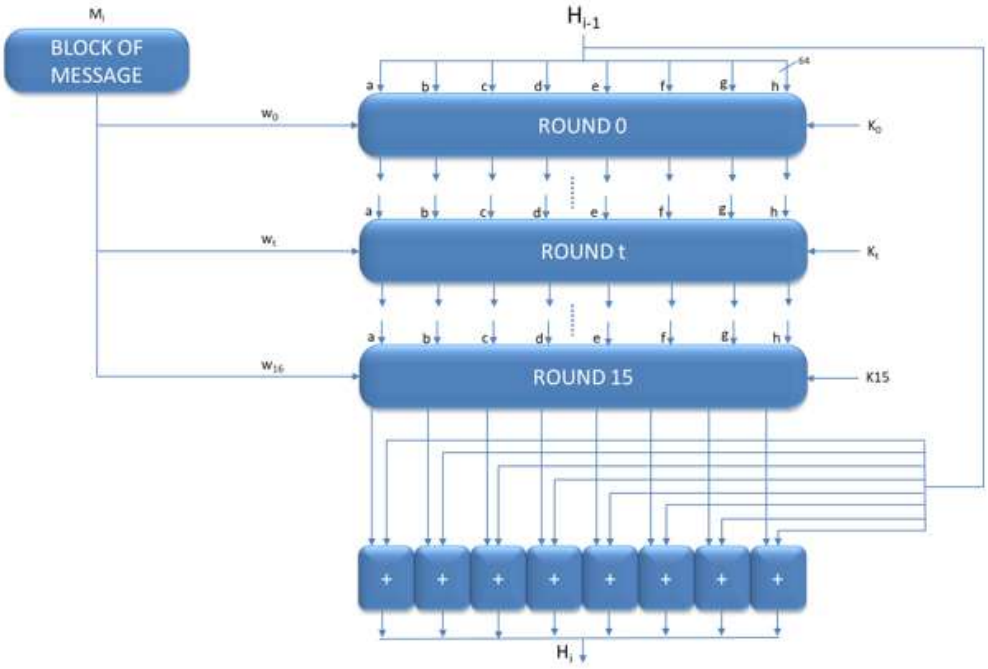


Figure 7: *Hash Function of LHA*

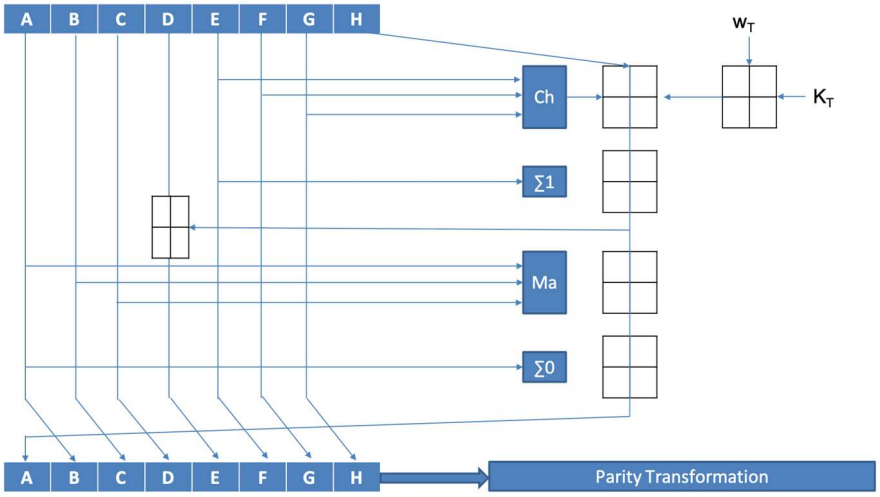


Figure 8: *Each Round of Hash function of LHA*

- **Ma:** Majority selects the next output in a bit-wise manner based on the majority bit forthe three input bytes.
- **Ch:** Choose selects the next output in a bit-wise manner based on the values in the 'x' variable; if $x[i] = 0$, $y[i]$ is selected, else $z[i]$ is selected.
- Σ0: Sigma A: XOR Summation on A.
- Σ1: Sigma E: XOR Summation on E.

- LHA - SECURITY PARAMETERS: To evaluate the security of the LHA algorithm, we consider Collision Resistance and Shannon Entropy as crucial security parameters.

Collision Resistance: This property in cryptographic hash functions ensures that it is challenging to find two inputs (a and b) such that they produce the same output (H(a) =H(b)), where $a \neq b$.

Shannon Entropy: This parameter is employed to gauge Confusion, specifically the obfuscation of the Input-Output Relationship.

4. Experimental Result Analysis

The evaluation and results analysis were conducted to compare the performance of the Lightweight Symmetric Algorithm (LSA) against the Advanced Encryption Standard (AES) in terms of execution time and memory consumption. Security parameters such as avalanche effect, Hamming distance, and entropy were considered.

4.1 LSA vs AES Execution Time Comparison

We conducted a thorough evaluation of the execution time for the LSA compared to the Advanced Encryption Standard (AES). This analysis provides insights into the relative efficiency of LSA and AES in handling cryptographic operations, with implications for real-world applications and system performance. Figure 9 represents the LSA vs AES in terms of execution time.

Figure 9: *LSA vs AES in term of Execution Time*

4.2 LSA vs AES Memory Consumption Comparison

Our study assessed the memory consumption of Lightweight Symmetric Algorithm (LSA) compared to Advanced Encryption Standard (AES), providing insights into their resource requirements for optimizing cryptographic algorithm performance. Figure 10 visually represents the LSA vs AES memory consumption dynamics.

4.3 Security Parameters Evaluation

The algorithm's security assessment centered on the avalanche effect, hamming distance, and entropy, evaluating diffusion, confusion, and bit changes. Parameters were derived from 10,000 randomly generated blocks to conduct a thorough security analysis.

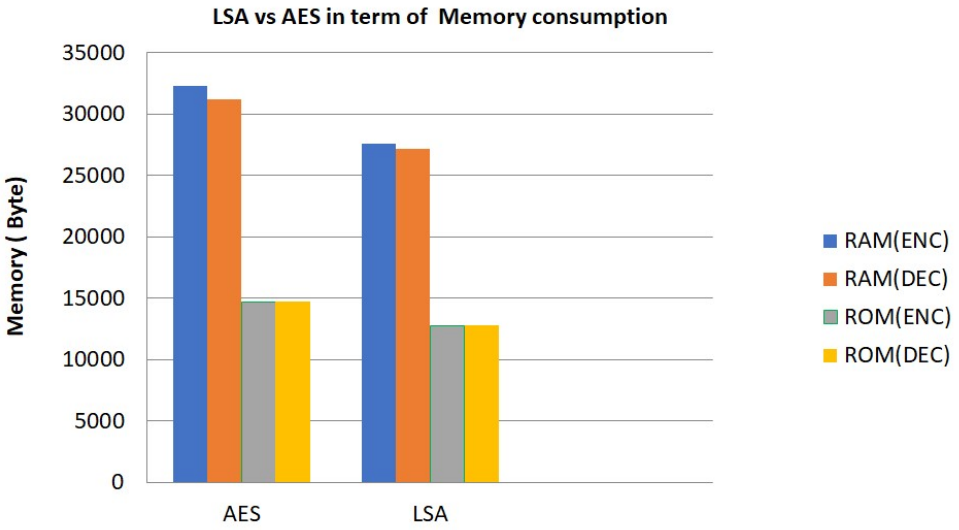


Figure 10: *LSA vs AES in term of Memory Consumption*

4.3.1 LSA Avalanche Criterion Evaluation

Our assessment of the Lightweight Symmetric Algorithm (LSA) scrutinizes its performance using the Avalanche Criterion (AC), considering first, second, and third-order criteria along with the average. This comprehensive analysis sheds light on LSA's efficacy in diffusing changes in input data, reflecting in output data, essential for robust cryptographic applications.

Table 1: *Performance of LSA in term of Avalanche Criterion*

	1st Order AC	2nd Order AC	3rd Order AC	Average AC
AES	49.00%	49.00%	49.00%	49.00%
LSA	49.00%	42.00%	49.00%	46.66%

Table 1 displays LSA's performance in Avalanche Criterion (AC) at various orders and the average AC, contrasting with AES's constant 49.00%. LSA shows variations, notably a 2nd Order AC drop to 42.00%, resulting in an average AC of 46.66%, indicating different levels of change propagation compared to AES. Figure 11 visually compares LSA and AES in terms of Avalanche Effect, offering insights into their respective responses to modifications in input data and highlighting their unique diffusion characteristics.

4.3.2 LSA Entropy Criteria

Table 2 reveals the performance of LSA and AES based on Entropy criteria, indicating a 50.00%

Hamming Distance for both and very close Shannon's Entropy values. LSA slightly edges higher at 3.612 compared to AES's 3.611, showcasing comparable security aspects related to confusion and the number of flipped bits after conversion.

Table 2: *Performance of LSA in term of Avalanche Criterion*

	Hamming Distance	Shannon's Entropy (Log2)
AES	50.00%	3.611
LSA	50.00%	3.612

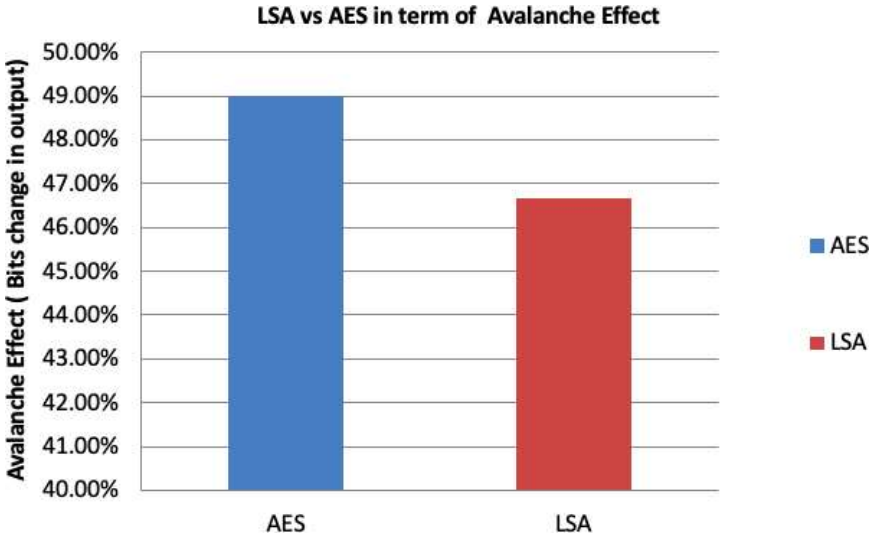


Figure 11: *LSA vs AES in Term of Avalanche Effect*

Figure 12 illustrates the evaluation of LSA versus AES based on Hamming Distance, showing an equivalent and favorable 50.00% for both algorithms. Hamming Distance, measuring bit differences between original and encrypted data, underscores their similar and consistent performance in this security parameter.

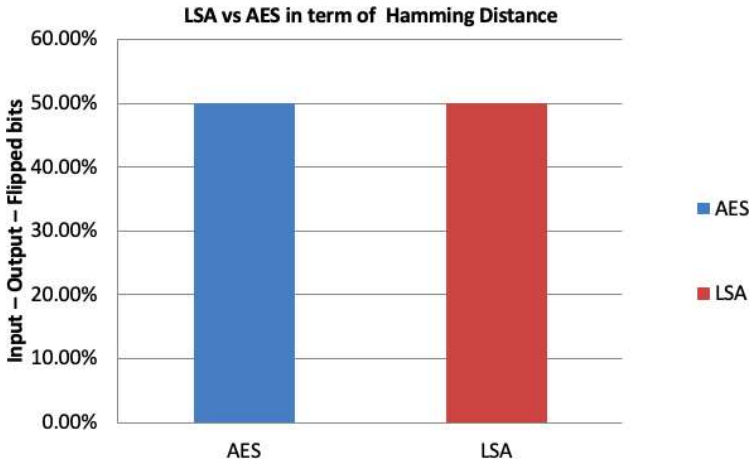


Figure 12: *LSA vs AES in Term of Hamming Distance*

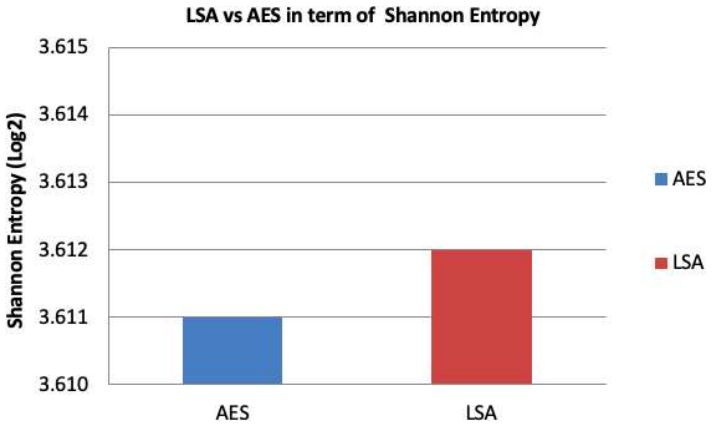


Figure 13: *LSA vs AES in term of Shannon Entropy*

In Figure 13, LSA and AES exhibit identical Shannon Entropy values of 3.612, emphasizing their equivalent performance in providing a high degree of confusion and obfuscation in the input-output relationship. These similar Shannon Entropy values highlight a strong level of security for both algorithms in this parameter.

4.1. Performance Analysis for SHA256 and LHA

4.4.1 Memory Consumption of SHA256 and LHA

Figure 14 presents the results of experiments evaluating the memory consumption of SHA256 and LHA, offering a comparison of their respective requirements under specified conditions. The analysis provides insights into the efficiency and resource utilization of SHA256 and LHA, highlighting their impacts on memory consumption in cryptographic operations.

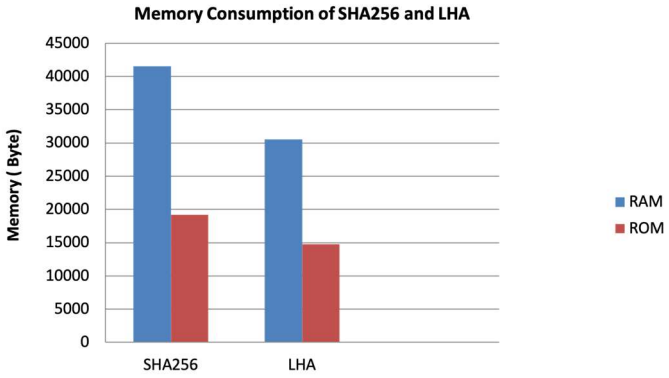


Figure 14: *Performance of SHA256 and LHA in Term of Memory Consumption*

4.4.2 Performance of SHA256 and LHA in Term of Execution time

Figure 15 illustrates the experimentally evaluated execution time performance of SHA256 and LHA, offering insights into the time efficiency of both algorithms, including processing time and overall execution speed.

4.4.3 Collision Count Analysis

The collision count analysis for SHA256 and LHA, considering different sizes of strings per hash, is presented in the Table 3. The collision count is an important metric to assess the robustness and security of hash functions.

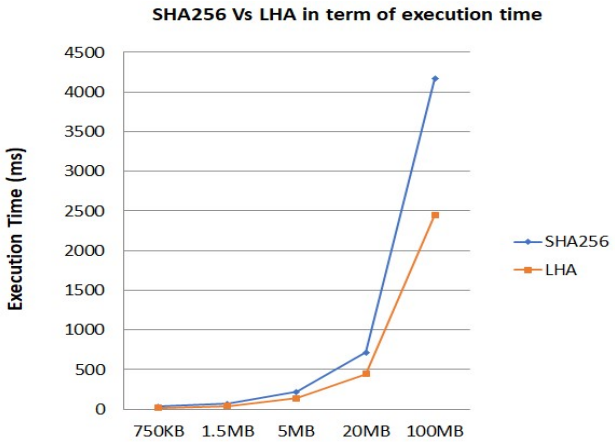


Figure 15: *Performance of SHA256 and LHA in Term of Execution time*

Table 3: *Collision Count Analysis for SHA256 and LHA*

Size of String per Hash	SHA256 Collision Count	LHA Collision Count
04 Hex	6.10621	6.10645
08 Hex	5.2e-05	3.25e-05
16 Hex	0	0
32 Hex	0	0
64 Hex (Full Hash code)	0	0

4.4.4 Shannon Entropy Analysis

The Shannon Entropy analysis for SHA256 and LHA, considering different sizes of strings per hash, is presented in the Table 4. Shannon Entropy is a measure of uncertainty or information content in a system.

