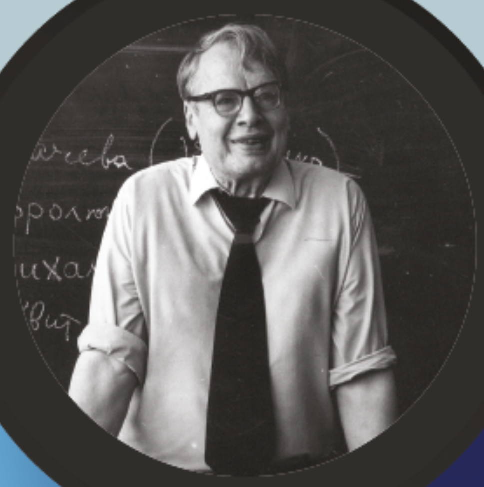
ISSN 1932-2321

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

RTA

RELIABILITY: THEORY & APPLICATIONS

INTERNATIONAL GROUP ON RELIABILITY



RELIABILITY
RISK ANALYSIS
MAINTENANCE
SAFETY

GNEDENKO FORUM PUBLICATIONS

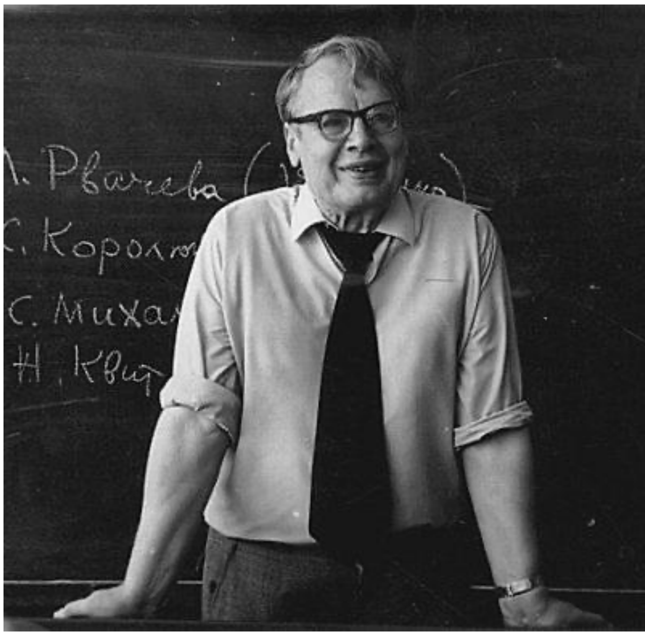
(72) VOL.18
#1
MARCH
2023

SAN DIEGO

ISSN 1932-2321

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

Journal of International Group on Reliability



RELIABILITY:

THEORY & APPLICATIONS

Vol.18 No.1 (72),
March 2023

San Diego
2023

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

Ranked set sampling is an approach to data collection originally combines simple random sampling with the field investigator's professional knowledge and judgment to pick places to collect samples. Alternatively, field screening measurements can replace professional judgment when appropriate and analysis that continues to stimulate substantial methodological research. The use of ranked set sampling increases the chance that the collected samples will yield representative measurements. This results in better estimates of the mean as well as improved performance of many statistical procedures. Moreover, ranked set sampling can be more cost-efficient than simple random sampling because fewer samples need to be collected and measured. The use of professional judgment in the process of selecting sampling locations is a powerful incentive to use ranked set sampling. Optimum stratification is the method of choosing the best boundaries that make strata internally homogeneous, given some sample allocation. In order to make the strata internally homogenous, the strata should be constructed in such a way that the strata variances for the characteristic under study be as small as possible. This could be achieved effectively by having the distribution of the study variable known and create strata by cutting the range of the distribution at suitable points. If the frequency distribution of the study variable is unknown, it may be approximated from the past experience or some prior knowledge (auxiliary information) obtained at a recent study. The present investigation deals with paper the problem of optimum stratification on an auxiliary variable for proportional allocation under ranked set sampling (RSS), when the form of the regression of the estimation variable on the stratification variable given the variance function is known. A cum rule of finding approximately optimum strata boundaries has been developed. Further, empirical study has been made and presented along with relative efficiency which showed remarkable gain in efficiency as compared to unstratified RSS.

High-altitude platform (HAP) are stations on an object at an altitude of around 15-50 km at a specified nominal fixed point relative to Earth. Tethered high-altitude platform (tHAP) are unmanned aerial vehicle that are connected to the ground via a tether with a lift height of 100 − 150 meters, and a multi-copter as high-altitude mode. The reliability of the tHAP can be assessed with a focus on the tether that connects it to the ground. This article proposes a Markov model which obtain the reliability of the tHAP. The tether is considered to be made up of multiple wires in such a way that the tether still operates for a given number of functioning wires. The failure rates of the wires are dependent on the number of failed wires. Through the reliability analysis of the proposed Markov model, the key performance measures such as reliability of the system, mean time between failures and the probability of the system being reliable are computed. The optimal number of wires is also obtained via the numerical computation of the performance measures.

The objective of this paper is to describe a particular case of the k-out-of-n: G system for k=2 and n=3 with different repair policies and to discuss the application of the proposed in a toxic waste incinerator. The system has all the three units identical in nature. The system model is developed using semi-Markov process and regenerative point technique. The preventive maintenance and repair activities of the units are carried out immediately by a single service facility whenever desires. The service facility is subjected to failure during repair of the units while it does preventive maintenance of the units without any problem. The failed service facility undergoes for treatment to restore its efficiency to perform the remaining jobs with full capacity. The provision of priority to preventive maintenance of the units has been made over the repair to avoid the earlier failure of the system. The measures that can affect and enhance the performance of the system have been discussed for arbitrary values of the rates which follow some arbitrary distributions including the negative exponential. The system is analysed in steady state and the graphs have been drawn to see the effect of different transition rates such as failure rate, preventive maintenance rate, treatment rate, and repair rate of the units on reliability measures and the profit. The study reveals that there is a decline in these measures with the increase of the rate by which unit undergoes for preventive maintenance, failure rates of the units and service facility. However, the values of reliability measures MTSF, availability and profit function keep on increasing with the increase of treatment rate, repair rate of the unit and preventive maintenance completion rate. The profit increases if the rate with which a unit completes its preventive maintenance. Hence, implementing the preventive maintenance repair policy for a 2-out-of-3 system is beneficial as it increases the availability and hence the profit of the system.

The study focuses on the E-Bayesian estimation of a Type-II censored sample from the Chen distribution. Three distinct prior distributions for the hyper-parameters and three different loss functions are considered here for deriving the E-Bayes estimators of the scale parameter and hazard rate of above said distribution under Type-II censoring. Also derived analytical expressions for the E-MSE of the proposed estimators. Additionally, several features of the E-Bayesian estimators and E-MSEs are derived. This paper compares E-Bayesian estimation with traditional estimation methods like MLE and Bayesian. The applicability of the proposed estimators is demonstrated using a real data application. Furthermore, the credible intervals of the scale parameter estimators are also provided. The numerical analysis demonstrates that the proposed method is simpler and more feasible than traditional techniques.

It has been noted in the literature on probability theory that the classical probability distributions do not adequately fit real-world data and do not exhibit non-monotonic hazard rate behavior. To overcome this limitation, researchers are focusing on the improvement of these distributions. In this manuscript, we have introduced a new probability model called Ratio Transformation Lomax Distribution (RTLD) as a new generalization of Lomax distribution. A thorough mathematical analysis of the new distribution is provided in closed form such as density function, distribution function, the r-th moment, survival function, hazard function, moment generating function, generalized entropy and also the order statistics. The new model's parameters are calculated using the method of maximum likelihood estimation. The proposed distribution's performance and adaptability is backed by three sets of real lifetime data as well as simulated data.

Based on the power Cauchy distribution, we have purposed a new distribution called inverse power Cauchy distribution that offers greater modeling flexibility for lifetime data. Real-world data can be efficiently analyzed using the suggested model because it is analytically sound. Its density function can take on a number of different shapes, including reversed-J, symmetrical, and right-skewed. Depending on the different values of the parameters, it can adapt to different hazard forms, such as an upside-down bathtub, a monotonically increasing or decreasing curve, and others. Its moments, quantile, reliability, hazard, order statistics with density function, moment generating function, and entropy are all given with various explicit forms. The observed information matrix is created when the new model's parameters are calculated through maximum likelihood technique. A simulation study is conducted to investigate the behaviour of maximum likelihood estimators. The proposed model gets a superior fit compared to certain well-known distributions, according to the test of goodness-of-fit we conducted. The significance of the purposed distribution is demonstrated empirically using two real-world data sets.

This paper analyses the performance and behavior of water circulation system (WCS) of a thermal power plant in fuzzy environment. For this purpose, fuzzy λ–τ technique coupled with petrinet modelling has been used. To address the vagueness in data, trapezoidal fuzzy numbers have been employed in fuzzy λ–τ technique. Various reliability indicators of WCS viz. failure rate, repair time, expected number of failures, mean time between failures, reliability and availability have been computed at ±15%, ±25% and ±40% spreads using fuzzy λ–τ technique. Further, fuzzy values of reliability indicators have been defuzzified employing COA method and the failure dynamics of WCS have been studied on account of decreasing / increasing trends of reliability indicators. The outcomes of this study are of great importance for plant personnel / management to enhance the availability of WCS.

Aim. The purpose of this paper is to find the reliability measures and profit of a two-unit gas turbine power generating system incorporated with one gas turbine and one steam turbine. Effects of different humidity condition (humidity ≤/> 50%) are taken into consideration by fixing the range of temperature (5℃-25℃) for developing the model. At initial stage, both units (gas turbine and steam turbine) are in operative mode. If steam turbine fails, gas turbine remains in operative mode but if gas turbine fails, system goes to down state and when both unit fails, system fails. In this system we assume that failure time distribution is exponentially distributed while repair time distribution is arbitrary. Methods. In this paper we use the Laplace transform for mathematical analysis, and semimarkov process and regenerative point technique to investigate reliability measures and profit of the system. Findings. The system is analysed in steady state and different reliability measures such as mean time to system failure, availability for different cycles and for different humidity conditions, busy period, down time of the system etc. are calculated and the graphs have been drawn to see the effect of different transition rates such as failure rate and repair rate of the units for different humidity conditions on reliability measures and the profit for particular case is evaluated using the information/data collected from gas turbine power generating system located at Bawana, Delhi, India. Conclusion. Our finding shows that mean time to system failure and availability when both turbines are working decreases with increase in any one of failure rate while availability when only gas turbine is working increases with increase in steam turbine failure rate and profit for plant decreases with increase in failure rates. From this we concluded that availability for the fixed range of temperature (5℃-25℃) is higher when humidity is >50% as compared to when humidity is ≤50%. Thus, a comprehensive study of gas turbine system may be helpful to those who are involved in power generating industry.

Khaim Borisovich Kordonsky. Educator and Researcher (1917-1999)

Andronov, A.[1], Shestakov, V.[2]
•
[1]Transport and Telecommunication Institute, Riga, Latvija,
aleksander.andronov1@gmail.com
[2]Riga Technical University, Institute of Aeronautics, Riga, Latvija,
shestakov@inbox.lv

Abstract

The article is about Khaim Borisovich Kordonsky (Tula, Russia, 1919 - Boston, USA, 1999) - Soviet mathematician, Doctor of Technical Sciences, Professor at the Riga Red Banner Institute of Civil Aviation Engineers; specialist in the field of probability and reliability theory, leader of the development of the first computer system for making aircraft schedules, Honoured Scientist of the Latvian SSR, Laureate of the State Prize of the Latvian SSR. Kordonsky's entire scientific and teaching career is connected with RCIIGA. He is one of the founders of reliability theory and the author of the first monograph on the application of probability theory to solving real-world problems. The book "Applications of Probability Theory in Engineering" was published in 1963 and was reviewed by Academician Yuri Vladimirovich Linnik, the greatest expert in probability theory and mathematical statistics in the USSR. Khaim Borisovich also worked on scientific problems in statistics, medicine and technical diagnostics. He conducted joint research with the Department of Mathematical Statistics of the Faculty of Informatics and Cybernetics of Moscow State University, headed by Academician Yuri Vasilievich Prokhorov, and collaborated with Academician A. N. Kolmogorov.

Keywords: memories, probability theory, reliability theory, mathematics.

1. Life's journey

Khaim Borisovich Kordonsky was born on March 28, 1919, in the city of Tula (USSR). In 1941, he graduated from the Faculty of Mathematics and Mechanics of Leningrad State University as a mechanic. On June 25 the same year he joined the militia. After completing accelerated courses at the Leningrad Air Force Academy named after Zhukovsky, he went to fight in the ranks of the Air Force. He went the way from a militia soldier to lieutenant colonel. He was Deputy Chief Engineer of the largest 218th Aircraft Repair Plant. He received the following awards: "For Defense of Soviet Polar Region", "For the victory over Germany in the Great Patriotic War of 1941-1945", "For services in war" and "Order of the Red Star".

From 1947 to 1950, he studied at the Leningrad Air Force Academy in the Mathematics Department under academician Yury Linnik, one of the world leading experts in mathematical statistics. Friendship bound Yuri Vladimirovich and Khaim Borisovich all their lives. Employees of the academician told me. "Usually serious and concentrated, Yuri Vladimirovich blossomed when he was expecting Khaim Borisovich's arrival from Riga."

After finishing his postgraduate studies and defending his doctoral thesis in 1950 he was sent to Riga, to Riga Higher Military Aviation School. Here Kordonskiy spent all his working life, changing only the name of the university, where for more than 30 years he was the head of the department "Aircraft repair and production technologies". The university was transformed several times and in 1992 it became Aviation University of Riga.

The Kordonsky Scientific School of talented and active young people was established in the department. His students are now working in various countries - from Canada to Australia, most of them in Latvia. They are the corresponding members of the Latvian Academy of Sciences N. Salinieks and J. Rudzitis, RTU professors Yu. Paramonov, A. Andronov, Yu. Martynov, etc. The total number of doctors and candidates of science trained under Kordonsky's supervision exceeded 50.

Khaim Borisovich treated his students like a father. The following comes to mind. In the sixties the buildings of the institute were situated on both sides of a beautiful park. Khaim Borisovich's apartment was in the same street, hundreds of metres from the institute. When he left the department in the evening, he liked to walk around the green square. He was not alone: one of his students was waiting for him, and they walked and talked about their dissertations. Often, hiding behind trees, another student would be waiting for his turn.

In the last years of his stay in Latvia, K. Kordonsky, together with Latvian doctors, led by academician J. Anshelevich, was engaged in research on the application of probabilistic-statistical methods in medicine, especially in cardiac diagnostics.

The merits of H.B. Kordonsky were highly appreciated: he has the title of Honoured Scientist and Technician of the Latvian SSR (1969) and is a laureate of the State Prize of the Latvian SSR (1985).

In 1993, H.B. Kordonsky emigrated to the United States. This was a difficult decision for him, and the desire to be reunited with his daughters and grandchildren who already lived there permanently prevailed. During his time in the United States, he was twice invited to be a visiting professor at Ber-Shev University, where he worked with his beloved student, Ilya Boruchovich Herzbach. Chaim Borisovich died in Boston in 1999.

We are going to talk about two areas of Khaim Borisovich Kordonsky's scientific activity. The first is his contribution to the development and implementation of probabilistic-statistical methods. The second is his contribution to the application of computer technology in civil aviation.

The Kordonsky Scientific School was created in the department, consisting of talented and active young people. His students are now working in various countries - from Canada to Australia, most of them in Latvia. They are the corresponding members of the Latvian Academy of Sciences N. Salinieks and J. Rudzitis, RTU professors Yu. Paramonov, A. Andronov, Yu. Martynov, etc. The total number of doctors and candidates of science trained under Kordonsky's supervision exceeds 50.

Khaim Borisovich treated his students like a father. The following comes to mind. In the sixties, the buildings of the institute were situated on either side of a beautiful park. Khaim Borisovich's apartment was on the same street, hundreds of metres from the institute. When he left the department in the evening, he liked to walk around the green square. He was not alone: one of his students was waiting for him, and they walked and talked about their dissertations. Often, hiding behind trees, another student would be waiting for his turn.

In the last years of his stay in Latvia, K. Kordonsky, together with Latvian doctors, led by academician J. Anshelevich, was engaged in research on the application of probabilistic-statistical methods in medicine, especially in cardiac diagnostics.

The merits of H.B. Kordonsky were highly appreciated: he has the title of Honoured Scientist and Technician of the Latvian SSR (1969) and is a laureate of the State Prize of the Latvian SSR (1985).

In 1993, H.B. Kordonsky emigrated to the United States. This was a difficult decision for him, and the desire to be reunited with his daughters and grandchildren who already lived there permanently prevailed. During his time in the United States, he was twice invited to be a visiting professor at Ber-Shev University, where he worked with his beloved student, Ilya Boruchovich Herzbach. Chaim Borisovich died in Boston in 1999.

We are going to talk about two areas of Khaim Borisovich Kordonsky's scientific activity. The first is his contribution to the development and implementation of probabilistic-statistical methods. The second is his contribution to the application of computer technology in civil aviation.

2. Probabilistic-Statistical Legacy

The first studies were devoted to the problems of statistical (sampling) control of product quality. At that time it was the main direction of application of mathematical statistics, and the leading place belonged to Leningrad mathematicians, because the fundamental article by A.N. Kolmogorov "Statistical acceptance control" with admissible number of defective products equal to zero (1951) was published in Leningrad, in the collection of the Leningrad House of Scientific and Technical Propaganda.

Related works:

1953 Kordonsky H.B. Statistical acceptance control on flow and conveyor lines. Vestnik Mashinostroenia, 7.
1955 Kordonsky H.B. Application of Markov chain theory to batch control. Vestnik of Leningrad University, 11.
1956 Kordonsky H.B. The simplest form of product control. Standardization, 5.
1958 Kutai A.K. and Kordonsky H.B. Analysis of accuracy and quality control in mechanical engineering, chapters 3, 4. Mashgiz, M-L.
1959 Kordonsky H.B. Probable product quality. Standardization, 10.
1961 Kordonsky H.B. The distribution of the number of defective units in lots of products. Probability theory and its applications, 3.

And today, references to these first works by Khaim Borisovich Kordonsky are frequent. Thus, in the proceedings of "The International Symposium on Stochastic Models in Reliability, Safety and Logistics", held from 15 to 17 February this year in Israel, in the article Sh. Formanov and T.A. Formanova "Optimal plans of statistical acceptance control taking into account Sheppard corrections" we read: "We consider the problem of optimal plans of statistical acceptance control (SAC) studied by Kh. Kordonsky, S.Kh. Sirajdinov, Van der Varden, K. Stang".

In 1950-1955, under the leadership of Khaim Borisovich Kordonsky, the introduction of statistical methods of quality control was carried out at the VEF Plant, he provided scientific advice to the reliability services of the Avtoelektropribor Plant, the Wagon Plant and the Diesel Plant.

In 1963 a book by Khaim Borisovich Kordonsky Applications of Probability Theory in Engineering was published by the Publishing House of Physical and Mathematical Literature (Moscow, Leningrad). It was the first book in Russian on probability theory and mathematical statistics aimed at engineers, enabling them to master probability and statistical methods and apply them to their work.

A characteristic of Khaim Borisovich Kordonsky was a rare combination of knowledge of the mechanism of the physical processes under consideration and mathematics capable of describing them adequately. As a result, his books and articles (as well as his lectures) were characterised by a striking simplicity and lucidity in the presentation of material that was quite complicated in content. For example, in his later book Models of Failure, he writes

The log-normal distribution describes the life behaviour of objects that have the property of "hardening" over time. "Hardening is manifested by a gradual decrease in the rate of wear. Therefore, before using the log-normal distribution to describe experimental data, it is necessary to determine, on the basis of the physical nature of the wear process and, if possible, by analysing the behaviour of wear realisations, whether the objects studied have the property of "hardening".

The Riga Aviation Institute, where Khaim Borisovich Kordonsky worked, trained engineers in aircraft operation, not aircraft manufacture. Heim Borisovich Kordonsky immediately saw a huge field for the application of probabilistic-statistical methods in conditions of the highest requirements for the reliability of aviation technology and flight safety. These requirements were ensured by a system of measures such as the use of lead aircraft (which had more flight hours than the entire fleet), inspections (to check the technical condition of the aircraft's power elements), maintenance and repair of aircraft. Many scientific problems arise: how many lead planes are needed and what should be the lead time (so that it is possible, with a certain degree of confidence, to assess the entire fleet in the conditions of its further operation), when to carry out inspections and how to predict the rate of development of the cracks detected, what are the terms and volumes of maintenance and repair work, etc.?

Khaim Borisovich Kordonsky soon became the undisputed authority in civil aviation on these issues. He carried out important research and made recommendations at the request of the Ministry of Civil Aviation, GosNII GA, operating and repair companies. At the same time, a large scientific team of talented and active young people was created in the department. His favourite student Ilya Gertsbach deserves special mention. In 1964 he brilliantly defended his doctoral thesis. His scientific supervisor was Khaim Borisovich, and his official opponent was Academician Boris Vladimirovich Gnedenko. The defence took place in the auditorium of the Academy of Sciences of the Latvian SSR and was a remarkable event in the scientific life of the Latvian capital.

Heim Borisovich Kordonski solved practical problems on a strictly mathematical level. As a result, the models, methods and algorithms he developed were universal and practically applicable to many technical systems (not only aviation and transport).

This fact is reflected in subsequent high-level scientific publications:

1964 Kordonsky H.B. Calculations and tests of fatigue durability. Proceedings of the 4th All-Union Mathematical Congress, Nauka, Moscow.
1966 Gertsbakh I.B., Kordonsky H.B. Failure Models. Soviet Radio, Moscow.
1969 Gertsbakh I. and Kordonsky Kh. Models of Failures. Springer, Berlin - Heidelberg - New York}
1967 Kordonsky Kh.B. Probabilistic analysis of stitching processes. Nauka, Moscow.
1969 Gertsbakh I.B. Models of Prevention. Soviet Radio, Moscow.
2000 Gertsbakh I.B. Reliability Theory with Applications to Preventive Maintenance, Springer, Berlin - Heidelberg - New York.

Let us dwell on two statistical problems posed by the needs of practice.

The first concerned methods of statistical processing of data on aircraft failures. In the literature on mathematical statistics and in data processing manuals, we have always considered the so-called case of complete sampling, where the estimation of the distribution of a random variable and its parameters is carried out on the basis of the exact values of the random variable recorded. In practice, this would mean that each product or performance element of a design would be operated to failure. In reality (to ensure reliability and safety), the operating time is limited by specified resources, premature shutdown (even of serviceable objects), etc. Khaim Borisovich Kordonsky introduced this class of problems into mathematical statistics and proposed the method of partitioning partitions for their solution, using the method of maximum likelihood.

Related publications:

1966 Gertsbakh I.B., Kordonsky H.B. Failure Models. Soviet Radio, Moscow.
1970 Artamanovsky A.V., Kordonsky H.B. Maximum likelihood estimation in simple data grouping. Probability Theory and its Applications, 1.
1985 Kordonski H.B., Rastrigin V.L. Random censoring on trajectories in phase space. Izv. of the Academy of Sciences of the USSR, Technical Cybernetics, 6.
1986 Kordonski Kh.B., Rastrigin V.L., Shulkin Z.A. Estimation of reliability indices under the action of several causes. Izv. of the Academy of Sciences of the USSR, Technical Cybernetics, 6.

Later, this problem developed into a whole branch of mathematical statistics called censored sampling. It is now one of the most important applied branches of mathematical statistics, with hundreds of famous mathematicians working on it, and there is a great deal of monograph and journal literature.

The second problem was the theory of unbiased estimation. At that time, when probability methods were just beginning to be widely used in practice, the models used were very simple. They usually involved one or more uniformly distributed random variables.

Therefore, the main effort of mathematicians was to develop methods to obtain the best estimates of the parameters of the main distributions of the random variables. The best estimates were understood as unbiased estimates with minimum variance. Much progress has been made and such estimates (when they exist) have been found (also for the cases of censored samples).

As experience grew, more complex situations were considered, where the subject of statistical analysis was large systems whose models included many random variables. The task was to estimate the performance of the system as a whole on the basis of statistical data relating to individual elements. This was done "the old-fashioned way": for each random variable (system element) we found the best estimate and substituted it for the corresponding unknown parameter of the probability model of the system. (In modern literature, this method is called the plug-in method).

This ignores the fact that (in the case of small samples) good properties of individual estimates are lost. This is natural, because in selecting these estimates we were not trying to optimise the system as a whole, but its individual elements. In order to get the best estimates for the system as a whole, it is necessary to keep it in mind at once, and not to consider separate problems of estimating individual parameters in isolation.

In mathematical statistics, fundamental results had already been obtained by S.R. Rao, A.N. Kolmogorov and D. Blackwell, but their practical use for the above purpose was out of the question. Under the leadership of Khaim Borisovich Kordonsky, the theory of unbiased estimation was applied for the first time in our country (and perhaps in the world) to the estimation of the performance of complex systems.

I remember a conversation we had after Khaim Borisovich Kordonsky returned from Moscow, a little excited after a business trip and full of energy: Sasha, how are you going to estimate, for example, the average waiting time in a unilinear mass service system with Poisson input flow and exponential service time?

Here are the most important publications on unbiased estimation:

1972 Andronov A.M., Kordonsky H.B., Rosenblit P.Y. Application of unbiased estimation theory to mass service problems. Izv. AS USSR Technical Cybernetics.

1976 Kordonsky H.B., Rosenblit P.Y. On unbiased estimation of polynomials from moments. Probability theory and its applications, 1.

1979 Rosenblit P.J. Statistical estimation of reliability and efficiency characteristics of complex systems. Zinatne, Riga.

1982 Larin M.M. Unbiased estimates of variance and some other characteristics of the inverse normal distribution. Izv. AS USSR, Technical Cybernetics.

1989 Voinov V.G., Nikulin M.S. Unbiased estimates and their applications. Nauka, Moscow.

In downloading the history of Khaim Borisovich Kordonsky as a scientist and educator in probability theory and especially in mathematical statistics, here is a quotation from S. Radhakrishn Rao from the preface to his book Linear statistical methods and their applications, Nauka, Moscow, 1968 (Linear statistical inference and its applications, John Wiley & Sons, New York, 1966): I wish to express my gratitude to Ronald A. Fisher and Professor Mahalanobis, under whose influence I have come to appreciate mathematical statistics as the new method of our century.

In the same words, we express our gratitude to Khaim Borisovich Kordonsky.

3. Computer scheduling of civil aviation aircraft

In 1963, the GA Scientific and Computer Centre was established on the basis of the Riga Flight Automation Laboratory, which was part of the RCII GA. After two years of confusion, Associate Professor Lev Fedorovich Krasnikov was appointed its director in 1965. He invited Khaim Borisovich Kordonsky to form the main scientific directions of the NVC. In 1971, the SAA NVC was transformed into the Central Research Institute of Automated Control Systems, headed by Gennady Tikhonovich Kalchenko.

For almost 35 years, Professor K.B. Kordonsky was the scientific director of the work of NVTS GA and the Central Research Institute of ACS GA on computer planning of civil aviation aircraft - the largest airline in the world at that time.

Some of the works of this period of great importance:

1969 Kordonsky H.B., Gerzbach I.B., V. Venyavtsev, Maxim M., Linis V. A heuristic method for aircraft scheduling. In Collection of Automation in Mechanical Engineering, USSR Academy of Sciences, Moscow.

1969 Kordonsky H.B., Linis V. et al. Algorithms for scheduling passenger airplanes. In Proceedings of the 4th Congress on Automatic Control, Warsaw.

1970 Kordonski H.B., Venyavtsev V., et al. Central aircraft scheduling as part of air traffic control. In Proceedings of the 1st International Symposium on Traffic Control, Versailles.

1999 Kordonsky Kh.B., Gertsbakh I.B.. Using Entropy Criterion for Job-Shop Scheduling Algorithm.

Of course, the main result was not the publications, but the central flight plan that civil aviation flew. It was the world's first computerised timetable. Given the state of computer technology at the time, one wonders how this was possible: bicast computers that ran continuously for more than 24 hours to compile a piece of the timetable; punched tapes with the compiled timetable; linotype machines that printed a hard copy of the timetable. It was made possible by the talents of Khaim Borisovich Kordonsky and the dedication of the young team who believed in him: Valery Venyavtsev, Ilya Gertsbakh, Misha Maxim, Yuri Paramonov and many others.

Ilya Gertsbakh wrote in the introductory article of the Proceedings of the 1999 Aviation Reliability-99 conference dedicated to the 80th birthday of Khaim Kordonsky:

Together with Yu. Paramonov, V. Venyavcev, M. Maksim and V. Linis, I worked on this project for seven years, which were probably the most productive and interesting years of my life. Now I realise that we were all extremely lucky to work under the leadership of such a brilliant scientist and outstanding personality as Khaim Borisovich.

The planning project was a very difficult and complex task. Nobody in the Ministry's top management had the slightest idea of how to approach it or what was meant by the term "computerised scheduling". In addition, the computers at that time were extremely primitive. The "Ural-4", which took up an entire floor of an old church, had less power than a modern pocket calculator. Even with modern computer power, it takes a man of exceptional intellectual courage to take on the challenge of leading such a project.

From Prof. Kordonsky we learnt important things for our whole life. The first lesson was: before you start doing the computerised schedule, you should be able to do it manually. This was wise advice, because it was only after a year of intensive contact with practitioners that we began to understand what scheduling was about, what was essential and what was secondary.

Prof Kordonsky was never a "boss" who gave orders and instructions. He created a stimulating atmosphere of intense exchange and discussion, sometimes heated, but always efficient. He was open to every suggestion and critical comment. Despite this enormous scientific authority, no one was afraid to ask questions or to insist on their opinion. I am convinced that the truly democratic nature of our group was a key factor in the success of the project.

4. American Period

The most recent works of Chaim Borisovich Kordonsky are devoted to the theory of calculating the time of degradation of systems whose operating time is measured in different scales (calendar time, number of cycles, hours of operation in different modes, etc.).

Most of these works have been published in leading foreign scientific journals:

1993 Kh.Kordonsky and I.Gertsbakh. Choice of the Best Time Scale for Reliability Analysis. Europ. J.Operat. Res., 65.
1994 Kh.Kordonsky and I.Gertsbakh. Best Time Scale for Age Replacement. Inter. J. of Reliab., Quality and Safety Engineering, 1.
1995 Kh.Kordonsky and I.Gertsbakh. System State Monitoring and Lifetime Scales. I, II. Reliab. Engineering and System Safety, 47, 49.
1997 Kh.Kordonsky and I.Gertsbakh. Multiple Time Scales and the Lifetime Coefficient of Variation: Engineering Applications. Lifetime Data Analysis, 3.
1997 Kh.Kordonsky and I.Gertsbakh. Fatigue Crack Monitoring on Parallel Time Scales. Proceedings of ESREL 97, Lisbon, June 17-20, 1997, 2.
1997 Kh.Kordonsky and I.Gertsbakh. Optimal Preventive Maintenance in Heterogeneous Environments. Europ. J.Operat. Res., 98.
1998 Kh.Kordonsky and I.Gertsbakh. Parallel Time Scales and Two-Dimensional Manufacturer and Individual Customer Warranties. IIE Transactions, 30.

RT&A, No 1 (72)
Volume 18, March 2023

The bright image of Chaim Borisovich Kordonsky will always remain in the memory of his grateful students and followers.

References

1. Report read by Prof. A.M. Andronov at International Conference Reliability And Statistics In Transportation And Communication (RELSTAT'05), Riga, Latvia, November, 2006. http://www.gnedenko.net/Memorial/Kordonsky/memories.htm.
2. Gnedenko Forum website, http://www.gnedenko-net.
3. Vinogradov R.I., Shestakov V.Z. "History of Aviation Science Development in Latvia," Riga, 1989.
4. Shestakov V.Z. "The contribution of Latvian scientists to aviation science and technology in the 20th century," LAP Lambert Academic Publishing, ISBN, EAN 978-3-659-64442-9, 2014.
5. Shestakov V., Zyamzin A., Kuleshov N. Triumph and Tragedy of the First in the History of Russian Aviation Technical School, "HOLDA", Riga, ISBN 978-9934-8072-5-1, 2019.

A MODIFIED INVERSE WEIBULL DISTRIBUTION USING KM TRANSFORMATION

Gauthami P., Mariyamma K. D. and Kavya P.

•

University of Calicut
gauthaminair34@gmail.com
srjosephinedavid@gmail.com
kavyapnair90@gmail.com

Abstract

In the subject of reliability engineering and statistics, a new reliability model is proposed, where survival analysis or life time data analysis is of major importance in the current scenario. The goal of this study is to introduce a new model that has applications to real data sets from the field of survival analysis. Deriving out the new model there are various methods to propose a new model, one of them is by using the method of transforming a variable to the variable of interest and there are numerous transformation methods which are in use right now. The newly proposed model is achieved by using the transformation method known as KM Transformation where it does not require any additional parameters to the baseline distribution which absolutely is an advantage. The model considered in this paper as baseline model is Inverse Weibull distribution with two parameters, one is a scale and other is a shape parameter. Inverse Weibull distribution is a continuous probability distribution which presently has great applications in real life phenomenon as well as so many modifications and advanced studies are introduced in this distribution from various fields. A proper study on the newly proposed model is done by deriving out its various functions and statistical properties such as Probability density function, Cumulative distribution function, Hazard rate function, Moments, Moment generating function, Characteristic function, Quantile function, Order statistics, etc. along with its Probability density function plot and Hazard rate function plot which have both upside-down and decreasing curves. Focusing on the inference procedures, the estimation of the parameters involved in this model is done by using the method of Maximum likelihood estimation. A simulation study for valuing the parameter consistency using two parameter combinations is carried out as well as a data analysis on an actual data set is also conducted. A comparison of the newly proposed model with other popular well-known models such as Inverse Weibull distribution (IW), KME distribution and KMW distribution using R programming language yielded that the new model is a better fit for the real data considered in this paper. The results and conclusions achieved throughout the paper are also mentioned at the last.

Keywords: KM Transformation, Inverse Weibull distribution, upside-down curve, decreasing curve

1. Introduction

In the current scenario of model building and real life data analysis there are so many lifetime distributions that are in use for which [14] and [15] explains the basic ideas and concepts. Among these well-known distributions, Weibull distribution and its various modified Weibull distributions plays a great role in fitting real data sets. Inverse Weibull (IW) distribution has great applications in reliability engineering, where [3] and [4] studies the various inference procedures in the distribution of consideration. A theoretical analysis of IW distribution is done by [12] and order statistics with inference is done by [16]. [5] studies the generalized modified weibull distribution along with applications to two real data sets and [2] applied the IW distribution to model the wind speed data. Extended studies are done in this distribution where [13] studies the Bayesian inference and prediction of the IW distribution for type-II censored data and [10] produced a new model, generalized IW distribution.

Deriving out a new model with a baseline distribution using a transformation is common in practice but finding a model that best fits for a real data set than other related models are of great relavance. There are so many transformations in existance where the transformation named KM Transformation that studied in [11] is a recent development in the subject in which Exponential distribution and Weibull distribution are used as baseline models to fit real data sets. In this paper, we consider two-parameter Inverse Weibull distribution (IW) as baseline distribution to perform KM Transformation.

This paper is about the KM-IW Model which consists of the introduction to the model with its probability density function (pdf), cumulative distribution function (cdf) and hazard function in section 2. The basic properties of a model such as Moments, Moment generating function, Quantile function and Distribution of order Statistic are studied in Section 3. In section 4, estimation of the model parameters are done by using the method of maximum likelihood estimation. Section 5 gives the results on the simulation study of the model and section 6 gives the results on the fitting of KM-IW model to a real data set. And the conclusions achieved throughout the paper is mentioned in Section 7.

2. KM-IW Distribution

Let X be the random variable of interest. Then using the KM Transformation method the pdf, cdf and hazard function of X is obtained by the formulae given below respectively.

$$f(x) = \frac{e}{e-1}\, g(x)\, e^{-G(x)}$$

$$F(x) = \frac{e}{e-1}\, [1 - e^{-G(x)}]$$

and

$$h(x) = \frac{g(x)\, e^{1-G(x)}}{e^{1-G(x)} - 1}$$

where, g(x) and G(x) are the pdf and cdf of IW distribution which are as given below.

$$g(x) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{-(\alpha+1)} e^{-\left(\frac{x}{\beta}\right)^{-\alpha}},\ x > 0,\ \alpha,\ \beta > 0$$

and

$$G(x) = e^{-\left(\frac{x}{\beta}\right)^{-\alpha}},\ x > 0,\ \alpha,\ \beta > 0$$

Then the following are being the pdf, cdf and hazard function of our new model KM-IW distribution.

$$f(x) = \frac{e}{e-1} \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{-(\alpha+1)} e^{-\left(\frac{x}{\beta}\right)^{-\alpha}} e^{-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}},\ x > 0,\ \alpha,\ \beta > 0 \tag{1}$$

$$F(x) = \frac{e}{e-1} \left[1 - e^{-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}}\right],\ x > 0,\ \alpha,\ \beta > 0 \tag{2}$$

and

$$h(x) = \frac{\frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{-(\alpha+1)} e^{1-\left(\frac{x}{\beta}\right)^{-\alpha} - e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}}}{e^{1-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}} - 1},\ x > 0,\ \alpha,\ \beta > 0 \tag{3}$$

This newly proposed KM-IW Distribution has the pdf and hazard function plots as given below which are obtained using the R programming language.

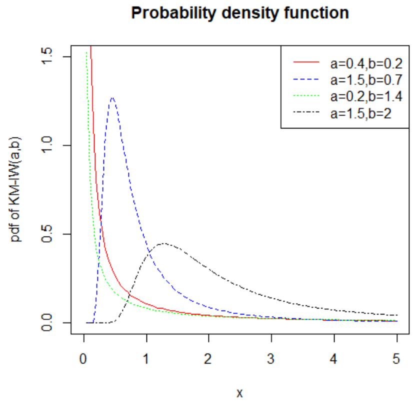


Figure 1: *Probability Density Function of KM-IW Distribution*

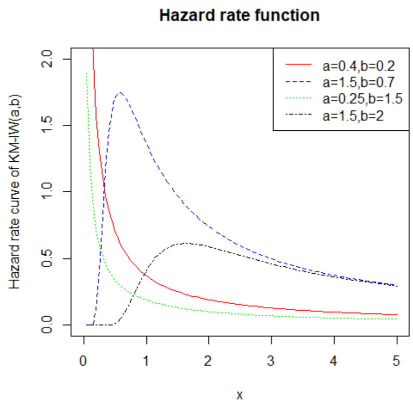


Figure 2: *Hazard Rate Function of KM-IW Distribution*

As seen in the plots, both pdf and hazard rate function have upside-down and decreasing curves for different combinations of α and β.

3. Properties of KM-IW Distribution

Here we discuss some statistical properties of the new model. They are Moments of the distribution, Moment generating function, Characteristic function, Quantile function and Order Statistic.

3.1. Moments of the Distribution

The r^{th} raw moment about origin of the distribution is obtianed as below.

$$\begin{aligned} \mu_r' &= E\left(X^r\right) \\ &= \int_0^\infty x^r f(x)\,dx \\ &= \int_0^\infty x^r \frac{e}{e-1}\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{-(\alpha+1)} e^{-\left(\frac{x}{\beta}\right)^{-\alpha}} e^{-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}}\,dx \end{aligned}$$

The exponential term is expanded, so we get,

$$\mu'_r = \frac{\alpha\, \beta^\alpha\, e}{e-1} \int_0^\infty x^r\, x^{-(\alpha+1)} e^{-\left(\frac{x}{\beta}\right)^{-\alpha}} \sum_{m=0}^{\infty} \frac{\left[-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}\right]^m}{m!}\, dx$$

On proper substitution and simplification, we get,

$$\mu'_r = \frac{e}{e-1} \sum_{m=0}^{\infty} \frac{(-1)^m\, \Gamma\left(-\frac{r}{\alpha}+1\right)(m+1)^{\frac{r}{\alpha}} \beta^r}{(m+1)!} \tag{4}$$

Putting $r = 1$, we get the 1^{st} raw moment about origin (mean of the distribution) and is given by,

$$\mu'_1 = E\,(X)$$

i.e.,

$$\mu'_1 = \frac{e}{e-1} \sum_{m=0}^{\infty} \frac{(-1)^m\, \Gamma\left(-\frac{1}{\alpha}+1\right)(m+1)^{\frac{1}{\alpha}} \beta}{(m+1)!} \tag{5}$$

3.2. Moment Generating Function

The moment generating function (mgf) of the distribution is obtianed as below.

$$\begin{aligned}
M_X(t) &= E\left(e^{tX}\right) \\
&= \int_0^\infty e^{tx} f(x)\, dx \\
&= \int_0^\infty e^{tx} \frac{e}{e-1} \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{-(\alpha+1)} e^{-\left(\frac{x}{\beta}\right)^{-\alpha}} e^{-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}}\, dx
\end{aligned}$$

on expanding the exponential term we get,

$$\begin{aligned}
M_X(t) &= \frac{\alpha\, \beta^\alpha\, e}{e-1} \int_0^\infty e^{tx}\, x^{-(\alpha+1)} e^{-\left(\frac{x}{\beta}\right)^{-\alpha}} \sum_{m=0}^{\infty} \frac{\left[-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}\right]^m}{m!}\, dx \\
&= \frac{\alpha\, \beta^\alpha\, e}{e-1} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \int_0^\infty x^{-(\alpha+1)} e^{-(m+1)\left(\frac{x}{\beta}\right)^{-\alpha}} \sum_{n=0}^{\infty} \frac{(tx)^n}{n!}\, dx
\end{aligned}$$

after simplification, we get,

$$M_X(t) = \frac{e}{e-1} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^m\, t^n\, \Gamma\left(-\frac{n}{\alpha}+1\right)(m+1)^{\frac{n}{\alpha}} \beta^n}{(m+1)!\,n!} \tag{6}$$

Differentiating the mgf will yield the moments about origin of the distribution.
Similarly, the characteristic function is obtianed as

$$\phi_X(t) = \frac{e}{e-1} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^m\, (it)^n\, \Gamma\left(-\frac{n}{\alpha}+1\right)(m+1)^{\frac{n}{\alpha}} \beta^n}{(m+1)!\,n!} \tag{7}$$

where, $i^2 = -1$

3.3. Quantile Function

For obtaining the p^{th} quantile function (Q(p)) of KM-IW distribution, we solve $F(Q(p)) = p$, where $0 < p < 1$. And is obtained as,

$$Q(p) = \beta \left[-\log\left[-\log\left(1 - \frac{p(e-1)}{e}\right)\right]\right]^{\frac{-1}{\alpha}} \tag{8}$$

3.4. Order statistic

Let the random sample of size n, $X_1, X_2, ..., X_n$ be from the new distribution. Then $X_{(1)}, X_{(2)}, ..., X_{(n)}$ are the order statistics respectively. The pdf and cdf of the r^{th} order statistic $f_r(x)$ and $F_r(x)$ are given by

$$f_r(x) = \frac{n!}{(r-1)!(n-r)!} F^{r-1}(x) \left[1 - F(x)\right]^{n-r} f(x)$$

and

$$F_r(x) = \sum_{j=r}^{n} \binom{n}{j} F^j(x) \left[1 - F(x)\right]^{n-j}$$

And those of our newly proposed model are,

$$\begin{aligned} f_r(x) \quad &= \quad \frac{n!\, \alpha\, \beta^{\alpha}}{(r-1)!(n-r)!(e-1)^n} x^{-(\alpha+1)} e^{\left[r - \left(\frac{x}{\beta}\right)^{-\alpha} - e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}\right]} \\ &\quad \left(1 - e^{-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}}\right)^{r-1} \left(e^{1-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}} - 1\right)^{n-r} \end{aligned} \tag{9}$$

and

$$\begin{aligned} F_r(x) \quad &= \quad \sum_{j=r}^{n} \binom{n}{j} \frac{e^j}{(e-1)^n} \left(1 - e^{-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}}\right)^{j} \\ &\quad \left(e^{1-e^{-\left(\frac{x}{\beta}\right)^{-\alpha}}} - 1\right)^{n-j} \end{aligned} \tag{10}$$

respectively.

4. Estimation of parameters

This section is about the estimates of the parameters involved in the distribution. Here we are using the maximum likelihood estimation method.
The likelihood function of KM-IW distribution is found by,

$$L(x, \alpha, \beta) = \prod_{i=1}^{n} f\left(x_i, \alpha, \beta\right)$$

i.e., we will have,

$$L(x) = \left(\frac{e}{e-1}\right)^n \frac{\alpha^n}{\beta^n} \left[\prod_{i=1}^{n} \left(\frac{x_i}{\beta}\right)^{-(\alpha+1)}\right] e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{-\alpha}} e^{-\sum_{i=1}^{n} e^{-\left(\frac{x_i}{\beta}\right)^{-\alpha}}}$$

So that the log-likelihood function becomes,

$$\begin{aligned} \log L \quad &= \quad n \log e - n \log(e-1) + n \log \alpha + n\alpha \log \beta - (\alpha+1) \sum_{i=1}^{n} \log x_i \\ &\quad - \sum_{i=1}^{n} \left(\frac{x_i}{\beta}\right)^{-\alpha} - \sum_{i=1}^{n} e^{-\left(\frac{x_i}{\beta}\right)^{-\alpha}} \end{aligned} \tag{11}$$

The partial derivatives of log L with respect to the unknown parameters α and β are obtianed. Equating these non-linear equations to zero gives the MLEs of α and β.

$$\frac{n}{\alpha} + n \log \beta - \sum_{i=1}^{n} \log x_i + \sum_{i=1}^{n} \left(\frac{x_i}{\beta}\right)^{-\alpha} \log\left(\frac{x_i}{\beta}\right) - \sum_{i=1}^{n} e^{-\left(\frac{x_i}{\beta}\right)^{-\alpha}} \left(\frac{x_i}{\beta}\right)^{-\alpha} \log\left(\frac{x_i}{\beta}\right) = 0 \tag{12}$$

$$\frac{\alpha}{\beta}\left[n - \sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{-\alpha} + \sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{-\alpha} e^{-\left(\frac{x_i}{\beta}\right)^{-\alpha}}\right] = 0 \tag{13}$$

These equations 12 and 13 cannot be solved analytically and by using statistical softwares it can be possible.

5. Simulation Study

For different combinations of α and β, samples of sizes 25, 50, 100, 500 and 1000 are generated from the KM-IW model.
We calculate the bias and mean square error (MSE)s of the estimates. Simulation is conducted for two different combinations of parameter values which are $\alpha = 2$, $\beta = 0.5$ and $\alpha = 0.5$, $\beta = 1.5$. As the sample size increases, the mse value decreases (see Table 1 and 2). The bias of the estimates approaches to zero as n increases. And the estimate values approaches to the true parameter values.

Table 1: *Simulation study at $\alpha = 2$ and $\beta = 0.5$*

n	Estimated value of Parameters	Bias	MSE
25	$\hat{\alpha} = 2.0849208$	0.1165587	0.1566783
	$\hat{\beta} = 0.5670579$	0.005577362	0.002953267
50	$\hat{\alpha} = 1.8655174$	0.0514391	0.06337958
	$\hat{\beta} = 0.5121194$	0.00162314	0.001382758
100	$\hat{\alpha} = 2.000336$	0.0207338	0.02929086
	$\hat{\beta} = 0.513538$	0.001644057	0.0006789977
500	$\hat{\alpha} = 1.9363471$	-0.01219562	0.004949019
	$\hat{\beta} = 0.5115159$	0.0007829622	0.0001313705
1000	$\hat{\alpha} = 2.021294$	-0.01847309	0.002710569
	$\hat{\beta} = 0.488996$	0.0008567411	6.075921×10^{-05}

Table 2: *Simulation study at $\alpha = 0.5$ and $\beta = 1.5$*

n	Estimated value of Parameters	Bias	MSE
25	$\hat{\alpha} = 0.5212292$	0.02913895	0.009792336
	$\hat{\beta} = 2.4815462$	0.1813135	0.6622222
50	$\hat{\alpha} = 0.4663787$	0.01285905	0.003961199
	$\hat{\beta} = 1.6508171$	0.07065759	0.2386087
100	$\hat{\alpha} = 0.5000835$	0.005182735	0.00183067
	$\hat{\beta} = 1.6691755$	0.04467576	0.1099399
500	$\hat{\alpha} = 0.4840861$	-0.003049672	0.0003093182
	$\hat{\beta} = 1.6430402$	0.01415897	0.01952109
1000	$\hat{\alpha} = 0.5053225$	-0.004619055	0.0001694182
	$\hat{\beta} = 1.3722554$	0.0124857	0.008962368

6. Real Data Analysis

Here we use KM-IW(α, β) distribution to fit a real data set and compare the results with IW distribution, KME distribution and KMW distribution. The data-set, considered here, represents survival times in Days, from a Two-Arm Clinical Trial considered by [8] and [18]. The survival time in days for the 31 patients from Arm B are:

Table 3: *survival time in days for the 31 patients from Arm B*

37	84	92	94	110	112	119	127	130
133	140	146	155	159	173	179	194	195
209	249	281	319	339	432	469	519	633
725	817	1557	1776					

The analysis is performed by using R programming language. Table 4 gives the estimates of the model parameters, AIC (Akaike information criterion) and the BIC (Bayesian information criterion) values.

$$AIC = -2l + 2k \tag{14}$$

$$BIC = -2l + k \log n \tag{15}$$

where, l denotes the log-likelihood function, k is the number of parameters and n is the sample size. And Kolmogrov-Simnorov (K-S) test is also performed and the p-value is used for comparison.

Table 4: *Results of the Data Analysis*

Model	Estimates	AIC	BIC	KS statistic	p-value
KM-IW	$\hat{\alpha} = 1.1949$ $\hat{\beta} = 190.7256$	417.3394	420.2074	0.0847	0.9657
IW	$\hat{\alpha} = 1.3375$ $\hat{\beta} = 150.4226$	417.4484	420.3164	0.0897	0.9452
KME	$\hat{\lambda} = 0.0022$	426.1304	427.5644	0.2112	0.1085
KMW	$\hat{\alpha} = 1.1864$ $\hat{\beta} = 464.2267$	426.3770	429.2450	0.1779	0.2493

From the Table 4, we can see that our model KM-IW has better AIC, BIC values, KS statistic value and p-value than IW, KMW and KME distributions. So we can conclude that the newly proposed model is a better fit for the data taken than the other models.

7. Conclusions

A new model is introduced by modifying the Inverse Weibull distribution using the KM Transformation. Its statistical properties are studied along with the estimation of parameters. A simulation study is carried out for 2 parameter combinations and a data set is fitted by the KM-IW model which yielded a better AIC, BIC, KS-statistic values and p-value than the models compared in the study.

References

[1] Adamidis, K., and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics and Probability Letters*, 39(1):35–42.

[2] Akgul, F. G., Senoglu, B. and Arslan, T. (2016). An alternative distribution to Weibull for modeling the wind speed data: Inverse Weibull distribution. *Energy Conversion and Management*, 114:234–240.

[3] Calabria, R. and Pulcini, G. (1989). Confidence Limits for Reliability and Tolerance Limits in the Inverse Weibull Distribution. *Reliability Engineering & System Safety*, 24:77–85.

[4] Calabria, R. and Pulcini, G. (1994). Bayes 2-sample prediction for the inverse weibull distribution. *Communication in Statistics-Theory and Methods*, 23:1811–1824.

[5] Carrasco, J. M. F., Ortega, E. M. M. and Cordeiro, G. M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics and Data Analysis*, 53:450–462.

[6] Chakrabarty, J. B. and Chowdhury, S. (2018). Compounded inverse Weibull distributions: Properties, inference and applications. *Communications in Statistics - Simulation and Computation*, 48(7):2012–2033.

[7] Drapella, A. (1993). The complementary Weibull distribution: Unknown or just forgotten?. *Quality and Reliability Engineering International*, 9:383–386.

[8] Efron, B (1988). Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of the American Statistical Association*, 83:414–425.

[9] Gauthami, P. and Chacko, V. M. (2021). Dus Transformation of Inverse Weibull Distribution: An Upside-Down Failure Rate Model. *Reliability Theory and Applications*, 16(2):58–71.

[10] Gusmao, F. R. S., Ortega, E. M. M. and Cordeiro, G. M. (2011). The generalized inverse Weibull distribution. *Statistical Papers*, 52:591–619.

[11] Kavya, P. and Manoharan, M. (2021). Some parsimonious models for lifetimes and applications. *Journal of Statistical Computation and Simulation*, 91(18):3693–3708.

[12] Khan, M. S., Pasha, G. R. and Pasha, A. H. (2008). Theoretical Analysis of Inverse Weibull Distribution. *WSEAS TRANSACTIONS on MATHEMATICS*, 7:30–38.

[13] Kundu, D. and Howlader, H. (2010). Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data *Computational Statistics and Data Analysis*, 54:1547–1558.

[14] Lawless, J. F. Statistical Models and Methods for Lifetime Data, John Wiley & Sons, New York, 1982.

[15] Lee, E. T. and Wang, J. W. Statistical Methods for Survival Data Analysis, John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.

[16] Mahmoud, M. A. W., Sultan, K. S. and Amer, S. M. (2003). Order statistics from inverse weibull distribution and associated inference. *Computational Statistics & Data Analysis*, 42:149–163.

[17] Manoharan, M. and Kavya, P. (2022). A New Reliability Model and Applications. *Reliability Theory and Applications*, 17(1):65–75.

[18] Mudholkar, G. S., Srivastava, D. K. and Kollia, G. D. (1996). A Generalization of the Weibull Distribution with Application to the Analysis of Survival Data. *Journal of the American Statistical Association*, 91.

[19] Nassar, M. M. and Eissa, F. H. (2003). On the Exponentiated Weibull Distribution. *Communications in Statistics - Theory and Methods*, 32(7):1317–1336.

[20] Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 293–297.

CLASSICAL AND BAYESIAN STOCHASTIC ANALYSIS OF A TWO UNIT PARALLEL SYSTEM WITH WORKING AND REST TIME OF REPAIRMAN

Vashali Saxena[1*], Rakesh Gupta[1], Prof. Bhupendra Singh[1]

•

[1]Department of Statistics,
Ch. Charan Singh University, Meerut-250004(India)
*vaishalistat412@gmail.com;smprgccsu@gmail.com;bhupendra.rana@gmail.com

Abstract

The aim of the present paper is to deal with the analysis of the classical and Bayesian estimation of various measures of system effectiveness in a two non-identical unit parallel system. Each unit has two possible modes Normal (N) and total failure (F). A single repairman is always available with the system and after working for a random period he goes for rest for a random period. After taking complete rest he again starts the repair of the failed unit on a pre-emptive repeat basis. The system failure occurs when both the units are in (F-mode). The distributions of failure time as well as working and rest time of repairman are assumed to be exponential whereas repair time and rest time distribution of repairman are taken as general. A simulation study is also conducted for analysing the considered system model both in Classical and Bayesian setups. Bayesian estimates of various measure of system effectiveness are also obtained by taking different priors. The comparative study is made to judge the performance of Maximum likelihood estimation and Bayesian estimation methods. A simulation study at the end exhibits the behaviour of such a system. The Monte-Carlo technique is employed to draw observations for this simulation study. To obtain various interesting measure of system effectiveness technique have used the Regenerative point technique, MCMC technique and Gibbs sampler technique. From the graphs and tables we have drawn various important conclusions such that a smaller value of failure rate α_1 introduces a larger value of Maximum likelihood estimate and Bayes estimates for fixed value of the parameter of the repairman rest time distribution β. Moreso, when the value of the failure rate α_1 increases the mean time to system failure and net expected profit are also decreases. To compare the performance of asymptotic confidence interval and highest posterior density interval with the maximum likelihood estimates technique, it has been observed that width of the highest posterior density interval is less than the width of an asymptotic confidence interval.

Keywords: Transition probabilities, Mean sojourn time, mean time to system failure, Pre-emptive repeat repair, Regenerative point, Bayesian estimation, highest posterior density intervals, Maximum likelihood estimation, Gibbs sampler.

1. Introduction

In the planning, design, and operation of different stages of complex systems, the evaluation of high reliability is an important criterion. A two-component redundant system is frequently used to improve reliability as well as availability with maximum expected revenue. A large number of authors have analyzed the two identical and non-identical unit parallel system models in respect of their classical estimates of various measures of system effectiveness. Gupta et.al. [6] analyzed a two non-identical unit parallel system with two independent repairman-skilled and ordinary. A failed unit is first attended to by a skilled repairman to perform first phase repair and then it goes for second phase repair by an ordinary repairman. Both types of repair discipline are FCFS. Chaudhary et.al. [2] analyzed a two non-identical unit parallel system model assuming that an administrative delay occurs in getting the repairman available with the system whenever needed. Chandra et. Al. [1] performed the reliability and cost-benefit analysis of the two identical and non-identical unit parallel system modes by using the Semi-Markov process in the regenerative point technique. A study of comparison is made between the reliability characteristics for both the system models under study.

Realistic situations may arise when a repairman can't repair a failed unit continuously for a long period due to his tiredness/ fatigue as after some time, the working efficiency of the repairman may reduce and he needs to rest for some time. In view of this K. Murari et. al. [9] has analyzed a 2-unit parallel system with a single repairman assuming the working as well as rest period of a repairman. Gupta et. al. [7] has analyzed a single-server two-unit (priority and ordinary) cold standby system with two modes—normal and total failure. The priority unit gets preference both for operation and repair. After working for a random amount of time, the operator of the system needs to rest for a random amount of time and during the rest period of the operator, the system becomes down but not failed. The system failure occurs when both the units are in total failure mode.

Kishan et. Al. [8] analyzed of reliability characteristics of a two-unit parallel system under classical and Bayesian setups. They assumed that the system consists of two non-identical units arranged in a parallel configuration. System failure occurs when both the units stop functioning. Gupta et.al. [5] Performed the cost-benefit and reliability analysis of a two-unit cold standby system assuming that a failed unit enters into the fault detection to identify whether the failed unit needs minor or major repair with fixed known probabilities. Keeping the above idea in view, the present study deals with the analysis of two non-identical units in a parallel system model assuming that the working and rest time of the repairman are uncorrelated random variables. A single repairman is always available with the system to repair a failed unit with priority given to one of the units. The repairman also goes for rest after some time as he is unable to work continuously for a long time. After taking complete rest he again starts the repair of a failed unit assuming that the time already spent in the repair of the failed unit goes to waste. For a more concrete study of the system model, a simulation study is also carried out:

The probability density function (PDF) of exponential is given by:

$$f(t) = \theta e^{-\theta x}; \quad x \geq 0$$

In addition, it has been seen in practice that lifetime experiments take a long period because the ambient circumstances cannot be the same during the trial. As a result, random variables are considered for parameters characterizing the system/dependability unit's characteristics. Therefore, a simulation study is conducted for analysing the considered system model both in classical and Bayesian setups. The Monte Carlo simulation technique has been used in conducting the numerical study in a classical setup, the maximum likelihood estimate of the parameters involved in the model and reliability characteristics along with their standard errors and width of confidence intervals are obtained. In the Bayesian setup, Bayes estimates of the parameters and reliability characteristics along with their posterior standard error and width of highest posterior density intervals are computed. In the end, the comparative conclusions are drawn to judge the performance of the maximum likelihood and Bayes estimates.

The following economic related measures of system effectiveness are obtained by using regenerative point technique:

- Transition probabilities and mean sojourn times in various states.
- Reliability and mean time to system failure.
- Point-wise and steady state availability of the system as well as expected up time of the system during interval (0, t).
- Expected busy period of the repairman during time interval (0, t).
- Net expected profit earned by the system during a finite and interval and in steady state.

2. Model Description and Assumptions

- The system consists of two non-identical units –unit 1 and unit 2. Both the units work in parallel configuration.
- Each unit has two possible modes: Normal (N) and total failure (F).
- A single repairman is always available with the system. After working for a random period, he goes for rest due to his tiredness as after some time, the working efficiency of the repairman may reduce and he needs rest for some time. After taking complete rest the repairman again starts the repair of the failed unit.
- The failure time distribution, as well as the working time of the repairman, are assumed to be exponential whereas the repair time distribution of each unit and rest time of the repairman is taken as general. The repair discipline for the repair of units is FCFS.
- The system failure occurs when both the unit are in total F-mode.
- A repaired unit always works as good as new.

The transition diagram of the system model is shown in Figure [1].

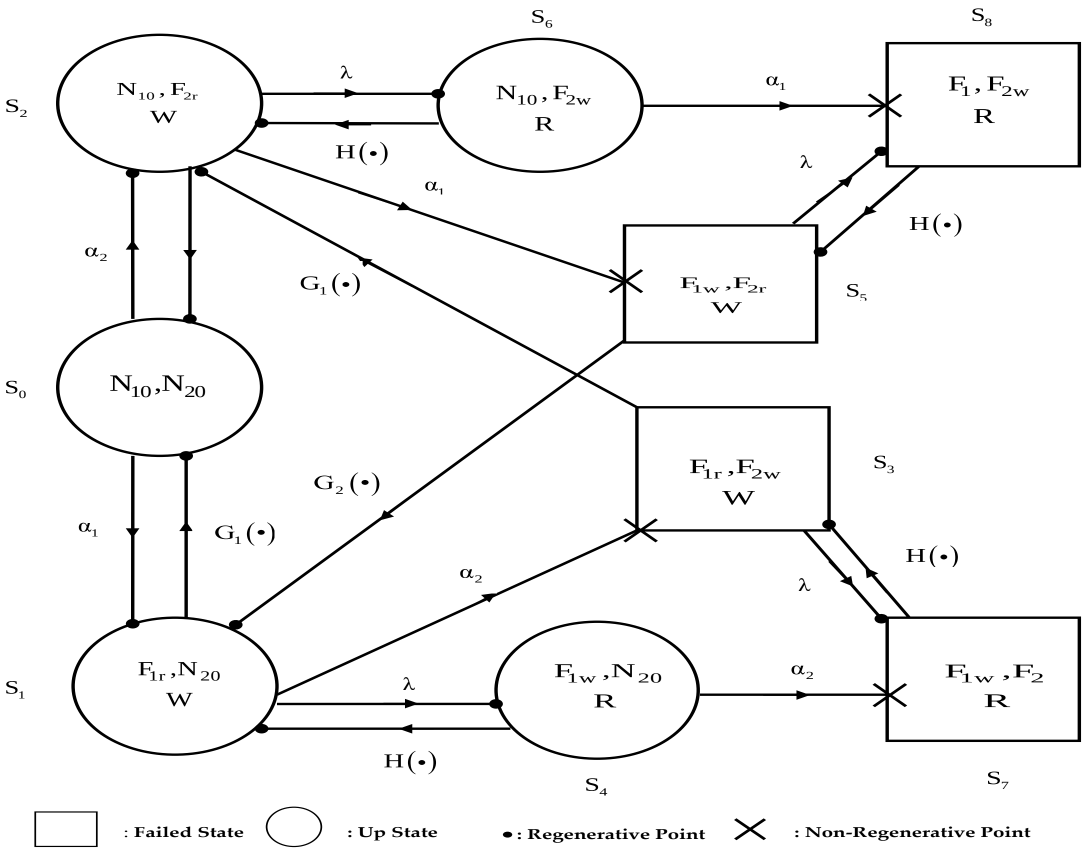


Figure. 1 Transition diagram

3. Notations and Sates of the system

3.1. Notations

α_1, α_2	:	Constant failure rates of first and second type of unit.
$G_1(\bullet), G_2(\bullet)$	:	c.d.f of time to complete repair for first and second unit respectively.
λ	:	Constant rate by which a repairman goes for rest.
$H(\bullet)$	:	General c.d.f rate which a repairman goes to working position after rest.

3.2. Symbols for the state of the system

N_{10}, N_{20}	:	Unit 1 and Unit 2 are N-mode and operative mode.
F_{1r}, F_{2r}	:	Unit 1 and Unit 2 are in F-mode and under repair respectively.
F_{1w}, F_{2w}	:	Unit 1 and Unit 2 are waiting for repair respectively.
F_1, F_2	:	Unit 1 and Unit 2 are failed respectively.

Considering the above symbols, we have the following states of the system:-

Up states: $S_0=(N_{10},N_{20})$, $S1=(F_{1r},N_{20})$, $S2=(N_{10},F_{2r})$, $S_4=(F_{1w},N_{20})$, $S_6=(N_{10},F_{2w})$

Failed states: $S_3=(F_{1r},F_{2w})$, $S_5=(F_{1w},F_{2r})$, $S_7=(F_{1w},F_2)$, $S_8=(F_1,F_{2w})$

4. Transition Probabilities and Sojourn times

The non-zero elements of one and more steps steady state transition probabilities from state S_i to S_j are as follows-

$$p_{ij}=\lim_{t\to\infty}Q_{ij}(t) \qquad p_{ij}^{(k)}=\lim_{t\to\infty}Q_{ij}^{(k)}(t)$$

$$p_{01}=\int\alpha_1 e^{-(\alpha_1+\alpha_2)t}dt=\frac{\alpha_1}{(\alpha_1+\alpha_2)} \qquad p_{02}=\int\alpha_2 e^{-(\alpha_1+\alpha_2)t}dt=\frac{\alpha_2}{(\alpha_1+\alpha_2)}$$

$$p_{10}=\tilde{G}_1(\lambda+\alpha_2) \qquad p_{14}=\frac{\lambda}{(\lambda+\alpha_2)}\left[1-\tilde{G}_1(\lambda+\alpha_2)\right]$$

$$p_{12}^{(3)}=\frac{\alpha_2}{(\lambda+\alpha_2)}-\left[\tilde{G}_1(\lambda)-\tilde{G}_1(\lambda+\alpha_2)\right] \qquad p_{17}^{(3)}=\tilde{G}_1(\lambda)-\tilde{G}_1(\lambda+\alpha_2)$$

$$p_{20}=\tilde{G}_2(\lambda+\alpha_1) \qquad p_{26}=\frac{\lambda}{(\lambda+\alpha_1)}\left[1-\tilde{G}_2(\lambda+\alpha_1)\right]$$

$$p_{21}^{(5)}=\frac{\lambda}{(\lambda+\alpha_1)}-\left[\tilde{G}_2(\lambda)-\tilde{G}_2(\lambda+\alpha_1)\right] \qquad p_{28}^{(5)}=\tilde{G}_1(\lambda)-\tilde{G}_1(\alpha_1+\lambda)$$

$$p_{32}=\tilde{G}_1(\lambda) \qquad p_{37}=\left[1-\tilde{G}_1(\lambda)\right]$$

$$p_{41}=\tilde{H}(\alpha_2)$$

$$p_{51}=\tilde{G}_2(\lambda) \qquad p_{58}=\left[1-\tilde{G}_2(\lambda)\right]$$

$$p_{73}=p_{85}=1$$

It can be easily verified that

$P_{01}+P_{02}=1,$ $\quad P_{10}+P_{14}+P_{12}^{(3)}+P_{17}^{(3)}=1$

$P_{20}+P_{26}+P_{21}^{(5)}+P_{28}^{(5)}=1,$ $\quad P_{32}+P_{37}=1$

$P_{41}+P_{43}^{(7)}=1,$ $\quad P_{51}+P_{58}=1$

$P_{62}+P_{65}^{(8)}=1,$ $\quad P_{73}=P_{85}=1$

Mean sojourn time ψ_i in state S_i is defined as the expected time taken by the system in state S_i before transition to any other state.

$$\psi_0=\int P\left(T_0>t\right)dt=\int e^{-(\alpha_1+\alpha_2)t}dt=\frac{1}{\left(\alpha_1+\alpha_2\right)}$$

Similarly,

$$\psi_1=\int e^{-(\lambda+\alpha_2)t}\,\bar{G}_1\left(t\right)dt, \qquad \psi_2=\int e^{-(\lambda+\alpha_1)t}\,\bar{G}_2\left(t\right)dt,$$

$$\psi_3=\int e^{-\lambda t}\bar{G}_1\left(t\right)dt, \qquad \psi_4=\int e^{-\alpha_2 t}\,\bar{H}\left(t\right)dt$$

$$\psi_5=\int e^{-\lambda t}\,\bar{G}_2\left(t\right)dt \qquad \psi_6=\int e^{-\alpha_1 t}\,\bar{H}\left(t\right)dt$$

$$\psi_7=\psi_8=\int \bar{H}\left(t\right)dt=m$$

5. Methodology for Developing Equations

To obtain various interesting measures of system effectiveness some techniques are available such as the semi-Markov process and regenerative-point technique [1-18]. The present study deals with the technique of regenerative point as it is easy to handle the problem when the behavior of the system at some epochs of entrance into the states is Non-Markovian. We develop the recurrence relations for reliability, availability, and busy period of repairman as follows-

5.1. Reliability of the system

Here we define $R_i\left(t\right)$ as the probability that the system does not fail up to t epochs 0,1, 2,…..(t-1) when it is initially started from up state S_i. To determine it, we regard the failed state S_3, S_5, S_7, S_8 as absorbing states. Now, for the expressions of $R_i\left(t\right)$; i=0, 1,2,4,6. By simple probabilistic arguments, the value of $R_0\left(t\right)$ in terms of its Laplace Transform (L.T.) is given by,

$$R_0^*\left(s\right)=\frac{\begin{array}{c}\left(1-q_{26}^*q_{62}^*-q_{14}^*q_{41}^*+q_{14}^*q_{41}^*q_{26}^*q_{62}^*\right)Z_0^*+\left(q_{01}^*-q_{01}^*q_{26}^*q_{62}^*\right)Z_1^*+\left(q_{02}^*-q_{02}^*q_{14}^*q_{41}^*\right)Z_2^*+\\ \left(q_{01}^*q_{14}^*+q_{01}^*q_{14}^*q_{26}^*q_{62}^*\right)Z_4^*+\left(q_{02}^*q_{26}^*-q_{14}^*q_{41}^*q_{26}^*q_{62}^*\right)Z_6^*\end{array}}{(1-q_{26}^*q_{62}^*-q_{14}^*q_{41}^*+q_{14}^*q_{41}^*q_{26}^*q_{62}^*)+q_{10}^*\left(q_{01}^*-q_{01}^*q_{26}^*q_{62}^*\right)-q_{20}^*\left(q_{02}^*-q_{02}^*q_{14}^*q_{41}^*\right)} \tag{1}$$

We have omitted the arguments from $q_{ij}^*\left(s\right)$ and $Z_i^*\left(s\right)$ for brevity, $Z_i^*\left(s\right)$; i=0,1,2,4,6 are the L.T. of

$Z_0\left(t\right)=e^{-(\alpha_1+\alpha_2)t}$, $\quad Z_1\left(t\right)=e^{-(\lambda+\alpha_2)t}\,\bar{G}_1\left(t\right)$,

$Z_2\left(t\right)=e^{-(\lambda+\alpha_1)t}\,\bar{G}_2\left(t\right)$, $\quad Z_4\left(t\right)=e^{-(\alpha_2)t}\,\bar{H}\left(t\right)$,

$Z_6\left(t\right)=e^{-(\alpha_1)t}\,\bar{H}\left(t\right)$,

Taking the inverse Laplace transform of [1], one can get the reliability of the system when system initially starts from S_0. The MTSF is given by

$$E\left(T_0\right)=\lim_{s\to 0}R_0^*\left(s\right)=\frac{N_1\left(0\right)}{D_1\left(0\right)}$$

$$R_0^*(s)=\frac{\left(1-p_{26}p_{62}-p_{14}p_{41}+p_{14}p_{41}p_{26}p_{62}\right)\psi_0+\left(p_{01}-p_{01}p_{26}p_{62}\right)\psi_1+\left(p_{02}-p_{02}p_{14}p_{41}\right)\psi_2+\left(p_{01}p_{14}+p_{01}p_{14}p_{26}p_{62}\right)\psi_4+\left(p_{02}p_{26}-p_{14}p_{41}p_{26}p_{62}\right)\psi_6}{\left(1-p_{26}p_{62}-p_{14}p_{41}+p_{14}p_{41}p_{26}p_{62}\right)+p_{10}\left(p_{01}-p_{01}p_{26}p_{62}\right)-p_{20}\left(p_{02}-p_{02}p_{14}p_{41}\right)}$$

5.2. Availability of the system

Let $A_i(t)$ be the probability that the system is up at epoch t, when initially it starts operation from state $S_i \in E$. Using Regenerative point technique and the tools of Laplace Transform, one can obtain the value of in terms of its Laplace Transforms i.e. $A_0^*(s)$ given as follows:

$$A_0^*(s)=\frac{N_2(s)}{D_2(s)}$$

The steady state values of equation (44) is given by

$$A_0=\lim_{t\to\infty}A_0(t)=\lim_{s\to 0}sA_0^*(s)=\lim_{s\to 0}s\frac{N_2(s)}{D_2(s)}$$

Where,

$$N_2(s)=U_0\psi_0+U_1\psi_1+U_2\psi_2+U_3\psi_3+U_4\psi_4+U_5\psi_5+U_6\psi_6$$

Where,

$$U_0=\left[\begin{array}{l}(1-p_{37})(1-p_{58})+p_{26}p_{62}(1-p_{37})(1-p_{58})-p_{21}^{(5)}p_{17}p_{32}(1-p_{58})-p_{21}^{(5)}p_{14}p_{43}p_{32}(1-p_{58})-p_{14}p_{41}(1-p_{37})(1-p_{58})+\\ \quad p_{26}p_{62}p_{14}(1-p_{37})(1-p_{58})-p_{17}^{(3)}p_{32}p_{26}p_{65}^{(8)}+p_{14}p_{43}^{(7)}p_{32}p_{26}p_{65}\end{array}\right]$$

$$U_1=\left[-p_{01}(1-p_{37})(1-p_{58})+p_{26}p_{62}(1-p_{37})(1-p_{58})-p_{02}p_{21}^{(5)}(1-p_{37})(1-p_{58})+p_{02}p_{26}p_{65}^{(8)}p_{51}(1-p_{37})\right]$$

$$U_2=\left[-p_{01}p_{17}^{(3)}p_{32}(1-p_{58})-p_{01}p_{14}p_{43}^{(7)}p_{32}(1-p_{58})-p_{02}(1-p_{37})(1-p_{58})-p_{14}p_{41}p_{02}(1-p_{37})(1-p_{58})\right]$$

$$U_3=\left[\begin{array}{l}-p_{01}p_{17}^{(3)}(1-p_{58})-p_{01}p_{14}p_{43}^{(7)}(1-p_{58})+p_{01}p_{17}^{(3)}p_{26}p_{62}(1-p_{58})+p_{01}p_{14}p_{43}^{(7)}p_{26}p_{62}(1-p_{58})+p_{02}p_{21}^{(5)}p_{17}^{(3)}(1-p_{58})\\ \quad -p_{02}p_{26}p_{65}^{(8)}p_{51}p_{17}^{(3)}-p_{02}p_{26}p_{65}^{(8)}p_{43}^{(7)}\end{array}\right]$$

$$U_4=\left[-p_{01}p_{14}(1-p_{37})(1-p_{58})-p_{01}p_{14}p_{26}p_{62}(1-p_{37})(1-p_{58})-p_{02}p_{21}^{(5)}p_{14}(1-p_{37})(1-p_{58})\right]$$

$$U_5=\left[p_{01}p_{17}^{(3)}p_{32}p_{26}p_{65}^{(8)}+p_{01}p_{14}p_{43}^{(7)}p_{32}p_{26}p_{65}^{(8)}-p_{02}p_{26}p_{65}^{(8)}(1-p_{37})\right]$$

$$U_6=\left[p_{01}p_{17}^{(3)}p_{32}p_{26}(1-p_{58})+p_{01}p_{14}p_{43}^{(7)}p_{32}p_{26}(1-p_{58})+p_{02}p_{26}(1-p_{37})(1-p_{58})-p_{14}p_{41}p_{02}p_{26}(1-p_{37})(1-p_{58})\right]$$

We observe that

$$D_2(0)=0$$

Therefore by using L. Hospital rule, we get

$$A_0=\lim_{s\to 0}\frac{N_2(s)}{D_2'(s)}=\frac{N_2(0)}{D_2'(0)} \tag{2}$$

Thus we have,

$$D_2'(0)=U_0\psi_0+(U_1+U_3)\psi_3+U_2\psi_5+(U_4+U_6)m+U_5\psi_5 \tag{3}$$

Using the relation $N_2(0)$ and $D_2'(0)$ in equation [2-3], we get the expression for A_0.