Table 4: *Shannon Entropy Analysis for SHA256 and LHA*

Size of String per Hash	SHA256 Shannon Entropy	LHA Shannon Entropy
04 Hex	1.81759	1.81727
08 Hex	2.59429	2.59478
16 Hex	3.2089	3.20719
32 Hex	3.61134	3.61357
64 Hex (Full Hash code)	3.8196	3.82103

4.4.5 Performance Analysis of AES vs LSA and SHA vs LHA on ESP32 Platform

In this configuration, the ESP32-CAM acts as an access point with a WebSocket server and DHT11 sensor, transmitting encrypted data. The M5Stack ESP32 serves as a client, connecting to ESP32-CAM, triggering operations with three buttons, ensuring secure communication and decoding of encrypted data for display.

Table 5: *Comparison of LSA and AES Encryption/Decryption Performance for Text-Based Data*

Parameter	**LSA** (microseconds)	**AES** (microseconds)
Encryption Time	268	1680
Encryption Time per Block	48	384
Decryption Time	267	2060
Decryption Time per Block	48	464
Data Length	69 bytes	69 bytes

Table 6: *Comparison of LSA and AES Encryption/Decryption Performance for Image-Based Data*

Parameter	**LSA** (microseconds)	**AES** (microseconds)
Encryption Time	7205	61438
Encryption Time per Block	32	320
Decryption Time	3090	73251
Decryption Time per Block	32	368
Data Length	3056 bytes	3056 bytes

Table 7: *Comparison between the Performance of the LSA and AES For Image-based Data*

Algorithm	Operation	Time (μs)	Time per Block (μs)	Data Length (bytes)	Hash Function	Hash Time (μs)	Hash Time per Block (μs)
AES with SHA	Encrypt	69226	320	3440	SHA256	2037	9
AES with SHA	Decrypt	82459	368	3440	SHA256	2037	9
LSA with LHA	Encrypt	8108	32	3440	LHA256	1259	5
LSA with LHA	Decrypt	8059	32	3440	LHA256	1259	5

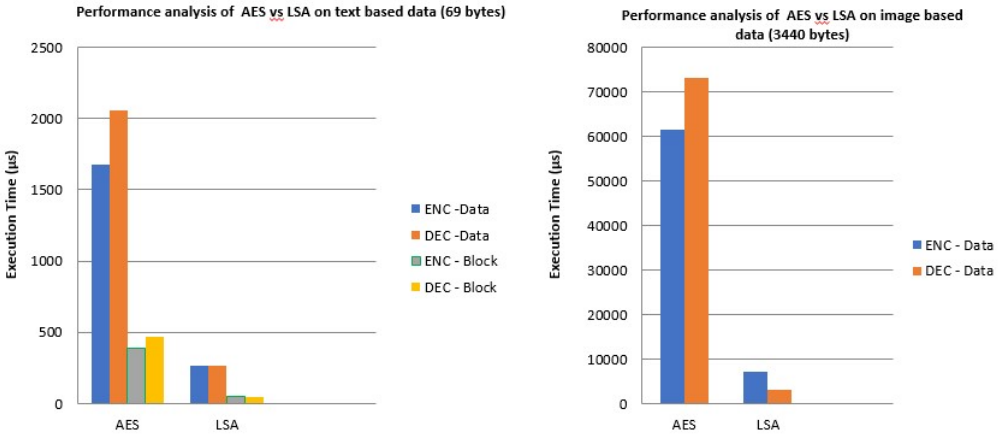


Figure 16: *Performance analysis of AES vs LSA on text/image based data on IoT Set up*

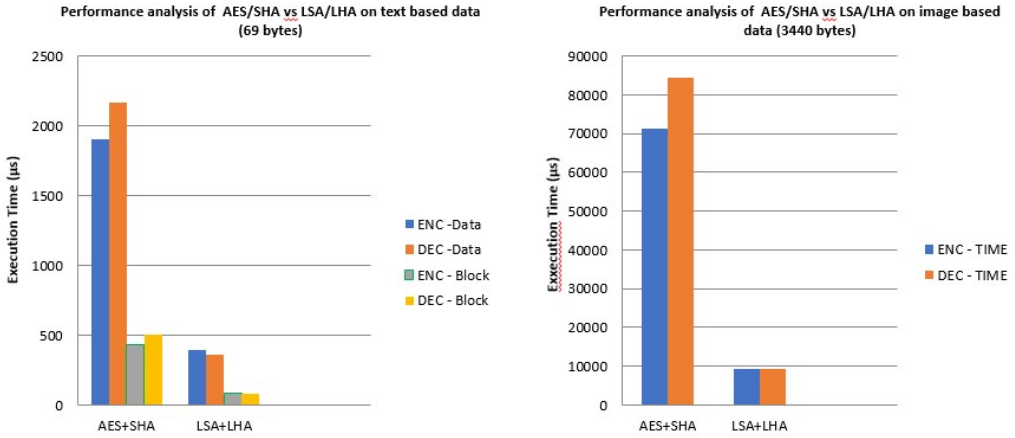


Figure 17: *Performance analysis of AES/SHA vs LSA/LHA on text/image based data on IoT Set up*

Tables 5 and 6 compare LSA and AES encryption/decryption performance for text-based and image-based data. Figure 16 illustrates the experimentally evaluated execution time performance of AES and LSA on IoT set up. Tables 7 and 8 compare performance of combination of LSA-LHA and AES-SHA encryption/decryption for text-based and image-based data. Figure 17 illustrates the experimentally evaluated execution time performance of AES-SHA and LSA-LHA on IoT setup.

Table 8: *Comparison between the Performance of the LSA and AES for Text based data*

Algorithm	Operation	Time (µs)	Time per Block (µs)	Data Length (bytes)	Hash Function	Hash Time (µs)	Hash Time per Block (µs)
AES with SHA	Encrypt	1755	400	69	SHA256	146	36
AES with SHA	Decrypt	2016	464	69	SHA256	146	36
LSA with LHA	Encrypt	311	64	69	LHA256	81	20
LSA with LHA	Decrypt	281	64	69	LHA256	81	20

Table 9 and 10 detail the throughput of AES and LSA for text-based and image-based data, respectively. Furthermore, Table 11 illustrates the throughput of AES and SHA for image data, while Table 12 showcases the throughput of LSA and LHA for image data.

Table 9: *Text-Based Data Throughput of LSA and AES*

Parameter	LSA Throughput (Mb/s)	AES Throughput (Mb/s)
Encryption Throughput	2.06	0.33
Decryption Throughput	2.07	0.27
Data Length	69 bytes	69 bytes

Table 10: *Image-Based Data Throughput LSA and AES*

Parameter	LSA Throughput (Mb/s)	AES Throughput (Mb/s)
Encryption Throughput	3.39	0.40
Decryption Throughput	3.41	0.33
Data Length	3056 bytes	3056 bytes

Table 11: *AES with SHA Throughput with Image Data*

Operation	Throughput (Mb/s)
AES Encryption	0.40
AES Decryption	0.33
SHA Hashing	13.51

Table 12: *LSA with LHA Throughput with Image Data*

Operation	Throughput (Mb/s)
LSA Encryption	3.39
LSA Decryption	3.415
LHA Hashing	21.86

The comparison tables show that LSA consistently outperforms AES in encryption and decryption performance for both text-based and image-based data, with lower times per block and higher throughput. When combined with LHA, LSA demonstrates remarkable efficiency, especially for image data, surpassing AES with SHA. These findings highlight LSA's potential for secure and efficient data transmission, particularly in applications like IoT and image processing.

5. Conclusion

Our Hybrid Lightweight Cryptographic System (LCS) is a significant advancement in IoT data security, employing LSA for encryption, ECC for key management, and LHA for integrity verification. Our LHA algorithm outperforms SHA-256, showcasing similar collision resistance in practical scenarios, ensuring enhanced performance without compromising security.Additionally, LSA with LHA exhibits notable performance improvements in comparison with AES128 with SHA256, particularly in real-time image-based data transmission on ESP32. When working with text data, there's an observed approximately 79.5% improvement in performance, and when

working with image data, the improvement is approximately 86.8%. Such improvements have potential implications for real-time applications, particularly where latency is a concern. This study paves the way for future cryptographic optimizations in IoT. Identifying and addressing potential security vulnerabilities in these areas will be a key focus of our future research.

References

[1] Hernndez-Ramos, Jos© L., et al. "Protecting personal data in IoT platform scenarios through encryption-based selective disclosure." Computer Communications 130 (2018): 20-37.
[2] Yao, Xuanxia, Zhi Chen, and Ye Tian. "A lightweight attribute-based encryption scheme for the Internet of Things." Future Generation Computer Systems 49 (2015): 104-112.
[3] Goyal, Tarun Kumar, and Vineet Sahula. "Lightweight security algorithm for low power IoT devices." 2016 international conference on advances in computing, communications and informatics (ICACCI). IEEE, 2016.
[4] Singh, Saurabh, et al. "Advanced lightweight encryption algorithms for IoT devices: survey, challenges and solutions." Journal of Ambient Intelligence and Humanized Computing (2017): 1-18.
[5] Naru, Effy Raja, Hemraj Saini, and Mukesh Sharma. "A recent review on lightweight cryptography in IoT." 2017 international conference on I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC). IEEE, 2017
[6] Dhanda, Sumit Singh, Brahmjit Singh, and Poonam Jindal. "Lightweight cryptography: a solution to secure IoT." Wireless Personal Communications 112 (2020): 1947-1980.
[7] Rana, Muhammad, Quazi Mamun, and Rafiqul Islam. "Lightweight cryptography in IoT networks: A survey." Future Generation Computer Systems 129 (2022): 77-89.
[8] Shah, Amita, et al. "LSA: A LIGHTWEIGHT SYMMETRIC ENCRYPTION ALGORITHM FOR RESOURCE-CONSTRAINED IOT SYSTEMS." Reliability: Theory & Applications 18.3 (74) (2023): 44-58.
[9] Bos, Joppe W., et al. "Elliptic curve cryptography in practice." Financial Cryptography andData Security: 18th International Conference, FC 2014, Christ Church, Barbados, March 3-7, 2014, Revised Selected Papers 18. Springer Berlin Heidelberg, 2014.
[10] Lara-Nino, Carlos Andres, Arturo Diaz-Perez, and Miguel Morales-Sandoval. "Lightweight elliptic curve cryptography accelerator for internet of things applications." Ad Hoc Networks 103 (2020): 102159.
[11] Yang, Xu, et al. "Blockchain-based secure and lightweight authentication for Internet of Things." IEEE Internet of Things Journal 9.5 (2021): 3321-3332.
[12] Mahlake, Ntebatseng, et al. "A Lightweight Encryption Algorithm to Enhance Wireless Sensor Network Security on the Internet of Things." J. Commun 18 (2023): 47-57.
[13] Alfrhan, Aishah, Tarek Moulahi, and Abdulatif Alabdulatif. "Comparative study on hash functions for lightweight blockchain in Internet of Things (IoT)." Blockchain: Research and Applications 2.4 (2021): 100036.
[14] Landge, Irfan A., and Hannan Satopay. "Secured IoT through hashing using MD5." 2018 fourth international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB). IEEE, 2018.
[15] Ragab, Ahmed, et al. "Robust hybrid lightweight cryptosystem for protecting IoT smart devices." Security, Privacy, and Anonymity in Computation, Communication, and Storage: SpaCCS 2019 International Workshops, Atlanta, GA, USA, July 14“17, 2019, Proceedings 12.Springer International Publishing, 2019.
[16] Sharma, Jayanti, and Deepali Koppad. "Low power and pipelined secure hashing algorithm-3 (SHA-3)." 2016 IEEE Annual India Conference (INDICON). IEEE, 2016.
[17] Al-Mashhadi, Haider M., Hala B. Abdul-Wahab, and Rehab F. Hassan. "Secure and time

efficient hash-based message authentication algorithm for wireless sensor networks." 2014 Global Summit on Computer & Information Technology (GSCIT). IEEE, 2014.

[18] Rao, Vidya, and K. V. Prema. "Light-weight hashing method for user authentication in Internet-of-Things." Ad Hoc Networks 89 (2019): 97-106.

[19] Shah, Amita, et al. "LSA: A LIGHTWEIGHT SYMMETRIC ENCRYPTION ALGORITHM FOR RESOURCE-CONSTRAINED IOT SYSTEMS." Reliability: Theory & Applications 18.3 (74) (2023): 44-58.

[20] Zargar, Ansah Jeelani, Mehreen Manzoor, and Taha Mukhtar. "ENCRYPTION/DECRYPTION USING ELLIPTICAL CURVE CRYPTOGRAPHY." International journal of Advanced Re- search in computer science 8.7 (2017).

[21] Keerthi, K., and B. Surendiran. "Elliptic curve cryptography for secured text encryption." 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT). IEEE, 2017.

[22] Seok, Byoungjin, Jinseong Park, and Jong Hyuk Park. "A lightweight hash-based blockchain architecture for industrial IoT." Applied Sciences 9.18 (2019): 3740.

[23] Pandya, Hetal B., and Tushar A. Champaneria. "Enhancement of security in IoTSyS framework." Proceedings of International Conference on Communication and Networks: ComNet 2016. Springer Singapore, 2017.

RT&A, No 2 (78)
Volume 19, June, 2024

THE EFFICIENT CLASSES OF ESTIMATORS FOR THE PRODUCT OF TWO POPULATION MEANS IN THE EXISTENCE OF NON-RESPONSE UNDER THE STRATIFIED POPULATION-A SIMULATION STUDY

Manish Mishra, B. B. Khare & Sachin Singh

•

School of Liberal Studies, UPES, Dehradun-248001, India.
Department of Statistics, Banaras Hindu University, Varanasi-221005, India.
Department of Mathematics, Sharda University, Greater Noida-201306, India.
manish.mishra@ddn.upes.ac.in, bbkhare56@yahoo.com, singhat619@gmail.com

Abstract

This paper focuses on estimating the product of two population means. Within this paper, we have introduced three distinct classes of estimators for product of two population means. These estimators take into account the known population mean of an auxiliary variable under the framework of stratified random sampling and the presence of non-response in the study variable. Basically, for case (I) we assume the non-response on the study variable and utilize the auxiliary information corresponding to the responding units of the study variable and in case (II), we utilize the complete dataset from the auxiliary variable while also accounting for non-response in the study variable. In case (III) we combined both the information of the auxiliary variable and assumed the non-response on the study variable. Expressions for bias and mean square error have been derived, extending up to the first-order derivative. We have also pinpointed some specific members of the proposed estimator. We have conducted a simulation study to evaluate the valuable insights into the performance of the suggested classes of estimators with the conventional estimator.

Keywords: Product of two population means, Auxiliary variable, Stratified Sampling, Non-response, Mean square error

1. Introduction

The product of two population means is a common parameter used in various areas, including agriculture, economics, social sciences, public health and other scientific investigations. The estimation problem of the product of two population means is very useful in practice. For instance, suppose we want to estimate the total production in a farm having N plants then we select some plants and observe the number of pods and the seeds in pods in a plant then from selected units we will calculate the average number of pods/plant and average number of seeds/plant. Multiplying the average number of seeds/plant and the number of pods/plant with the total number of plants we get the total number of seeds.

In the case of finite population utilizing the auxiliary information(s), the estimation problems related to the product of population means have been examined by several researchers such as Singh[14], Singh[15], Singh[16], Singh[17], Ray and Singh[11], Srivastava et al.[20] and Khare[3] Kumar and Srivastava [8].

Any researcher or statistician may typically encounter the phenomena of the non-response in a scientific investigation or in a sample survey. The reasons for the occurrence of the non-response

are the respondent's unwillingness to respond to some questions that are sentimental, the person not at home, Lack of enthusiasm, etc.

To cope with the complications of non-response, Hansen and Hurwitz [2] have put forward a method of sub-sampling from non-responding units. Using information from responding and sub-sampling units picked from non-responding units in the sample, he suggested a population mean estimator.

In the existence of non-response, the estimation of the product of population means utilizing auxiliary information has been considered by Khare and Sinha[4, 5] and Khare et al.[6].

For stratified populations, Khare and Jha[7], Singh et al.[18] and Mishra et al.[9] have suggested precise estimators of population parameters in the situation of non-response on study character using known and unknown population mean of auxiliary variable.

Motivated by Khare and Jha[7], Three classes of estimators are proposed to estimate the product of two population means under the stratified random sampling with the known population mean of auxiliary character having the non-response on study character. The expressions for bias and mean square error (MSE) of the suggested estimators have been obtained. For the numerical study, the data generated by simulation using R-software has also been given to validate the supremacy of the suggested estimators.

2. Notations and Sampling procedure:

In the study of the population $\eta : (\eta_1, \eta_2, \eta_3 ... \eta_N)$ of the size N. The population is divided into L strata. We denote y_h $(h = 1, 2)$ as study characters of the population having population means $\overline{Y}_{hi}$ $(h = 1, 2)$ in the i^{th} stratum, $(i = 1, 2, 3..L)$ and for the auxiliary character (x), the population mean is denoted by $\overline{X}_i$ which is known for each stratum. For each stratum, the stratum proportion and size is also known. Let N_{i1} and $N_{i(2)}$ are the number of units belonging to the responding and non-responding part of the i^{th} stratum such that $N_{i1} + N_{i(2)} = N_i$. For the i^{th} stratum, $W_{i1} = \frac{N_{i1}}{N}$ and $W_{i2} = \frac{N_{i(2)}}{N}$ represent the response and non-response rates.

Here, we are considering the problem in stratified population using non-response in each stratum, the estimation of $P\ (= \overline{Y}_1 \overline{Y}_2)$ for known $\overline{X}$ has been considered. The sampling procedure we use for the study is as:
we select n_i units from i^{th} stratum, n_{i1} units are selected from N_{i1} units using simple random sampling without replacement (SRSWOR) method of sampling, Here, we obtain n_{i1} units as respondent and $n_{i(2)}$ units as non-responding units out of n_i selected units. We select r_i units, $\left(r_i = \frac{n_{i(2)}}{k_i}, k_i > 1\right)$ from $n_{i(2)}$ units of the i^{th} stratum.
Using the information available on $(n_{i1} + r_i)$ units, the estimator $\overline{y}^*_{hi}$, $(h = 1, 2)$ is given by Hansen and Hurwitz[2] method as follows:

$$\overline{y}^*_{hi} = \frac{n_{i1}}{n_i}\overline{y}_{hi(1)} + \frac{n_{i(2)}}{n_i}\overline{y}_{hi(2)}, \tag{1}$$

where, $\overline{y}_{hi(1)}$ denotes the sample means of $\overline{y}_h$ for n_{i1} responding units in the i^{th} stratum and $\overline{y}_{hi(2)}$ is the sample mean of r_i units drawn from non-responding units $(n_{i(2)})$ in the i^{th} stratum. The estimator $\overline{y}^*_{hi}$ $(h = 1, 2)$ is unbiased and has variance given by:

$$V(\overline{y}^*_{hi}) = \frac{(1 - f_i)}{n_i}S^{*2}_{yhi} + \frac{W_i(k_i - 1)}{n_i}S^{*2}_{yhi(2)} \tag{2}$$

where, S^{*2}_{yhi} and $S^{*2}_{yhi(2)}$ are the population mean square for n_{i1} responding units and non-responding units of i^{th} stratum of the population.

The Hansen and Hurwitz [2] estimator for the auxiliary variable x, we have

$$\overline{x}^*_i = \frac{n_{i1}}{n_i}\overline{x}_{i(1)} + \frac{n_{i(2)}}{n_i}\overline{x}_{i(2)} \tag{3}$$

where, $\overline{x}_{i(1)}$ and $\overline{x}_{i(2)}$ are the sample means of x for n_{i1} responding units and based on r_i units sub-sampled from $n_{i(2)}$ non-responding units of the i^{th} stratum.

In the presence of non-response, the stratified sample means for $\overline{Y}_h$ $(h = 1, 2)$ and $\overline{X}$ are given as follows:

$$\overline{y}^*_{hst} = \sum_{i=1}^{L} W_i \overline{y}^*_{hi} \text{ and } \overline{x}^*_{st} = \sum_{i=1}^{L} W_i \overline{x}^*_i \tag{4}$$

In each stratum utilizing the information on n_i units, the stratified sample mean to estimate $\overline{X}$ is given by:

$$\overline{x}_{st} = \sum_{i=1}^{L} W_i \overline{x}_i \tag{5}$$

where, $W_i = \frac{N_i}{N}$.

3. Proposed Classes of Estimators:

Let $\hat{P}_{st} = \overline{y}^*_{1st}\overline{y}^*_{2st}$ denotes an estimator for P using stratified random sampling in the existence of non-response. We propose three different estimators for P under different situations of non-response which are given as follows:

Case (A): **Utilizing Incomplete information on y_h, $(h = 1, 2)$ and corresponding values of x when $\overline{X}$ is known**.

$$l_1 = d_{(1)}(\hat{P}_{st}, m_1) \tag{6}$$

Case (B): **Incomplete information on y_h, $(h = 1, 2)$ and complete information on x for the known $\overline{X}$.**

$$l_2 = d_{(2)}(\hat{P}_{st}, m_2) \tag{7}$$

Case (C): **Utilizing Incomplete information on y_h, $(h = 1, 2)$ and corresponding values of x and complete information on auxiliary variable.**

$$l_3 = h(\hat{P}_{st}, m_1, m_2) \tag{8}$$

such that,

$$\begin{aligned} &d_{(j)}(P,1) = P,\ d_{1(j)}(P,1) = 1,\ h(P,1,1) = P \text{ and } h_1(P,1,1) = 1 \text{ where,} \\ &d_{1(j)}(P,1) = \Big[\frac{\delta}{\delta \hat{P}_{st}} d_{(j)}(\hat{P}_{st}, m_j)\Big]_{(P,1)}, d_{2(j)}(P,1) = \Big[\frac{\delta}{\delta m_j} d_{(j)}(\hat{P}_{st}, m_j)\Big]_{(P,1)}, \\ &h_1(P,1,1) = \Big[\frac{\delta}{\delta \hat{P}_{st}} h(\hat{P}_{st}, m_1, m_2)\Big]_{(P,1,1)}, h_2(P,1,1) = \Big[\frac{\delta}{\delta m_1} h(\hat{P}_{st}, m_1, m_2)\Big]_{(P,1,1)} \\ &h_3(P,1,1) = \Big[\frac{\delta}{\delta m_2} h(\hat{P}_{st}, m_1, m_2)\Big]_{(P,1,1)} \text{ for j=1,2.} \end{aligned} \tag{9}$$

We denote, $\hat{P}_{st} = \overline{y}^*_{1st}\overline{y}^*_{2st}$, $m_1 = \frac{\overline{x}^*_{st}}{\overline{X}}$ and $m_2 = \frac{\overline{x}_{st}}{\overline{X}}$.

We assume that the function $d_{(1)}(\hat{P}_{st}, m_1)$, $d_{(2)}(\hat{P}_{st}, m_2)$ and $h(\hat{P}_{st}, m_1, m_2)$ are fulfilling the regularity conditions which are given as follows:

- For any sample chosen from any design, the selected functions $d_{(j)}(\hat{P}_{st}, m_j)$, and $h(\hat{P}_{st}, m_1, m_2)$ take the values in two-dimensional and three-dimensional real space G_1 and G_2 having the points $(P, 1)$ and $(P, 1, 1)$.
- The partial derivatives of the first-order and second-order with respect to $\hat{P}$, m_1 and m_2 for the functions $d_{(j)}(\hat{P}_{st}, m_j)$, and $h(\hat{P}_{st}, m_1, m_2)$ are continuous and bounded in G_1 and G_2.

(10)

Using the regularity conditions (10) and expanding the function $d_{(j)}(\hat{P}_{st}, m_j)$, for $j = 1, 2$ and $h(\hat{P}_{st}, m_1, m_2)$ about the point $(P, 1)$ and $(P, 1, 1)$ respectively by utilizing Taylor's series up to the second-order partial derivatives, we get

$$l_j = d_{(j)}(\theta_1) + (\hat{P}_{st} - P)d_{1(j)}(\theta_1) + (m_j - 1)d_{2(j)}(\theta_1) + \frac{1}{2}\Big[(\hat{P}_{st} - P)^2 d_{11(j)} (\theta_1^*) + (m_j - 1)^2 d_{22(j)}(\theta_1^*) + 2(\hat{P}_{st} - P)(m_j - 1)d_{12(j)}(\theta_1^*)\Big] \tag{11}$$

$$\begin{aligned} l_3 = {} & h(\theta_2) + (\hat{P}_{st} - P)h_1(\theta_2) + (m_1 - 1)h_2(\theta_2) + (m_2 - 1)h_3(\theta_2) + \\ & \frac{1}{2}\Big[(\hat{P}_{st} - P)^2 h_{11}(\theta_2^*) + (m_1 - 1)^2 h_{22}(\theta_2^*) + (m_2 - 1)^2 h_{33}(\theta_2^*) \\ & + 2(\hat{P}_{st} - P)(m_1 - 1)h_{12}(\theta_2^*) + 2(\hat{P}_{st} - P)(m_2 - 1)h_{13}(\theta_2^*) \\ & + 2(m_1 - 1)(m_2 - 1)h_{23}(\theta_2^*)\Big] \end{aligned} \tag{12}$$

Where, $\theta_1 = (P, 1)$, $\theta_1^* = (\hat{P}_{st}^*, m_j^*)$, $\theta_2 = (P, 1, 1)$, $\theta_2^* = (\hat{P}_{st}^*, m_1^*, m_1^*)$, $\hat{P}^* = P + \alpha_1(\hat{P}_{st} - P)$ and $m_j^* = 1 + \alpha_2(m_j - 1)$, $0 < \alpha_j > 1$ for $j = 1, 2$.

Here, $d_{1(j)}(\theta_1)$ and $d_{11(j)}(\theta_1)$ are the first and second-order partial derivatives with respect to $\hat{P}_{st}$ and $d_{2(j)}(\theta_1)$ and $d_{22(j)}(\theta_1)$ are the first and second-order partial derivatives with respect to m_j for the function $d_{(j)}(\theta_1)$. Similarly $h_1(\theta_2)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ are the first-order partial derivatives with respect to $\hat{P}_{st}$, m_1 and m_2 respectively and using the condition given in equation (9) and presume that the second-order derivatives are very small in equations (11) and (12), the expression for l_j and l_3 are given as follows:

$$l_j = \hat{P}_{st} + (m_j - 1)d_{2(j)}(\theta_1) + P^{-1}(\hat{P}_{st} - P)(m_j - 1)d_{2(j)}(\theta_1^*) \tag{13}$$

$$l_3 = \hat{P}_{st} + (m_1 - 1)h_2(\theta_2) + (m_2 - 1)h_3(\theta_2) \tag{14}$$

For the function $d_{(j)}(\hat{P}_{st}, m_j)$, for $j = 1, 2$ and $h(\hat{P}_{st}, m_1, m_2)$, under the conditions given in equation (10) the Bias(l_j), bias(l_3), MSE(l_j) and MSE$()l_3)$ are always exist.

4. Properties of the Proposed Classes of Estimators

Utilizing the large sample approximation, we assume that:

$$\begin{aligned} & \overline{y}_{1st}^* = \overline{Y}_1(1 + \epsilon_0), \overline{y}_{2st}^* = \overline{Y}_2(1 + \epsilon_1), \overline{x}_{st}^* = \overline{X}(1 + \epsilon_2) \text{ and} \\ & \overline{x}_{st} = \overline{X}(1 + \epsilon_3) \end{aligned} \tag{15}$$

Such that $|\epsilon_i| < 1$ and $E(\epsilon_i) = 0 \; \forall \; i = 0, 1, 2, 3$.

We also have,

$$E(\epsilon_0^2) = \frac{V(\overline{y}_{1st}^*)}{\overline{Y}_1^2} = \frac{1}{\overline{Y}_1^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{y_{1i}}^2 + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{1i(2)}}^2\right\},$$

$$E(\epsilon_1^2) = \frac{V(\overline{y}_{2st}^*)}{\overline{Y}_2^2} = \frac{1}{\overline{Y}_2^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{y_{2i}}^2 + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{2i(2)}}^2\right\},$$

$$E(\epsilon_3^2) = E(\epsilon_2\epsilon_3) = \frac{V(\overline{x}_{st})}{\overline{X}^2} = \frac{1}{\overline{X}^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{x_i}^2\right\},$$

$$E(\epsilon_2^2) = \frac{V(\overline{x}_{st}^*)}{\overline{X}^2} = \frac{1}{\overline{X}^2}\sum_{i=1}^{L}\left\{W_i^2 f_i S_{x_i}^2 + \frac{(k_i-1)}{n_i}W_{i2}S_{x_{i(2)}}^2\right\},$$

$$E(\epsilon_0\epsilon_1) = \frac{Cov(\overline{y}_{1st}^*\overline{y}_{1st}^*)}{\overline{Y}_1\overline{Y}_2} = \frac{1}{\overline{Y}_1\overline{Y}_2}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{1i}y_{2i}} + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{1i}y_{2i(2)}}\right\},$$

$$E(\epsilon_0\epsilon_2) = \frac{Cov(\overline{y}_{1st}^*\overline{x}_{st}^*)}{\overline{Y}_1\overline{X}} = \frac{1}{\overline{Y}_1\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{1i}x_i} + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{1i}x_{i(2)}}\right\},$$

$$E(\epsilon_1\epsilon_2) = \frac{Cov(\overline{y}_{2st}^*\overline{x}_{st}^*)}{\overline{Y}_2\overline{X}} = \frac{1}{\overline{Y}_2\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{2i}x_i} + \frac{(k_i-1)}{n_i}W_{i2}S_{y_{2i}x_{i(2)}}\right\},$$

$$E(\epsilon_0\epsilon_3) = \frac{Cov(\overline{y}_{1st}^*\overline{x}_{st})}{\overline{Y}_1\overline{X}} = \frac{1}{\overline{Y}_1\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{1i}x_i}\right\},$$

$$E(\epsilon_1\epsilon_3) = \frac{Cov(\overline{y}_{2st}^*\overline{x}_{st})}{\overline{Y}_2\overline{X}} = \frac{1}{\overline{Y}_2\overline{X}}\sum_{i=1}^{L}W_i^2\left\{f_i S_{y_{2i}x_i}\right\} \tag{16}$$

Where, $(S_{y_{1i}}^2, S_{y_{2i}}^2)$ represent the population mean square error of (y_1, y_2) for i^{th} stratum and $(S_{y_{1i(2)}}^2, S_{y_{2i(2)}}^2)$ are the population mean square of r_i units sub-sampled from $n_{i(2)}$ units of the study variable for i^{th} stratum and $(S_{x_i}^2, S_{x_{i(2)}}^2)$ represent the population mean square error of x for i^{th} stratum and r_i units sub-sampled from $n_{i(2)}$ for i^{th} stratum respectively.

$(S_{y_{1i}y_{2i}}, S_{y_{1i}x_i}, S_{y_{2i}x_i})$ and $(S_{y_{1i}y_{2i(2)}}, S_{y_{1i}x_{i(2)}}, S_{y_{2i}x_{i(2)}})$ are the covariance between $(S_{y_{1i}y_{2i}}, S_{y_{1i}x_i}, S_{y_{2i}x_i})$ and $(S_{y_{1i}y_{2i(2)}}, S_{y_{1i}x_{i(2)}}, S_{y_{2i}x_{i(2)}})$stratum of the population and r_i units sub sampled from $n_{i(2)}$ units of the population for the i^{th} stratum respectively and $f_i = \left(\frac{1}{n_i} - \frac{1}{N_i}\right)$.