The expected up time of the system in interval (0, t) is given by

$$\mu_{up}(t)=\int_0^t A_0(u)\,du$$

So that

$$\mu_{up}^{*}(s)=\frac{A_0^{*}(s)}{s}$$

5.3. Busy Period analysis

Let us define $B_i(t)$ be the probability that the repairman is busy in the repair of a failed unit at epoch t, when the system starts form state .Here by using the basic probabilistic arguments, we have the following relations for B_i (t-1), i=0,1,2,3,4,5,6,7,8. The dichotomous variable δ taken value 0 and 1.

In the long run, fraction of time for which the system is under repair, starting from state S_0 is given by,

$$B_0^{*}=\lim_{t\to\infty}B_0(t)=\lim_{s\to 0}sB_0^{*}(s)=\lim_{s\to 0}s\frac{N_3(s)}{D_2(s)}$$

$D_2(0)=0$, therefore by L Hospital rule, we have

$$B_0=\lim_{s\to 0}\frac{N_3(s)}{D_2'(s)}=\frac{N_3(0)}{D_2'(0)}$$

$$N_3(s)=U_1\psi_1+U_2\psi_2-U_3\psi_3-U_5\psi_5$$

And $D_2'(0)$ is same as given in section [b].

Now by using $N_3(0)$ and $D_2'(0)$, the expression of B_0, can be obtained. The expected busy period of the repairman in repairing in time interval (0, t) is given by

$$\mu_b(t)=\int_0^t B_0(u)\,du$$

So that,

$$\mu_b^{*}(s)=\frac{B_0^{*}(s)}{s}$$

5.4. Profit function analysis

Let us define

K_0 = per-unit up time revenue by the system due to the operation of any unit.

K_1 = repair cost per-unit of time when a unit is under repair.

Here we assume that repair cost of the unit-1 and unit-2 are same.

Then, the net expected total cost incurred in time interval (0, t) is given by

$$P(t)=K_0\mu_{up}(t)-K_1\mu_b(t)$$

The expected total cost incurred in unit interval of time is

$$\frac{P(t)}{t}=K_0\frac{\mu_{up}(t)}{t}-K_1\frac{\mu_b(t)}{t}$$

The expected total cost per-unit time in steady state is given by-

$$\begin{aligned}P&=\lim_{t\to\infty}\frac{P(t)}{t}\\&=K_0\lim_{t\to\infty}\frac{\mu_{up}(t)}{t}-K_1\lim_{t\to\infty}\frac{\mu_b(t)}{t}\\&=K_0\lim_{s\to 0}s^2\mu_{up}^{*}(s)-K_1\lim_{s\to 0}s^2\mu_b^{*}(s)\\&=K_0\lim_{s\to 0}s^2\frac{A_0^{*}(s)}{s}-K_1\lim_{s\to 0}s^2\frac{B_0^{*}(s)}{s}\end{aligned}$$

$$= K_0A_0 - K_1B_0$$

Where A_0 and B_0 have been already defined.

6. Estimation of Parameters, MTSF and Profit Function

6.1. Classical Estimation

In this section, we consider the classical estimation of the model parameters. Suppose that the failure, repair, repairman rest time and working time distribution are independently distributed as Exponential with respective PDF defined in section [1]. Let

$$\underset{\sim}{T}_1 = \left(t_{11}, t_{12}, \ldots\ldots. t_{1n_1}\right), \underset{\sim}{T}_2 = \left(t_{21}, t_{22}, \ldots\ldots. t_{2n_2}\right), \underset{\sim}{T}_3 = \left(t_{31}, t_{32}, \ldots\ldots. t_{3n_3}\right),$$
$$\underset{\sim}{T}_4 = \left(t_{41}, t_{42}, \ldots\ldots. t_{4n_4}\right), \underset{\sim}{T}_5 = \left(t_{51}, t_{52}, \ldots\ldots. t_{5n_5}\right) \text{and } \underset{\sim}{T}_6 = \left(t_{61}, t_{62}, \ldots\ldots. t_{6n_6}\right)$$

Be the random samples respectively drawn from their respective PDF. Then the joint likelihood function is

$$L = L\left(\underset{\sim}{T}_1, \underset{\sim}{T}_2, \underset{\sim}{T}_3, \underset{\sim}{T}_4, \underset{\sim}{T}_5, \underset{\sim}{T}_6 / \alpha_1, \lambda_1, \beta, \alpha_2, \theta, \lambda\right) = \alpha_1^{n_1} e^{-\alpha_1 \sum x_1} * \lambda_1^{n_2} e^{-\lambda_1 \sum x_2} * \beta^{n_3} e^{-\beta \sum x_3} * \alpha_2 * \theta^{n_5} e^{-\theta \sum x_5} * \lambda^{n_6} e^{-\lambda \sum x_6}$$

6.2. Maximum Likelihood Estimation

The log-likelihood function is

$$\log L = n_1 \log \alpha_1 - \alpha_1 \sum x_1 + n_2 \log \lambda_1 - \lambda_1 \sum x_2 + n_3 \log \beta - \beta \sum x_3 + n_4 \log \alpha_2 - \alpha_2 \sum x_4 + n_5 \log \theta - \theta \sum x_5$$
$$+ n_6 \log \lambda - \lambda \sum x_6 \quad (4)$$

By applying the Maximum likelihood approach, the maximum likelihood estimators (MLEs) $\left(\hat{\alpha}_1, \hat{\lambda}_1, \hat{\beta}, \hat{\alpha}_2, \hat{\theta}, \hat{\lambda}\right)$ of the parameters ($\alpha_1, \lambda_1, \beta, \alpha_2, \theta, \lambda$) can be obtained as

$$\hat{\alpha}_1 = \left(1+T_1\right)^{-1},\ \hat{\lambda}_1 = \left(1+T_2\right)^{-1},\ \hat{\beta} = \left(1+T_3\right)^{-1},\ \hat{\alpha}_2 = \left(1+T_4\right)^{-1},\ \hat{\theta} = \left(1+T_5\right)^{-1},\ \hat{\lambda} = \left(1+T_6\right)^{-1}$$

Where,

$$\hat{\alpha}_1 = \frac{n_1}{\sum x_1},\ \hat{\lambda}_1 = \frac{n_2}{\sum x_2},\ \hat{\beta} = \frac{n_3}{\sum x_3},\ \hat{\alpha}_2 = \frac{n_4}{\sum x_4},\ \hat{\theta} = \frac{n_5}{\sum x_5},\ \hat{\lambda} = \frac{n_6}{\sum x_6}$$

The asymptotic distribution of $\left(\hat{\alpha}_1 - \alpha_1, \hat{\lambda}_1 - \lambda_1, \hat{\beta} - \beta, \hat{\alpha}_2 - \alpha_2, \hat{\theta} - \theta, \hat{\lambda} - \lambda\right) \sim N_6\left(0, I^{-1}\right)$ is 6-variate normal distribution, where I is the Fisher information matrix with diagonal elements as

$$i_{11} = E\left[-\frac{\partial^2 \log L}{\partial \alpha_1^2}\right] = \frac{n_1}{\alpha_1^2},\ i_{22} = E\left[-\frac{\partial^2 \log L}{\partial \lambda_1^2}\right] = \frac{n_2}{\lambda_1^2},\ i_{33} = E\left[-\frac{\partial^2 \log L}{\partial \beta}\right] = \frac{n_3}{\beta^2},$$
$$i_{44} = E\left[-\frac{\partial^2 \log L}{\partial \alpha_2^2}\right] = \frac{n_4}{\lambda_4^2},\ i_{55} = E\left[-\frac{\partial^2 \log L}{\partial \theta^2}\right] = \frac{n_5}{\theta^2},\quad i_{66} = E\left[-\frac{\partial^2 \log L}{\partial \lambda^2}\right] = \frac{n_6}{\lambda^2}$$

All off- diagonal elements of I are zero. The maximum likelihood estimator $\hat{M}$ and $\hat{P}$ of MTSF and Profit function can be obtained on using the invariance property of MLE. Also, the asymptotic distribution of $\left(\hat{M} - M\right) \sim N_6\left(0, \Delta_1' I^{-1} \Delta_1\right)$ and $\left(\hat{P} - P\right) \sim N_6\left(0, \Delta_2' I^{-1} \Delta_2\right)$ are respectively,

Where,

$$\Delta_1 = \left(\frac{\partial M}{\partial \alpha_1}, \frac{\partial M}{\partial \lambda_1}, \frac{\partial M}{\partial \beta}, \frac{\partial M}{\partial \alpha_2}, \frac{\partial M}{\partial \theta}, \frac{\partial M}{\partial \lambda}\right) \text{ And } \Delta_2 = \left(\frac{\partial P}{\partial \alpha_1}, \frac{\partial P}{\partial \lambda_1}, \frac{\partial P}{\partial \beta}, \frac{\partial P}{\partial \alpha_2}, \frac{\partial P}{\partial \theta}, \frac{\partial P}{\partial \lambda}\right)$$

6.3. Bayesian Estimation

The Bayesian estimation is used to measure the impact of prior information along with the sample information. Therefore, in this section, the Bayesian method of estimation is also considered for estimating the model parameter. As discussed above, the Bayesian method of the estimation considers the parameters involved in the model as random variable. Here, we estimate the unknown parameters considering the prior distribution of gamma with respective PDFs

$$\alpha_1 \sim \text{Gamma}(a_1,b_1); \qquad (\alpha_1,a_1,b_1)>0, \tag{5}$$

$$\lambda_1 \sim \text{Gamma}(a_2,b_2); \qquad (\lambda_1,a_2,b_2)>0, \tag{6}$$

$$\beta \sim \text{Gamma}(a_3,b_3); \qquad (\beta,a_3,b_3)> 0, \tag{7}$$

$$\alpha_2 \sim \text{Gamma}(a_4,b_4); \qquad (\alpha_2,a_4,b_4)>0, \tag{8}$$

$$\theta \sim \text{Gamma}(a_5,b_5); \qquad (\theta,a_5,b_5)>0, \tag{9}$$

$$\lambda \sim \text{Gamma}(a_6,b_6); \qquad (\lambda,a_6,b_6)>0, \tag{10}$$

Here a_i and b_i (i=1, 2, 3, 4, 5, 6) respectively denote the scale and shape parameters.

Now by using the likelihood in [4] and the priors in [5-10], the posterior distribution of the parameters $\alpha_1,\lambda_1,\beta,\alpha_2,\theta,\lambda$ given data are:

$$w_1(\alpha_1|x_1,\lambda_1,\beta,\alpha_2,\theta,\lambda) \sim \text{Gamma}(n_1+a_1,b_1+x_1), \tag{11}$$

$$w_2(\lambda_1|x_1,\alpha_1,\beta,\alpha_2,\theta,\lambda) \sim \text{Gamma}(n_2+a_2,b_2+x_2), \tag{12}$$

$$w_3(\beta|x_1,\alpha_1,\lambda_1,\alpha_2,\theta,\lambda) \sim \text{Gamma}(n_3+a_3,b_3+x_3), \tag{13}$$

$$w_4(\alpha_2|x_1,\alpha_1,\lambda_1,\beta,\theta,\lambda) \sim \text{Gamma}(n_4+a_4,b_4+x_4), \tag{14}$$

$$w_5(\theta|x_1,\alpha_1,\lambda_1,\beta,\alpha_2,\lambda) \sim \text{Gamma}(n_5+a_5,b_5+x_5), \tag{15}$$

$$w_6(\lambda|x_1,\alpha_1,\lambda_1,\beta,\alpha_2,\theta) \sim \text{Gamma}(n_6+a_6,b_6+x_6), \tag{16}$$

All the prior parameters (also known as hyper parameters) are assumed to be known. These parameters are those whose values are set before the Bayesian learning process start. We utilized the Markov Chain Monte Carlo (MCMC) techniques available that can be used to simulate draws from the posterior distribution and also use the Gibbs sampler, a well-known MCMC algorithm proposed by [1]. It allows us to generate posterior samples for all the parameters using their full conditional posterior distributions.

Now we proceed as follows:

- Simulate α_1 from $w_1(\alpha_1|x_1,\lambda_1,\beta,\alpha_2,\theta,\lambda)$, given in equation [11].
- Simulate λ_1 from $w_2(\lambda_1|x_1,\alpha_1,\beta,\alpha_2,\theta,\lambda)$, given in equation [12].
- Simulate β from $w_3(\beta|x_1,\alpha_1,\lambda_1,\alpha_2,\theta,\lambda)$, given in equation [13].
- Simulate α_2 from $w_4(\alpha_2|x_1,\alpha_1,\lambda_1,\beta,\theta,\lambda)$, given in equation [14].
- Simulate θ from $w_5(\theta|x_1,\alpha_1,\lambda_1,\beta,\alpha_2,\lambda)$, given in equation [15].
- Simulate λ from $w_6(\lambda|x_1,\alpha_1,\lambda_1,\beta,\alpha_2,\theta)$, given in equation [16].

Repeat steps 1-6, N times and record the sequence of $\alpha_1,\lambda_1,\beta,\alpha_2,\theta,\lambda$ after discarding the burn-in-sampler of size, say N_0, from the sample so that the effect of the initial values is neutralised.

Under the squared error loss function, Bayes estimates of $\alpha_1,\lambda_1,\beta,\alpha_2,\theta,\lambda$ are, respectively, the means of posterior distribution given in equations [11-16] and as follows:

$$\hat{\alpha}_1=\frac{n_1}{\sum x_1},\ \hat{\lambda}_1=\frac{n_2}{\sum x_2},\ \hat{\beta}=\frac{n_3}{\sum x_3},\ \hat{\alpha}_2=\frac{n_4}{\sum x_4},\ \hat{\theta}=\frac{n_5}{\sum x_5},\ \hat{\lambda}=\frac{n_6}{\sum x_6}$$

7. Simulation Study

In this section, a simulation study is carried out to investigate the behavior of an assumed system in steady state. Estimates for the parameters of interest as well as the reliability measures were obtained using both the classical maximum likelihood estimation technique and the Bayesian approach. For simulation purposes, random samples of size n were generated for every iteration from the assumed distribution after setting n1=n2=n3=n4=30, 50, 100, 150 to obtain the maximum likelihood estimate and Bayes estimator of the parameters using an expression in 6.1, 6.2 & 6.3 respectively. Then using the asymptotic convergence of maximum likelihood estimate of the standard errors (SE) and the confidence intervals for the maximum likelihood estimate of the reliability and profit have been obtained. For studying the posterior performance of the MTSF and Profit function, the values of hyper-parameters have been so chosen that $E(\alpha_1)=a_1/b_1$, $E(\lambda_1)=a_2/b_2$, $E(\beta)=a_3/b_3$, $E(\alpha_2)=a_4/b_4$, $E(\theta)=a_5/b_5$, $E(\lambda)=a_6/b_6$ to generate the samples from posterior distribution of the parameters. Here simulated results have been obtained by 10,000 iterations for the priors. These simulated posterior values have been used to obtain the posterior estimates of the reliability measures as the means of these simulated values and their posterior standard errors (PSEs). Also, highest posterior density [HPD] for mean time to system failure MTSF and Profit of the assumed system have been obtained using these posterior values.

The True values, maximum likelihood estimates MLE's, and Bayes estimates of mean time to system failure MTSF and Profit function for fixed repair rate α_1 and repairman rest time distribution β respectively, and varying failure rates α_1 of mean time to system failure and profit function have been plotted in fig.[2-3]and [4-5] and also shown in tables [1-4]. The 95% asymptotic confidence intervals and highest posterior density [HPD] intervals of mean time to system failure MTSF and Profit function are plotted in fig. [6-7] and [8-9].

8. Concluding Remarks

From the simulation results in Table [1-4] and various figures [2-9], it is observed that:

- Table [1] and Table [2] give the maximum likelihood estimates and Bayes estimates of mean time to system failure MTSF for various values of failure rate fixed repairman rest time β =0.04 and 0.09, its clear that a smaller value of failure rate α_1 introduces a larger value of maximum likelihood estimates and Bayes estimates for the given value of repairman rest time distribution β . As the value of the failure rate increases the MTSF decreases which shows in fig [2] and fig [3].
- Comparing table [1] and table [2], the value of a parameter of rest time distribution β increased from 0.04 to 0.09, which results in the decay of mean time to system failure MTSF for the given value and which shows in fig. [2-3].
- Table [3] and Table [4] observed that C and Bayes estimates of Profit function for the various value of failure rate α_1 and rest time distribution β= 0.04 and 0.09, it observed that the value of failure rate α_1 increases so the estimate of the true value of Profit function is decreased which shows in fig. [3-4].
- Comparing Table [3] and Table [4] shows that an increment in the value of the parameter of rest time distribution β from 0.04 to 0.09 results in a decay of the Profit function.
- To compare the performance of asymptotic confidence interval and higher posterior density [HPD] interval with the maximum likelihood estimates technique, from all the tables [1-4] and fig. [6-9] it has been seen that width of the highest posterior density [HPD] interval is less than the width of an asymptotic confidence interval.

- Fig [6-9] gives the posterior distribution of reliability measure to show the performance of mean time to system failure MTSF and Profit function at different value of Failure rate α_1 and fixed value of rest time distribution β .

- Comparing fig [6-7] shows that when sample size n=30, 50,100,150 i.e. increases then highest posterior density [HPD] intervals of mean time to system failure MTSF are less than the asymptotic confidence interval of mean time to system failure MTSF. The same trend can see in the highest posterior density [HPD] interval and Confidence interval of the Profit function in fig [8-9].

Graphs

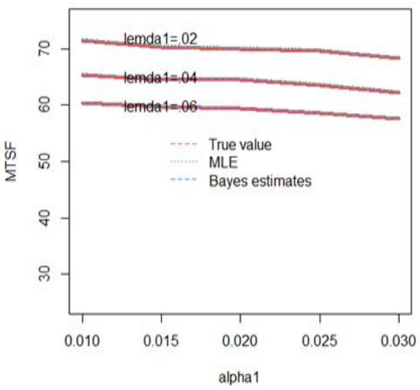


Fig.2: *Plot of MTSF for fixed beta=.04*

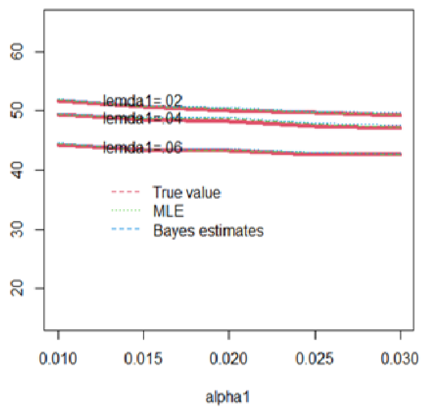


Fig.3: *Plot of MTSF for fixed beta=.09*

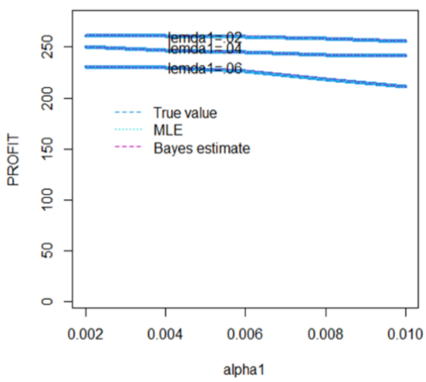


Fig.4: *Plot of PROFIT for fixed beta=0.04*

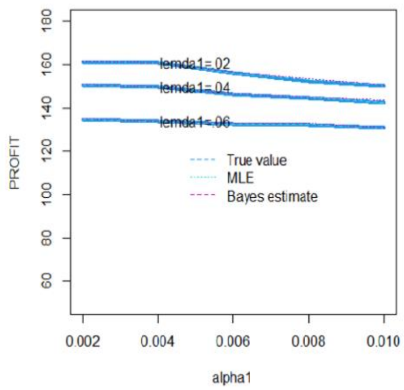


Fig.5: *Plot of PROFIT for fixed beta=0.09*

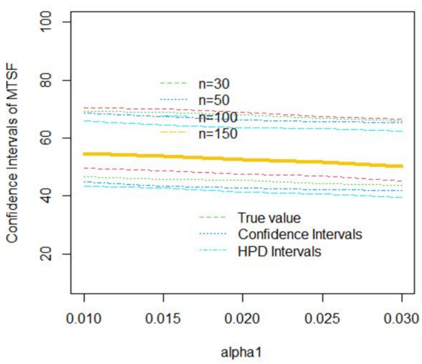


Fig.6: *Plot of CI for MTSF*

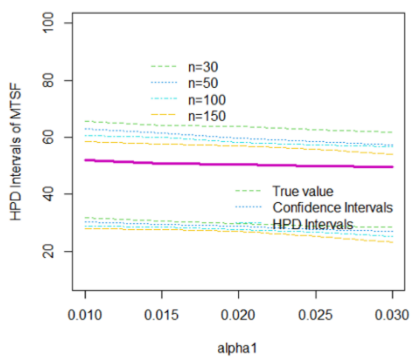


Fig.7: *Plot of HPD for MTSF*

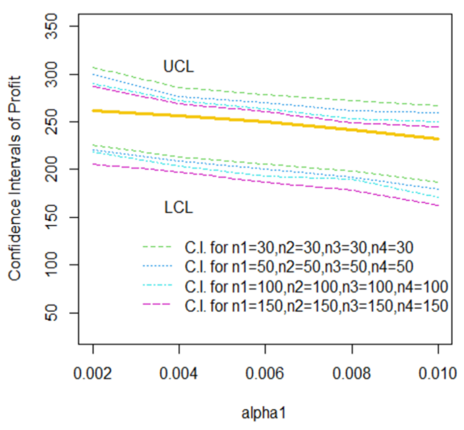


Fig.8: *Plot of CI for Profit*

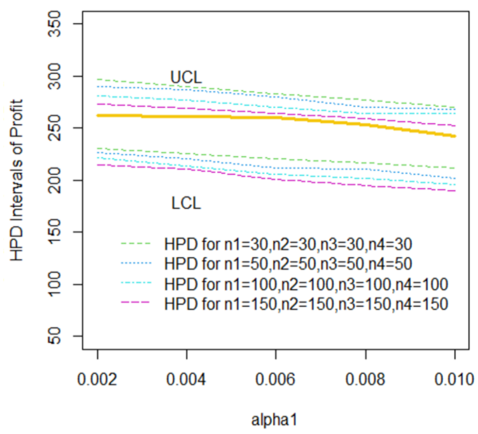


Fig.9: *Plot of HPD Intervals for Profit*

Tables

Table 1. *Various estimates of MTSF for fixed β=0.04 and varying α_1*

α_1	0.01	0.015	0.02	0.025	0.03
True Value	71.46	70.29	70.02	69.67	68.26
ML Estimates	71.58	70.45	70.1	69.74	68.3
Bayes Estimates	71.69	70.66	70.25	69.98	68.49
C.I. Width	22.25	21.14	20.24	19.89	17.04
HPD Width	21.76	20.53	20.07	19.23	16.65
MSE	0.5427	0.4284	0.3721	0.3486	0.2453

Table 2. *Various estimates of MTSF for fixed β=0.09 and varying α_1*

α_1	0.01	0.015	0.02	0.025	0.03
True Value	51.75	50.64	50.02	49.67	49.26
ML Estimates	51.89	50.76	50.3	49.74	49.4
Bayes Estimates	51.95	50.87	50.53	49.8	49.76
C.I. Width	21.56	21.15	20.47	19.43	18.51
HPD Width	20.43	19.65	19.54	18.08	17.22
MSE	0.4329	0.3897	0.3237	0.2983	0.1984

Table 3. *Various estimate of Profit for fixed β=0.04 and varying α_1*

α_1	0.002	0.004	0.006	0.008	0.01
True Value	261.39	260.74	260.19	258.27	256.11
ML Estimates	261.75	260.82	260.52	258.54	256.54
Bayes Estimates	261.83	260.95	260.67	258.67	256.35
C.I. Width	50.05	43.86	42.84	41.34	40.67
HPD Width	44.85	42.53	41.79	39.05	38.49
MSE	1.5678	1.5232	1.5137	1.5108	1.5099

Table 4. *Various estimate of Profit for fixed β=0.09 and varying α_1*

α_1	0.002	0.004	0.006	0.008	0.01
True Value	161.47	160.74	156.19	152.27	150.11
ML Estimates	161.8	160.89	156.52	152.23	150.56
Bayes Estimates	161.93	160.97	156.84	153.56	150.75
C.I. Width	50.05	43.86	42.84	41.34	40.67
HPD Width	52.85	51.53	50.27	49.05	48.82
MSE	1.2784	1.2643	1.262	1.2539	1.2487

9. Conclusion

This study report investigates the current work's value in several dimensions. First, this work investigates the steady-state reliability measures MTSF and Profit function of the subjected system, which are produced under failure time and repairman rest time distribution of the system, making the study applicable to a wide range of real-world circumstances. Second, the system's unknown parameters and reliability measures were estimated using the ML technique, along with their appropriate asymptotic confidence ranges. In addition, the Bayesian technique to estimate allows practitioners to use any prior information for better results, which is immediately seen when gamma priors are applied. Third, the supposed system's behaviour is evaluated using the Monte Carlo simulation approach to gain a better understanding of the system's behaviour. As a result, this study proves that the Bayesian method with appropriate prior is extremely utilitarian and simple to implement for analysing the redundant repairable system waiting for a repair facility.

References

[1] Chandra, A., Naithani, A., Gupta, S., and Jaggi, C.K. (2020). Reliability and cost analysis comparison between two-unit parallel systems with non-identical and identical consumable units. Journal of Critical Reviews, 7(7):695-702.

[2] Chaudhary, A., Jaiswal, S., and Gupta, R. (2015). A two unit parallel load sharing system with administrative delay in repair. Journal of Reliability and statistical studies, 8(1):1-13.

[3] Chen, M.H. and Q.M. Shao (1999), Monte Carlo estimation of Bayesian Credible and HPD intervals. J. Comput. Graph. Statist., 8:69-92.

[4] Goel, L.R., Gupta ,R. and Shrivastava, P.(1990). Profit analysis of a two unit cold standby system with varying physical conditions of the repairman. Microelectronics Reliability, 30(4).

[5] Gupta, R., Chaudhary, A., Jaiswal, S., and Singh, B. (2018). Classical and Bayesian analysis of a two identical unit standby system with fault detection, Minor and Major repair under Geometric distribution. Int. J. Agricult. Stat. Sci., 14(2): 791-801.

[6] Gupta, R., Kujal, S., and Mahi, M. (2014). A two dissimilar unit parallel system with two phase repair by skilled and ordinary repairman. Int .J.of S.A.E&M., 5(4):554-561.

[7] Gupta, R., Bansal, S., and Goel, L.R. (1990). Profit analysis of two unit priority standby system with rest period of the operator. Microelectronic Reliability, 30(4): 649-654.

[8] Kishan, R., & Jain, D. (2014). Classical and Bayesian analysis of reliability characteristics of a two unit parallel system with Weibull failure and repair laws. Int.J.S.A.E&M., 5, 252-261.

[9] Murari, K., & Muruthachalom. (1981). Two unit parallel system with period of working and rest. IEEE Transactions on Reliab., 30(91).

[10] Sherbeny-El, Mohamed S., and Hussien, Z.M. (2022). Reliability Analysis of a Two Non identical Unit Parallel System with Optional Vacations under Poisson Shocks. Mathematical Problems in Engineering, https://doi.org/10.1155/2022/24881.

PROBABILISTIC ANALYSIS OF A TWO UNIT COLD STANDBY SYSTEM WITH REPAIR AND REPLACEMENT POLICIES

Alka Chaudhary[1], Suman Jaiswal[2] and Nidhi Sharma

•

Department of Statistics
Meerut College, Meerut-250001
Ch. Charan Singh University, Meerut-250004 (India)
[1]E-mail ID: alka_813@yahoo.com,
[2]E-mail ID: jaiswalsuman85@gmail.com,
E-mail ID: 18nidhi94@gmail.com

Abstract

The present paper deals with two identical units, one is operative and the other of which other is kept on cold standby. If the operative unit fails, it goes under repair and after repair, it is not considered as good as new. If the unit fails after the first repair, it is replaced with a new unit. A single repairman is always available with the system to repair a failed unit. Failure time, repair time and replacement time distributions are taken as exponential to reduce the complexity of the system model. By using the regenerative point technique, the various important measures of system effectiveness have been obtained and are shown with the help of graph.

Keywords: Cold standby, replacement policy, MTSF, regenerative point, availability, transition probability and mean sojourn time.

1. Introduction

Various researchers have widely used the redundancy technique in systems of identical units to improve system reliability. Since the demand for improving system reliability is increasing, the field of research in reliability theory is becoming more advanced. Many researchers in the field of reliability theory, including [2, 5], have analyzed the model of a two-unit cold standby system. The performance measure of the model of a two-unit system with repair and replacement policies has been studied by various authors [1, 3, 4, 6, 7].

The current paper deals with the study of a probabilistic analysis of a two-unit cold standby system with repair and replacement policies. One unit is operational, while the other is kept on cold standby. Whenever the operative unit fails, it goes into repair. A repaired unit is not considered as good as new. If the repaired unit fails after the first repair, a new unit is installed to replace it. For the purpose of repairing the failed units, a single repairman is always ready with the system. To reduce the complexity of the system, the exponential form of failure time, repair time, and replacement time is taken. After using the regenerative point technique, the following measures of system effectiveness are obtained:

- Transition probabilities and mean sojourn times in various states of the system.

- Reliability and mean time to system failure (MTSF).
- Point-wise and steady-state availabilities of the system as well as expected up time of the system during time interval (0, t).
- The expected busy period of repairman during time interval (0, t).
- Net expected profit earned by the system in time interval (0, t) and in steady-state.

2. Model description and assumptions

- The system consists of two identical units, one operational (o) and the other in standby (s).
- A single repairman is always available with the system to repair and replace a failed unit.
- After repair, it is not considered as good as new.
- If the unit fails after the first repair, replace it with a new one.
- The failures of the units are independent and the failure time, repair time and replacement time distributions of the units are as taken exponentials.

3. Notations and states of the system

3.1. Notations

E : Set of regenerative states $=\{S_0 \text{ to } S_3\}$.

θ : Constant failure rate of operative unit.

β : Constant failure rate after one time repair of a unit.

α : Repair rate of a unit.

η : Replacement rate of failed unit.

$q_{ij}(\bullet)$: p.d.f. of transition time from regenerative state S_i to S_j.

$q_{ij}^{(k)}(\bullet)$: p.d.f. of transition time from regenerative state S_i to S_j via non-regenerative state S_k.

~ : Symbol for Laplace-stieltjes transforms i.e.

$\tilde{A}(s)=\int e^{-st}dA(t)$

* : Symbol for Laplace-transform i.e.

$A^*(s)=\int e^{-st}A(t)dt$

© : Symbol for ordinary convolution i.e.

$A(t)©B(t)=\int_0^t A(u)B(t-u)du$

Here, $\beta>\theta$, Also the limits of the integration are not mentioned whenever they are 0 to ∞.

3.2. Symbols for the states of the systems

N_o, N_s : Unit-1 and Unit-2 in operative and in standby mode.

$\overline{N}s, N_O$: Unit-1 in standby mode and Unit-2 is waiting for repair.

$\overline{N}o, F_r$: Unit-1 and Unit-2 in operative and standby mode.

F_w, F_{wr} : Unit-1 in repair mode and Unit-2 is waiting for repair.

F_R , F_{wr} **:** Replaced unit-1 after repair and unit-2 waiting for repair.

Considering the above symbols and keeping in view the assumptions stated in section-2, the possible states of the system model are shown in transition diagram (Figure.1).

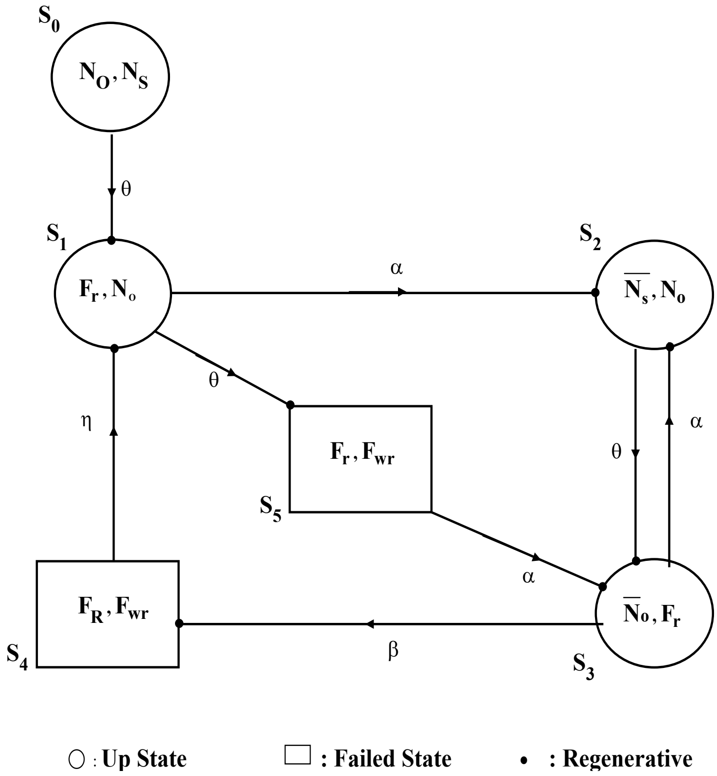


Figure.1: *Transition diagram*

4. Transition probabilities

The direct or one-step steady state transition probabilities can be obtained as follows:

$$p_{ij} = \lim_{t\to\infty} Q_{ij}(t)$$

Similarly,

$$p_{01} = \int \theta e^{-\theta t}\,dt, \qquad p_{12} = \int \alpha e^{-(\alpha+\theta)t}\,dt = \frac{\alpha}{\alpha+\theta}$$

$$p_{15} = \int \theta e^{-(\alpha+\theta)t}\,dt = \frac{\theta}{\alpha+\theta}, \qquad p_{01} = \int \theta e^{-\theta t}\,dt = 1$$

$$p_{32} = \int \alpha e^{-(\alpha+\beta)t}\,dt = \frac{\alpha}{\alpha+\beta}, \qquad p_{34} = \int \beta e^{-(\alpha+\beta)t}\,dt = \frac{\beta}{\alpha+\beta}$$

$$p_{41} = \int \eta e^{-\eta t}\,dt = 1, \qquad p_{53} = \int \alpha e^{-\alpha t}\,dt = 1$$

It can be easily verified that,

$$p_{01} = p_{23} = p_{41} = p_{53} = 1, \qquad p_{12} + p_{15} = 1, \qquad p_{32} + p_{34} = 1 \tag{1-3}$$

5. Mean sojourn times

The mean sojourn time ψ_i in state S_i is defined as the expected time taken by the system in state S_i before making the transition into any other state. If T_i denotes the sojourn time in state S_i, then mean sojourn time in state S_i is;

$$\psi_i = \int P\left(T_i > t\right)dt \tag{4}$$

Therefore,

$$\psi_0 = \int e^{-\theta t}dt = \frac{1}{\theta} = \psi_2$$
$$\psi_1 = \int e^{-(\alpha+\theta)t}dt = \frac{1}{\alpha+\theta}$$
$$\psi_3 = \int e^{-(\alpha+\beta)t}dt = \frac{1}{\alpha+\beta}$$
$$\psi_4 = \int e^{-\eta t}dt = \frac{1}{\eta}$$
$$\psi_5 = \int e^{-\alpha t}dt = \frac{1}{\alpha} \tag{5-9}$$

6. Analysis of characteristics

6.1 Reliability of the system and MTSF

Let $R_i(t)$ be the probability that the system is operative during (0, t) given that at t = 0, it starts from state $S_i \in E$..To obtain it, we assume the failed states S_4 and S_5 as absorbing. Now using simple probabilistic arguments we have the following recurrence relations in $R_i(t)$; i =0, 1,2,3.

$$R_0(t) = Z_0(t) + q_{01}(t)©R_1(t)$$
$$R_1(t) = Z_1(t) + q_{12}(t)©R_2(t)$$
$$R_2(t) = Z_2(t) + q_{23}(t)©R_3(t)$$
$$R_3(t) = Z_3(t) + q_{32}(t)©R_2(t) \tag{10-13}$$

Where, $Z_0 = e^{-\theta t}$, $Z_1 = e^{-(\alpha+\theta)t}$, $Z_2 = e^{-\theta t}$, $Z_3 = e^{-(\alpha+\beta)t}$

Taking Laplace transforms of relations (10-13) and simplifying the resulting set of algebraic equations for $R_0^*(s)$, we get;

$$R_0^*(s) = \frac{N_1(s)}{D_1(s)}$$

Where,

$$N_1(s) = (1-q_{23}^*q_{32}^*)Z_0^* + q_{01}^*(1-q_{23}^*q_{32}^*)Z_1^* + q_{01}^*q_{12}^*Z_2^* + q_{01}^*q_{12}^*q_{23}^*Z_3^*$$
$$D_1(s) = 1 - q_{23}^*q_{32}^* \tag{14}$$

Here, also for brevity we have omitted the argument's' from $q_{ij}^{*}(s), Z_0^{*}(s)$.

The expression of mean time to system failure is given by,

$$E(T_0) = \lim_{s\to 0} R_0^{*}(s)$$

Observing $q_{ij}^{*}(0) = p_{ij}, Z_0^{*}(0) = \psi_0$,

We have,

$$E[T_0] = \frac{p_{34}(\psi_0 + \psi_1) + p_{12}(\psi_2 + \psi_3)}{p_{34}} \quad (15)$$

6.2. Availability analysis

Let $A_i(t)$ be the probability that the system is up (operative) at epoch t, when system initially starts from state $S_i \in E$. Using the basic probabilistic concepts in regenerative point technique as in case of reliability, one can obtain the recurrence relations for $A_i(t)$; i=0, 1, ---, 5. Taking the Laplace Transformations and solving the resulting set of algebraic equations for $A_0^{*}(s)$, we get,

$$A_0^{*}(s) = \frac{N_2(s)}{D_2(s)} \quad (16)$$

Where,

$$N_2(s) = (1 - q_{23}^{*}q_{32}^{*} - q_{41}^{*}q_{12}^{*}q_{23}^{*}q_{34}^{*} - q_{41}^{*}q_{15}^{*}q_{53}^{*}q_{34}^{*})Z_0^{*} + (q_{01}^{*} - q_{01}^{*}q_{32}^{*}q_{23}^{*})Z_1^{*}$$
$$+ (q_{01}^{*}q_{12}^{*} + q_{01}^{*}q_{15}^{*}q_{53}^{*}q_{32}^{*})Z_2^{*} + (q_{01}^{*}q_{12}^{*}q_{23}^{*} + q_{01}^{*}q_{15}^{*}q_{53}^{*})Z_3^{*}$$

and

$$D_2(s) = 1 - q_{23}^{*}q_{32}^{*} - q_{12}^{*}q_{23}^{*}q_{34}^{*}q_{41}^{*} - q_{15}^{*}q_{53}^{*}q_{34}^{*}q_{41}^{*}$$

The steady state availability of the system is given by,

$$A_0 = \lim_{s\to 0} sA_0^{*}(s) = \lim_{s\to 0} s\frac{N_2(s)}{D_2(s)}$$

As $D_2(0) = 0$, so by using L. Hospitals rule, we get

$$A_0 = \frac{N_2(0)}{D_2'(0)} \quad (17)$$

Where,

$$N_2(0) = p_{34}\psi_1 + (p_{12} + p_{15}p_{32})\psi_2 + \psi_3$$

and

$$D_2'(0) = p_{34}(\psi_1 + \psi_4) + (p_{12} + p_{15}p_{32})\psi_2 + \psi_3 + p_{15}p_{34}\psi_5 \tag{18}$$

The expected up time of the system in interval (0, t) are given by;

$$\mu_{up}(t) = \int_0^t A_0(u)du$$

So that

$$\mu_{up}^*(s) = \frac{A_0^*(s)}{s} \tag{19}$$

6.3. Busy period analysis

Let $B_i^1(t)$ and $B_i^2(t)$ be the respective probabilities that the unit is under repair and under replacement at time t due to the failure unit, when the system initially starts from regenerative states $S_i \in E$. Using the probabilistic arguments as in case of availability, we develop the recurrence relations for $B_i^1(t)$ and $B_i^2(t)$; i=0,1,2,3,4,5. Then, taking the Laplace Transforms of these recurrence relations and solving the resulting algebraic equations for $B_0^{1*}(s)$ and $B_0^{2*}(s)$ we get;

$$B_0^{1*}(s) = \frac{N_3(s)}{D_2(s)} \text{ and } B_0^{2*}(s) = \frac{N_4(s)}{D_2(s)} \tag{20-21}$$

Where,

$$N_3(s) = (q_{01}^* - q_{01}^* q_{32}^* q_{23}^*)Z_1^* + (q_{01}^* q_{12}^* q_{23}^* + q_{01}^* q_{53}^* q_{15}^*)Z_3^* \tag{22}$$

$$N_4(s) = (q_{01}^* q_{12}^* q_{23}^* q_{34}^* + q_{01}^* q_{53}^* q_{15}^* q_{34}^*)Z_4^* \tag{23}$$

In the long run, the probabilities that the repairman will be busy are respectively given by:-

$$B_0^1 = \frac{N_3(0)}{D_2'(0)} \text{ and } B_0^2 = \frac{N_4(0)}{D_2'(0)} \tag{24}$$

Where,

$$N_3(0) = p_{34}\psi_1 + \psi_3$$

$$N_4(0) = p_{34}\psi_1$$

The value of $D_2'(0)$ is same as given by expression (18).

The expected busy period of skilled repairman and regular repairman during (0,t) are given by,

$$\mu_b^1(t) = \int_0^t B_0^1(u)du \text{ and } \mu_b^2(t) = \int_0^t B_0^2(u)du$$

So that

$$\mu_b^{1*}(s) = \frac{B_0^{1*}(s)}{s} \text{ and } \mu_b^{2}(s) = \frac{B_0^{2*}(s)}{s} \qquad (25\text{-}26)$$

7. Profit function analysis

We are now in the position to obtain the net expected profit incurred during time (0, t) by considering the characteristics obtained in earlier sections.
Let us consider,

K_0 = revenue per unit time by the system when it is operative.

K_1 = cost per unit time when repairman is busy in repair of failed unit.

K_2 = cost per unit time when repairman is busy in replacement of failed unit.

Then, the net expected profit incurred during time (0, t),

$$P(t) = \text{Expected total revenue in } (0,t) - \text{Expected amount spent on repair in } (0,t)$$
$$= K_0\mu_{up}(t) - K_1\mu_b^1(t) - K_2\mu_b^2(t)$$

The expected profit per-unit time in steady state is given by,

$$P = K_0A_0 - K_1B_0^1 - K_2B_0^2$$

8. Graphical study of system behaviour

The Behavioral characteristics of MTSF and Profit function for different values of Failure rate θ shown in figures 2 and 3 respectively.

In figure 2, the graphical analysis of MTSF in relation to failure rate θ for three different values of repair rate α i.e. 0.09, 0.35 and 0.85 and for two fixed values of β at 0.05 and 0.09 has been plotted.

The graphical analysis of Profit function in relation to failure rate θ for different values of repair rate α i.e. 0.35, 0.55 and 0.95 and for fixed values of β at 0.32 and 0.45 has been shown in figure 3.

In the analysis at figure 2 and 3, the MTSF and Profit function shows a decline as the value of failure rate increases and when the value of repair rate increases, the graph shows the increase in the value of MTSF and Profit function.

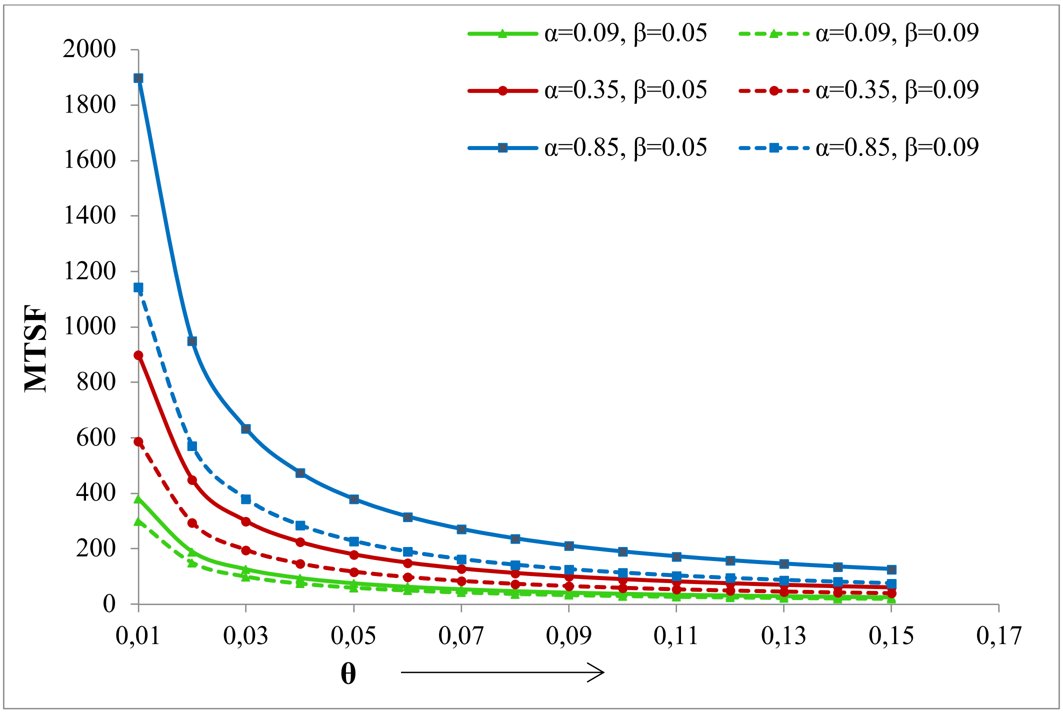


Figure.2: *Behavior of MTSF with respect to* α *,* β *and* θ

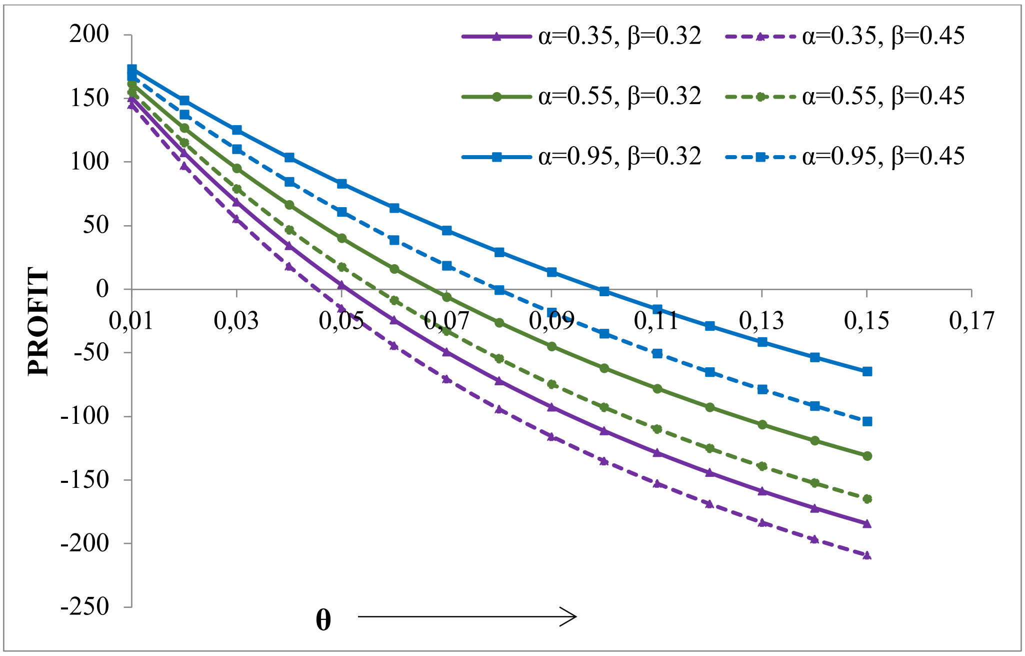


Figure.3: *Behavior of Profit with respect to* α *,* β *and* θ

According to the dotted curves in the MTSF graph in relation to failure rate, in order to achieve MTSF for at least 300 units, the failure rate θ must be less than 0.01, 0.02, and 0.04 for values of α 0.09, 0.35, and 0.85 when β is fixed at 0.09.

Similarly, for smooth curves the upper bounds for failure rate (θ) 0.012, 0.031 and 0.081 corresponding to α i.e., 0.09, 0.35and 0.85 when β is fixed at 0.05.

Figure 3 shows that the system is profitable only when the failure rate θ is less than 0.045, 0.058, and 0.078, respectively, for values of α 0.35, 0.55, and 0.95 when β fixed at 0.45.

Similarly, we conclude from smooth curves that the system is profitable only when failure rate (θ) is less than 0.05, 0.067 and 0.098 respectively for values of repair rate α i.e., 0.35, 0.55and 0.95 when β is fixed at 0.032.

References

[1] Chander, S. (2007). MTSF and profit evaluation of an electric transformer with inspection for online repair and replacement. *Journal of Indian Soc. Stat. Operation Research*, 28(1-4):33-43.

[2] Gupta, R. and Tyagi, A. (2014). A two identical unit cold standby system with switching device and geometric failure and repair time distributions. *Aligarh Journal of Statistics*, 34: 55-66.

[3] Gupta, U. and Kumar, P. (2019). Performance measures of a two non-identical unit system model with repair and replacement policies. *RT & A*, 14(1):53-62.

[4] Kumar, P., Bharti, A. and Gupta, A. (2012). Reliability analysis of a two non-identical unit system with repair and replacement having correlated failures and repairs. *Journal of Informatics and Mathematical Sciences*, 4(3):339-350.

[5] Mahmoud, M.A.W and Moshref, M.E. (2010). On a two-unit cold standby system considering hardware, human error failures and preventive maintenance. *Mathematical and Computer Modeling*, 51(5-6):736-745.

[6] Sharma, N., joorel, J. P. S. (2019). Study of stochastic model of a two unit system with inspection and replacement under multi failure. *RT & A*, 14(3):31-38.

[7] Singh, M. and Chander, S. (2005). Stochastic analysis of reliability models of an electric transformer and generator with priority and replacement. *Jr. Decision and Mathematical Sciences*, 10(1-3):79-100.

DESIGN OF INERTIAL DELAY OBSERVER BASED MODEL FOLLOWING DYNAMIC SLIDING MODE CONTROL

S. S. Nerkar

•

Department of Instrumentation Engineering,
Government College of Engineering, Jalgaon,
Maharashtra, India.
sachinnerkar@rediffmail.com

•

B. M. Patre

Department of Instrumentation Engineering,
S.G.G.S. Institute of Engineering and Technology,
Nanded, Maharashtra, India.
bmpatre@yahoo.com

Abstract

This paper proposes the design and implementation of Inertial delay observer (IDO) based model following dynamic sliding mode control (DSMC). The Inertial delay observer estimates the states as well as the uncertainties and disturbances in an integrated manner. The DSMC is provides smooth control signal with the mechanism of chattering elimination while maintaining the accuracy of control. The efficacy of the proposed technique is demonstrated with numerical simulation of uncertain second order system. The observer based model following DSMC technique is also validated through experimentation on Quanser DC servo motor. Results show the effectiveness of the combination of the controller-observer design for position control of DC motor against uncertainties and sensor noise. The technique is robust due to appropriate estimation and follows the model precisely which improves overall life of the system. The stability of the designed observer based control scheme is provided by Lyapunov theory.

Keywords: Dynamic sliding mode control, Inertial delay observer, DC servo motor, Lyapunov theory

1. Introduction

The purpose of design of model following control is to generate control such that the controlled system behaves like a model, which specifies the design objectives. It has the objective to minimize the error between the states of the model and the controlled plant despite presence of uncertanity, disturbances and measurement noise. The design guarantees that the error between the states of the model and the controlled plant goes to zero. Over the past decades many model following approches are designed, for controlling the output close to the model output and to satisfy the closed-loop stability as well as regulation. [1, 2, 3, 4, 5].

SMC is a robust control technique to counter the presence of uncertainties and disturbances in the system. But the main drawback of the SMC is that it uses the discontinuous control to achieve the control objective. Thus chattering in the SMC restricts it for the real life applications. SMC requires that, the full state vector to be available for the control to apply effectively. But states may not be available always. The range of uncertainty, if it can not be determined or not known exactly, sliding condition may not be satisfied. Various methods have been proposed to eliminate the chattering like continuous approximation by boundary layer technique[6], to use higher order sliding mode control (HOSM) [7]. Dynamic sliding mode control (DSMC) is also designed for chattering elimination, where the control developed by sliding mode is filtered and then applied to the actual plant [8].

The DSMC adopts a special control structure, in which a integrator as a filter is placed in front of the system as depicted in figure(1). A sliding mode control w is designed for the augmented system consist of system and the filter. Being a sliding mode control, the auxiliary control signal w has chattering, however the actual control signal u, applied to the system is smooth. Here, a low pass filter eliminates the chattering along with maintaining the control accuracy. In case of a system where measurement noise is present DSMC effectively filter out the chattering due to it. Hence, the mechanism of chattering elimination but with maintaining the accuracy of control, both are decoupled in the design of DSMC. [9, 10, 11].

In the design of DSMC, the main problem is to form the sliding surface, as we are designing the SMC for the augmented system, sliding surface should be the function of the states of the augmented system not the actual system. Since the augmented system is one dimension larger than the original system, the new sliding variable in DSMC contains an uncertainty term due to the external disturbance and/or parametric uncertainty. And also from the structure of DSMC one can observe that the lumped uncertainty involves in the augmented system. Evaluation of the new sliding variable in DSMC becomes difficult because of this reason. To over come the problem of bounds of uncertainty, instead of using the bounds of uncertainty in the control one can estimate the uncertainty and disturbance in the system by using the methods like time delay control(TDC), inertial delay control(IDC) and can use it in the control design. Hence, there is a need to design an observer to estimate the lumped uncertainty for the design of sliding surface in DSMC. Different types of state observers are developed to estimate the states as well as uncertainties in the systems like uncertanity and disturbance estimator(UDE), supertwisting disturbance observer(STDO), inertial delay observer(IDO), extended state observer(ESO) etc. [12, 13, 14, 15].

In this paper inertial delay observer (IDO) based model following control is designed. Here the states estimation is achieved by linear observer while uncertainties are estimated by inertial delay method. The proposed strategy estimatate the states as well as the lumped uncertainty in the system. This state and uncertainty observers are more useful for the control design in real time applications where it is difficult to measure all the states using the sensors. The control strategy is verified by the practical experimentation on Quanser's servo plant SRV02 under influence of sensor measurement noise and uncertanities in terms of unmodeled dynamics. The main contributions of the work are summarized as follows:

- Design of model following DSMC by state estimation using IDO for achieving smooth control.
- The effectiveness of the designed controller is verified by simulation of uncertain plant.
- Due to simultaneous estimation of states and uncertainty, IDO based model following DSMC design reduces the need sensors.
- Experimental results validates the efficacy of robustness of control strategies against the presence of disturbance, uncertainties in the system parameters and sensor noise.

The paper is organised as follows. Sect. 2 presents the mathematical model of system. Sect. 3 presents the observer based design approach with numerical simulation. Sect. 4 contains stability analysis. Application of the proposed strategy in Sect. 3 to the system mentioned in Sect. 2 along with the experimental results and perfomance analysis is presented in Sect. 5. Finally in Sect. 6 paper is concluded.

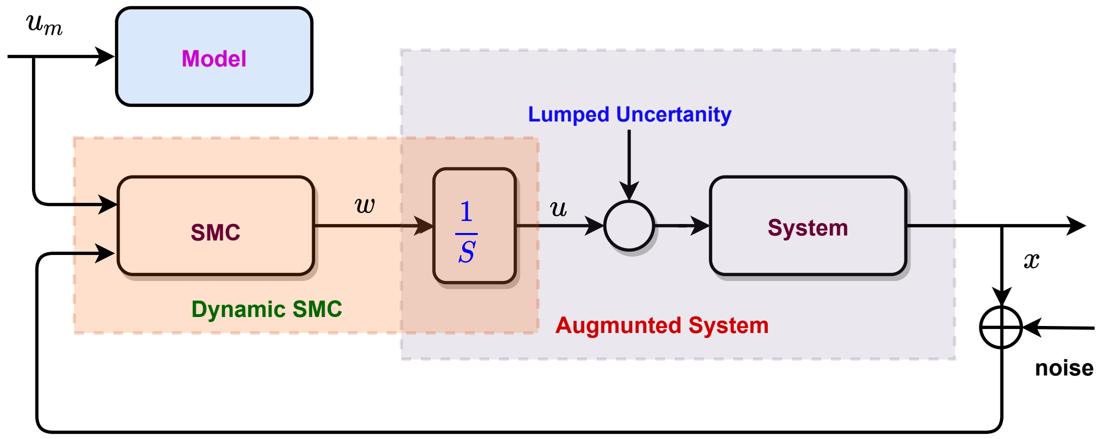


Figure 1: *Model following DSMC design with measurement noise*

2. Mathematical Modelling

Due to excellent speed control characteristics, the DC motors are used in many industrial applications. The most common device used as an actuator in mechanical control is the DC motor. In Quanser set up it is a base unit for controlling rotary inverted pendulum, double inverted pendulum, 2-dof robot etc. As a result, it becomes necessary to control the amount of electric voltage supplied to the servomotor by continuously detecting the position and speed of shaft. It has attracted researchers and considerable research has been done and several methods are proposed and implemented for it.

2.1. Transfer function representation of SRV-02

The dynamic equations and transfer function of SRV-02 servo plant using first principles is discussed in [16]. The servomotor transfer function is given by ,

$$\frac{\theta_l(s)}{V_m(s)} = \frac{K}{s(s\tau + 1)} \tag{1}$$

where

- $\theta_l(s)$ is Laplace transform of position of load,
- $V_m(s)$ is Laplace transform of input voltage,
- K is the steady state gain,
- τ is the system time constant.

2.2. State space representation of SRV-02

For implementation of the proposed control strategy to servo plant we convert transfer function into the state space form. Considering two states for state space model, θ load shaft position as first state x_1 and $\dot{\theta}$ load shaft speed as second state x_2.

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \theta(s) \\ \dot{\theta}(s) \end{bmatrix} \tag{2}$$

The equation (1) is represented in system state variable form as becomes

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -(1/\tau) \end{bmatrix} \tag{3}$$

Input matrix is form as

$$B = \begin{bmatrix} 0 \\ K/\tau \end{bmatrix} \tag{4}$$

while ouput matrix come as

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{5}$$

The output matrix is represented as,

$$y = C \begin{bmatrix} \theta(s) \\ \dot{\theta}(s) \end{bmatrix} \tag{6}$$

We are measuring the position of the load shaft x_1 The design specifications and parameters for the model are considered from [17].

3. Problem formulation

In the design of DSMC, as the lumped uncertainties are included in between the filter and the system, the sliding surface requires the unknown lumped uncertainty in its design. Hence design of sliding surface in DSMC is a tedious and requires the unknown parameter called lumped uncertainty. Therefore it is required to design an observer for estimation of lumped uncertainty for the design of sliding surface. For the systems where the states are unavailable for measurement, there is a need to design an observer which estimates the states of the systems. The main motivation for the simultaneous estimation of states and uncertainty is to reduce the number of sensors and to robustify the control in the presence disturbance. To overcome the problem of bounds of uncertainty, instead of using the bounds of uncertainty in the control one can estimate the uncertainty and disturbance in the system by using IDO and can use it in the control design. The IDO based model following DSMC is designed for experimentation as follows.

3.1. Design of inertial delay observer based model following dynamic sliding mode control

In the real time applications, it is complicated to measure all the states except the output. Here the design of model following DSMC by estimating the system states is proposed. Since the design of sliding surface requires the lumped uncertainty in the system, and here in addition to that system states are also required. Hence, IDO is proposed for estimation of the system states as well as the uncertainty.
Consider the DSMC design for linear time invariant uncertain system with relative degree two

$$\begin{aligned} \dot{x} &= Ax + Bu + Be \\ y &= Cx \end{aligned}$$

The model to be followed is,

$$\begin{aligned} \dot{x}_m &= A_m x_m + B_m u_m \\ y_m &= C x_m \end{aligned}$$

The objective is to design control u so as force the system to follow the sliding dispite of the parameter variations. The design of DSMC by estimating the system states where all states other than output are not available.

$$\begin{aligned} &\dot{x}(t) = Ax(t) + Bu(t) + Be(x, u, t) \\ &y(t) = Cx(t) \end{aligned} \tag{7}$$

Where, $x \in \Re^n$ is the state vector, $u \in \Re^m$ is the control input and $y \in \Re^p$ is the output of the system. $e(x, u, t)$ is lumped uncertainties which includes uncertainty in A matrix, B matrix and

disturbances. The output and input of the system are measurable (available) for all the time $t = 0$. The current state and initial state $x(0)$ are supposed to be non available. $d(x, t)$ is external unmeasurable disturbances.
Assumption:

- Uncertainties of the system and input matrices are lumped into the $e(x, u, t)$
- Lumped uncertainty varies slowly $\dot{e}(x, u, t) \cong 0$.
- The model parameters are completely known and no uncertainty or disturbances in the model
- The model matrices A_m and B_m are stable.

Inertial delay observer for the estimation of states of the system is given as,

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + B\hat{e}(x, u, t) + L(y(t) - \hat{y}(t)) \\ \hat{y}(t) &= C\hat{x}(t) \end{aligned} \tag{8}$$

Where, L is observer gain matrix, $\hat{e}(x, u, t)$ is estimation of lumped uncertainty $\hat{x}(t)$, $\hat{y}(t)$ are states and output of observer respectively
From eqn.(7), we get,

$$Be(x, u, t) = \dot{x}(t) - Ax(t) - Bu(t) \tag{9}$$

Defining the pseudo inverse of matrix B as,
$B^+ = (B^T B)^{-1} B^T$, then the equation (7) is,

$$e(x, u, t) = B^+[\dot{x}(t) - Ax(t) - Bu(t)] \tag{10}$$

If all the states are available for measurement then one can directly use above eqn (10) for estmation but, all the states are not available so replacing $x(t)$ with $\hat{x}(t)$ in eqn. (10) which give us,

$$\tilde{e}(x, u, t) = B^+[\dot{\hat{x}}(t) - A\hat{x}(t) - Bu(t)] \tag{11}$$

uncertanity can be estimated as,

$$\hat{e}(x, u, t) = G_f(s)\ e(x, u, t) \tag{12}$$

$$G_f(s) = \frac{1}{\tau s + 1} \tag{13}$$

where $G_f(s)$ is first order filter with high bandwidth. Note that, for sake of simplicity, the laplace domain notations are used to represent $G_f(s)$ in the time domain equations.

$$\hat{e}(x, u, t) = B^+[\dot{x}(t) - Ax(t) - Bu(t)]G_f(s) \tag{14}$$

Using equations (8) and (11) and simplifying

$$\begin{aligned} \hat{e}(x, u, t) &= B^+[A\hat{x}(t) + Bu(t) + B\hat{e}(x, u, t) + L(y(t) - \hat{y}(t)) - A\hat{x}(t) - Bu(t)]G_f(s) \\ &= B^+ LC\tilde{x}(t)\frac{G_f(s))}{1 - G_f(s)} \\ &= \frac{B^+ LC\tilde{x}(t)}{\tau s} \end{aligned} \tag{15}$$

$$\dot{\hat{e}}(x, u, t)) = -\frac{1}{\tau} B^+ LC\tilde{x}(t) \tag{16}$$

Observer error dynamics is derived as,

$$\dot{\tilde{e}}(x,u,t)) = -\frac{1}{\tau}B^{+}LC\tilde{x}(t) - \dot{\hat{e}}(x,u,t) \tag{17}$$

We assumed that the uncertainties varies slowly,
$\dot{\tilde{e}}(x,u,t)) \cong 0$, considering it equation(17) becomes,

$$\dot{\tilde{e}}(x,u,t) = -\frac{1}{\tau}B^{+}L\tilde{x}(t) \tag{18}$$

In combined form we can write observer dynamics (as matrix form)

$$\begin{bmatrix} \dot{\tilde{x}}(t) \\ \dot{\tilde{e}}(x,u,t) \end{bmatrix} = \begin{bmatrix} A-LC & B \\ -\frac{1}{\tau}B^{+}LC & 0 \end{bmatrix} \cdot \begin{bmatrix} \tilde{x}(t) \\ \tilde{e}(x,u,t) \end{bmatrix} \tag{19}$$

Converting this matrix form into the form which is similar to the pole placement problem. Using closed loop poles of overall system at desired location,

$$H = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}, G = \begin{bmatrix} C & 0 \end{bmatrix} \text{ and } K_{ob} = \begin{bmatrix} L \\ \frac{1}{\tau}B^{+}L \end{bmatrix}$$

The eigen values of the matrix $A - K_{ob}G$ can be made arbitrary by appropriate choice of the observer gain K_{ob}, when the pair (H, G) is observable.

3.2. Sliding surface

Now in model following, one can consider the sliding surface as the error between the model and actual system. So that when ever sliding surface becomes zero the system should follow the model. Taking this idea as a backend sliding surface is chosen as

$$\sigma = y_e^{''} + \lambda_1 y_e^{'} + \lambda_0 y_e \tag{20}$$

Where y_e, y_e', y_e'' are the errors of outputs and its derivatives between system and model

$$\begin{aligned} y_e &= Cx - Cx_m \\ y_e' &= CAx - CA_m x_m \\ y_e'' &= CA^2x + CABu + CABe - CA_m^2 x_m - CA_m B_m u_m \end{aligned} \tag{21}$$

From the inertial delay observer one can get the estimate of the states $\hat{x}$ and lumped uncertainty $\hat{e}$. Here SMC is designed for the augmented system, so one can choose the sliding surface for the model following as,

$$\hat{\sigma} = \hat{y_e}'' + \lambda_1 y_e' + \lambda_0 y_e \tag{22}$$

The y_e, y_e', y_e'' can be defined by using estimated states as,

$$\begin{aligned} y_e &= Cx - Cx_m \\ y_e' &= CA\hat{x} - CA_m x_m \\ y_e'' &= CA^2\hat{x} + CABu + CAB\hat{e} - CA_m^2 x_m - CA_m B_m u_m \end{aligned} \tag{23}$$

3.3. Control design

The control for the actual plant can be obtained from the auxiliary control as,

$$u = \int w \, dx \tag{24}$$

Sliding surface is designed for the model following DSMC, From the equation (22) and taking it's derivative one can get,

$$\dot{\hat{\sigma}} = CA^3\hat{x} + CA^2Bu + CABw - CA_m^3x_m - CA_m^2B_mu_m + CA_mB_m\dot{u}_m + \hat{e}^* + \lambda_1\hat{y}_e{}'' + \lambda_0 y_e' \quad (25)$$

where, $\hat{e}^* = CA^2B\hat{e} + CAB\hat{\dot{e}}$
Assume the auxiliary control w as

$$w = w_{eq} + w_n \quad (26)$$

where w_{eq} addresses the known dynamics while w_n is for the unknown dynamics. Then, the for known dynamics w_{eq} is selected as,

$$\begin{aligned} w_{eq} =& -(CAB)^{-1}(CA^3\hat{x} + CA^2Bu + CABw - CA_m^3x_m - CA_m^2B_mu_m - CA_mB_m\dot{u}_m + \hat{e}^* \\ &+ \lambda_1\hat{y}_e{}'' + \lambda_0 y_e' + K\hat{\sigma}) \end{aligned} \quad (27)$$

where k is a small positive constant.
To satsisfy sliding condition the control w_n is selected as,

$$w_n = -(CAB)^{-1}K_1 sign(\hat{\sigma}) \quad (28)$$

where, $K1$ is a positive constant $> |\tilde{e}^*|$.

3.4. Numerical simulation

To demonstrate the efficacy of the design of IDO based model following DSMC with state estimation, a second order system is considered. The system matrices are as follows,

$$A = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

The reference model to be followed is

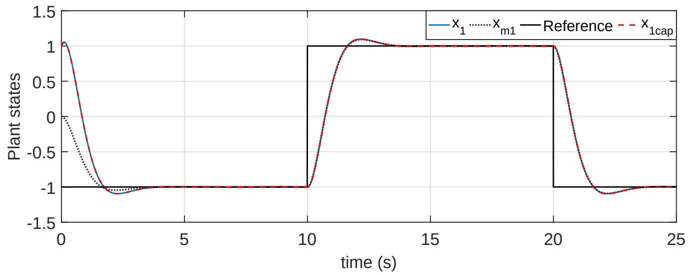
$$A_m = \begin{bmatrix} 0 & 1 \\ -4 & -2.8 \end{bmatrix}, \quad B_m = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

Initial conditions are

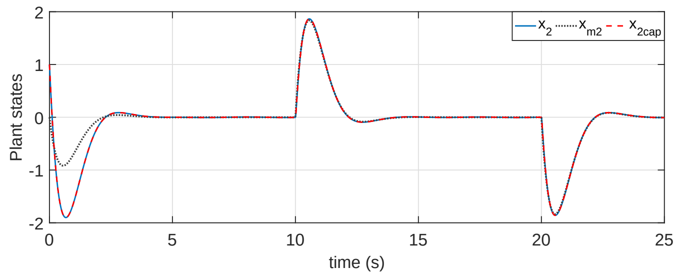
$$x(0) = \begin{bmatrix} 1 & 1 \end{bmatrix}^T, x_m(0) = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$$

With the 40% uncertainties in the system parameters and $sin(2t)$ as a external disturbance. while the low pass filter time constant $tau = 0.01$, the constant parameters are $\lambda_0 = 15$, $\lambda_1 = 5$, $K = 5$ and $K_1 = 10$. The observer pols are at $\begin{bmatrix} -5 & -9 & -11 \end{bmatrix}^T$. Apart from the disturbance here uniform measurement noise of absolute magnitude 0.01 is added with the lumped uncertainty. The reference input u_m is square wave switching at $10s$.

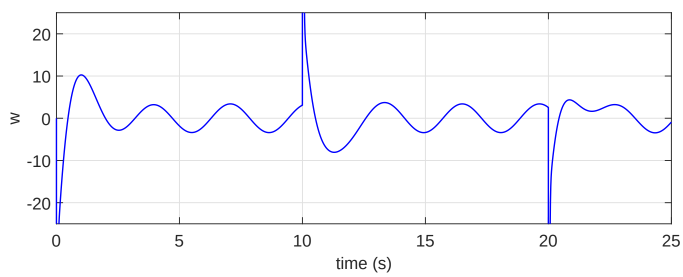
The states of the system are estimated using the IDO and accuracy of the estimation depends on the observer poles. The observer poles are selected such that, it estimates the states and the uncertainty precisely. The simulation results in Fig. (2) depicts that, system is following the model precisely despite the presence of uncertanities, disturbance and noise. The plot of plant and observer states are as shown in Fig.(2a) and (2b). The Fig. (2c) shows auxiliary control w, designed for the augmented system, while the actual control u to the system is smooth as shown in Fig.(2d). Finaly the sliding surface is in Fig.(2e).