The Bias and MSE of l_j and l_3 by using the equation (81) and (82) respectively up to the n^{-1} terms of order are given as below:

$$\begin{aligned} Bias(l_j) = E(l_j - P) &= E\Big[\hat{P}_{st} + (m_j-1)d_{2(j)}(\theta_1) \\ &\quad + P^{-1}(\hat{P}_{st} - P)(m_j-1)d_{2(j)}(\theta_1^*) - P\Big] \\ &= E(\hat{P}_{st} - P) + P^{-1}E(\hat{P}_{st} - P)(m_j-1)d_{2(j)}(\theta_1) \end{aligned} \tag{17}$$

$$\begin{aligned} MSE(l_j) = E(l_j - P)^2 &= E(\hat{P}_{st} - P)^2 + E(m_j-1)^2 d_{2(j)}^2(\theta_1) \\ &\quad + 2E(\hat{P}_{st} - P)(m_j-1)d_{2(j)}(\theta_1) \end{aligned} \tag{18}$$

$$\begin{aligned} Bias(l_3) = E(l_3 - P) &= E\Big[\hat{P}_{st} + (m_1-1)h_2(\theta_2) + (m_2-1)h_3(\theta_2) - P\Big] \\ &= E(\hat{P}_{st} - P) \end{aligned} \tag{19}$$

$$\begin{aligned} MSE(l_3) = E(l_3 - P)^2 &= E(\hat{P}_{st} - P)^2 + E(m_1-1)^2 h_2^2(\theta_2) + E(m_2-1)^2 h_3^2(\theta_2) \\ &\quad - 2E(\hat{P}_{st} - P)(m_1-1)h_2(\theta_2) + 2E(\hat{P}_{st} - P)(m_2-1)h_3(\theta_2) \\ &\quad + 2E(m_1-1)(m_2-1)h_2(\theta_2)h_3(\theta_2) \end{aligned} \tag{20}$$

The equation (86) can be written as:

$$MSE(l_3) = E(l_3 - P)^2 = MSE(\hat{P}_{st}) + Ah_2(\theta_2) + Bh_3(\theta_2) + 2Ch_2(\theta_2) + 2Dh_3(\theta_2) + 2Eh_2(\theta_2)h_3(\theta_2) \tag{21}$$

Where, $MSE(\hat{P}_{st}) = E(\hat{P}_{st} - P)^2$, $A = E(m_1 - 1)^2$, $B = E(m_2 - 1)^2$, $C = E(\hat{P}_{st} - P)(m_1 - 1)$, $D = E(\hat{P}_{st} - P)(m_2 - 1)$ and $E = E(m_1 - 1)(m_2 - 1)$.

Now, to get the optimum value of $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ We partially differentiate the expression (18) w.r.t. $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$ and equation (21) w.r.t $h_2(\theta_2)$ and $h_3(\theta_2)$ and equating to zero. Assuming that the partial derivatives of order second are positive, we get

$$d_{2(1)}(\theta_1)_{(opt)} = -\frac{\overline{X}}{V(\overline{x}^*_{st})}\left[\overline{Y}_2 Cov(\overline{y}^*_{1st}, \overline{x}^*_{st}) + \overline{Y}_1 Cov(\overline{y}^*_{2st}, \overline{x}^*_{st})\right] \tag{22}$$

$$d_{2(2)}(\theta_1)_{(opt)} = -\frac{\overline{X}}{V(\overline{x}_{st})}\left[\overline{Y}_2 Cov(\overline{y}^*_{1st}, \overline{x}_{st}) + \overline{Y}_1 Cov(\overline{y}^*_{2st}, \overline{x}_{st})\right] \tag{23}$$

$$h_2(\theta_2)_{opt} = \frac{(DE - BC)}{(AB - E^2)} \tag{24}$$

$$h_3(\theta_2)_{opt} = \frac{(CE - AD)}{(AB - E^2)} \tag{25}$$

The minimum mean square errors after putting the optimum values of $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ in equation (18) and (21) are given by:

$$MSE(l_1)_{min} = MSE(\hat{p}_{st}) - \frac{P^2}{V(\overline{x}^*_{st})}\left[\frac{Cov(\overline{y}^*_{1st}, \overline{x}^*_{st})}{\overline{Y}_1} + \frac{Cov(\overline{y}^*_{2st}, \overline{x}^*_{st})}{\overline{Y}_2}\right]^2 \tag{26}$$

$$MSE(l_2)_{min} = MSE(\hat{p}_{st}) - \frac{P^2}{V(\overline{x}_{st})}\left[\frac{Cov(\overline{y}^*_{1st}, \overline{x}_{st})}{\overline{Y}_1} + \frac{Cov(\overline{y}^*_{2st}, \overline{x}_{st})}{\overline{Y}_2}\right]^2 \tag{27}$$

$$MSE(l_3)_{min} = MSE(\hat{p}_{st}) - \left[\frac{BC^2 + AD^2 - 2CDE}{AB - E^2}\right] \tag{28}$$

Members of the proposed classes l_1, l_2 and l_3:

For the given condition in (9), any parametric function $d_{(1)}(\hat{P}_{st}, m_1)$, $d_{(2)}(\hat{P}_{st}, m_2)$ and $h(\hat{P}_{st}, m_2)$ can produce a class of asymptotic estimators. Such types of estimators have a very large number of classes. Some of the members are given below for the proposed classes of estimators:

$$\begin{aligned} &l_{j1} = \hat{P}_{st}(\lambda_1 + (1 - \lambda_1)m_j), l_{j2} = (\hat{P}_{st} + a_1(m_j - 1))m_j^{\beta} \\ &l_{j3} = \hat{P}_{st}(m_j)^{\alpha_2}, l_{j4} = \hat{P}_{st}(2 - m_j^{\alpha_3}), l_{j5} = \hat{P}_{st}\exp\left[\frac{m_j - 1}{m_j + 1}\right] \\ &l_{31} = \hat{P}_{st}(\gamma_1 + m_1\gamma_2 + (1 - \gamma_1 - \gamma_2)m_2), l_{32} = \hat{P}_{st}\exp[\alpha_4(m_1 - 1) \\ &+ \alpha_5(m_2 - 1)], l_{33} = \hat{P}_{st}m_1^{\alpha_6}m_2^{\alpha_7}, l_{32} = \hat{P}_{st}\frac{1}{2}\left[1 + \alpha_8 m_1^{\gamma_3} + (1 - \alpha_8)m_2^{\gamma_4}\right] \end{aligned} \tag{29}$$

The MSEs of the suggested classes of the estimators will attain the minimum value of the MSE given in equation (26), (27), (28) for the optimum value $d_{2(1)}(\theta_1)$, $d_{2(2)}(\theta_1)$, $h_2(\theta_2)$ and $h_3(\theta_2)$ given in equation (22), (23), (24) and (25). The member of the proposed classes of the estimators l_1, l_2 and l_3 given in equation (29) will also attain the same minimum MSE. The optimum values are occasionally in the form of some unknown parameters and occasionally in the form of the value of unknown constants these values can be obtained from past data (Reddy [13]) or can be estimated by sample values that do not affect the minimum MSE of the estimator up to the term of order $(\frac{1}{n})$ (Srivastva and Jhajj [19]).

5. Comparison of (l_1, l_2 and l_3) with Pertinent Estimators [$\hat{P}_{st}$]

In stratified random sampling, the estimator for P in the case of non-response is defined as:

$$\hat{P}_{st} = \overline{y}^*_{1st}\overline{y}^*_{2st} \tag{30}$$

The MSE of the $\hat{P}_{st}$ is given as:

$$MSE(\hat{P}) = P^2\left[\frac{V(\overline{y}^*_{1st})}{\overline{Y}_1^2} + \frac{V(\overline{y}^*_{2st})}{\overline{Y}_2^2} + 2\frac{Cov(\overline{y}^*_{1st}, \overline{y}^*_{2st})}{\overline{Y}_1\overline{Y}_2}\right] \tag{31}$$

6. Simulation Study:

To evaluate the characteristics of suggested classes of estimators, we perform a simulation study to artificially generate the population using normal distribution. For this study, we generate a population of size 4500. The following table (1) shows the distributions of the population generated to perform the study.

Table 1: *Sample Size and Distribution*

Strata No.	Stratum Size(N_i)	Sample Size (n_i)	Distribution of $y_{1h_i} \sim N(\mu, \sigma^2)$	Distribution of $y_{2h_i} \sim N(\mu, \sigma^2)$	Distribution of $x_{h_i} \sim N(\mu, \sigma^2)$
1	1800	600	N(500,81)	N(900,121)	N(600,100)
2	1200	400	N(300,64)	N(500,100)	N(300,121)
3	900	300	N(400,100)	N(400,64)	N(500,81)
4	600	200	N(500,121)	N(300,81)	N(300,64)

Here, we employ some transformations suggested by Reddy et al.[12]. To generate the population for study variables and auxiliary variable with some association. The following table (2) shows the required transformation and correlation to generate the random variables. Now, to estimate the approximate mean square error of the suggested classes of the estimators and estimator $\hat{P}_{st}$. We average the outcomes after 6000 iterations of the loop.

To calculate the approximate mean square error (AMSE) the formula is given as follows:

$$AMSE(m_1^*) = \frac{1}{6000}\sum_{t=1}^{6000}(m_1^* - P)^2$$

where, $m_1^* = \hat{P}_{st}$, l_1 l_2 *and* l_3.

Table 2: *Transformation and correlation:*

Strata no.	cor(y_{1h_i}, x_{h_i})	cor(y_{1h_i}, x_{h_i})	Transformed auxiliary variable x_{h_i} using y_{1h_i}	Transformed auxiliary variable x_{h_i} using y_{2h_i}
I	−0.58	−0.60	$x_{11i} = r_{y_1x_1}y_{11i} + x_{1i}\sqrt{1-r^2_{y_1x_1}}$	$x_{21i} = r_{y_2x_1}y_{21i} + x_{1i}\sqrt{1-r^2_{y_2x_1}}$
II	−0.62	−0.72	$x_{12i} = r_{y_1x_2}y_{12i} + x_{2i}\sqrt{1-r^2_{y_1x_2}}$	$x_{22i} = r_{y_2x_2}y_{22i} + x_{2i}\sqrt{1-r^2_{y_2x_2}}$
III	−0.70	−0.58	$x_{13i} = r_{y_1x_3}y_{13i} + x_{3i}\sqrt{1-r^2_{y_1x_3}}$	$x_{23i} = r_{y_2x_3}y_{23i} + x_{3i}\sqrt{1-r^2_{y_2x_3}}$
IV	−0.65	−0.65	$x_{14i} = r_{y_1x_4}y_{14i} + x_{4i}\sqrt{1-r^2_{y_1x_4}}$	$x_{24i} = r_{y_2x_4}y_{24i} + x_{4i}\sqrt{1-r^2_{y_2x_4}}$

Table 3: *Percentage relative efficiency (PRE) of the proposed estimator with respect to relevant estimators.*

Rate of Non-response	Estimators	$1/k$		
		1/2	1/3	1/4
10%	$\hat{P}_{st}$	100(3991.83)	100(4484.76)	100(5022.54)
	l_1	117.97(3383.59)	118.25(3792.38)	118.49(4238.51)
	l_2	115.26(3463.26)	113.39(3954.87)	111.83(4491.21)
	l_3	118.22(3376.57)	118.80(3774.79)	119.45(4204.63)
20%	$\hat{P}_{st}$	100(4509.60)	100(5508.47)	100(6616.83)
	l_1	118.06(3819.70)	118.10(4664.08)	118.08(5603.45)
	l_2	113.21(3983.25)	110.62(4979.58)	108.69(6087.46)
	l_3	118.26(3813.21)	118.40(4652.04)	115.57(5580.08)
30%	$\hat{P}_{st}$	100(5034.18)	100(6578.39)	100(8146.68)
	l_1	117.98(4266.66)	118.24(5563.51)	118.54(6872.50)
	l_2	111.72(4505.72)	108.73(6049.82)	106.93(7618.18)
	l_3	118.14(4261.11)	118.47(5552.53)	118.85(6854.30)

(Note: The AMSE of the estimators are shown in parenthesis.)

Table (3) depicts the $AMSE$ of the estimators $\hat{P}_{st}$, l_1, l_2 and l_3 for different values of k and non-response rates.

From the above Table (3) we can see that as we increase the value of k and non-response rates from 10% to 30% the approximate means square error of the estimators $\hat{P}_{st}$, l_1, l_2 and l_3 increases.

7. Conclusion

The proposed estimators l_1, l_2 and l_3 are found to be more efficient than the estimator $\hat{P}_{st}$. And it is also observed that from Table (3) the proposed classes of estimators l_1 and l_3 are found to be almost equally efficient for the different choices of sub-sampling fraction and non-response rates.

The $AMSE(l_1)$ less than or greater than $AMSE(l_2)$ depending upon the situation of correlation between study variables y_1, y_2 and auxiliary variable x. Rao[10] has shown the situation when the conventional estimator using $\overline{x}^*$ and alternate estimator using $\overline{x}$ will be less than each other depending upon the situations given by him. This theory also works here in the case of l_1, l_2 and l_3.

Here, some times $AMSE(l_1) < AMSE(l_2)$ then $AMSE(l_3)$ will be less than $AMSE(l_2)$ and almost equal or less than $AMSE(l_1)$. But if $AMSE(l_1) > AMSE(l_2)$ then $AMSE(l_3)$ will be less than $AMSE(l_1)$ and almost equal to the $AMSE(l_2)$. So the use of $AMSE(l_3)$ is advisable if the condition when $AMSE(l_1) < AMSE(l_2)$ or $AMSE(l_1) > AMSE(l_2)$ is unknown.

Hence, for the estimation of P under the stratified random sampling in the existence of non-response for the Known $\overline{X}$. We suggest to use the estimators l_1, l_2 and l_3 depending on the situations discussed in the results.

References

[1] El-Badry, M. A. (1956).A sampling procedure for mailed questionnaires. *Journal of the American Statistical Association* , 51(274), 209-227.

[2] Hansen, M. H., Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, 41(236), 517-529.

[3] Khare, B. B. (1992). On class of estimators for product of two population means using multi-auxiliary characters with known and unknown means, *.Jour. Appl. Stat*, 1, 56-67.

[4] Khare, B. B., and R. R. Sinha(2010). On class of estimators for the product of two population means using auxiliary character in presence of non-response."*Inter. Trans. Appl. Sci* 2.4 841-846.

[5] Khare, B. B., & Sinha, R. R. (2012).Combined class of estimators for ratio and product of two population means in presence of non-response.*International Journal of Statistics and Economics*, 8(12), 12-20.

[6] Khare, B. B., Jha, P. S., & Kumar, K. (2014). Improved generalized chain estimators for ratio and product of two population means using two auxiliary characters in the presence of non-response. *International J. Stats & Economics*, 13(1), 108-121.

[7] Khare, B. B., Jha, P. S. (2017). Classes of estimators for population mean using auxiliary variable in stratified population in the presence of non-response. *Communications in Statistics-Theory and Methods*, 46(13), 6579-6589.

[8] Kumar, K. Srivastava, U. (2018). Estimation of ratio and product of two population means using exponential type estimators in sample surveys. *International Journal of Mathematics and Statistics*, 19‘‘3, pp. 102‘‘109.

[9] Mishra, M., Khare, B. B., Singh, S. (2023). TWO-CLASSES FOR REGRESSION TYPE OF ESTIMATORS FOR THE RATIO OF TWO POPULATION MEANS IN TWO-PHASE SAMPLING IN THE PRESENCE OF NON-RESPONSE FOR STRATIFIED POPULATION. *Reliability: Theory Applications*, 18(3 (74)), 610-620.

[10] Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the non-respondents. *Survey methodology*, 12(2), 217-230.

[11] Ray, S. K., & Singh, R. K. (1985). Some estimators for the ratio and product of population parameters. *Journal of the Indian Society of Agricultural Statistics*, 37(1), 1-10.

[12] Reddy, M. K., Rao, K. R., & Boiroju, N. K. (2010). Comparison of ratio estimators using Monte Carlo simulation. *International Journal of Agriculture and Statistical Sciences*, 6(2), 517-527.

[13] Reddy, V. N. (1978). A study on the use of prior knowledge on certain population parameters in estimation. *Sankhya C*, 40, 29-37.

[14] Singh, M. P. (1965). On the estimation of ratio and product of the population parameters. *Sankhy´: The Indian Journal of Statistics*, Series B, 321-328.

[15] Singh, M. P. (1967). cum product method of estimation. *Metrika*, 12(1), 34-42.

[16] Singh, M. P. (1969). of some ratio-cum-product estimators. *Sankhy´: The Indian Journal of Statistics, Series B*, 375-378.

[17] Singh, R.K. (1982). the estimating ratio and product of population parameters. . *Stat. Assoc. Bull*, 31, 69-76.

[18] Singh, R. , Khare, S. , Khare, B.B. Jha , P.S. (2020). The general classes of estimators for population mean under stratified two phase random sampling in the presence of non-response. *Int. J. Agricult. Stat. Sci.*, Vol. 16, No. 2, pp. 557-565, 2020.

[19] Srivastava, S. K., Jhajj, H. S. (1983). A class of estimators of the population mean using multiauxiliary information. *Calcutta Statistical Association Bulletin*, 32(1-2), 47-56.

[20] Srivastava, S. R., Khare, B. B., & Srivastava, S. R. (1988). On generalized chain estimator for ratio and product of two population means using auxiliary characters. *Assam Statistical Review*, 2(1), 21-29.

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

Keerthiga S[1] 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 4^{th}-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 queuing 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 1^{st} and 2^{nd} 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 1^{st} and 2^{nd} 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 $\mathcal{B}(0) = 0$, $\mathcal{B}(\infty) = 1$, $\mathcal{C}(0) = 0$, $\mathcal{C}(\infty) = 1$, 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

$$\eta(\tilde{\varphi})d\tilde{\varphi} = \frac{d\mathcal{B}(\tilde{\varphi})}{1-\mathcal{B}(\tilde{\varphi})}$$
$$\zeta(\tilde{\varphi})d\tilde{\varphi} = \frac{d\mathcal{C}(\tilde{\varphi})}{1-\mathcal{C}(\tilde{\varphi})}$$
$$\kappa(\tilde{\varphi})d\tilde{\varphi} = \frac{d\mathcal{H}_v(\tilde{\varphi})}{1-\mathcal{H}_v(\tilde{\varphi})}$$
$$v(\tilde{\varphi})d\tilde{\varphi} = \frac{d\mathcal{I}(\tilde{\varphi})}{1-\mathcal{I}(\tilde{\varphi})}$$

$$L(\xi) = \begin{cases} 0, & \text{if the server is free} \\ 1, & \text{if the server is active period} \\ 2, & \text{if the server is operative mode on WV period} \\ 3, & \text{if the server is on repair} \end{cases}$$

Thus, the state of the system $\mathcal{B}^0(\xi)$,$\mathcal{C}^0(\xi)$, $\mathcal{H}_v^0(\xi)$, 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 $\{\xi_\sigma; \sigma = 1, 2, ...\}$ 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} x N.

3.2. Theorem

The embedded Markov chain $\{U_\sigma; \sigma \epsilon N\}$ is ergodic iff $\rho < \bar{\mathcal{B}}(\mu)$ for our system to be stable, where $\rho = \lambda\mu\mathcal{C}_1 + \bar{\lambda}(1 + \mu\mathcal{I}_1) + \lambda\omega$.

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 $\chi_0 = lim_{\xi\to\infty}\chi_0(\xi)$ and limiting densities are
$\chi_\sigma(\tilde{\varphi}) = lim_{\xi\to\infty}\chi_\sigma(\tilde{\varphi}, \xi)$ for $\tilde{\varphi} \geq 0$ and $\sigma \geq 1$.
$\Psi_\sigma(\tilde{\varphi}) = lim_{\xi\to\infty}\Psi_\sigma(\tilde{\varphi}, \xi)$ for $\tilde{\varphi} \geq 0$ and $\sigma \geq 0$.
$\Lambda_{v,\sigma}(\tilde{\varphi}) = lim_{\xi\to\infty}\Lambda_{v,\sigma}(\tilde{\varphi}, \xi)$ for $\tilde{\varphi} \geq 0$ and $\sigma \geq 0$.
$\Pi_\sigma(\tilde{\varphi}) = lim_{\xi\to\infty}\Pi_\sigma(\tilde{\varphi}, \xi)$ for $\tilde{\varphi} \geq 0$ and $\sigma \geq 1$.

Applying the SVT, we create the following system of equations.

$$\mu\chi_0 = \gamma r_1\phi_0 \tag{1}$$

$$(\mu+\gamma)\phi_0 = \gamma r_2\phi_0\int_0^\infty \Lambda_{v,0}(\tilde{\varphi})\kappa(\tilde{\varphi})d\tilde{\varphi} + \int_0^\infty \Psi_\sigma(\tilde{\varphi})\zeta(\tilde{\varphi})d\tilde{\varphi} \tag{2}$$

$$\frac{d}{d\tilde{\varphi}}\chi_\sigma(\tilde{\varphi}) + (\mu+\eta(\tilde{\varphi}))\chi_\sigma(\tilde{\varphi}) = 0, \sigma \geq 1 \tag{3}$$

$$\frac{d}{d\tilde{\varphi}}\Psi_0(\tilde{\varphi}) + (\mu+\zeta(\tilde{\varphi}))\Psi_0(\tilde{\varphi}) = 0, \sigma = 0. \tag{4}$$

$$\frac{d}{d\tilde{\varphi}}\Psi_\sigma(\tilde{\varphi}) + (\mu+\zeta(\tilde{\varphi}))\Psi_\sigma(\tilde{\varphi}) = \mu\Psi_{\sigma-1}(\tilde{\varphi}), \sigma \geq 1 \tag{5}$$

$$\frac{d}{d\tilde{\varphi}}\Lambda_{0,v}(\tilde{\varphi}) + (\mu+\gamma+\kappa(\tilde{\varphi}))\Lambda_{0,v}(\tilde{\varphi}) = 0, \sigma = 0. \tag{6}$$

$$\frac{d}{d\tilde{\varphi}}\Lambda_{\sigma,v}(\tilde{\varphi}) + (\mu+\gamma+\kappa(\tilde{\varphi}))\Lambda_{\sigma,v}(\tilde{\varphi}) = \mu\Lambda_{v,\sigma-1}(\tilde{\varphi}), \sigma \geq 1. \tag{7}$$

$$\frac{d}{d\tilde{\varphi}}\Pi_0(\tilde{\varphi}) + (\mu+v(\tilde{\varphi}))\Pi_0(\tilde{\varphi}) = 0, \sigma = 0. \tag{8}$$

$$\frac{d}{d\tilde{\varphi}}\Pi_\sigma(\tilde{\varphi}) + (\mu+v(\tilde{\varphi}))\Pi_\sigma(\tilde{\varphi}) = \mu\Pi_{\sigma-1}(\tilde{\varphi}), \sigma \geq 1. \tag{9}$$

At $\tilde{\varphi} = 0$ the steady state boundary conditions are as follows:

$$\chi_\sigma(0) = \bar{\omega}\int_0^\infty \Psi_\sigma(\tilde{\varphi})\zeta(\tilde{\varphi})d\tilde{\varphi} + \omega\int_0^\infty \Psi_{\sigma-1}(\tilde{\varphi})\zeta(\tilde{\varphi})d\tilde{\varphi} + \bar{\omega}\int_0^\infty \Lambda_{v,\sigma}(\tilde{\varphi})\kappa(\tilde{\varphi})d\tilde{\varphi} + \omega\int_0^\infty \Lambda_{v,\sigma-1}(\tilde{\varphi})\kappa(\tilde{\varphi})d\tilde{\varphi} + \int_0^\infty \Pi_\sigma(\tilde{\varphi})v(\tilde{\varphi})d\tilde{\varphi} \tag{10}$$

$$\Psi_0(0) = \lambda\int_0^\infty \chi_1(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \lambda\bar{\mu}\chi_0 + \gamma\int_0^\infty \Lambda_{0,v}(\tilde{\varphi})d\tilde{\varphi}, \sigma = 0 \tag{11}$$

$$\Psi_\sigma(0) = \lambda\int_0^\infty \chi_{\sigma+1}(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \lambda\mu\int_0^\infty \Psi_\sigma(\tilde{\varphi})d\tilde{\varphi} + \gamma\int_0^\infty \Lambda_{\sigma,v}(\tilde{\varphi})d\tilde{\varphi}, \sigma \geq 1 \tag{12}$$

$$\Lambda_{v,\sigma}(0) = \begin{cases} \mu\phi_0, & \sigma = 0 \\ 0, & \sigma \geq 1 \end{cases} \tag{13}$$

$$\Pi_1(0) = \bar{\lambda}\int_0^\infty \chi_1(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \bar{\lambda}\mu\chi_0 \tag{14}$$

$$\Pi_\sigma(0) = \bar{\lambda}\int_0^\infty \chi_\sigma(\tilde{\varphi})\eta(\tilde{\varphi})d\tilde{\varphi} + \bar{\lambda}\mu\int_0^\infty \chi_{n-1}(\tilde{\varphi})d\tilde{\varphi}, \sigma \geq 2 \tag{15}$$

The normalizing condition is

$$\chi_0 + \phi_0 + \sum_{\sigma=1}^{\infty}\int_0^\infty \chi_\sigma(\tilde{\varphi})d\tilde{\varphi} + \sum_{\sigma=0}^{\infty}\int_0^\infty \Psi_\sigma(\tilde{\varphi})d\tilde{\varphi} + \sum_{\sigma=0}^{\infty}\int_0^\infty \Lambda_{\sigma,v}(\tilde{\varphi})d\tilde{\varphi} + \sum_{\sigma=1}^{\infty}\int_0^\infty \Pi_\sigma(\tilde{\varphi})d\tilde{\varphi} = 1 \tag{16}$$

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:

$$\chi(\tilde{\varphi},\vartheta) = \sum_{\sigma=1}^{\infty} \chi_\sigma(\tilde{\varphi})\vartheta^\sigma; \chi(0,\vartheta) = \sum_{\sigma=1}^{\infty} \chi_\sigma(0)\vartheta^\sigma;$$
$$\Psi(\tilde{\varphi},\vartheta) = \sum_{\sigma=0}^{\infty} \Psi_\sigma(\tilde{\varphi})\vartheta^\sigma; \Psi(0,\vartheta) = \sum_{n=0}^{\infty} \Psi_0(0)\vartheta^\sigma; i = 1,2$$
$$\Lambda_v(\tilde{\varphi},\vartheta) = \sum_{\sigma=0}^{\infty} \Lambda_{v,\sigma}(\tilde{\varphi})\vartheta^\sigma; \Lambda_v(0,\vartheta) = \sum_{\sigma=0}^{\infty} \Lambda_{v,\sigma}(0)\vartheta^\sigma;$$
$$\Pi(\tilde{\varphi},\vartheta) = \sum_{\sigma=1}^{\infty} \Pi_\sigma(\tilde{\varphi})\vartheta^\sigma; \Pi(0,\vartheta) = \sum_{\sigma=1}^{\infty} \Pi_\sigma(0)\vartheta^\sigma$$

Next multiply the SS eqn. and SS boundary conditions from (3) to (15) by ϑ^σ and adding over σ, $(\sigma = 0,1,2,...)$

$$\frac{\partial}{\partial\tilde{\varphi}}\chi(\tilde{\varphi},\vartheta) + [\mu + \eta(\tilde{\varphi})]\chi(\tilde{\varphi},\vartheta) = 0 \tag{17}$$
$$\frac{\partial}{\partial\tilde{\varphi}}\Psi(\tilde{\varphi},\vartheta) + [\mu(1-\vartheta) + \zeta(\tilde{\varphi})]\Psi(\tilde{\varphi},\vartheta) = 0 \tag{18}$$
$$\frac{\partial}{\partial\tilde{\varphi}}\Lambda_v(\tilde{\varphi},\vartheta) + [\gamma + \mu(1-\vartheta) + \kappa(\tilde{\varphi})]\Lambda_v(\tilde{\varphi},\vartheta) = 0 \tag{19}$$
$$\frac{\partial}{\partial\tilde{\varphi}}\Pi(\tilde{\varphi},\vartheta) + [\mu(1-\vartheta) + v(\tilde{\varphi})]\Pi(\tilde{\varphi},\vartheta) = 0 \tag{20}$$

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

$$\chi(0,\vartheta) = (\bar{\omega} + \omega\vartheta)\int_0^\infty \Psi(\tilde{\varphi},\vartheta)\zeta(\tilde{\varphi})d\tilde{\varphi} + (\bar{\omega} + \omega\vartheta)\int_0^\infty \Lambda_v(\tilde{\varphi},\vartheta)\kappa(\tilde{\varphi})d\tilde{\varphi} + \int_0^\infty \Pi(\tilde{\varphi},\vartheta)v(\tilde{\varphi})d\tilde{\varphi} - (\mu + \gamma r_1)\phi_0 \tag{25}$$

$$\Psi(0,\vartheta) = \frac{\lambda}{\vartheta}\int_0^\infty \chi(\tilde{\varphi},\vartheta)\eta(\tilde{\varphi})d\tilde{\varphi} + \lambda\mu\int_0^\infty \chi(\tilde{\varphi},\vartheta)d\tilde{\varphi} + \gamma\int_0^\infty \Lambda_v(\tilde{\varphi},\vartheta)d\tilde{\varphi} + \lambda\mu\chi_0 \tag{26}$$
$$\Lambda_v(0,\vartheta) = \mu\phi_0 \tag{27}$$
$$\Pi(0,\vartheta) = \bar{\lambda}\vartheta\mu\int_0^\infty \chi(\tilde{\varphi},\vartheta)d\tilde{\varphi} + \bar{\lambda}\int_0^\infty \chi(\tilde{\varphi},\vartheta)\eta(\tilde{\varphi})d\tilde{\varphi} + \vartheta\mu\bar{\lambda}\chi_0 \tag{28}$$

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)

$$\Pi(0,\vartheta) = \vartheta\gamma r_1\bar{\lambda}\chi_0 + \bar{\lambda}\chi(0,\vartheta)[\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)] \tag{30}$$

Substituting equations (22),(23) and (24) in (25), we obtain

$$\chi(0,\vartheta) = (\bar{\omega}+\omega\vartheta)\Psi(0,\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + (\bar{\omega}+\omega\vartheta)\Lambda_v(0,\vartheta)\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \Pi(0,\vartheta)\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0 \tag{31}$$

Using equations (27),(29) and (30) in equation (31)

$$\chi(0,\vartheta) = \vartheta\left\{\frac{(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]}\right\} \tag{32}$$

substituting equation (32) in (29) and (30),we obtain

$$\Psi(0,\vartheta) = \left\{\frac{\lambda[(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)] + \mu\phi_0\mathcal{V}(\vartheta) + \lambda\gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]}\right\} \tag{33}$$

$$\Pi(0,\vartheta) = \left\{\frac{\bar{\lambda}[(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0] + \vartheta\bar{\lambda}\gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]}\right\} \tag{34}$$