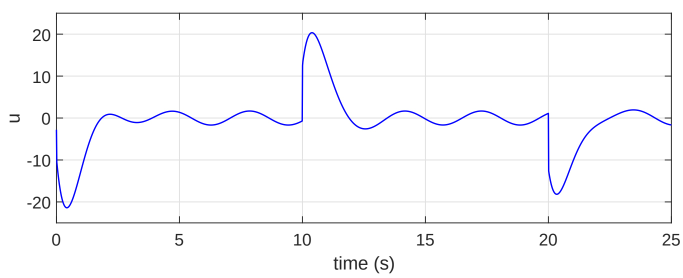
(a) *State, it's estimate and observer state* x_1



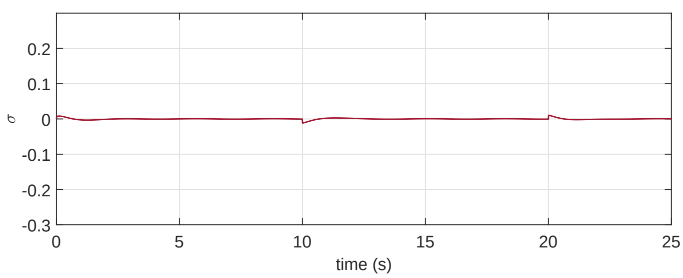
(b) *State it's estimate and observer state* x_2



(c) *Auxiallary control* w



(d) *Control* u



(e) *Sliding surface* σ

Figure 2: *IDO based model following DSMC control of a uncertain plant.*

4. Stability analysis

The model is considered to be stable by design so for the designed control, stability of the system is proved. Consider the Lyapunov function as,

$$V = \frac{1}{2}\sigma^2 \tag{29}$$

Taking derivative we get,

$$\dot{V} = \sigma\dot{\sigma} \tag{30}$$

From the equation(21)

$$\begin{aligned}\dot{\sigma} =& CA^3x + CA^2Bu + CA^2Be + CABw + CAB\dot{e} \\ &+ \lambda_1 y'' + \lambda_0 y'\end{aligned} \tag{31}$$

During sliding motion, the control term w_{eq} substituting in the equation(31) we get,

$$\begin{aligned}\dot{\sigma} =& CA^3\tilde{x} + CA^2B\tilde{e} + CAB\dot{\tilde{e}} + \lambda_1\tilde{y}'' + \lambda_0\tilde{y}' - K\hat{\sigma} \\ &+ CABw_n\end{aligned} \tag{32}$$

The observer error $\tilde{x}$ in equation (19) converges to zero asymptotically, if K_{ob} is chosen appropriatly. Then, $\dot{\sigma}$ is,

$$\dot{\sigma} = CABw_n - K\hat{\sigma} + \tilde{e}^* \tag{33}$$

So the equation (30) becomes,

$$\begin{aligned}\dot{V} =& \sigma\tilde{e}^* - K\sigma\hat{\sigma} - \sigma CAB(CAB)^{-1}K_1 sign(\hat{\sigma}) \\ \dot{V} =& \sigma\tilde{e}^* - K\sigma\hat{\sigma} - \sigma K_1\frac{|\hat{\sigma}|}{\hat{\sigma}} \\ \dot{V} \leq& |\sigma|\left(|\tilde{e}^*| - K|\hat{\sigma}| - K_1\frac{|\hat{\sigma}|}{|\hat{\sigma}|}\right) \\ \dot{V} \leq& |\sigma|\left(|\tilde{e}^*| - K|\hat{\sigma}| - K_1\right)\end{aligned} \tag{34}$$

As $K_1 > |\tilde{e}^*|$,
K_1 and K are both positive constants, $\dot{V}$ is negative definite.
Thus the sliding manifold is stable.

5. Application of IDO based model following DSMC control for position control of DC servomotor

The real time implementation of the model following DSMC strategy discussed in Sect. 3 has been done for position control of the Quanser servomotor SRV02. The block diagram of IDO based model following DSMC control of the servomotor is as shown in Fig.(3). It has potentiometer, encoder and tachometer. The potentiometer and encoder sensors measures the angular position of the load gear and tachometer is used to measure its velocity. The different parameters of DC servomotor are considered as mentioned in Quanser setup manual [17]. The constant parameters are $\lambda_0 = 15$, $\lambda_1 = 7$, $K = 1$ and $K_1 = 20$. Here uniform measurement noise as a sensor noise of absolute magnitude 0.01 is added. The observer poles are selected at $1.5 * [-5\ \ -9\ \ -11]$. The inital condition for states are considered as $[0\ \ 0]^T$ and $[1\ \ 1]^T$ respectively for the consideration.

Experimentation has done under different considerations like with and without noise for the observer based model following DSMC strategy. Fig.(4), (5) shows the system and model states where the initial mismatch is due to the initial conditions of the model and system. The experimental results have proved the efficacy of proposed design for real time position control of DC motor.

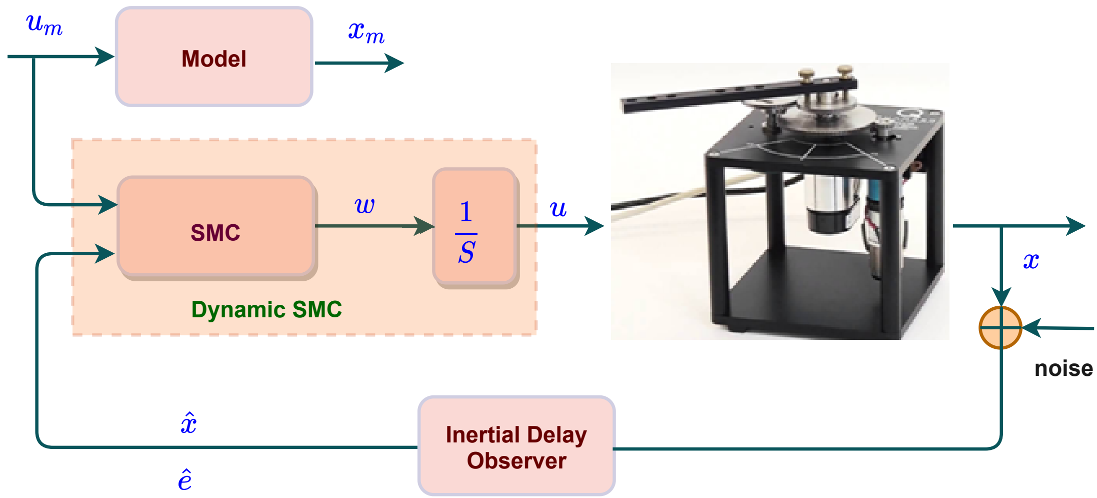


Figure 3: *IDO based model following dynamic sliding mode control of DC servomotor with measurement noise*

The effectiveness of the scheme is assessed through the results shown in Fig. (4) are taken from the encoder for output without any measuring noise. Fig.(4a) and Fig.(4b) shows the states of the model, system and observer, where one can observe that the observer has followed the model dynamics as the estimation is accurate. Fig.(4c) shows the chattering in the auxiliary control w while the actual control to the system u is smooth control shown in Fig.(4d). The sliding surface σ is shown in Fig. (4e).

IDO based DSMC design estimates the states and uncertainty simultaneously, It reduces reduces the need of sensors. In order to demonstrate robustness of the controller, the noisy output is taken from the potentiometer that is the control is tested for the system with measurement noise and uncertainty. The results are shown in Fig.(5a). As the output from the potentiometer is noisy, the second state calculated consists heavy chattering due to the addition of measurement noise as shown in Fig. (5b), model and the observer states are smooth at the same time. Due to it, the auxiliary control is affected as shown in Fig.(5c) but the actual control driving the system is smooth signal as shown in Fig.(5d). The sliding surface σ is shown in Fig. (5e). With the design of DSMC the actual system is driven by the smooth control though the states are contaminated with the measurement noise.

6. Conclusion

In this paper, performance analysis of observer based estimation based techniques has been carried out for position control of DC motor. The effectiveness of the proposed schemes has been validated in presence of uncertainty due to unmodeled dynamics and disturbance along with measurement noise. DSMC is made more robust by augmenting a observer. IDO based model following DSMC estimates the system states as well as the uncertainty. Due to the simultaneous estimation of states and uncertainty, it reduce the number of sensors and robustify the control despite unknown disturbance and uncertainties in the system parameters or load changes. The simulation and experimental results show that the proposed IDO based model following DSMC is robust and follows the model precisely. The essential boundness of lumped uncertainty estimation error and sliding manifold is demonstrated by Lyapunov theory.

References

[1] K.-K. Young, ,"Design of variable structure model-following control systems. *IEEE Transactions on Automatic Control*, vol. 23, no. 6, pp. 1079-1085, 1978.

[2] K. Furuta and K. Komiya, Design of model-following servo controller, *IEEE Transactions on Automatic Control*, vol. 27, no. 3, pp. 725-727, 1982.

[3] S. Talole and S. Phadke, Model following sliding mode control based on uncertainty and disturbance estimator, *Journal of Dynamic Systems, Measurement and Control*, vol. 130, no. 3, 2008.

[4] P. Shendge and B. Patre, Robust model following load frequency sliding mode controller based on ude and error improvement with higher order filter, *IAENG International Journal of Applied Mathematics,* vol. 37, no. 1,pp. 1-6, 2007.

[5] M.-C. Pai, Observer-based adaptive sliding mode control for robust tracking and model following, *International Journal of Control, Automation and Systems,* vol. 11, no. 2, pp. 225-232, 2013.

[6] M.S. Chen, Y.-R. Hwang, and M. Tomizuka, A state dependent boundary layer design for sliding mode control, *IEEE transactions on automatic control,* vol. 47, no. 10, pp. 1677-1681, 2002.

[7] A. Levant, Higher-order sliding modes, differentiation and output-feedback control, *International journal of Control*,vol. 76, no. 9-10, pp. 924-941, 2003.

[8] A. J. Koshkouei, K. J. Burnham, and A. S. Zinober, Dynamic sliding mode control design, *IEEE Proceedings Control Theory and Applications,* vol. 152, no. 4, pp. 392-396, 2005.

[9] M. R. Shokoohinia, M. M. Fateh, and R. Gholipour, Design of an adaptive dynamic sliding mode control approach for robotic systems via uncertainty estimators with exponential convergence rate *SN Applied Sciences,* vol. 2, no. 2, pp. 1-11, 2020.

[10] M.S. Chen, C.H. Chen, and F.Y. Yang, An ltr observer based dynamic sliding mode control for chattering reduction, *Automatica,* vol. 43, no. 6, pp. 1111-1116, 2007.

[11] H. Sira-Ramirez, On the dynamical sliding mode control of nonlinear systems, *International journal of control,* vol. 57, no. 5, pp. 1039-1061, 1993.

[12] S. Nerkar, P. Londhe, and B. Patre, Design of super twisting disturbance observer based control for autonomous underwater vehicle, *International Journal of Dynamics and Control,* pp. 1-17, 2021.

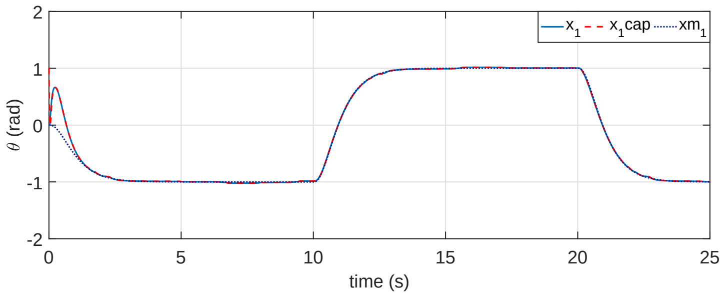
[13] A. Patil, D. Ginoya, P. Shendge, and S. Phadke, Uncertainty estimation-based approach to antilock braking systems, *IEEE Transactions on Vehicular Technology,* vol. 65, no. 3, pp. 1171-1185, 2015.

[14] D. L. Ginoya, T. R. Patel, P. Shendge, and S. Phadke, Design and hardware implementation of model following sliding mode control with inertial delay observer for uncertain systems, *International Conference on Electronics Computer Technology,* vol. 3, pp. 192-196, IEEE, 2011.

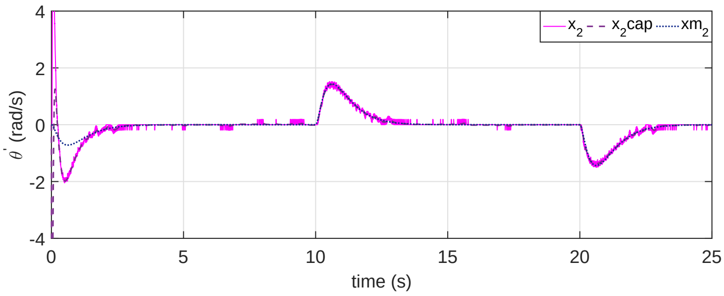
[15] S. N. Pawar, B. Patre, et al., Extended state observer based robust sliding mode control for fourth order nonlinear systems with experimental validation, *Journal of Dynamics and Control,* pp. 1-12, 2021.

[16] Introduction to quarc 2.0 & srv 02 and instructor manual, Quanser Inc., Markham, ON, Canada.

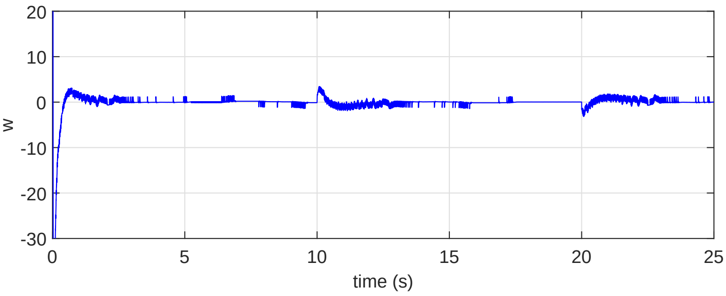
[17] Rotary servo plant srv 02 user manual, set up and configuration, Quanser Inc., Markham, Canada.



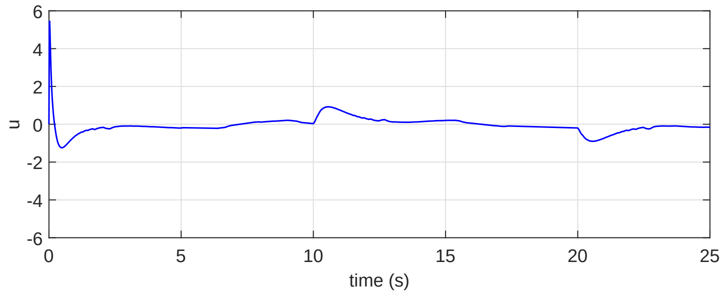
(a) *Actual position of shaft and its estmation: state* x_1



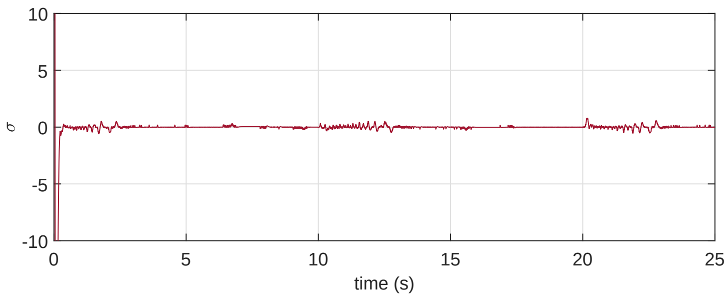
(b) *Actual speed of shaft and its estmate: state* x_2



(c) *Auxiallary control* w

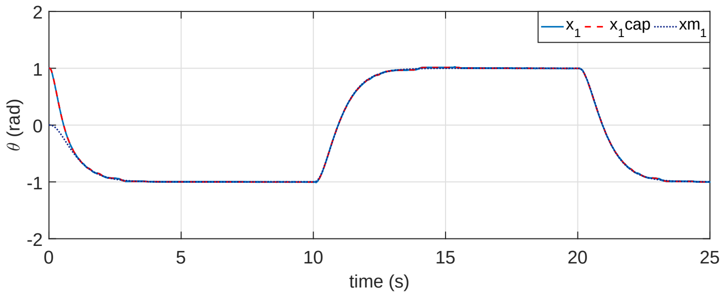


(d) *Control* u

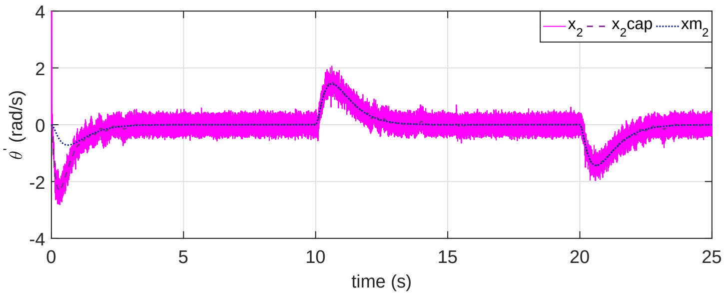


(e) *Sliding surface* σ

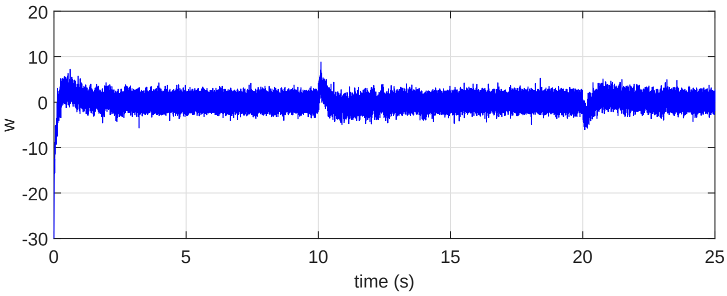
Figure 4: *IDO based model following DSMC control of DC servomotor without measuring noise.*



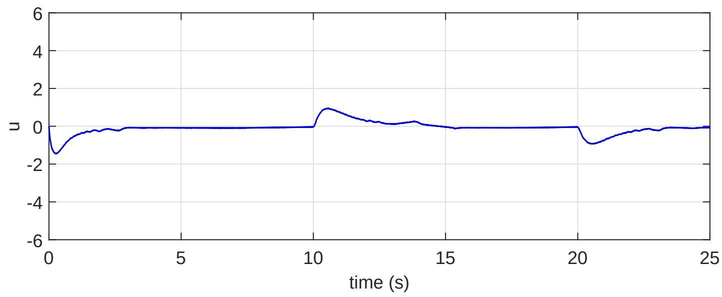
(a) *Actual position of shaft and its estmation: state* x_1



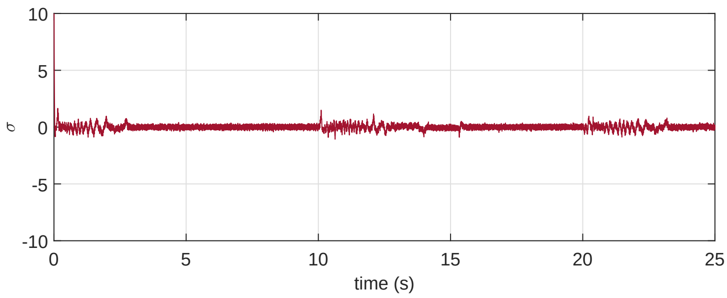
(b) *Actual speed of shaft and its estmate: state* x_2



(c) *Auxiallary control* w



(d) *Control* u



(e) *Sliding surface* σ

Figure 5: *IDO based model following DSMC control of DC servomotor with measuring noise.*

RT&A, No 1 (72)
Volume 18, March 2023

On Reliability: A Mathematical Fault Tree

M. S. Fahmy, A. I. Ahmed, M. Khalil

•

Faculty of Engineering, October University for Modern Sciences and Arts (MSA), Egypt.
msfahmy@msa.edu.eg; ahmed.ibrahim22@msa.edu.eg; mkibrahim@msa.edu.eg

Abstract

Fault tree analysis (FTA) is a top down approach that was initially used and developed in Bell laboratories in the year 1962 by H Watson and A Mearns for the intercontinental ballistic missile (ICBM) system for the US air force called the Minuteman System. Since then, the technique has been adopted and adapted by many companies who are interested in reliability engineering and dangerous technology. Today FTA is widely used in system safety and reliability engineering, aerospace, nuclear power, chemical and process, pharmaceutical, petrochemical and other high-hazard industries; but is also used in fields as diverse as risk factor identification relating to social service system failure and in software engineering for debugging purposes and is closely related to cause-elimination technique used to detect bugs. Now FTA is considered as one of the most important system reliability and safety analysis techniques. Fault tree analysis has proved to be a useful analytical tool to analyze the potential for system or machine failure by graphically and mathematically representing the system itself. It is a top-down approach that reverse-engineers the root causes of a potential failure through the root cause analysis process. Our main contribution is to develop a mathematical theory of fault tree analysis using some statistical concepts relating to probability of series and parallel systems to set up a mathematical model that represent any hierarchical control system to calculate its reliability for both homogeneous and nonhomogeneous structures. A Fault Tree is a hierarchical model used to analyze the probability that an event will occur. Fault Tree provides all the tools needed to build graphic representations of large-scale problems gracefully so we can use it to set up a mathematical model that represent any hierarchical control system and evaluate its reliability using our general mathematical formula that represent the structure in its two cases. The graphical representation (fault tree diagram) for a hierarchical controlled system enabled us to set up a mathematical general formula that help us to evaluate the reliability of the system in general case (nonhomogeneous structure) and another derived formula for the special case (homogeneous structure). This analysis may help to understand how one or more small failure events lead to a catastrophic failure.

Keywords: Reliability, Mathematical Modeling, Analysis of Fault Tree, Serial and Parallel Systems

1. Introduction

Reliability as a state can be a factor of of many parameters; such as, Mean Time To Failure, Reliability, Availability and few others. These terms have been developing over the last six decade. Its evident that such a concept will be portrayed in a structure of system or systems, Barlow in 1973 and Fussel in 1974, discussed the fault tree construction and concept respectively [1 and 2]. In some articles it has been observed that many of the quantities computed by fault tree analysis can also be computed using the concepts and techniques of reliability theory. Henceforth, this paper aims to build and construct an intuitive rigor of understanding reliability in the realm of Mathematical Statistics, in form of a fault tree mathematical model, whereas we aim to introduce the concepts in a matter of detail, building the theory upon prior establishments, and presenting a a general model by the end, incorporating the aforementioned usage of the prior sections. It serves, to add, that general assumptions will be accounted for and discussed, also mentioned when done otherwise. The Introduction section will serve as a block of terms, dissected into multi-topics that adds up to the required definition this paper aims to deliver. For in Section 1 we get to understand what reliability as a concept with their basic definitions. In Section 2 we discuss how system connections influence their reliability, and least in Section 3; We present and elaborate the homogeneity of a system of fault tree representation, giving a derived general model by the end.

1.1. Basic Definitions and Concepts

In this section we introduce some basic terms as reliability and its components, as well as its rigor definitions.

1.1.1 Reliability: What we quantify as reliable

Reliability purposefulness can be interpreted as; it serves as a developmental method for Engineering. In which it helps in: cost, effort, and time efficiency. It is essential to grasp that reliability is not a property, but a characteristic of an item, in which it is expressed by the probability that the said system will function as expected for a stated time interval [3]. It is usually denoted by R. One can view reliability in other terms, namely, a quantitative approach, as from said approach, the one thing we care about the most, is how long a system stays operational with no interruption. However, this does not imply that redundant parts might not fail, as they can fail and be repaired without causing an operational interruption at system level. Thus, one may conclude that, the concept of reliability can be applied to both repairable and non-repairable systems.

Definition 1.1 (Reliability)**.** Given n statistical identical systems, which starts into operations at $t = 0$, $\bar{v} < n$ is to accomplish them successfully, where $\bar{v} \in \mathbb{R}^n$.Then, We can write that the ratio $\frac{\bar{v}}{n}$ is a mere random variable which converges to the true value of the reliability as n increases. Namely, $\lim_{n\to\infty} \frac{\bar{v}}{n} = R$: $|\frac{\bar{v}}{n} - R| < \epsilon$

Definition 1.2 (Reliability)**.** It is the ability of an object (or process or service) to function as expected to fulfil the demanded tasks under given conditions. In other accurate words reliability is the probability a component or system will perform as designed. In which the value would range from [0,1]. reliability is related to failure rate by a simple exponential function: $R = e^{-\lambda t}$ Where R is the reliability, e is the Euler constant, λ is the failure rate, and t is the time.

1.1.2 Essential Metrics: Basic Definitions

Few Metrics need to be defined in order to be able to calculate the reliability of a system [3, 4, and 5].

Definition 1.3 (Failure rate (λ))**.** The frequency of a component failure per unit time, it is an essential metric that is used to calculating either reliability or availability.

$$\lambda = \frac{1}{MTBF} \equiv \frac{1}{MTTF}$$

Definition 1.4 (Repair rate (μ))**.** The frequency of successful repair operations performed on a failed component per unit time.

$$\mu = \frac{1}{MTTR}$$

Definition 1.5 (Mean time to failure (MTTF))**.** The average time duration before a non-repairable system component fails.

$$\text{MTTF} = \frac{\sum Hours\ of\ Operation}{\sum Units} \equiv \frac{1}{\lambda}$$

Definition 1.6 (Mean time between failure (MTBF))**.** The average time duration between inherent failures of a repairable system component.

$$\text{MTBF} = \frac{\sum Hours\ of\ Operation}{\sum Failurs} \equiv \frac{1}{\lambda} \equiv \text{MTTF+ MTTR}$$

2. System Types: Analysis

What we are concerned about mathematically, when we are assessing the reliability of a system is its own basic sub-systems. The purpose of this paper, is to detail the explanation to the process, given a system or a network of systems that are adjoined together. The physical layout or rather the connection of the system is of a cruciality to both its functioning, and its reliability.

2.1. Series Systems

A network with N systems or blocks in a series link (Figure 1 [3]) where the failure of any one item causes the entire system to fail [3]. As a result, for a series system to perform properly, all of its components must function properly within the time range specified t.

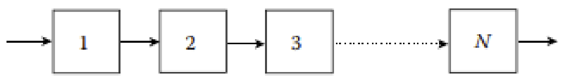


Figure 1: *Series System*

Definition 2.1. (Reliability of a Series System) The reliability of a system in a series connection is the probability that all N items succeed during its intended interval of time t.

$$R_s(t) = R_1(t) \cdot R_2(t) \dots R_N(t) = \prod_{i=1}^{N} R_i(t)$$

A practical conclusion is that the reliability of a series system is always lower than the reliability of any of its components.

We are also concerned with the instantaneous failure rate, one can conclude such an outcome by recalling the definition of $\lambda(t)$:

$$\lambda_s(t) = \frac{-d \ln \prod_{i=1}^{N} R_i(t)}{dt} \equiv \sum_{i=1}^{N} \frac{-d \ln R_i(t)}{dt} \equiv \sum_{i=1}^{N} \lambda_i(t)$$

2.2. Parallel Systems

In a parallel arrangement or link, a system failure is caused by the failure of all components, not just one. As a result, the performance of only one unit will be enough to ensure the system's overall success.

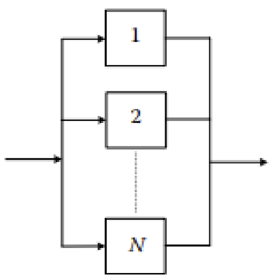


Figure 2: *Parallel System*

Definition 2.2 (Reliability of a Parallel System)**.** For a set of N independent items connected in parallel (Figure 2 [3]), their failure rate is be given by:

$$F_s(t) = F_1(t) \cdot F_2(t) \ldots F_N(t) = \prod_{i=1}^{N} F_i(t)$$

Since $R_i(t) = 1 - F_i(t)$

$$R_s(t) = 1 - F_s(t) = 1 - \prod_{i=1}^{N} [1 - R_i(t)]$$

The instantaneous failure rate is still an essential metric, however in parallel configuration, it is not as trivial to come up with one. one can start with the definition that the failure rate is $h(t) = \frac{-d \ln R(t)}{dt}$, yet it will lead to a complicated formula, for instance, let a system of two units with constant failure rate be connected in parallel, their failure rate can be given by;

$$\lambda_s(t) = \frac{\lambda_1 \exp(-\lambda_1 t) + \lambda_2 \exp(-\lambda_2 t) - (\lambda_1 + \lambda_2) \exp(-(\lambda_1 + \lambda_2)t)}{\exp(-\lambda_1 t) + \exp(-\lambda_2 t) - \exp(-(\lambda_1 + \lambda_2)t)}$$

For the same instance, or case; of an N identical units in parallel with a constant failure rate, their reliability can be put to as:

$$R_s(t) = 1 - [1 - \exp(-\lambda t)]^N$$

2.3. General Structure

Consider the K-out-of-N system, which is a more general structure of series and parallel systems (Figure 3 [3]). If any combination of K units out of N independent units works in this type of system, the system is guaranteed to succeed. Assume that all units are identical for the sake of simplicity. The likelihood that the system will work is represented by the binomial distribution [7]:

$$R_s(t) = \sum_{r=K}^{N} \binom{N}{r} [R(t)]^r [1 - R(t)]^{N-r} = 1 - \sum_{r=0}^{K-1} \binom{N}{r} [R(t)]^r [1 - R(t)]^{N-r}$$

The majority of practical systems are neither parallel nor series, but rather a blend of the two. Parallel-series systems are a common name for these systems. A complex system that is neither series nor parallel alone, nor parallel-series, is another sort of complex system. Which is referred to as nonparallel-series system (Figure 4 [3]).

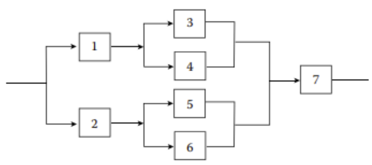


Figure 3: *A parallel-series system*

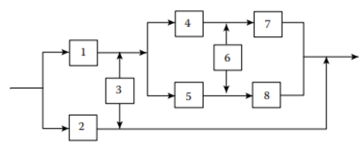


Figure 4: *A nonparallel-series system*

3. Homogeneity of a System

In this section we discuss the homogeneity of a system in form of a fault tree, where their nodes branching, and rate of distribution affect their reliability. We start by introducing what a fault tree is, then how they are analyzed.

3.0.1 Fault Trees and Reliability Block Diagrams

It is important to understand the essential difference between RBDs and fault tree diagrams; RBDs work in success space while, FTDs work in failure space; the FTDs addresses the failure combinations while the RBDs address the success combinations. Also FTDs are traditionally used in analyzing fixed probabilities, while RBDs may include time-varying distributions for the blocks' failure or success, in addition to other properties like restoration or repair distributions [5 and 6].

3.0.2 Fault Tree Analysis

Fault Tree Analysis can be explained simply as an analytical technique that describes an undesired state of the system. normally that is in a critical state from a safety standpoint. The system is inspected in the environmental context and operation to extract all credible ways for the undesired event to occur [3,4, and 5]. It is also important to point out that a fault tree is a graphic model of the different sequential and parallel combinations of faults that will happen in the investigated model. In fact, using the model of fault tree is more convenient to deal with, because it is qualitative model that enable us to evaluate it quantitatively and do not change the qualitative nature of the model itself.

3.1. Homogeneous and Non-homogeneous systems: Analysis

The homogeneity of a system influences the reliability of the whole system [3] and that can be seen evident in a fault tree, where some demonstrate a homogeneous structure, and some do not. We represent in Figure 5 the model of general fault tree, and define how said figure can be interpreted to be homogeneous or non-homogeneous.

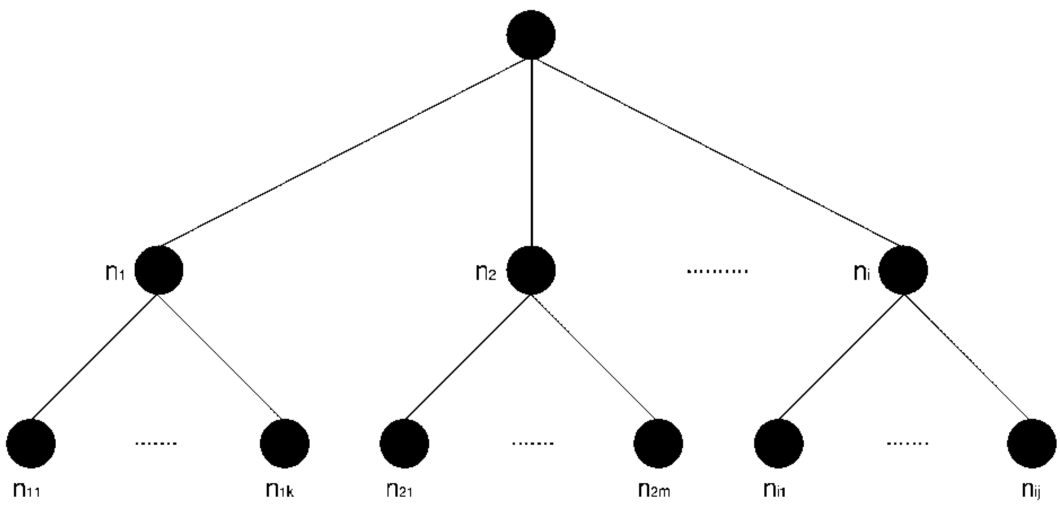


Figure 5: *Generic Tree*

Definition 3.1 (Homogeneity of a tree)**.** A tree is said to be homogeneous if and only if, the number of its sub nodes is equal to the number of every other sub node on any level from the root. Mainly, $n_{1k} = n_{2m} = n_{ij}$

Definition 3.2 (Non-Homogeneity of a tree)**.** A tree is said to be non-homogeneous if and only if, the number of its sub nodes is not equal to at least one other sub node on any level from the root. Mainly, $n_{1k} \neq n_{2m} \neq n_{ij}$

3.2. General Case

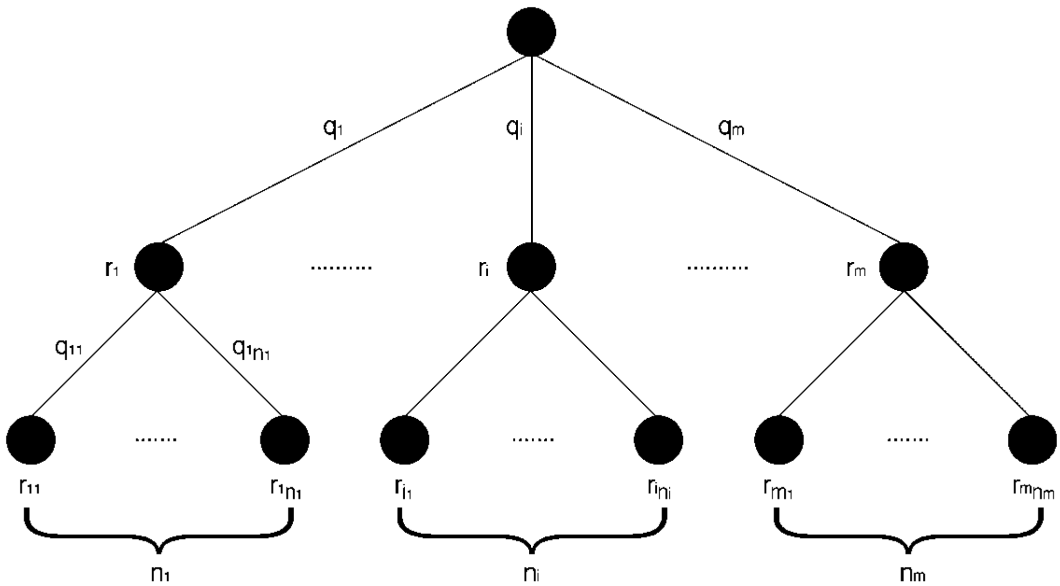


Figure 6: $\varphi - Tree$

In Figure 6, let P_0 be the reliability of the root; $q_1, \ldots, q_m$ are to be the reliability of the edges, and $r_1, \ldots, r_m$ be the reliability of the nodes of the first level, where m is the number of nodes of the first level; $q_{i_1}, \ldots, q_{in_i}$ be the reliability of edges of second level in the i-th subtree; $r_{i_1}, \ldots, r_{in_i}$ be the reliability of nodes of the second level in the i-th subtree, where n_i is the number of edges (nodes) in the i-th subtree. It goes evident to see that the number of leaves is equal to N, where; $N = n_1 + n_2 + \ldots + n_m$ such that $n_1 \leq n_2 \leq \ldots \leq n_m$. For the purpose of generality, consider that an arbitrary path from the root to the end is unfailing (successful), if all nodes and edges on said path is unfailing (successful). Now we find the reliability of the tree (φ) through V paths, throughout the derivation such reliability is detonated $\wp(\varphi; V)$; where; $V = 1, 2, \ldots, N$. $N \in \mathbb{N}$-$\{0\}$.

Theorem 1.

$$\wp(\varphi; V) = P_0 \times \sum_{k=0}^{m} \sum_{\substack{A, A \subset \{1,\ldots,m\} \\ |A|=k}} \left[\prod_{i \in A} \left(1 - P_i \left(1 - \prod_{j=1}^{n_i} (1 - P_{ij}) \right) \right) \right] \times$$
$$\left[\sum_{\substack{a_1+\ldots+a_m=V \\ i \in A \implies a_i=0}} \left(\prod_{i \notin A} (P_i \times \left(\sum_{\substack{B, B \subset \{1,\ldots,n_i\} \\ |B|=a_i}} \left(\prod_{j \in B} P_{ij} \times \left(\prod_{\substack{j \notin B \\ 1 \leq j \leq n_i}} (1 - P_{ij}) \right) \right) \right) \right) \right]$$

Where:

$$P_i = r_i q_i, P_{ij} = r_{ij} q_{ij}, i = 1, \ldots, m \quad i \leq j \leq n_j$$
$$\prod_{\phi} = 1, \quad \sum_{\phi} = 0$$

Proof.
We proof the theorem by showing that the probability of failure throughout all paths from the

root to the i-th subtree with n_i nodes is equal to

$$1 - P_i\left(1 - \prod_{j=1}^{n_j}\left(1 - P_{ij}\right)\right)$$

The reliability of the fault tree through exactly a paths (1≤ a ≤n$_i$) from the root to the i-th subtree is equal to

$$P_i \times \left(\sum_{\substack{B,B\subset\{1,\ldots,n_i\}\\|B|=a_i}}\left(\prod_{j\in B} P_{ij} \times \left(\prod_{\substack{j\notin B\\1\le j\le n_i}}\left(1 - P_{ij}\right)\right)\right)\right)$$

Further we remark that the event of having exactly V paths to operate successfully is considered the sum of all mutually exclusive events of the types; there are exactly a_1 paths operate simultaneously successfully ending in the first subtree, a_2 paths operate simultaneously successfully ending in the second subtree, and a_m paths operate simultaneously successfully ending in the m-th subtree; where $a_1 + a_2 + \ldots + a_m = V$. From here it follows

$$\wp(\varphi; V) = P_0 \times \left(\sum_{\substack{a_1+\ldots+a_m=V\\0\le a_j\le n_j}}\prod_{i=1}^{m}\Omega(i, a_i)\right)$$

where

$$\Omega(i,a_i)=\begin{cases} 1 - P_i\left(1 - \prod_{j=1}^{n_j}\left(1 - P_{ij}\right)\right) & a_i = 0 \\ P_i \times \left(\sum_{\substack{B,B\subset\{1,\ldots,n_i\}\\|B|=a_i}}\left(\prod_{j\in B} P_{ij} \times \left(\prod_{\substack{j\notin B\\1\le j\le n_i}}\left(1 - P_{ij}\right)\right)\right)\right) & a_i \neq 0 \end{cases}$$

■

3.3. Special Case

For the special case that the system behaves in a homogeneous matter as per Definition 3.1. Namely, this implies that: $q_1 = q_2 = \ldots = q_m = q_0$, and for the second level operating on the same type and identical conditions: $q_{11} = \ldots = q_{1n_1} = q_{21} = \ldots = q_{2n_2} = \ldots = q_{m1} = \ldots = q_{mn_m} = q_1$. It trivially follows that $r_{11} = \ldots = r_{1n_1} = r_{21} = \ldots = r_{2n_2} = \ldots = r_{m1} = \ldots = r_{mnm} = r_1$, and $n_1 = n_2 = \ldots = n_m = n$.
By transforming the formula of Theorem 1, by change of assertions; namely: $P_1 = P_2 = P_m = q_0 r = p_1; P_{11} = P_{12} = \ldots = P_{m_1} = \ldots = P_{mn_m} = q_1 r = p_2$.
Let $\wp(\ \varphi_1\ ;\ V)$ denote the reliability of such case, from Section 2.3 it follows:

Corollary 1.

$$\wp\left(\ \varphi_1\ ;\ V\right) = p_0 \times \sum_{k=\lceil \frac{V}{n}\rceil}^{m}\left\{\binom{m}{k} \times p_1^k \times \left[1 - p_1 \times \left(1 - (1 - p_2)^n\right)\right]^{m-k} \times \binom{k \times n}{V} \times p_2^V \times (1 - p_2)^{(k\times n)-V}\right\}$$

Where; $\binom{m}{k} = \frac{m!}{k!(m-k)!}$ And $\left\lceil \frac{V}{n}\right\rceil$ is the smallest integer greater than $\frac{V}{n}$

4. Conclusion

A theoretical study is introduced for a mathematical model that can describe a hierarchical control system (homogeneous and non-homogeneous) in order to evaluate its reliability. This investigation of the reliability is highly suitable for study of many Engineering applications ranging from industrial process control, through production management to Economic and other systems.

References

[1] Barlow, E. (1973). Introduction to Fault Tree Analysis. Research Air Force Systems Command.

[2] Fussell, J. B. (1974). Fault Tree Analysis - Concepts and Techniques. Nato Advanced Study Institute On Generic Techniques Of System Reliability Assessment, Nordhoff Publishing Company.

[3] Rausand, M., Hyland, A. (2004). System Reliability Theory: Models, Statistical Methods, and Applications (2nd ed.). Wiley.

[4] Epstein, Weissman. (2009). Mathematical Models for Systems Reliability. CRC Press.

[5] Mavko, B., V., A., Marko, C. (2009). Application of the fault tree analysis for assessment of power system reliability (Vol. 94). Reliability Engineering System Safety.

[6] Curcuru, G., Galante, G. M., La Fata, C. M. (2013). An imprecise Fault Tree Analysis for the estimation of the Rate of Occurrence Of Failure. J Loss Prev Process Ind.

[7] Carpitella, S., Certa, A., Izquierdo Sebastin, J., La Fata, C. M. (2018). k-out-of-n systems: an exact formula for the stationary availability and multi-objective configuration design based on mathematical programming and TOPSIS. Journal of Computational and Applied Mathematics.

A TESTING-EFFORT BASED SRGM INCORPORATING IMPERFECT DEBUGGING AND CHANGE POINT

Umashankar Samal[1*], Shivani Kushwaha[2], Ajay Kumar[3]

•

[1,2,3]Department of Applied Sciences, ABV-IIITM Gwalior, Gwalior-474015, MP, India
[1]umashankar@iiitm.ac.in, [2]shivanik@iiitm.ac.in, [3]ajayfma@iiitm.ac.in
*Corresponding Author

Abstract

In this paper, a scheme for constructing software reliability growth model based on Non-Homogeneous Poisson Process is proposed. Here, we consider the software reliability growth model that incorporates with imperfect debugging, change point and testing effort. However, most researchers assume a constant detection rate per fault in deriving their software reliability models. They suppose that all faults have equal probability of being detected during the software testing process, and the rate remains constant over the intervals between fault occurrences. In reality, the fault detection rate strongly depends on the skill of test teams, program size, and software testability. Also in most realistic situations, fault repair has associated with a fault re-introduction rate due to imperfect debugging phenomenon. In this case, the fault detection rate and fault introduction rate will be changed during the software development process. Therefore, here we incorporate both generalized logistic testing-effort function, change-point parameter into software reliability modelling. The Least Square Estimation approach is used to estimate the unknown parameters of the new model. So in our new proposed model we collect software testing data from real application and utilize it to illustrate the proposed model. Experimental results show that the proposed framework to incorporate both testing-effort and change-point for Imperfect-Debugging SRGM has a fairly accurate prediction capability.

Keywords: Stochastic software reliability model , Non-homogeneous Poisson process, Imperfect debugging, Testing effort, Change point.

Acronyms and Notations

NHPP	nonhomogeneous Poisson process
LSE	least squares estimation
SRGM	software reliability growth model
TEF	testing effort function
ID	imperfect debugging
$m(t)$	mean value function, i.e., the expected number of software failures by time t
$a(t)$	error content function
$b(t)$	error detection rate per error at time t
$R(t)$	reliability function
$R(x/t)$	conditional software reliability
β	fault introduction rate
$w(t)$	current testing-effort estimated by a logistic testing effort function
$W(t)$	cumulative testing effort estimated by a logistic testing effort function
τ	change point
$\lambda(t)$	failure density function

1. Introduction

There are many reasons that a software fails, but usually a software fails because of design issue. Other failures occur when the code is written, or when changes are introduced to a working system. For the past several decades, various statistical models have been proposed to assess the software reliability. So, we have reviewed many previous well known models. The NHPP based software reliability growth models are proved quite successful in practical software reliability engineering [1] . The main issue of the NHPP model is to determine an appropriate mean value function to denote the expected number of failures experienced up to a certain time point. Model parameters can be estimated by using maximum likelihood or least square estimate. Once the mean value function is determined the software reliability and the related measurements can be easily derived.

We have found that most SRGMS use calendar time as a unit to fault detection and removal [2, 3, 5, 8, 16], but very few use human power, number of test case runs, or CPU time as a unit [12]. Researchers have proposed SRGMs that incorporates the concept of TEF into an NHPP model to get a better description on the software fault phenomenon [4, 6, 7, 9, 10, 11, 13, 15, 17, 18, 19, 20]. TEF has the advantage of relating the work profile more directly to the natural structure of software development. Many researchers assume that all faults have equal probability of being detected during the software testing process, and the rate remains constant over the intervals between fault occurrences, constant detection rate per fault, but we, in our model, assume that the rate changes before and after a fixed point. Most authors also assume that the once the fault is removed there will be no new faults introduced but we assume that even during debugging new faults will be introduced, and also introduction rate may not be same during overall testing. So in our model, there will be a imperfect debugging with change point and testing effort scenario.

2. Related work

In this section, we reviewe the well-known NHPP SRGMs and then introduce our new general model that incorporates both the imperfect debugging, change-point problem and testing effort. So we first start with some well known SRGMs.

2.1. A general NHPP model

Let $N(t)$ be a counting process representing the cumulative number of software failure by time t. The counting process is shown to be a NHPP with a mean value function $m(t)$. Mean value function represents the expected number of failures by time t. Goel and Okumoto [2] assume that number of software failures is time independent and software failure density is proportional to residual fault content. Thus $m(t)$ can be solved by solving the following differential equation.

$$\lambda(t) = \frac{dm(t)}{dt} = b(a - m(t)) \tag{1}$$

Where a denotes the initial number of faults contained in a program and b represents the fault detection rate. The result shows that

$$m(t) = a(1 - e^{-b\,t}) \tag{2}$$

The conditional software reliability, $R(x|t)$, is defined as the probability that there is no failure observed in the time period $(t, t+x)$, given that the last failure occurred at a time point $t(t \geq 0; x > 0)$. Given the mean value function $m(t)$, the conditional software reliability can be shown as

$$R(x|t) = e^{-[m(t+x)-m(t)]} = exp[-a(e^{-bt} - e^{-b(t+x)}] \tag{3}$$

2.2. Imperfect software debugging models

In the general NHPP model, a constant fault content function implies the perfect debugging assumption,i.e., no new faults are introduced during the debugging process. Pham introduced an NHPP SRGM in imperfect debugging environment. He assumed if detected faults are removed, then there is a possibility of introduction of new faults with a constant rate β. Let $a(t)$ be the number of faults to be eventually detected (denoted by a) plus the number of new faults introduced to the program by time t, the mean value function $m(t)$ can be given as the solution of the following system of differential equations:

$$\frac{\partial m(t)}{\partial t} = b[a(t) - m(t)], \frac{\partial a(t)}{\partial t} = \beta \frac{\partial m(t)}{\partial t} \tag{4}$$

Solving the above equations, we can obtain the mean value function and conditional software reliability, respectively, as follows

$$m(t) = \frac{a}{1-\beta}\left[1 - e^{-(1-\beta)bt}\right], R(x|t) = exp\left(-m(x)e^{-(1-\beta)bt}\right) \tag{5}$$

2.3. An NHPP model with change-point

Many SRGMs suppose the fault detection rate is a constant, or a monotonically increasing function. The failure intensity is expected to be a continuous function of time. An increasing fault detection rate function represents the debugging process with the learning phenomenon. But the fault detection rate can be affected by many factors such as the testing strategy and resources allocation. During a software testing process, there is a possibility that the underlying fault detection rate function is changed at some time moments τ, called as change-point.

Chang [7] considered the change-point problems in the NHPP SRGMs. The parameters of the NHPP with change-point models are estimated by the weighted least square method. Let the parameter τ be the change point that is considered unknown and is to be estimated from the data. The fault detection rate function is defined as :

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

On solving for $m(t)$

$$m(t) = \begin{cases} a(1 - e^{-b_1 t}) & when \quad 0 \le t \le \tau, \\ a(1 - e^{-b_1\tau - b_2(t-\tau)}) & when \quad t > \tau \end{cases}$$

Reliability function will be

$$R(x|t) = \begin{cases} exp\left\{-a\left(e^{-b_1 t} - e^{-b_1(t+x)}\right)\right\} & when \quad t \le t + x \le \tau, \\ exp\left\{-a\left(e^{-b_1 t} - e^{-b_1\tau - b_2(t+x-\tau)}\right)\right\} & when \quad t \le \tau \le t + x \\ exp\left\{-a\left(e^{-b_1\tau - b_2(t-\tau)} - e^{-b_1(t+x)}\right)\right\} & when \quad \tau < t \end{cases}$$

2.4. Imperfect software debugging model with change point

To consider the NHPP SRGM that integrates imperfect debugging with change-point problem, the following assumptions are made:

(a) When detected faults are removed at time t, there is a possibility of introduction new faults at a rate $\beta(t)$.

$$\beta(t) = \begin{cases} \beta_1 & when \quad 0 \le t \le \tau, \\ \beta_2 & when \quad t > \tau \end{cases}$$

(b) The fault detection rate represented as following is a step function

$$b(t) = \begin{cases} b_1 & when \quad 0 \leq t \leq \tau, \\ b_2 & when \quad t > \tau \end{cases}$$

(c) A NHPP model of fault detection phenomenon in the software system.

The testing strategy and resource allocation can be tracked all the time during the fault detection process. It may be more reasonable to reconsider that the change-point (τ) is given. According to these assumptions, one can derive the new set of differential equations to obtain the new mean value function.

$$\frac{\partial m(t)}{\partial t} = b[a(t) - m(t)], \frac{\partial a(t)}{\partial t} = \beta \frac{\partial m(t)}{\partial t}, a(0) = a, m(0) = 0 \tag{6}$$

Solving the differential equation (6)

$$m(t) = \begin{cases} \frac{a}{1-\beta_1}[1 - e^{-(1-\beta_1)b_1 t}] & when \quad 0 \leq t \leq \tau, \\ \frac{a}{1-\beta_2}[1 - e^{-(1-\beta_1)b_1\tau-(1-\beta_2)b_2(t-\tau)}] + \frac{m(\tau)(\beta_1-\beta_2)}{1-\beta_2} & when \quad t > \tau \end{cases}$$

$$\lambda(t) = \frac{\partial m(t)}{\partial t} = \begin{cases} ab_1 e^{-(1-\beta_1)b_1 t} & when \quad 0 \leq t \leq \tau, \\ ab_2 e^{-(1-\beta_1)b_1\tau-(1-\beta_2)b_2(t-\tau)} & when \quad t > \tau \end{cases}$$

3. Proposed model

In this section, we propose a new Software Reliability Model that incorporates imperfect debugging with change-point and testing effort. Beginning with the necessity of testing in software reliability, we will make some assumptions for our model to construct it. In the beginning of the testing phase, many faults can be discovered by inspection and the fault detection rate depends on the fault discovery efficiency, the fault density, the testing effort, and the inspection rate. Later, the fault detection rate depends on some more additional parameters such as the failure-to-fault relationship, the code expansion factor, the skill of test teams, program size, and software testability.

Here we use a NHPP model with TEF and The following assumptions are made for the same.

(a) The fault removal process follows the Non-Homogeneous Poisson Process (NHPP).

(b) The software system is subject to failures at random times caused by the manifestation of remaining faults in the system.

(c) The mean number of faults detected in the time interval $(t, t + \lambda t)$ by the current testing-effort expenditures is proportional to the mean number of remaining faults in the system.

(d) The consumption curve of testing-effort is modelled by a generalized logistic TEF.

$$W(t) = \frac{N}{1 + A\, exp[-\alpha t]}$$

where

N = total amount of testing-effort to be eventually consumed,

α = consumption rate of testing-effort expenditures,

A = constant

(e) When detected faults are removed at time t, it is possible to introduce new faults with introduction rate $\beta(t)$.

$$\beta(t) = \begin{cases} \beta_1 & when \quad 0 \le t \le \tau, \\ \beta_2 & when \quad t > \tau \end{cases}$$

(f) The fault detection rate represented as following is a step function.

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

We can describe an SRGM based on generalized logistic TEF with fault introduction rate and change-point as follow:

$$\frac{dm(t)}{dt} * \frac{1}{w(t)} = b(t) * (a - m(t))$$
$$\frac{\partial a(t)}{\partial t} = \beta \frac{\partial m(t)}{\partial t} \tag{7}$$
$$a(0) = a, m(0) = 0$$

$W(t)$ is defined as-

$$W(t) = \int_0^t w(\tau)\, d\tau$$

From the above differential equation, the mean value function will be

$$m(t) = \begin{cases} \frac{a}{1-\beta_1}\left[1 - e^{-(1-\beta_1)b_1(W(t)-W(0))}\right] & when \quad 0 \le t \le \tau, \\ \frac{a}{1-\beta_2}\left[1 - e^{-(1-\beta_1)b_1(W(\tau)-W(0))-(1-\beta_2)b_2(W(t)-W(\tau))}\right] + \frac{m(\tau)(\beta_1-\beta_2)}{1-\beta_2} & when \quad t > \tau \end{cases}$$

Failure density function is

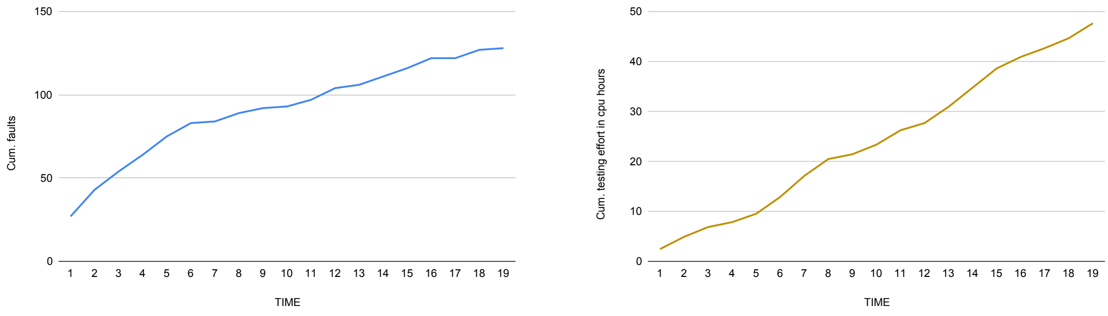
$$\lambda(t) = \frac{\partial m(t)}{\partial t} = \begin{cases} ab_1 w(t) e^{-(1-\beta_1)b_1(W(t)-w(0))} & when \quad 0 \le t \le \tau, \\ ab_2 w(t) e^{-(1-\beta_1)b_1(W(\tau)-W(0))-(1-\beta_2)b_2((W(t)-W(\tau))} & when \quad t > \tau \end{cases}$$

The results for mean value function and failure intensity function obtained have integrated the imperfect debugging change point problem and testing effort problem into a single NHPP SRGM. The unknown parameters for the above equations can be calculated using LSE.

4. Numerical and Data Analysis

4.1. Description of real dataset

To verify the new proposed model and to evaluate the performance of the SRGM, We have taken a dataset from Ohba [14]. The total testing time, cumulative number of software failures and cumulative testing consumption of generalized logistic TEF are recorded.
To assume the change point for the given set of data given in Table (1), graph (1a) of the cumulated number of faults versus time has been considered. It is found that it is not differentiable around 11.2.
During the testing period, 20 Hours of experiments, 47.65 CPU Hours were consumed and about 128 software errors are removed. To find the unknown parameters of the logistic testing effort function, we have used LSE. Using LSE, the unknown parameters are found to be $N = 54.823$, $A = 13.033$, $\alpha = 0.226$.



(a) *Cumulative faults of $N(t)$ process*

(b) *Testing efforts vs time*

Figure 1: *CF and TEF versus time*

Table 1: *Ohba data set*

time	cumulative failure	testing effort consumption	time	cumulative failure	testing effort consumption
1.00	27.00	2.45	11.00	97.00	26.23
2.00	43.00	4.90	12.00	104.00	27.67
3.00	54.00	6.86	13.00	106.00	30.93
4.00	64.00	7.84	14.00	111.00	34.77
5.00	75.00	9.52	15.00	116.00	38.61
6.00	83.00	12.89	16.00	122.00	40.91
7.00	84.00	17.10	17.00	122.00	42.67
8.00	89.00	20.47	18.00	127.00	44.66
9.00	92.00	21.45	19.00	128.00	47.65
10.00	93.00	23.35			

4.2. Model Comparison

In the paper, we have compared the accuracy with Delayed S-shaped model [3] and Huan-Jyh Shyur's imperfect debugging and change point model [1].

No.	Model	m(t)
1	Delayed S-shaped [3]	$m(t) = a\left(1-(1+bt)\,e^{-bt}\right)$
2	Huan-Jyh Shyur's model [1]	$m(t) = \begin{cases} \frac{a}{1-\beta_1}\left[1-e^{-(1-\beta_1)b_1 t}\right] & \textit{when} \quad 0 \le t \le \tau, \\ \frac{a}{1-\beta_2}\left[1-e^{-(1-\beta_1)b_1\tau-(1-\beta_2)b_2(t-\tau)}\right] + \frac{m(\tau)(\beta_1-\beta_2)}{1-\beta_2} & \textit{when} \quad t > \tau \end{cases}$
3	New model	$m(t) = \begin{cases} \frac{a}{1-\beta_1}\left[1-e^{-(1-\beta_1)b_1(W(t)-W(0))}\right] & \textit{when} \quad 0 \le t \le \tau, \\ \frac{a}{1-\beta_2}\left[1-e^{-(1-\beta_1)b_1(W(\tau)-W(0))-(1-\beta_2)b_2(W(t)-W(\tau))}\right] + \frac{m(\tau)(\beta_1-\beta_2)}{1-\beta_2} & \textit{when} \quad t > \tau \end{cases}$

4.3. Comparision Criteria

In order to compare the performance of the proposed model with other models, we have used MSE [17]. MSE is defined as:

$$MSE = \sum_{i=1}^{n} \frac{[m(t_i) - m_i]^2}{D}$$

where $m(t_i)$, m_i and D represent estimated values, observed values and degrees of freedom respectively.

5. Results

The unknown parameters in the proposed model are a, b_1, b_2 and the unknown parameters in the testing effort function are α, N, A. We have examined the proposed model with the given dataset,

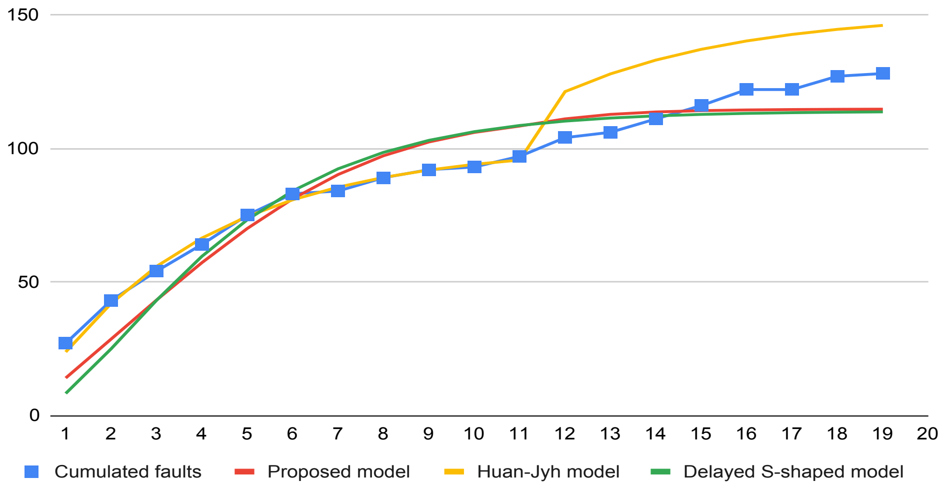


Figure 2: *Mean value functions $m(t)$ of all models for dataset1*

and the results from the TEF i.e. $N = 54.826$, $\alpha = 13.033$, $A = 0.226$. The fault introduction rate before change point β_1 is taken as 0.2 and the fault introduction rate after change point β_2 is taken as 0.5. Using LSE, the values of a, b_1 and b_2 are found to be 90.0059 0.1834 and 0.388 respectively.

Table 2: *Estimation of parameters for datasets*

No.	Model	Parameter	MSE
1	Delayed S-shaped [3]	$a = 113.9062$, $b = 0.438$	127.0564
2	Huan-Jyh Shyur's model [1]	$a = 80.75$, $b_1 = 0.335$, $b_2 = 0.5$	195.4614
3	New model	$a = 90.75$, $b_1 = 0.1834$, $b_2 = 0.3887$	101.8713

We have compared the proposed model with previous two well known models and the result shows that with introduction of testing effort function, performance of the model has increased.

6. Conclusion

In this paper, we present a new change point software reliability model considering the testing effort function based on NHPP and imperfect debugging environment. A generalized logistic testing effort function and effect of change point in a imperfect debugging environment are discussed and the explicit mean value function for the new model is presented. Furthermore, comparisons of this model with several existing change point, imperfect debugging models have also been provided in terms of values of MSE on Ohba data set. Numerical results demonstrate that the proposed model can give a better goodness-of-fit. It seems that this proposed model, though a little more sophisticated but By means of incorporating both ID, testing effort and the effect of change point, provides a more powerful property to model the changing fault detection rate, which describes more realistically actual effects of the real testing process.

References

[1] Shyur, Huan-Jyh. "A stochastic software reliability model with imperfect-debugging and change-point." *Journal of Systems and Software* 66, no. 2 (2003): 135-141.

[2] Goel, A.L. and Okumoto, K., 1979. Time-dependent error-detection rate model for software reliability and other performance measures. *IEEE transactions on Reliability*, 28(3), pp.206-211.

[3] Yamada, Shigeru, Mitsuru Ohba, and Shunji Osaki. "S-shaped software reliability growth models and their applications." *IEEE Transactions on Reliability* 33, no. 4 (1984): 289-292.

[4] Chiu, Kuei-Chen, Yeu-Shiang Huang, and Tzai-Zang Lee. "A study of software reliability growth from the perspective of learning effects." *Reliability Engineering* & *System Safety* 93, no. 10 (2008): 1410-1421.

[5] Pham, Hoang, and Xuemei Zhang. "NHPP software reliability and cost models with testing coverage." *European Journal of Operational Research* 145, no. 2 (2003): 443-454.

[6] Zhao, Ming. "Change-point problems in software and hardware reliability." *Communications in Statistics-Theory and Methods* 22, no. 3 (1993): 757-768.

[7] Chang, I. P. "An analysis of software reliability with change-point models." Taipei, Taiwan: National Science Council (1997).

[8] Pham, Hoang, Lars Nordmann, and Zuemei Zhang. "A general imperfect-software-debugging model with S-shaped fault-detection rate." *IEEE Transactions on reliability* 48, no. 2 (1999): 169-175.

[9] Huang, Chin-Yu, Sy-Yen Kuo, and Michael R. Lyu. "An assessment of testing-effort dependent software reliability growth models." *IEEE transactions on Reliability* 56, no. 2 (2007): 198-211.

[10] Huang, Chin-Yu. "Performance analysis of software reliability growth models with testing-effort and change-point." *Journal of Systems and Software* 76, no. 2 (2005): 181-194.

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

[12] Yamada, Shigeru, and Shunji Osaki. "Software reliability growth modeling: Models and applications." *IEEE Transactions on software engineering* 12 (1985): 1431-1437.

[13] Pradhan, Sujit K., Anil Kumar, and Vijay Kumar. "An Optimal Resource Allocation Model Considering Two-Phase Software Reliability Growth Model with Testing Effort and Imperfect Debugging." *Reliability: Theory* & *Applications* SI 2 (64) (2021): 241-255.

[14] Ohba, Mitsuru. "Software reliability analysis models." *IBM Journal of research and Development* 28, no. 4 (1984): 428-443.

[15] Khurshid, Shozab, A. K. Shrivastava, and Javaid Iqbal. "Effort based software reliability model with fault reduction factor, change point and imperfect debugging." *International Journal of Information Technology* 13, no. 1 (2021): 331-340.

[16] Li, Qiuying, and Hoang Pham. "A testing-coverage software reliability model considering fault removal efficiency and error generation." PloS one 12, no. 7 (2017): e0181524.

[17] Shrivastava, Avinash K., and P. K. Kapur. "Change points based software scheduling." *Quality and Reliability Engineering International* 37, no. 8 (2021): 3282-3296.

[18] 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, no. 1 (2018): 1-12.

[19] 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, no. 6 (2021): e2344.

[20] Panwar, Saurabh, Vivek Kumar, P. K. Kapur, and Ompal Singh. "Software reliability prediction and release time management with coverage." *International Journal of Quality* & *Reliability Management* (2021).

A New Pi-Exponentiated Method for Constructing Distributions with an Application to Weibull Distribution

M. A. Lone*

•

Department of Statistics, University of Kashmir, Srinagar, India
murtazastat@gmail.com

T. R. Jan

•

Department of Statistics, University of Kashmir, Srinagar, India
drtrjan@gmail.com

Abstract

A novel method for generating families of continuous distributions is presented by introducing a new parameter referred as Pi-Exponentiated Transformation (PET). Various properties of the PET method have been obtained. The method has been specialized on two-parameter Weibull distribution, and a new distribution called Pi-Exponentiated Weibull (PEW) is attained. A comprehensive mathematical treatment of the new proposal is provided. Closed-form expressions for the density function, distribution function, reliability function, hazard rate function have been provided. The PEW distribution is quite flexible, and it can be used to model data with decreasing, increasing or bathtub shaped hazard rates. Simulation study has been carried out to assess the behavior of the model parameters. Finally, the effectiveness of the suggested method is demonstrated by examining two real-life data sets.

Keywords: Pi-Exponentiated Transformation; Quantile Function; Reliability Function; Mean Waiting Time; Maximum Likelihood Estimation.

1. Introduction

Classical distributions are extensively employed in many applicable domains, including engineering, environmental studies, medical sciences, economics, actuarial, finance, insurance etc. to represent lifetime data. These distributions have been successfully implemented in all the fields listed above. However, in many domains, like reliability engineering and medical science, these conventional distributions do not offer the perfect fit when the data follow non-monotonic failure rates. As a result, generalized versions of these classical distributions are required to model reliability engineering and medical science data. Therefore, researchers became inspired to develop new modifications to theses existing distributions. These modified distributions offer more flexibility to the baseline model by introducing one or more extra parameters. In recent advances in distribution theory, researchers have shown a keen interest in proposing new methods for expanding the family of lifetime distributions. This has been accomplished through a variety of methods. Some well-known methods are:

- The exponentiated transformation initiated by Mudholkar and Srivastava [16], and is given by

$$F(x;\alpha) = (\psi(x))^{\alpha};\ \alpha > 0, x \in \mathbb{R}.$$

Where $\psi(x)$ is the cumulative distribution function (cdf) of baseline model.

- The beta-generated technique was proposed by Eugene et al. [7] that makes use of the beta distribution as the generator with parameters a and b to establish the beta generated distributions.

$$F(x) = \int_0^{\psi(x)} r(s)ds.$$

Where $r(s)$ is the probability density function (pdf) of a beta random variable (rv) and $\psi(x)$ is the cdf of any rv X.

- The quadratic rank transmutation map approach proposed by Shaw and Buckley [19] and is given as

$$F(x;\xi) = (1+\xi)\psi(x) - \xi\psi(x)^2,\ |\xi| \le 1,\ x \in \mathbb{R}.$$

Where $\psi(x)$ is the cdf of an existing distribution.

- Minimum Guarantee distribution proposed by Kumar et al. [9] and is given by

$$F(x) = e^{1-\frac{1}{\psi(x)}},\ x \in \mathbb{R}.$$

Where $\psi(x)$ is the cdf of an existing distribution.

- Log-transformation proposed by Maurya et al. [15] and is given by

$$F(x) = 1 - \frac{log(2-\psi(x))}{log2},\ x \in \mathbb{R}.$$

Where $\psi(x)$ is the cdf of an existing distribution.

- A new transmuted cumulative distribution function based on the Verhulst logistic function proposed by Kyurkchiev [10] and is given by

$$F(x) = \frac{2\psi(x)}{1+\psi(x)}.$$

Where $\psi(x)$ is the cdf of an existing distribution.

- Marshall and Olkin [14] proposed a general method for generating a new family of life distributions defined in terms of survival function as:

$$\bar{F}(x;\alpha) = \frac{\alpha\bar{\psi}(x)}{1-\bar{\alpha}\bar{\psi}(x)} = \frac{\alpha\bar{\psi}(x)}{\psi(x)+\alpha\bar{\psi}(x)} \quad ;\alpha > 0, x \in \mathbb{R}.$$

Where $\bar{\alpha} = 1-\alpha$ and $\bar{\psi}(x) = 1-\psi(x)$ is the survival function of the random variable X.

- Anwar et al. [8] presented a new method based on trigonometric function called Sine-Exponentiated-Transformation (SET). The cdf of SET family of distributions for $x \in \mathbb{R}$ is defined as

$$F_{SET}(x,\alpha) = \psi(x)\sin\left(\frac{\pi}{2}\psi^{\alpha}(x)\right)\ ;\alpha \ge 0.$$

Where $\psi(x)$ is the cdf of a continuous rv X.

- Lone et al. [11] proposed a new method for generating a family of continuous distributions called ratio trasformation (RT) method. The cdf of RT method for $x \in \mathbb{R}$ is defined as

$$F_{RT}(x;\alpha) = \frac{\psi(x)}{1+\alpha-\alpha^{\psi(x)}}\ ;\alpha > 0.$$

  Where $\psi(x)$ is the cdf of a continuous rv X.

- Recently, Lone et al. [12] introduced an innovative method for generating a family of continuous distributions called the MTI method. They employed MTI method on Weibull distribution and derived a new three-parameter MTI Weibull (MTIW) distribution. The cdf of MTI method for $x \in \mathbb{R}$ is defined as

$$F_{MTI}(x;\alpha) = \frac{\alpha\psi(x)}{\alpha - log\alpha\ \bar{\psi}(x)}\ ;\alpha > 0.$$

  Where $\bar{\psi}(x) = 1 - \psi(x)$ is the survival function of the random variable X.

In this manuscript a novel method for introducing greater flexibility to a family of distribution functions by bringing in new parameter to the given family has been introduced. This novel method has been refereed as PET. The proposed PET transformation is very simple and efficient method for introducing a new parameter to generalize the existing distributions. Some general properties of this class of distribution functions have been discussed. Then PET method has been specialized to a two-parameter Weibull distribution and generated a three-parameter PEW distribution, several statistical and reliability measures of PEW distribution have been obtained.

In section 2, the pdf and the cdf of the novel method have been obtained and various general properties of this method have been discussed. In section 3, the method has been specialized on two-parameter Weibull distribution and its structural properties as well as reliability measures have been obtained. In section 4, estimates of unknown parameters and simulation study have been performed. In section 5, two real data sets were analyzed to illustrate the efficacy of the suggested model. In section 6, the conclusion is stated.

2. General properties of PET method

Let X be a continuous rv, then the cdf of PET for $x \in \mathbb{R}$, is defined as

$$F_{PET}(x) = \frac{\pi^{(F(x))^{\alpha}} - 1}{\pi - 1}\ ; \qquad \alpha > 0. \tag{1}$$

Obviously, $F_{PET}(x)$ is a valid cdf only if $F(x)$ is a valid cdf. The corresponding pdf of PET for $x \in \mathbb{R}$, is defined as

$$f_{PET}(x) = \frac{\alpha log\pi}{\pi - 1}\pi^{(F(x))^{\alpha}}(F(x))^{\alpha-1}f(x)\ ; \qquad \alpha > 0. \tag{2}$$

Clearly, $f_{PET}(x)$ is a weighted version of $f(x)$, the weight function is given by

$$v(x) = \pi^{(F(x))^{\alpha}}(F(x))^{\alpha-1}.$$

Therefore, $f_{PET}(x)$ can be written as

$$f_{PET}(x) = \frac{f(x)v(x)}{k}.$$

Where, $k = E[v(X)]$ is the normalizing constant.