Substituting equations (27) and (32) to (34) in (21) to (24)

$$\chi(\tilde{\varphi},\vartheta) = \vartheta\left\{\frac{(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]}\right\} \times [1-\mathcal{B}(\tilde{\varphi})]e^{-\mu\tilde{\varphi}} \tag{35}$$

$$\Psi(\tilde{\varphi},\vartheta) = \left\{\frac{\lambda[(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)] + \mu\phi_0\mathcal{V}(\vartheta) + \lambda\gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]}\right\} \times [1-\mathcal{C}(\tilde{\varphi})]e^{-\mathcal{F}(\vartheta)\tilde{\varphi}} \tag{36}$$

$$\Lambda_v(\tilde{\varphi},\vartheta) = \mu\phi_0[1-\mathcal{H}_v(\tilde{\varphi})]e^{-\mathcal{F}_v(\vartheta)\tilde{\varphi}} \tag{37}$$

$$\Pi(\tilde{\varphi},\vartheta) = \left\{\frac{\bar{\lambda}[(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0] + \vartheta\bar{\lambda}\gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]}\right\} \times [1-\mathcal{I}(\tilde{\varphi})]e^{-\mathcal{F}(\vartheta)\tilde{\varphi}} \tag{38}$$

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 $\rho < \bar{\mathcal{B}}(\mu)$ under the stability condition

$$\chi(\vartheta) = \vartheta \left\{ \frac{(\bar{\omega}+\omega\vartheta)\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\frac{\bar{\lambda}}{\mu}\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\frac{\bar{\lambda}}{\mu}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \phi_0 - \frac{\gamma r_1}{\mu}\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]} \right\} \times [1-\bar{\mathcal{B}}(\mu)] \tag{39}$$

$$\Psi(\vartheta) = \left\{ \frac{\lambda[(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)] + \mu\phi_0\mathcal{V}(\vartheta) + \lambda\gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]} \right\} \times \frac{[1-\bar{\mathcal{C}}(\mathcal{F}(\vartheta))]}{\mu(1-\vartheta)} \tag{40}$$

$$\Lambda_v(\vartheta) = \frac{\mu\phi_0\mathcal{V}(\vartheta)}{\gamma} \tag{41}$$

$$\Pi(\vartheta) = \left\{ \frac{\bar{\lambda}[(\bar{\omega}+\omega\vartheta)\mu\phi_0[\bar{\mathcal{H}}_v(\mathcal{F}_v(\vartheta)) + \mathcal{V}(\vartheta)\bar{\mathcal{C}}(\mathcal{F}(\vartheta))] + (\bar{\omega}+\omega\vartheta)\lambda\gamma r_1\phi_0\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\gamma r_1\phi_0\bar{\mathcal{I}}(\mathcal{F}(\vartheta)) - \mu\phi_0 - \gamma r_1\phi_0] + \vartheta\bar{\lambda}\gamma r_1\phi_0}{\vartheta - [(\bar{\omega}+\omega\vartheta)\lambda\bar{\mathcal{C}}(\mathcal{F}(\vartheta)) + \vartheta\bar{\lambda}\bar{\mathcal{I}}(\mathcal{F}(\vartheta))][\vartheta + (1-\vartheta)\bar{\mathcal{B}}(\mu)]} \right\} \times \frac{[1-\bar{\mathcal{I}}(\mathcal{F}(\vartheta))]}{\mu(1-\vartheta)} \tag{42}$$

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 $\rho < \bar{\mathcal{B}}(\mu)$ 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

$$H_s(\vartheta) = \frac{Ne_s(\vartheta)}{De_s(\vartheta)} \tag{43}$$

$$H_0(\vartheta) = \frac{Ne_0(\vartheta)}{De_s(\vartheta)} \tag{44}$$

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 $\vartheta \to 1$ 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)

$$\chi(1) = \phi_0[1-\bar{\mathcal{B}}(\mu)]\left\{\frac{[[\frac{\mu}{\gamma}[1-\bar{\mathcal{H}}_v(\gamma)]-\mu\mathcal{C}_1[1-\bar{\mathcal{H}}_v(\gamma)]]-\frac{\lambda}{\mu}\gamma r_1[\omega+\mu\mathcal{C}_1]+\frac{\bar{\lambda}}{\mu}\gamma r_1[1-\mu\mathcal{I}_1]}{\bar{\mathcal{B}}(\mu)+\lambda\omega-\lambda\mu\mathcal{C}_1-\bar{\lambda}(1+\mu\mathcal{I}_1)}\right\} \tag{45}$$

(ii)Pr(The server is operative on usual service period)

$$\Psi(1) = \left\{\frac{\phi_0\lambda\mathcal{C}_1[[[\frac{\mu}{\gamma}[1-\bar{\mathcal{H}}_v(\gamma)]-\mu\mathcal{C}_1[1-\bar{\mathcal{H}}_v(\gamma)]]]-\lambda\gamma r_1[\omega+\mu\mathcal{C}_1+\bar{\lambda}\gamma r_1[1-\mu\mathcal{I}_1]+\mu[\mu\mathcal{H}_1+\frac{\mu}{\gamma}[1-\bar{\mathcal{H}}_v(\gamma)]]]}{\bar{\mathcal{B}}(\mu)+\lambda\omega-\lambda\mu\mathcal{C}_1-\bar{\lambda}(1+\mu\mathcal{I}_1)}\right\} \tag{46}$$

(iii)Pr(The server is on WV)

$$\Lambda_v(1) = \frac{\phi_0\mu[1-\bar{\mathcal{H}}_v(\gamma)]}{\gamma} \tag{47}$$

(iv)Pr(The server is under repair time during usual active period)

$$\Pi = \Pi(1) = \left\{\frac{\phi_0\bar{\lambda}\mathcal{I}_1[[[\frac{\mu}{\gamma}[1-\bar{\mathcal{H}}_v(\gamma)]-\mu\mathcal{C}_1[1-\bar{\mathcal{H}}_v(\gamma)]]]-\lambda\gamma r_1[\omega+\mu\mathcal{C}_1]+\bar{\lambda}\gamma r_1[1-\mu\mathcal{I}_1]+\bar{\lambda}\gamma r_1]}{\bar{\mathcal{B}}(\mu)+\lambda\omega-\lambda\mu\mathcal{C}_1-\bar{\lambda}(1+\mu\mathcal{I}_1)}\right\} \tag{48}$$

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 $\vartheta = 1$.

$$L_q = H_o^{'}(1) = \lim_{\vartheta\to 1}\frac{d}{d\vartheta}H_o(\vartheta) = \phi_0\left[\frac{Ne_q^{'''}(1)De_q^{''}(1)-De_q^{'''}(1)Ne_q^{''}(1)}{3(De_q^{''}(1))^2}\right] \tag{49}$$

(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.

$$L_s = H_s^{'}(1) = \lim_{\vartheta\to 1}\frac{d}{d\vartheta}H_s(\vartheta) = \phi_0\left[\frac{Ne_s^{'''}(1)De_q^{''}(1)-De_q^{'''}(1)Ne_q^{''}(1)}{3(De_q^{''}(1))^2}\right] \tag{50}$$

(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

$$\phi_0 = \frac{A(T_0)}{A(T_b)+A(T_0)}; A(T_b) = \frac{1}{\mu}\left(\frac{1}{\phi_0}-1\right); A(T_c) = \frac{1}{\mu\phi_0} = A(T_0)+A(T_b). \tag{51}$$

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 $\bar{\mathcal{B}}(\mu) \to 1$. 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 $\omega = 0.19$, $\mu = 0.9$, $\lambda = 0.8$, $\gamma = 3$, $\bar{\mathcal{H}}_v(\gamma) = 0.9$, $r_1 = 0.5$.
Table 2 demonstrates that the vacation rate (γ) mounts,ϕ_0 increases,L_q, W_q subsides for the value of $\omega = 0.19$, $\mu = 1.5$, $\lambda = 0.19$, $\bar{\mathcal{H}}_v(\gamma) = 0.9$, $r_1 = 0.9$.
Table 3 clearly displays that feedback rate (ω) mounts, χ_0, L_q, W_q diminshes for the value of $\mu = 0.9$, $\lambda = 0.10$, $\bar{\mathcal{H}}_v(\gamma) = 0.9$, $r_1 = 0.19$, $\gamma = 0.9$.

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

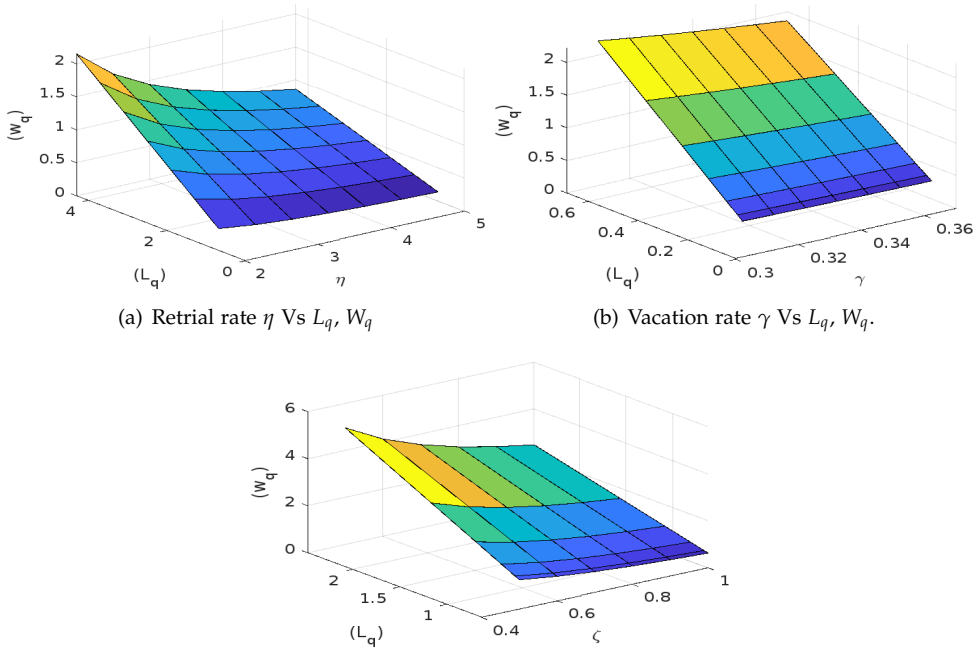
The Figure 1 (a) indicates that retrial rate (η) escalates, (L_q) and (W_q) increases. The Figure 1 (b) displays that vacation rate (γ) escalates, (L_q) and (W_q) decreases. The Figure 1 (c)

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



(a) Retrial rate η Vs L_q, W_q

(b) Vacation rate γ Vs L_q, W_q.

(c) Feedback rate ω Vs L_q, W_q.

Figure 1: *Effects of a few parameters on 3D representation.*

demonstrates that feedback rate (ω) increases, (L_q) and (W_q) diminshes.

We may use the numerical findings above to determine the influence of features on the

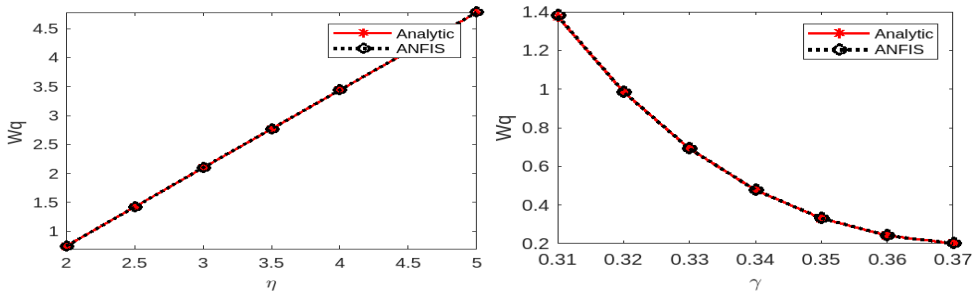
system's assessment criteria with certainty, the outcomes correspond to actual circumstances.

6.1. ANFIS Computing

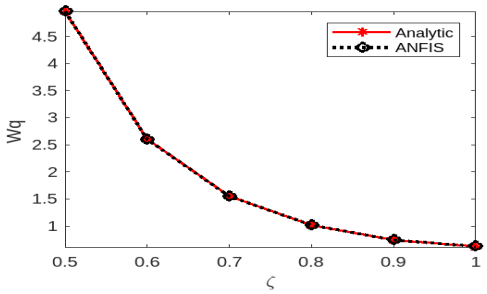
Using the fuzzy toolbox of MATLAB software, an ANFIS network can be executed to compare outcomes from analyses. ANFIS, a soft computing approach, is an effective tool for identifying important results that are useful in busy everyday environments. This approach aids in the identification of approximate 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 $(\omega))$, as shown in the numerical results in Figure 2 $(a-c)$. We consider the parameters $(\eta, \gamma$ and $\omega)$ 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.



(a) Retrial rate η Vs W_q (b) vacation rate γ Vs W_q



(c) Feedback rate ω Vs W_q

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.*

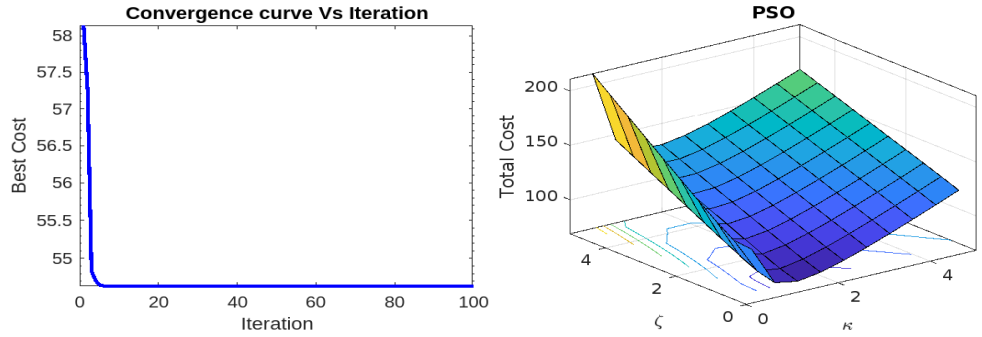
Parameters		(TEC^*)		
		Cost set 1	Cost set 2	Cost set 3
μ	0.20	\$75.1471	\$59.7276	\$66.7365
	0.25	\$115.1554	\$87.6434	\$92.1116
	0.30	\$151.7660	\$114.6657	\$119.3708
ω	1.00	\$73.5533	\$58.6272	\$65.7761
	1.15	\$74.5054	\$59.2872	\$66.3535
	1.20	\$74.8255	\$59.5074	\$66.5452
γ	1.25	\$78.6575	\$61.4278	\$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.



(a) Convergence in PSO (b) Expected Total Cost

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., Saravanarajan,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 Saravanarajan,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.,Saravanarajan,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 Queuing 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 retrial queue with feedback and starting failures.*Applied Mathematical Modelling*26:1057-1075.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.,Saravanarajan,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 working 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,Amita 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

METHOD AND ALGORITHM FOR QUANTITATIVE ANALYSIS OF AVERAGE MONTHLY VALUES OF THE OPERATIONAL RELIABILITY OF OVERHEAD POWER LINES

Farhadzadeh E.M., Muradaliyev A.Z., Abdullayeva S.A.
•
Azerbaijan Scientific-Research and Design-Prospecting Institute of Energetic
e-mail: elmeht@rambler.ru

Abstract

The relevance of ensuring the efficiency of equipment, devices and installations (object) of electric power systems increases every year and becomes the most important problem of maintaining energy security. The decrease in work efficiency is due to a number of factors, but, first, an increase in the relative number of objects, the service life of which exceeds the standard value. An illustration of the methodology for quantifying, comparing and ranking the monthly average values of indicators of the operational reliability of 110 kV overhead power transmission lines and above given in order to identify and restore the wear of the least reliable lines.

Keywords. Analysis, operational reliability, overhead power lines, automated system, classification, features, varieties

I. Introduction

At present, the service life of more than half of overhead power transmission lines (hereinafter - OHL) of electric power systems (hereinafter - EPS) exceeds the standard value, which leads to a decrease in their efficiency [1]. To limit the consequences of this change, risk oriented approaches are being developed for organizing their maintenance and repair (hereinafter - MRO) [2-4]. Their essence boils down to the theory of production assets management by ensuring a balance between operating costs and the risk of damage. It is known, that [5]:

– there is no methodology for calculating the technical condition index (hereinafter - TC). There is no monitoring of the TC overhead power lines, the methods used to assess the risk of damage are subjective;

– there are no operational recommendations to improve the efficiency of the overhead transmission lines.

TC OHL determines the reliability and safety of their work. Therefore, the possibility of assessing the indicators of operational reliability by analogy with the individual reliability of power units is relevant [6]. The apparent simplicity of this solution is deceptive, and, first of all, because there are no statistical data on continuous monitoring of the TC of overhead transmission lines, and according to statistical data on the failure within one month, it is impossible to assess the reliability indicators of specific overhead transmission lines due to their small number. When comparing and ranking indicators of operational reliability, the use of the mathematical apparatus for analyzing homogeneous statistical data is unacceptable, because the data are multidimensional and scarce [7].

We also shall remind that:

– the efficiency of functioning of EPS facilities today is understood as a joint accounting of efficiency, reliability of operation and safety of service;
– such an expanded concept of efficiency is due to the increase in the number of objects requiring operational inspection not only for efficiency, but also for reliability and safety;
– operational survey is understood as a quantitative assessment of work efficiency;
– if, for design purposes, a methodology for quantitative assessment of reliability indicators been developed and is widely used in practice to compare the design options of objects under design, then there is no methodology for calculating operational reliability indicators for solving operational problems. The science intensity, cumbersomeness and laboriousness of calculating operational reliability indicators requires a transition to automated systems for assessing, comparing and ranking, with monthly submission to the management of EPS facilities and electrical enterprises of guidelines for increasing the TC. The implementation of automated systems for analyzing TC of EPS facilities allows a large power grid enterprise to save tens of millions of rubles [5];
– unfortunately, there is no methodology for quantitative assessment of service safety indicators not only when solving operational problems, but also when designing EPS facilities [5];
– identity of the methodological approach to the assessment, comparison and ranking of EPS objects does not affect the discrepancy between the calculation algorithms due to the fundamental features of the functioning of these objects.

II. Initial data and indicators of operational reliability of OHL of EPS.

Initial data on OHL represented by constant and variable parts. The permanent part is compiled on the basis the passport data of the OHL and includes the name of the line, the name of the electro network enterprise (hereinafter - ENE), the nominal voltage, the year of commissioning, the material of the supports, the length of the line, the number of circuits. The variable part is compiled on the basis of operational logs and includes the following information about the change in the state of the OHL: name of the OHL, rated voltage, date (month, day, hour) of shutdown and activation, type of shutdown (emergency, by emergency or planned request). This information is entered into special tables of the database (we will designate them, respectively, as tables A and B) and are used in assessing the indicators of the operational reliability of the OHL EPS. These averaged indicators theoretically allow us to compare the operational reliability of a number of EPS and are necessary to control the nature of changes in the reliability of OHL in the calculated month in comparison with the reliability in the previous month. For illustrative purposes, table 1 shows the results of the assessment of indicators characterizing the initial data on the OHL of the EPS in one of the calculated months and their operational reliability.

Table 1. *Illustration of the initial data and estimates of the operational reliability indicators of OHL 110 kV and above EPS*

Parameters	Symbol	Unit of measure	Quantitative estimation	The formula of calculation
Initial data OHL				
Number OHL	$n_{t,l}$	unit	273	
Total length	$L_{t,l}$	км	7918,5	$L_t=\Sigma L_i$
Number of automatic switching-off.	$n_{t,a}$	unit	58	
- the same, but with successful RUE	$n_{t,r}$	unit	50	
Number of switching-off under the emergency request	$n_{t,e,r,}$	unit	94	

Parameters	Symbol	Unit of measure	Quantitative estimation	The formula of calculation
Duration of emergency idle time	$T_{t,e}$	hour.	407	$T_{t,e}=T_e+T_{e.r.}$
Number AT OHL	$n_{t,l}^{AT}$	unit	164	
Total length AT	L_t^{AT}	км	4853	$L_t^{AT}=\Sigma L_i^{AT}$
Total service life AT	ΔT_t^{AT}	years	3718	$\Delta T_t^{AT}=\Sigma\Delta T_i^{AT}$
Parameters of operative reliability				
Specific number of automatic switching-off	$\lambda_{t,a}^*$	Sw-off /years	12,9	$\lambda_{t,a}^*=12\cdot10^2 n_{t,a}/L_{t,l}$
Specific number of automatic switching-off with successful RUE	$\lambda_{t,r}^*$	Sw-off /years	7,6	$\lambda_{t,r}^*=12\cdot10^2 n_{t,r}/L_{t,l}$
Specific number of switching-off under the emergency request	$\omega_{t,e.r.}^*$	Sw-off /years	14,2	$\omega_{t,e.r}^*=12\cdot10^2 n_{t,e.r}/L_{t,l}$
Average duration of idle time under repair under the emergency request	$M^*(\tau_{e.r.})$	hour	4,3	$M^*(\tau_{e.r.})=\sum_{i=1}^{n_{e.r.3}}\tau_{e.r,i}/n_{t,e.r}$
Average duration of idle time in emergency repair	$M^*(\tau_{e.rp.})$	hour	3,3	$M^*(\tau_{e.rp.})=\sum_{i=1}^{n_{e.rp}}\tau_{e.rp,i}/n_{e.rp}$
Relative duration of idle time in emergency repair	$K_{e.r.}^*$	%	0,71	$K_{e.r.}^*=10^4 T_{e,r}/(T_m\cdot L_{e,r})$
Average length OHL	$M_t^*(L)$	км	29	$L_{t,av}=L_{t,l}/n_{t,l}$
Relative number AT OHL	$\delta n_{t,l}^{AT}$	%	60,1	$\delta n_{t,l}^{AT}=10^2 n_{t,l}^{AT}/n_{t,l}$
Relative length AT OHL	$\delta L_{t,l}^{AT}$	%	61,3	$\delta L_{t,l}^{AT}=10^2 L_{t,l}^{AT}/L_{t,l}$
Average service life AT OHL	$M_t^*(\Delta T^{AT})$	years	22,7	$M_t^*(\Delta T^{AT})=\Delta T_t^{AT}/n_{t,l}^{AT}$
Average length AT OHL	$M_t^*(L^{AT})$	км	23,6	$M_t^*(L^{AT})=L_t^{AT}/n_{t,l}^{AT}$

Note: Te<Te.r.; Tm - the duration of the calculated month, AT - symbolic designation of OHL, the service life of which exceeds the standard value

To compare these indicators with the reliability indicators given in reference books and literature on the reliability of EPS facilities, the monthly average estimates of the reliability indicators of OHL, multiplied by the number of months in a year (12) and reduced to a conventional line 100 km long.

As expected, the given monthly average values of operational reliability indicators may differ significantly from the average annual values due to the uneven distribution of the intensity of the impact of disturbing factors (for example, thunderstorm activity) throughout the year. Nevertheless, a significant excess of the reduced average monthly value of the estimate of the operational reliability indicator of the average annual value indicates insufficient protection of the OHL from the main influencing factor in the calculated month.

III. Initial data and indicators of operational reliability of OHL ENE EPS

The possibility of comparing the operational reliability of OHL ENE EPS is one of the most urgent tasks of the EPS and, first of all, because it allows you to optimize the total operating costs of the EPS. The methodology for assessing the operational reliability indicators of OHL ENE EPS is similar to the methodology for assessing the operational reliability indicators of OHL EPS as a whole. The essential difference is that information on passport data and changes in the technical condition of OHL must classify according to the "name of the ENE". Here, similar to the data in

table A, tables AN are compiled, where N - is the conditional serial number of the ENE EPS, which are also practically unchanged, but unlike table A, they do not contain the column "name of the ENE". The automated generation of AN tables is not difficult.

IV. Initial data and indicators of operational reliability OHL ENE ESP.

The ability to compare the operational reliability of the OHL ENE ESP is one of the most pressing issues for EPS and, above all, because it allows you to optimize the total operating costs of the EPS. The methodology for assessing operational reliability indicators OHL ENE ESP is similar to the methodology for assessing operational reliability indicators OHL ESP as a whole. The essential difference is that information on the passport data and changes in the technical condition of the OHL classified according to the "ENE name" attribute. Here, similar to the data in Table A, tables AN compiled, where N is the conditional serial number ENE ESP, which are also practically unchanged, but unlike Table A, they do not contain the column "ENE name".

The automated generation of AN tables is not difficult. Of course, when classifying passport data according to ENE manually, the grouping process is laborious and cumbersome. But it is carried out only once. Formation of BN tables turns out to be much more difficult, since in table B, and of course, in the operational logs, there is no information about the name of the ENE, to which the OHL belongs, the state of which has changed. Searching for the passport data of a specific OHL among hundreds of considered OHL is tedious, and the risk of a wrong decision turns out to be unacceptably high. An automated search can, of course, solve this problem without error. But even in this case, the time spent turns out to be unacceptably large.

Offered:

– transform the adopted sequence of OHL placement (as a rule - by voltage class) into a sequence of OHL names in alphabetical order, indicating the name of the ENE of each OHL (analogue - telephone directory);
– define the OHL group, the first letter of the name of which coincides with the first letter of the name of the recognized OHL;
– among a relatively small number of OHL of this group (maximum - several tens of OHL), it is quite simple to identify the desired ENE, on the balance of which this OHL is located, manually.

Table 2. *Results of calculation of indicators of initial data and estimates of indicators of operational reliability*

Parameters	Unit of measure	ENE N							
		1	2	3	4	5	6	7	8
Initial data									
$n_{t,l}$	unit	49	22	28	7	29	34	69	18
$L_{t,l}$	км	1600	561	850	220	1026	604	1253	1802
$n_{t,a}$	unit	31	0	4	1	14	6	23	18
$n_{t,r}$	unit	25	0	3	1	7	3	9	3
$n_{t,e,r,}$	unit	20	2	7	0	15	13	23	14
$T_{t,e}$	hour	97,4	2,5	21,5	0	97,5	60	83,7	44,4
$n_{t,l}^{AT}$	unit	38	16	20	7	13	18	41	11
L_{t}^{AT}	км	1214	392	633	220	403	470	610	911
ΔT_{t}^{AT}	years	653	256	403	109	291	442	1340	224
Parameters of operative reliability									
$\lambda_{t,a}^{*}$	Sw-off /years	**23,3**	0	5,6	5,5	16,4	8,2	17,6	4,0
$\lambda_{t,r}^{*}$	Sw-off /years	**18,8**	0	4,5	5,5	8,2	6,0	8,6	2,0
$\omega_{t,e.r.}^{*}$	Sw-off /years	15	4,2	9,9	0	17,6	**25,8**	22	9,3

Parameters	Unit of measure	ENE N							
		1	2	3	4	5	6	7	8
$M^{*}(\tau_{e.r.})$	hour	4,9	1,5	3	0	**6,5**	4,6	3,6	3,8
$K^{*}_{e.r.}$	%	0,85	0,061	0,35	0	1,32	**1,38**	0,93	0,34
$M^{*}_{t}(L)$	км	327	25,8	30,4	31,4	**35,4**	17,8	12,2	100,1
$\delta n^{AT}_{t,l}$	%	77,6	72,7	71,4	**100**	44,8	52,9	59,4	61,1
$\delta L^{AT}_{t,l}$	%	45,9	69,1	74,5	**100**	39,3	77,8	48,7	50,6
$M^{*}_{t}(\Delta T^{AT})$	hour	17,2	16,0	20,2	15,6	22,4	24,6	**32,7**	20,4
$M^{*}_{t}(L^{AT})$	км	31,9	24,5	31,7	31,4	31	26,1	14,9	***82,8***

Table 2 shows the results of calculations of operational reliability indicators OHL ENE ESP. This data allows you to:

1. Compare and rank ENE. For example, according to the indicator, the ranking of ENE in order of increasing reliability is: ENE1, ENE7, ENE5, ENE6, ENE3, ENE4, ENE8 and ENE2. If we take into account that for EPS as a whole, the value is 12.9 sw-off / years (see Table 1), then we can conclude that the least reliable OHL are ENE1, ENE7 and ENE5.
2. Draws attention to the fact that the ranking results depend on the operational reliability indicator. For example, the OHL in ENE6 is the least reliable for the indicator and in ENE7 for the indicator. To overcome this ambiguity of the decision, it is necessary either to choose for comparison one of ten indicators of operational reliability or to calculate an integral indicator that takes into account the significance of each of the 10 indicators. The second method is more reliable, but it also requires solving a number of tasks, such as assessing the degree of relationship between indicators of operational reliability, overcoming the difference in their dimensions and scale, and preserving the physical essence [8].
3. Increasing the reliability of comparison and ranking of operational reliability indicators requires taking into account their random nature. As a first approximation, the ENE list classified into three groups. The indicators of the first group OHL ENE accidentally differ from the same indicator for OHL ESP as a whole, the second group is not accidentally less than the indicator for OHL ESP, and the third group is not accidentally higher than the indicator for OHL ESP.
4. An illustration of the solution to these problems requires special consideration.

III. Analysis of OHL ENE, the operational reliability of which is the lowest

Before carrying out this analysis, let us answer one non-standard question: how much can the operational reliability of the OHL ESP increase if the reliability of the OHL ENE1 increased at least to the level of the operational reliability of the OHL ESP in the calculated month? According to table 1, and according to table 2 for ENE1.

It is easy to see that approximate equality achieved by reducing the value by about half. At the same time, the specific number of automatic shutdowns OHL ESP will decrease by 102·15/85=17.6%. Such a dramatic change is certainly tempting.

The purpose of the analysis of OHL ENE1 is to recognize the types of signs for which the specific number of automatic shutdowns of OHL ENE1 most higher than the estimate for OHL ENE1. For illustrative purposes, table 3 shows the results of calculating the specific number of automatic trips when classifying OHL ENE1 by voltage class, service life, support material and OHL length. Analysis of these data shows:

– the dependence of the specific number of automatic shutdowns OHL on the voltage class, known from reference books, remains unchanged according to long-term data - with an increase in the nominal voltage, the specific number of automatic shutdowns OHL decreases;

– the dependence of the specific number of automatic OHL shutdowns on the service life is also clearly confirmed. It is somewhat overestimated in the initial period of operation and significantly increases when the service life is exceeded ΔT=53 years.
– at the initial stage, the classification of OHL according to the material of the supports [metal or mixed (metal, reinforced concrete or wood)] turned out to be inappropriate, since their significance is approximately the same;
– most often automatically shut off OHL, the length of which is in the range (31 ÷ 60) km.

Table 3. *Illustration of the significance of the varieties of features that characterize the operational reliability of OHL ENE1.*

The name		Parameters		
attribute	varieties	n_a, unit	$L_л$, км	λ^*, sw-off /years
ENE1		31	1600	23,3
Voltage class, кV	330	2	268	9,0
	220	5	284	21,1
	110	24	1048	27,5
Service life, years	≤ 17	6	228	31,6
	18-35	-	7,2	-
	36-52	4	654,8	7,5
	≥ 53	21	710	35,5
Material of support	Metal	10	513,3	23,4
	Mixed	21	1086,7	23,2
Length of a line, км	≤ 30	4	330,2	14,5
	30-60	20	613,3	39,2
	60-90	6	542,5	13,3
	≥ 90	1	114	10,5

Table 4. *Recognition of the most significant variety of signs for OHL ENE1 with L = 30-60 km*

The name		Parameters		
attribute	varieties	n_a, unit	$L_л$, км	λ^*, sw-off /years
ENE1, $L_л$=30-60 км		20	613,3	39,2
Voltage class, кV	330	-	-	
	220	2	78,7	30,4
	110	18	534,6	40,4
Service life, years	≤ 17	4	94,9	50,6
	18-35	-	-	-
	36-52	2	202	119
	≥ 53	14	316,4	53,1
Material of support	Metal	5	252,7	19,7
	Mixed	15	360,6	41,5

Table 5. *Results of the third stage of classification OHL ENE1*

The name		Parameters		
attribute	varieties	n_a, unit	$L_л$, км	λ^*, Sw-off /years
ENE1, $L_л$=30-60 км; $\Delta T_{сл}$>53 years		14	316,4	53,1
Voltage class, kV	220	-	-	-
	110	14	316,4	53,1
Material of support	Metal	5	173,3	34,6
	Mixed	9	143,1	75,5

To recognize the OHL features, with a length from 31 to 60 km, Table 4 shows the results of their classification according to the characteristics: stress class, service life and material of supports. The calculation results allow us to conclude:

– the significance of the varieties of the attribute "stress class" has changed little. Only the excess of the specific number of automatic shutdowns OHL 110 kV over OHL 220 kV became clearer;
– the dependence of the specific number of automatic OHL shutdowns on the service life has also remained unchanged;
– but the higher reliability of OHL on metal supports known from the operational experience was confirmed - the specific number of automatic shutdowns is almost two times less;
– the largest specific number of automatic shutdowns is observed at OHL, the service life of which is 1.5 times higher than the standard value. Since the length of these lines is 316.4 km, and the OHL number is 7, let us clarify the significance of these OHL by classifying them according to the remaining two features, stress class and support material. The calculation results are shown in Table 5.