By using the following power series

$$\alpha^u = \sum_{j=0}^{\infty} \frac{(log\alpha)^j}{j!} u^j, \tag{3}$$

the linear representation for the cdf and the pdf in (1) and (2) are respectively given by

$$F_{PET}(x) = \frac{1}{\pi - 1}\left[\sum_{j=0}^{\infty} a_j (F(x))^{\alpha j} - 1\right]$$

and

$$f_{PET}(x) = b \sum_{j=0}^{\infty} a_j (F(x))^{\alpha(j+1)-1} f(x).$$

Where, $a_j = \frac{(log\alpha)^j}{j!}$ and $b = \frac{\alpha log\pi}{\pi-1}$.

The reliability function $R_{PET}(x)$ is given by

$$R_{PET}(x) = \frac{\pi}{\pi - 1}\left(1 - \pi^{(F(x))^{\alpha}-1}\right) \; ; \qquad \alpha > 0. \tag{4}$$

The hazard rate function $h_{PET}(x)$ is given by

$$h_{PET}(x) = \frac{\alpha log\pi f(x)(F(x))^{\alpha-1}}{\pi^{1-(F(x))^{\alpha}} - 1} \; ; \qquad \alpha > 0. \tag{5}$$

If $h(x)$ and $R(x)$ are the hazard rate function and reliability function of f then the hazard rate $h_{PET}(x)$ is given by

$$h_{PET}(x) = \alpha log\pi h(x) R(x) \frac{(F(x))^{\alpha-1}}{\pi^{1-(F(x))^{\alpha}} - 1} \; ; \qquad \alpha > 0. \tag{6}$$

From (6), it is clear that

$$\lim_{x\to-\infty} h_{PET}(x) = \begin{cases} 0 & \forall\ \alpha > 1 \\ \frac{log\pi}{\pi-1} \lim_{x\to-\infty} h(x) & \forall\ \alpha = 1 \\ \infty & \forall\ \alpha < 1 \end{cases}$$

and

$$\lim_{x\to\infty} h_{PET}(x) = \lim_{x\to\infty} h(x).$$

If $F^{-1}(x)$ exists, then for $\alpha > 0$, a random sample from $F_{PET}(x)$ can be obtained as

$$X = F^{-1}\left\{\left(\frac{log(1+U(\pi-1))}{log\pi}\right)^{\frac{1}{\alpha}}\right\}$$

where U is a uniform rv, $0 < u < 1$.

3. PEW distribution and its properties

A rv X has a three-parameter PEW distribution denoted by PEW(α, β, λ) with parameters α, β and λ, if the cdf and the pdf of X for $x > 0$, are respectively, given by

$$F_{PEW}(x) = \frac{\pi^{(1-e^{-\lambda x^{\beta}})^{\alpha}} - 1}{\pi - 1} \; ; \qquad \alpha, \beta, \lambda > 0 \tag{7}$$

and

$$f_{PEW}(x) = \frac{\alpha\lambda\beta log\pi}{\pi-1} x^{\beta-1} e^{-\lambda x^{\beta}} \pi^{(1-e^{-\lambda x^{\beta}})^{\alpha}} (1-e^{-\lambda x^{\beta}})^{\alpha-1} \; ; \qquad \alpha, \beta, \lambda > 0. \tag{8}$$

The linear representations for the cdf in (7)is given by (9).

$$F_{PEW} = \frac{1}{\pi-1} \left(\sum_{k=0}^{\infty} a_m e^{-k\lambda x^{\beta}} - 1 \right). \tag{9}$$

Where

$$a_m = \sum_{j=0}^{\infty} (-1)^k \binom{j\alpha}{k} \frac{(log\pi)^j}{j!}.$$

The linear representations for the pdf in (8) is given by (10).

$$f_{PEW} = \sum_{k=0}^{\infty} b_m g(x). \tag{10}$$

Where

$$b_m = \sum_{j=0}^{\infty} \frac{(-1)^k \alpha (log\pi)^{j+1}}{(\pi-1)(k+1)j!} \binom{\alpha(j+1)-1}{k}$$

and

$$g(x) = (k+1)\lambda\beta x^{\beta-1} e^{-(k+1)\lambda x^{\beta}}.$$

Clearly, $g(x)$ is the Weibull distribution with scale parameter $(k+1)\lambda$ and shape parameter β.

The reliability and the hazard rate of PEW distribution for $x > 0$ are given by (11) and (12), respectively

$$R_{PEW}(x) = \frac{\pi}{\pi-1} \left(1 - \pi^{(1-e^{-\lambda x^{\beta}})^{\alpha}-1} \right) \; ; \qquad \alpha, \beta, \lambda > 0 \tag{11}$$

and

$$h_{PEW}(x) = \frac{\alpha\lambda\beta log\pi x^{\beta-1} e^{-\lambda x^{\beta}} (1-e^{-\lambda x^{\beta}})^{\alpha-1}}{\pi^{1-(1-e^{-\lambda x^{\beta}})^{\alpha}} - 1} \; ; \; \alpha, \beta, \lambda > 0. \tag{12}$$

Figure 1 shows some PEW density graphs for various selected parameter values. Figure 2 depicts graphs of the hazard rate of the PEW distribution for different parameter values.

3.1. Simulation and Quantile

The PEW distribution can be simulated using inverse cdf method

$$X = \left\{ -\frac{1}{\lambda} log \left[1 - \left(\frac{log(1+U(\pi-1))}{log\pi} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{1}{\beta}}.$$

Where U is a uniform rv, $0 < u < 1$. The q^{th} quantile of PEW distribution is given by

$$x_q = \left\{ -\frac{1}{\lambda} log \left[1 - \left(\frac{log(1+q(\pi-1))}{log\pi} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{1}{\beta}}$$

The median can be obtained as

$$x_{0.5} = \left\{ -\frac{1}{\lambda} log \left[1 - \left(\frac{log(\frac{1}{2}(\pi+1))}{log\pi} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{1}{\beta}}$$

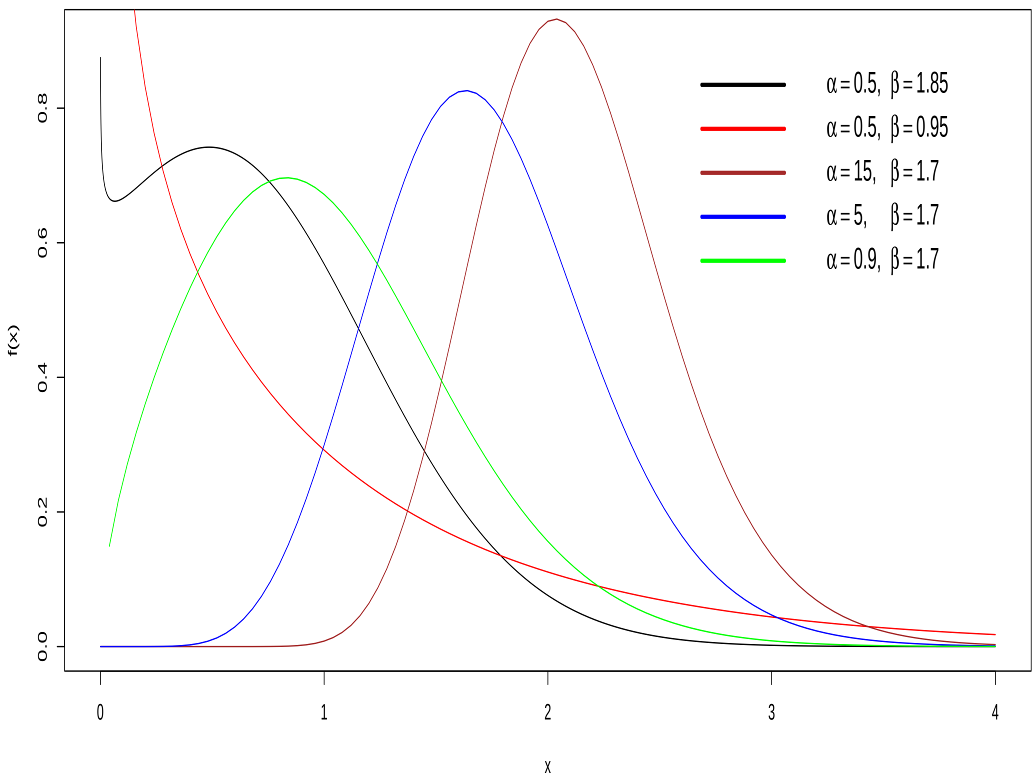


Figure 1: *Density plots of PEW for different combinations of* α, β *and* $\lambda = 1$.

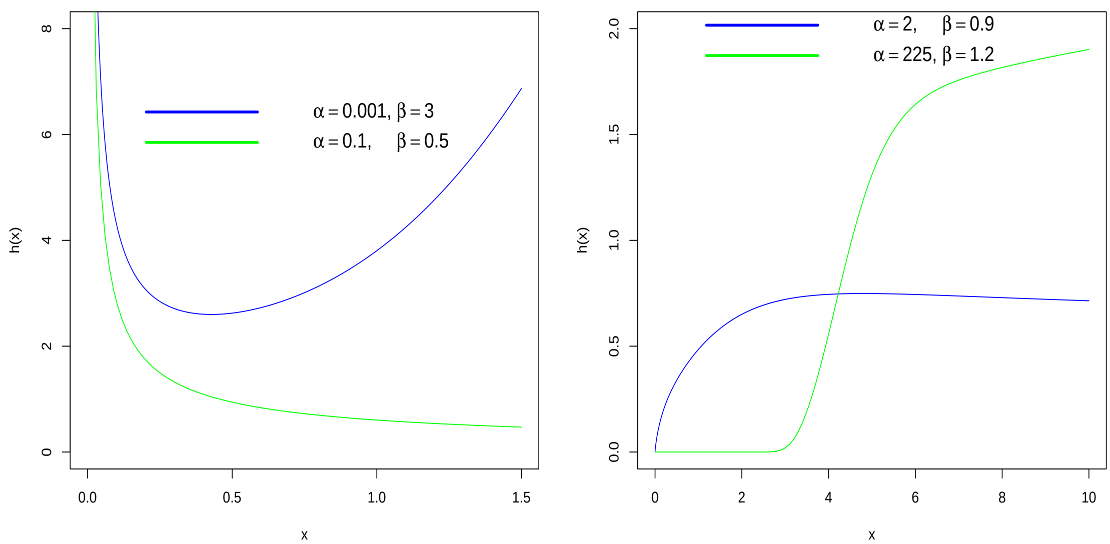


Figure 2: *Hazard rate plots of PEW for different combinations of* α, β *and* $\lambda = 1$.

3.2. Moments and generating function

The r^{th} moment of PEW distribution is obtained by using the following series representation.

$$\alpha^x = \sum_{k=0}^{\infty} \frac{(log\alpha)^k x^k}{k!} \tag{13}$$

$$(1-x)^{b-1} = \sum_{m=0}^{\infty} (-1)^m \binom{b-1}{m} x^m \; ; \qquad |x| < 1, \; b > 0. \tag{14}$$

The r^{th} moment of X can be obtained as

$$\begin{aligned} E(X^r) &= \int_0^{\infty} x^r f(x) dx \\ &= \frac{\alpha\lambda\beta log\pi}{\pi - 1} \int_0^{\infty} x^{r+\beta-1} e^{-\lambda x^\beta} \pi^{(1-e^{-\lambda x^\beta})^\alpha} (1 - e^{-\lambda x^\beta})^{\alpha-1} dx. \end{aligned} \tag{15}$$

Using (13) and (14) in (15), we have

$$E(X^r) = \frac{\alpha\lambda\beta}{\pi - 1} \sum_{a,m=0}^{\infty} \frac{(log\pi)^{a+1}(-1)^m}{a!} \binom{\alpha(a+1)-1}{m} \int_0^{\infty} x^{r+\beta-1} e^{-\lambda(m+1)x^\beta} dx. \tag{16}$$

By applying the transformation $x^\beta = y$ in (16), we get the final expression as

$$E(X^r) = \frac{\alpha}{\pi - 1} \sum_{a,m=0}^{\infty} \frac{(-1)^m (log\pi)^{a+1}}{\lambda^{\frac{r}{\beta}} a! (m+1)^{\frac{r}{\beta}+1}} \binom{\alpha(a+1)-1}{m} \Gamma(\frac{r}{\beta} + 1).$$

The moment generating function of PEW distribution is obtained as

$$M_X(t) = \int_0^{\infty} e^{tx} f(x) dx.$$

By using the same procedure as above, we get the final expression for moment generating function as

$$M_X(t) = \frac{\alpha}{\pi - 1} \sum_{a,l,m=0}^{\infty} \frac{(-1)^m t^l (log\pi)^{a+1}}{\lambda^{\frac{l}{\beta}} l! a! (m+1)^{\frac{l}{\beta}+1}} \binom{\alpha(a+1)-1}{m} \Gamma(\frac{l}{\beta} + 1)$$

3.3. The Mean residual life of PEW distribution

The mean residual life function, say $\mu(t)$ of PEW distribution can be obtained as

$$\mu(t) = \frac{1}{R(t)} \left(E(t) - \int_0^t x f(x) dx \right) - t. \tag{17}$$

Where

$$E(t) = \frac{\alpha}{\pi - 1} \sum_{a,m=0}^{\infty} \frac{(-1)^m (log\pi)^{a+1}}{\lambda^{\frac{1}{\beta}} a! (m+1)^{\frac{1}{\beta}+1}} \binom{\alpha(a+1)-1}{m} \Gamma(\frac{1}{\beta} + 1) \tag{18}$$

and

$$\begin{aligned} \int_0^t x f(x) dx = & \frac{\alpha}{\pi - 1} \sum_{a,m=0}^{\infty} \frac{(-1)^m (log\pi)^{a+1}}{\lambda^{\frac{1}{\beta}} a! (m+1)^{\frac{1}{\beta}+1}} \\ & \times \binom{\alpha(a+1)-1}{m} \gamma \left(\lambda(m+1)t^\beta, \frac{1}{\beta} + 1 \right). \end{aligned} \tag{19}$$

Substituting (11), (18) and (19) in (17), we have

$$\mu(t) = \frac{\alpha}{\pi - \pi^{(1-e^{-\lambda x^{\beta}})^{\alpha}}} \sum_{a,m=0}^{\infty} \frac{(-1)^m (log\pi)^{a+1}}{\lambda^{\frac{1}{\beta}} a!(m+1)^{\frac{1}{\beta}+1}} \binom{\alpha(a+1)-1}{m}$$
$$\times \left[\Gamma\left(\frac{1}{\beta}+1\right) - \gamma\left(\lambda(m+1)t^{\beta}, \frac{1}{\beta}+1\right)\right] - t.$$

Where $\gamma(p,q) = \int\limits_0^p x^{q-1}e^{-x}dx$, is called lower incomplete gamma function.

The mean waiting time $\bar{\mu}(t)$ of PEW distribution, can be obtained as

$$\bar{\mu}(t) = t - \frac{1}{F(t)} \int\limits_0^t xf(x)dx. \tag{20}$$

Substituting (7) and (19) in (20), we get

$$\bar{\mu}(t) = t - \frac{\alpha}{\pi^{(1-e^{-\lambda x^{\beta}})^{\alpha}} - 1} \sum_{a,m=0}^{\infty} \frac{(-1)^m (log\pi)^{a+1}}{\lambda^{\frac{1}{\beta}} a!(m+1)^{\frac{1}{\beta}+1}}$$
$$\times \binom{\alpha(a+1)-1}{m} \gamma\left(\lambda(m+1)t^{\beta}, \frac{1}{\beta}+1\right)$$

3.4. Renyi Entropy

Renyi entropy of PEW distribution, say$RE_X(u)$ can be obtained as

$$RE_X(u) = \frac{1}{1-u} log\left(\int\limits_{-\infty}^{\infty} f(x)^u dx\right); \quad u > 0, \quad u \neq 1.$$
$$= \frac{1}{1-u} log\left(\int\limits_0^{\infty} \left(\frac{\alpha\lambda\beta log\pi}{\pi-1}\right)^u x^{u(\beta-1)} e^{-u\lambda x^{\beta}}\right.$$
$$\left. \times (1-e^{-\lambda x^{\beta}})^{u(\alpha-1)} \pi^{u(1-e^{-\lambda x^{\beta}})^{\alpha}} dx\right). \tag{21}$$

Using (13) in (21), we have

$$RE_X(u) = \frac{u}{1-u} log\left(\frac{\alpha\lambda log\pi}{\pi-1}\right) - log(\beta) + log\left(\sum_{a=0}^{\infty} \frac{(ulog\pi)^a}{a!}\right.$$
$$\left. \times \int\limits_0^{\infty} \beta x^{u(\beta-1)} e^{-u\lambda x^{\beta}} (1-e^{-\lambda x^{\beta}})^{\alpha(a+u)-u} dx\right). \tag{22}$$

Using (14) and applying the transformation $y = x^{\beta}$ in (22) , then the final expression for $RE_X(u)$ is given by

$$RE_X(u) = \frac{u}{1-u} log\left(\frac{\alpha\lambda log\pi}{\pi-1}\right) - log(\beta) + log\left(\sum_{a,m=0}^{\infty} \frac{(-1)^m (ulog\pi)^a}{a!}\right.$$
$$\left. \times \binom{\alpha(a+u)-u}{m} \frac{\Gamma\left(u+\frac{1-u}{\beta}\right)}{(\lambda(m+u))^{u+\frac{1-u}{\beta}}}\right)$$

3.5. Order Statistics

Let $X_1, X_2, ..., X_n$ be a random sample of size n, and let $X_{r:n}$ denote the r^{th} order statistic, then, the pdf of $X_{r:n}$, say $f_{r:n}(x)$ is given by

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} F(x)^{r-1} f(x) (1-F(x))^{n-r}. \tag{23}$$

Substituting (7) and (8) in (23) and using(14), we get

$$\begin{aligned} f_{r:n}(x) =& \frac{\alpha\lambda\beta log\pi}{B(r, n-r+1)} \sum_{a=0}^{n-r} \frac{(-1)^a \binom{n-r}{a}}{(\pi-1)^{a+r}} \left(\pi^{(1-e^{-\lambda x^\beta})^\alpha} - 1 \right)^{a+r-1} \\ & \times x^{\beta-1} e^{-\lambda x^\beta} \pi^{(1-e^{-\lambda x^\beta})^\alpha} (1-e^{-\lambda x^\beta})^{\alpha-1}. \end{aligned}$$

Where $B(a, m)$ is a beta function.

3.6. Stress Strength Reliability

If $X_1 \sim PEW(\alpha_1, \lambda_1, \beta)$ and $X_2 \sim PEW(\alpha_2, \lambda_2, \beta)$, where X_1 and X_2 are independent strength and stress rv's respectively, then, the stress strength reliability $P(X_1 > X_2)$, say SSR, can be obtained as

$$SSR = \int_{-\infty}^{\infty} f_1(x) F_2(x) dx. \tag{24}$$

Using (7) and (8) in (24), we have

$$\begin{aligned} SSR = \int_0^\infty & \left(\frac{\alpha_1\lambda_1\beta log\pi}{(\pi-1)^2} x^{\beta-1} e^{-\lambda_1 x^\beta} \pi^{(1-e^{-\lambda_1 x^\beta})^{\alpha_1}} \right. \\ & \left. \times (1-e^{-\lambda_1 x^\beta})^{\alpha_1-1} \pi^{(1-e^{-\lambda_2 x^\beta})^{\alpha_2}} \right) dx - \frac{1}{\pi-1}. \end{aligned} \tag{25}$$

Using (13), (14) and applying the transformation $y = x^\beta$ in (25), then the final expression for SSR is given by

$$SSR = \frac{1}{\pi-1} \left(\frac{\alpha_1\lambda_1}{(\pi-1)} \sum_{a,b=0}^{\infty} \sum_{m,n=0}^{\infty} \frac{(-1)^{m+n} (log\pi)^{a+b+1}}{a!b!(\lambda_1(1+m)+n\lambda_2)} \binom{\alpha_1(a+1)-1}{m} \binom{b\alpha_2}{n} - 1 \right)$$

4. Estimation

4.1. Maximum Likelihood Estimation

Let $x_1, x_2, ..., x_n$ be a random sample from PEW distribution, then the logarithm of the likelihood function is

$$\begin{aligned} l =& nlog(\alpha\lambda\beta) + nlog\left(\frac{log\pi}{\pi-1}\right) + (\beta-1)\sum_{i=1}^n x_i - \lambda \sum_{i=1}^n x_i^\beta \\ & + log\pi \sum_{i=1}^n \left(1-e^{-\lambda x_i^\beta}\right) + (\alpha-1)\sum_{i=1}^n log\left(1-e^{-\lambda x_i^\beta}\right). \end{aligned} \tag{26}$$

The MLEs of α, λ and β are obtained by partially differentiating (26) with respect to the corresponding parameters and equating to zero, we have

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} log\left(1 - e^{-\lambda x_i^{\beta}}\right)\left(log\pi(1 - e^{-\lambda x_i^{\beta}})^{\alpha} + 1\right) = 0 \tag{27}$$

$$\begin{aligned}\frac{\partial l}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^{n} x_i - \lambda \sum_{i=1}^{n} x_i^{\beta} log x_i \\ &+ \sum_{i=1}^{n} \frac{\lambda x_i^{\beta} log x_i}{e^{-\lambda x_i^{\beta}} - 1}\left(\alpha + \alpha log\pi(1 - e^{-\lambda x_i^{\beta}})^{\alpha} - 1\right) = 0\end{aligned} \tag{28}$$

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i^{\beta} + \sum_{i=1}^{n} \frac{x_i^{\beta}}{e^{-\lambda x_i^{\beta}} - 1}\left(\alpha + \alpha log\pi(1 - e^{-\lambda x_i^{\beta}})^{\alpha} - 1\right) = . \tag{29}$$

Since, the above equations (27), (28) and (29) are not in closed form and are difficult to solve analytically. As a result, it is difficult to calculate the estimates of the parameters α, β and λ. However, R software can be used to solve the equations numerically.

4.2. Simulation study

The simulation study has been conducted using R Software to demonstrate the behaviour of the MLEs in terms of the sample size. Two sets of sample (n=50, n=100) each repeated 1000 times with different combinations of parameters $\lambda = (1, 2)$, $\alpha = (0.5, 1.5, 3)$ and $\beta = (0.5, 1.5, 3, 5)$ were achieved from PEW. In each setting, the average values of MLEs and the corresponding empirical mean squared errors (MSEs) were obtained. The simulation results are presented in tables 1 and 2. Tables 1 and 2 show that the estimates are stable and reasonably close to the true parameter values. As the sample size increases the MSE decreases in all the cases.

5. Applications

In this section, we examine two data sets in order to describe the significance and flexibility of PEW distribution. The first data set has been taken from (Cordeiro and Brito [6]), consist of 48 rock samples from a petroleum reservoir. The dataset corresponds to twelve core samples from petroleum reservoirs that were sampled by four cross-sections. Each core sample was measured for permeability and each cross-section has the following variables: the total area of pores, the total perimeter of pores and shape. We analyze the shape perimeter by squared (area) variable. The observations are: 0.0903296, 0.2036540, 0.2043140, 0.2808870, 0.1976530, 0.3286410, 0.1486220, 0.1623940, 0.2627270, 0.1794550, 0.3266350, 0.2300810, 0.1833120, 0.1509440, 0.2000710, 0.1918020, 0.1541920, 0.4641250, 0.1170630, 0.1481410, 0.1448100, 0.1330830, 0.2760160, 0.4204770, 0.1224170, 0.2285950, 0.1138520, 0.2252140, 0.1769690, 0.2007440, 0.1670450, 0.2316230, 0.2910290, 0.3412730, 0.4387120, 0.2626510, 0.1896510, 0.1725670, 0.2400770, 0.3116460, 0.1635860, 0.1824530, 0.1641270, 0.1534810, 0.1618650, 0.2760160, 0.2538320, 0.2004470.

The second set of data is taken from (Aydin [2]) representing a random sample of average daily wind speed data for March, collected in 2015 from the Turkish Meteorological Services for Sinop, Turkey.The data are recorded as follows
2.8, 1.8, 3.2, 5.0, 2.4, 4.8, 2.9, 2.9, 2.3, 3.2, 2.3, 2.0, 1.9, 3.3, 4.4, 6.7, 4.3, 1.9, 2.2, 3.3, 2.1, 4.0, 2.0, 3.1, 3.8, 3.1, 3.2, 3.4, 2.8, 2.1, 3.1.

We compare the fit of the proposed PEW distribution with its sub-model Weibull (W) (see [20]) and a number of other competing models, namely Alpha Power Weibull (APW) (see [13]), Alpha Power Inverse Weibull (APIW) (see [3]), Modified Weibull (MW) (see [18]), Transmuted

Table 1: *mean values of ML estimates and their corresponding mean square errors(n=50).*

Parameter			MLE			MSE		
λ	α	β	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$
1	0.5	0.5	1.10179	0.50379	0.50421	0.33514	0.01926	0.01998
		1.5	1.09164	0.50369	1.48989	0.29139	0.02022	0.07911
		3	1.09261	0.50384	2.96632	0.29186	0.01920	0.27951
		5	1.10208	0.50390	4.93128	0.33049	0.01722	0.77969
	1.5	0.5	1.10144	1.48760	0.50812	0.34598	0.06202	0.02153
		1.5	1.10143	1.48689	1.49968	0.34594	0.05021	0.05021
		3	1.10492	1.47913	3.07588	0.35744	0.07182	0.30989
		5	1.09812	1.48823	4.97413	0.34323	0.06181	0.92916
	3	0.5	1.06841	2.92528	0.51901	0.34669	0.26838	0.02320
		1.5	1.06708	2.92446	1.53312	0.35017	0.26879	0.11011
		3	1.05536	2.92289	3.06026	0.27696	0.26986	0.39581
		5	1.06038	2.92293	5.08217	0.27756	0.26961	1.06087
2	0.5	0.5	2.05707	0.50408	0.50614	0.58601	0.01735	0.01989
		1.5	2.0553	0.50405	1.49456	0.58328	0.02217	0.07999
		3	2.05262	0.50411	2.97686	0.58133	0.01935	0.28283
		5	2.06155	0.50419	4.94529	0.58885	0.01855	0.76788
	1.5	0.5	2.07548	1.48263	0.51078	0.47035	0.06455	0.02193
		1.5	2.07572	1.48192	1.52755	0.47079	0.06453	0.09875
		3	2.07602	1.48205	3.10344	0.46824	0.06445	0.35757
		5	2.06563	1.48288	5.190317	0.39804	0.06414	0.93515
	3	0.5	2.08168	2.92146	0.51719	0.50571	0.27205	0.02432
		1.5	2.08232	2.92217	1.5266	0.50634	0.27417	0.12172
		3	2.06791	2.92024	3.04978	0.44313	0.27343	0.43695
		5	2.06542	2.92144	5.07546	0.44743	0.27186	1.19552

Table 2: *mean values of ML estimates and their corresponding mean square errors(n=100).*

Parameter			MLE			MSE		
λ	α	β	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$
1	0.5	0.5	1.0506	0.50247	0.50289	0.23614	0.01735	0.01694
		1.5	1.05062	0.50287	1.49297	0.21722	0.01631	0.05371
		3	1.04942	0.50287	2.97447	0.20015	0.01601	0.17807
		5	1.04802	0.50289	4.96282	0.19792	0.01496	0.47121
	1.5	0.5	1.08086	1.49116	0.50075	0.28492	0.05099	0.01989
		1.5	1.09562	1.49818	1.501702	0.26724	0.04504	0.08107
		3	1.08072	1.48961	3.06019	0.27247	0.06647	0.30124
		5	1.09101	1.49015	4.98213	0.28726	0.05136	0.76879
	3	0.5	1.06714	2.95807	0.51053	0.23006	0.19014	0.02037
		1.5	1.06075	2.95628	1.50918	0.19142	0.19102	0.08285
		3	1.04935	2.95675	3.00808	0.19031	0.19078	0.29455
		5	1.04988	2.95522	5.01794	0.14934	0.19213	0.77231
2	0.5	0.5	2.01161	0.50271	0.50106	0.46357	0.01587	0.01686
		1.5	2.01208	0.50251	1.49633	0.46451	0.01733	0.05311
		3	2.01008	0.50253	2.98083	0.46424	0.01524	0.17549
		5	2.01285	0.50246	4.96023	0.45973	0.01634	0.46167
	1.5	0.5	2.01167	1.48439	0.49989	0.39586	0.05389	0.01959
		1.5	2.01133	1.48356	1.49515	0.39526	0.05386	0.07749
		3	2.01735	1.48402	2.98175	0.36287	0.05382	0.27414
		5	2.01093	1.48444	4.92742	0.36888	0.05417	0.72391
	3	0.5	2.06175	2.95508	0.51010	0.31644	0.19261	0.02031
		1.5	2.05964	2.95508	1.50872	0.31542	0.19263	0.08399
		3	2.06125	2.95504	3.00326	0.31317	0.19257	0.29708
		5	2.05041	2.95512	5.00925	0.28678	0.19251	0.77898

Weibull (TW) (see [1]), Odd Weibull (OW) (see [4]), Lindley Weibull (LW) (see [5]), Alpha Power Within Weibull Quantile (APWQ) (see [17]), Marshall Olkin Weibull (MOW) (see [14]) and Alpha Power exponential (APE) ([13]). The corresponding density functions for $x > 0$ are presented in

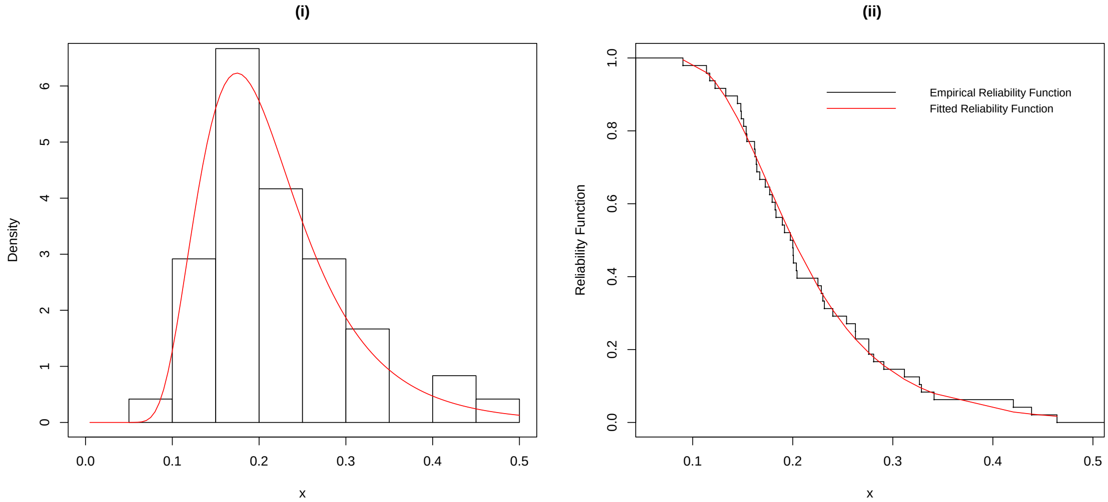


Figure 3: *(i) Fitted PEW density & relative histogram. (ii) Fitted PEW reliability & empirical reliability for first data set.*

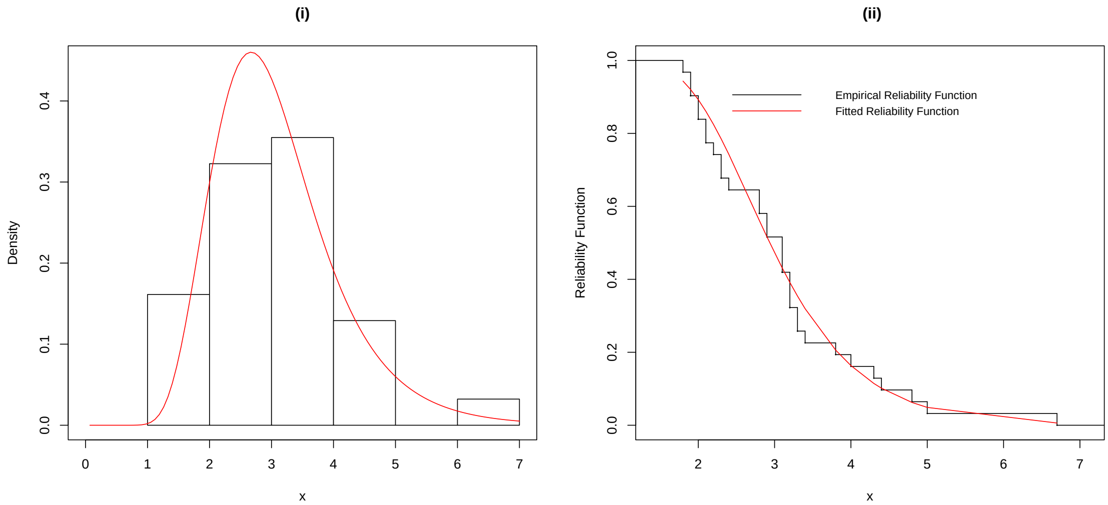


Figure 4: *(i) Fitted PEW density & relative histogram. (ii) Fitted PEW reliability & empirical reliability for second data set.*

the Appendix.

Tables 3, 4, 5 and 6 show that the PEW distribution has the minimum $-2l(\hat{\beta})$, AIC, AICC, BIC and K-S values, as well as the greatest p-value, of all the competing models. As a result, the suggested model fits both the data sets better than the other competitive models. Also the Figures 3, 4, 5 and 6 definitely confirm the conclusions presented in Tables 3, 4, 5,& 6.

6. Conclusion

In this manuscript, a novel method known as PET has been presented. The PET approach has been applied to the Weibull distribution, and a new three-parameter PEW distribution is

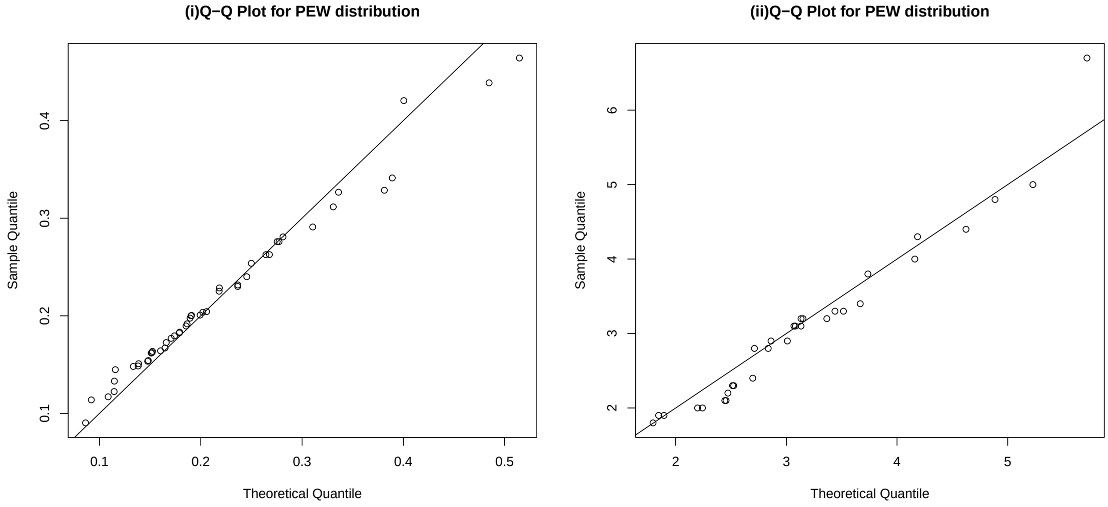


Figure 5: *q-q plot for first and second data set.*

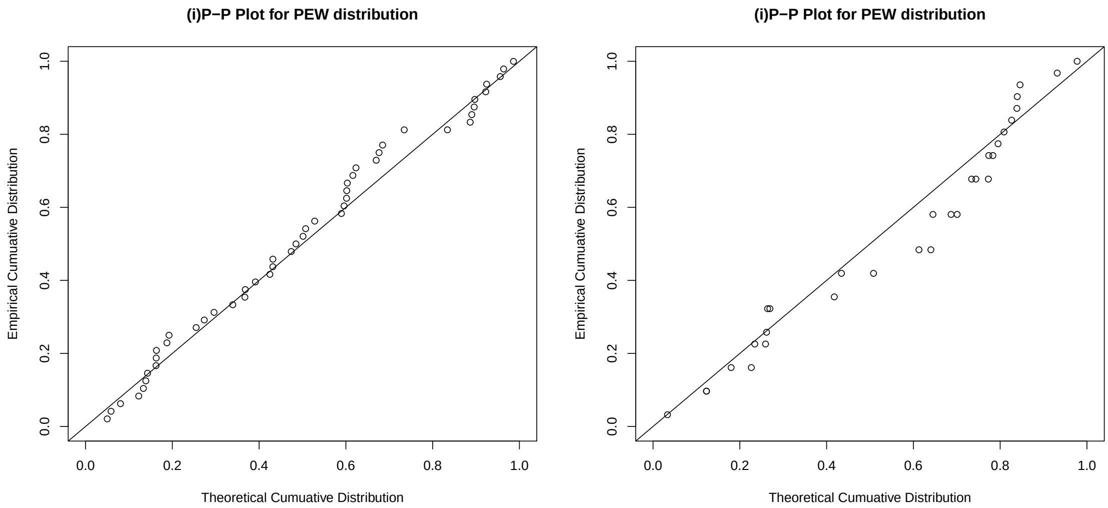


Figure 6: *p-p plot for first and second data set.*

established. Various structural properties as well as reliability measures of the PEW distribution have been highlighted. The reason for adopting this method is that its cdf has a closed form and can represent data with monotone and non-monotone failure rates. It has been revealed that the three-parameter PEW distribution offers more flexibility in respect of hazard rate function and the density function. The suggested model is fitted to two distinct real-life data sets, and the figures demonstrate that it fits both data sets better than any other competing models.

Table 3: *Estimates (standard errors) and kolmogorov smirnov test statistic for the first data set.*

Model	Estimates			Statistics	
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	K-S	p-value
PEW	358.7757 (24.4872)	0.5380 (0.3304)	15.8364 (2.0124)	0.08433	0.8844
APW	0.0320 (0.0508)	3.4096 (0.3606)	4.4898 (2.4429)	0.12804	0.4108
APIW	4.5086 (2.2779)	3.0823 (0.4720)	0.0029 (0.0017)	0.10264	0.6927
MW	0.0010 (2.0711)	2.7475 (0.3700)	47.5555 (7.9292)	0.14985	0.2313
TW	0.6464 (0.2711)	3.0077 (0.3111)	0.2796 (0.0213)	0.14075	0.2976
OW	27.13668 (15.5796)	0.1312 (0.0737)	3.2941 (5.1657)	0.08862	0.8452
LW	17.0146 (22.4843)	2.7406 (0.2854)	1.4788 (1.4712)	0.15011	0.2296
APWQ	64.6499 (9.0106)	6.8937 (0.2609)	65.4380 (0.7298)	0.17289	0.1134
MOW	0.0224 (0.0362)	4.8044 (0.6295)	2.2389 (9.9652)	0.09189	0.8124
APE	100.4597 (16.7779)	-	15.4005 (0.8223)	0.10423	0.6741
W	-	2.7475 (0.2844)	47.5560 (17.9142)	0.14990	0.2310

Table 4: *Information measures for the first data set.*

Model	$-2l(\hat{\beta})$	AIC	AICC	BIC
PEW	-116.4881	-110.4881	-109.9427	-104.8745
APW	-110.56961	-104.56961	-104.02416	-98.95601
APIW	-113.1797	-107.1797	-106.6342	-101.5661
MW	-105.4775	-99.4775	-98.9321	-93.8639
TW	-107.8930	-101.8930	-101.3476	-96.2794
OW	-114.7898	-108.7898	-108.2443	-103.1762
LW	-105.42378	-99.42378	-98.87832	-93.81017
APWQ	-111.7091	-105.7091	-105.1636	-100.0955
MOW	-115.3954	-109.3954	-108.8500	-103.7818
APE	-111.3370	-107.3370	-106.7915	-103.5946
W	-105.48441	-101.48441	-101.21774	-97.74201

Appendix

$$\text{APW} \quad f(x) = \frac{log\alpha}{\alpha-1}\lambda\beta\alpha^{1-e^{-\lambda x^{\beta}}}x^{\beta-1}e^{-\lambda x^{\beta}}$$

$$\text{APIW} \quad f(x) = \frac{log\alpha}{\alpha-1}\lambda\beta x^{-(\beta+\alpha)}e^{-\lambda x^{-\beta}}\alpha^{e^{-\lambda x^{-\beta}}}$$

$$\text{MW} \quad f(x) = (\alpha+\lambda\beta x^{\beta-1})e^{-\alpha x-\lambda x^{\beta}}$$

$$\text{TW} \quad f(x) = \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1}e^{-\left(\frac{x}{\lambda}\right)^{\beta}}\left(1-\alpha+2\alpha e^{-\left(\frac{x}{\lambda}\right)^{\beta}}\right)$$

$$\text{OW} \quad f(x) = \frac{\alpha\beta}{x}\left(\frac{x}{\lambda}\right)^{\beta}e^{\left(\frac{x}{\lambda}\right)^{\beta}}\left(e^{\left(\frac{x}{\lambda}\right)^{\beta}}-1\right)^{\alpha-1}\left[1+\left(e^{\left(\frac{x}{\lambda}\right)^{\beta}}-1\right)^{\alpha}\right]^{-2}$$

$$\text{LW} \quad f(x) = \frac{\beta\alpha^{2}}{\alpha+1}\lambda^{\beta}x^{\beta-1}+\lambda^{2\beta}x^{2\beta-1}e^{-\alpha(\lambda x)^{\beta}}$$

Table 5: *Estimates (standard errors) and kolmogorov smirnov test statistic for the second data set.*

Model	Estimates			Statistics	
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	K-S	p-value
PEW	48.9866 (71.8227)	0.8570 (0.2715)	1.8620 (1.1181)	0.10299	0.8974
APW	0.5344 (2.3730)	1.2214 (1.5742)	7.8545 (0.1492)	0.10759	0.8655
APIW	2.6751 (4.6311)	4.0481 (0.7247)	32.1024 (16.6057)	0.13443	0.6297
MW	0.0010 (0.2108)	2.9427 (0.4632)	0.0255 (0.0249)	0.16492	0.3680
TW	0.7341 (0.2973)	3.2334 (0.4079)	4.0055 (0.3343)	0.14982	0.4897
OW	56.6837 (34.145)	6.9111 (4.0662)	5.8591 (1.8199)	0.10573	0.8789
LW	0.0146 (0.0189)	2.1105 (0.2638)	3.0945 (2.1745)	0.15024	0.4860
APWQ	7.9249 (7.7298)	3.7269 (0.3934)	0.0047 (0.0034)	0.16588	0.3611
MOW	0.0139 (0.0249)	5.3051 (0.8008)	0.1529 (0.0459)	0.10472	0.8859
APE	183.6176 (22.3726)	-	1.0341 (0.7071)	0.12275	0.7385
W	-	2.9413 (0.3668)	0.0256 (0.0140)	0.16544	0.3642

Table 6: *Information measures for the second data set.*

Model	$-2l(\hat{\beta})$	AIC	AICC	BIC
PEW	83.70147	89.70147	90.59036	94.00343
APW	84.90786	90.90786	91.79675	95.20982
APIW	85.51669	91.51669	92.40557	95.81865
MW	92.26848	98.26848	99.15737	102.57044
TW	90.26095	96.26095	97.14984	100.56291
OW	85.10416	91.10416	91.99305	95.40612
LW	89.30294	95.30294	96.19183	99.60490
APWQ	90.44442	96.44442	97.33331	100.74638
MOW	84.89439	90.89439	91.78328	95.19636
APE	88.34464	92.34464	93.23353	95.21262
W	92.19582	96.19582	96.62439	99.06379

$$\text{APWQ} \quad f(x) = \frac{(\alpha-1)\lambda\beta x^{\beta-1}e^{-\lambda x^{\beta}}}{log\alpha\left(1+(\alpha-1)(1-e^{-\lambda x^{\beta}})\right)}$$

$$\text{MW} \quad f(x) = \frac{\alpha\lambda\beta(\lambda x)^{\beta-1}e^{-(\lambda x)^{\beta}}}{1-(1-\alpha)e^{-(\lambda x)^{\beta}}}$$

$$\text{APE} \quad f(x) = \frac{log\alpha}{\alpha-1}\lambda e^{-\lambda x}\alpha^{1-e^{-\lambda x}}$$

where $\alpha,\beta,\lambda > 0$ and $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$ is the gamma function.

Declaration

Conflict of interest: The authors declare that they have no Conflict of interest.

References

[1] Gokarna R Aryal and Chris P Tsokos. Transmuted weibull distribution: A generalization of theweibull probability distribution. *European Journal of pure and applied mathematics*, 4(2):89–102, 2011.

[2] Demet Aydin. The new weighted inverse rayleigh distribution and its application. *Mathemtics and Informatics*, 34(3):511–523, 2019.

[3] Abdulkareem M Basheer. Alpha power inverse weibull distribution with reliability application. *Journal of Taibah University for Science*, 13(1):423–432, 2019.

[4] Kahadawala Cooray. Generalization of the weibull distribution: the odd weibull family. *Statistical Modelling*, 6(3):265–277, 2006.

[5] Gauss M Cordeiro, Ahmed Z Afify, Haitham M Yousof, Selen Cakmakyapan, and Gamze Ozel. The lindley weibull distribution: properties and applications. *Anais da Academia Brasileira de Ciências*, 90:2579–2598, 2018.

[6] Gauss Moutinho Cordeiro and Rejane dos Santos Brito. The beta power distribution. *Brazilian journal of probability and statistics*, 26(1):88–112, 2012.

[7] Nicholas Eugene, Carl Lee, and Felix Famoye. Beta-normal distribution and its applications. *Communications in Statistics-Theory and methods*, 31(4):497–512, 2002.

[8] Anwar Hassan, I. H Dar, and M. A Lone. A novel family of generating distributions based on trigonometric function with an application to exponential distribution. *Journal of Scientific Research*, 65(5), 2021.

[9] Dinesh Kumar, U Singh, and Umesh Singh. Life time distributions: Derived from some minimum guarantee distribution. *Sohag Journal of Mathematics*, 4(1):7–11, 2017.

[10] Nikolay Kyurkchiev. A new transmuted cumulative distribution function based on the verhulst logistic function with application in population dynamics. *Biomath Communications*, 4(1), 2017.

[11] MA Lone, IH Dar, and TR Jan. A new method for generating distributions with an application to weibull distribution. *Reliability: Theory & Applications*, 17(1 (67)):223–239, 2022.

[12] Murtiza Ali Lone, Ishfaq Hassain Dar, and TR Jan. An innovative method for generating distributions: Applied to weibull distribution. *Journal of Scientific Research*, 66(3), 2022.

[13] Abbas Mahdavi and Debasis Kundu. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557, 2017.

[14] Albert W Marshall and Ingram Olkin. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3):641–652, 1997.

[15] Sandeep K Maurya, Arun Kaushik, Rajwant K Singh, Sanjay K Singh, and Umesh Singh. A new method of proposing distribution and its application to real data. *Imperial Journal of Interdisciplinary Research*, 2(6):1331–1338, 2016.

[16] Govind S Mudholkar and Deo Kumar Srivastava. Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, 42(2):299–302, 1993.

[17] M Nassar, A Alzaatreh, O Abo-Kasem, M Mead, and M Mansoor. A new family of generalized distributions based on alpha power transformation with application to cancer data. *Annals of Data Science*, 5(3):421–436, 2018.

[18] Ammar M Sarhan and Mazen Zaindin. Modified weibull distribution. *APPS. Applied Sciences*, 11:123–136, 2009.

[19] William T Shaw and IR Buckley. The alchemy of probability distributions: Beyond gram-charlier & cornish-fisher expansions, and skew-normal or kurtotic-normal distributions. *Submitted, Feb*, 7:64, 2007.

[20] Waloddi Weibull. A statistical distribution function of wide applicability. *Journal of applied mechanics*, 1951.

A COMPREHENSIVE CASE STUDY ON INTEGRATED REDUNDANT RELIABILITY MODEL USING k-out-of-n CONFIGURATION

[1]Srinivasa Rao Velampudi

•

Department of Sciences and Humanities, Raghu Institute of Technology, Visakhapatnam, Andhra Pradesh, India, vsr.rit@gmail.com

[2]Sridhar Akiri *

•

Department of Mathematics, GITAM School of Science, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India, sakiri@gitam.edu

[3]Pavan Kumar Subbara

•

Department of Mathematics, GITAM School of Science, GITAM (Deemed to be University), Bangalore Campus, Karnataka, India, psubbar@gitam.edu

[4]Yadavalli V S S

•

Department of Industrial & Systems Engineering, Pretoria University, 0002 Pretoria, South Africa, sarma.yadavalli@up.ac.za

Abstract

Designers may introduce a system with multiple technologies in series to improve system efficiency. The configuration can be applied to k out of n systems if each technology contains k out of n factors. The k out of n configuration method is successful until every component of the system is successful. The efficiency of the entire system is more in amount than that of a single system factor in a k out of n shape. An Integrated Reliability Model (IRM) for the k out of n, here, an additional system is suggested to account for both the efficiencies of the factors and the number of factors in every phase and the different constraints to optimize the efficiency of the system. To enhance system efficiency, the authors employed the numerous methods of Lagrangean approach to determine the numbers and efficiency of the factors as well as the reliabilities of the phase under different parameters namely load, size, and cost. The dynamic programming approach and simulation method have been adapted to attain an integer result as well as to see the values real.

Keywords: Reliability Theory, IRM, Lagrangean Approach, Stage Efficiency, D P Approach, System Efficiency

1. Introduction

The structure's reliability [1] can be improved by either placing superfluous units, applying the element of greater reliability or by adopting the two methods at a time and both of them use extra resources. Optimizing structure reliability, and conditions to resource availability viz. size, value, load, are examined. In general, reliability is tested as an element of value; But, when tested with real-world problems, the invisible effect of other restraints such as load, size [4], etc. has a special effect on improving structural reliability. The specific functionality of the over-reliability model with several limitations to optimize the recommended setup was examined to maximize the recommended setup. The problem examines the unknowns that is, various elements ($\boldsymbol{X_{ej}}$), the element reliability (r_{ej}), and the stage reliability (R_{sj}) at a specific point for disposing of multiple restraints to magnify the structure reliability that is described as a [14] United Reliability Model (URM). In literature, United Reliability Models [8] are enhanced by applying value restraints where there is a fixed association between value and reliability. A unique pattern of planned work is a deliberation of the load and size as supplementary restraints along with value to form and improve the superfluous reliability system for [15] k out of n structure composition [6, 7].

2. Methods

2.1. Assumptions and Notations:

- Each stage's elements are believed to be identical, i.e., all elements have the same level of reliability.
- All elements are supposed to be statistically independent, meaning that their failure has no bearing on the performance of other elements in the structure.

R_{SR}	=	Structure Efficiency
R_{sj}	=	Efficiency of phase $'sj'$, $0 < R_{sj} < 1$
r_{ej}	=	Efficiency of each component in phase $'ej'$; $0 < r_{ej} < 1$
X_{ej}	=	Number of components in phase $'ej'$
C_{ej}	=	Worth coefficient of each component in phase $'ej'$
L_{ej}	=	Load coefficient of each component in phase $'ej'$
S_{ej}	=	Size coefficient of each component in phase $'ej'$
C_{e0}	=	Greatest allowable structure - Value
L_{e0}	=	Greatest allowable structure - Load
S_{e0}	=	Greatest allowable structure – Size
LMM		Lagrangean Multiplier Method
DPA		Dynamic Programming Approach
IRRM		Integrated Redundant Reliability Model

c_j, d_j, i_j, k_j, m_j, n_j are Constants.

2.2 Mathematical Analysis:

The efficiency of the system to the provided worth function

$$R_{SR} = \sum_{i=1}^{n} B\,(m, i) p^i\,(1 - p)^{m-i} \qquad (1)$$

The following relationship between worth and efficiency is used to calculate the worth coefficient of each unit in phase $'ej'$.

$$r_{ej} = sinh^{-1}\left[\frac{C_{ej}}{f_j}\right]^{\frac{1}{dj}} \qquad (2)$$

Therefore $C_{ej} = f_j \sinh [r_{ej}]^{d_j}$ (2a)

Similarly, $L_{ej} = p_j \sinh [r_{ej}]^{k_j}$ (2b)

$S_{ej} = q_j \sinh [r_{ej}]^{n_j}$ (2c)

Since value-components are linear in ej,

$$\sum_{j=1}^{n} c_{ej}.\ X_{ej} \leq C_{e0} \quad (3a)$$

Similarly load-components and size-components are also linear in ej,

$$\sum_{j=1}^{n} l_{ej}.\ X_{ej} \leq L_{e0} \quad (3b)$$

$$\sum_{j=1}^{n} s_{ej}.\ X_{ej} \leq S_{e0} \quad (3c)$$

Substituting (2) in (3)

$$\sum_{j=1}^{n} f_j \sinh [r_{ej}]^{d_j}.\ X_{ej} - C_{e0} \leq 0 \quad (4a)$$

$$\sum_{j=1}^{n} p_j \sinh [r_{ej}]^{k_j}.\ X_{ej} - L_{e0} \leq 0 \quad (4b)$$

$$\sum_{j=1}^{n} q_j \sinh [r_{ej}]^{n_j}.\ X_{ej} - S_{e0} \leq 0 \quad (4c)$$

The transformed equation through the relation $X_{ej} = \frac{\ln R_{SJ}}{\ln r_{ej}}$ (5)

Where $R_{sj} = \sum_{k=2}^{n} B\,(ej,k)(r_{ej})^k\,(1-r_{ej})^{ej-k}$ (6)

Subject to the constraints

$$\sum_{j=1}^{n} f_j \sinh [r_{ej}]^{d_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - C_{e0} \leq 0 \quad (7a)$$

$$\sum_{j=1}^{n} p_j \sinh [r_{ej}]^{k_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - L_{e0} \leq 0 \quad (7b)$$

$$\sum_{j=1}^{n} q_j \sinh [r_{ej}]^{n_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - S_{e0} \leq 0 \quad (7c)$$

Positivity restrictions $ej \geq 0$

A Lagrangean function is defined as

$$L_F = R_{SR} + δ_1\left[\sum_{j=1}^{n} f_j \sinh [r_{ej}]^{d_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - C_{e0}\right] + δ_2\left[\sum_{j=1}^{n} p_j \sinh [r_{ej}]^{k_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - L_{e0}\right] + δ_3\left[\sum_{j=1}^{n} q_j \sinh [r_{ej}]^{n_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - S_{e0}\right] \quad (8)$$

The Lagrangean function can be used to find the ideal point and separating it by R_{sj}, r_{ej}, $δ_1$, $δ_2$ and $δ_3$.

$$\frac{\partial L_F}{\partial R_{SR}} = 1 + δ_1[\sum_{j=1}^{n} f_j \sinh [r_{ej}]^{d_j}.\frac{1}{\ln r_{ej}}\frac{1}{R_{sj}}] + δ_2\left[\sum_{j=1}^{n} p_j \sinh [r_{ej}]^{k_j}.\frac{1}{\ln r_{ej}}\frac{1}{R_{sj}}\right] + δ_3\left[\sum_{j=1}^{n} q_j \sinh [r_{ej}]^{n_j}.\frac{1}{\ln r_{ej}}\frac{1}{R_{sj}}\right] \quad (9)$$

$$\frac{\partial L_F}{\partial r_{ej}} = δ_1\left[\sum_{j=1}^{n} f_j \sinh [r_{ej}]^{d_j}.\frac{\ln R_{sj}}{\ln r_{ej}}\right]\left[d_j.\ \coth r_{ej} - \frac{1}{r_{ej}.\ln r_{ej}}\right] + δ_2\left[\sum_{j=1}^{n} p_j \sinh [r_{ej}]^{k_j}.\frac{\ln R_{sj}}{\ln r_{ej}}\right]\left[d_j.\ \coth r_{ej} - \frac{1}{r_{ej}.\ln r_{ej}}\right] + δ_3\left[\sum_{j=1}^{n} q_j \sinh [r_{ej}]^{n_j}.\frac{\ln R_{sj}}{\ln r_{ej}}\right]\left[d_j.\ \coth r_{ej} - \frac{1}{r_{ej}.\ln r_{ej}}\right] \quad (10)$$

$$\frac{\partial L_F}{\partial δ_1} = \sum_{j=1}^{n} f_j \sinh [r_{ej}]^{d_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - C_{e0} \quad (11)$$

$$\frac{\partial L_F}{\partial δ_2} = \sum_{j=1}^{n} p_j \sinh [r_{ej}]^{k_j}.\frac{\ln R_{sj}}{\ln r_{ej}} - L_{e0} \quad (12)$$

$$\frac{\partial L_F}{\partial \delta_3} = \sum_{j=1}^{n} q_j \sinh\ [r_{ej}]^{\,n_j} \cdot \frac{\ln R_{sj}}{\ln r_{ej}} - S_{e0} \tag{13}$$

Where δ_1, δ_2 *and* δ_3 are Lagrangean multipliers.

The number of elements in each phase (X_{ej}), the best element reliability (r_{ej}), the phase reliability (R_{sj}) and the structure reliability (R_{SR})are derived by using the Lagrangean method [12]. This method provides a real (valued) solution concerning worth, load, and size.

2.3 Case Problem

To derive the multiple parameters of a given mechanical system using optimization techniques [9], where all the assumptions like value, weight, and volume are directly proportional to system reliability has been considered in this research work. The same logic may not be true in the case of electronic systems. Hence, the optimal element accuracy (r_{ej}), phase reliability (R_{sj}), Number of elements in each phase (X_{ej}), and structure accuracy (R_{SR}) can be evaluated in any given mechanical system [10]. In this work, an attempt has been made to evaluate the Structure accuracy [13] of a special purpose machine that is used for single phase industrial power generators assembly.

The machine is used for the assembly of 3 or 4 components on the base of the power generator. The machine's approximate worth was $3000, which is considered a structure value, the load of the machine is 120 pounds which is the load of the structure, and the space occupied by the machine is 100 cm^3, which is the volume or size of the structure. To attract the authors from different cross sections, the authors attempted to use hypothetical numbers, which can be changed according to the environment.

2.4 Constants

The data required for the constants for the case problem are provided in Table 1.

Table1: *Worth, Load and Size Pre-fixed Constant Values*

Phase	Worth Constants		Load Constants		Size Constants	
	f_j	d_j	p_j	k_j	q_j	n_j
1	2200	0.85	100	0.92	100	0.94
2	2400	0.88	80	0.88	90	0.89
3	2600	0.91	60	0.91	80	0.86

The efficiency of each factor, phase, and number of factors in each stage, as well as the structural efficiency [2, 3], are shown in the tables below.

2.4.1 The Details of Component-Worth Constraint by using Lagrangean Multiplier Method without Rounding-Off

The value-related efficiency design is described in the Table 2.

Table2: *Worth Constraint Analysis by using Lagrangean Multiplier Method*

Phase	f_j	d_j	r_{ej}	$Log\ r_{ej}$	R_{sj}	$Log\ R_{sj}$	X_{ej}	C_{ej}	$C_{ej} . X_{ej}$
01	2200	0.85	0.8741	-0.0584	0.6777	-0.1690	2.89	2233	6456
02	2400	0.88	0.8445	-0.0734	0.6487	-0.1880	2.56	2334	5977
03	2600	0.91	0.8456	-0.0728	0.5461	-0.2627	3.61	2516	9077
Final Worth									21510

2.4.2 The Details of Component-Load Constraint by using Lagrangean Multiplier Method without Rounding-Off

The equivalent results for the load are shown in the Table 3.

Table3: *Load Constraint Analysis by using Lagrangean Multiplier Method*

Phase	p_j	k_j	r_{ej}	$Log\ r_{ej}$	R_{sj}	$Log\ R_{sj}$	X_{ej}	L_{ej}	$L_{ej}.X_{ej}$
01	100	0.92	0.8741	-0.0584	0.6777	-0.1690	2.89	100	290
02	80	0.88	0.8445	-0.0734	0.6487	-0.1880	2.56	78	199
03	60	0.91	0.8456	-0.0728	0.5461	-0.2627	3.61	58	209
Final Load									698

2.4.3 The Details of Component-Size Constraint by using Lagrangean Multiplier Method without Rounding-Off

Equivalent results for size are described in the Table 4.

Table4: *Size Constraint Analysis by using Lagrangean Multiplier Method*

Phase	q_j	n_j	r_{ej}	$Log\ r_{ej}$	R_{sj}	$Log\ R_{sj}$	X_{ej}	S_{ej}	$S_{ej}.X_{ej}$
01	100	0.94	0.8741	-0.0584	0.6777	-0.1690	2.89	100	289
02	90	0.89	0.8445	-0.0734	0.6487	-0.1880	2.56	87	224
03	80	0.86	0.8456	-0.0728	0.5461	-0.2627	3.61	78	282
Final Size									795

3. Efficiency Design by using Lagrangean Multiplier Method

The efficiency design [11] summarizes the $e_j$values as integers (rounding the value of e_j to the nearest integer), and the acceptable outcomes for the worth, load, and size are listed in the tables. Calculate variance due to worth, load, size, and construction capacity (before and after rounding off e_j to the nearest integer) to obtain information.

3.1 Efficiency Design by using Lagrangean Multiplier Method Concerning Worth, Load and Size with Rounding-Off

Table5: *Efficiency design relating to Worth, Load and Size Constraint Analysis by using Lagrangean Multiplier Method with Rounding Off is shown in the following table*

Phase	r_{ej}	R_{sj}	X_{ej}	C_{ej}	$C_{ej}.X_{ej}$	L_{ej}	$L_{ej}.X_{ej}$	$S_{ej}S_{ej}$	$S_{ej}.X_{ej}$
01	0.8741	0.6777	3	2233	6699	100	300	100	300
02	0.8445	0.6487	3	2334	7002	78	234	87	261
03	0.8456	0.5461	4	2516	10066	58	232	78	312
Total Worth, Load and Size				23767		766		873	
Structure Efficiency (R_{SR})								0.9987	

Variation in Worth $= \frac{\text{Total Worth with rounding off} - \text{Total Worth without rounding off}}{\text{Total Value without rounding off}}$ = 10.49%

Variation in Load $= \frac{\text{Total Load with rounding off} - \text{Total Load without rounding off}}{\text{Total Load without rounding off}}$ = 09.72%

Variation in Size $= \frac{\text{Total Size with rounding off} - \text{Total Size without rounding off}}{\text{Total Size without rounding off}}$ = 05.00%

Variation in Efficiency $= \frac{\text{Effciency with rounding off} - \text{Efficiency without rounding off}}{\text{Structure Efficiency without rounding off}}$ = 10.06%.

4. Dynamic Programming Approach

Using the Lagrangean technique [5], which has a number of drawbacks, such as having to provide the amount of components needed at each stage ($X_{ej}('ej')$) in real numbers, which may be difficult to apply. The generally used approach of rounding down the value of results in changes in worth, load, and size, affecting system reliability and having a significant impact on the model's efficiency design. This flaw could be considered, for which the author suggests a substitute empirical implementation that uses the dynamic programming method to obtain an integer solution by using the solutions produced from the Lagrangean approach as parameters for the proposed dynamic programming method.

Table6: *Phase I of the Dynamic Programming*

Phase–I($'ej'$)	Phase - I - Reliability ($\boldsymbol{R_{sj}}$)
01	0.5789
02	0.7183
03	0.8577
04	0.9265
05	0.9605
06	**0.9945**

Table7: *Phase II of the Dynamic Programming*

Phase – II ($'ej'$)	Phase - II - Reliability ($\boldsymbol{R_{sj}}$)							
04	0.4370	0.4490	0.4540					
05	0.4666	0.4916	0.4978	0.5238				
06	0.4962	0.5342	0.5416	0.6125	0.3991			
07	0.5258	0.5768	0.5854	0.7012	0.5496	0.4285		
08	0.5406	0.5981	0.6073	0.7899	0.7571	0.6614	0.6734	0.4523
09	0.5554	0.6194	0.6292	0.8786	0.8936	0.7258	0.7254	0.4835
10	0.5628	0.6407	0.6511	**0.9229**	0.9187	0.9021	0.8852	0.7264

Table8: *Phase III of the Dynamic Programming*

Phase – III ($'ej'$)	Phase - III - Reliability ($\boldsymbol{R_{sj}}$)							
09	0.4932	0.6235	0.6461	0.5938	0.5814	0.4235	0.2824	-
10	0.5383	0.7045	0.7044	0.6451	0.6287	0.4516	0.3125	0.1818
11	0.5824	0.7436	0.7654	0.6821	0.7345	0.5216	0.4514	0.2345
12	0.6265	0.7874	0.7952	0.7514	0.7841	0.6834	0.5387	0.3678
13	0.6706	0.7951	0.8164	0.8354	0.8547	0.7893	0.6868	0.4864
14	0.7147	0.8356	0.8573	0.9125	**0.9571**	0.8514	0.7256	0.6454

5. Results

The application of the Lagrangean multiplier technique yielded a real-valued solution for the suggested Integrated Redundant Reliability Systems for the k-out-of-n configuration mathematical models under investigation, as well as the much-required integer solution. The author obtained new phase reliability ($\boldsymbol{R_{sj}}$) by using a dynamic programming approach. The new values for stage reliability ($\boldsymbol{R_{sj}}$) are (0.9445, 0.9229, and 0.9571). The Dynamic Programming Approach is utilised, and the results for the given mathematical function are outlined in tables 9, 10, and 11 that follow in order to derive the required conclusions.

5.1 The Details of Component-Worth Constraint by using Dynamic Programming Approach

The value-related efficiency design is described in the Table 9.

Table9: *The Details of Component-Worth constraint by using Dynamic Programming Approach*

Phase	f_j	d_j	r_{ej}	$Log\ r_{ej}$	R_{sj}	$Log\ R_{sj}$	X_{ej}	C_{ej}	$C_{ej}\ .X_{ej}$
01	2200	0.85	0.9982	-0.0008	0.9945	-0.0024	3	2580	7740
02	2400	0.88	0.9736	-0.0116	0.9229	-0.0348	3	2735	8205
03	2600	0.91	0.9891	-0.0048	0.9571	-0.0190	4	3016	12064
Final Worth									28009

Mutation in Worth -Component = 30.21%
Mutation in Structure Efficiecy = 01.23%

5.2 The Details of Component-Load Constraint by using Dynamic Programming Approach

The equivalent results for the load are shown in the Table 10.

Table10: *The Details of Component-Load constraint by using Dynamic Programming Approach*

Phase	p_j	k_j	r_{ej}	$Log\ r_{ej}$	R_{sj}	$Log\ R_{sj}$	X_{ej}	L_{ej}	$L_{ej}\ .X_{ej}$
01	100	0.92	0.9982	-0.0008	0.9945	-0.0024	3	117	351
02	80	0.88	0.9736	-0.0116	0.9229	-0.0348	3	91	273
03	60	0.91	0.9891	-0.0048	0.9571	-0.0190	4	70	280
Final Load									904

Mutation in Load-Component = 29.32%
Mutation in Structure Efficiecy = 01.23%

5.3 The Details of Component-Size Constrain by using Dynamic Programming Approach

The equivalent results for size are described in the Table 11.

Table11: *The Details of Component-Size constraint by using Dynamic Programming Approach*

Phase	q_j	n_j	r_{ej}	$Log\ r_{ej}$	R_{sj}	$Log\ R_{sj}$	X_{ej}	S_{ej}	$S_{ej}.X_{ej}$
01	100	0.94	0.9982	-0.0008	0.9945	-0.0024	3	117	351
02	90	0.89	0.9736	-0.0116	0.9229	-0.0348	3	103	309
03	80	0.86	0.9891	-0.0048	0.9571	-0.0190	4	93	372
Final Size									1032
Structure Efficiency (R_{SR})									0.9864

Mutation in Size-Component = 29.81%
Mutation in Structure Efficiecy = 01.23%

5.4 Comparison of Optimization of Integrated Redundant Reliability k out of n systems – LMM with rounding-off and Dynamic Programming approach for Worth

Table12: *Results Correlated LMM with rounding off approach and Dynamic programming approach for Worth*

		With Rounding Off				Dynamic Programming			
Phase	X_{ej}	r_{ej}	R_{sj}	C_{ej}	$C_{ej}.X_{ej}$	r_{ej}	R_{sj}	C_{ej}	$C_{ej}.X_{ej}$
01	3	0.8741	0.6777	2233	6699	0.9982	0.9945	2580	7740
02	3	0.8445	0.6487	2334	7002	0.9736	0.9229	2735	8205
03	4	0.8456	0.5461	2516	10066	0.9891	0.9571	3016	12064
Total Worth		23767				28009			
Structure Efficiency		Using With LMM Approach (R_{SR})			0.9987	Using DP Approach (R_{SR})			0.9999

5.5 Comparison of Optimization of Integrated Redundant Reliability k out of n systems – LMM with rounding-off and Dynamic Programming approach for Load

Table13: *Results Correlated with LMM rounding off approach and Dynamic programming approach for Load*

		With Rounding Off				Dynamic Programming			
Phase	X_{ej}	r_{ej}	R_{sj}	L_{ej}	$L_{ej}.X_{ej}$	r_{ej}	R_{sj}	L_{ej}	$L_{ej}.X_{ej}$
01	3	0.8741	0.6777	100	300	0.9982	0.9945	117	352
02	3	0.8445	0.6487	78	234	0.9736	0.9229	91	273
03	4	0.8456	0.5461	58	232	0.9891	0.9571	70	278
Total Load		767				904			
Structure Efficiency		Using With LMM Approach (R_{SR})			0.9987	Using DP Approach (R_{SR})			0.9999

5.6 Comparison of Optimization of Integrated Redundant Reliability k out of n systems – LMM with rounding-off and Dynamic Programming approach for Size

Table14: *Results Correlated LMM with rounding off approach and Dynamic programming approach for Size*

		With Rounding Off				Dynamic Programming			
Phase	X_{ej}	r_{ej}	R_{sj}	S_{ej}	$S_{ej}.X_{ej}$	r_{ej}	R_{sj}	S_{ej}	$S_{ej}.X_{ej}$
01	3	0.8741	0.6777	100	300	0.9982	0.9945	117	352
02	3	0.8445	0.6487	87	261	0.9736	0.9229	103	307
03	4	0.8456	0.5461	78	312	0.9891	0.9571	93	372
Total Size		873				1031			
Structure Efficiency		Using With LMM Approach (R_{SR})			0.9987	Using DP Approach (R_{SR})			0.9999

6. Discussion

This work proposes an integrated reliability model for a k out of n configuration system with many efficiency criteria. When the data are discovered to be in reals, the Lagrangean multiplier approach is used to compute the number of components (X_{ej}), component efficiencies (r_{ej}), phase efficiencies (R_{sj}), and system efficiency (R_{SR}). To obtain practical applicability, the dynamic way of programming approach is employed to construct an integer solution using the inputs from the Lagrangean method.

This work proposes an integrated reliability model for a k out of n configuration system with many efficiency criteria, when the data are discovered to be in real solution. The Lagrangean multiplier approach is used to compute the number of components (X_{ej}) and the respective component efficiencies (r_{ej}) are 0.8741, 0.8445 & 0.8456, stage reliabilities (R_{sj}) are 0.6777, 0.6487 & 0.5461, and structure reliability (R_{SR}) is 0.9987. To obtain practical applicability, Dynamic programming approach is employed to construct an integer solution whereas component reliabilities (r_{ej}) are 0.9982, 0.9736 & 0.9891, stage reliabilities (R_{SJ}) are 0.9945, 0.9229 & 0.9571, and the system reliability (R_{SR}) is 0.9999, Using the inputs from the Lagrangean method. Finally, we observed that the worth, load and size components changed slightly, but compare with stage reliability, resulting in increased system reliability.

The IRM generated in this manner is quite valuable, particularly in real-world settings when a k from n configuration IRM with reliability engineer redundancy is required. In circumstances where the system value is low, the proposed model is especially valuable for the dependability design engineer to build high-quality and efficient materials.

In future study, the authors recommend utilizing a unique approach that limits the minimum and maximum component reliability values while maximizing system dependability using any of the current heuristic processes to build similar IRMs with redundancy.

References

[1] Agarwal, K. K. and Gupta, J. S. (1975). On minimizing the expense of reliable systems. *IEEE Transactions on Reliability*, 24(3):205-215.

[2] Agarwal, K. K., Mishra, K. B. and Gupta, J. S. (1975). Reliability evaluation a comparative study of different techniques. *Micro Electronics Reliability*, 14(1): 49-56.

[3] Fan, L. T. and Wang, T. (1997). Optimization of system reliability. *IEEE Transactions on Reliability*, 16(2):81-86.

[4] Kumar, S. P., Sridhar, A. and Sasikala, P. (2020). Optimization of Integrated Redundant Reliability Coherent Systems – Heuristic Method. *Journal of Critical Reviews*, 7(5):2480-2491.

[5] Kumar, Y. V. (2004). Optimized a set of integrated reliability models for redundant systems by applying Lagrange-an and Dynamic programming approaches; Ph.D., Thesis, Sri Krishnadevaraya University, India.

[6] Kuo, W. and Prasad, V. R. (2000). An annotated overview of system-reliability optimization. *IEEE Transactions on Reliability*, 49(2):176-187.