Analysis of this data shows:
– the greatest significance of the varieties of signs established at the second stage of the classification - OHL with $\Delta T>53$ years is entirely related to OHL with U_H=110 kV.
– at the third stage, a significant excess of the operational reliability of OHL on metal supports compared to OHL on mixed supports manifests itself.

Let's summarize the results. Determined that:
– 110 kV OHL on mixed towers with a length of 30 to 60 km, the service life of which exceeds one and a half of the rated service life, are subject to increase in the reliability of operation on ENE1;
– it is easy to see that it is for these types of signs that an intuitive choice of OHL is made, subject to TC certification and overhaul;
– the recommended method allows to set the list of OHL to be restored in ENE automatically according to the statistical data of operation. It refers to risk-based approaches since it significantly reduces the risk of an erroneous decision;
– with all the apparent cumbersomeness and laboriousness, the apparent simplicity of the analysis of the operational reliability of OHL is deceptive, primarily because when comparing and ranking estimates of operational reliability indicators, their random nature was not taken into account, and thus the recommendations were not specified. The possibility of an accidental discrepancy is objective and indicates the inexpediency of classification, and the use of the recommended methods and algorithms requires an unconditional transition to automated systems for analyzing operational reliability

Conclusion

1. The developed methods and algorithms for assessing, comparing and ranking indicators of operational (average monthly) performance (economy, reliability and safety), practical testing of individual stages of their application according to statistical operating data indicate real possibilities for improving the management of production assets;
2. This result is due to a significant increase in the number of objects, the service life of which has exceeded the standard values;
3. For example OHL with voltage 110 kV and above:

– calculation formulas and quantitative estimates of monthly average values of operational reliability indicators characterizing their TC are given. These estimates can be compared with similar estimates calculated for the month preceding the calculated one;
– calculated and compared the monthly average values of OHL operational reliability indicators for ENE EPS. These estimates allowed for the first time to rank the operational TC OHL EPS, to identify enterprises with the least operational reliability;
– since this enterprise may include dozens of OHL, not all of which do not meet the requirements of operational reliability, an illustration of the OHL recognition method that requires immediate (prompt) recovery is given

The use of the developed algorithms in automated systems for assessing, comparing and ranking production assets eliminates the risk of erroneous decisions of an intuitive approach when organizing operation, maintenance and repair.

References

[1] Voropai N.I., Kovalev G.F. and others. The concept of ensuring reliability in the electric power industry. - M., LLC Publishing House "Energia", 2013, 304 p.
[2] STO-34.01-24-003-2017. UNEG production assets management system. 2017, FGC.
[3] STO-34.01-24-002-2018 Organization of maintenance and repair of power facilities of the UNEG, 2018, FGC
[4] STO-34.01-35-001-2020. Guidelines for the technical examination of substation equipment, power lines. "Rosseti", 2020, 20p.
[5] Gromova G.A., Ismailova L.A. Management of production assets of a power grid company in the life cycle concept. Bulletin of the Altai Academy of Economics and Law No. 1, part 1, 2019, pp. 37-44.
[6] Farhadzadeh E.M., Muradaliyev A.Z., Farzaliyev Y.Z., Abdullayeva S.A. Comparison and ranking of steam turbine units of TPP power units by operating efficiency. Heat power engineering, 2018, No. 10, p.41-49. DOI: 10. 1134 / SOO 40363618100028
[7] Farhadzadeh E.M., Muradaliyev A.Z., Rafiyeva T.K., Rustamova A.A. Improving the efficiency of the power units of thermal power plants. "Power stations" No. 8, 2019, pp. 14-17, DOI - 10. 34831 / EP.2019.1054.44206
[8] Farhadzadeh E.M., Muradaliyev A.Z., Rafiyeva T.K., Abdullayeva S.A. Method and algorithm for calculating reliability indicators from multidimensional data. Minsk Power Engineering No. 1, 2017. p.16-29. DOI-: 10 .21122 / 1029-7448-2017-60-1-16-29.

APPLICATIONS OF SIMULATION AND QUEUING THEORY IN SCOOTER INDUSTRY

Mohit Yadav[1], Shruti Gupta[2] and Sandeep Singh[3]

•

[1]Department of Mathematics, University Institute of Sciences
Chandigarh University, Mohali, Chandigarh
[2]Poornima Institute of Engineering and Technology, Jaipur
[3]SOP, Raffles University, Neemrana, Rajasthan
mohit.e15793@cumail.in, shruti.gupta@poornima.org, sandeepnimbhera1986@gmail.com

Abstract

This paper describes the role of queuing theory in developing queuing networks in companies. Queuing networks can be considered as a collection of nodes where each node stands for a service facility. It is a powerful and versatile tool for modeling facilities in manufacturing products. In the realm of service industries like scooter manufacturing, the queuing theory and simulation play a vital role. These concepts help in predicting queue lengths and waiting durations when multiple scooters are manufactured and distributed using first come first serve discipline. Tables are used to explore the availability of furnished scooters in the companies and their comparative study analyzes the waiting scooters and space availability in the companies.

Keywords: Queue network, space availability, service facility, simulation and production analysis.

I. Introduction

Most works on queuing models are restricted to deriving the formulations for transient or stationary states. The past study has not paid much attention to the analysis of queuing systems in the directions such as waiting times and space availability. Frequently, some information about the distribution of scooters in lorry and number of waiting scooters is analyzed with the help of simulation. Generally, simulation is the representation of reality through the use of a model/device which will react in the same manner as reality under a given set of conditions such as testing the performance of a scooter under different conditions.

Baskett et al. [1] analyzed the different networks of queues under different classes of customers using joint equilibrium distribution. Reddy et al. [11] evaluated the queue system with multiple vacations and a server leaves for a vacation of random length when the queue length is less than requirement. Divya and Indhira [6] threw light on the queuing model with working vacation where during the vacation period, the server provides service at a slower pace. Baba [16] evaluated the queue model with multiple working vacations where the server works with different rates. Chakravarthy [4] examined a single server stochastic queue model with Markovian arrivals and impatient customers where customers can become impatient after waiting a random amount of time.

Chinwuko et al. [3] analyzed the effects of queuing theory in a banking system with first come first serve discipline where the organization needs five servers instead of three available servers. Igwe et al. [2] evaluate the performance of queue management in supermarkets subject to average service rate using first come first serve descipline. Roy and Sinha [12] examined the effects of internet banking on customer acceptance of electronic payment systems where credit card and

debit card provide variety of services to the customers. Zhang and Liu [9] analyzed the queue with server breakdown, working vacations and vacation interruption using the supplementary variable method.

Rajadurai [15] analyzed the sensitivity of a retrial queuing system with disasters under working vacations where a system may become defective by disasters at any point of time under availability of a regular server. Saraswat [7] evaluated the effects of a single counter Markovian queuing model with multiple inputs. Bura [8] examined the M/M/∞ queue subject to impatient customers where customers are impatient due to low quality of service. Chakravarthy and Kulshrestha [13] described a queuing model with backup server in the absence of the main server to continued the process.

Agrawal et al. [10] described the steady state probability distribution of a queuing model with working vacation under two types of server breakdown using first come first serve discipline. Narmadha and Rajendran [14] analyzed the queuing network through many developments which made its existence in many fields. Daş et al. [5] described the fluctuation in the participation of companies and two stage stochastic industrial networks under uncertain demand.

The production of two scooter companies is analyzed with simulation to calculate the number of waiting scooters and available space. It is considered that both companies have the same production per day but have different probability values and random number values. Tables are used to explore the availability of furnished scooters in the companies and their comparative study analyzes the waiting scooters and space availability in the companies.

II. Assumptions

To describe the performance of the scooter industry, there are following assumptions

- There are two scooter companies such that company (A) and company (B).
- It is considered that companies (A) and (B) may produce 150 scooters.
- The daily production varies from 146 to 154.
- The probabilities of production per day of both companies are different.
- The average number of scooters waiting in the factory and average number of available spaces in the lorry are analyzed by using simulation.

III. Descriptions of Company (A) and (B)

(I) Analysis of Scooter Company (A)

A scooter company (A) manufactures 150 scooters. The daily production varies from 146 to 154. Now, it is observed that

Table 1: *Production per day of company (A)*

Production Per Day	146	147	148	149	150	151	152	153	154
Probability	0.04	0.09	0.12	0.14	0.11	0.10	0.20	0.12	0.08

Then the furnished (or tested good position) scooters are transported in a lorry accordingly 150 scooters using the following random variables,
X=80, 81, 76, 75, 64, 43, 18, 26, 10, 12, 65, 68, 69, 61, 57.
Then simulate the following:
(i) Average number of scooters waiting in the factory.
(ii) Average number of available space in the lorry.

Solution. From the given data, get

Table 1.1: *Probability Distribution of Company (A)*

Production	Probability	Cumulative Probability	Random Number
146	0.04	0.04	0-3
147	0.09	0.13	4-12
148	0.12	0.25	13-24
149	0.14	0.39	25-38
150	0.11	0.50	39-49
151	0.10	0.60	50-59
152	0.20	0.80	60-79
153	0.12	0.92	80-91
154	0.08	1.00	92-99

Table 1.2: *Number of waiting Scooter and Available Space in lorry of Company (A)*

Sr. No.	Random No.	Production Per Day	No. of Scooters Waiting	No. of available Space in Lorry
1	80	153	3	
2	81	153	3	
3	76	152	2	
4	75	152	2	
5	64	152	2	
6	43	150		
7	18	148		2
8	26	149		1
9	10	147		3
10	12	147		3
11	65	152	2	
12	68	152	2	
13	69	152	2	
14	61	152	2	
15	57	151	1	
			Sum=21	Sum=9

Here, production per day is calculated from (Table-1.2).
So, (i) Average number of scooters waiting in the factory=21/15.
(ii) Average number of available space in the lorry=9/15.

(II) Analysis of Scooter Company (B)

A scooter company (B) manufactures 150 scooters. The daily production varies from 146 to 154. Now, it is observed that

Table 2: *Production per day of company (B)*

Production Per Day	146	147	148	149	150	151	152	153	154
Probability	0.05	0.08	0.10	0.16	0.10	0.08	0.20	0.15	0.08

Then the furnished (or tested good position) scooters are transported in a lorry accordingly 150 scooters using the following random variables,
X=95, 70, 60, 85, 65 48, 25, 15, 10, 18, 75, 78, 87, 65, 54.
Then simulate the following:
(i) Average number of scooters waiting in the factory.
(ii) Average number of available space in the lorry.
Solution. From the given data, get

Table 2.1: *Probability Distribution of Company (B)*

Production	Probability	Cumulative Probability	Random Number
146	0.05	0.05	0-4
147	0.08	0.13	5-12
148	0.10	0.23	13-22
149	0.16	0.39	23-38
150	0.10	0.49	39-48
151	0.08	0.57	49-56
152	0.20	0.77	57-76
153	0.15	0.92	77-91
154	0.08	1.00	92-99

Table 2.2: *Number of waiting Scooter and Available Space in lorry of Company (B)*

Sr. No.	Random No.	Production Per Day	No. of Scooters Waiting	No. of Available Space in Lorry
1	95	154	4	
2	70	152	2	
3	60	152	2	
4	85	153	3	
5	65	152	2	
6	48	150		
7	25	149		1
8	15	148		2
9	10	147		3
10	18	148		2
11	75	152	2	
12	78	153	3	
13	87	153	3	
14	65	152	2	
15	54	151	1	
			Sum=24	Sum=8

Here, production per day is calculated from (Table-1.4).
So, (i) Average number of scooters waiting in the factory=24/15.
(ii) Average number of available space in the lorry=8/15.

IV. Discussion

From tables 1.1 and 1.2 it is clear that the number of scooters waiting (21) is more than the available spaces (9) in lorry. So, it is concluded that the number of production is more than the available space in lorry. From tables 2.1 and 2.2 it is clear that the number of scooters waiting (24) is more than available space (8) in the lorry. So, it is concluded that the number of production is more than the available space in lorry.

V. Conclusion

From tables, it is concluded that the average number of waiting scooters in company (A) is less than the average number of waiting scooters in company (B). Thus, the production of scooters in company (B) is more than the production of company (A).

References

[1] Baskett, F., Chandy, K.M., Muntz, R. and Palacios, M.G. (1975). Open, closed, and mixed networks of queues with different classes of customers. Journal of the Association for Computing Machinery, 22(2), 248-260.

[2] Igwe, A., Onwuere, J. U. J. and Egbo, O. P. (2014). Efficient queue management in supermarkets: A case study of Makurdi town, Nigeria. European Journal of Business and Management, 6(39), 185-192.

[3] Chinwuko, C. E., Ezeliora, C. D., Okoye, P. U. and Obiora, J. O. (2014). Analysis of a queuing system in an organization (A case study of First Bank PLC, Nigeria). American Journal of Engineering Research, 3(2), 63-72.

[4] Chakravarthy, S. R.(2009). A disaster queue with Markovian arrivals and impatient customers. Applied Mathematics and Computation, 214, 48–59.

[5] Daş, G. S., Yeşilkaya, M. and Birgören, B. (2024). A two-stage stochastic model for an industrial symbiosis network under uncertain demand. Applied Mathematical Modelling, 125, 444-462.

[6] Divya, K. and Indhira, K. (2024). A literature survey on queuing model with working vacation. Reliability Theory and Applications, 1(77), 40-49.

[7] Saraswat, G. K. (2018). Single counter Markovian queuing model with multiple inputs. International Journal of Mathematics Trends and Technology, 60(4), 205-219.

[8] Bura, G. S. (2024). M/M/∞ queue with impatient customers. Reliability Theory and Applications, 1(77), 228-237.

[9] Zhang, M. and Liu, Q. (2015). An M/G/1 G-queue with server breakdown, working vacations and vacation interruption. Operation Research, 52(2), 256-270.

[10] Agrawal, P. K., Jain, A. and Jain, M. (2021). M/M/1 Queuing model with working vacation and two type of server breakdown. Journal of Physics, 1849(1), 012021.

[11] Reddy, G. K., Nadarajan, R. and Arumuganathan, R. (1998). Analysis of a bulk queue with N-policy multiple vacations and setup times. *Computers and Operations Research*, *25*(11), 957-967.

[12] Roy, S. and Sinha, I. (2014). Determinants of customer's acceptance of electronic payment system in Indian Banking Sector. International Journal of Scientific & Engineering Research, 5(1), 177-187.

[13] Chakravarthy, S. R. and Kulshrestha, R. (2020). A queuing model with server breakdowns, repairs, vacations, and backup server. Operations Research Perspectives, 7, 100131.

[14] Narmadha, V. and Rajendran, P. (2024). A literature review on development of queuing networks. Reliability Theory and Applications, 1(77), 696-702.

[15] Rajadurai, P. (2018). Sensitivity analysis of an M/G/1 retrial queuing system with disaster under working vacations and working breakdowns. *RAIRO-Operations Research*, *52*(1), 35-54.

[16] Baba, Y. (2005). Analysis of a GI/M/1 queue with multiple working vacations. Operations Research Letters, 33(2), 201-209.