[7] Mettas, A. (2000). Reliability allocation and optimization for complex systems - In Annual reliability and maintainability symposium. Proceedings of International symposium on product quality and integrity (Cat. No. 00CH37055). *IEEE Transactions on Reliability*, 216-221.

[8] Mishra, K. B. (1971). A method of solving redundancy optimization problems. *IEEE Transactions on Reliability*, 20(3):117-20.

[9] Mishra, K. B. (1972). Reliability optimization of a series-parallel system. *IEEE Transactions on Reliability*, 21(4):230-238.

[10] Sankaraiah, G., Sarma, B. D., Umashankar, C. and Sarma, Y. V. S. (2011). Designing, modelling and optimizing of an integrated reliability redundant system. *South African Journal of Industrial Engineering*, 22(2):100-106.

[11] Sasikala, P., Sridhar, A. and Kumar, S. P. (2020). Optimization of IRM Series-Parallel and Parallel-Series Redundant System. *Journal of Critical Reviews*, 7(17):1801-1811.

[12] Sasikala, P., Sridhar, A., Kumar, S. P. and Umashankar, C. (2013). Optimization of IRM Parallel-Series Redundant System, *International Journal of Engineering Research & Technology*, 2(2):1-6.

[13] Sridhar, A., Kumar, S. P., Reddy, Y. R., Sankaraiah, G. and Umashankar, C. (2013). The k-out-of-n Redundant IRM Optimization with multiple constraints. *International Journal for Research in Science & Advanced Technologies*, 3(4):436-438.

[14] Sridhar, A., Sasikala, P., Kumar, S. P. and Sarma, Y. V. S. (2021). Design and evaluation of coherent redundant system reliability. Systems Reliability Engineering, edited by Amit Kumar and Mangey Ram, De Gruyter, 137-152.

[15] Sridhar, A., Sasikala, P., Kumar, S. P. and Sarma, Y. V. S. (2022). Design and evaluation of parallel-series IRM system. System Assurances Modeling and Management, edited by Prashant Johri, Adarsh Anand, Juri Vain, Jagvinder Singh, and Mohammad Quasim, Academic Press (Elsevier), 189-208.

WEIBULL COMPARISON BASED ON RELIABILITY, AVAILABILITY, MAINTAINABILITY, AND DEPENDABILITY (RAMD) ANALYSIS

Anas Sani Maihulla

•

Department of Mathematics, Sokoto State University, Sokoto- Nigeria
anas.maihulla@ssu.edu.ng

Ibrahim Yusuf

•

Department of Mathematical sciences, Bayero University, Kano- Nigeria
Iyusuf.mth@buk.edu.ng

Saminu I. Bala

•

Department of Mathematical sciences, Bayero University, Kano- Nigeria
sibala.mth@buk.edu.ng

Abstract

As a continuous probability distribution, the Weibull distribution is widely used in the study of reliability, availability and other life data. In this research, we propose the RAMD analysis to estimate the three-parameter Weibull distribution. The estimation of the distribution parameters is an important problem that has received a lot of attention from researchers because of their effects in several measurements. The real data results indicate that our proposed estimation method is significantly consistent in estimation compared to the RAMD analysis method. The numerical values of filtration system reliability and availability were calculated using Maple software. The system of first-order differential equations is formulated using a mnemonic approach and solved recursively. Several scenarios were examined to determine the impact of the models under consideration. The calculations were done with Maple 13 software. Other reliability measures such as mean time to failure (MTTF), mean time to repair (MTTR), and dependability ratio was estimated. The comparative analysis was conducted using a reverse osmosis (RO) filtration system.

Keywords: Availability, life data, Weibull-distribution, Maple, Reliability, Reverse Osmosis

1. Introduction

A reliability study's main goal is to provide information that can be used to make decisions. Performance Analysis of K out of N Reverse Osmosis (RO) system in the treatment of wastewater using RAMD analysis was carried out by Ibrahim et al. [19]. Based on the analyses, Jun et al.[9] Concluded The key to the life model is the estimating parameters; it can accurately predict the life of a product in terms of reliability. Because obtaining the estimated parameters is difficult, the process of estimating 3-parameter Weibull distribution is important. Iterative methods are required for the estimation process using the maximum likelihood function, which takes a long time and effort. RAMD analysis is the proposed method for the estimation of the 3- parameter weibull distribution. In comparison to the maximum likelihood method, the real data results show that our proposed estimation method is significantly more consistent in estimation. Maihulla A. S. et al. [10] investigated the Reliability, Availability, Maintainability, and Dependability Analysis of a Complex Reverse Osmosis Machine System in Water Purification. The swedish physicist Weibull proposed the Weibull distribution (Walooddi weibull, 1939). He used it to determine the tensile strength of various materials. Since then, it's been widely used in reliability and life testing problems, such as determining a component's time to failure or life length, which is measured from a specified time until it fails as stated by [14].

Jukić and Marković, [8] and Walpole et al., [18] studied the Reliability Estimation of Three Parameters Weibull Distribution based on Particle Swarm Optimization. Basheer et al. [5] demonstrated how to treat the pollutants found in olive mill wastewater (total organic carbon (TOC), dissolved organic carbon (DOC), total phosphorus (TP), total nitrogen (TN), and total polyphenols), a sequential Direct Contact Membrane Distillation (DCMD) and a Reverse Osmosis (RO) hybrid membrane system were used. Similar study was also conducted by Teresa [3] but with recoveries of some economical merits. The influence of permeate flow and pressure on pollutant parameter removals was also included in his study. One of the biggest challenges that humanity must address in the twenty-first century is the scarcity of freshwater. That is why Biniaz et al. [17] investigated potential development of an environmentally friendly, economical, and energy-efficient membrane distillation process. Industrial and household wastes can lessen pollution than to the membrane distillation (MD) method. The Gumbel–Hougaard family copula was used to model the reliability and performance of a solar photovoltaic system was analyzed by Maihulla A. S. et al. [12]. Failure of the purification filter is a serious issue that the water purification industries are facing. Modern technical challenges include removing human-caused contaminants from drinking water. After a brief discussion of the treatment stages that municipal water through before it reaches your tap, you'll learn about the detectable contamination of drinking water caused by anthropogenic (human-made) toxins that is still present. These were all examined by Maihulla et al. [11]. The issue surrounding the comparison of the RAMD analysis and the 3-parameter Weibull was not covered in any of the aforementioned publications.

Our motivation for using a reverse osmosis filtration system as a test bed for the two methods was to see how well they worked together (3-parameter Weibull and RAMD analysis) stems from a serious problem that the water purification industry is facing due to purification filter failure. As a result, there has been a slow advancement in technological advancement in water purification, as well as its importance in the lives of people all over the world. The filtration industry is working hard to keep up with the increasing complexity of the systems. RAMD analysis was used to examine the filtration system's strength, efficiency, and performance improvement, according to the findings of the paper. If the strength, efficiency, and performance of the filtration system are evaluated, users will be able to save money on medical care due to contaminated water. Protect yourself from pollution in the water. The study is divided into five parts, one of which is the current introduction. Modeling of filtration is discussed in the section. Materials and methods are

Covered in Section 2. Section 3 discussed the proposed RAMD analysis method, which was compared to the 3-parameter Weibull, as well as an analytical analysis of the system. Section 4 presented the result. Section 5 included a discussion and explanation of the findings, as well as a conclusion.

1.1. Filtration

One of the most common methods for removing these materials is gravity filtration. This procedure involves passing water containing solid impurities (for example, precipitates after water softening) through a porous material, usually sand and gravel layers. The force of gravity pushes the water through the medium. The gaps between the sand and gravel grains allow small water molecules to flow through. Precipitation-derived solids, on the other hand, become trapped in the pores and thus remain in the porous medium. The solid contaminants have been removed from the water that passes through the bottom of the filter.

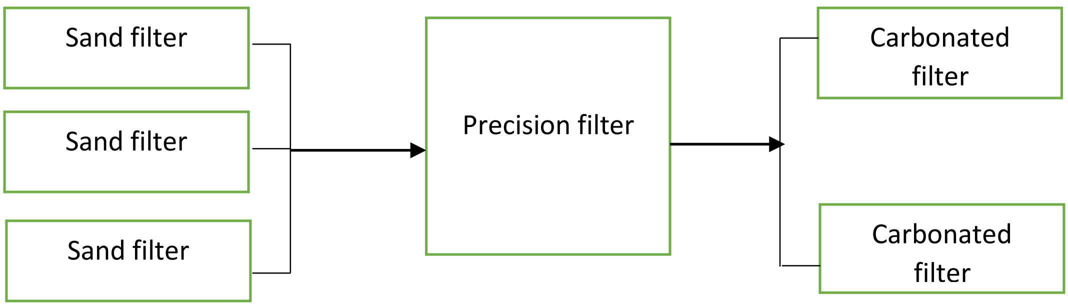


Figure 1: RO Filtration System's Block diagram

1.2. RAMD indices for subsystem 1 (sand filter)

Sand Filter: George et al. [6] stated that, Sand filters are widely used in water purification and work on a completely different mechanism to remove suspended matter. Rather than passing through small orifices through which particles cannot pass, the water passes through a 750 mm deep bed of filter medium, typically 0.75 mm sand.

1.3. RAMD indices for subsystem 2. (Precision filter)

The cylinder shell is usually made of stainless steel, and the internal filter elements are made of PP melt-blown, wire burning, folding, titanium filter, activated carbon filter, and other tubular filter elements, which are chosen based on the different filter media and design process to meet the effluent quality requirements.

1.4. RAMD indices for subsystem 3 (activated carbon filter)

Activated carbon filters are used to purify water without leaving any harmful chemicals behind according to Y. K. Siong et al. [21]. For water treatment, a prototype is being created using activated carbon and an ultraviolet radiation system. Analysis of surface area and porosity. For comparison of surface morphology, scanning electron microscopy (SEM) was used to obtain magnified images of GAC-A and GAC-B.

2. Methods

With the probability density function, let T be a random variable that represents time to failure. f(t), where f(t) is the 3-parameter Weibull distribution's probability density function, which can be written in (1) below, as analyzed by: Jukić and Marković, [8] and Jun et al., [9].

$$f(t;\alpha,\beta,\Upsilon) = \begin{cases} \frac{\alpha}{\beta}\left(\frac{t-\Upsilon}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{t-\Upsilon}{\beta}\right)^{\alpha}\right] & t \geq 0 \\ 0 & t \leq 0 \end{cases} \quad (1)$$

Where $\alpha > 0$ denotes the shape parameter, $\beta > 0$ denotes the scale parameter, and $\Upsilon \leq t$ denotes the location parameter.
Integrating the probability density function yields the cumulative distribution function for the 3-parameter Weibull distribution in (2).

$$F(t) = P(T \leq t) = \int_0^t f(t;\alpha,\beta,\Upsilon)dt = 1 - e^{-\left(\frac{t-\Upsilon}{\beta}\right)^{\alpha}} \quad (2)$$

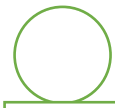
The three-parameter Weibull distribution's mean and variance are (3) and (4) below, as in (Muraleedharan, [14])

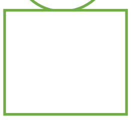
$$\boldsymbol{\mu} = \Upsilon + \beta\Gamma\left(\frac{\alpha+1}{\alpha}\right) \quad (3)$$

Where $\boldsymbol{\mu} = E(t) = MTTF$ (Mean time to failure)

$$\sigma^2 = \beta^2\left[\Gamma\left(\frac{\alpha+2}{\alpha}\right) - \Gamma^2\left(\frac{\alpha+1}{\alpha}\right)\right] \quad (4)$$

2.1. List of notations and definitions

 : Represent the system in working state

 : Represent the system in failed state

P_0 : Represent the initial state of the system working in full capacity state.
P_1 : Represent the state in which one parallel unit is failed
P_2 : Represent the state in which two parallel unit is failed
P_3 : Represent the state in which three parallel unit is failed
α_i $_{i=i,2,3}$ Represent the failure rates subsystems
β_j $_{j=1,2,3}$ Represent the repair rates subsystems
$P_x(t)$ Probability to remain at xth state at time t
$\frac{d}{dt}P_x(t)$, $x = 0,1,2,3$. Represent the derivative with respect to time t

2.2. Reliability

The chance that a device will run without failure for a particular period of time is referred to as reliability under the operational conditions indicated.

$R(t) = \int_t^\infty f(x)dx$ Reliability function (5)

2.3. MTBF

Mean Time between Failures (MTBF): The mean time between failures refers to the average period of good system functioning. The MTBF is the reciprocal of the constant failure rate or the ratio of the test time to the number of failures when the failure rate is reasonably consistent over the operating period. It is given by (6), as in Ashish K. and Monika S. [13].

MTBF = $\int_0^\infty R(t)dt = \int_0^\infty e^{-\theta t} = \frac{1}{\theta}$ Mean Time between Failure (6)

2.4. MTTR

Mean Time between repairs (MTTR): is the reciprocal of the system repair rate, as in (7) below:

MTTR = $\frac{1}{\beta}$ Mean Time to Repair (7)

2.5. Availability

Availability: For repairable systems, availability is a performance criterion that considers both the system's dependability and maintainability. It's defined as the probability that the system will work correctly when it's needed. JG Wohl [20]. That is the ratio of life time to total time, as in (8) below.

Availability = $\frac{Life\ time}{Total\ time} = \frac{Life\ time}{Life\ time+Repair\ time} = \frac{MTTF}{MTTF+MTTR}$ (8)

2.6. Maintainability

Maintainability: When maintenance is necessary, is a design, installation, and operation feature that is often stated as the possibility that a machine can be kept in, or returned to, a particular operational condition within a specified time frame, according to Wohl, JG [20]. It is given by (9) below.

$M(t) = 1 - e^{(-\frac{-t}{MTTR})}$ Maintainability function (9)

2.7. Dependability A. Ebeling [4] defined dependability as a measure of a system's availability, reliability, maintainability. The advantage of dependability is that it allows for cost, reliability, and maintainability comparisons. For random variables with exponential distribution, the dependability ratio is as follows:

$\beta = Repair\ rate$, α = Failure rate

$$d = \frac{\alpha}{\beta} = \frac{MTBF}{MTTR}$$

The importance of maintenance is reflected in the high value of the dependability ratio. The dependability value increases if availability is greater than 0.9 and decreases if availability is less than 0.1, according to C. Li, S. Besarati, and colleagues [2]. The following formula in (10) calculates the minimum value of dependability:

$$D_{min} = 1 - (\frac{1}{d-1})(e^{-\frac{lnd}{d-1}} - e^{-\frac{d\ln d}{d-1}}) \quad (10)$$

2.8. Functions of reliability and failure rate

$R(t)$ Is the reliability function (also known as the survival function). The probability that the device will operate without failure for a given period of time under the specified operating conditions is known as reliability. This included in the study carried out by Haldar and Mahadevan, [7] and

Walpole et al., [18]. The R(t) in (11) below is in terms of the failure and repair rates of the three parameter weibull distribution, unlike that of (5).

$$R(t) = P(T \leq t) = \int_0^t f(t; \alpha, \beta, \Upsilon)dt = 1 - F(t) = e^{-\left(\frac{t-\Upsilon}{\beta}\right)^{\alpha}} \quad (11)$$

$h(t)$ Represents the failure rate (also known as the hazard rate). The failure rate is expressed in terms of failures per unit time. It is computed as the ratio of number of failures of the items undergoing the test time. From $T = t$ to $T = t + \Delta t$, given that it survived to time t, this was investigated by Nachlas, [15] and Walpole et al.,[18].

$$P(t < T < t + \Delta t | T > t) = \frac{P(t<T<t+\Delta t)}{P(T>t)} = \frac{F(t+\Delta t)-F(t)}{R(t)} = \frac{\Delta F(t)}{R(t)} \quad (12)$$

We get the failure rate by dividing this ratio by Δt and finding the limit as Δt→0: in the three parameter weibull distribution, as below in (13). The computation of availability, MTTF, MTTR, MTBF as well as dependability analyses for RAMD against three parameter weibull distribution were presented in table 2.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T < t + \Delta t | T > t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta F(t)}{\Delta R(t)} = \frac{F(t)}{R(t)} \quad (13)$$

And for each correspondent subsystem: by substituting h(t) as failure rate, we got (14) below:

$$MTTF = \frac{1}{Failure\ rate} = \frac{1}{h(t)} \quad \text{(Mean time between failures)} \quad (14)$$

And

$$MTTR = \frac{1}{Repair\ rate} \quad \text{(Mean time to repair)} \quad (15)$$

And therefore the Availability of the system can be calculated using the relation below as in Monika et al. [16]. Availability, also using three parameter weibull distribution can be computed as (16) below:

$$\text{Availability} = \frac{MTBF}{MTBF+MTTR} \quad (16)$$

3. The proposed RAMD Analysis method

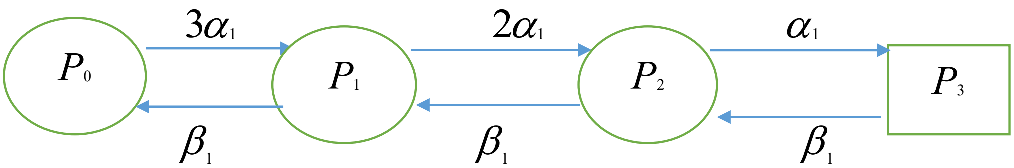


Figure 2: *Transition diagram for subsystem 1*

$$\frac{d}{dt}P_0(t) = -3\alpha_1 P_0 + \beta_1 P_1 \quad (17)$$
$$\frac{d}{dt}P_1(t) = -(2\alpha_1 + \beta_1)\, P_1 + 3\alpha_1 P_0 + \beta_1 P_2 \quad (18)$$
$$\frac{d}{dt}P_2(t) = -(\alpha_1 + \beta_1)\, P_2 + 2\alpha_1 P_1 + \beta_1 P_3 \quad (19)$$
$$\frac{d}{dt}P_2(t) = -\beta_1 P_2 + \alpha_1 P_1 \quad (20)$$

The steady-state probabilities of the system are obtained by imposing the following restrictions: $\frac{d}{dt} \to 0$, as $t \to \infty$. see Arora and Kumar [1].

Under steady state, equation (17) - (20) reduces to

$$P_1 = \frac{\alpha_1}{\beta_1} P_0 \quad (21)$$

Substituting (21) into (18)

$$P_2 = \frac{\alpha_1^2}{\beta_1^2} P_0 \quad (22)$$

Substituting (22) into (19)

$$P_3 = \frac{\alpha_1^3}{\beta_1^3} P_0 \quad (23)$$

Using normalization condition

$$P_0 + P_1 + P_2 + P_3 = 1 \quad (24)$$

Substituting (21) and (22) and (23) into (24) we have:

$$P_0 + = \frac{\alpha_1}{\beta_1} P_0 + \frac{\alpha_1^2}{\beta_1^2} P_0 + \frac{\alpha_1^3}{\beta_1^3} P_0 = 1$$
$$P_0 = \frac{\beta_1^3}{\beta_1^3 + 6\alpha_1^3 + 3\beta_1^2\alpha_1 + 6\,\beta_1\alpha_1^2} \quad (25)$$

Table 2 contains important device output metrics that have been extracted

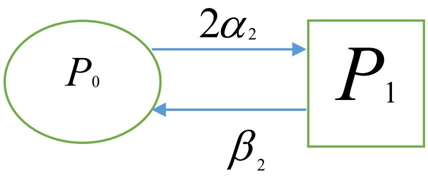


Figure 3: *Transition diagram for subsystem 2*

$$\frac{d}{dt}P_0(t) = -2\alpha_4 P_0 + \beta_4 P_1 \quad (26)$$
$$\frac{d}{dt}P_1(t) = 2\alpha_4 P_0 - \beta_4 P_1 \quad (27)$$

Under steady state, equation (20) and (21) reduces to:

$$-2\alpha_4 P_0 + \beta_4 P_1 = 2\alpha_4 P_0 - \beta_4 P_1$$

And

$$P_1 = \frac{2\alpha_4}{\beta_4} P_0 \quad (28)$$

Using normalization condition

$$P_0 + P_1 = 1 \quad (29)$$

Substituting (22) into (23) we have:

$$P_0 + \frac{2\alpha_4}{\beta_4} P_0 = 1$$
$$P_0 = \frac{\beta_4}{\beta_4 + 2\alpha_4} \quad (30)$$

Table 2 contains important device output metrics that have been extracted.

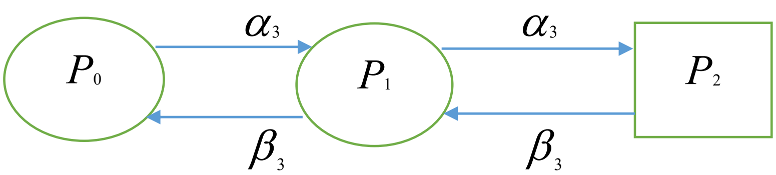


Figure 4: *Transition diagram for subsystem 3*

$$\frac{d}{dt}P_0(t) = -\alpha_3 P_0 + \beta_3 P_1 \quad (31)$$
$$\frac{d}{dt}P_1(t) = -(\beta_3 + \alpha_3)P_1 + \alpha_3 P_0 + \beta_3 P_2 \quad (32)$$
$$\frac{d}{dt}P_2(t) = -\beta_3 P_2 + \alpha_3 P_1 \quad (33)$$

Under steady state, equation (25)- (27) yield

$$P_1 = \frac{\alpha_3}{\beta_3} P_0 \quad (34)$$

Substituting (28) into (26) under steady state we have:

$$P_2 = \frac{\alpha_3^2}{\beta_3^2} P_0 \tag{35}$$

Using normalization condition

$$P_0 + P_1 + P_2 = 1 \tag{36}$$

It follows that:

$$P_0 = \frac{\beta_3^2}{\beta_3^2 + \alpha_3^2 + \beta_3\alpha_3} \tag{37}$$

3.1. System Reliability

$$R_{sys}(t) = R_{s1}(t) \times R_{s2}(t) \times R_{s3}(t) \tag{38}$$

$$R_{sys}(t) = e^{-\alpha_1(t)} \times e^{-\alpha_2(t)} \times e^{-\alpha_3(t)} \tag{39}$$

$$R_{sys}(t) = e^{-(\alpha_1+\alpha_2+\alpha_3)t} \tag{40}$$

3.2. System Availability

Arranged in series, failure of one cause the complete failure of the system.

$$A_{sys} = A_{s1} \times A_{s2} \times A_{s3} \tag{41}$$

$$A_{sys} = \left(\frac{\beta_1^3}{\beta_1^3 + 6\alpha_1^3 + 3\beta_1^2\alpha_1 + 6\,\beta_1\alpha_1^2}\right) \times \left(\frac{\beta_4}{\beta_4 + 2\alpha_4}\right) \times \left(\frac{\beta_3^2}{\beta_3^2 + \alpha_3^2 + \beta_3\alpha_3}\right) \tag{42}$$

$$A_{sys} = \left(\frac{0.07}{0.08005}\right) \times \left(\frac{0.09}{0.11}\right) \times \left(\frac{0.0121}{0.01585}\right) \tag{43}$$

$$A_{sys} = 0.87445 \times 0.81818 \times 0.76341 \tag{44}$$

$$A_{sys} = 0.18249 \tag{45}$$

Table 2 contains important device output metrics that have been extracted.

4. Results

The numerical outcomes from RAMD and Weibull analysis performed using Maple software are presented in this section. The discussion of the comparative analysis in this section is found in the part that follows. The tables' contents are translated into graphs. Table 1 below shows the failure and repair rates for the corresponding subsystems. The comparison findings for the two aforementioned methodologies were shown in Tables 2, 3, and 4.

Table 1: *Failure and repair rates for subsystems*

Subsystem	Failure Rate (α)	Repair Rate (β)
S_1	α_1= 0.004	β_1 = 0.54
S_2	α_2= 0.006	β_2 = 0.67
S_3	α_3= 0.008	β_3 = 0.90

Table 2: *RAMD/Weibull indices for the RO filtration system*

RAMD indices of Subsystems	Subsystem S_1	Subsystem S_2	Subsystem S_3	System
Reliability	$e^{-0.009t}$	$e^{-0.005t}$	$e^{-0.014t}$	$e^{-0.028t}$
Availability (Weibull)	0.96755	0.77853	0.89706	0.67573
Availability (RAMD)	0.84556	0.69453	0.77453	0.45486
MTBF (Weibull)	104.34	198.92	88.06	391.32
MTBF(RAMD)	83.33 21.01	166.67 32.25	62.50 25.56	312.50 78.82
MTTR (Weibull)	4.77	3.21	1.96	9.94
MTTR(RAMD)	1.85	1.49	1.11	4.45
Dependability (Weibull)	0.99302	0.99746	0.99753	0.98805
Dependability (RAMD)	0.97497	0.94301	0.92891	0.85405
Dependability ratio (Weibull)	21.87	61.97	44.93	
Dependability ratio (RAMD)	45.04	111.86	56.31	

Table 3: *Availability of a system*

Subsystems	3-Parameter weibull distribution	RAMD (Exponential distribution)
1	0.96755	0.84556
2	0.77853	0.69453
3	0.89706	0.77453

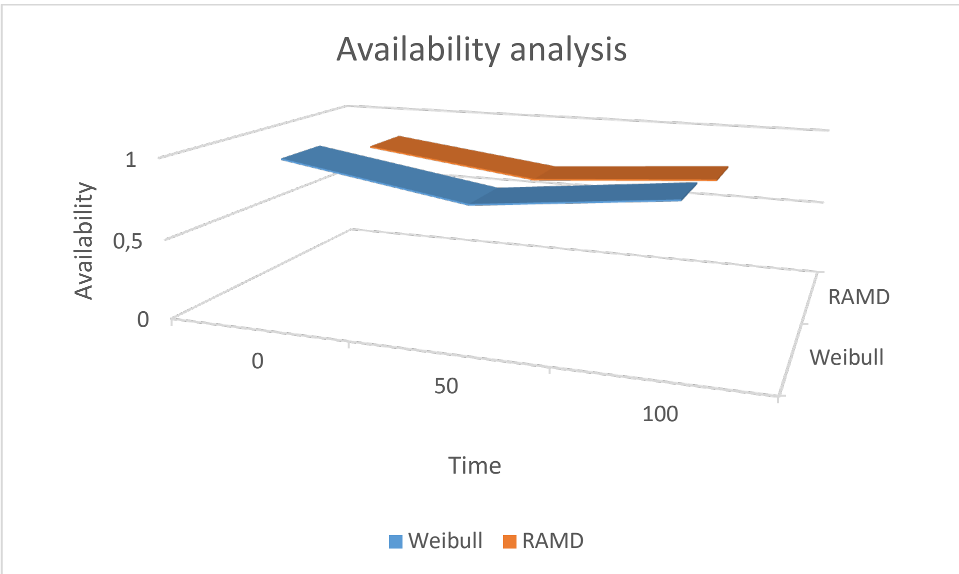


Figure 5: *Variation of Availability with time*

Table 4: *Reliability of a system*

Time(in days)	3-Parameter weibull distribution	RAMD (Exponential distribution)
0	1.00000	0.99999
10	0.79918	0.79384
20	0.71540	0.70030
30	0.66808	0.64285
40	0.63288	0.59827
50	0.60240	0.55945
60	0.57436	0.52410
70	0.54801	0.49140
80	0.52308	0.46100
90	0.49942	0.43268
100	0.47695	0.40628

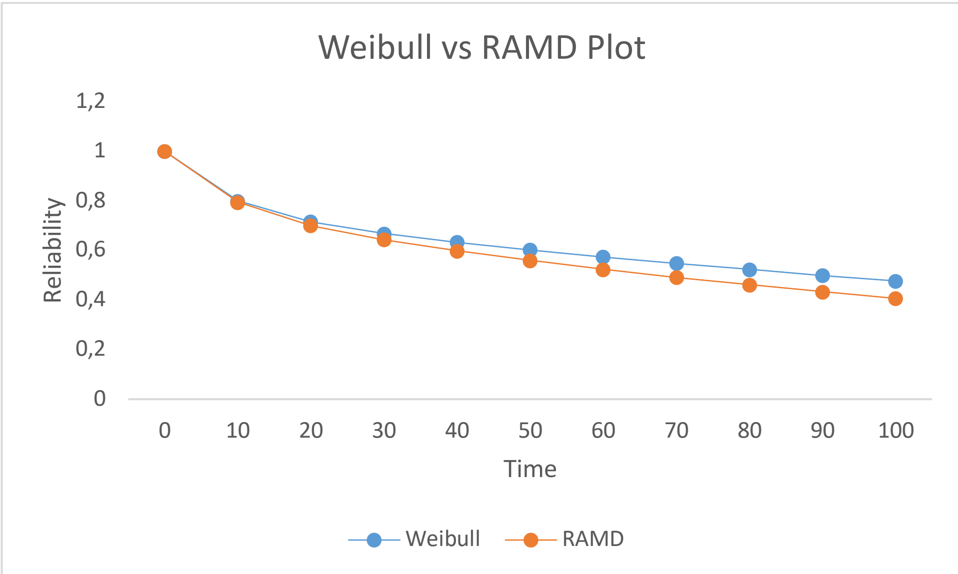


Figure 6: *Variation of Reliability with time*

.

5. Discussion

Figures 2, 3, and 4 show the transition diagrams for the three reverse osmosis filtering units/subsystems, namely the sand filters, precision filters, and carbonated filters. A comparison of the efficiency of two methods (RAMD analysis and 3-parameter Weibull) was conducted. Table 2 lists all of the additional RAMD-Weibull metrics. It demonstrated how effective the Weibull is when compared to the RAMD. The variation in the system's availability is 0.22087, which is due to the fact that subsystem 3 has a larger variation than subsystems 1 and 2. When the MTBF of the entire system is considered, it is found that the weibull is more efficient in terms of the average operating time from the time a failed device is restored to the time it becomes failed again than RAMD analysis, with a variation of nearly 78.82. When compared to 1 and 3, the same subsystem 2, contributed significantly. Despite the fact that it had fewer units than 1 and 3, it was still a good choice. Using the RAMD analysis, the system's reliability after 50 days of operation is only 0.55945. The reliability of the same system using a 3-parameter Weibull distribution after the same period is 0.60240.

Figure 5 depicted the availability of the two methods as a function of time (3-parameter Weibull and RAMD analysis). The efficiency of 3-parameter Weibull over RAMD analysis is clearly observed. According to the numerical analysis in Table 4 and the corresponding values in Figure 6.

6. Conclusion

For desalinating saltwater to produce drinking water, reverse osmosis (RO) filtration system is an important technology. The performance of a desalination system is determined by the failure behavior of its components. Because the RO filtration system was designed to be low-power, the subsystems' dependability must be maintained through proper design and material selection in order for the plant to operate continuously. Furthermore, the estimation of the system's strength in terms of reliability characteristics using the 3-parameter Weibull distribution is more efficient than the RAMD analysis method. As a result, the study investigated the system's availability and reliability, as well as other features such as MTBF, MTTR, and dependability analysis. To address the issue of precision filter redundancy, more research can be done.

7. Acknowledgement

We sincerely appreciate the distinguished Sokoto State University. This is because I was given a study leave and other benefits to allow me to finish my PhD in mathematics. This research fits inside that.

References

[1] Arora N, Kumar D (1997) Availability analysis of steam ans power generation system in thermal power plant. Macroelectron Reliab 37(5):795-799

[2] C. Li, S. Besarati, Y. Goswami, E. Stefanakos, H. Chen, (2013) "Reverse osmosis desalination driven by low temperature supercritical organic Rankine cycle". Appl. Energy, Vol. 102 Pp 1071–1080.

[3] D. Teresa Sponza (2021). "Treatment of Olive Mill Effluent with Sequential Direct Contact Membrane Distillation (DCMD)/Reverse Osmosis (RO) Hybrid Process and Recoveries of Some Economical Merits" Res Appl. 2021 vol. 1(1): 001-010. Advanced Journal of Physics Research and Applications. Vol. 1(1): 001-010.

[4] Ebeling A (2000) An introduction to reliability and maintainability engineering. Tata Mcgraw Hill Company Ltd, New Delhi

[5] G. T. Basheer, Zakariya and Y Algamal. (2021). Reliability Estimation of Three Parameters Weibull Distribution based on Particle Swarm Optimization. *Pakistan journal of statistics and operational research*. DOI: http://dx.doi.org/10.18187/pjsor.v17i1.2354 Pak.j.stat.oper.res. Vol.17 No. 1 2021 pp 35-42.

[6] George Solt CEng, FIChemE, (2002) in Plant Engineer's Reference Book (Second Edition).

[7] Haldar A., Mahadevan S., (2000). Probability, Reliability and Statistical Methods in Engineering Design. John Wiley and Sons, Inc.

[8] Jukić D., Marković D, (2010). On nonlinear weighted total least squares parameter estimation problem for the three-parameter weibull density. Applied Mathematical Modelling, 34, 1839-1848.

[9] Jun W.C, Xin L, Xin C., (2017). Parameter Evaluation of 3-parameter weibull Distribution based on Adaptive Genetic Algorithm. 2nd International Conference on Machinery, Electronics and Control Simulation, Advances in Engineering Research, volume 138.

[10] Maihulla, A. S., Yusuf, I., and Bala, S. I. (2021). Performance Evaluation of a Complex Reverse Osmosis Machine System in Water Purification using Reliability, Availability, Maintainability and Dependability Analysis. *Reliability: Theory & Applications*, *16*(3), 115-131.

[11] Maihulla, A. S., & 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.

[12] 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

[13] Monika S. and Ashish K. (2019). "Performance analysis of evaporation system in sugar industry using RAMD analysis" journal of Brazilian society of mechanical science and engineering Vol. 41. (175). https://doi.org/10.1007/s40430-019-1681-3

[14] Muraleedharan G., (2013). Characteristic and Moment Generating Functions of Three Parameter Weibull Distribution-an Independent Approach. Research Journal of Mathematical and Statistical Sciences, Vol.1(8), 25-27.

[15] Nachlas J.A., (2005). Reliability Engineering Probabilistic Models and Maintenance Methods. Taylor & Francis Group.

[16] Nivedita Gupta, Monika Saini and Ashish Kumar (2020). Behavioral Analysis of Cooling Tower in Steam Turbine Power Plant using Reliability, Availability, Maintainability and Dependability Investigation. Journal of Engineering Science and Technology Review 13 (2) (2020) 191 – 198. doi:10.25103/jestr.132.23

[17] P. Biniaz, N. T. Ardekani, M. A. Makarem and M. R. Rahimpour (2019). "Water and Wastewater Treatment Systems by Novel Integrated Membrane Distillation (MD)" ChemEngineering, vol. 3, no. 8

[18] Walpole R.E., Myers R. H., Myers S.L. and Ye K., (2012). Probability and statistics for Engineers and Scientists. Pearson Education, Inc.

[19] Walpole R.E., Myers R. H., Myers S.L. and Ye K., (2012). Probability and statistics for Engineers and Scientists. Pearson Education, Inc.

[20] Wohl JG (1996) system operational readiness and equipment dependability. IEEE Trans Reliability 15 (1): 1-6.

[21] Y. K. Siong , J. Idris a , M. Mazar Atabaki , (2013) "Performance of Activated Carbon in Water Filters" See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/234060484 Article in Water Resources.

New Cosine-Generator With an Example of Weibull Distribution: Simulation and Application Related to Banking Sector

Aijaz Ahmad*

•

Department of Mathematics, Bhagwant University, Ajmer, India
aijazahmad4488@gmail.com

Muzamil jallal

•

Department of Mathematics, Bhagwant University, Ajmer, India
muzamiljallal@gmail.com

Sh.A.M. Mubarak

•

The Higher Institute of Engineering and Technology, Ministry of Higher Education, El-Minia, Egypt
Shaimaa.Mubarak@mhiet.edu.eg

Abstract

In this work, we propose a novel trigonometric-based generator entitled the "New Cosine-Generator" to acquire elevated distribution adaptability. This generator is formed without the insertion of extra parameters. Adopting the Weibull distribution as the baseline distribution, and this distribution is referred to as the New Cosine-Weibull Distribution. Several statistical features of the investigated distribution were studied, including moments, moment generating functions, order statistics, and reliability measures. For different parameter values, a graphical representation of the probability density function (pdf) and the cumulative distribution function (cdf) is provided. The distribution's parameters are determined using the well-known maximum likelihood estimation approach. Finally, simulation analysis and an application is used to evaluate the effectiveness of the distribution.

Keywords: Cosine function, moments, maximum likelihood estimation, reliability indicators, simulation.

1. Introduction

In applied sciences such as biological sciences, medical sciences, environmental sciences, engineering, finance, and actuarial science, among others, statistical evaluation of lifetime data is unpredictable, and statistical modelling is the finest and most effective technique to examine the ambiguity of any occurrence. Because of the complex nature and distinctive characteristics of data, life time data serves a critical function in sectors such as insurance and finance. Thus, there is an apparent need for expansion and modification of current traditional statistical distributions. Indeed, various initiatives have been made to develop additional classes of lifetime distributions in order to extend various families of distributions and offer more adaptability to the novel model. Numerous investigators have added new classes of life time distributions throughout the last few years, which are now available in the statistical literature. The implementation of trigonometric functions to construct new statistical distributions is becoming a prevalent approach, and employing these trigonometric distributions for data interpretation demonstrates greater versatility. Looking back over the literature, we can see that numerous authors employed many generators or transformations. For instance, Eugene et al. [16], the gamma-G family by Zagrofos and Balakrishana [20], the transformedtransformer(T-X) by Alzaatrah et al [1], the Weibull-G by Bourguignon et al. [6], Morad Alizadeh et al. [17] constructed the Gompertz-G distribution family, Brito et al. [7], formulated the Topp-Leone odd log-Logistic family of distributions and Aijaz et al. [4] a noval aproach for constructing distributions with an example of Rayleigh distribution, SS-transformation based on trigonometric functions is proposed by kumar et al. [2], Chesneau et al.[10], Mahmood. Z and Chesneau. C [18], Souza.l et al.[13], Jammal.F et al. [11], M.A.Lone et al.[19],I.H. Dar et al [5] and Aijaz Ahmad et al. [3]. This work aims to present the cosine-generator distributions, a novel family of trigonometric function-based generator. The benefit of this generator is that flexibility is achieved without the insertion of further parameters.

Let us suppose $F(x;\zeta)$ be cdf of a random variable X, then the cumulative distribution function of new cosine-generator family of distributions is described as.

$$\begin{aligned} F(x;\zeta) &= -\int_0^{\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}} sinxdx \\ &=1-cos\left(\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}\right) \quad ; \quad x\in\mathbb{R},\zeta>0 \end{aligned} \tag{1}$$

The related probability density function of equation (1) is stated as

$$f(x;\zeta)=\frac{\pi}{2}log(2)2^{\bar{G}(x;\zeta)}g(x;\zeta)sin\left(\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}\right); \quad x\in\mathbb{R},\zeta>0 \tag{2}$$

Where $\bar{G}(x;\zeta))=1-G(x;\zeta)$ and $\frac{dG(x;\zeta)}{dx}=g(x;\zeta)$.
Futhermore, the reliability function represented as $R(x;\zeta)$, hazard rate function represented as $H(x;\zeta)$ and reverse hazard rate function represented as $h(x;\zeta)$ are respectively stated in general form by

$$\begin{aligned} R(x;\zeta) &=1-F(x;\zeta)=cos\left(\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}\right) \\ H(x;\zeta) &=\frac{f(x;\zeta)}{R(x;\zeta)}=\frac{\pi}{2}log(2)2^{\bar{G}(x;\zeta)}g(x;\zeta)tan\left(\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}\right) \\ h(x;\zeta) &=\frac{f(x;\zeta)}{F(x;\zeta)}=\frac{\frac{\pi}{2}log(2)2^{\bar{G}(x;\zeta)}g(x;\zeta)sin\left(\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}\right)}{1-cos\left(\frac{\pi(2-2^{\bar{G}(x;\zeta)})}{2}\right)} \end{aligned}$$

The Weibull distribution has been employed in a wide range of domains and applications. The hazard function of the Weibull distribution can only be monotone. As a consequence, it cannot be used to replicate lifespan data with a bathtub-shaped hazard function. We adopt the Weibull distribution as the baseline distribution for the newly formed generator and exhibit its numerous characteristics.
Suppose X denotes a random variable that follows the Weibull distribution, then its cumulative distribution function is stated as

$$G(x;\alpha,\beta)=1-e^{-\alpha x^{\beta}}; \quad x>0,\alpha,\beta>0 \tag{3}$$

The related probability density function is given as

$$g(x;\alpha,\beta) = \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}}; \quad x > 0, \alpha, \beta > 0 \tag{4}$$

1.1. Usefull expansion

We apply Taylor's series of $sinx = \sum_{p=0}^{\infty}(-1)^p \frac{x^{2p+1}}{(2p+1)!}$ in equation (2) to get its mixture form

$$f(x;\zeta) = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!}\frac{\pi}{2}log(2)g(x;\zeta)2^{\bar{G}(x;\zeta)}2^{2p+1}\left(1-2^{-G(x;\zeta)}\right)^{2p+1} \tag{5}$$

Now, we apply $(1-u)^a = \sum_{q=0}^{\infty}(-1)^q\binom{a}{q}u^q$ in equation (5), we obtain

$$f(x;\zeta) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{(-1)^{p+q}}{(2p+1)!}log(2)\binom{2p+1}{q}\pi^{2p+2}g(x;\zeta)2^{-(q+1)G(x;\zeta)} \tag{6}$$

Again we apply Taylor's series of $a^y = \sum_{r=0}^{\infty} \frac{(log(a))^r}{r!}y^r$in equation (6), we have

$$f(x;\zeta) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty} \frac{(-1)^{p+q+r}}{(2p+1)!}\frac{(log(2))^{r+1}}{r!}\binom{2p+1}{q}\pi^{2(p+1)}(q+1)^r g(x;\zeta)\left(G(x;\zeta)\right)^r \tag{7}$$

Equations (3) and (4) in equation (7) enable us to construct the baseline model's probability density function in mixture form, which has been employed as an illustration for the established generator.

$$f(x;\alpha,\beta) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\alpha\beta x^{\beta-1}e^{-\alpha(r+1)x^{\beta}} \tag{8}$$

$$\Psi_{pqrs} = \frac{(-1)^{p+q+r+s}}{(2p+1)!}\frac{(log(2))^{r+1}}{r!}\binom{2p+1}{q}\binom{r}{s}\pi^{2(p+1)}(q+1)^r$$

2. New Cosine-Weibull Distribution and its Mathematical properties

By applying the new cosine-generator, we exhibit the probability density function (pdf) and cumulative distribution function (cdf) of a newly formed distribution called new Cosine-Weibull distribution (NCWD) in this part and strengthen certain of its mathematical features. Using equation (3) in equation(4), we obtain the cdf of desired distribution as

$$F(x;\alpha,\beta) = 1 - cos\left(\frac{\pi(2-2^{e^{-\alpha x^{\beta}}})}{2}\right) \quad ; \quad x > 0, \alpha, \beta > 0 \tag{9}$$

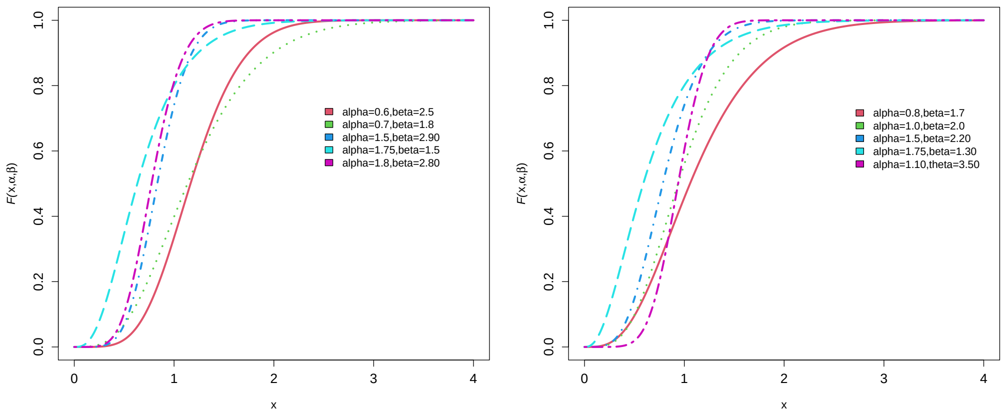


Figure 1: *cdf plot of NCWD for different choise of parameters*

The related probability density function is stated as

$$f(x;\alpha,\beta) = \frac{\pi log(2)}{2}\alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}} sin\left(\frac{\pi(2-2^{e^{-\alpha x^{\beta}}})}{2}\right) \quad ; \quad x>0,\alpha,\beta>0 \tag{10}$$

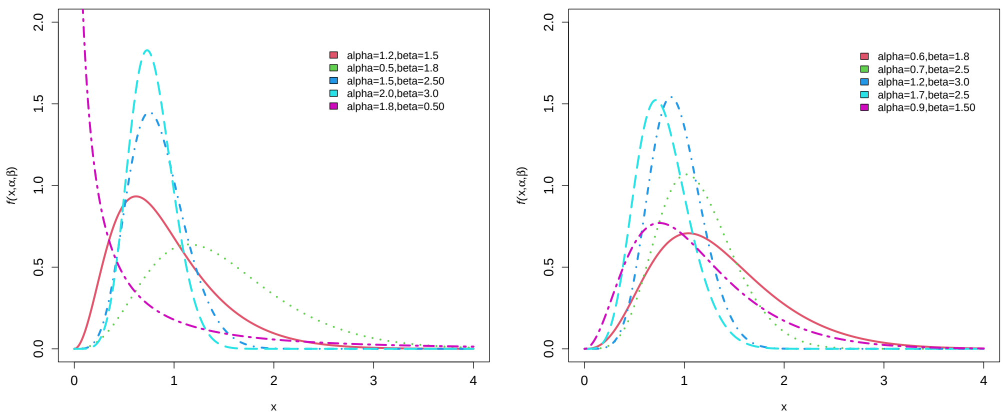


Figure 2: *pdf plot of NCWD for different choise of parameters*

2.1. Moments of new cosine-Weibull distribution

Let x denotes a random variable, then the k^{th} moment of NCWD is denoted as μ_k' and is given by

$$\mu_k' = E(x^k) = \int_0^{\infty} x^k f(x;\alpha,\beta)dx \tag{11}$$

Using equation (8) in equation (11), it yields

$$\mu_k' = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\alpha\beta\int_0^{\infty} x^{k+\beta-1}e^{-\alpha(r+1)x^{\beta}}dx$$

Making substitution $\alpha(r+1)x^{\beta} = z$, sothat $0 < z < \infty$, we have

$$\mu_k^{'} = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\frac{\alpha}{(\alpha(r+1))^{\frac{k+\beta}{\beta}}}\int_0^{\infty} z^{\frac{k}{\beta}}e^{-z}dz$$

After solving the integral, we obtain

$$\mu_k^{'} = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\frac{\alpha}{(\alpha(r+1))^{\frac{k+\beta}{\beta}}}\Gamma\left(\frac{k+\beta}{\beta}\right)$$

Where $\Gamma(.)$ denotes the gamma function.

2.2. Moment generating function of new cosine-Weibull distribution

Suppose x denotes a random variable follows NCWD. Then the moment generating function of the distribution denoted by $M_X(t)$is given

$$\begin{aligned}
M_X(t) = E(e^{tx}) &= \int_0^{\infty} e^{tx} f(x;\alpha,\beta)dx \\
&= \int_0^{\infty}\left(1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + ...\right) f(x;\alpha,\beta)dx \\
&= \sum_{k=0}^{\infty}\frac{t^k}{k!}\int_0^{\infty} x^k f(x;\alpha,\beta)dx \\
&= \sum_{k=0}^{\infty}\frac{t^k}{k!}E(x^k) \\
&= \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{k=0}^{\infty}\Psi_{pqrs}\frac{\alpha t^k}{k!\,(\alpha(r+1))^{\frac{k+\beta}{\beta}}}\Gamma\left(\frac{k+\beta}{\beta}\right)
\end{aligned}$$

2.3. Incomplete moments of new cosine-Weibull distribution

The v^{th} incomplete moment for density function in general is stated as

$$I(v) = \int_0^{v} x^k f(x;\alpha\beta)dx$$

Using the equation (8), we have

$$I(v) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\alpha\beta\int_0^{v} x^{k+\beta-1}e^{-\alpha(r+1)x^{\beta}}dx$$

Making substitution $\alpha(r+1)x^{\beta} = z$, sothat $0 < z < \alpha(r+1)v^{\beta}$, we have

$$I(v) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\frac{\alpha}{(\alpha(r+1))^{\frac{k+\beta}{\beta}}}\int_0^{\alpha(r+1)v^{\beta}} z^{\frac{k}{\beta}}e^{-z}dz$$

After solving the integral, we obtain

$$I(v) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\frac{\alpha}{(\alpha(r+1))^{\frac{k+\beta}{\beta}}}\gamma\left(\frac{k+\beta}{\beta},\alpha(r+1)v^{\beta}\right)$$

Where $\gamma(a,x) = \int_0^x u^{a-1}e^{-u}du$ denotes lower incomplete gamma function

2.4. Quantile function of new cosine-Weibull distribution

The quantile function of any distribution may be described as follows:

$$Q(u) = X_q = F^{-1}(u)$$

Where $Q(u)$ denotes the quantile function of $F(x)$ for $u \in (0,1)$.
Let us suppose

$$F(x) = 1 - cos\left(\frac{\pi(2 - 2^{e^{-\alpha x^{\beta}}})}{2}\right) = u \tag{12}$$

After simplifying equation (12), we obtain quantile function of NCWD distribution as

$$Q(u) = X_q = \left\{\frac{-1}{\alpha} log\left[\frac{1}{log(2)} log\left(\frac{2\pi - 2\arccos(1-u)}{\pi}\right)\right]\right\}^{\frac{1}{\beta}}$$

3. Reliability Measures of new cosine-Weibull distribution

This section is focused on researching and developing distinct ageing indicators for the formulated distribution.

3.1. Survival function

Suppose X be a continuous random variable with cdf $F(x)$.Then its Survival function which is also called reliability function is defined as

$$S(x) = p_r(X > x) = \int_x^{\infty} f(x)dy = 1 - F(x)$$

Therefore, the survival function for NCWD is given as

$$\begin{aligned} S(x;\alpha,\beta) =& 1 - F(x;\alpha,\beta) \\ =& cos\left(\frac{\pi(2 - 2^{e^{-\alpha x^{\beta}}})}{2}\right) \end{aligned} \tag{13}$$

3.2. Hazard rate function

The hazard rate function of a random variable x is denoted as

$$H(x;\alpha,\beta) = \frac{f(x;\alpha,\beta)}{S(x;\alpha,\beta)} \tag{14}$$

Using equation (10) and (13) in equation (14), then the hazard rate function of NCWD is given as

$$H(x;\alpha,\beta) = \frac{\pi log(2)}{2} \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}} 2^{e^{-\alpha x^{\beta}}} tan\left(\frac{\pi}{2}\left(2 - 2^{e^{-\alpha x^{\beta}}}\right)\right)$$

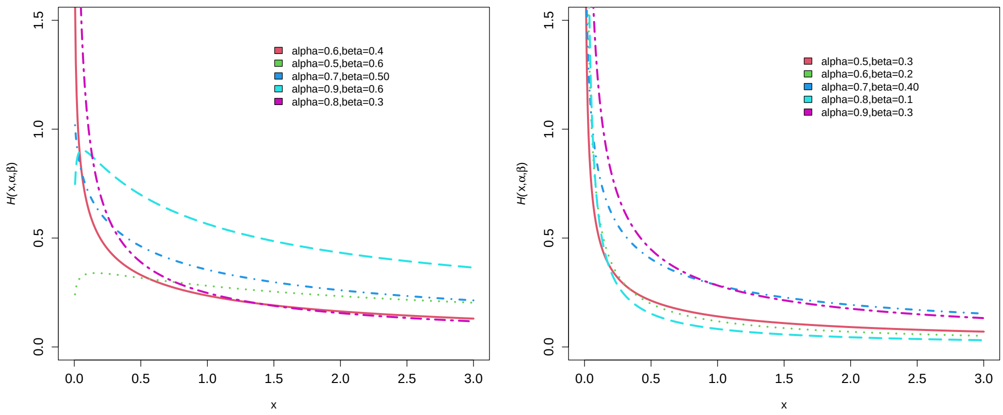


Figure 3: *HRF plot of NCWD for different choise of parameters*

3.3. Reverse hazard rate function

The reverse hazard rate function of a random variable x is given as

$$h(x;\alpha,\beta) = \frac{f(x;\alpha\beta)}{F(x;\alpha\beta)} \tag{15}$$

using equations (9) and (10) in equation (15), then we obtain reverse hazard rate function as

$$h(x;\alpha,\beta) = \frac{\frac{\pi log(2)}{2}\alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}} sin\left(\frac{\pi(2-2^{e^{-\alpha x^{\beta}}})}{2}\right)}{1-cos\left(\frac{\pi(2-2^{e^{-\alpha x^{\beta}}})}{2}\right)}$$

3.4. Mean residual function

The mean residual lifetime is the predicted residual life or the average completion period of the constituent after it has exceeded a certain duration x. It is extremely significant in reliability investigations.
Mean residual function of random x variable can be obtained as

$$\begin{aligned} m(x;\alpha,\beta) &= \frac{1}{S(x;\alpha,\beta)}\int_x^{\infty} tf(t,\alpha,\beta)dt - x \\ &= \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\Psi_{pqrs}\alpha\beta\sec\left((\pi/2)(2-2^{e^{-\alpha x^{\beta}}})\right)\int_x^{\infty} t^{\beta}e^{-\alpha(r+1)t^{\beta}}dt - x \end{aligned}$$

Making substitution $\alpha(r+1)t^{\beta} = z$, sothat $\alpha(r+1)x^{\beta} < z < \infty$, we have

$$m(x;\alpha,\beta) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\Psi_{pqrs}\alpha\sec\left((\pi/2)(2-2^{e^{-\alpha x^{\beta}}})\right)}{(\alpha(r+1))^{\frac{1+\beta}{\beta}}}\int_{\alpha(r+1)x^{\beta}}^{\infty} z^{\frac{1}{\beta}}e^{-z}dz - x$$

After solving the integral, we get

$$m(x;\alpha,\beta) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\Psi_{pqrs}\alpha\sec\left((\pi/2)(2-2^{e^{-\alpha x^{\beta}}})\right)}{(\alpha(r+1))^{\frac{1+\beta}{\beta}}}\Gamma\left(\frac{\beta+1}{\beta},\alpha(r+1)x^{\beta}\right) - x$$

Where $\Gamma(a,x) = \int_x^{\infty} u^{a-1}e^{-u}du$ denotes the upper incomplete gamma function

4. Order Statistics and Maximum likelihood Estimation of New cosine-Weibull distribution

Let us suppose $X_1, X_2, ..., X_n$ be random samples of size n from NCWD with pdf $f(x)$ and cdf F(x). Then the probability density function of the k^{th} order statistics is given as

$$f_X(k) = \frac{n!}{(k-1)!(n-1)!} f(x) \left[F(x)\right]^{k-1} \left[1 - F(x)\right]^{n-1} \tag{16}$$

Using equation (3) and (4) in equation (10), we have

$$f_X(k) = \frac{n!}{(k-1)!(n-1)!} \frac{\pi log(2)}{2} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} sin\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right) \left[cos\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right)\right]^{k-1}$$
$$\left[1 - cos\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right)\right]^{n-1}$$

The pdf of the first order X_1 and nth order X_n statistics of new cosine-Weibull distribution are respectively given as

$$f_X(1) = \frac{n\pi log(2)}{2} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} sin\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right) \left[1 - cos\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right)\right]^{n-1}$$

And

$$f_X(n) = \frac{n\pi log(2)}{2} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} sin\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right) \left[cos\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right)\right]^{n-1}$$

Let the random samples $x_1, x_2, x_3, ..., x_n$ are drawn from new cosine-Weibull distribution. The likelihood function of n observations is given as

$$L = \prod_{i=1}^{n} \frac{\pi log(2)}{2} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} sin\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right)$$

The log-likelihood function is given as

$$l = nlog(log(2)) + nlog(\alpha) + nlog(\beta) - \alpha \sum_{i=1}^{n} x_i^\beta + (\beta - 1) \sum_{i=1}^{n} log(x_i) + log(2) \sum_{i=1}^{n} e^{-\alpha x_i^\beta}$$
$$+ \sum_{i=1}^{n} log\left(sin\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right)\right) \tag{17}$$

The partial derivatives of the log-likelihood function with respect to α and β are given as

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} x_i^\beta - log(2) \sum_{i=1}^{n} x_i^\beta e^{-\alpha x_i^\beta} + (\frac{\pi}{2}) log(2) \sum_{i=1}^{n} x_i^\beta 2^{e^{-\alpha x_i^\beta}} cot\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right) \tag{18}$$

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} - \alpha\beta \sum_{i=1}^{n} x_i^{\beta-1} + \sum_{i=1}^{n} log(x_i) + log(2) - \alpha \sum_{i=1}^{n} x_i^\beta log(x_i) e^{-\alpha x^\beta}$$
$$+ (\frac{\alpha\pi}{2}) log(2) \sum_{i=1}^{n} x_i^\beta 2^{e^{-\alpha x_i^\beta}} e^{-\alpha x_i^\beta} cot\left(\frac{\pi(2-2^{e^{-\alpha x^\beta}})}{2}\right) log(x_i) \tag{19}$$

For interval estimation and hypothesis tests on the model parameters, an information matrix is required. The 2 by 2 observed matrix is

$$I(\xi) = \frac{-1}{n}\begin{bmatrix} E\left(\frac{\partial^2 logl}{\partial\alpha^2}\right) & E\left(\frac{\partial^2 logl}{\partial\alpha\partial\beta}\right) \\ E\left(\frac{\partial^2 logl}{\partial\beta\partial\alpha}\right) & E\left(\frac{\partial^2 logl}{\partial\beta^2}\right) \end{bmatrix}$$

The elements of above information matrix can be obtain by differentiating equations (18) and (19) again partially. Under standard regularity conditions when $n \to \infty$ the distribution of $\hat{\zeta}$ can be approximated by a multivariate normal $N(0, I(\hat{\zeta})^{-1})$ distribution to construct approximate confidence interval for the parameters. Hence the approximate $100(1-\psi)\%$ confidence interval for α and β are respectively given by

$$\hat{\alpha} \pm Z_{\frac{\psi}{2}}\sqrt{I_{\alpha\alpha}^{-1}(\hat{\zeta})} \quad and \quad \hat{\beta} \pm Z_{\frac{\psi}{2}}\sqrt{I_{\beta\beta}^{-1}(\hat{\zeta})}$$

5. Simulation Analysis

The bias, variance and MSE were all addressed to simulation analysis. From NCWD taking N=500 with samples of size n=50,150,150,250,350,450 and 500. The following expression has been used to produce random numbers.

$$X = \left\{\frac{-1}{\alpha}log\left[\frac{1}{log(2)}log\left(\frac{2\pi - 2\arccos(1-u)}{\pi}\right)\right]\right\}^{\frac{1}{\beta}}$$

Where u is uniform random numbers with $u \in (0,1)$. For various parameter combinations, simulation results have been achieved. The bias, variance and MSE values are calculated and presented in table 1 and 2. As the sample size increases, this becomes apparent that these estimates are relatively consistent and approximate the actual values of parameters. Interestingly, with all parameter combinations, the bias and MSE reduce as the sample size increases.

Table 1: *Bias, variance and their corresponding MSE's for different parameter values* $\alpha = 1.2, \beta = 0.5$

Sample size	Parameters	Bias	Variance	MSE
50	α	0.01450	0.02467	0.02488
	β	0.01371	0.00268	0.00287
150	α	0.00415	0.00614	0.00616
	β	0.00213	0.00091	0.00091
250	α	-0.00031	0.00411	0.00411
	β	0.00406	0.00049	0.00051
350	α	-0.00411	0.00276	0.00277
	β	0.00248	0.00037	0.00038
450	α	0.00223	0.00218	0.00219
	β	0.00255	0.00025	0.00026
500	α	0.00388	0.00214	0.00216
	β	0.00096	0.00024	0.00024

6. Data Aanalysis

This subsection evaluates a real-world data set to demonstrate the new cosine-Weibull distribution's applicability and effectiveness. The new cosine-Weibull distribution (NCWD) adaptability

Table 2: *Bias, variance and their corresponding MSE's for different parameter values $\alpha = 1.5, \beta = 1.2$*

Sample size	Parameters	Bias	Variance	MSE
50	α	0.01450	0.03773	0.03794
	β	0.03003	0.01942	0.02032
150	α	0.01324	0.01125	0.01143
	β	0.01132	0.00508	0.00521
250	α	0.00982	0.00698	0.00708
	β	0.00446	0.00324	0.00326
350	α	0.00286	0.00517	0.00518
	β	0.01186	0.00243	0.00257
450	α	0.00248	0.00370	0.00371
	β	0.00319	0.00165	0.00166
500	α	0.00160	0.00311	0.00311
	β	0.00073	0.00140	0.00140

is determined by comparing its efficacy to that of other analogous distributions such as Weibull distribution (BD), Frechet distribution (FD), Inverse Burr distribution (IBD), Nadrajah Haghigi distribution (NHD),Rayleigh distribution (RD) and Exponential distribution (ED).
To compare the versatility of the explored distribution, we consider the criteria like AIC (Akaike information criterion), CAIC (Consistent Akaike information criterion), BIC (Bayesian information criterion), HQIC (Hannan-Quinn information criterion), Kolmogorov-Smirnov tes (K.S), the Cramer-Van Mises criteria (W*) and the Anderson-Darling test (A*). Distribution having lesser AIC, CAIC, BIC,HQIC, K.S, W* and A* values is considered better.
Data set: The data set given below represents the waiting times (in minutes) before service of 100 bank customers, information provided by Ghitany et al.[10].
0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2,4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9,5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2,6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8.0, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9,12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27.0, 31.6, 33.1, 38.5
The ML estimates with corresponding standard errors in parenthesis of the unknown parameters are presented in Table 3 and the comparison statistics, AIC, BIC, CAIC, HQIC and the goodness-of-fit statistic for the data set are displayed in Table 4.

Table 3: *Descriptive statistics for data set*

Min.	Max.	Ist Qu.	Med.	Mean	3rd Qu.	kurt.	Skew.
0.800	38.500	4.675	8.100	9.877	13.025	5.5402	1.4727

Table 4: *The ML Estimates (standard error in parenthesis) for data set*

Model	$\hat{\alpha}$	$\hat{\beta}$
NCWD	0.0803	1.1474
	(0.0174)	(0.0830)
WD	0.0304	1.4585
	(0.0094)	(0.1087)
FD	6.535	1.1629
	(0.8918)	(0.0799)
IBD	8.8661	1.2900
	(1.2087)	(0.0827)
NHD	0.0212	3.3292
	(0.0138)	(1.8244)
ED	0.1012	....
	(0.0101)	....
RD	0.0066	....
	(0.00065)	....

Table 5: *Comparison criterion and goodness-of-fit statistics for data set*

Model	-l	AIC	CAIC	BIC	HQIC	K.S statistic	W*	A*	p-value
NCWD	316.98	637.96	638.09	643.17	640.07	0.0353	0.0168	0.1243	0.9996
WD	318.73	641.46	641.58	646.67	643.57	0.0577	0.0629	0.3962	0.8922
FD	334.38	672.76	672.88	677.97	674.87	0.1167	0.3832	2.505	0.1312
IBD	330.42	664.85	664.97	670.06	666.96	0.1026	0.2922	1.9478	0.2425
NHD	323.44	650.89	651.02	656.10	653.00	0.1076	0.1111	0.6958	0.197
ED	329.02	660.04	660.08	662.64	661.09	0.1730	0.0270	0.1790	0.0050
RD	329.24	660.48	660.52	663.08	661.53	0.1734	0.1268	0.7877	0.0048

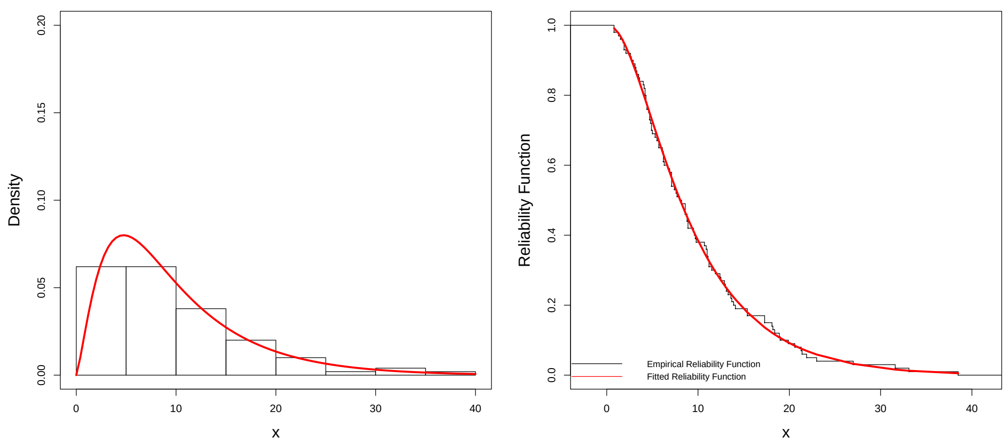


Figure 4: *Estimated pdf of the fitted model and Empirical versus fitted reliability function for data set.*

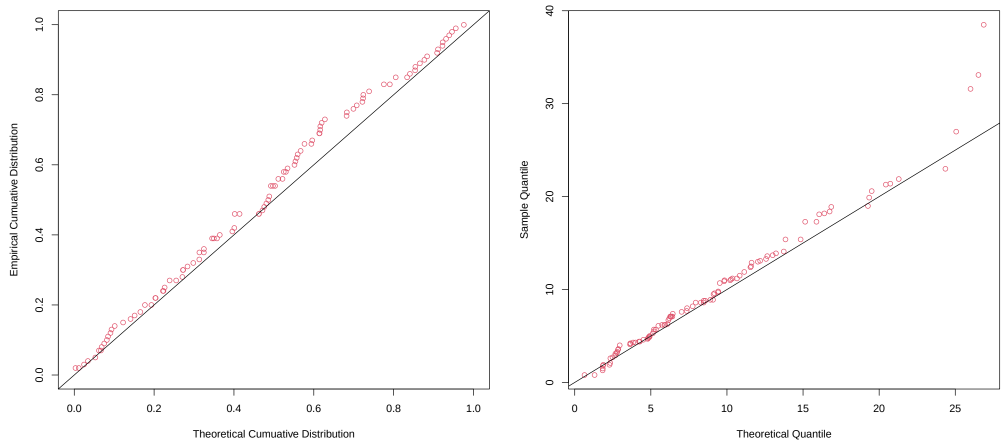


Figure 5: *PP and QQ plot for NCWD.*

It is observed from table 4 that NCWD provides best fit than other competative models based on the measures of statistics, AIC, BIC, CAIC, HQIC, K-S statistic W* and A*. Along with p-values of each model.

7. Conclusion

There is a growing concern among both statisticians and applied researchers in constructing versatile lifetime models to enhance the modelling of survival data. In this research, we established a two-parameter new cosine-Weibull distribution, which is created by employing the weibull distribution as the baseline. Several structural properties of the proposed distribution including moments, moment generating function, order statistics and reliability measures has been discussed. The parameters of the distribution are estimated by famous method of maximum likelihood estimation. Finally the efficiency of the distribution is examined through an application.

References

[1] Alzaatreh ,A., Lee, C., Famoye,F. A new method for generating families of distributions. *Metron*, *71 (2013),https://doi.org/10.1002/wics.1255*, 63-79.

[2] Aijaz, A., Qurat Ul-Ain, S., Rajnee, T and Afaq, A . Inverse Weibull-Rayleigh distribution distribution: characterization with application related cancer data. *Reliability Theoryy and Applications* , *16(4),(2021)*, 364-382.

[3] Aijaz, A. Muzmil, J., Ishfaq, H. D., and Rajnee, T. Inverse Weibull-Burr III distribution with properties and application related to survival rates in animals*Reliability Theoryy and Applications*, *2(68),(2022)*, 340-355.

[4] Aijaz, A., Mujamil,J and Afaq, A. A novel approach for constracting distributions with an example of the Rayleigh distribution. *Reliability theory and Applications*, *17 (2022)*, 42-64.

[5] Anwar, H., Ishfaq, H. D., and Murtaza, L. A novel family of generating distributions based on trigonometric functions with an application to exponential distribution. *Journal of scientific research*, *65(2) (2021)*, 173-179.

[6] Bourguingnon,M., Silva,R.B and Corderio,G.M. The Weibull-G family of probability distributions *Journal of data science*,*12 (2014), http://www.jds-online.com*, 53-68.

[7] Brito,E., Cordeiro,G.M., Yousuf,H.M., Alizadeh,M., Silva,G.O. The Topp-Leone odd log-logistic family of distributions *J stat comput simul*, *87(15) (2017), https://doi.org/10.1080/00949655.2017.1351972*, 3040-3058.

[8] Chesneau, C., Bakouch, H.S., and Hussian, T. A new class of probability distributions via cosine and sine functions with applications. *Comminications in Statistics-Simulation and Computation*, *48(8), (2019)*, 2287-2300.

[9] Faton. M and Ibrahim. E . Weibull-Rayleigh distribution theory and applications. *Applied mathematics and information sciences an international journal*, *9(5), (2015)*, 1-11.

[10] Ghitney, M. E., Atieh, B., Nadrajah, S. Lindley distribution and its applications. *Math. Comput. Simul.*, *78 (2008)*, 493-506.

[11] Jamal, F., and Chesneau, C. A new family of polyno-trigonometric distributions with applications. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, *22 (4), (2020)*, 1950027.

[12] Kumar, D., Singh, U and Singh, S.K. A new distribution using sine function its application to bladder cancer patients data. *Journal of Statistics and Probability*, *4(3) (2015)*, 417-427.

[13] Souza, L., Junior, W.R.O., Brito, C.C.R., Chesneau, C., Ferreira, T.A.E., and Soares, L. General properties of cos-G class of distributions with applications. *Eurasian Bull.Math*, *2, (2019)*, 63-79.

[14] Terna G .I, Sauta A, Issa A. A . Odd Lindley- Rayleigh distribution its properties and applications to simulated and real life datasets. *Journal of advances in mathematics and computer science*, *35(1) (2020)*, 68-88, DOI: 10.12785/jsap/030215.

[15] Voda V.G . Inferential procedures on generalized Rayleigh variate. *Applications of mathe-matics*, *21, (1976)*,395-412.

[16] Eugene,N., Lee,C and Famoye,F. Beta-normal distribution and its applications. *Communication in statistics- theory and methods*,*31 (2002)*,*https://doi.org/10.108/STA-120003130*, 497-512.

[17] Morad,A and Gauss,M.C. The Gompertz-G family of distributions *Journal of statistical theory and practice*,*11(1)(2017)*,*https:doi.org/10.108/15598608.2016.1267668*, 179-207.

[18] Mahood, Z., and Chesneau, C. A new sine-G family of distributions properties and applications.

[19] Murtaza, L., Ishfaq, H.D and Tariq, R. J. A new method for generating distributions with an application to Weibull distribution *Relibility theory and applications*,*1(67)(2022)*, 223-239.

[20] Zagrafos,K and Balakrishnan,N. On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, *6 (2009)*,*https://doi.org/10.1016/j.stamet.2008.12.003*, 344-362.

ON THE INFERENCES AND APPLICATIONS OF WEIBULL HALF LAPLACE{EXPONENTIAL} DISTRIBUTION

Adeyinka S. Ogunsanya*
[1]Department of Statistics, University of Ilorin, Kwara State, Nigeria
Correspondence email address:ogunsanyaadeyinka@yahoo.com,

Obalowu Job
[2]Department of Statistics, University of Ilorin, Kwara State, Nigeria
oobalowu@unilorin.edu.ng

Abstract

The study of probability distribution has expanded the field of statistical modelling of real life data. It has also provided solution to the problems of skewed data which often violate the normality. This research work introduces a new T - Half-Lapalace{Exponential} family with a novel Half-Laplace distribution as baseline distribution with specific interest in three-parameter lifetime model called the Weibull-Half-Lapalace{Exponential} (W-HLa{E}) distribution. The W-HLa{E} model is capable of modeling various shapes of aging events. The W-HLa{E} distribution is derived by combining Half-Laplace and Weibull distribution using the quartile function of Exponential distribution. Some of its statistical properties such as the mean, mode, quantile function, median, variance, standard deviation, skewness, and kurtosis are derived. Other statistical properties such as survival function, hazard rate, moments, asymptotic limit, order statistics, and entropy which is the measure of uncertainty of a random variable are derived and studied. The parameter estimation method adopted in this study is the maximum likelihood method. The graphs of W-HLa{E} at different values of shape and scale parameters show that the distribution is unimodal hence the mode is given as $mode = \theta + (\frac{k-1}{k})^{\frac{1}{k}}$ *and it is positively skewed with a steep peak. A simulation study is carried on the new proposed distribution using maximum likelihood estimation. The simulation also supported the theoretical expression of the statistical properties of the proposed distribution such as the location parameter does not affect the variance, skewness, and kurtosis of the new distribution. The importance and the flexibility of the proposed distribution in modeling some real life data sets is demostrated inn the research. The results of the sudy shows that the proposed W-HLa{E} distribution perform better than other disribbutions in the literature.*

Keywords: Laplace distribution; Half-Laplace distribution; Weibull Half-Laplace distribution; Censored data; Lifetime data; Maximum likelihood estimation

1. INTRODUCTION

Numerous classical distributions have been extensively used over the past decades for modeling data in several areas such as engineering, actuarial, environmental and medical sciences, biological studies, demography, economics, finance, and insurance. However, there is a clear need for extended forms of these classical distributions. Due to that reason, researchers have developed and studied several methods for generating new families of distributions. The most outstanding characteristics of this distribution are that it is unimodal and symmetric. Laplace distribution is a mixture of normal laws (see Kotz *[10]*), as a possible explanation of the wide applicability of this distribution for modeling growth rates. Let $f(x,\theta,\beta)$ be the probability density function of Laplace distribution.

$$f(x,\theta,\beta)=\frac{1}{2\beta}e^{(-|x-\theta|/\beta)}, x\in(-\infty,\infty) \quad (1)$$

and CDF, F(x) is given as:

$$F(x;\theta,\beta)=\begin{cases}\frac{1}{2}e^{\left(-\frac{x-\theta}{\beta}\right)}, & if\ x<\theta\\ 1-\frac{1}{2}e^{\left(-\frac{x-\theta}{\beta}\right)} & if\ x\geq\theta\end{cases} \quad (2)$$

The expected value of a Laplace distribution is given as $E(x)=\theta$. The expected value of a Laplace distribution is the same as the location parameter and a symmetric situation means that the mean is the same as mode and median. Laplace distribution can be compared with other symmetric distributions like Normal, Logistic, etc. except that Laplace has a higher spike and slightly thicker tails.

In recent years, Asymmetric Laplace distribution of [10] has received much attention in modeling currency exchange rates, interests, stock price changes which is a modification of Laplace distribution but not for survival data. Many researchers have developed compound distributions using different methods to fit survival data. In this research work, we shall reduce the classical Laplace distribution to a non-negative function. Most real-life quantities to be measured are non-negative values. With the assumption that the data is non-negative, therefore, it is necessary to reduce the Laplace distribution to one-sided (positively skewed distribution). Thus, the one-sided Laplace distribution, otherwise called the half-Laplace distribution is the positive side of the Laplace distribution.

Let X be a random variable on $R_+=(0,\infty)$ given by the density function of Laplace distribution equation (1), where $x\geq\theta\geq 0 and\ \beta>0$, then x is said to have a half Laplace distribution, denoted by $HL(\theta,\beta)$

$$f(x,\theta,\beta)=\frac{1}{\beta}e^{(-\frac{x-\theta}{\beta})}, x>\theta\geq 0;\beta>0 \quad (3)$$

and the CDF of the half-Laplace distribution is

$$F(x,\theta,\beta)=1-\frac{1}{\beta}e^{(-\frac{x-\theta}{\beta})}, x>\theta\geq 0;\beta>0 \quad (4)$$

Where θ is the location parameter and β is the shape parameter The half-Laplace distribution reduces to the exponential distribution when $\theta=0$, we have the exponential distribution.

Recently, many researchers have developed and studied compound distributions using T-X which was introduced by Alzaatreh [5] and T-X{Y} by Aljarrah [3]. This was later modified by Alzaatreh [6] as T-gamma family, Alzaatreh [7] also constructed T-normal families. Almheidat [4] studied the T-Weibull family. Amalare [8] derived Lomax-Cauchy {Uniform}. Ogunsanya [13] developed and studied the extension of Cauchy distribution named Rayleigh Cauchy distribution, Ogunsanya [14] studied Weibull-Inverse Rayleigh distribution: Classical/ Bayesian approach and Job [11] applied Weibull Loglogistic{Exponential} distribution on some survival data.

In Section 2, we derive the new Weibull Half Laplace distribution, with statistical properties such as hazard function, survival function, Moments, skewness, kurtosis, order statistics, and Shannon entropy are determined. Section 3 shows the simulation study. Estimation of the parameters of W-HLa{E} distribution by maximum likelihood is performed in Section 4. In Section 5, the performance of the new W-HLa{E} distribution is demonstrated on two real data sets and the conclusion and summary of the work are expressed in Section 6.

2. DERIVATION OF WEIBULL-HALFLAPLACE {EXPONENTIAL} DISTRIBUTION

In this section, we investigate in details the properties, parameters estimation, and applications of a new distribution of the T-Half Laplace $\{Y\}$ family called Weibull- Half Laplace $\{Exponential\}$ (W-HLa{E}) distribution.

Let t be a random variable that follows a two-parameter Weibull distribution, then PDF is given as

$$f_T(t;\lambda,k)=\left\{\frac{k}{\lambda}\left(\frac{t}{\lambda}\right)^{k-1}e^{-\left(\frac{t}{\lambda}\right)^k}\right\};t\geqslant 0,\lambda,k\geqslant 0$$

And the CDF is

$$F_T(t;\lambda,k)=1-e^{-\left(\frac{t}{\lambda}\right)^k};t\geqslant 0,\lambda,k\geqslant 0 \tag{5}$$

Recall Equation (4), and given the quantile function of exponential distribution as *-blog[1-x]*

$$F_X(x)=F_T\{-blog[1-(F_R(x))]\}$$

then the CDF of proposed W-HLa{E} distribution is

$$F_X(x)=F_T\left\{-blog\left[1-\left(1-e^{\left(-\frac{x-\theta}{\beta}\right)}\right)\right]\right\} \tag{6}$$

Substituting the CDF of Weibull distribution in equation (6)

$$F_X(x)=1-e^{-\left(\frac{b}{\lambda\beta}\right)^k(x-\theta)^k} \tag{7}$$

Let $\frac{b}{\lambda\beta}=\gamma$ in (7), then we have

$$F_X(x)=1-e^{-(\gamma)^k(x-\theta)^k} \tag{8}$$

Hence the corresponding PDF using equation (8)

$$f_X(x)=k\left(\frac{b}{\lambda\beta}\right)^k(x-\theta)^{k-1}e^{\left(\frac{b}{\lambda\beta}\right)^k(x-\theta)^k} \tag{9}$$

Let $\frac{b}{\lambda\beta}=\gamma$ in (9) or differentiate equation (8) with respect to x, then we have

$$f_X(x)=k\gamma^k(x-\theta)^{k-1}e^{\gamma^k(x-\theta)^k} \tag{10}$$

where $k,\theta,\gamma\geq 0$ are parameters of W-HLa{E} distribution

3. STATISTICAL PROPERTIES OF W-HLA{E} DISTRIBUTION

The statistical properties of the W-HLa{E} distribution including quantile function, ordinary moments, and Shannon entropy are provided in this section

Proposition 1 (Quantile Function) If X is a random variable that has W-HLa{E} distribution $(x;k,\theta,\gamma)$ and let $Q_X(p)$, such that $0\leq p\leq 1$ denote the quantile function for the W-HLa{E} distribution. Then $Q_X(p)$ is given by

$$Q_X(p)=\theta+\frac{1}{\gamma}\{-log(1-p)\}^{\frac{1}{k}} \tag{11}$$

where $k,\theta,\gamma\geq 0$ are parameters of W-HLa{E} distribution

Proof:

From Equation (8) replace $F_X(x)$ with p and solve for x, we obtain (11), the quantile function of W-HLa{E} distribution.

Setting$p=0.25,0.50, and 0.75$ in (30)he quartiles of the W-HLa{E} distribution can be obtained.

Lower Quartile

$$Q_X(0.25)=\theta+\frac{1}{\gamma}\{-log(1-0.25)\}^{\frac{1}{k}}$$

Median

$$Q_X(0.5)=\theta+\frac{1}{\gamma}\{-log(1-0.5)\}^{\frac{1}{k}}$$

Upper Quartile

$$Q_X(0.75) = \theta + \frac{1}{\gamma}\{-log(1-0.75)\}^{\frac{1}{k}}$$

Proposition 2 (Modal Function) *If X is a random variable that has W-HLa{E} distribution*$(x; k, \theta, \gamma)$ *and let* $X_{Mode}(x)$*, such that* $0 \le p \le 1$ *denote the mode function for the W-HLa{E} distribution. Then* $Q_X(p)$ *is given by*

$$X_{mode}(x) = \theta + \frac{1}{\gamma}\left[\frac{k-1}{k}\right]^{\frac{1}{k}} \tag{12}$$

where $k, \theta, \gamma > 0$ are parameters of W-HLa{E} distribution

Proof:

Differentiate (10) and equate to zero

$$\frac{d}{dx}f(x) = 0$$

$$\frac{d}{dx}\left[k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^{k-1}}\right] = 0$$

$$\frac{k(k-1)\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^{k-1}}}{(x-\theta)} - \frac{k^2\gamma^{2k}(x-\theta)^{2k-1}e^{-\gamma^k(x-\theta)^{k-1}}}{(x-\theta)} = 0$$

$$\frac{k(k-1)\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^{k-1}}}{(x-\theta)} = \frac{k^2\gamma^{2k}(x-\theta)^{2k-1}e^{-\gamma^k(x-\theta)^{k-1}}}{(x-\theta)}$$

Solving for x, we have

$$X_{mode}(x) = \theta + \frac{1}{\gamma}\left[\frac{k-1}{k}\right]^{\frac{1}{k}}$$

3.1. Shape Properties of W-HLa{E} Distribution

Cumulative Distribution Function (CDF) of W-HLa{E} Distribution.

Equation (9) is now the CDF of the new probability distribution called W-HLa{E} Distribution.

3.2 Hazard Function

The hazard function of the W-HLa{E} distribution is derived from this definition

$$h_X(x) = f_X(x)/1 - F_X(x)$$

where $f_X(x)$ and $F_X(x)$ are the PDF and CDF of W-HLa{E} distribution given in (9) and (10) respectively. The hazard function h(x) can be written as

$$h_X(x) = \frac{k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^k}}{1-\{1-e^{-\gamma^k(x-\theta)^k}\}} \tag{13}$$

Simplifying (13), we have

$$h_X(x) = k\gamma^k(x-\theta)^{k-1} \tag{14}$$

The log hazard of W-HLa{E}D, which is frequently used in modeling is given by

$$\lambda_X(t) = log(h_X(x))$$

$$\lambda_X(t) = log(k\gamma^k(x-\theta)^{k-1}) \tag{15}$$

Expanding equation (15)

$$\lambda_X(t) = log(k\gamma^k) + log(x-\theta)^{k-1}$$

$$\lambda_X(t) = log(k\gamma^k) + (k-1)log(x-\theta) \tag{16}$$

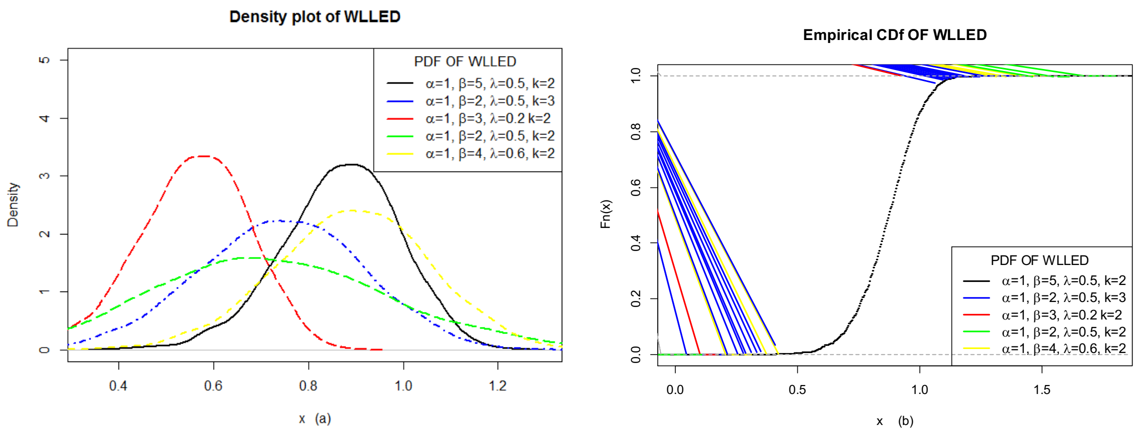


Figure 1: *(a) Density plot and (b) CDF plot of WLE distribution for sample size = 1000 and for various values of* k, θ *and* γ.

Figure 1(a) ahows that the PDF is positively skewed. Figure 1(b) shows that the CDF of W-HLa{E}D increases as x increases and remains constant as it approaches 1.

3.3 Survival Function

The survival function of the W-HLa{E} distribution is derived from this definition

$$S_X(x) = 1 - F_X(x)$$

where F(x) is the CDF of the W-HLa{E} distribution as defined in equation(9) The survival function S_x can be written as $S_X(x) = 1 - \{1 - e^{-\gamma^k (x-\theta)^k}\}$

$$S_X(x) = e^{-\gamma^k (x-\theta)^k} \quad (17)$$

For $x > 0, k, \theta and \gamma > 0 \; and \; t > 0$, the probability that a system having age x units of time will survive up to $x + t$ units of time is given by

$$S_X(x) = \frac{e^{-\gamma^k (x+t-\theta)^k}}{e^{-\gamma^k (x-\theta)^k}}$$

3.4 Cumulative Hazard Function

The cumulative hazard function, of the W-HLa{E} distribution, is given as

$$H_X(x) = -log_e\{e^{-\gamma^k (x-\theta)^k}\} \quad (18)$$

Simplifying (18) we have

$$H_X(x) = \gamma^k (x-\theta)^k \quad (19)$$

3.5 Asymptotic Behavior of W-HLa{E} Distribution

To investigate the asymptotic behavior of the proposed distribution model W-HLa{E}, we find the limit as $x \longrightarrow \theta$ and as $x \longrightarrow \infty$ of the W-HLa{E} distribution

$$lim_{x \to \theta} f(x) = lim_{x \to \theta} k\gamma^k (x-\theta)^{k-1} e^{-\gamma^k (x-\theta)^k} = 0$$

Proof: Since $lim_{x \to \theta}(x - \theta) = 0$, then

$lim_{x \to \theta} f(x) = 0$

Hence as x tend to a minimum value of the distribution, W-HLa{E} distribution becomes zero

Similarly,

$$lim_{x \to \infty} f(x) = lim_{x \to \theta} k\gamma^k (x-\theta)^{k-1} e^{-\gamma^k (x-\theta)^k} = 0$$

Proof: Since $lim_{x \to \infty} = \left(e^{-\gamma^k (x-\theta)^k}\right) = 0$, then

$lim_{x \to \infty} f(x) = 0$

As x tends to infinity the W-HLa{E} distribution becomes zero

3.6 Moments and Variance

In this subsection, we shall determine r^{th} moment of about the origin and n^{th} moment about the mean. Given

$$E(X) = \int_{\theta}^{\infty} x f_X(x) dx \tag{20}$$

Proposition 3 (First Moment about Origin) *If X is a random variable that has W-HLa{E} distribution$(x; k, \theta, \gamma)$ and let X_M denote the first moment about the origin of the W-HLa{E} distribution. Then $\mu_1^{`}$ is given by*

$$X_M = \mu_1^{`} = \theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)$$

where $k, \theta, \gamma \geq 0$ are parameters of W-HLa{E} distribution

Proof: Substituting (10) in (20)

$$E(X) = \int_{\theta}^{\infty} xk\gamma^k (x-\theta)^{k-1} e^{-\gamma^k (x-\theta)^k} dx \tag{21}$$

Let

$$z = \gamma^k (x-\theta)^k \tag{22}$$

then differentiate Equation (22) with respect to x

$$dx = \frac{dz}{k\gamma^k (x-\theta)^{k-1}} \tag{23}$$

Substitute (23) in (21), we have

$$E(X) = \int_{\theta}^{\infty} xk\gamma^k (x-\theta)^{k-1} e^{-\gamma^k (x-\theta)^k} \frac{dz}{k\gamma^k (x-\theta)^{k-1}} \tag{24}$$

$$E(X) = \int_{\theta}^{\infty} x e^{-\gamma^k (x-\theta)^k} dz \tag{25}$$

From (22), we make x the subject,

$$x = \theta + \frac{z^{1/k}}{\gamma} \tag{26}$$

$0 \leq z \leq \infty$ Substitute (22) and (26) in Eqn. (25), we have

$$E(X) = \int_{0}^{\infty} \left(\theta + \frac{z^{1/k}}{\gamma}\right) e^{-z} dz \tag{27}$$

$$E(X) = \int_{0}^{\infty} \theta e^{-z} dz + \int_{0}^{\infty} \frac{z^{1/k}}{\gamma} e^{-z} dz \tag{28}$$

By evaluating the limits, we have

$$E(X) = -\theta e^{-\gamma^k (x-\theta)^k} |_{\theta}^{\infty} + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) \tag{29}$$

Hence equation (29) becomes

$$E(X) = \theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) \tag{30}$$

Proposition 4 (Second Moment about Origin) *If X is a random variable that has W-HLa{E} distribution$(x; k, \theta, \gamma)$ and let $\mu_2^{`}$ denote the second moment about the origin of the W-HLa{E} distribution. Then $\mu_2^{`}$ is given by*

$$\mu_2^{`} = \theta + 2\frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right)$$

where $k, \theta, \gamma \geq 0$ are parameters of W-HLa{E} distribution

Proof: Given

$$E(X^2) = \mu_2^{`} = \int_{\theta}^{\infty} x^2 f_X(x) dx \tag{31}$$

Substitute equation (10) in equation (31)

$$\mu_2^{`} = \int_{\theta}^{\infty} x^2 k\gamma^k (x-\theta)^{k-1} e^{-\gamma^k (x-\theta)^k} dx \tag{32}$$

Using the above method adopted in proposition 4.4, we have,

$$\mu_2^{`} = \theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) \tag{33}$$

Corollary 1 *(rth moment) Let X be a random variable that follows W-HLa{E} distribution$(x; k, \theta, \gamma)$ and let $\grave{\mu_r}$ denote the rth moment about the origin of the W-HLa{E} distribution. Then $\grave{\mu_r}$ is given by*

$$\grave{\mu_r} = \theta^r + \sum_{i=1}^{r} \binom{r}{i} \frac{\theta^{r-1}}{\gamma^i} \Gamma\left(1 + \frac{i}{k}\right) \tag{34}$$

Where i=1,2,3,…,r

Proof: By Mathematical induction, it follows from equations (30) and (33) of propositions 3 and 4. respectively

Hence the first four moments of the proposed distribution are given

$$\grave{\mu_1} = \theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)$$

$$\grave{\mu_2} = \theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right)$$

$$\grave{\mu_3} = \theta^3 + 3\frac{\theta^2}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + 3\frac{\theta}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) + \frac{1}{\gamma^3}\Gamma\left(1 + \frac{3}{k}\right)$$

$$\grave{\mu_4} = \theta^4 + 4\frac{\theta^3}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + 6\frac{\theta^2}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) + 4\frac{\theta}{\gamma^3}\Gamma\left(1 + \frac{3}{k}\right) + \frac{1}{\gamma^4}\Gamma\left(1 + \frac{4}{k}\right)$$

Again using the relationship between moments about mean and moments about the origin, the moments about the mean of W-HLa{E} distribution are obtained as

$$Variance(\mu_2) = E(X^2) - [E(X)]^2$$

$$Variance(\mu_2) = \theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) - \left[\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]^2$$

$$= \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) - \left[\frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]^2$$

$$= \frac{1}{\gamma^2}\left\{\Gamma\left(1 + \frac{2}{k}\right) - \left[\Gamma\left(1 + \frac{1}{k}\right)\right]^2\right\}$$

Therefore the standard deviation is given as

$$StandardDeviation(X) = \sqrt{\frac{1}{\gamma^2}\left\{\Gamma\left(1 + \frac{2}{k}\right) - \left[\Gamma\left(1 + \frac{1}{k}\right)\right]^2\right\}}$$

$$= \frac{1}{\gamma}\sqrt{\left\{\Gamma\left(1 + \frac{2}{k}\right) - \left[\Gamma\left(1 + \frac{1}{k}\right)\right]^2\right\}}$$

$$Co - efficient \text{ of } Variation(C.V) = \frac{StandardDeviation}{E(X)}$$

$$C.V = \frac{\frac{1}{\gamma^2}\sqrt{\left\{\Gamma\left(1 + \frac{2}{k}\right) - \left[\Gamma\left(1 + \frac{1}{k}\right)\right]^2\right\}}}{\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)}$$

Again using the relationship between moments about mean and moments about origin where $\mu_3 = \grave{\mu_3} - 3\grave{\mu_1}\grave{\mu_2} + 2\grave{\mu_1}^3$ and $\mu_4 = \grave{\mu_4} - 4\grave{\mu_1}\grave{\mu_3} + 6\grave{\mu_1}^2\grave{\mu_2} + 3\grave{\mu_1}^4$ are third and fourth moments about the mean respectively. The moments about the mean of W-HLa{E} distribution are obtained as

$$\mu_3 = \theta^3 + 3\frac{\theta^2}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + 3\frac{\theta}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) + \frac{1}{\gamma^3}\Gamma\left(1 + \frac{3}{k}\right) -$$
$$3\left[\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]\left[\theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right)\right] + 2\left[\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]^3$$

$$\mu_4 = \theta^4 + 4\frac{\theta^3}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + 6\frac{\theta^2}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) + 4\frac{\theta}{\gamma^3}\Gamma\left(1 + \frac{3}{k}\right) + \frac{1}{\gamma^4}\Gamma\left(1 + \frac{4}{k}\right)$$
$$-4\left[\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]\left[\theta^3 + 3\frac{\theta^2}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + 3\frac{\theta}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right) + \frac{1}{\gamma^3}\Gamma\left(1 + \frac{3}{k}\right)\right]$$
$$+6\left[\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]^2\left[\theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1 + \frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1 + \frac{2}{k}\right)\right] + 3\left[\theta + \frac{1}{\gamma}\Gamma\left(1 + \frac{1}{k}\right)\right]^4$$

Hence, the skewness and kurtosis are determined as follows;

$$Skewness = \frac{E(x-\mu)^3}{\sigma^2} = \frac{\mu_3}{\sigma^2}$$

$$Skewness = \frac{1}{\frac{1}{\gamma^2}\left\{\Gamma\left(1+\frac{2}{k}\right)-\left[\Gamma\left(1+\frac{1}{k}\right)\right]^2\right\}}\theta^3 + 3\frac{\theta^2}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + 3\frac{\theta}{\gamma^2}\Gamma\left(1+\frac{2}{k}\right) + \frac{1}{\gamma^3}\Gamma\left(1+\frac{3}{k}\right) -$$

$$3\left[\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]\left[\theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1+\frac{2}{k}\right)\right] + 2\left[\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]^3$$

Further simplification of above we have

$$= \frac{\frac{2\Gamma\left(1+\frac{1}{k}\right)^2 - 3\Gamma\left(1+\frac{1}{k}\right)\Gamma\left(1+\frac{2}{k}\right) + \Gamma\left(1+\frac{3}{k}\right)}{\gamma^2}}{\frac{1}{\gamma^2}\Gamma\left(1+\frac{2}{k}\right) - \left[\frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]^2}$$

$$= \frac{2\Gamma\left(1+\frac{1}{k}\right)^2 - 3\Gamma\left(1+\frac{1}{k}\right)\Gamma\left(1+\frac{2}{k}\right) + \Gamma\left(1+\frac{3}{k}\right)}{\Gamma\left(1+\frac{2}{k}\right) - \left[\frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]^2}$$

$$\text{Kurtosis} = \frac{\mu_4}{\sigma^4}$$

$$Kurtosis = \left\{\frac{1}{\frac{1}{\gamma^4}\left\{\Gamma\left(1+\frac{2}{k}\right)-\left[\Gamma\left(1+\frac{1}{k}\right)\right]^2\right\}^2}\right\}\left\{\theta^4 + 4\frac{\theta^3}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + 6\frac{\theta^2}{\gamma^2}\Gamma\left(1+\frac{2}{k}\right) + \right.$$

$$4\frac{\theta}{\gamma^3}\Gamma\left(1+\frac{3}{k}\right) + \frac{1}{\gamma^4}\Gamma\left(1+\frac{4}{k}\right) - 4\left[\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]\left[\theta^3 + 3\frac{\theta^2}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + 3\frac{\theta}{\gamma^2}\Gamma\left(1+\frac{2}{k}\right) + \frac{1}{\gamma^3}\Gamma\left(1+\frac{3}{k}\right)\right] +$$

$$\left. 6\left[\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]^2\left[\theta^2 + 2\frac{\theta}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + \frac{1}{\gamma^2}\Gamma\left(1+\frac{2}{k}\right)\right] + 3\left[\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)\right]^4\right\}$$

3.7 Order Statistics

Order statistics is an important concept in probability theory. Let a random sample $X_1, X_2, \ldots X_n$, from the distribution function F(x) and corresponding pdf f(x), therefore the pdf of ith order statistic is given as

Proposition 5 *If X is a random variable that has W-HLa{E} distribution*$(x; k, \theta, \gamma)$ *and let* $f(x_i)$ *denote the pdf of ith order statistic which is given as*

$$f(X_i) = \frac{n!}{(i-1)!(n-i)!} k\theta^p\gamma^k \sum_{p,q=0}^{\infty} \binom{k-1}{p}\binom{i-1}{q}(-1)^{p+q}\left(e^{-\gamma^k(x-\theta)^k}\right)^{1+q+n-i} \quad (35)$$

where $k, \theta, \gamma \geq 0$ are parameters of W-HLa{E} distribution

Proof: Given

$$f(X_i) = \frac{n!}{(i-1)!(n-i)!} f(x)F(x)^{i-1}[1-F(x)]^{n-1} \quad (36)$$

hence the pdf of ith order statistic of W-HLa{E} distribution is determined by substituting equation (7) and (10) in equation (35) we have

$$f(X_i) = \frac{n!}{(i-1)!(n-i)!} k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^k}\left[1-e^{-\gamma^k(x-\theta)^k}\right]^{i-1}\left[e^{-\gamma^k(x-\theta)^k}\right]^{n-1} \quad (37)$$

$$f(X_i) = \frac{n!}{(i-1)!(n-i)!} k\gamma^k \sum_{p=0}^{\infty}\binom{k-1}{p}(-1)^p\theta^p e^{-\gamma^k(x-\theta)^k}\sum_{q=0}^{\infty}\binom{i-1}{q}(-1)^q\left(e^{-\gamma^k(x-\theta)^k}\right)^q \quad (38)$$

$$f(X_i) = \frac{n!}{(i-1)!(n-i)!} k\theta^p\gamma^k \sum_{p,q=0}^{\infty}\binom{k-1}{p}\binom{i-1}{q}(-1)^{p+q}\left(e^{-\gamma^k(x-\theta)^k}\right)^{1+q+n-i}$$

Therefore the first and nth order statistics for W-HLa{E} distribution can be determined as thus

Corollary 2 *(nth Order Statistic) Let X be a random variable that follows W-HLa{E} distribution*$(x; k, \theta, \gamma)$ *and let* $f(x_1)$ *denote the first (1st) order statistic of the W-HLa{E} distribution. Then* $f(x_1)$ *is given by*

$$f(X_1) = \frac{n}{(q)!(-q)!} k\theta^p\gamma^k \sum_{p,q=0}^{\infty}\binom{k-1}{p}(-1)^{p+q}\left(e^{-\gamma^k(x-\theta)^k}\right)^{n+q}) \quad (39)$$

Where i=1,2,3,…,n

Proof: From equation (35), replace i with 1 we have equation (39)

Corollary 3 *(nth Order Statistic) Let X be a random variable that follows W-HLa{E} distribution*$(x; k, \theta, \gamma)$ *and let* $f(x_n)$ *denote the last (nth) order statistic of the W-HLa{E} distribution. Then* $f(x_n)$ *is given by*

$$f(X_i) = \frac{n}{(n-1-q)!(q)!} k\theta^p \gamma^k \sum_{p,q=0}^{\infty} \binom{k-1}{p} (-1)^{p+q} \left(e^{-\gamma^k(x-\theta)^k}\right)^{1+q}) \quad (40)$$

Where i=1,2,3,…,n

Proof: From equation (35), replace i with n we have equation (40)

3.8 Entropy

In information theory, entropy is an important concept and can be defined as a measure of the randomness or uncertainty associated with a random variable. However, the Shannon entropy for a random variable X with pdf $f_X(x)$ is defined as $E\{-log(f_X(x))\}$

Proposition 6 *If X is a random variable that has W-HLa{E} distribution*$(x; k, \theta, \gamma)$ *then Shannon's entropy is given as*

$$E\{-log(f_X(x))\} = \gamma^k - (k-1)\left(\frac{0.57722}{k}\right) - log(\gamma) - log(k)$$

where $k, \theta, \gamma \geq 0$ are parameters of W-HLa{E} distribution and $\Phi = -\int_0^{\infty} logze^{-z}dz \approx 0.57722$ is the Euler gamma constant

Proof: Substitute equation (10) in the definition of entropy we have

$$E\{-log(f_X(x))\} = E\left\{-log\left(k\gamma^k(x-\theta)^{k-1}e^{\gamma^k(x-\theta)^k}\right)\right\} \quad (41)$$

$$E\{-log(f_X(x))\} = E\{-[log(k) + klog(\gamma) + (k-1)log(x-\theta) - \gamma^k(x-\theta)^k]\}$$
$$= E\{\gamma^k(x-\theta)^k - [log(k) + klog(\gamma) + (k-1)log(x-\theta)]\} \quad (42)$$

Finding the expectation of $(x-\theta)$ with the proposed distribution W-HLa{E}, we have,

$$E((x-\theta)^k) = \int_{\theta}^{\infty} (x-\theta)^k k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^k}dx \quad (43)$$

Recall Equation (22) and (23)and substitute them in equation (43)

$$E((x-\theta)^k) = \int_{\theta}^{\infty} (x-\theta)^k k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^k} \times \frac{dz}{k\gamma^k(x-\theta)^{k-1}} \quad (44)$$

$$E((x-\theta)^k) = \int_0^{\infty} (x-\theta)^k e^{-\gamma^k(x-\theta)^k}dz \quad (45)$$

Substituting Equation (26) in Equation (45) and integrating the expression

$$E((x-\theta)^k) = \int_0^{\infty} ze^z dz = [-ze^z - e^z]_0^{\infty} = 1 \quad (46)$$

Hence

$$E(\gamma^k(x-\theta)^k) = \gamma^k E((x-\theta)^k) = \gamma^k \quad (47)$$

Also

$$E((x-\theta)^k) = \int_{\theta}^{\infty} log(x-\theta)^k f_X(x)dx \quad (48)$$

Therefore

$$E(log(x-\theta)^k) = \int_{\theta}^{\infty} log(x-\theta)^k k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^k}dx \quad (49)$$

Substitute Equation (22) and (23)and substitute them in equation (49)

$$E(log(x-\theta)^k) = \int_{\theta}^{\infty} log(x-\theta)^k k\gamma^k(x-\theta)^{k-1}e^{-\gamma^k(x-\theta)^k} \times \frac{dz}{k\gamma^k(x-\theta)^{k-1}} \quad (50)$$

Then equation (50) becomes

$$= \int_{\theta}^{\infty} log(x-\theta)^k e^{-z}dz \quad (51)$$

$$= \int_{\theta}^{\infty} log\left(\frac{z^{\frac{1}{k}}}{\gamma}\right)e^{-z}dz \quad (52)$$

$$= \int_0^{\infty} log\left(z^{\frac{1}{k}}\right)e^{-z}dz + \int_0^{\infty} log\left(\frac{1}{\gamma}\right)e^{-z}dz$$

$$= \int_0^{\infty} log\left(z^{\frac{1}{k}}\right)e^{-z}dz - \int_0^{\infty} log(\gamma)e^{-z}dz \quad (53)$$

$$= \int_0^{\infty} log\left(z^{\frac{1}{k}}\right)e^{-z}dz - log(\gamma)\int_0^{\infty} e^{-z}dz \quad (54)$$

$$= \frac{1}{k}\int_0^{\infty} log(z)e^{-z}dz - log(\gamma)\int_0^{\infty} e^{-z}dz \quad (55)$$

Let $\Phi = -\int_0^{\infty} log(z)e^{-z}dz$ then equation (55) becomes

$$E(log(x-\theta)^k) = -\frac{\Phi}{k} - log(\gamma) \quad (56)$$

Substitute Equations (47) and (56) in Equation (42), we have

$$E\{-log(f_X(x))\} = \gamma^k - (k-1)\left(-\frac{\Phi}{k} - log(\gamma)\right) - \left(log(k) + klog(\gamma)\right) \quad (57)$$

$$E\{-log(f_X(x))\} = \gamma^k + (k-1)\left(\frac{\Phi}{k} + log(\gamma)\right) - log(k) - klog(\gamma)$$

$$E\{-log(f_X(x))\} = \gamma^k + (k-1)\left(\frac{\Phi}{k}\right) - log(\gamma) - log(k)$$

where $k, \theta, \gamma \geq 0$ are parameters of W-HLa{E} distribution and $\Phi = -\int_0^\infty logze^{-z}dz \approx 0.57722$ is the Euler gamma constant from Abramowitz [1]

$$E\{-log(f_X(x))\} = \gamma^k + (k-1)\left(\frac{0.57722}{k}\right) - log(\gamma) - log(k) \quad (58)$$

3.9 Mean Residual life

The Mean Residual Life (MRL) at a given time *t* measures the expected remaining life of an individual of age t. it otherwise called the life expectancy.

Proposition 7: *Let T be a random variable that follows W-HLa{E} distribution*(t, k, γ, θ) *and let* MRL *represents mean residual life at a given time t W-HLa{E} distribution. Then* MRL *is given by*

$$\text{MRL} = \frac{1}{e^{-(\gamma)^k(t-\theta)^k}}\left\{\frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + \theta e^{-t} - \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}, t\right)\right\} - \text{t}$$

where $k, \gamma, \theta \geq 0$ are parameters of *t W-HLa{E}* distribution

Proof: Given

$$\text{MRL} = \frac{1}{1-S(t)}\left\{E(x) - \int_0^t \text{t}f_X(t)dt\right\} - \text{t} \quad (59)$$

And $S_X(x) = 1 - \text{F(x)}$

From equation (8) and (34)

$$F_X(t) = 1 - e^{-(\gamma)^k(t-\theta)^k}, \quad E(X) = \theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right)$$

$$\text{MRL} = \frac{1}{e^{-(\gamma)^k(x-\theta)^k}}\left\{\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right) - \int_0^t \text{t}f_X(t)dt\right\} - \text{t} \quad (60)$$

$$\int_0^t \text{t}f_X(t)dt = \int_0^t \left(\theta + \frac{z^{1/k}}{\gamma}\right)e^{-t}dt$$

$$= \int_0^t \theta e^{-t}dt + \int_0^t \frac{t^{1/k}}{\gamma}e^{-t}dt$$

$$= -\theta e^{-t}{}_0^t + \frac{1}{\gamma}\gamma\left(1+\frac{1}{k}, t\right)$$

$$\int_0^t \text{t}f_X(t)dt = \theta - \theta e^{-t} + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}, t\right) \quad (61)$$

Substituting equations (61) in equation (60)

$$\text{MRL} = \frac{1}{e^{-(\gamma)^k(x-\theta)^k}}\left\{\theta + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right) - \left(\theta - \theta e^{-t} + \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}, t\right)\right)\right\} - \text{t}$$

Where $\Gamma\left(1+\frac{1}{k}, t\right)$ is an incomplete gamma function of variable t.

$$\text{MRL} = \frac{1}{e^{-(\gamma)^k(t-\theta)^k}}\left\{\frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}\right) + \theta e^{-t} - \frac{1}{\gamma}\Gamma\left(1+\frac{1}{k}, t\right)\right\} - \text{t}$$

4. SIMULATION STUDY

The simulation was done using equation (11) which is the quantile function of W-HLa{E}. Let *p be* a uniform random variable on (0,1), then $X = \theta + \frac{1}{\gamma}\{-log(1-p)\}^{\frac{1}{k}}$, the descriptive summaries were obtained through Statistical software R3.4.4.version.

Table 1: *Descriptive summaries of simulation of W-HLa{E} distribution with various parameters*

Model	Parameters	Mean	Median	Max.	Variance	skewness	kurtosis	Cv
y_1	k=0.5; γ=0.5; θ=5	6.801	5.453	21.337	14.239	3.115	12.309	0.554
y_2	k=0.5; =0.5;θ=10	11.8	10.45	26.34	14.239	3.115	12.309	0.320
y_3	k=1; γ=3; θ=5	10.22	10.16	10.95	0.055	1.748	5.827	0.023
y_4	k=1.5; γ=1; θ=5	5.67	5.609	7.014	0.251	1.040	3.751	0.088
y_5	k=2; γ=2; θ=10	10.35	10.34	10.85	0.042	0.612	3.008	0.036
y_6	k=2; γ=2; θ=5	5.352	5.345	5.845	0.042	0.612	3.008	0.036
y_7	k=3; γ=1; θ=5	5.76	5.781	6.419	0.097	0.127	2.625	0.054
y_8	k=3; γ=1; θ=10	10.76	10.78	11.42	0.097	0.127	2.625	0.029

From table 1 it is observed that as parameter k increases the variance, skewness, and kurtosis function decreases for different values of the scale and location parameters of W-HLa{E} distribution hence, the skewness and kurtosis are decreasing functions of k.
When k=2 and γ=2, the mean is approximately equal to the median hence, the distribution tends to be symmetric. Table 1 also shows that the coefficient of variation is deeply affected by γ.
When the γ increases positively the value of the coefficient of variation decreases

5. ESTIMATION OF PARAMETERS FOR THE W-HLA{E} DISTRIBUTION

The Maximum Likelihood estimates of W-HLa{E} distribution will be obtained in this section
Definition: Let $x_1, x_2, \dots, x_n$ denote a random sample drawn from W-HLa{E} distribution with parameters k, γ, θ. The likelihood function l (x, k, γ, θ) of W-HLa{E} distribution is defined to be the joint density of the random variables $x_1, x_2, \dots, x_n$

$$l(x,k,\gamma,\theta) = \prod k\gamma^k (x-\theta)^{k-1} e^{\gamma^k (x-\theta)^k} \quad (62)$$

$$= k^n \gamma^{nk} \prod (x-\theta)^{k-1} e^{-\gamma^k (x-\theta)^k} \quad (63)$$

finding the loglikelihood function of equation (63), we have

$$logl(x,k,\gamma,\theta) = \log k^n + \log\gamma^{nk} + \sum_i^n \log(x-\theta)^{k-1} + \sum_i^n -\gamma^k((x-\theta)^k) \quad (64)$$

where L=logl(x, k, γ, θ)

$$L = n\log k + nk\log\gamma + (k-1)\sum_i^n \log(x-\theta) - \gamma^k \sum_i^n ((x-\theta)^k) \quad (66)$$

$$L = n\log k + nk\log\gamma + (k-1)\sum_i^n \log(x-\theta) - \gamma^k \sum_i^n ((x-\theta)^k) \quad (66)$$

Differentiating the log-likelihood function in (66) with respect to the parametersK, γ, θ we have

$$\frac{dL}{d\gamma} = \frac{nk}{\gamma} - k\gamma^{k-1} \sum_i^n (x_i-\theta)^k \quad (67)$$

$$\frac{dL}{dk} = \frac{n}{k} + n\log\gamma \sum_i^n (x_i-\theta)^k - \gamma^k \sum_i^n (x_i-\theta)^k \log(x_i-\theta) \quad (68)$$

$$\frac{dL}{d\theta} = k\gamma^k \sum_i^n (x_i-\theta)^{k-1} - (k-1)\sum_i^n \frac{1}{(x_i-\theta)} \quad (69)$$

The maximum likelihood estimates (MLE), $\hat{k}, \hat{\gamma}, \hat{\theta}$ for the parameters k, γ, θ respectively, are obtained by setting (67) - (69) to zero and solving them simultaneously.

6 APPLICATIONS

In this section, we shall be investing the importance of the new distribution W-HLa{E} distribution and apply it to three real-life data. In these applications, the maximum likelihood estimation method is used in these two applications to estimate the parameters of fitted distributions. The maximized log-likelihood, the Kolmogorov-Smirnov test (K-S) along with the corresponding p-value, the Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) are reported to compare the W-HLa{E} distribution with the other distributions.

DATA SET I: Remission times of bladder cancer patients

Remission Time is the time taken for the signs and symptoms of a particular disease, in this case, cancer, to decrease or disappear after treatment. Though cancer may be considered in remission, cancer cells may remain in the body.

Table 2: *Remission times (in months) of bladder cancer patients' data*

0.080 0.200 0.400 0.500 0.510 0.810 0.900 1.050 1.190 1.260 1.350 1.400 1.460 1.760 2.020
2.020 2.070 2.090 2.230 2.260 2.460 2.540 2.620 2.640 2.690 2.690 2.750 2.830 2.870 3.020
3.250 3.310 3.360 3.360 3.480 3.520 3.570 3.640 3.700 3.820 3.880 4.180 4.230 4.260 4.330
4.340 4.400 4.500 4.510 4.870 4.980 5.060 5.090 5.170 5.320 5.320 5.340 5.410 5.410 5.490
5.620 5.710 5.850 6.250 6.540 6.760 6.930 6.940 6.970 7.090 7.260 7.280 7.320 7.390 7.590
7.620 7.630 7.660 7.870 7.930 8.260 8.370 8.530 8.650 8.660 9.020 9.220 9.470 9.740 10.06
10.34 10.66 10.75 11.25 11.64 11.79 11.98 12.02 12.03 12.07 12.63 13.11 13.29 13.80 14.24
14.76 14.77 14.83 15.96 16.62 17.12 17.14 17.36 18.10 19.13 20.28 21.73 22.69 23.63 25.74
25.82 26.31 32.15 34.26 36.66 43.01 46.12 79.05

Table 2 shows data of remission times(in month) of 128 bladder cancer patients selected at random as reported by Lee, et al [12], which was studied by Zea [16] to compare the fits of a different family of beta-Pareto(BP) and beta exponentiated Pareto (BEP) distributions. Almheidat [4] also applied to four parameters Cauchy-Weibull $\{logistic\}$ $(C-W\{L\})$ distribution in fitting this same data and just of recent Aldeni [2] applied uniform-exponential$\{generalisedlambda\}$ distribution $(U-E\{GL\})$ distribution to fit the same data

Table 3: *Descriptive Statistics of remission times of bladder cancer patients distributions*

Min.	Max.	Mean	1st Qu.	Median	3rd Qu.	Skewness	Kurtosis	SD
0.080	79.050	9.366	3.348	6.395	11.838	3.287	18.483	10.508

Table 4: *Performance of the distributions remission times of bladder cancer patients distributions Parameter estimates: Log-likelihood, AIC, and p-value (Standard errors in parentheses)*

Distributions	$WHLa\{E\}$	$U-E\{GL\}$	$C-W\{L\}$	BEP	BP
	k = 0.4847 (0.0267)	θ=0.2757 (0.0665)	α=2.3040 (1.0937)	a = 0.348 (0.0970)	a = 4.805 (0.0550)
	$\gamma = 0.4850$ (0.0651)	$\lambda_3 = 2.504$ (0.9285)	$\beta = 2.0205$ (0.4585)	b =159831 (183.7501)	b = 100.502 (0.2510)
	$\mu = 0.0785$ (-)	$\lambda_4 = 0.2894$ (0.0858)	k=3.0673 (0.7319)	k= 0.051 (0.0190)	k = 0.011 (0.0010)
			$\lambda = 12.663$ (2.6326)	$\beta = 0.080$ (2.0930)	$\beta = 0.080$

Table 5: *Performance of the distributions remission times of bladder cancer patients distributions Parameter estimates: Log-likelihood, AIC, and p-value*

Distributions	$WHLa\{E\}$	$U-E\{GL\}$	$C-W\{L\}$	BEP	BP
-Loglikelihood	271.3276	409.45	416.0965	432.41	480.446
AIC	548.6552	824.9	840.2	874.819	968.893
K-S	0.0078125	0.02876	0.06672	0.142	0.217
P-value	0.9922	0.9999	0.6189	0.0121	1.11E-05

Table 3 shows the summary of the dataset I and table 4 displaces the parameters estimates of W-HLa$\{E\}$ and four other distributios. Table 5 shows the values of parameter estimates, log-likelihood, AIC, K-S, and its p-value at 95%. Based on the above test statistics, W-HLa$\{E\}$ has the least AIC with 548.6552 and K-S Statistic (0.0078125) hence W-HLa{E} performed best among the five distribution models applied to remission time of bladder cancer. This implies that the new distribution can fit skewed data with long-tail better than any distribution. After an appropriate distribution has been identified and parameters estimated, we can estimate the probability of having a given duration of remission and other probabilities. For example, the probability of having a remission time longer than 10 months can be predicted as $P(X > x) = e^{-\gamma^k(x-\theta)^k}$ When $k = 0.4847, \gamma = 0.4850, \theta = 0.0785$

$$P(X > x) = e^{-0.4850^{0.4847}(10-0.0785)^{0.4847}}$$

$$P(X > 10) = 0.117$$

Data Set II: 72 pigs infected by virulent tubercle Bacilli (Bjerkedal, T, 1960)

The data in Table 6 are survival times (in days) of seventy-two pigs infected by virulent tubercle bacilli [9] The data in Table 6 are survival times (in days) of seventy-two pigs infected by virulent tubercle bacilli reported by Tahir [15] The Tables 4.4 shows the performance of W-HLa{E} and three other models (Logistic Frechet (LFr), Marshall-Olkin Frechet (MOFr), exponentiated-Frechet (EFr) and Frechet (Fr).

Table 6: *Infected Pigs data (in day)*

10 33 44 56 59 72 74 77 92 93 96 100 100 102 105 107 107 108 108 108 109 112
113 115 116 120 121 122 122 124 130 134 136 139 144 146 153 159 160 163 163 168
171 172 176 183 195 196 197 202 213 215 216 222 230 231 240 245 251 253 254 254
278 293 327 342 347 361 402 432 458 555

Table 7: *Descriptive Statistics of Infected Pigs data*

Min.	Max.	Mean	1st Qu.	Median	3rd Qu.	Skewness	Kurtosis	SD
43.00	598.00	141.85	82.75	102.50	149.25	2.5153	9.332	109.209

In this application, we obtain the descriptive statistics, maximum likelihood estimates of the parameters of the fitted distributions, and the values of the following statistics: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), and HQIC (Hanna Quinn Information Criterion).

In addition, we compute some goodness of fit statistics to verify which distribution provides the best fit to the data sets. We apply Kolmogorov-Smirnov (K-S) statistics. These statistics are described in detail in Tables 8 and 9. In general, the smaller the value of these statistics, the better the fit of the data by the distribution.

Table 8: *Parameter Estimates and Standard errors in parentheses for W-HLa{E} Distribution*

Distributions	$WHLa\{E\}$	LFr	$MOFr$	EFr
	$\theta = 42.99956$ (16.8814)	λ=32.5054 (0.0665)	a=212.7251 (115.7064)	a= 155.680 (5.7063)
Parameter Estimates	$\gamma = 0.4850$ (0.0651)	$\lambda_3 = 2.504$ (0.9285)	$\beta = 2.0205$ (0.4585)	$b = 159831$ (183.7501)
	$\mu = 0.0785$ (-)	$\lambda_4 = 0.2894$ (0.0858)	k=3.0673 (0.7319)	k= 0.051 (0.0190)
			$\lambda = 12.663$ (2.6326)	$\beta = 0.080$ (2.0930)

$skewness = 1.841076$ and $kurtosis = 7.49277$

Table 9: *Log-likelihood, AIC, BIC, HQIC, KS Statistic and p-value of 72 pigs infected by virulent tubercle Bacilli of different distributions*

Distributions	$WHLa\{E\}$	LFr	$MOFr$	EFr
-Loglikelihood	410.3371	426.2306	426.6764	449.7452
AIC	826.6741	858.4612	859.3527	903.4903
BIC	833.5041	865.2912	866.1827	908.0437
$HQIC$	829.3932	861.1803	866.1827	905.3030
K-S	0.027778	0.0695	0.1213	0.0923
P-value	0.9460	0.8773	0.2403	0.5710

From table 9, the new proposed W-HLa{E} model corresponds to the lowest values of the loglikelihood, AIC, BIC, HQIC, and K-S statistics among the fitted LFr, MOFr, and EFr models and therefore the W-HLa{E} model can be chosen as the best for the data set above. The distribution with the lowest Akaike Information Criteria (AIC) or BIC and the lowest Log-likelihood value is declared as ˜" best fit" distribution. In this case, W-HLa{E} distribution has the lowest Log-likelihood of -410.3371 with the lowest corresponding lowest AIC value of 826.6741. Hence, W-HLa{E}D is regarded as a best-fit model for this particular data used.

7. CONCLUSION

The research work introduced a new probability distribution called Weibull Half Laplace exponential distribution. Expressions for the probability density function, cumulative distribution function, survival function, and hazard function, and cumulative hazard function of the proposed distribution are derived. Some properties of the proposed distribution such as moments, order of statistics, and Shannon entropy have been studied. The simulation also supported the theoretical expression of the statistical properties of the proposed distribution such as the location parameter does not affect the variance, skewness, and kurtosis of the new distribution. From table 1, when k=2 and γ=2, the mean is approximately equal to the median hence, the distribution tends to be symmetric. The maximum likelihood method is adopted to estimate the parameters of the distribution. Coefficient of variation is deeply affected by γ. When the γ increases positively, the value of the coefficient of variation decreases It is shown, by means of two real data sets.

Two life data were applied to the new distribution and others established distributions by researchers in the field of probability distribution and we found out that W-HLa{E} distribution has the lowest log-likelihood and the smallest AIC, BIC, and HQIC for the two data sets in tables 6 and 9. W-HLa{E} distribution is a better fit for the two data sets.

References

[1] Abramowitz, M, Stegun, IA (1964). Handbook of mathematical functions with formulas, Graphs and mathematical tables, Dover Publication, Inc., New York.

[2] Aldeni, M., Lee, C. and Famoye, F. (2017). Families of distributions arising from the quantile of Generalized lambda distribution. Journal of Statistical Distributions and Applications.

[3] Aljarrah, MA, Lee, C. and Famoye, F (2014): On generating T-X family of distributions using quantile functions. J. Stat. Distrib. Appl. 1, 1–17

[4] Almheidat, M., Famoye, F. and Lee, C. (2015). Some Generalized Families of Weibull Distribution: Properties and Applications, International Journal of Statistics and Probability;4(3). 18-35.

[5] Alzaatreh, A., Lee, C., Famoye, F. (2013). A new method for generating families of continuous distributions, METRON 71,63–79. DOI 10.1007/s40300-013-0007-y

[6] Alzaatreh, A., Lee, C. and Famoye, F. (2014). T-normal family of distributions: a new approach to generalize the normal distribution. J. Stat. Distrib. Appl. 1, pp. 1–16.

[7] Alzaatreh, A, Lee, C. and Famoye, F. (2015). Family of generalized gamma distributions: properties and applications. To appear in Hacettepe Journal of Mathematics and Statistics

[8] Amalare, A. A., Ogunsanya, A. S. Ekum, M. I. and Owolabi, T. (2020). Lomax Cauchy{Uniform} Distribution: Properties and Application to Exceedances of Flood Peaks of Wheaton River, Benin Journal of Statistics, 3, 66– 81.

[9] Bjerkedal, T (1960): Acquisition of Resistance in Guinea Pigs Infected with Different Doses of Virulent Tubercle Bacilli. Am. J. Epidemiol. 72, 130–148.

[10] Job, O. and Ogunsanya, A. S. (2022). Weibull Log Logistic {Exponential} Distribution: Some Properties and Application to Survival Data. International Journal of Statistical Distributions and Applications. 8(1), 1-13. doi: 10.11648/j.ijsd.20220801.11

[11] Kotz, S. Kozubowski, T. J. and Podgorski, K. (2001): The Laplace Distribution and Generalizations A Revisit with New Applications, Springer, ISBN: 0-8176-4166-1

[12] Lee, E.T. and Wang, J.W. (2003) Statistical Methods for Survival Data Analysis. John Wiley: New York

[13] Ogunsanya, A. S., Akarawak, E. E. and Ekum, M. I. (2021a) On some properties of Rayleigh Cauchy distribution, Journal of Statistics and Management Systems, DOI: 10.1080/09720510.2020.1822499

[14] Ogunsanya, A.S., Yahya, W. B., Adegoke, T. M., Iluno, C., Aderele, O.R. and Ekum M. I. (2021b). A New Three-parameter Weibull Inverse Rayleigh Distribution: Theoretical Development and Applications. Mathematics and Statistics, 9(3), 249 - 272. DOI:10.13189/ms.2021.090306.

[15] Tahir, M. H., Mansoor, M. and Zubair, M. (2014). McDonald Log-Logistic Distribution with an Application to Breast Cancer Data. Journal of Statistical Theory and Applications. 13(1): 65-82

[16] Zea, L.M., Silva, R.B., Bourguignon, M., Santos, A.M & Cordeiro, G.M (2012): The Beta Exponentiated Pareto distribution with application to bladder cancer susceptibility. International Journal of Statistics and Probability. 1(2), 8–21.

APPROXIMATE OPTIMUM STRATA BOUNDARIES FOR PROPORTIONAL ALLOCATION USING RANKED SET SAMPLING

Khalid Ul Islam Rather[1]*, S.E.H Rizvi[2], Manish Sharma[3], M. Iqbal Jeelani[4], Faizan Danish[5]

•

Division of Statistics and Computer Science, SKUAST-Jammu, India[1,2,3]
Division of Social and Basic Sciences Faculty of Forestry, SKUAST-Kashmir, India [4]
Department of Mathematics, School of Advanced Science(SAS) VIT-AP India[5]
khalidstat34@gmail.com
sehrizvistat@gmail.com
manshstat@gmail.com
jeelani.miqbal@gmail.com
danishstat@gmail.com

Abstract

Ranked set sampling is an approach to data collection originally combines simple random sampling with the field investigator's professional knowledge and judgment to pick places to collect samples. Alternatively, field screening measurements can replace professional judgment when appropriate and analysis that continues to stimulate substantial methodological research. The use of ranked set sampling increases the chance that the collected samples will yield representative measurements. This results in better estimates of the mean as well as improved performance of many statistical procedures. Moreover, ranked set sampling can be more cost-efficient than simple random sampling because fewer samples need to be collected and measured. The use of professional judgment in the process of selecting sampling locations is a powerful incentive to use ranked set sampling. Optimum stratification is the method of choosing the best boundaries that make strata internally homogeneous, given some sample allocation. In order to make the strata internally homogenous, the strata should be constructed in such a way that the strata variances for the characteristic under study be as small as possible. This could be achieved effectively by having the distribution of the study variable known and create strata by cutting the range of the distribution at suitable points. If the frequency distribution of the study variable is unknown, it may be approximated from the past experience or some prior knowledge (auxiliary information) obtained at a recent study. The present investigation deals with paper the problem of optimum stratification on an auxiliary variable for proportional allocation under ranked set sampling (RSS), when the form of the regression of the estimation variable on the stratification variable given the variance function is known. A cum rule of finding approximately optimum strata boundaries has been developed. Further, empirical study has been made and presented along with relative efficiency which showed remarkable gain in efficiency as compared to unstratified RSS.

Keywords: Ranked set Sampling, approximately optimum strata boundaries, auxiliary variable optimum strata width

1. Introduction

The main aim of stratification is to produce estimators with more precision when a population characteristic is under study. The problem of obtaining optimum strata boundaries (OSB) taking study variable itself as stratification variable was first considered [1]. The study showed that for a given method of allocation, the variance is clearly a function of the strata boundaries. By minimizing the variance of the estimate of population mean sets of equations were obtained, solutions to which gave optimum strata boundaries. Due to the implicit nature of the minimal equations, their exact solutions could not be obtained. Subsequently, various authors gave methods of obtaining approximations to the exact solutions of the minimal equations. However, the ideal situation is that the distribution of such a study variable is known and the OSB can be determined by placing boundaries on the range of this distribution at suitable cut points. For an excellent account of these investigations reference may be made [2]. When the information about the study variables are not known, the utilization of auxiliary variable as stratification variable may be considered and obtained approximate OSB under simple random sampling (SRS) studied [3]. The problem of optimum stratification for two characters under study using auxiliary information [4]. A methodology developed for obtaining AOSB using auxiliary information under compromise method of allocation [5]. The problem of obtaining OSB under proportional allocation with varying cost of each unit studied [6]. A technique proposed under Neyman allocation when the stratification is done on the two auxiliary variables under consideration [7]. Recently, the problem considered of optimum stratification for a model-based allocation under a super population model [8]. Situation considered of optimum stratification of heteroscedastic populations in stratified selection for a known allocation under SRS strategy [9]. Estimator for the population Mean under Ranked Set Sampling [10] and [11]. Situation of optimum stratification under RSS considered by [12] and [13]. Further, the selection of sample from each stratum could be taken by utilizing any sampling technique. Therefore, in the present investigation we have taken the procedure of ranked set sampling (RSS) as a method of selection of units from each stratum, which is more efficient as compared to SRS. A stratified ranked set sample (SRSS) is a sampling plan in which a population is divided into mutually exclusive strata and a RSS of n elements is quantified within each stratum. The sampling is performed independently across the strata. Therefore, one can think of an SRSS scheme as a collection of L separate ranked set samples. Originally, RSS was first suggested to estimate mean pasture and forage yields [14]. The necessary mathematical theory provided [15].

Let the population under consideration be divided into L strata and a sample $n_{0h} = (R_h \times n_h)$ units which is selected from h^{th} stratum is drawn using RSS, where R_h is the number of cycles and n_h is sample size of each cycle. Each sample element is measured with respect to some variable Y, and estimator of the population mean is given by

$$\bar{y}_{SRSS} = \sum_{h=1}^{L} \frac{W_h}{n_{0h}} \left[\sum_{j=1}^{R_h} \sum_{i=1}^{n_h} \bar{y}_{ij(r)} \right] \tag{1.1}$$

where W_h is the weight of the h^{th} stratum and $\bar{y}_{ij(r)}$ is the sample mean based on n_{0h} units drawn from the h^{th} stratum.
If the finite correction is ignored, the variance of the estimate will be:

$$V\left(\bar{y}_{SRSS}\right)_{prop} = \frac{1}{n}\left[\sum_{h=1}^{L} W_h \sigma^2_{h(r)}\right] \tag{1.2}$$

$\sigma^2_{h(r)} = \left(\sigma^2{}_{hc} - \frac{1}{n}(\mu_i - \mu)^2\right)$ denotes the variance of r^{th} order statistics in h^{th} stratum of the random sample of size n_h.

In most of these investigations related to optimum stratification, both the estimation and stratification variables are taken to be the same. Since the distribution of the estimation variable $'Y'$ is rarely known in practice, it is desirable to stratify on the basis of some suitably chosen concomitant variable $'X'$. An investigation has considered the general problem of optimum stratification based on auxiliary variables for the case of proportional allocation [16]. We consider the problem of optimum stratification on the auxiliary variable 'X', assuming knowledge about the form of the regression of 'Y' on 'X' and the variance function $V(y \mid x)$, minimal equations giving optimum strata boundaries have been obtained for proportional allocations under ranked set sampling. Since these equations cannot be solved easily, various methods of finding approximations to the exact solutions have been given.

In this paper, the problem of construction of strata boundaries will be dealt using classical approach when the sample is selected from the strata using RSS.

MINIMAL EQUATIONS UNDER PROPORTIONAL ALLOCATION

If the regression of the estimation variable 'Y' on the stratification variable 'X', in the infinite super population is given by

$$y = c(x) + e \tag{2.1}$$

Where '$c(x)$' is a function of auxiliary variable, 'e' is the error term such that $E(e \mid x) = 0$ and $V(e \mid x) = \eta(x) > 0 \ \forall x \in (a,b)$ with $(b-a) < \infty$. Let $f(x,y)$ and $f(x)$ be the joint density function and marginal density function of (x,y) and x respectively. Then, we have

$$W_h = \int_{x_{h-1}}^{x_h} f_i(x) \ , \ \mu_{hc} = \frac{1}{W_h} \int_{x_{h-1}}^{x_h} c(x) f_i(x) \ \text{ and } \ \sigma^2{}_{hy} = \sigma^2{}_{hc} + \mu_{h\eta} \ , \ (h = 1,2,3\ldots,L) \tag{2.2}$$

where (x_{h-1}, x_h) are lower and upper boundaries of the h^{th} stratum with $(x_0 = a)$ and $(x_L = b)$, $\mu_{h\eta}$ is the expected value of $\eta(x)$ and σ^2_{hc} is the variance of $c(x)$ in the h^{th} stratum.

Using these relations, the variance expression (1.2), under proportional allocation

$$V(\bar{y}_{SRSS})_{prop} = \frac{1}{n}\left[\sum_{h=1}^{L} W_h \sigma^2_{hc(r)} + \mu_{h\eta}\right] \tag{2.3}$$

Let $[x_h]$ denote the set of optimum points of stratification on the range (a,b), for which the $V(\bar{y}_{SRSS})$ is minimum. These points $[x_h]$ are the solutions of the minimal equations which are obtained by equating to zero the partial derivatives of $V(\bar{y}_{SRSS})$ with respect to $[x_h]$.

Minimization of the variance expression given in (2.3), is equivalent to the minimization of the expression $\sum_{h=1}^{L} W_h \sigma^2_{hc(r)}$, since $\sum_{h=1}^{L} W_h \mu_{h\eta} = \mu_{h\eta}$ is a population parameter and is a constant. On equating to zero the partial derivative of this expression with respect to $[x_h]$, we get

$$\frac{W_h f(x_h)\left[(c(x_h) - \mu_{hc(r)})^2 - \sigma^2_{hc(r)}\right]}{W_h + f(x_h)\sigma^2_{hc(r)}} = \frac{W_i f(x_h)\left[(c(x_h) - \mu_{ic(r)})^2 - \sigma^2_{ic(r)}\right]}{W_i + f(x_h)\sigma^2_{ic(r)}}$$

Therefore, using these results we get the minimal equations on simplification as

$$c(x_h) = \left(\frac{\mu_{hc(r)} + \mu_{ic(r)}}{2}\right), i = h+1, h = 1,2,\ldots,L-1 \tag{2.4}$$

These equations are implicit functions of the strata boundaries $[x_h]$ and their exact solutions are somewhat difficult to find. Therefore, we proceed to find the method of solving these minimal equations by conducting approximations.

APPROXIMATE EXPRESSIONS FOR CONDITIONAL MEAN AND VARIANCE

Let the functions $f_i(x)$, $c(x)$ and $\eta(x)$ are bounded away from zero and possess first two derivatives continuous $\forall x \in (a,b)$. Then, we have the following identities due to [17].

$$I_i(y,x) = \int_Y^X (t-y)^i f_i(t)dt = \sum_{j=0}^{3} \frac{(k)^{i+j+1}}{j!(i+j+1)} f^{(j)} + O(k^{i+5}) \tag{3.1}$$

where $f^{(j)}$ is the j^{th} derivative of $f_i(t)$ at $t = y$ and $k = x - y$

$$I_i(y,x) = \int_Y^X (t-y)^i f_i(t)dt = \sum_{j=0}^{3} \frac{(-k)^{i+j+1}}{j!(i+j+1)} f^{(j)} + O(k^{i+5}) \tag{3.2}$$

$O(k^i)$ is the higher order terms with power $\geq i$

Let $\mu_\eta(y,x)$ denote the conditional expectation of function $\eta(t)$ in the interval (y,x), so that

$$\mu_\eta(y,x) = \frac{\int_y^x \eta(t) f_i(t)dt}{\int_y^x f_i(t)dt} \tag{3.3}$$

we have from the definition of $\mu_\eta(y,x)$

$$\mu_\eta(y,x)\int_y^x f_i(t)dt = \int_y^x \eta(t) f_i(t)dt)$$

therefore, we have

$$\mu_\eta(y,x).I_0(y,x) = \left[\eta I_0(y,x) + \sum_{i=1}^{3} \eta^{(i)} I_i(y,x)/i!\right] + O(k^5)$$

Using the Taylor series expansions for $I_i(y,x)$ and $I_i(x,y)$ from (3.2) and simplifying the result at point $t = y$, we have

$$\mu_\eta(y,x) = \eta\left[1 + \frac{\eta'}{2\eta}k + \frac{(\eta' f' + 2f\eta'')}{12f\eta}k^2 + \frac{(ff''\eta' + ff'\eta'' + f^2\eta''' - \eta' f'^2)}{24f^2\eta}k^3 + O(k^4)\right] \tag{3.4}$$

Proceeding in the same fashion using Taylor series expansions about the point 'x', the expression for $\mu_\eta(y,x)$ is obtained as

$$\mu_\eta(y,x) = \eta\left[1 - \frac{\eta'}{2\eta}k + \frac{(\eta' f' + 2f\eta'')}{12f\eta}k^2 - \frac{(ff''\eta' + ff'\eta'' + f^2\eta''' - \eta' f'^2)}{24f^2\eta}k^3 + O(k^4)\right] \tag{3.5}$$

Let $\sigma^2{}_\eta(y,x)$ denotes the conditional variance of the function $\eta(t)$ in the interval (y,x), we have

$$\sigma^2{}_\eta(y,x) = \mu_{\eta^2}(y,x) - (\mu_\eta(y,x))^2$$

substituting values, we get

$$\sigma^2{}_\eta(y,x) = \frac{k^2(\eta'(y))^2}{12}\left[1 + (\eta''(y)/\eta'(y))k + O(k^2)\right] \tag{3.6}$$

Using the above results, several other approximations can be obtained. Multiplying the series expansions for $\mu_\eta(y,x)$ about the points $t = y$, $t = x$ and taking the square root, we obtain

$$\mu_\eta(y,x) = \sqrt{\eta(y)\eta(x)}\left[1 + O(k^2)\right] \tag{3.7}$$

From (3.3), we have

$$\mu_\eta(y,x).I_0(y,x) = \int_y^x \eta(t) f_i(t)dt$$

Taking $\eta(t) = t^2$ and using (3.7), we get

$$\int_y^x f_i(t)dt = \frac{1}{xy}\int_y^x t^2 f_i(t)dt\left[1 + O(k^2)\right] \tag{3.8}$$

Similarly expanding $\sqrt[\lambda]{f_i(t)}$ about the point t=y, we have

$$\left[\int_Y^X \sqrt[\lambda]{f_i(t)}dt\right]^\lambda = k^{\lambda-1}\int_Y^X f_i(t)dt\left[1 + O(k^2)\right] \tag{3.9}$$

APPROXIMATE SOLUTIONS OF THE MINIMAL EQUATIONS

To find approximate solutions to the minimal equation (2.4), we shall obtain the series expansions of system of equations about the point $[x_h]$, the common boundary of h^{th} and $(h+1)^{th}$ strata. The expansions for the two sides of the equation (2.4) are obtained by using various results proved in the preceding section. For the expansion of the right hand side about the point x_h, (y,x) is replaced by (x_{h-1}, x_h) while for the left hand side we replace (y,x) by (x_{h-1}, x_h).

we have

$$[\mu_{hc(r)} - c(x_h)]^2 - [\mu_{ic(r)} - c(x_h)]^2 = 0$$

We have from (2.13) after replacing $'y'$ and $'x'$ by x_h and x_{h+1} respectively.

$$\mu_{ic(r)} = c\left[1 + \frac{c'}{2c}k_i + \frac{(c'f' + 2fc'')}{12fc}k_i^2 + \frac{(ff''c' + ff'c'' + f^2c''' - c'f'^2)}{24f^2c}k_i^3 + O(k_i^4)\right]$$

The derivatives of c and f are evaluated at x_h. Therefore, we get

$$[\mu_{ic(r)} - c(x_h)] = \frac{k_i}{2}\left[c' + \frac{(c'f' + 2fc'')}{6f}k_i + O(k_i^2)\right] \tag{4.1}$$

Similarly, we get

$$\left[\mu_{hc(r)} - c(x_h)\right] = \frac{k_i}{2}\left[c^{'} + \frac{\left(c^{'} f^{'} + 2fc''\right)}{6f} k_i + O(k_i^{\,2})\right] \tag{4.2}$$

Therefore from (4.1) and (4.2), we have

$$\frac{k_i}{2}\left[c^{'} + \frac{\left(c^{'} f^{'} + 2fc''\right)}{6f} k_i + O(k_i^{\,2})\right] = \frac{k_i}{2}\left[c^{'} + \frac{\left(c^{'} f^{'} + 2fc''\right)}{6f} k_i + O(k_i^{\,2})\right]$$

Now let us consider an expansion of the function

$$B_h = \int_{x_{h-1}}^{x_h} c^{'2} f_i(t)dt$$

$$\therefore B_h = c^{'2} f k_h \left[1 - \frac{\left(c^{'} f^{'} + 2fc''\right)}{fc^{'}} \frac{k_h}{2} + O(k_h^2)\right]$$ (Using Taylors expansion)

multiplying it by $\left(\frac{k_h^2}{8}\frac{c^{'}}{f}\right)$ and taking cube root both sides, we get

$$\left[\frac{k_h^2}{8}\frac{B_h c^{'}}{f}\right]^{\frac{1}{3}} = \frac{c^{'} k_h}{2}\left[1 - \frac{\left(c^{'} f^{'} + 2fc''\right)}{6fc^{'}} k_h + O(k_h^2)\right] \tag{4.3}$$

similarly, we have

$$\left[\frac{k_i^2}{8}\frac{B_i c^{'}}{f}\right]^{\frac{1}{3}} = \frac{c^{'} k_i}{2}\left[1 - \frac{\left(c^{'} f^{'} + 2fc''\right)}{6fc^{'}} k_i + O(k_i^2)\right] \tag{4.4}$$

From the equations (4.3) and (4.4), we have

$$\left[\frac{k_h^2}{8}\frac{B_h c^{'}}{f}\right]^{\frac{1}{3}} = \left[\frac{k_i^2}{8}\frac{B_i c^{'}}{f}\right]^{\frac{1}{3}} \tag{4.5}$$

Or

$$k_h^2 B_h = \text{constant} \tag{4.6}$$

In case it is possible to find a function $Q_2\left(x_{h-1}, x_h\right)$ such that

$$k_h^2 B_h = k_h^2 \int_{x_{h-1}}^{x_h} c^{'2} f_i(t)dt$$

$$= Q_2\left(x_{h-1}, x_h\right)\left[1 + O(k_h)^2\right] \tag{4.7}$$

Thus the system of equations (4.6) to the same degree of accuracy can be put as

$Q_2\left(x_{h-1}, x_h\right) =$ Constant

the above results can be put in the form of theorem as follows.

THEOREM :-

If the regression of the estimation variable $'Y'$ on the stratification variable $'X'$, in the infinite super population is given by

$$y = c(x) + e$$

where $'c(x)'$ is a function of auxiliary variable, $'e'$ is the error term such that $E(e \mid x) = 0$ and $V(e \mid x) = \eta(x) > 0 \ \forall x \in (a,b)$ with $(b-a) < \infty$, and further if the function $g_1(x)\ f_i(x) \ \in \Omega$: then the system of equations (2.4) give strata boundaries (x_h) which correspond to the minimum of $V(\overline{Y}_{stRSS})_{prop}$ can be written as

$$\left[k_h^2 \int_{x_{h-1}}^{x_h} c'^2 f_i(t)dt.\left[1+O(k_h^2)\right]\right]^{\frac{1}{3}} = \left[k_i^2 \int_{x_h}^{x_{h+1}} c'^2 f_i(t)dt.\left[1+O(k_i^2)\right]\right]^{\frac{1}{3}}$$

Neglecting the terms of order $O\left(Sup_{(a,b)}(k_h)\right)^3$ can be neglected; these equations can be replaced by the approximate system of equations

$$k_h^2 \int_{x_{h-1}}^{x_h} c'^2 f_i(t)dt = \text{constant}$$

Or equivalently by

$$Q_2(x_{h-1}, x_h) = \text{constant} \quad , \quad k_h = (x_h - x_{h-1})$$

$$Q_2(x_{h-1}, x_h)\left[1+O(k_i^2)\right] = k_h^2 \int_{x_{h-1}}^{x_h} c'^2 f_i(t)dt \quad , \ i = h+1, h = 1,2,\ldots,L$$

The similar results can also be obtained by minimizing the function

$$\sum_{h=1}^{L} \int_{x_{h-1}}^{x_h} c'^2 f_i(t)dt\left[1+O(k_h^2)\right]$$

Thus we find that if the function $c'^2 f_i(x)$ belongs to Ω, the minimum value of $\sum_{h=1}^{L} W_h \sigma_{hc}^2$ and therefore $V(\overline{Y}_{stRSS})_{prop}$, exists and the solutions of the system of equations (2.4) or equivalently of (4.5). These equations as such are very difficult to solve and therefore it is essential to find some way out of this difficulty. It is done by replacing these systems of equations by other systems of equations which are comparatively easier to solve but are only asymptotically equivalent to the exact minimal equations. The error factor is introduced because we neglect the terms of higher powers of strata widths which is of course justifiable if the number of strata is large. We have obtained these systems of equations after neglecting the terms of order $O\left(Sup_{(a,b)}(k_h)\right)^4 = O\left(m^4\right)$ where $m = Sup_{(a,b)}(k_h)$, on both sides of the equation (4.5). If the number of strata is large and therefore terms of order $O\left(m^4\right)$ are quite small, the error involved in the approximate systems of equations is expected to be quite small and the set of points (x_h) obtained from them shall be quite near the optimum values.

Now we proceed to develop the approximate systems of equations given in (4.6) and (4.7). Here, in finding various forms of the function $Q_2(x_{h-1}, x_h)$, we shall keep in mind that the function $Q_2(x_{h-1}, x_h)$ is such that

$$k_h^2 \int_{x_{h-1}}^{x_h} g_1(t) f_i(t) dt = Q_1(x_{h-1}, x_h)[1 + O(k_i^2)]$$

If in (2.4) we retain only the first term on both sides of the equation and neglect the others, the two sides are equalized if

$$k_h = \text{constant}, \ \left(\frac{b-a}{L}\right)h, \quad \text{for} \quad h = 1,2,\ldots,L \tag{4.8}$$

and therefore $x_h = a + \left(\frac{b-a}{L}\right)h$, with $(x_0 = a)$ and $(x_L = b)$

This set of solutions cannot be expected to yield very good results as we have neglected terms of order $O(m^3)$ on both sides of the exact minimal equations. This solution holds for all $c'^2 f_i(x)$ provided they belong to Ω and all density functions with finite range. Due to its universality of application it can be recommended in case of less information about c'^2 and $f_i(x)$. Apart from this, it gives the strata boundaries at once without any difficulty that may arise even in solving the approximate systems of equations. This approximate method fails if the range of $'x'$ is infinite, but one can resort to truncation of the density function to any suitable probability level before using this approximation.

We obtain next approximate systems of equations, the optimum points of stratification are such that

$$k_h^2 \int_{x_{h-1}}^{x_h} c'^2 f_i(t) dt = c_1, \quad h = 1,2,\ldots,L \tag{4.9}$$

The solutions obtained from this approximation are expected to be quite close to the optimum points as only terms of $O(m^4)$ have been neglected. All the approximate systems that will now follow also give the points of stratification to the same degree of accuracy.

From (3.8) and equation (4.9) is obtained the following class of approximate equations. The approximations to optimum (x_h) are obtained from

$$\left[k_h^{(3\lambda-1)} \int_{x_{h-1}}^{x_h} c'^2 f_i(t)^\lambda dt\right]^{\frac{1}{\lambda}} = \text{constant}, \quad h = 1,2,\ldots,L$$

For λ = 1/2 and 1/3 we have

$$k_h\left[\sqrt{\int_{x_{h-1}}^{x_h} c'^2 f_i(t)^\lambda dt}\right] = c_2 = \text{constant}, \quad h = 1,2,\ldots,L \tag{4.10}$$

For λ =1/3, we have the system of equation as

$$\sqrt[3]{\int_{x_{h-1}}^{x_h}\left[c'^2 f_i(t)dt\right]} = c_3\,, \qquad h = 1,2,\ldots,L \tag{4.11}$$

In all these systems of equations, $c's$ are constants to be determined and in some cases, the few equations may be meaningless such as for $h=1$, i,e $x_{h-1}=0$.

Cum$\sqrt[3]{K_1(x)}$ Rule

If the $K_1(x) = c'^2 f_i(t)$ is bounded and its first two derivative exists $\forall x \in [a,b]$, then for given value of L taking equal intervals on the cumulative cube root of $K_1(x)$ will give AOSB.

EMPIRICAL STUDY:

We shall consider following distributions of auxiliary variable for evaluating the efficiency of the proposed method for obtaining optimum points of stratification.

Let us assume the auxiliary variable 'x' follows certain distributions as follows:

I. Rectangular $f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, otherwise \end{cases}$

II. Right-triangular $f(x) = \begin{cases} \dfrac{2(2-x)}{(b-a)^2}, & a \le x \le b \\ 0, otherwise \end{cases}$

III. Exponential $f(x) = e^{-x+1}$, $1 \le x \le \infty$

IV. Standard normal $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ $-\infty \le x \le \infty$

If the stratification variable follows the uniform distribution with pdf $f(x) = \dfrac{1}{b-a}$ $x \in [1,2]$, utilizing the cum $\sqrt[3]{K_1(x)}$ rule, we get the stratification points as given in table I.

Table I: *AOSB and Variance for uniformly distributed auxiliary variable*

L	AOSB	Total variance $n\{V(\bar{y}_{stRSS})_{Prop}\}$	R.E.%
2	1.4967	0.75521	102.069
3	1.3257, 1.6624	0.75232	102.461
4	1.2483, 1.4959, 1.7466	0.75151	102.572
5	1.1983, 1.3965, 1.5898, 1.7967	0.75084	102.663
6	1.1631, 1.3289, 1.4922, 1.6543, 1.8236	0.75058	102.698

If the stratification variable follows the right triangular distribution with pdf $f(x) = 2(2-x)$ $x \in [1,2]$, utilizing the cum $\sqrt[3]{K_1(x)}$ rule, we get the stratification points as given in table II.

Table II: *AOSB and Variance for Right-triangular distributed auxiliary variable*

L	AOSB	Total Variance $n\{V(\bar{y}_{stRSS})_{Prop}\}$	R.E.%
2	2.9662	1.01136	111.979
3	2.2935, 3.6419	0.97727	115.886
4	2.0325, 3.0001, 3.9986	0.96505	117.353
5	1.7954, 2.5879,3.4289, 4.2247	0.95773	118.250
6	1.6565, 2.3143, 2.4771, 3.6057, 4.2593	0.95589	118.477

If the stratification variable follows the Exponential distribution with pdf $f(x)=e^{-x+1}$ $x\in[1,5]$, utilizing the cum $\sqrt[3]{K_1(x)}$ rule, we get the stratification points as given in table III.

Table III: *AOSB and Variance when the auxiliary variable is exponentially distributed:*

L	AOSB	Total Variance $n\{V(\bar{y}_{stRSS})_{Prop}\}$	R.E.%
2	1.4949	0.67123	101.389
3	1.3257, 1.6558	0.66879	101.760
4	1.2589, 1.5173, 1.7814	0.66801	101.878
5	1.1981, 1.3971, 1.6452, 1.8494	0.66758	101.944
6	1.1638, 1.3277, 1.4911, 1.6562, 1.8217	0.66722	101.999

If the stratification variable follows the standard Normal distribution with pdf $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ $x\in[0,1]$, utilizing the $cum\sqrt[3]{K_1(x)}$ rule, we get the stratification points as given in table IV.

Table IV: *AOSB and Variance for Standard normally distributed*

L	AOSB	Total Variance $n\{V(\bar{y}_{stRSS})_{Prop}\}$	R.E.%
2	0.4939	0.08024	121.102
3	0.3282, 0.6531	0.07927	122.584
4	0.2448, 0.4947, 0.7509	0.07893	123.111
5	0.1973, 0.3943, 0.5993, 0.7986	0.07877	123.360
6	0.1687, 0.3286, 0.4944, 0.6596, 0.8269	0.07868	123.496

CONCLUSION

The AOSB are determined for this distribution by using cum $\sqrt[3]{K_1(x)}$ method. For each L = 2, 3, 4, 5 and 6 the variance $n\{V(\bar{y}_{stRSS})_{Prop}\}$ is calculated, which is used for the efficiency of the stratification. The results of this investigation are given in Table I-IV. When the auxiliary variable follows uniform, right triangular, exponential and standard normal distributions the stratification points obtained has been presented in Table I-IV and percent RE as compared unstratified RSS have been worked out. The standard normal distribution shows highest % R.E. Overall results show that the increase in the number of strata is directly proportional to the decrease in total variance. From the last column of tables it can be seen that the AOSB obtained by the proposed method are more efficient for all L = 2,3,...,6. Thus, the proposed method of cum $\sqrt[3]{K_1(x)}$ shows increases gain in precision in obtaining AOSB while selecting samples using RSS.

References

[1] Dalenius, T. (1950). The problem of optimum stratification. *Scandinavian Actuarial Journal, 33*(3-4): 203-213.

[2] Cochran, W. G. (1961). "Comparison of methods of determining strata boundaries," Bulletin of the International Statistical Institute, 38: 345-358.

[3] Singh, R. and Sukhatme, B.V., (1969). Optimum stratification. Annals of the Institute of Statistical Mathematics, 21, 515-528.

[4] Rizvi, S.E.H., Gupta, J.P. and Singh, R. (2000): Approximately optimum Stratification for two study variables using auxiliary information. *Journal of the Indian Society of the Agricultural Statistics.* 53(3): 287-298.

[5] Rizvi, S. E. H., Gupta, J. P. and Bhargava, M. (2002): Optimum stratification based on auxiliary variable for compromise allocation. *Metron*, 28(1), 201-215.

[6] Danish, F., Rizvi, S. E. H., Jeelani, M. I. and Reashi, J. A. (2017). Obtaining strata boundaries under proportional allocation with varying cost of every unit. Pakistan Journal of Statistics and Operation Research 13 (3):567–74.

[7] Danish, F. and Rizvi, S.E.H. (2018) Optimum stratification in Bivariate auxiliary variables under Neyman allocation, *Journal of Modern Applied Statistical Methods*, 17(1).

[8] Gupt. K, B, and Ahamed. I, M, (2021). Construction of Strata for a Model-Based Allocation Under a Super population Model. *Journal of Statistical Theory and Applications.* 20(1): 46-60.

[9] Gupt. K. B., Swer, M, Ahamed, I. M., Singh, B. K. and Singh, K. H., (2022). Stratification Methods for an Auxiliary Variable Model-Based Allocation under a Super-population Model. *Mathematics and Statistics,* 10(1): 15-24.

[10] Rather, K.U.I. and Kadilar, C. (2021). Exponential Type Estimator for the population Mean under Ranked Set Sampling. *Journal of Statistics: Advances in Theory and Applications,* 25(1), 1–12.

[11] Rather, K.U.I., Eda, K. G. Unal, C. & Jeelani, M. I. (2022). New exponential ratio estimator in Ranked set sampling. *Pakistan Journal of Statistics and operation research,* 18(2), 403–409.

[12] Rather, K.U.I., Rizvi, S.E.H., Sharma, M. and Jeelani, M. I. (2022): Optimum Stratification for Equal Allocation using Ranked Set Sampling as Method of Selection. International Journal of Statistics and applied mathematics. 7(2): 63-67.

[13] Rather, K.U.I. and Rizvi, S.E.H. (2022): Approximate Optimum Strata Boundaries for Equal Allocation under Ranked Set Sampling. *Journal of Scientific Research,* 66(3): 302-307.

[14] Mcintyre, G. A. (1952). A method for unbiased selective sampling, using ranked set sampling and stratified simple random sampling. *Journal of Applied Statistics* 23, 231-255.

[15] Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the stratified sampling by means of ordering. Annals of the Institute of Statistical Mathematics, 20, 1-31.

[16] Hayashi, C., Maruyama F. and Isida, M. D. (1951). "On some criteria for stratification," *Annals of the Institute of Statistical Mathematics,* 2: 77-86.

[17] Ekman, G. (1959). "Approximate expressions for the conditional mean and variance over small intervals of a continuous distribution," Annals of the Institute of Statistical Mathematics, 30: 1131-1134.

Study of reliability of the on-tether subsystem of a tethered high-altitude unmanned telecommunication platform

Dharmaraja Selvamuthu[a], Adwaith H Sivam[a], Raina Raj[a], Vladimir Vishnevsky[b]

•

Indian Institute of Technology Delhi, New Delhi, India[a]
dharmar@maths.iitd.ac.in
adwaithhs@gmail.com
rainaraj.curaj@gmail.com
V.A. Trapeznikov Institute of Control Science, Moscow, Russia[b]
vishn@inbox.ru

Abstract

High-altitude platform (HAP) are stations on an object at an altitude of around 15-50 km at a specified nominal fixed point relative to Earth. Tethered high-altitude platform (tHAP) are unmanned aerial vehicle that are connected to the ground via a tether with a lift height of $100 - 150$ *meters, and a multi-copter as high altitude mode. The reliability of the tHAP can be assessed with a focus on the tether that connects it to the ground. This article proposes a Markov model which obtain the reliability of the tHAP. The tether is considered to be made up of multiple wires in such a way that the tether still operates for a given number of functioning wires. The failure rates of the wires are dependent on the number of failed wires. Through the reliability analysis of the proposed Markov model, the key performance measures such as reliability of the system, mean time between failures and the probability of the system being reliable are computed. The optimal number of wires is also obtained via the numerical computation of the performance measures.*

Keywords: Markov chain, reliability, tethered high-altitude platform, unmanned aerial vehicle

1. Introduction

Autonomous unmanned aerial vehicles (UAVs) are currently in widespread use. The biggest disadvantage of UAVs is their limited operational period due to a limited energy resource of batteries or fuel-carrying capacity. High-altitude platform (HAP)s are stations on an object at an altitude of around 15-50 km at a specified nominal fixed point relative to Earth. The long-term operation can be provided by a tethered high-altitude platform (tHAP), with a lift height of 100-150 meters, in which power supply of engines and payload equipment are provided from ground-based power sources via copper cables [1].

tHAPs have several advantages when compared to terrestrial stations or satellite stations ([2] [3]). tHAPs have an intermediate geographical range between that of terrestrial stations and a satellite stations. tHAPs can be deployed in a matter of hours. They can be operated for long duration and can return to ground easily. Furthermore they have a secure and efficient back-haul, since tHAPs are connected to the core network via wire. tHAPs are also extensively used in the case of disaster situation in the concerned region due to the easy deployment.

Components of a tHAP are shown in Fig. 1. tHAPs consists of a flying platform connected to the Ground Control Station (GCS) via a tether. The GCS houses a system for diagnostics and control of the tHAP from the ground. The tether also includes power transmission cable and data transmission cable. The navigation and stabilization system comprises of on-board sensors, on-ground anchor points, etc., that measures various parameters for navigation, and the circuits that stabilize the platform based on these parameters. The main power source is located on the ground. The intelligent wench is responsible for controlling the length of the tether.

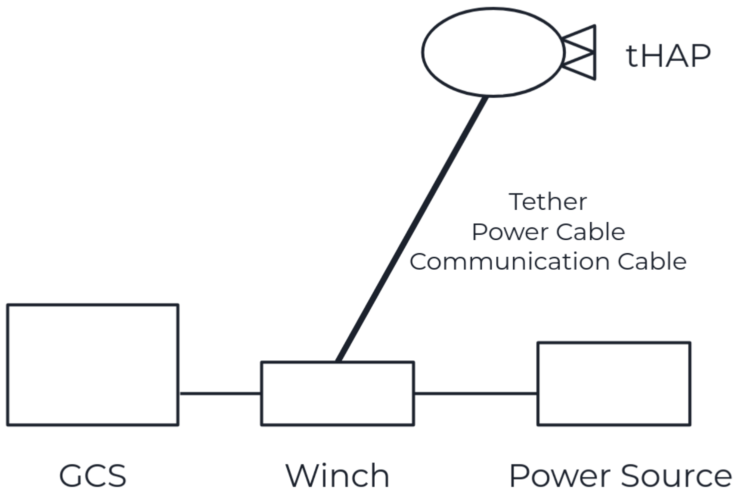


Figure 1: *Components of a tHAP*

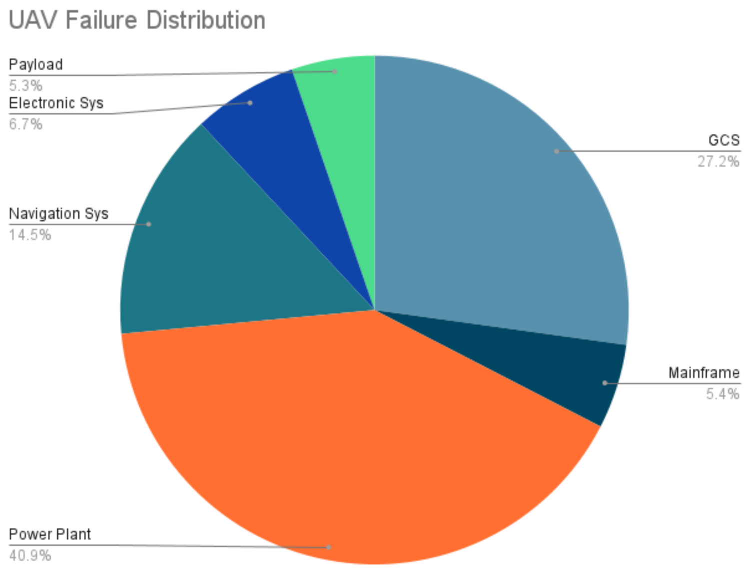


Figure 2: *UAV Failure Causes*

Fig. 2 shows the contribution of various factors in the failure of an ordinary UAV. Most of the failures can be attributed to the power system. These failures should be less frequent in the case of tHAP, since the power source is stationed on the ground and more reliable. The failures of navigation system should also be fewer since the tHAP is mostly stationary. The failure rate of GCS, electric system, mainframe and payload should not differ much. This data does not include the failures of the components in the tether [5]. The tether will be around 15-50km long. It will also have a total mass around a few hundred kg. There is a need to focus on improving the reliability of the tether that connects the tHAP to the ground as well as the power and communication cables. Vishnevsky et al. [7] studied the reliability of the functioning of a flight module of a tethered high-altitude telecommunication platform, utilizing the k-out-of-n : F model. The authors of [9] proposed an analytical model of the k-out-of-n:G system under two system failure scenarios. Ivanova [10] presented a hot standby repairable k-out-of-n system.

This study proposes a Markov model setup for the prediction of reliability of the tether. It has been assumed that the tether is constructed of multiple wires. The functioning of the tether depends on the number of working wires. If the number is below some level, the tether will stop working. To this end, numerical examples have been carried out. The obtained results have been found to be very close to the existing results of [6] which uses a different criterion - minimizing the weight of the cable.

The paper is divided into four parts. Section 2 provides a general overview and a broad description of the approach taken. Section 3 comprises of the analysis of the given approach using Markov model. Section 4 shows the various numerical results obtained based on the analysis. Section 5 provides the conclusions and future work.

2. Model description

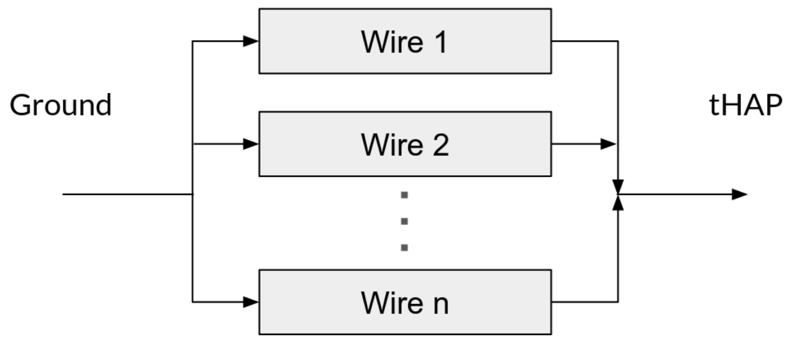


Figure 3: *Diagram of the proposed model*

Consider the cable be made of multiple parallel wires, such that the cable still operates provided a given number of these wires are still functioning. As seen in Fig. 3, the cable consists of n wires running parallel to each other. The wires have current flowing through them with different phases. Their phases will be symmetrically distributed such that the total current is zero. Whenever a wire has a fault, the phases of the currents in the remaining wires will be adjusted such that the total current is still zero.

Assume that a cable made up of n wires will still function reasonably provided k wires are still working. Whenever a wire fails there is, in general, a higher probability for the next wire to fail. In this work, take this into consideration by assuming that the failure rates of the wires are dependant on the number of wires already failed. Furthermore, assume that the failure rates of the remaining wires are the same. Let λ_i be failure rate of each wire after i wires have failed.

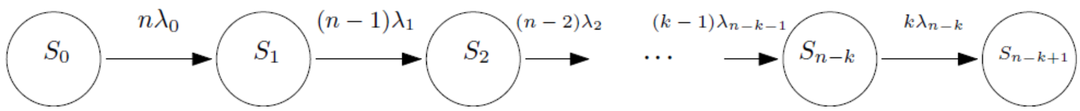


Figure 4: *State transition diagram for the proposed Markov model*

The system can be assessed using a Markov model [8]. The Markov model of the system after applying state aggregation is shown in Fig. 4. Here S_i $(0 \le i \le n-k)$ is the state in which exactly i wires failed and $n-i$ wires are working. S_{n-k+1} is the state in which more than k wires have faulted, i.e., the whole system fails.

3. Reliability Analysis

Let $P_n(t)$ be the probability that the system in the state S_n at time t. The governing equations are given by

$$\frac{dP_0(t)}{dt} = -n\lambda_0 P_0(t) \quad (1)$$

$$\frac{dP_i(t)}{dt} = (n-i+1)\lambda_{i-1}P_{i-1}(t) - (n-i)\lambda_i P_i(t), \;\; 1 \leq i \leq n-k \quad (2)$$

$$\frac{dP_{n-k+1}(t)}{dt} = k\lambda_{n-k}P_{n-k}(t). \quad (3)$$

Initial condition is given by $P_0(0) = 1, P_i(0) = 0, \;\; 1 \leq i \leq n-k+1$. Solving the above Equations (1- 3) with the initial condition, we get,

$$P_0(t) = e^{-n\lambda_0 t} \quad (4)$$

$$P_i(t) = \sum_{j=0}^{i} A_{ij} e^{-(n-j)\lambda_j t}, 0 < i \leq n-k \quad (5)$$

$$P_{n-k+1}(t) = 1 - \sum_{i=0}^{n-k} P_i(t) \quad (6)$$

where

$$A_{ij} = \frac{\prod_{h=0}^{i-1}(n-h)\lambda_h}{\prod_{h=0,h\neq j}^{i}(n-h)\lambda_h - (n-j)\lambda_j}.$$

3.1. Measures

1. Reliability of the system is computed by using Equations (4) - (5) and is given by,

$$R_n(t) = \sum_{i=0}^{n-k} P_i(t)$$
$$= \sum_{j=0}^{n-k} \frac{\prod_{h=0}^{n-k}(n-h)\lambda_h}{\prod_{h=0,h\neq j}^{n-k}(n-h)\lambda_h - (n-j)\lambda_j} e^{-(n-j)\lambda_j t}. \quad (7)$$

2. Mean time between failures ($MTBF_n$) is computed using Equation (7) and is given by,

$$MTBF_n = \int_0^{\infty} R_n(t)dt.$$

3. Let T_{conf} be the largest value of t for which probability of system being reliable for time t is greater than or equal to p_{conf}.

$$T_{conf} = max(t|R_n(t) \geq p_{conf}).$$

3.2. Special Case

Suppose λ_i are equal to λ, Equations (4) - (6) become

$$P_i(t) = {}^nC_i e^{-n\lambda t}(e^{\lambda t}+1)^i$$

$$P_{n-k+1}(t) = 1 - \sum_{i=0}^{n-k} P_i(t).$$

$$R_n(t) = \sum_{i=0}^{n-k}\sum_{j=0}^{i} {}^nC_i {}^iC_j e^{-(n-j)\lambda t}.$$

Further,

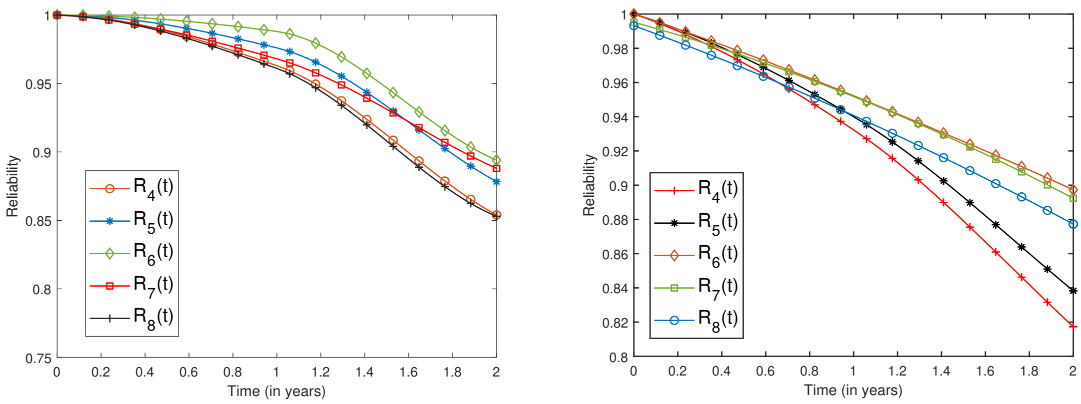
$$MTBF_n = \frac{1}{\lambda}\sum_{i=0}^{n-k}\sum_{j=0}^{i}\frac{{}^{n}C_i{}^{i}C_j}{n-j}.$$

4. Numerical Results

In this section, numerical illustration is presented for the proposed model. In order to plot the graphs, the parameters' values are considered depending on the stability conditions of the proposed model. The graphs have been plotted considering $\lambda = 0.1 year^{-1}$ for various values of ρ.

Fig. 6 shows the plot of the reliability function for various n taking $\lambda = 0.1 year^{-1}$ for the special case $\lambda_i = \lambda$. It can be observed from the figure that the value of $R_n(t)$ decreases with respect to the time as there will be more chances for the system to fail due to the circumstances. As the value of n increases, the reliability can be seen to increase. After $n = 6$, the reliability starts to decrease which provides the optimal value of n.

The similar behaviour can be observed from Fig. 5(b) for $\rho = 1.5$. Here, it can be observed that $R_n(t)$ is exhibiting the decreasing behaviour with respect to t. Also, $R_n(t)$ starts decreasing after $n = 6$ which shows that $n = 6$ is the optimal value. On the similar track, Fig. 5(a) demonstrates the similar behaviour of $R_n(t)$ with respect to t for $\rho = 0.8$. Here also the obtained optimal value of n is 6. Similarly, Table 1 gives the values of $MTBF_n$ for various n. The values of T_{conf} for various values of n and p_{conf} taking $\lambda = 0.1 year^{-1}$ is shown in Table 2.



(a) Reliability function for various n for $\rho = 0.8$ (b) Reliability function for various n for $\rho = 1.5$

Figure 5: *Reliability function for various n.*

n \ ρ	1	1.5	2	2.5
4	5.8333	4.7222	4.1667	3.8333
5	6.7193	5.0724	5.3225	4.4667
6	6.9698	5.4921	5.8811	4.5921
7	4.4421	3.7880	3.5552	3.2986
8	1.3850	1.4006	1.3080	1.7920
9	1.1108	1.3276	1.2102	1.3677

Table 1: *$MTBF_n$ vs n taking $\lambda = 0.1$ year^{-1}, for various ρ*

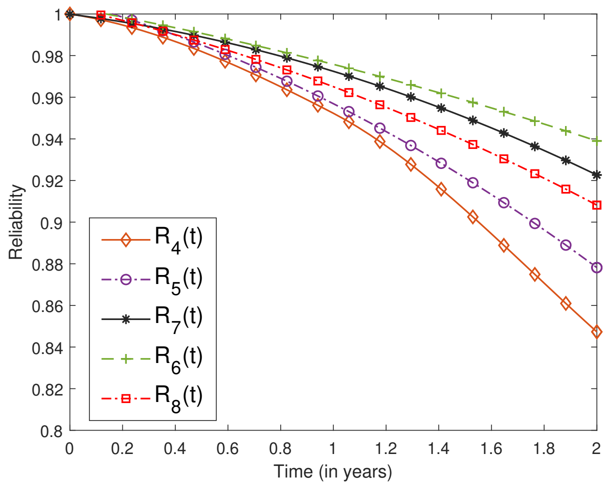


Figure 6: *Reliability function for various n for $\lambda_i = \lambda$.*

n \ p_{conf}	0.9	0.95	0.99	0.995	0.999
4	0.334	0.221	0.189	0.096	0.024
5	1.221	0.967	0.746	0.544	0.100
6	1.975	1.464	1.002	0.866	0.377
7	1.556	1.115	0.988	0.545	0.323
8	1.215	0.966	0.766	0.432	0.307
9	0.977	0.851	0.676	0.402	0.211

Table 2: *Values of T_{conf} (years) for $\rho = 1.5$ and $k = 3$ for various n and p_{conf}.*

5. Conclusions and Future Work

Based on the reliability graphs, Table 1 and Table 2, the reliability improves as increase the number of wires n until $n = 6$, after which the reliability decreases. Thus, a 6- core wire will be the best configuration for the chosen parameters. Vishnevsky et al. [6] obtained in their study of tether HAP system, by a different criterion, i.e., minimizing the weight of the cable. The obtained results are very close to the results obtained by [6].

In this study, all distributions in the proposed model are considered as independent exponential distributions. Hence, the underlying stochastic process is a Markov process. By extension of the above consideration, for future work, this research work can be considered with non Markov processes. As a future work, the non Markov model can be simulated to test whether Markov process is a good approximation for more complex non Markov processes.

Acknowledgment

The publication has been prepared with the support of DST-RSF research project no. 22-49-02023 (RSF) and research project no. 64800 (DST).

References

[1] Vishnevsky, V., Meshcheryakov, R. (2019). Experience of Developing a Multifunctional Tethered High-Altitude Unmanned Platform of Long-Term Operation. In: Ronzhin, A., Rigoll, G., Meshcheryakov, R. (eds) Interactive Collaborative Robotics. ICR 2019. Lecture

Notes in Computer Science, vol 11659. Springer, Cham. https://doi.org/10.1007/978-3-030-26118-4_23

[2] https://idstch.com/space/high-altitude-platforms-haps-for-communications-and-persistent-wide-area-real-time-intelligence-surveillance-and-reconnaissance-isr/

[3] J. Gavan, S. Tapuchi and D. Grace, "Concepts and main applications of high-altitude-platform radio relays," in URSI Radio Science Bulletin, vol. 2009, no. 330, pp. 20-31, Sept. 2009, doi: 10.23919/URSIRSB.2009.7909716.

[4] B. E. Y. Belmekki and M. -S. Alouini, "Unleashing the Potential of Networked Tethered Flying Platforms: Prospects, Challenges, and Applications," in IEEE Open Journal of Vehicular Technology, vol. 3, pp. 278-320, 2022, doi: 10.1109/OJVT.2022.3177946.

[5] E. Petritoli, F. Leccese and L. Ciani, "Reliability assessment of UAV systems," 2017 IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace), 2017, pp. 266-270, doi: 10.1109/MetroAeroSpace.2017.7999577.

[6] Vishnevsky, V., Tereschenko, B., Tumchenok, D., Shirvanyan, A. (2017). Optimal Method for Uplink Transfer of Power and the Design of High-Voltage Cable for Tethered High-Altitude Unmanned Telecommunication Platforms. In: Vishnevskiy, V., Samouylov, K., Kozyrev, D. (eds) Distributed Computer and Communication Networks. DCCN 2017. Communications in Computer and Information Science, vol 700. Springer, Cham. https://doi.org/10.1007/978-3-319-66836-9_20.

[7] Vishnevsky, V., Dharmaraja S., Rykov V., Kozyrev D. and Ivanova N. (2022). "Reliability modeling of a flight module of a tethered high-altitude telecommunication platform," 2022 6th International Scientific Conference on Information, Control, and Communication Technologies (ICCT), doi: 10.1109/ICCT56057.2022.9976764.

[8] Castaneda, L.B., Arunachalam V., Dharmaraja S. (2012). Introduction to Probability and Stochastic Processes with Applications. Wiley, New Jersey.

[9] Ivanova N. and Vishnevsky, V. (2021). "Application of k-out-of-n: G System and Machine Learning Techniques on Reliability Analysis of Tethered Unmanned Aerial Vehicle," In Information Technologies and Mathematical Modelling. Queueing Theory and Applications: 20th International Conference, ITMM 2021, Named after AF Terpugov, Tomsk, Russia, doi: https://doi.org/10.1007/978-3-031-09331.

[10] Ivanova N. (2020). "Modeling and Simulation of Reliability Function of ak-out-of-n: F System," In Distributed Computer and Communication Networks: Control, Computation, Communications: 23rd International Conference, DCCN 2020, Moscow, Russia, doi:https://doi.org/10.1007/978-3-030-66242.

TRUNCATED PRANAV DISTRIBUTION: PROPERTIES AND APPLICATIONS

Kamlesh Kumar Shukla

•

Department of Community Medicine,
Noida International Institute of Medical Sciences,
Noida International University, Gautam Budh Nagar, India
Email: kkshukla22@gmail.com

Abstract

In this paper, truncated Pranav distribution has been proposed. The behavior of truncated Pranav distribution has been presented graphically. Moment based measures including coefficient of variation, skewness, kurtosis, and Index of dispersion have been derived and presented graphically. Nature of survival and hazard rate functions are presented graphically. Maximum likelihood method has been used to estimate the parameter of proposed model. Simulation based study of proposed distribution has also been discussed. It has been applied on two data sets and its superiority has been compare and checked using goodness of fit (AIC and K. S. test) over other truncated distributions as well as one parameter distribution, such as exponential, Lindley, Pranav, Ishita, truncated Akash, truncated Lindley, and truncated Akash distribution. It was found good fit over above-mentioned distributions. It can be considered as good lifetime distribution especially for non-skewed data.

Keywords: Akash distribution, Lindley distribution, Moments, Right Truncated, Left Truncated

1. Introduction

In the recent past decades, lifetime modeling has been becoming popular in distribution theory, where many statisticians are involved in introducing new models. Some of the life time models are very popular and applied in biological, engineering and agricultural areas, such as Lindley distribution of Lindley [1], weighted Lindley distribution introduced by [2], Akash distribution suggested by Shanker[3], Ishita distribution proposed by Shanker and Shukla [4], Pranav distribution introduced by Shukla [5], are some among others and extension of above mentioned distribution has also been becoming popular in different areas of statistics.

Shukla [5] proposed Pranav distribution convex combination of exponential and gamma distributions which is defined by its pdf and cdf

$$f_1(y;\theta) = \frac{\theta^4}{\theta^4+6}(\theta + y^3)e^{-\theta y}; y > 0, \theta > 0 \tag{1}$$

$$F_2(y;\theta) = 1 - \left[1 + \frac{\theta y(\theta^2 y^2 + 3\theta y + 6)}{\theta^4 + 6}\right] e^{-\theta y}; y > 0, \theta > 0 \quad (2)$$

The rth moment about origin $\mu_r{}'$ of Pranav distribution given as

$$\mu_r{}' = \frac{r!\{\theta^4 + (r+1)(r+2)(r+3\}}{\theta^r(\theta^4+6)}; r = 1,2,3,\ldots \quad (3)$$

Shukla [5] has discussed in detailed about its mathematical and statistical properties, estimation of parameters and applications to model lifetime data from engineering and biomedical engineering. Truncated type of distribution is more effective for modeling lifetime data because its limits used as bound either upper or lower or both according to the given data. Truncated normal distribution is proposed by Johnson et al. [6]. It has wide application in economics and statistics. Many researchers have been proposed truncated type of distribution and applied in different areas of statistics, especially in censor data such as truncated Weibull distribution of Zange and Xie [7], truncated Lomax distribution of Aryuyuen and Bodhisuwan [8], truncated Pareto distribution of Janinetti and Ferraro [9], truncated Lindley distribution of Singh et al. [10]. Some researchers have proposed distribution such as Sindhu and Hussai [11] proposed a Mixture of two generalized inverted exponential distributions with censored sample model and applied on cancered data, and Shukla [12] proposed Inverse Ishita distribution and applied on using the data sets of bladder cancer patient and failure times of the air conditioning system. Recently truncated version of Akash distribution introduced by Shukla and Shanker [13] and truncated version of two parameter Pranav distribution has been proposed by Shukla [14] and its superiority has been shown in their paper over other truncated distribution. Some distributions and their introducer names have been listed in the table1.

Table1: *Pdf of some selected distribution and their introducer's name*

Distribution name	Pdf (probability density function)	Introducer' name
Truncated Lindley (TLD)	$f(x;\theta) = \dfrac{\theta^2(x+1)e^{-\theta x}}{(a\theta+1)e^{-\theta a} - (b\theta+1)e^{-\theta b} + (\theta+1)(e^{-\theta a} - e^{-\theta b})}$	Singh et al. [10]
Truncated Akash (TAD)	$f(x;\theta) = \dfrac{\theta^3(x^2+1)e^{-\theta x}}{a\theta(a\theta+2)e^{-\theta a} - b\theta(b\theta+2)e^{-\theta b} + (\theta^2+2)(e^{-\theta a} - e^{-\theta b})}$	Shukla & Shanker [13]
Akash	$f(x;\theta) = \dfrac{\theta^3}{\theta^2+2}(1+x^2)e^{-\theta}; x > 0, \theta > 0$	Shanker [3]
Ishita	$f(x;\theta) = \dfrac{\theta^3}{\theta^3+2}(\theta+x^2)e^{-\theta}; y > 0, \theta > 0$	Shanker& Shukla [4]
Lindley	$f(x;\theta) = \dfrac{\theta^2}{\theta+1}(1+x)e^{-\theta x}; x > 0, \theta > 0$	Lindley [1]
Exponential	$f(x;\theta) = \theta e^{-\theta x}; x > 0, \theta > 0$	

Truncated version of a continuous distribution can be defined as:

Definition1. Let X be a random variable distributed according to some pdf $g(x;\theta)$ and cdf $G(x;\theta)$, where θ is a parameter vector of X. Let X lies within the interval $[a,b]$, where $-\infty < a \le x \le b < \infty$, then X, conditional on $a \le x \le b$ is distributed as truncated distribution. The pdf of truncated distribution as reported by Singh et al (2014) defined by:

$$f(x;\theta) = g(x/a \le x \le b;\theta) = \frac{g(x;\theta)}{G(b;\theta) - G(a;\theta)} \quad (4)$$

where $f(x;\theta) = g(x;\theta)$ for all $a \le x \le b$ and $f(x;\theta) = 0$ elsewhere. Note that $f(x;\theta)$ in fact is a pdf of X on interval $[a,b]$. We have

$$f(x;\theta)=\int_a^b f(x;\theta)\,dx=\frac{1}{G(b;\theta)-G(a;\theta)}\int_a^b g(x;\theta)\,dx$$

$$=\frac{1}{G(b;\theta)-G(a;\theta)}G(b;\theta)-G(a;\theta)=1 \quad (5)$$

The cdf of truncated distribution is given by

$$F(x;\theta)=\int_a^x f(x;\theta)dx=\frac{G(x;\theta)-G(a;\theta)}{G(b;\theta)-G(a;\theta)} \quad (6)$$

The main objective of this paper is to propose new truncated distribution using Pranav distribution which is called as truncated Pranav distribution, and to know the behavior and properties of proposed distribution over other truncated as well as parent distributions. Present paper has been divided into eight sections. Introduction about the paper is described in the first section. In the second section, truncated Pranav distribution has been derived. Behavior of hazard rate has been presented in third section Statistical properties including its moment have been discussed in the fourth section. Estimation of parameters of the proposed distribution has been discussed in the fifth section. Simulation study of proposed distribution has been discussed in the sixth section. Its application and comparative study with one parameter lifetime distribution have been illustrated in the section seven. Finally, the conclusion of the paper has been given in the eighth section.

2. Truncated Pranav Distribution

In this section, pdf and cdf of new truncated distribution is proposed and named Truncated Pranav distribution, using (5) & (6) of definition1 and from (1) & (2) , which is defined as :

Definition 2: Let X be random variable which is distributed as Truncated Pranav distribution (TPD) with scale parameter θ and location parameters a &b, and denoted by TPD(a,b,θ). The pdf and cdf of X are derived respectively as:

$$f(x;\theta)=g(x/a\le x\le b;\theta)=\frac{g(x;\theta)}{G(b;\theta)-G(a;\theta)}$$

$$F(x;\theta)=\frac{G(x;\theta)-G(a;\theta)}{G(b;\theta)-G(a;\theta)}$$

Where $g(x;\theta)=\frac{\theta^4}{\theta^4+6}(\theta+x^3)e^{-\theta x}$

$$G(b;\theta)=1-\left[1+\frac{\theta b(\theta^2b^2+3\theta b+6)}{\theta^4+6}\right]e^{-\theta b}$$

$$G(a;\theta)=1-\left[1+\frac{\theta a(\theta^2a^2+3\theta a+6)}{\theta^4+6}\right]e^{-\theta a}$$

$$f(x;\theta)=\frac{\theta^4(x^3+\theta)e^{-\theta x}}{(a^3\theta^3+3a^2\theta^2+\theta^4+6a\theta+6)e^{-\theta a}-(b^3\theta^3+3b^2\theta^2+\theta^4+6b\theta+6)e^{-\theta b}} \quad (7)$$

$$F(x;\theta)=\frac{(a^3\theta^3+3a^2\theta^2+\theta^4+6a\theta+6)e^{-\theta a}-(x^3\theta^3+3x^2\theta^2+\theta^4+6x\theta+6)e^{-\theta x}}{(a^3\theta^3+3a^2\theta^2+\theta^4+6a\theta+6)e^{-\theta a}-(b^3\theta^3+3b^2\theta^2+\theta^4+6b\theta+6)e^{-\theta b}} \quad (8)$$

where $-\infty<a\le x\le b<\infty$, and $\theta>0$

Three cases can be considered of proposed doubly truncated distribution as:

a) When $a=0$and $b=\infty$, it reduced to parent model (Pranav distribution).
b) When $a=0$, it is known as right (upper) truncated distribution of the parental model (Right truncated Pranav distribution)
c) When $b=\infty$, it is known as left (lower) truncated distribution of the parent model (left truncated Pranav distribution.

Properties of TPD are explained as follows:

i. TPD has three parameters, where two parameters a and b were considered lowest and largest value from the data.

ii. Parameter θ is considered as scale parameter, for the fixed value of a and b, TPD is observed decreasing and increasing as increased value of θ for $\theta < 1$ and $\theta > 1$ respectively.

Performance of pdf of TPD for varying values of parameters has been illustrated in the figure 1.

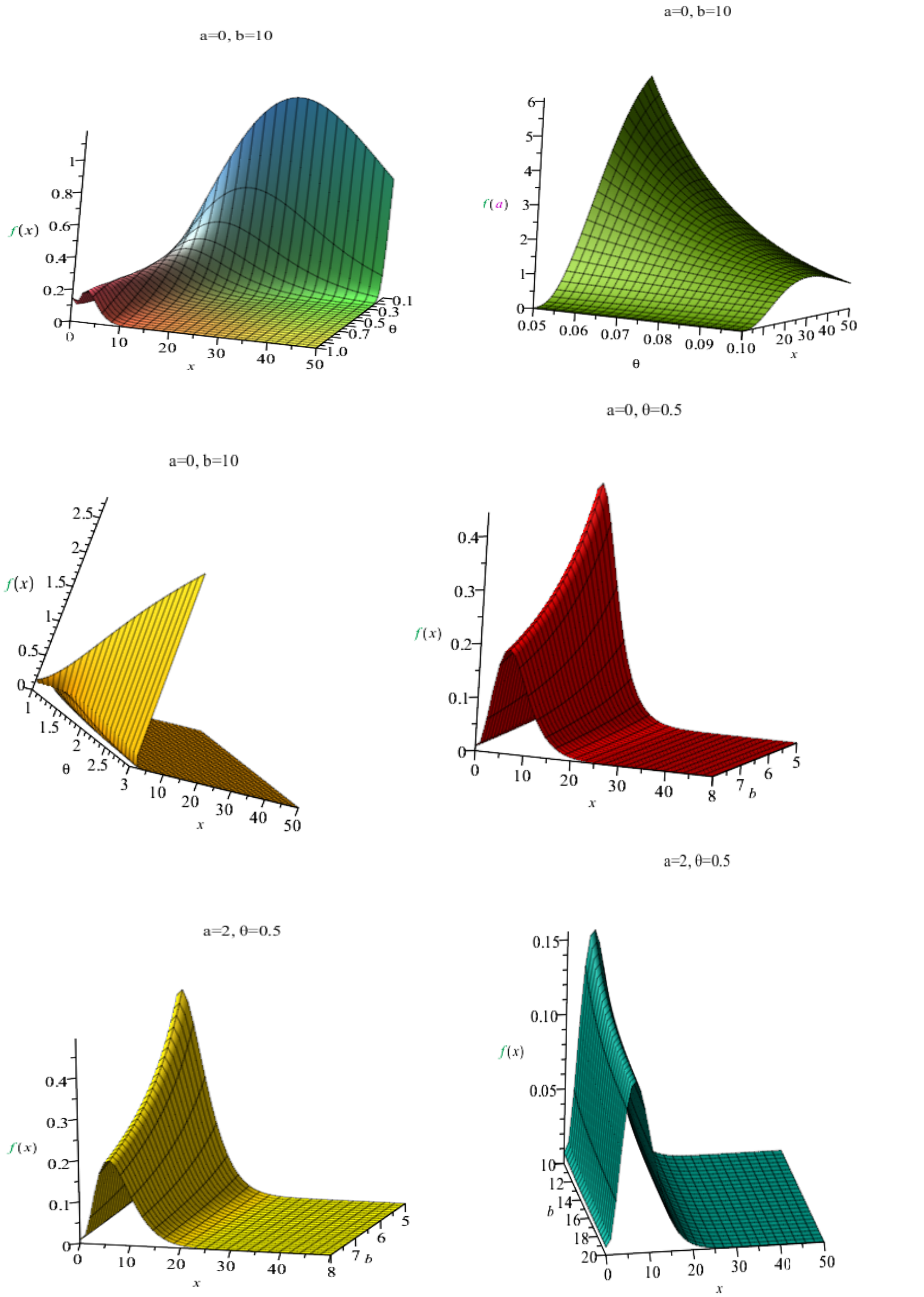


Figure 1: *pdf plots of TPD for varying values of parameters*

3. Survival and Hazard function

The survival function $S(x)$ and the hazard function $h(x)$ of TPD are defined as

$$S(x) = 1 - F(x) = \frac{(x^3\theta^3 + 3x^2\theta^2 + \theta^4 + 6x\theta + 6)e^{-\theta x} - (b^3\theta^3 + 3b^2\theta^2 + \theta^4 + 6b\theta + 6)e^{-\theta b}}{(a^3\theta^3 + 3a^2\theta^2 + \theta^4 + 6a\theta + 6)e^{-\theta a} - (b^3\theta^3 + 3b^2\theta^2 + \theta^4 + 6b\theta + 6)e^{-\theta b}}$$

$$h(x) = \frac{f(x)}{S(x)} = \frac{\theta^4(x^3 + \theta)e^{-\theta x}}{(x^3\theta^3 + 3x^2\theta^2 + \theta^4 + 6x\theta + 6)e^{-\theta x} - (b^3\theta^3 + 3b^2\theta^2 + \theta^4 + 6b\theta + 6)e^{-\theta b}}$$

It is obvious that $h(x)$ is independent from parameter a. Behavior of Survival and hazard function of TPD for varying values of parameter are presented in figures 2&3. It was observed from the figure 3 that value of hazard rate is increasing as increased value of parameter of θ when value of rest of the parameters (a, b)fixed.

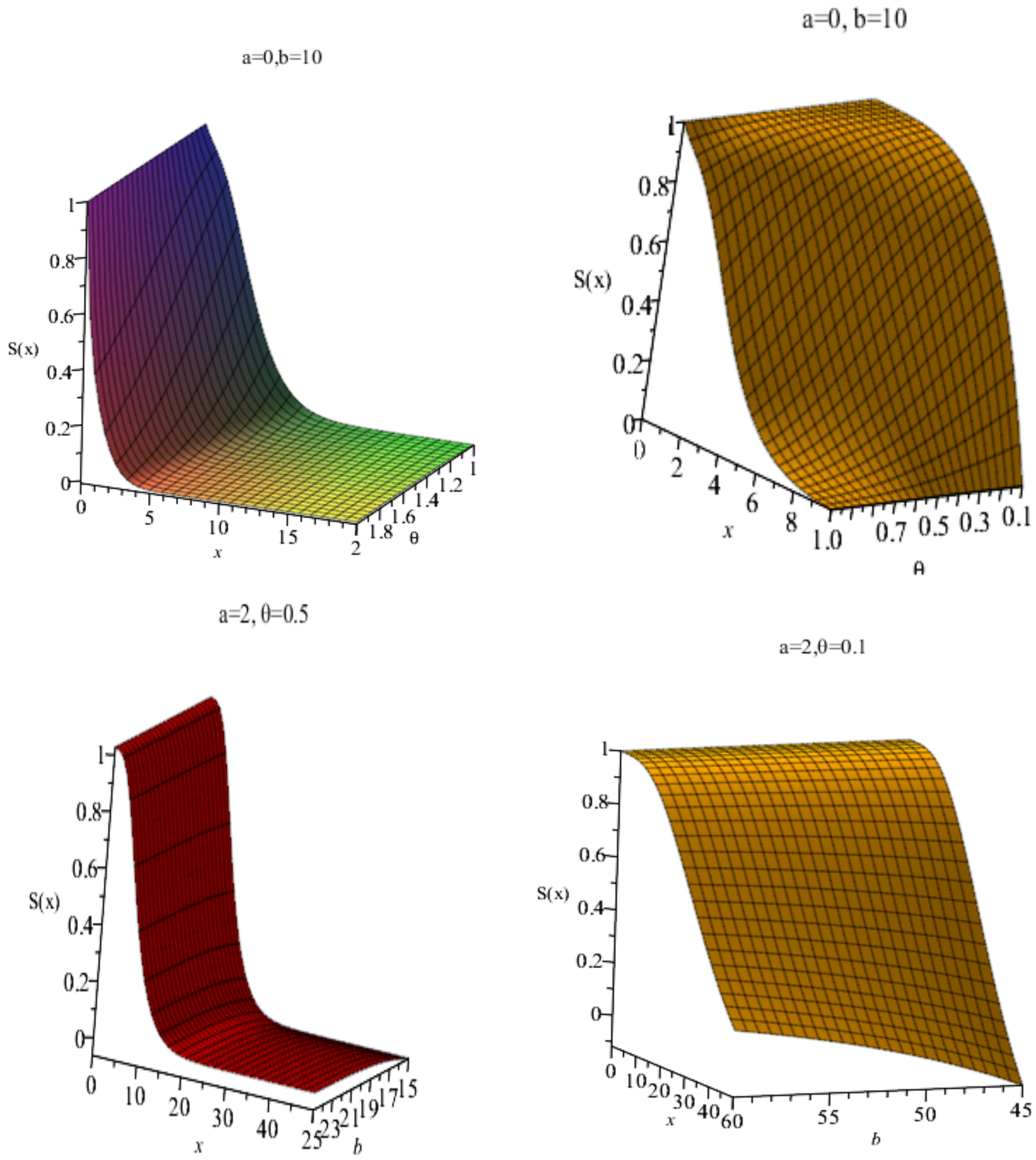


Figure 2: *S(x) plots of TPD for varying values of parameter*

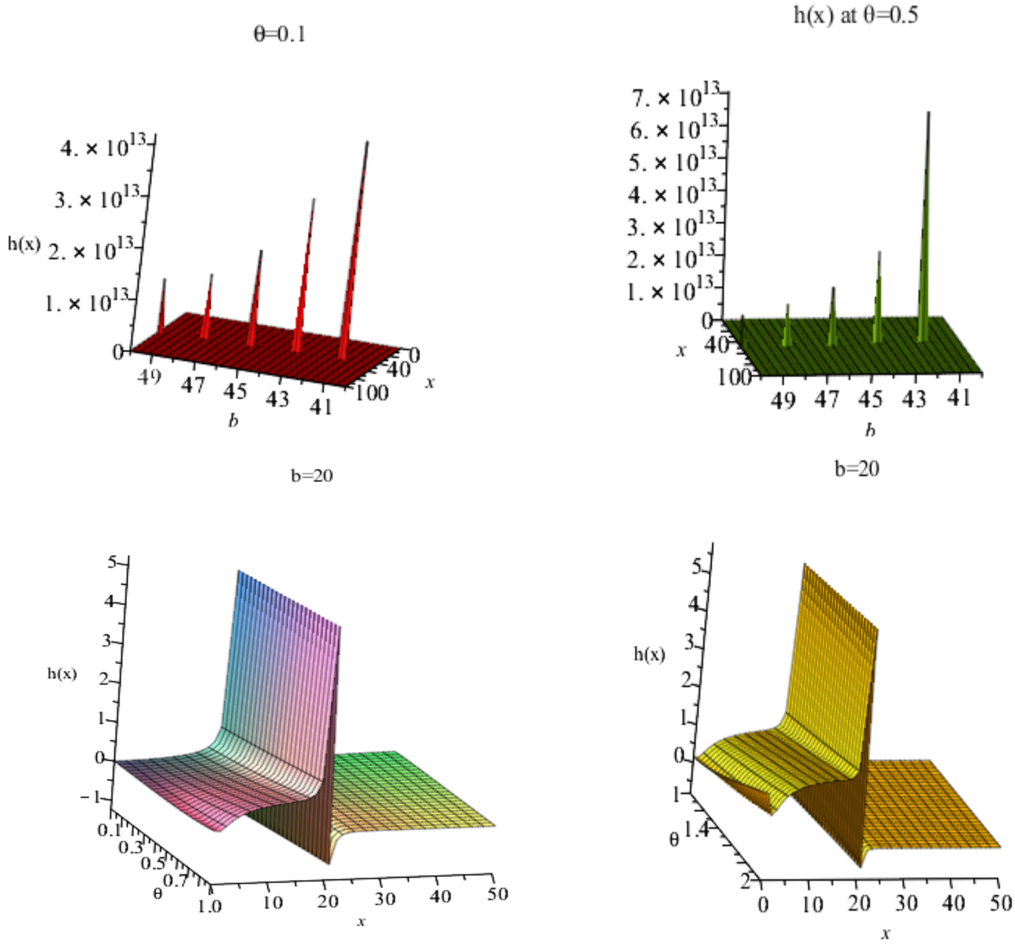


Figure3: *h(x) plots of TPD for varying values of parameter*

4. Moments and mathematical properties

Moments of a distribution are used to study the most important characteristics of the distribution including mean, variance, skewness, kurtosis, etc. The rth moment of origin $\mu_r{'}$ of TPD can be expressed in explicit expression in terms of incomplete gamma functions.

Theorem1: Suppose $X_i(a > 0, b \sim \infty)$ follows left (lower) truncated TPD (θ, a, ∞). Then the rth moment about origin $\mu_r{'}$ of TPD is

$$\mu_r{'} = \frac{\theta^4 \gamma(r+1,\theta a) + \gamma(r+4,\theta a)}{\theta^r[(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}-(\theta^4+6)]}; r = 1,2,3, \ldots$$

Proof: Considering $K = \{(a^3\theta^3 + 3a^2\theta^2 + 6a\theta + \theta^4 + 6)e^{-\theta a} - (\theta^4 + 6)\}$ in (7), we have

$$\mu_r{'} = \frac{\theta^4}{K}\int_a^\infty x^r(\theta + x^3)\, e^{-\theta x} dx$$

$$= \frac{\theta^4}{K}\left[\int_a^\infty \theta e^{-\theta x} x^r dx + \int_a^\infty e^{-\theta x} x^{r+3} dx\right]$$

Taking $u = \theta x, x = \frac{u}{\theta}$

$$= \frac{\theta^4}{K}\left[\frac{\theta}{\theta^{r+1}}\left\{\int_0^{\theta a} e^{-u} x^r du\right\} + \frac{1}{\theta^{r+4}}\left\{\int_0^{\theta a} e^{-u} x^{r+3} du\right\}\right]$$

Where $\gamma(\alpha, z) = \int_z^\infty e^{-x} x^{\alpha-1} dx$, $\alpha > 0, x > 0$ is the upper incomplete gamma function

$$= \frac{\theta^4}{K}\left[\frac{\gamma(r+1,\theta a)}{\theta^r} + \frac{\gamma(r+4,\theta a)}{\theta^{r+4}}\right]$$

$$=\frac{1}{K}\left[\frac{\theta^4\gamma(r+1,\theta a)+\gamma(r+4,\theta a)}{\theta^r}\right]$$

$$\mu_r{}' = \frac{\theta^4\gamma(r+1,\theta a)+\gamma(r+4,\theta a)}{\theta^r[(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}-(\theta^4+6)]}$$

Now taking $r = 1,2$, mean and variance can be obtained as

$$\mu_r{}' = \frac{\theta^4\gamma(2,\theta a)+\gamma(5,\theta a)}{\theta[(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}-(\theta^4+6)]}$$

$$\mu_r{}' = \frac{\theta^4\gamma(3,\theta a)+\gamma(6,\theta a)}{\theta^2[(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}-(\theta^4+6)]}$$

Variance $\mu_2 = \mu_2' - (\mu_1')^2$

Theorem2. Suppose $X_i(b > 0, a\sim 0)$ follows upper (right) truncated TPD$(\theta, 0, b)$. Then the rth moment about origin$\mu_r{}'$ of TPD is

$$\mu_r{}' = \frac{\theta^4\{\gamma(r+1,\theta b)\}+\gamma(r+4,\theta b)}{\theta^r[(\theta^4+6)-(b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}]}; r = 1,2,3,\ldots$$

Proof: Considering $K = \{(\theta^4+6)-(b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\}$ in (7), we have

$$\mu_r{}' = \frac{\theta^4}{K}\int_a^b x^r(\theta+x^3)\,e^{-\theta x}dx$$

$$= \frac{\theta^4}{K}\left[\int_0^b \theta e^{-\theta x}x^r dx + \int_0^b e^{-\theta x}x^{r+3}dx\right]$$

Taking $u = \theta x, x = \frac{u}{\theta}$

$$= \frac{\theta^4}{K}\left[\frac{\theta}{\theta^{r+1}}\left\{\int_0^{\theta b} e^{-u}x^r du\right\} + \frac{1}{\theta^{r+4}}\left\{\int_0^{\theta b} e^{-u}u^{r+3}du\right\}\right]$$

Where $\gamma(\alpha,z) = \int_0^z e^{-x}x^{\alpha-1}dx\,, \alpha > 0, x > 0$ is the lower incomplete gamma function

$$=\frac{\theta^4}{K}\left[\frac{\gamma(r+1,\theta b)}{\theta^r} + \frac{\gamma(r+4,\theta b)}{\theta^{r+4}}\right]$$

$$=\frac{1}{K}\left[\frac{\theta^4\{\gamma(r+1,\theta b)\}+\gamma(r+4,\theta b)}{\theta^r}\right]$$

$$=\mu_r{}' = \frac{\theta^4\{\gamma(r+1,\theta b)\}+\gamma(r+4,\theta b)}{\theta^r[(\theta^4+6)-(b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}]}$$

Now taking$r = 1,2$, mean and variance can be obtained as

$$\mu_1' = \frac{\theta^4\gamma(2,\theta b)+\gamma(5,\theta b)}{\theta[(\theta^4+6)-(b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}]}$$

$$\mu_2' = \frac{\theta^4\gamma(3,\theta b)+\gamma(6,\theta b)}{\theta^2[(\theta^4+6)-(b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}]}$$

Variance $\mu_2 = \mu_2' - (\mu_1')^2$

Theorem3: Suppose X follows doubly TPD (θ, a, b). Then the rth moment about origin $\mu_r{}'$ of TPD is

$$\mu_r{}' = \frac{\theta^4\{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)\}+\{\gamma(r+4,\theta b)-\gamma(r+4,\theta a)\}}{\theta^r\begin{pmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}- \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{pmatrix}}; r = 1,2,3,\ldots$$

Proof: Considering $K = \begin{Bmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a} - \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{Bmatrix}$

in (7), we have

$$\mu_r{}' = \frac{\theta^4}{K}\int_a^b x^r(\theta+x^3)\,e^{-\theta x}dx$$

$$= \frac{\theta^4}{K}\left[\int_a^b \theta e^{-\theta x}x^r dx + \int_a^b e^{-\theta x}x^{r+3}dx\right]$$

Taking $u = \theta x, x = \frac{u}{\theta}$

$$= \frac{\theta^4}{K}\begin{bmatrix}\frac{\theta}{\theta^{r+1}}\left\{\int_0^{\theta b} e^{-u}x^r du - \int_0^{\theta a} e^{-u}x^r du\right\} + \\ \frac{1}{\theta^{r+4}}\left\{\int_0^{\theta b} e^{-u}u^{r+3} du - \int_0^{\theta a} e^{-u}x^{r+3} du\right\}\end{bmatrix}$$

Where $\gamma(\alpha, z) = \int_0^z e^{-x}x^{\alpha-1}dx, \alpha > 0, x > 0$ is the lower incomplete gamma function

$$=\frac{\theta^4}{K}\left[\frac{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)}{\theta^r} + \frac{\gamma(r+4,\theta b)-\gamma(r+4,\theta a)}{\theta^{r+4}}\right]$$

$$=\frac{1}{K}\left[\frac{\theta^4\{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)\}+\{\gamma(r+4,\theta b)-\gamma(r+4,\theta a)\}}{\theta^r}\right]$$

$$=\frac{\theta^4\{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)\}+\{\gamma(r+4,\theta b)-\gamma(r+4,\theta a)\}}{\theta^r\begin{pmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}- \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{pmatrix}} \tag{9}$$

Now taking$r = 1,2$, mean and variance can be obtained as

$$\mu_1' = \frac{\theta^4\{\gamma(2,\theta b)-\gamma(2,\theta a)\}+\{\gamma(5,\theta b)-\gamma(5,\theta a)\}}{\theta\begin{pmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}- \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{pmatrix}}$$

$$\mu_2' = \frac{\theta^4\{\gamma(3,\theta b)-\gamma(3,\theta a)\}+\{\gamma(6,\theta b)-\gamma(6,\theta a)\}}{\theta^2\begin{pmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}- \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{pmatrix}}$$

Variance $\mu_2 = \mu_2' - (\mu_1')^2$

Similarly rest two moments of origin as well as coefficient of variation, coefficient of skewness, coefficient of kurtosis and Index of dispersion can be obtained, substituting $r = 3,4$ in the equation (9), which are as follows:

$$\mu_3' = \frac{\theta^4\{\gamma(4,\theta b)-\gamma(4,\theta a)\}+\{\gamma(7,\theta b)-\gamma(7,\theta a)\}}{\theta^3\begin{pmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}- \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{pmatrix}}$$

$$\mu_4' = \frac{\theta^4\{\gamma(5,\theta b)-\gamma(5,\theta a)\}+\{\gamma(8,\theta b)-\gamma(8,\theta a)\}}{\theta^4\begin{pmatrix}(a^3\theta^3+3a^2\theta^2+6a\theta+\theta^4+6)e^{-\theta a}- \\ (b^3\theta^3+3b^2\theta^2+6b\theta+\theta^4+6)e^{-\theta b}\end{pmatrix}}$$

Coefficient of Variation$=\frac{\left(\mu_2'-(\mu_1')^2\right)^{1/2}}{\mu_1'}$, Coefficient of Skweness$=\frac{\left(\mu_3'+3\mu_2'\mu_1'-(\mu_1')^2\right)}{(\mu_2'-(\mu_1')^2)^{3/2}}$, Coefficient of Kurtosis$=\frac{\left(\mu_4'-4\mu_3'\mu_1'+6\mu_2'(\mu_1')^2-3(\mu_1')^4\right)}{(\mu_2'-(\mu_1')^2)^2}$,

Index of dispersion$=\frac{\left(\mu_2'-(\mu_1')^2\right)}{\mu_1'}$, graph of above measures is presented in figures 4 to 9. From the figure 4 &5, it was observed that mean and variance are decreasing with increased value of θ and slightly increasing with *b* while parameter *a* is kept constant.

Coefficient of variation of TPD was found deceasing with increased value of θ and *b*.

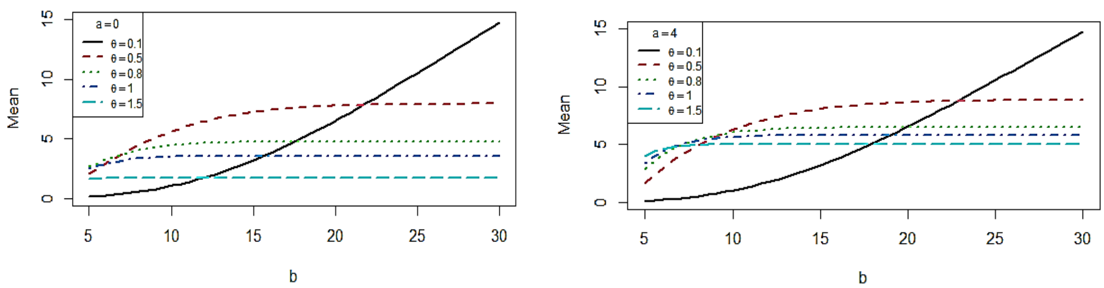


Figure 4: *Mean of TPD (doubly truncated) for varying value of parameter*

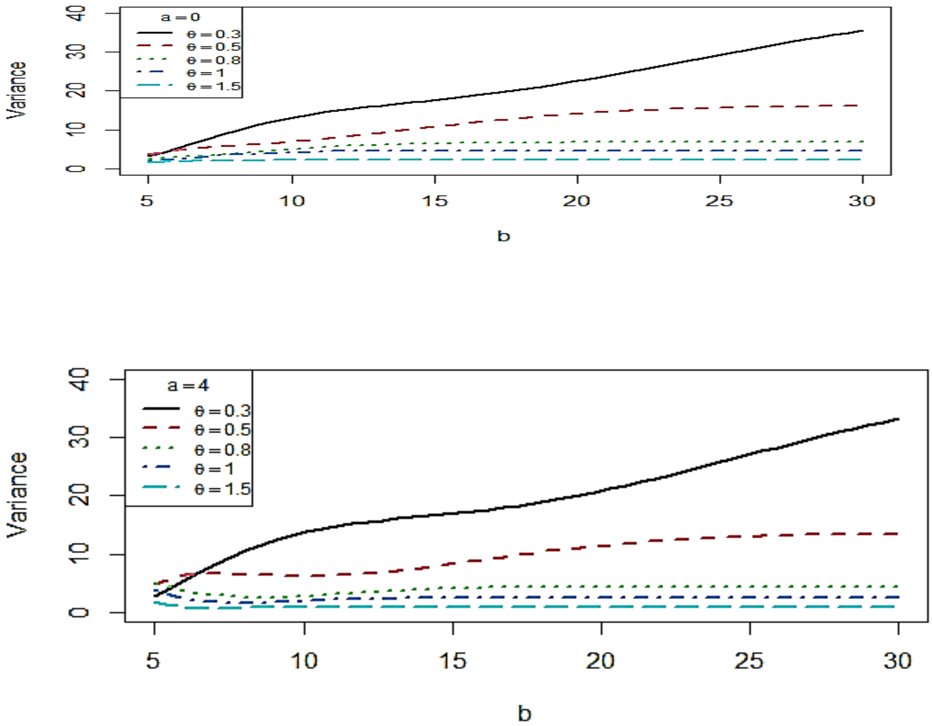


Figure5: *Variance of TPD for varying value of parameter*

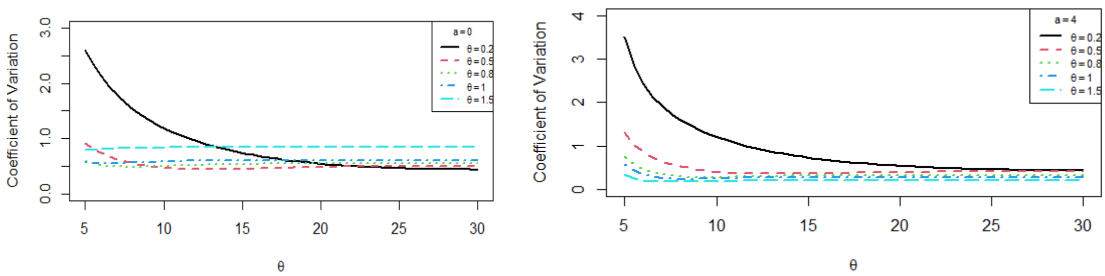


Figure 6: *Coefficient of variation of TPD (Doubly truncated) for varying value of parameter*

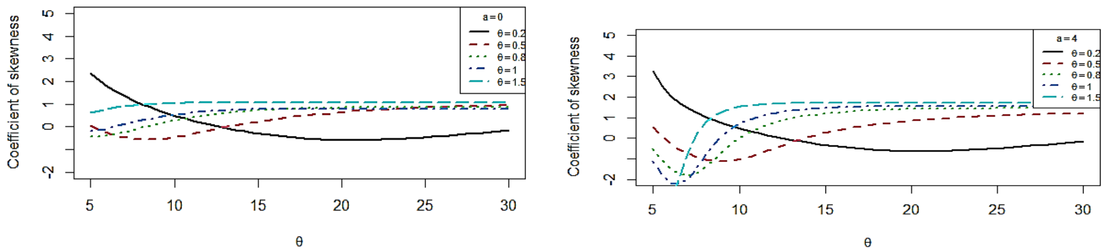


Figure 7: *Coefficient of skewness of TPD (Doubly truncated) for varying value of parameter*

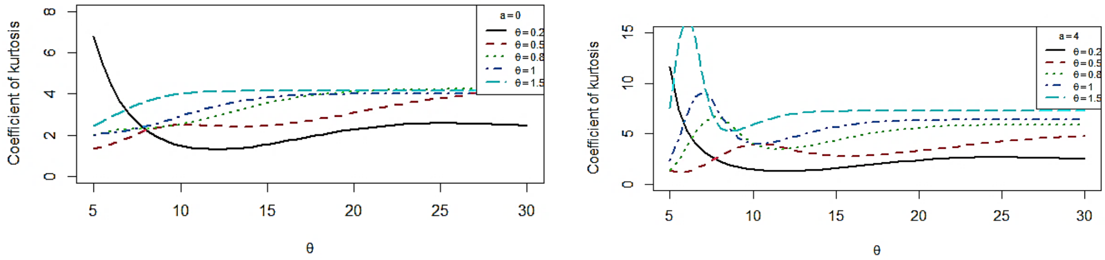


Figure 8: *Coefficient of kurtosis of TPD (Doubly truncated) for varying value of parameter*

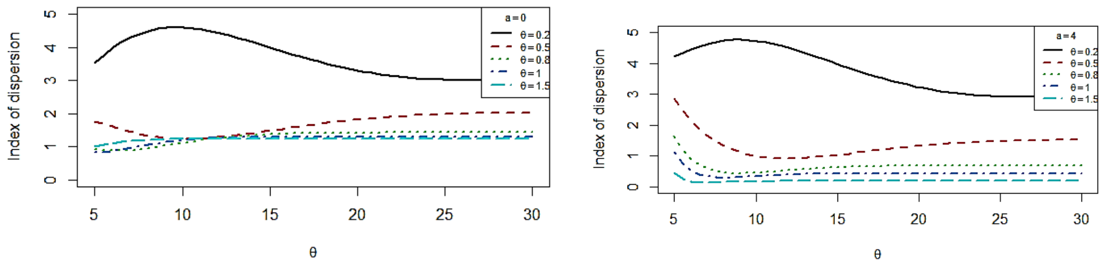


Figure9: *Index of dispersion of TPD (doubly truncated) for varying value of parameter*

5. Maximum likelihood Method Estimation

Let $(x_1, x_2, x_3, \dots, x_n)$ be a random sample of size n from (7).
The likelihood function, L of TPD is given by

$$L = \left(\frac{\theta^4}{(a^3\theta^3+3a^2\theta^2+\theta^4+6a\theta+6)e^{-\theta a}-(b^3\theta^3+3b^2\theta^2+\theta^4+6b\theta+6)e^{-\theta b}}\right)^n \prod_{i=1}^{n}(\theta + x_i^{\,3})\, e^{-n\theta\bar{x}}$$

The log likelihood function is thus obtained as

$$\ln L = n\ln\left(\frac{\theta^4}{(a^3\theta^3+3a^2\theta^2+\theta^4+6a\theta+6)e^{-\theta a}-(b^3\theta^3+3b^2\theta^2+\theta^4+6b\theta+6)e^{-\theta b}}\right) + \sum_{i=1}^{n}\ln(\theta + x_i^{\,3}) - n\theta\bar{x}$$

Taking $\hat{a} = min(x_1, x_2, x_3, \dots, x_n)$, $\hat{b} = max(x_1, x_2, x_3, \dots, x_n)$, the maximum likelihood estimate $\hat{\theta}$of parameter θ is the solution of the log-likelihood equation $\frac{\partial \log L}{\partial\theta} = 0$.
It is obvious that $\frac{\partial \log L}{\partial\theta} = 0$ will not be in closed form and hence some numerical optimization technique can be used e the equation forθ. In this paper the nonlinear method available in R software has been used to find the MLE of the parameterθ.

6. Simulation study

In this section, simulation study has been carried out using R-software. Acceptance and rejection method is used to generate random number, where sample size, $n = 20,40,60,80,100$, value of $\theta = 0.1,0.5,1.0,1.5$& $(a = 10, b = 100)$ have been used for calculating Bias error and MSE (Mean square error) of parameter θ, which is presented in table2.

General algorithm for generating the data which was given by Robert and Casella [15] are as follows: The constraints we impose on this candidate density f_Y are that:

(i) Y be simulate-able (the data simulation from Y be actually possible).

(ii) There is a constant c with $\frac{f_X(x)}{f_Y(x)} \leq c$ for all $x \in S_x = \{x: f_X(x) > 0\}$

(iii) $f_X(x)$ and $f_Y(x)$ have compatible supports (i.e., $S_X \subseteq S_Y$).

In this case, X can be simulated as follows by Accept-Reject method. First, we generate y from $Y \sim f_Y$ and, independently, we generate u from $U \sim U(0,1)$.

If $u \leq \frac{f_X(y)}{c f_y(y)}$ then we set $x = y$. If the inequality is not satisfied, we then discard/reject y and u and start again.

Table2: *Average Bias (AB) and Average MSE (AM) of the simulated MLEs of* θ *at fixed value of* $a = 10, b = 100$

Sample	θ	A B	AM	AB	AM
20	0.1	0.0606620	0.07359774	0.04834145	0.04673793
	0.5	0.0509281	0.05187350	0.05092813	0.00518735
	1.0	0.02592813	0.01344536	0.02592813	0.01344536
	1.5	0.00092813	0.000017228	0.00092813	0.00001722
40	0.1	0.02391497	0.02287705	0.02420517	0.02343561
	0.5	0.02546406	0.02593675	0.02546406	0.0259367
	1.0	0.01296406	0.006722681	0.01296406	0.00672268
	1.5	0.00046406	0.0000086143	0.00046406	0.00008614
60	0.1	0.01535922	0.01415435	0.01619160	0.01573008
	0.5	0.016976044	0.01729117	0.01697604	0.00172911
	1.0	0.008642711	0.004481787	0.00864271	0.00044817
	1.5	0.000309377	0.0000057428	0.00030937	0.00005742
80	0.1	0.010465770	0.008762596	0.01131883	0.01024928
	0.5	0.012732033	0.01296837	0.01273203	0.01296837
	1.0	0.006482033	0.003361341	0.00648203	0.00336134
	1.5	0.00023203	0.000004.307	0.00023204	0.00000430
100	0.1	0.00882808	0.007793502	0.00890854	0.07936211
	0.5	0.010185626	0.01037470	0.01018562	0.01037470
	1.0	0.005185626	0.002689072	0.00518562	0.00268907
	1.5	0.000185626	0.0000034457	0.00018562	0.00000344

From the above table, it is observed that Bias and Mean square error are decreasing as increased value of sample size.

7. Applications on lifetime data

In this section, TPD has been applied to two datasets using maximum likelihood estimates. Parameter θ estimated whereas another parameters a, and b are considered as lowest and highest values of data, i.e.

$\hat{a} = min(x_1, x_2, x_3, \ldots, x_n)$& $\hat{b} = max(x_1, x_2, x_3, \ldots, x_n)$. Goodness of fit has been decided using Akaike information criteria (AIC), Bayesian Information criteria (BIC) and Kolmogorov Simonov test (KS) values respectively, which are calculated for each distribution and compared with p-value. As we know that best goodness of fit of the distribution can be decided based on minimum value of KS, AIC and BIC.

Data Set 1: The data is given by Birnbaum and Saunders [16] on the fatigue life of 6061 – T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second.

Data Set 2: This data set is the strength data of glass of the aircraft window reported by Fuller et al. [17]:

18.83 20.8 21.657 23.03 23.23 24.05 24.321 25.5 25.52 25.8 26.69 26.77 26.78 27.05 27.67 29.9 31.11 33.2 33.73 33.76 33.89 34.76 35.75 35.91 36.98 37.08 37.09 39.58 44.045 45.29 45.381

Table 3: *MLE's, Standard Errors, - 2ln L, AIC, K-S and p-values of the fitted distributions for data set-1*

Distributions	ML Estimates	Standard Errors	$-2\ ln\,L$	AIC	BIC	K-S	p-value
TPD	$\hat{\boldsymbol{\theta}} = \mathbf{0.055278}$	**0.00330**	**927.37**	**929.37**	**928.76**	**0.136**	**0.048**
TAD	$\hat{\theta} = 0.03917$	0.00303	939.13	941.13	942.05	0.153	**0.017**
TLD	$\hat{\theta} = 0.02199$	0.00273	958.88	960.88	962.31	0.186	**0.001**
Pranav	$\hat{\theta} = 0.04387$	0.00253	950.97	952.97	954.40	0.194	**0.001**
Ishita	$\hat{\theta} = 0.04390$	0.002533	950.92	9952.92	954.35	0.194	**0.001**
Lindley	$\hat{\theta} = 0.02886$	0.002038	983.10	985.10	986.54	0.252	**0.000**
Exponential	$\hat{\boldsymbol{\theta}} = \mathbf{0.01463}$	**0.001457**	**1044.87**	**1046.87**	**1048.30**	**0.336**	**0.000**

Table 4: *MLE's, Standard Errors, - 2ln L, AIC, K-S and p-values of the fitted distributions for data set-2*

Distributions	ML Estimates	Standard Errors	$-2\ ln\,L$	AIC	BIC	K-S	p-value
TPD	**0.12067**	**0.02455**	**201.80**	**203.80**	**203.19**	**0.107**	**0.829**
TAD	$\hat{\theta} = 0.08776$	0.024241	201.96	203.96	205.58	0.112	**0.786**
TLD	$\hat{\theta} = 0.05392$	0.023917	202.18	204.18	205.61	0.117	**0.738**
Pranav	$\hat{\theta} = 0.09706$	0.01004	240.68	242.68	242.67	0.298	**0.005**
Ishita	$\hat{\theta} = 0.097328$	0.01008	240.48	242.48	243.48	0.297	**0.006**
Lindley	$\hat{\theta} = 0.06299$	0.00800	253.98	255.98	256.98	0.365	**0.000**
Exponential	$\hat{\boldsymbol{\theta}} = \mathbf{0.032452}$	**0.00582**	**274.52**	**276.52**	**277.52**	**0.458**	**0.000**

From the above tables 3&4 , it was observed that TPD gives good fit over other selected distributions. Fitted plots of the considered distributions are presented in figure 10 and 11 respectively.

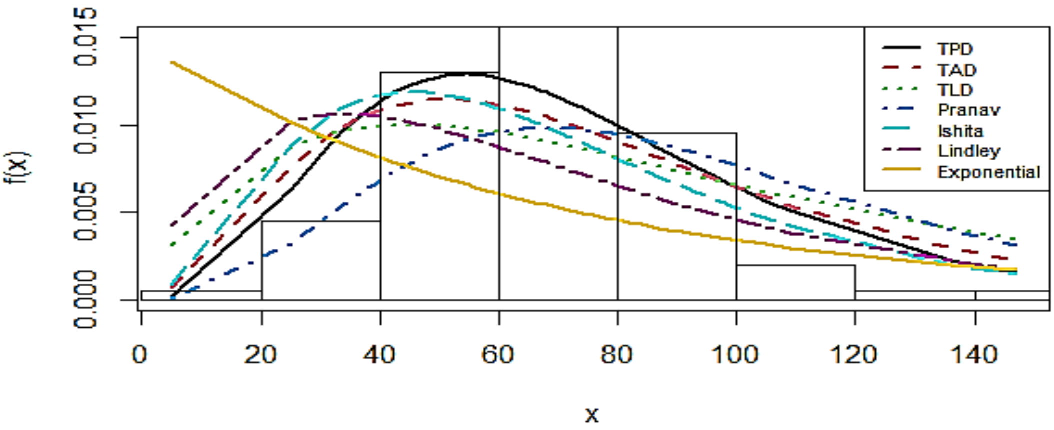


Figure10: *Fitted plots of distributions for the dataset 1*

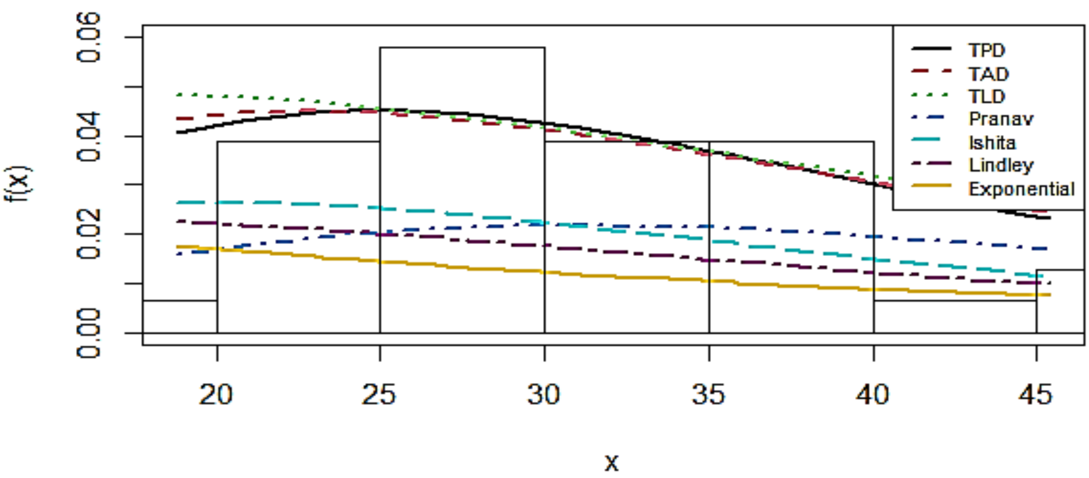


Figure11: *Fitted plot of distributions for dataset 2*

7. Conclusions

In this paper, truncated Pranav distribution (TPD) has been proposed. Its statistical and mathematical properties have been discussed. Maximum likelihood method has been used for estimation of its parameter. Simulation study has also been conducted to know behavior of proposed distribution. Goodness of fit of TPD has been discussed with two lifetime datasets and superiority has been checked with truncated Akash, truncated Lindley, Pranav, Ishita, Lindley, and exponential distributions. It has been observed that TPD gives good fit on both the data sets. In the first data set, value of AIC 929.37 and KS value- 0.136, p-value =0.048 <0.01, and for the second data set, value of AIC-203.80 and KS value- 0.107 (p-value =0.829 >0.05) were observed which were compared over two parameter TAD (truncated Akash Distribution), TLD (truncated Lindley Distribution) and one parameter Pranav, Ishita, Lindley and exponential distribution. Therefore, it may be considered good distribution for the lifetime data specially on fixed values (lower limit, upper limit).

Acknowledgement

Author is thankful to the Editor in chief and anonymous reviewers for the given fruitful comments to improve the quality and presentation of the paper.

References

[1] Lindley, D.V. (1958), Fiducial distributions and Bayes' Theorem, Journal of the Royal Statistical Society, Series B, 20:102 – 107.

[2] Ghitany ME, Atieh B, Nadarajah S (2008)., Lindley distribution and its applications. Mathematics Computing and Simulation, 78: 493– 506.

[3] Shanker, R. (2015), Akash distribution and its application, International Journal of Probability and Statistics, 4 (3):65–75.

[4] Shanker, R. and Shukla, K.K. (2017): Ishita distribution and its applications, Biometrics & Biostatistics International Journal, 5(2), 1-9.

[5] Shukla, K. K., Pranav distribution with Properties and Applications, Biometrics and Biostatistics International Journal,2018; 7(3):244-254.

[6] Johnson, N. L., Kotz,S. , Balakrishnan, N., Continuous Univariate Distributions, Wiley, New York (1994).

[7] Zhange, Z., and Xie, M. (2010), On the upper truncated Weibull distribution and its reliability implications, Reliability Engineering & System Safety, 96:194-200.

[8] Aryuyuen, S., Bodhisuwan, W. (2018), The truncated power Lomax distribution: Properties and applications, Walailak Journal of Science and Technology,16: 655-668.

[9] Zaninetti, L. and Ferraro,M., On the truncated Pareto distribution with applications, Central European Journal of Physics, 2008; 6:1-6.

[10] Singh, S. K., Singh,U. and Sharma, V.K.(2014),The truncated Lindley distribution: Inference and application, Journal of Statistics Applications & Probability, 3 :219-228.

[11] Sindhu, T. N. and Husai, Z. (2018), Mixture of two generalized inverted exponential distributions with censored sample: properties and estimation. *Statistica Applicata-Italian Journal of Applied Statistic*s 30 (3): 373-391.

[12] Shukla, K. K. (2021): Inverse Ishita distribution: properties and applications RT&A 16 (1) ,98-108

[13] Shukla, K. K. and Shanker, R., The Truncated Akash distribution: properties and applications, Biom Biostat Int J. (2020); 9(5):179–184. DOI: 10.15406/bbij.2020.09.00317

[14] Shukla, K. K. (2020): A truncated two parameter Pranav distribution and its applications RT&A, 15(3),103-116

[15] Robert, C., Casella, G., Introducing Monte Carlo Methods with R. In: Use R!, Springer (2010)

[16] Birnbaum, Z.W., Saunders, S.C., Estimation for a family of life distribution with applications to fatigue. Journal of applied probability, 1969; 6(2): 328-347.

[17] Fuller, E.J., Frieman, S., Quinn, J., Quinn, G., Carter, W. (1994), Fracture mechanics approach to the design of glass aircraft windows: A case study, SPIE Proc 1994; 2286:419-430.

MARKOV APPROACH FOR RELIABILITY AND AVAILABILITY ANALYSIS OF A FOUR UNIT REPAIRABLE SYSTEM

A. D. Yadav[1], N. Nandal[2], S.C. Malik[*]

•

[*,1,2]Department of Statistics, M.D. University, Rohtak
[1] amitdevstats@gmail.com,
[2]nsinghnandal@gmail.com,
[*]sc_malik@rediffmail.com
*

Abstract

Efforts have been made to analyze reliability and availability of a repairable system using Markov approach. The system has four non-identical units which work simultaneously. The system is assumed as completely non-functional at the failure of all the units. The failure and repair times as usual follow negative exponential distribution. The reliability measures of the system have been obtained by solving the Chapman-Kolmogorov equations using Laplace transform technique. The values of availability, reliability and mean time to system failure have been evaluated for particular values of the parameters considering all the units identical in nature. The effect of failure rate, repair rate and operating time on reliability, MTSF and availability has been studied. The application of the work has also been discussed with a real life example.

Keywords: Repairable System, Non-Identical Units, Markov Approach, Reliability Measures, Chapman-Kolmogorov Equations, Laplace Transform Technique

I. Introduction

In the era of fast-growing technology, everyone is interested to buy such a system which has a smaller number of design defectives and works as per the expectations. Therefore, the prime responsibility of the manufacturers is to produce the reliable products in order to stay for long in the competitive market. Today, the purpose of the system designers and reliability engineers is not only to develop a reliable system but also to identify the techniques that can be used to improve the system reliability. Over the years, several reliability improvement techniques have been evolved including provision of spare units, proper structure of the components, appropriate repair-maintenance policies and use of high-quality components. As a result of which researchers have also been succeeded in pointing out the guidelines for enhancing availability of the systems. It has been revealed that the availability of systems can be improved using the concept of redundancy in cold standby or in parallel mode.

There exist many systems in which functioning of the components (or units) are required in parallel mode not only to share the working stress but also minimize the failure risk of the systems. It has been observed that the use of some important methodologies including semi-Markov process and regenerative point technique has been made extensively by the researchers to analyze the well known reliability measures such as reliability, MTSF and availability of the repairable and non- repairable systems. The use of Markov approach has also been made by the researchers to assess reliability of systems with different structural designs of components. However, not much attention has been given by the researchers for analyzing reliability and availability of repairable systems using Markov approach. Chao et al. [1] carried out the reliability of large series system using Markov structure. El-Damcese et al. [2] analyzed a parallel repairable system with different failure modes. Umamaheshwari et al. [9] discussed a Markov model with human error and common cause. Li [4] obtained the reliability of a redundant system. Kalaiarasi et al. [3] analyzed system reliability using Markov Technique. Nandal and Bhardwaj [5] analyzed the profit of a parallel cold standby system using Lindley distribution. Saritha et al. [8] considered the reliability and availability for non-repairable & repairable systems using Markov modeling. Wang et al. [10] evaluated the reliability for multi state Markov repairable system. Reni et al. [7] studied the reliability of Markov models. Rathi et al. [6] discussed the reliability characteristics of a parallel system with priority concept. In that paper, authors considered that at least two, three and four modules must operate for the successful operation of the system. But they have not considered the case of repairable system and also when at least one module must work as we know that parallel system will work until it's all units fail.

Here we describe reliability and availability of a repairable system using Markov approach. The system has four non-identical units which work simultaneously. The system is assumed as completely non-functional at the failure of all the units. The failure and repair times as usual follow negative exponential distribution. The reliability measures of the system have been obtained by solving the Chapman-Kolmogorov equations using Laplace transform technique. The values of availability, reliability and mean time to system failure have been evaluated for particular values of the parameters considering all the units identical in nature. The effect of failure rate, repair rate and operating time on reliability, MTSF and availability has been studied. The application of the work has also been discussed with a real life example.

II. Assumptions and State Descriptions

1. The transition rates of the units follow negative exponential distribution
2. The repair of the failed unit is done immediately at the availability of the repair facility.
3. The system is declared failed at the failure of all units.
4. All the units are operative at time 't' in state '0' of the system.
5. The system is in state '1' at time 't' upon the failure of one unit, the repair is made immediately and other units are still in operation.
6. The system is in state '2' at time 't' upon the failure of two units and one of the failed units went into repair immediately and third unit is still operating.
7. The system is in state '3' at time 't' upon the failure of three units and one of the failed units went into repair immediately and fourth unit is still operating
8. All the units are failed at time 't' in state '4' of the system.

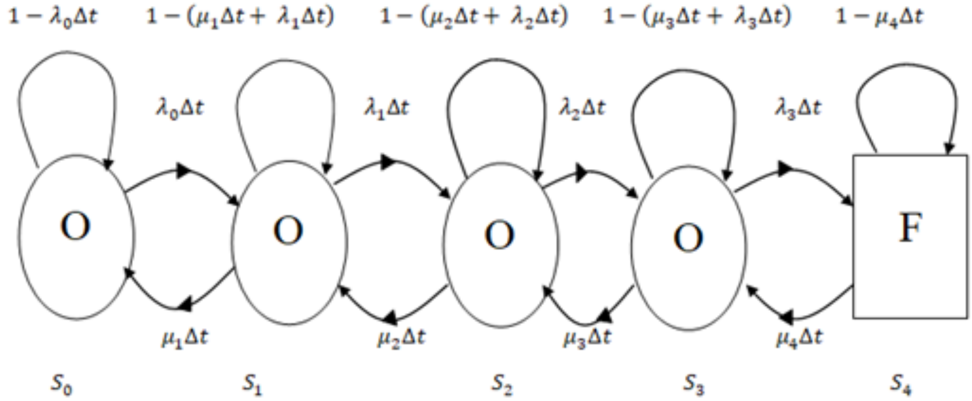


Figure 1: *State Transition Diagram*

a) Notations and Abbreviations

$\boldsymbol{\lambda_x}$ Failure rate of the system at state where x units have failed (x=0,1,2,3)
$\boldsymbol{\mu_y}$ Repair rate of the system at state where y units have failed (y=1,2,3,4)
$\boldsymbol{S_x}$ States (x=0,1,2,3,4)
t Time
$\boldsymbol{P_x(t)}$ Probability that the system is in state x at time t (x=0,1,2,3,4)
$\boldsymbol{A}(\infty)$ Steady State Availability of the system
R(t) Reliability of the system
MTSF Mean Time to System Failure

III. Reliability Measures of the System

a) Reliability

In passing from state i at time s to state j at time t (s<t), we must pass through some intermediate state k at some intermediate time u. When the continuous-time Markov chain is homogeneous, the Chapman-Kolmogorov equation may be written as:

$$P_{ij}(t+\Delta t) = \sum_{all\ k} P_{ik}(t)P_{kj}(\Delta t) \qquad for\ t,t \geq 0$$
$$= \sum_{k\neq j} P_{ik}(t)P_{kj}(\Delta t) + P_{ik}(t)P_{kj}(\Delta t) \tag{1}$$

The Chapman- Kolmogorov equations of the system can also be obtained from expression (1) as

$$P_0(t+\Delta t) = P_0(t)(1-\lambda_0\Delta t) + P_1(t)\mu_1\Delta t$$
$$P_1(t+\Delta t) = P_0(t)\lambda_0\Delta t + P_1(t)\big(1-(\mu_1\Delta t+\lambda_1\Delta t)\big) + P_2(t)\mu_2\Delta t$$
$$P_2(t+\Delta t) = P_1(t)\lambda_1\Delta t + P_2(t)\big(1-(\mu_2\Delta t+\lambda_2\Delta t)\big) + P_2(t)\mu_3\Delta t$$
$$P_3(t+\Delta t) = P_2(t)\lambda_2\Delta t + P_3(t)\big(1-(\mu_3\Delta t+\lambda_3\Delta t)\big)$$
$$P_4(t+\Delta t) = P_3(t)\lambda_3\Delta t + P_4(t) \tag{2-6}$$

These Markov equations are being developed by taking the probability of each state at time $t+\Delta t$. Above equations (2-6) can be rewritten as

$$\frac{P_0(t+\Delta t)-P_0(t)}{\Delta t} = P_0(t)(-\lambda_0) + P_1(t)\mu_1$$
$$\frac{P_1(t+\Delta t)-P_1(t)}{\Delta t} = P_0(t)\lambda_0 + P_1(t)\big(-(\mu_1+\lambda_1)\big) + P_2(t)\mu_2$$
$$\frac{P_2(t+\Delta t)-P_2(t)}{\Delta t} = P_1(t)\lambda_1 + P_2(t)\big(-(\mu_2+\lambda_2)\big) + P_2(t)\mu_3$$

$$\frac{P_3(t+\Delta t)-P_3(t)}{\Delta t} = P_2(t)\lambda_2 + P_3(t)\big(-(\mu_3 + \lambda_3)\big)$$

$$\frac{P_4(t+\Delta t)-P_4(t)}{\Delta t} = P_3(t)\lambda_3 \tag{7-11}$$

Converting these equations (7-11) to a differential equation and taking Lim $\Delta t \to 0$, we get

$$P_0'(t) = P_0(t)(-\lambda_0) + P_1(t)\mu_1$$
$$P_1'(t) = P_0(t)\lambda_0 + P_1(t)\big(-(\mu_1 + \lambda_1)\big) + P_2(t)\mu_2$$
$$P_2'(t) = P_1(t)\lambda_1 + P_2(t)\big(-(\mu_2 + \lambda_2)\big) + P_3(t)\mu_3$$
$$P_3'(t) = P_2(t)\lambda_2 + P_3(t)\big(-(\mu_3 + \lambda_3)\big)$$
$$P_4'(t) = P_3(t)\lambda_3 \tag{12-16}$$

Above equations (12-16) can be solved by using LT method.

$$sp_0(s) - P_0(0) = p_0(s)(-\lambda_0) + p_1(s)\mu_1$$
$$sp_1(s) - P_1(0) = p_0(s)\lambda_0 + p_1(s)\big(-(\mu_1 + \lambda_1)\big) + p_2(s)\mu_2$$
$$sp_2(s) - P_2(0) = p_1(s)\lambda_1 + p_2(s)\big(-(\mu_2 + \lambda_2)\big) + p_3(s)\mu_3$$
$$sp_3(s) - P_3(0) = p_2(s)\lambda_2 + p_3(s)\big(-(\mu_3 + \lambda_3)\big)$$
$$sp_4(s) - P_4(0) = p_3(s)\lambda_3 \tag{17-21}$$

Boundary conditions are

$$P_0(0) = 1, P_1(0) = 0, P_2(0) = 0, P_3(0) = 0, P_4(0) = 0$$

So, the system of equations will be

$$(s+\lambda_0)p_0(s) - p_1(s)\mu_1 = 1$$
$$-\lambda_0 p_0(s) + (s+\mu_1+\lambda_1)p_1(s) - \mu_2 p_2(s) = 0$$
$$-\lambda_1 p_1(s) + (s+\mu_2+\lambda_2)p_2(s) - \mu_3 p_3(s) = 0$$
$$-\lambda_2 p_2(s) + (s+\mu_3+\lambda_3)p_3(s) = 0$$
$$-\lambda_3 p_3(s) + (s)p_4(s) = 0$$

It can be written as

$$\begin{bmatrix} s+\lambda_0 & -\mu_1 & 0 & 0 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & s \end{bmatrix} \begin{bmatrix} p_0(s) \\ p_1(s) \\ p_2(s) \\ p_3(s) \\ p_4(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, solving for $p_0(s), p_1(s), p_2(s), p_3(s)$ *and* $p_4(s)$ using Cramer's Rule, we have

$$\Delta_R = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 0 & 0 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & s \end{vmatrix}$$

$$= s(s^4 + s^3\big(\lambda_0+\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3\big) + s^2\big(\lambda_0\lambda_1+\lambda_0\lambda_2+\lambda_0\lambda_3+\lambda_0\mu_2+\lambda_0\mu_3+\lambda_1\lambda_2+ \lambda_1\lambda_3+\lambda_1\mu_3+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_3\mu_1+\lambda_3\mu_2+\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3\big) + s\Big(\lambda_0\lambda_1\lambda_2+\lambda_0\lambda_1\lambda_3+ \lambda_0\lambda_2\lambda_3+\lambda_0\lambda_3\mu_2+\lambda_0\lambda_1\mu_3+\lambda_0\mu_2\mu_3+\lambda_1\lambda_2\lambda_3+\lambda_2\lambda_3\mu_1+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3\Big)+\lambda_0\lambda_1\lambda_2\lambda_3)$$

$$= s(s-m)(s-n)(s-o)(s-p)$$

Here m, n ,o and p are roots of Δ_R.

$$\Delta_{R_1} = \begin{vmatrix} 1 & -\mu_1 & 0 & 0 & 0 \\ 0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & s \end{vmatrix}$$

$$\Delta_{R_1} = s(s^4 + s^3\big(\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3\big) + s^2\big(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\mu_3+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_3\mu_1+ \lambda_3\mu_2+\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3\big) + s\Big(\lambda_1\lambda_2\lambda_3+\lambda_2\lambda_3\mu_1+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3\Big))$$

$$\Delta_{R_2} = \begin{vmatrix} s+\lambda_0 & 1 & 0 & 0 & 0 \\ -\lambda_0 & 0 & -\mu_2 & 0 & 0 \\ 0 & 0 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & s \end{vmatrix}$$

$$\Delta_{R_2} = \lambda_0 s^3 + \lambda_0 s^2(\lambda_2 + \lambda_3 + \mu_2 + \mu_3) + \lambda_0 s(\lambda_2\lambda_3 + \lambda_3\mu_2 + \mu_2\mu_3)$$

$$\Delta_{R_3} = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 1 & 0 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & 0 & 0 & 0 \\ 0 & -\lambda_1 & 0 & -\mu_3 & 0 \\ 0 & 0 & 0 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & s \end{vmatrix}$$

$$\Delta_{R_3} = \lambda_0\lambda_1 s^2 + s(\lambda_0\lambda_1\lambda_3 + \lambda_0\lambda_1\mu_3)$$

$$\Delta_{R_4} = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 0 & 1 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & 0 & 0 \\ 0 & 0 & -\lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & s \end{vmatrix}$$

$$\Delta_{R_4} = \lambda_0\lambda_1\lambda_2 s$$

$$\Delta_{R_5} = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 0 & 0 & 1 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & 0 \end{vmatrix}$$

$$\Delta_{R_5} = \lambda_0\lambda_1\lambda_2\lambda_3$$

Now,

$$p_0(\mathrm{s}) = \frac{\Delta_{R_1}}{\Delta_R}$$

$$= \frac{\begin{array}{c}[s(s^4 + s^3(\lambda_1 + \lambda_2 + \lambda_3 + \mu_1 + \mu_2 + \mu_3) \\ +s^2(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\mu_3 + \lambda_2\lambda_3 + \lambda_2\mu_1 + \lambda_3\mu_1 + \lambda_3\mu_2 + \mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3) \\ + s\left(\lambda_1\lambda_2\lambda_3 + \lambda_2\lambda_3\mu_1 + \lambda_3\mu_1\mu_2 + \mu_1\mu_2\mu_3\right))]\end{array}}{\Delta_R}$$

$$p_1(\mathrm{s}) = \frac{\Delta_{R_2}}{\Delta_R}$$

$$= \frac{\lambda_0 s^3 + \lambda_0 s^2(\lambda_2 + \lambda_3 + \mu_2 + \mu_3) + \lambda_0 s(\lambda_2\lambda_3 + \lambda_3\mu_2 + \mu_2\mu_3)}{\Delta_R}$$

$$p_2(\mathrm{s}) = \frac{\Delta_{R_3}}{\Delta_R}$$

$$= \frac{\lambda_0 s^3 + \lambda_0 s^2(\lambda_2 + \lambda_3 + \mu_2 + \mu_3) + \lambda_0 s(\lambda_2\lambda_3 + \lambda_3\mu_3 + \mu_2\mu_3)}{\Delta_R}$$

$$p_3(\mathrm{s}) = \frac{\Delta_{R_4}}{\Delta_R}$$

$$= \frac{\lambda_0\lambda_1 s^2 + s(\lambda_0\lambda_1\lambda_3 + \lambda_0\lambda_1\mu_3)}{\Delta_R}$$

$$p_4(\mathrm{s}) = \frac{\Delta_{R_5}}{\Delta_R}$$

$$= \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{\Delta_R}$$

Taking Laplace Inverse of $p_0(\mathrm{s}), p_1(\mathrm{s})$, $p_2(\mathrm{s})$ and $p_3(\mathrm{s})$, we get

$$p_0(\mathrm{t}) = e^{mt}\left[\frac{\begin{array}{c}m^3 + m^2(\lambda_1 + \lambda_2 + \lambda_3 + \mu_1 + \mu_2 + \mu_3) \\ +m(\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3 + \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\mu_3 + \lambda_2\lambda_3 + \lambda_2\mu_1 + \lambda_3\mu_1 + \lambda_3\mu_2) \\ + \left(\lambda_1\lambda_2\lambda_3 + \lambda_2\lambda_3\mu_1 + \lambda_3\mu_1\mu_2 + \mu_1\mu_2\mu_3\right)\end{array}}{(m-n)(m-o)(m-p)}\right]$$

$$+e^{nt}\left[\frac{\begin{array}{c} n^3+n^2(\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3) \\ +n(\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3+\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\mu_3+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_3\mu_1+\lambda_3\mu_2) \\ +(\lambda_1\lambda_2\lambda_3+\lambda_2\lambda_3\mu_1+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3)\end{array}}{(n-m)(n-o)(n-p)}\right]$$

$$+e^{ot}\left[\frac{\begin{array}{c} o^3+o^2(\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3) \\ +o(\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3+\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\mu_3+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_3\mu_1+\lambda_3\mu_2) \\ +(\lambda_1\lambda_2\lambda_3+\lambda_2\lambda_3\mu_1+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3)\end{array}}{(o-m)(o-n)(o-p)}\right]$$

$$+e^{pt}\left[\frac{\begin{array}{c} p^3+p^2(\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3) \\ +p(\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3+\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\mu_3+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_3\mu_1+\lambda_3\mu_2) \\ +(\lambda_1\lambda_2\lambda_3+\lambda_2\lambda_3\mu_1+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3)\end{array}}{(p-m)(p-n)(p-o)}\right]$$

$$p_1(\mathrm{t}) = e^{mt}\Big[\frac{\lambda_0 m^2+\lambda_0 m^1(\lambda_2+\lambda_3+\mu_2+\mu_3)+\lambda_0(\lambda_2\lambda_3+\lambda_3\mu_3+\mu_2\mu_3)}{(m-n)(m-o)(m-p)}$$
$$+e^{nt}\frac{\lambda_0 n^2+\lambda_0 n^1(\lambda_2+\lambda_3+\mu_2+\mu_3)+\lambda_0(\lambda_2\lambda_3+\lambda_3\mu_3+\mu_2\mu_3)}{(n-m)(n-o)(n-p)}$$
$$+e^{ot}\frac{\lambda_0 o^2+\lambda_0 o^1(\lambda_2+\lambda_3+\mu_2+\mu_3)+\lambda_0(\lambda_2\lambda_3+\lambda_3\mu_3+\mu_2\mu_3)}{(o-m)(o-n)(o-p)}$$
$$+e^{pt}\frac{\lambda_0 p^3+\lambda_0 p^2(\lambda_2+\lambda_3+\mu_2+\mu_3)+\lambda_0(\lambda_2\lambda_3+\lambda_3\mu_3+\mu_2\mu_3)}{(p-m)(p-n)(p-o)}\Big]$$

$$p_2(\mathrm{t}) = \lambda_0\lambda_1\Big[\frac{e^{mt}(\mu_3+\lambda_3+m)}{(m-n)(m-o)(m-p)}+\frac{e^{nt}(\mu_3+\lambda_3+n)}{(n-m)(n-o)(n-p)}+\frac{e^{ot}(\mu_3+\lambda_3+o)}{(o-m)(o-n)(o-p)}+\frac{e^{nt}(\mu_3+\lambda_3+p)}{(p-m)(p-n)(p-o)}\Big]$$

$$p_3(\mathrm{t}) = \lambda_0\lambda_1\lambda_2\Big[\frac{e^{mt}}{(m-n)(m-o)(m-p)}+\frac{e^{nt}}{(n-m)(n-o)(n-p)}+\frac{e^{ot}}{(o-m)(o-n)(o-p)}+\frac{e^{nt}}{(p-m)(p-n)(p-o)}\Big]$$

$$p_4(\mathrm{t}) = \lambda_0\lambda_1\lambda_2\lambda_3\Big[\frac{e^{mt}}{(m-n)(m-o)(m-p)}+\frac{e^{nt}}{(n-m)(n-o)(n-p)}+\frac{e^{ot}}{(o-m)(o-n)(o-p)}+\frac{e^{nt}}{(p-m)(p-n)(p-o)}+\frac{1}{mnop}\Big]$$

Thus, Reliability of the system is

$$\mathrm{R(t)} = p_0(t)+p_1(t)+p_2(t)+p_3(t) = 1-p_4(t)$$
$$= 1-\lambda_0\lambda_1\lambda_2\lambda_3\Big[\frac{e^{mt}}{(m-n)(m-o)(m-p)}+\frac{e^{nt}}{(n-m)(n-o)(n-p)}+\frac{e^{ot}}{(o-m)(o-n)(o-p)}+\frac{e^{nt}}{(p-m)(p-n)(p-o)}+\frac{1}{mnop}\Big] \quad (22)$$

b) Mean Time to System Failure (MTSF)

On integrating equation (22), we get the expression for MTSF as

$$MTSF = \int_0^\infty R(t)dt$$

$$MTSF = \frac{\lambda_1\lambda_2\lambda_3 + \mu_1(\lambda_3(\lambda_2+\mu_2)+\mu_2\mu_3) + \lambda_0(\lambda_3(\lambda_2+\mu_2)+\mu_2\mu_3+\lambda_1(\lambda_2+\lambda_3+\mu_3))}{mnop}$$

c) Availability

Now, using equation (1), we get the Chapman- Kolmogorov equations for the system to derive the expression for availability as

$$P_0(t+\Delta t) = P_0(t)(1-\lambda_0\Delta t) + P_1(t)\mu_1\Delta t$$
$$P_1(t+\Delta t) = P_0(t)\lambda_0\Delta t + P_1(t)\big(1-(\mu_1\Delta t + \lambda_1\Delta t)\big) + P_2(t)\mu_2\Delta t$$
$$P_2(t+\Delta t) = P_1(t)\lambda_1\Delta t + P_2(t)\big(1-(\mu_2\Delta t + \lambda_2\Delta t)\big) + P_2(t)\mu_3\Delta t$$
$$P_3(t+\Delta t) = P_2(t)\lambda_2\Delta t + P_3(t)\big(1-(\mu_3\Delta t + \lambda_3\Delta t)\big) + P_3(t)\mu_4\Delta t$$
$$P_4(t+\Delta t) = P_3(t)\lambda_3\Delta t + P_4(t)(1-\mu_4\Delta t) \quad (23\text{-}27)$$

These Markov equations are being developed by taking the probability of each state at time $t + \Delta t$. Above equations (23-27) can be rewritten as

$$\frac{P_0(t+\Delta t)-P_0(t)}{\Delta t} = P_0(t)(-\lambda_0) + P_1(t)\mu_1$$
$$\frac{P_1(t+\Delta t)-P_1(t)}{\Delta t} = P_0(t)\lambda_0 + P_1(t)\big(-(\mu_1+\lambda_1)\big) + P_2(t)\mu_2$$
$$\frac{P_2(t+\Delta t)-P_2(t)}{\Delta t} = P_1(t)\lambda_1 + P_2(t)\big(-(\mu_2+\lambda_2)\big) + P_2(t)\mu_3$$
$$\frac{P_3(t+\Delta t)-P_3(t)}{\Delta t} = P_2(t)\lambda_2 + P_3(t)\big(-(\mu_3+\lambda_3)\big) + P_4(t)\mu_4$$
$$\frac{P_4(t+\Delta t)-P_4(t)}{\Delta t} = P_3(t)\lambda_3 + P_4(t)(-\mu_4) \quad (28\text{-}32)$$

Converting equations (28-32) to a differential equation and taking Lim $\Delta t \to 0$, we get

$$P_0'(t) = P_0(t)(-\lambda_0) + P_1(t)\mu_1$$
$$P_1'(t) = P_0(t)\lambda_0 + P_1(t)\big(-(\mu_1+\lambda_1)\big) + P_2(t)\mu_2$$
$$P_2'(t) = P_1(t)\lambda_1 + P_2(t)\big(-(\mu_2+\lambda_2)\big) + P_3(t)\mu_3$$
$$P_3'(t) = P_2(t)\lambda_2 + P_3(t)\big(-(\mu_3+\lambda_3)\big) + P_4(t)\mu_4$$
$$P_4'(t) = P_3(t)\lambda_3 + P_4(t)(-\mu_4) \quad (33\text{-}37)$$

Above equations (33-37) can be solved by using LT method.

$$sp_0(s) - P_0(0) = p_0(s)(-\lambda_0) + p_1(s)\mu_1$$
$$sp_1(s) - P_1(0) = p_0(s)\lambda_0 + p_1(s)\big(-(\mu_1+\lambda_1)\big) + p_2(s)\mu_2$$
$$sp_2(s) - P_2(0) = p_1(s)\lambda_1 + p_2(s)\big(-(\mu_2+\lambda_2)\big) + p_3(s)\mu_3$$
$$sp_3(s) - P_3(0) = p_2(s)\lambda_2 + p_3(s)\big(-(\mu_3+\lambda_3)\big) + p_4(s)\mu_4$$
$$sp_4(s) - P_4(0) = p_3(s)\lambda_3 + p_4(s)(-\mu_4) \quad (38\text{-}42)$$

Boundary conditions are

$$P_0(0) = 1, P_1(0) = 0, P_2(0) = 0, P_3(0) = 0, P_4(0) = 0$$

So, the system of equations will be

$$(s+\lambda_0)p_0(s) - p_1(s)\mu_1 = 1$$
$$-\lambda_0 p_0(s) + (s+\mu_1+\lambda_1)p_1(s) - \mu_2 p_2(s) = 0$$
$$-\lambda_1 p_1(s) + (s+\mu_2+\lambda_2)p_2(s) - \mu_3 p_3(s) = 0$$
$$-\lambda_2 p_2(s) + (s+\mu_3+\lambda_3)p_3(s) - \mu_4 p_4(s) = 0$$
$$-\lambda_3 p_3(s) + (s+\mu_4)p_4(s) = 0$$

It can be written as

$$\begin{bmatrix} s+\lambda_0 & -\mu_1 & 0 & 0 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & -\mu_4 \\ 0 & 0 & 0 & -\lambda_3 & s+\mu_4 \end{bmatrix} \begin{bmatrix} p_0(s) \\ p_1(s) \\ p_2(s) \\ p_3(s) \\ p_4(s) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, solving for $p_0(s), p_1(s), p_2(s)$ *and* $p_3(s)$using Cramer's Rule, we have

$$\Delta_A = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 0 & 0 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & -\mu_4 \\ 0 & 0 & 0 & -\lambda_3 & s+\mu_4 \end{vmatrix}$$

$$\Delta_A = s(s^4 + s^3(\lambda_0+\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3+\mu_4) + s^2(\lambda_0\lambda_1+\lambda_0\lambda_2+\lambda_0\mu_3+\lambda_0\mu_3+\lambda_0\mu_4+\lambda_1\lambda_3+\lambda_1\lambda_2+\lambda_1\mu_3+\lambda_1\mu_4+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_2\mu_4+\lambda_3\mu_1+\lambda_3\mu_2+\mu_1\mu_2+\mu_1\mu_3+\mu_1\mu_4+\mu_2\mu_3+\mu_2\mu_4+\mu_3\mu_4) + s\left(\lambda_0\lambda_1\lambda_2+\lambda_0\lambda_1\lambda_3+\lambda_0\lambda_2\lambda_3+\lambda_0\lambda_3\mu_2+\lambda_0\lambda_1\mu_3+\lambda_0\mu_2\mu_3+\lambda_0\lambda_1\mu_4+\lambda_0\lambda_2\mu_4+\lambda_0\mu_2\mu_4+\lambda_0\mu_3\mu_4+\lambda_1\lambda_2\mu_3+\lambda_1\lambda_2\mu_4+\lambda_1\mu_3\mu_4+\lambda_2\lambda_3\mu_1+\lambda_2\mu_1\mu_4+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3+\mu_1\mu_2\mu_4+\mu_2\mu_3\mu_4\right) + (\mu_1\mu_2\mu_3\mu_4+\lambda_0\mu_2\mu_3\mu_4+\lambda_0\lambda_1\mu_3\mu_4+\lambda_0\lambda_1\lambda_2\mu_4+\lambda_0\lambda_1\lambda_2\lambda_3))$$

$$= s(s-a)(s-b)(s-c)(s-d)$$

Here, a ,b ,c and d are the roots of Δ_A.

$$\Delta_{A_1} = \begin{vmatrix} 1 & -\mu_1 & 0 & 0 & 0 \\ 0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & -\mu_4 \\ 0 & 0 & 0 & -\lambda_3 & s+\mu_4 \end{vmatrix}$$

$$\Delta_{A_1} = s^4 + s^3(\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3+\mu_4) + s^2(\lambda_1\lambda_3+\lambda_1\lambda_2+\lambda_1\mu_3+\lambda_1\mu_4+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_2\mu_4+\lambda_3\mu_1+\lambda_3\mu_2+\mu_1\mu_2+\mu_1\mu_3+\mu_1\mu_4+\mu_2\mu_3+\mu_2\mu_4+\mu_3\mu_4) + s\left(\lambda_1\lambda_2\mu_3+\lambda_1\lambda_2\mu_4+\lambda_1\mu_3\mu_4+\lambda_2\lambda_3\mu_1+\lambda_2\mu_1\mu_4+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3+\mu_1\mu_2\mu_4+\mu_2\mu_3\mu_4\right) + \mu_1\mu_2\mu_3\mu_4$$

$$\Delta_{A_2} = \begin{vmatrix} s+\lambda_0 & 1 & 0 & 0 & 0 \\ -\lambda_0 & 0 & -\mu_2 & 0 & 0 \\ 0 & 0 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & -\mu_4 \\ 0 & 0 & 0 & -\lambda_3 & s+\mu_4 \end{vmatrix}$$

$$\Delta_{A_2} = as^3 + as^2(\lambda_2+\lambda_3+\mu_2+\mu_3+\mu_4) + as(\lambda_2\lambda_3+\lambda_2\mu_4+\lambda_3\mu_2+\mu_2\mu_3+\mu_2\mu_4+\mu_3\mu_4) + \lambda_0\mu_2\mu_3\mu_4$$

$$\Delta_{A_3} = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 1 & 0 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & 0 & 0 & 0 \\ 0 & -\lambda_1 & 0 & -\mu_3 & 0 \\ 0 & 0 & 0 & s+\lambda_3+\mu_3 & -\mu_4 \\ 0 & 0 & 0 & -\lambda_3 & s+\mu_4 \end{vmatrix}$$

$$\Delta_{A_3} = \lambda_0\lambda_1 s^2 + s(\lambda_0\lambda_1\lambda_3+\lambda_0\lambda_1\mu_3+\lambda_0\lambda_1\mu_4) + \lambda_0\lambda_1\mu_3\mu_4$$

$$\Delta_{A_4} = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 0 & 1 & 0 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & 0 & 0 \\ 0 & 0 & -\lambda_2 & 0 & -\mu_4 \\ 0 & 0 & 0 & 0 & s+\mu_4 \end{vmatrix}$$

$$\Delta_{A_4} = \lambda_0\lambda_1\lambda_2\mu_4 + \lambda_0\lambda_1\lambda_2 s$$

$$\Delta_{A_5} = \begin{vmatrix} s+\lambda_0 & -\mu_1 & 0 & 0 & 1 \\ -\lambda_0 & s+\lambda_1+\mu_1 & -\mu_2 & 0 & 0 \\ 0 & -\lambda_1 & s+\lambda_2+\mu_2 & -\mu_3 & 0 \\ 0 & 0 & -\lambda_2 & s+\lambda_3+\mu_3 & 0 \\ 0 & 0 & 0 & -\lambda_3 & 0 \end{vmatrix}$$

$$\Delta_{A_5} = \lambda_0\lambda_1\lambda_2\lambda_3$$

Now,

$$p_0(\mathrm{s}) = \frac{\Delta_{A_1}}{\Delta_A}$$

$$= [s^4 + s^3(\lambda_1+\lambda_2+\lambda_3+\mu_1+\mu_2+\mu_3+\mu_4) + s^2(\lambda_1\lambda_3+\lambda_1\lambda_2+\lambda_1\mu_3+\lambda_1\mu_4+\lambda_2\lambda_3+\lambda_2\mu_1+\lambda_2\mu_4+\lambda_3\mu_1+\lambda_3\mu_2+\mu_1\mu_2+\mu_1\mu_3+\mu_1\mu_4+\mu_2\mu_3+\mu_2\mu_4+\mu_3\mu_4) + s(\lambda_1\lambda_2\mu_3+\lambda_1\lambda_2\mu_4+\lambda_1\mu_3\mu_4+\lambda_2\lambda_3\mu_1+\lambda_2\mu_1\mu_4+\lambda_3\mu_1\mu_2+\mu_1\mu_2\mu_3+\mu_1\mu_2\mu_4+\mu_2\mu_3\mu_4) + \mu_1\mu_2\mu_3\mu_4]/\Delta_A$$

$$p_1(\mathrm{s}) = \frac{\Delta_{A_2}}{\Delta_A}$$

$$= \frac{[\lambda_0 s^3 + \lambda_0 s^2(\lambda_2+\lambda_3+\mu_2+\mu_3+\mu_4) + \lambda_0 s(\lambda_2\lambda_3+\lambda_2\mu_4+\lambda_3\mu_2+\mu_2\mu_3+\mu_2\mu_4+\mu_3\mu_4) + \lambda_0\mu_2\mu_3\mu_4]}{\Delta_A}$$

$$p_2(\mathrm{s}) = \frac{\Delta_{A_3}}{\Delta_A} = \frac{[\lambda_0\lambda_1 s^2 + s(\lambda_0\lambda_1\lambda_3+\lambda_0\lambda_1\mu_3+\lambda_0\lambda_1\mu_4)+\lambda_0\lambda_1\mu_3\mu_4]}{\Delta_A}$$

$$p_3(\mathrm{s}) = \frac{\Delta_{A_4}}{\Delta_A} = \frac{[\lambda_0\lambda_1\lambda_2\mu_4 + \lambda_0\lambda_1\lambda_2 s]}{\Delta_A}$$

$$p_4(\mathrm{s}) = \frac{\Delta_{A_5}}{\Delta_A} = \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{\Delta_A}$$

Taking Laplace Inverse of $p_0(\mathrm{s}), p_1(\mathrm{s}), p_2(\mathrm{s}), p_3(\mathrm{s})$ and $p_4(\mathrm{s})$, we get

$$p_0(\mathrm{t}) = \frac{e^{at}(a\lambda_1(a\lambda_3+(a+\mu_3)(a+\mu_4)+\lambda_2(a+\lambda_3+\mu_4)) +(a+\mu_1)(a\lambda_2(a+\lambda_3+\mu_4))+(a+\mu_2)(a\lambda_3+(a+\mu_3)(a+\mu_4)))}{a(a-b)(a-c)(a-d)}$$
$$+\frac{e^{bt}(b\lambda_1(b\lambda_3+(b+\mu_3)(b+\mu_4)+\lambda_2(b+\lambda_3+\mu_4)) +(b+\mu_1)(b\lambda_2(b+\lambda_3+\mu_4)+(b+\mu_2)(b\lambda_3+(b+\mu_3)(b+\mu_4))))}{b(b-a)(b-c)(b-d)}$$
$$+\frac{e^{ct}(c\lambda_1(c\lambda_3+(c+\mu_3)(c+\mu_4)+\lambda_2(c+\lambda_3+\mu_4)) +(c+\mu_1)(c\lambda_2(c+\lambda_3+\mu_4)+(c+\mu_2)(c\lambda_3+(c+\mu_3)(c+\mu_4))))}{c(c-a)(c-b)(c-d)}$$
$$+\frac{e^{dt}(d\lambda_1(d\lambda_3+(d+\mu_3)(d+\mu_4)+\lambda_2(d+\lambda_3+\mu_4)) +(d+\mu_1)(d\lambda_2(d+\lambda_3+\mu_4)+(d+\mu_2)(d\lambda_3+(d+\mu_3)(d+\mu_4))))}{d(d-a)(d-b)(d-c)}$$
$$+\frac{\mu_1\mu_2\mu_3\mu_4}{abcd}$$

$$p_1(\mathrm{t}) = \lambda_0\Big(\frac{\mu_2\mu_3\mu_4}{abcd} + \frac{e^{at}(a\lambda_2(a+\lambda_3+\mu_4)+(a+\mu_2)(a\lambda_3+(a+\mu_3)(a+\mu_4)))}{a(a-b)(a-c)(a-d)}$$
$$+\frac{e^{bt}(b\lambda_2(b+\lambda_3+\mu_4)+(b+\mu_2)(b\lambda_3+(b+\mu_3)(b+\mu_4)))}{b(b-a)(b-c)(b-d)}$$
$$+\frac{e^{ct}(c\lambda_2(c+\lambda_3+\mu_4)+(c+\mu_2)(c\lambda_3+(c+\mu_3)(c+\mu_4)))}{c(c-a)(c-b)(c-d)}$$
$$+\frac{e^{dt}(d\lambda_2(d+\lambda_3+\mu_4)+(d+\mu_2)(d\lambda_3+(d+\mu_3)(d+\mu_4)))}{d(d-a)(d-b)(d-c)}\Big)$$

$$p_2(\mathrm{t}) = \lambda_0\lambda_1\Big(\frac{\mu_3\mu_4}{abcd} + \frac{e^{at}(a\lambda_3+(a+\mu_3)(a+\mu_4))}{a(a-b)(a-c)(a-d)} + \frac{e^{bt}(b\lambda_3+(b+\mu_3)(b+\mu_4))}{b(b-a)(b-c)(b-d)}$$
$$+\frac{e^{ct}(c\lambda_3+(c+\mu_3)(c+\mu_4))}{c(c-a)(c-b)(c-d)} + \frac{e^{dt}(d\lambda_3+(d+\mu_3)(d+\mu_4))}{d(d-a)(d-b)(d-c)}\Big)$$

$$p_3(\mathrm{t}) = \lambda_0\lambda_1\lambda_2\Big(\frac{\mu_4}{abcd} + \frac{e^{at}(a+\mu_4)}{a(a-b)(a-c)(a-d)} - \frac{e^{bt}(b+\mu_4)}{b(b-a)(b-c)(b-d)}$$
$$-\frac{e^{ct}(c+\mu_4)}{c(c-a)(c-b)(c-d)} - \frac{e^{dt}(d+\mu_4)}{d(d-a)(d-b)(d-c)}\Big)$$

$$p_4(\mathrm{t}) = \lambda_0\lambda_1\lambda_2\lambda_3\Big(\frac{1}{abcd} + \frac{e^{at}}{a(a-b)(a-c)(a-d)} + \frac{e^{bt}}{b(b-a)(b-c)(b-d)}$$
$$+\frac{e^{ct}}{c(c-a)(c-b)(c-d)} + \frac{e^{dt}}{d(d-a)(d-b)(d-c)}\Big)$$

The availability is given by

$$A(t) = p_0(\mathrm{t}) + p_1(\mathrm{t}) + p_2(\mathrm{t}) + p_3(\mathrm{t}) = 1 - p_4(t)$$
$$= 1 - \lambda_0\lambda_1\lambda_2\lambda_3\Big(\frac{1}{abcd} + \frac{e^{at}}{a(a-b)(a-c)(a-d)} + \frac{e^{bt}}{b(b-a)(b-c)(b-d)}$$
$$+\frac{e^{ct}}{c(c-a)(c-b)(c-d)} + \frac{e^{dt}}{d(d-a)(d-b)(d-c)}\Big)$$

The steady state availability is given by

$$A(\infty) = \lim_{s\to 0} s(p_0(s) + p_1(s) + p_2(s) + p_3(s))$$

$$= \lim_{s\to 0} [\, s[s^4 + s^3(\lambda_1 + \lambda_2 + \lambda_3 + \mu_1 + \mu_2 + \mu_3 + \mu_4) + s^2(\lambda_1\lambda_3 + \lambda_1\lambda_2 + \lambda_1\mu_3 + \lambda_1\mu_4 + \lambda_2\lambda_3 + \lambda_2\mu_1 + \lambda_2\mu_4 + \lambda_3\mu_1 + \lambda_3\mu_2 + \mu_1\mu_2 + \mu_1\mu_3 + \mu_1\mu_4 + \mu_2\mu_3 + \mu_2\mu_4 + \mu_3\mu_4) + s(\lambda_1\lambda_2\mu_3 + \lambda_1\lambda_2\mu_4 + \lambda_1\mu_3\mu_4 + \lambda_2\lambda_3\mu_1 + \lambda_2\mu_1\mu_4 + \lambda_3\mu_1\mu_2 + \mu_1\mu_2\mu_3 + \mu_1\mu_2\mu_4 + \mu_2\mu_3\mu_4) + \mu_1\mu_2\mu_3\mu_4] [\lambda_0 s^3 + \lambda_0 s^2(\lambda_2 + \lambda_3 + \mu_2 + \mu_3 + \mu_4) + \lambda_0 s(\lambda_2\lambda_3 + \lambda_2\mu_4 + \lambda_3\mu_2 + \mu_2\mu_3 + \mu_2\mu_4 + \mu_3\mu_4) + \lambda_0\mu_2\mu_3\mu_4] + [\lambda_0\lambda_1 s^2 + s(\lambda_0\lambda_1\lambda_3 + \lambda_0\lambda_1\mu_3 + \lambda_0\lambda_1\mu_4) + \lambda_0\lambda_1\mu_3\mu_4] + \lambda_0\lambda_1\lambda_2\mu_4 + \lambda_0\lambda_1\lambda_2 s]/\Delta_A$$

$$= \frac{\mu_1\mu_2\mu_3\mu_4 + \lambda_0(\mu_2\mu_3\mu_4 + \lambda_1(\lambda_2\mu_4 + \mu_3\mu_4))}{abcd}$$

Or it can also be expressed in the form of

$$= \frac{\mu_1\mu_2\mu_3\mu_4 + \lambda_0\mu_2\mu_3\mu_4 + \lambda_0\lambda_1\mu_3\mu_4 + \lambda_0\lambda_1\lambda_2\mu_4}{\mu_1\mu_2\mu_3\mu_4 + \lambda_0\mu_2\mu_3\mu_4 + \lambda_0\lambda_1\mu_3\mu_4 + \lambda_0\lambda_1\lambda_2\mu_4 + \lambda_0\lambda_1\lambda_2\lambda_3}$$

b) Particular Case

Now, if all the components are taken as identical units

For Availability we have,

$$\lambda_0 = 4\lambda, \lambda_1 = 3\lambda, \lambda_2 = 2\lambda, \lambda_3 = \lambda, \mu_4 = 4\mu, \mu_3 = 3\mu, \mu_2 = 2\mu \text{ and } \mu_1 = \mu$$

$$A(t) = 1 - 24\lambda^4\left(\frac{1}{abcd} + \frac{e^{at}}{a(a-b)(a-c)(a-d)} + \frac{e^{bt}}{b(b-a)(b-c)(b-d)} + \frac{e^{ct}}{c(c-a)(c-b)(c-d)} + \frac{e^{dt}}{d(d-a)(d-b)(d-c)}\right)$$

$$A(\infty) = \frac{48\lambda^3\mu + 144\lambda^2\mu^2 + 96\lambda\mu^3 + 24\mu^4}{12\lambda^4 + 48\lambda^3\mu + 144\lambda^2\mu^2 + 96\lambda\mu^3 + 24\mu^4}$$

For Reliability we have

$$\lambda_0 = 4\lambda, \lambda_1 = 3\lambda, \lambda_2 = 2\lambda, \lambda_2 = \lambda, \mu_4 = 0, \mu_3 = 3\mu, \mu_2 = 2\mu \text{ and } \mu_1 = \mu$$

$$R(t) = 1 - 24\lambda^4\left[\frac{e^{mt}}{(m-n)(m-o)(m-p)} + \frac{e^{nt}}{(n-m)(n-o)(n-p)} + \frac{e^{ot}}{(o-m)(o-n)(o-p)} + \frac{e^{nt}}{(p-m)(p-n)(p-o)} + \frac{1}{mnop}\right]$$

For MTSF we have

$$MTSF = \frac{6\lambda^3 + \mu(6\mu^2 + \lambda(2\lambda + 2\mu)) + 4\lambda(6\mu^2 + \lambda(2\lambda + 2\mu) + 3\lambda(3\lambda + 3\mu))}{mnop}$$

VI. Numerical and Graphical Presentation

Here, we evaluate the reliability, availability and MTSF for the arbitrary values of repair rate (μ) and failure rate (λ) with operating time (t) of the components. The numerical and graphical representation of the results is given below:

From Figure 2, it is observed that the reliability of the system declines with the increase of failure rate and operating time. Also, the reliability of the system is increases with an increase in repair rate of the units.

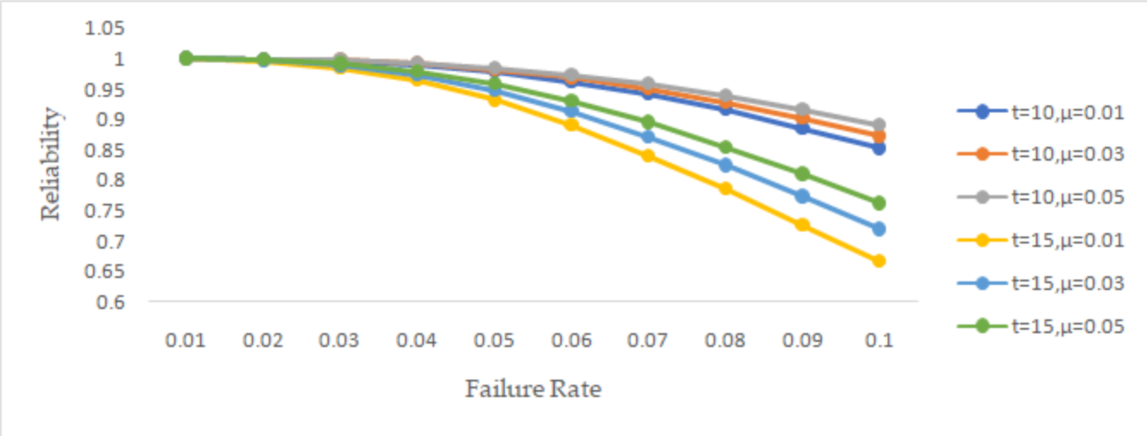


Figure 2: *Reliability V/s Repair Rate (μ) and Failure Rate(λ) with Operating Time(t)*

From Figure 3, it is observed that the MTSF of the system declines with the increase of failure rate. Also, the MTSF of the system is increases with an increase in repair rate of the units.

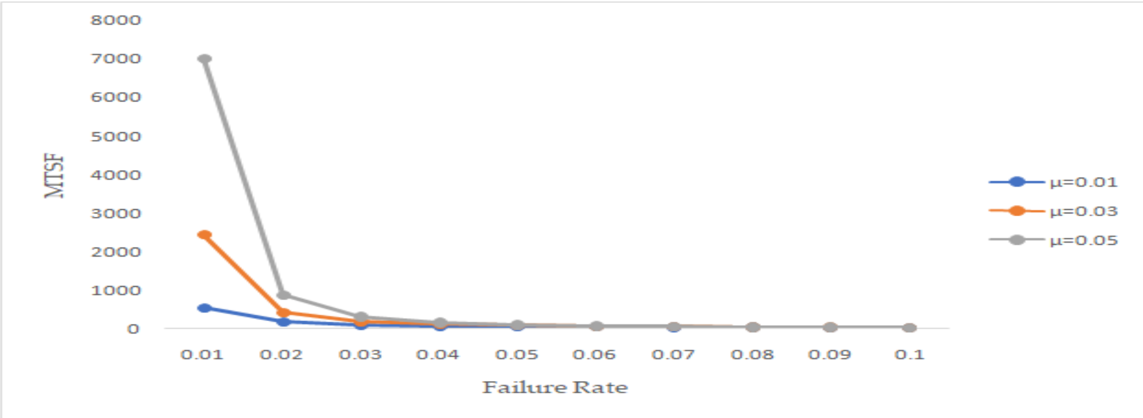


Figure 3: *MTSF V/s Repair Rate (μ) and Failure Rate(λ)*

From Figure 4, it is observed that the availability (A(t)) of the system declines with the increase of failure rate and operating time. Also, the availability of the system is increases with an increase in repair rate of the units.

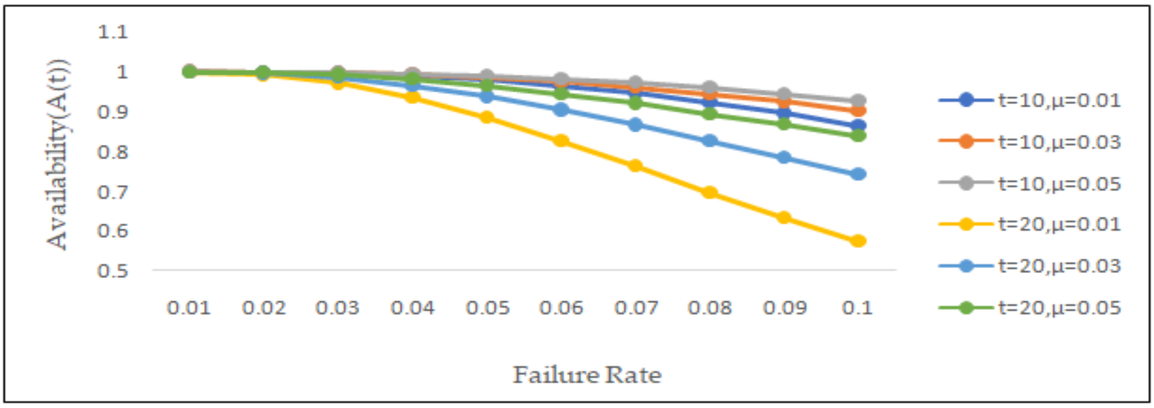


Figure 4: *Availability V/s Repair Rate (μ) and Failure Rate(λ) with Operating Time(t)*

From Figure 5, it is observed that the availability (A(∞)) of the system declines with the increase of failure rate. Also, the availability of the system is increases with an increase in repair rate of the units.

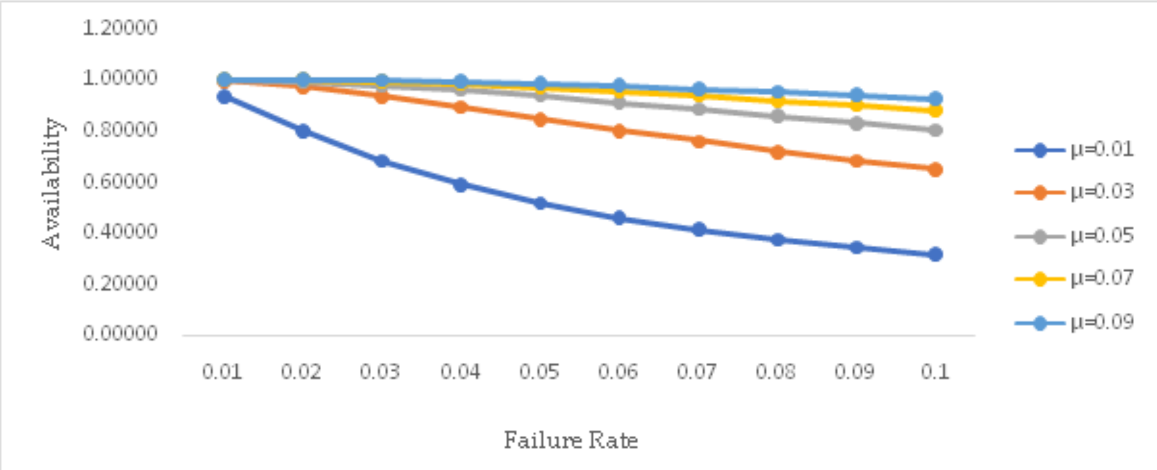


Figure 5: *Availability V/s Repair Rate (μ) and Failure Rate (λ)*

VII. Application

The real-life application of this study can be visualized at toll plaza. On a toll plaza, vehicles enter and exit the mainline roadway from certain locations along the highway. All vehicles must stop and pay the toll at any of the tollbooths that work simultaneously to reduce the gathering of traffic vehicles. The drivers prefer a tollbooth that has a shorter lane rather than a longer lane to minimize their own travel time. .

It is a common fact that delivery of services at the tollbooths plays an important role at the toll plaza. Nowadays, at some toll plazas traffic delay has become a common problem because of failure of the timely services at the tollbooths. And, therefore to improve the working efficiency of the toll plaza it becomes necessary to use more than one tollbooth at a time simultaneously. Here, we have considered a system in which four tollbooths are connected in parallel which works simultaneously as shown in the figure 6.

Figure 6: *Toll Plaza*

VIII. Conclusion

Reliability, MTSF, availability and steady state availability have been evaluated for a four-unit repairable parallel redundant system by using Markov approach. The values for these measures have been obtained for particular values of repair rate, failure rate and operating time. The study reveals that the availability and mean time to system failure keeps on increasing with the increase of the repair rate while they decline sharply when failure rate increases. The reliability and availability of the system is declines with the increase of failure rate and operating time. On the other hand, it is also analyzed that these measures keep on increasing with the increase of repair rate. Hence, it is observed that system is more reliable and available to use by increasing the repair rate of the units.

References

[1] Chao M.T., James C.F. (1991). The Reliability of a Large Series System under Markov Structure. Advances in Applied Probability, 23(4):894–908.

[2] El-Damcese M.A., Temraz N.S. (2012). Analysis for a Parallel Repairable system with different Failure modes. Journal of Reliability and Statistical Studies, 5(1):95-106.

[3] Kalaiarasi S., Merceline Anita A., Geethanjalii R. (2017). Analysis of system reliability using Markov Technique. Global Journal of Pure and Applied Mathematics, 13(9):5265-5273

[4] Li J. (2016): Reliability calculation of a parallel redundant system with different failure rate & repair rate using Markov modeling. Journal of Reliability and Statistical Studies, 9(1):1-10.

[5] Nandal N., Bhardwaj R.K. (2020). Profit Analysis of a Three Unit Parallel-Cold Standby System Using Lindley Distribution. International Journal of Statistics and Reliability Engineering, 7(3):364-370.

[6] Rathi Puran, Malik S.C., Nandal N. (2021). Reliability Characteristics of a Parallel Cold Standby System of Four Units with Priority Proviso. International Journal of Agricultural and Statistical Sciences, 17(2):563-572.

[7] Reni S. M., Merceline Anita A., Chandrababu A., Gowtam Prakash S. (2014). Markov Models in System Reliability with Applications. International Journal of Innovative Research and Development, 3(11):328-336

[8] Saritha G., Tirumala Devi M., Sumathi Uma Maheswari T. (2020). The Reliability and Availability for Non-Repairable & Repairable Systems using Markov Modelling. International Journal of Engineering Research and Technology, .9(2):169-17

[9] Umamaheshwari K., Srilakshmi S. (2020). Markov model for a system with 'n' Elements with Human Error and Common Cause. Journal of Xi'an University of Architecture and Technology, 12(5):2611-2618

[10] Wang L., Jia X., Zhang J. (2013). Reliability Evaluation for Multi State Markov Repairable Systems with Redundant Dependencies. Quality Technology and Quantitative Management, 10(3):277-289

A Novel Extended Version of the Ailamujia Inverted Weibull Distribution

Idzhar A. Lakibul

•

Department of Mathematics and Statistics
Mindanao State University - Iligan Institute of Technology
Iligan City, Philippines
idzhar.lakibul@g.msuiit.edu.ph

Abstract

Statististical distributions with support on the set of non-negative real numbers are important in modelling and describing the behaviour of lifetime data. Ailamujia distribution is one of the non-negative continuous distribution that has an application in lifetime data. In this paper, a new three-parameter non-negative continuous distribution which is an extension of the Ailamujia Inverted Weibull distribution is introduced. This extended distribution is labeled as the Cubic Transmuted Ailamujia Inverted Weibull distribution. The proposed distribution is derived from the cubic transmuted family of distributions by specifying Ailamujia Inverted Weibull distribution as a baseline distribution. The probability density function of the proposed distribution is derived and some of its plots are presented. It can be observed that the proposed distribution can model the data which are exponentially and skewed unimodal right tailed data. In addition, survival and hazard functions of the proposed distribution are derived. It reveals that the hazard function of the proposed distribution can model both monotonic and non-monotonic decreasing failure rate behaviour of the data. Some properties of the proposed distribution such as its moments, moment generating function, mean, variance are derived. The Maximum Likelihood approach is used to estimate the proposed distribution parameters. Furthermore, parameter estimates as well as the performance of the proposed distribution is investigated by utilizing two sets of lifetime data. For point of comparison, this paper uses the following criteria: Akaike Information Criterion, Bayesian Information Criterion, Kolmogorov - Smirnov statistics, Anderson-Darling and Cramer-von Mises. Results show that for both sets of data, the proposed distribution produce better estimate as compared to the Quasi Suja and the Weibull-Lindley distributions. So, the proposed distribution consider as the best model for modelling the given two real datasets.

Keywords: Ailamujia distribution, Ailamujia Inverted Weibull distribution, Cubic transmuted family of distributions

1. Introduction

Ailamujia distribution is one of the newly lifetime distributions that has many application in engineering [5] . Due to its flexibility in modelling lifetime data then it becomes an area of interest for some researchers.

Different extensions were considered in literature. Uzma [11] introduced the weighted analogue of Ailamujia distribution and derived some of its properties. Other identified extensions such as the area biased distribution [3] , the size biased Ailamujia distribution [8] and the inverse analogue of Ailamujia distribution [1].

Moreover, Smadi [10] proposed an extended version of the Ailamujia distribution called as Ailamujia inverted Weibull distribution (AIWD) and various properties such as reliability function, hazard function, moments, moment generating function, order statistics, mean residual function and Shannon's entropy were studied.

In this paper, a novel extension of the Ailamujia Inverted Weibull distribution called Cubic Transmuted Ailamujia Inverted Weibull distribution (CTAIWD) is derived. Some of its properties such as moments, moment generating function, mean and variance are studied. Maximum Likelihood approach is also discussed to estimate the proposed distribution parameters. Two real data sets are analyzed to evaluate the performance of the proposed distribution.

The rest of paper is organized as follows: The Cubic Transmuted Ailamujia Inverted Weibull (CTAIW) distribution is introduced in section 2. In section 3, some statistical properties of CTAIW distribution are presented. Maximum Likelihood Estimation of the proposed distribution is discussed in section 4. In section 5, the application of proposed distribution is illustrated. Some concluding remarks is presented in section 6.

2. Cubic Transmuted Ailamujia Inverted Weibull distribution

This section presents the derivation of Cubic Transmuted Ailamujia Inverted Weibull (CTAIW) distribution. The cumulative distribution function of the Ailamujia Inverted Weibull distribution is given by

$$F(x,\theta,\alpha) = (1+2\theta x^{-\alpha})e^{-2\theta x^{-\alpha}}, x>0, \alpha>0, \theta>0 \tag{1}$$

with corresponding probability density function

$$f(x,\theta,\alpha) = 4\alpha\theta^2 x^{-2\alpha-1}e^{-2\theta x^{-\alpha}}.$$

Rahman [7] proposed a cubic transmuted family of distributions for extending any continuous distribution. The cumulative distribution function of the cubic transmuted family of distributions is given by

$$F(x) = (1-\lambda)G(x) + 3\lambda G^2(x) - 2\lambda G^3(x), x \in R \tag{2}$$

where $\lambda \in [-1,1]$. The cumulative distribution function of the Cubic Transmuted Ailamujia Inverted Weibull distribution is obtain by inserting (1) into (2) then, we have

$$F(x,\theta,\alpha,\lambda) = (1-\lambda)ze^{-2\theta x^{-\alpha}} + 3\lambda z^2 e^{-4\theta x^{-\alpha}} - 2\lambda z^3 e^{-6\theta x^{-\alpha}}, \tag{3}$$

where $z = 1+2\theta x^{-\alpha}$. The corresponding probability density function is

$$f(x,\theta,\alpha,\lambda) = 4\alpha\theta^2 x^{-2\alpha-1}e^{-2\theta x^{-\alpha}}(1-\lambda+6\lambda z e^{-2\theta x^{-\alpha}} - 6\lambda z^2 e^{-4\theta x^{-\alpha}}). \tag{4}$$

Note that if $\lambda = 0$ then Cubic Transmuted Ailamujia Inverted Weibull distribution reduces to Ailamujia Inverted Weibull distribution.

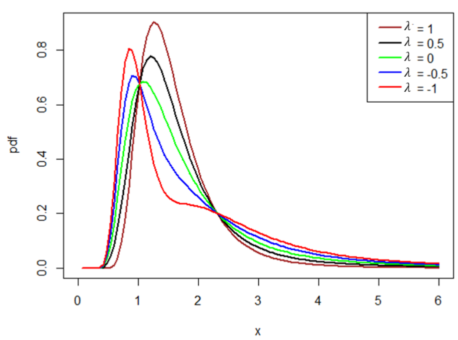


Figure 1: *pdf plots of CTAIW distribution for $\alpha = 1.5$, $\theta = 1.5$ and for varying values of λ*

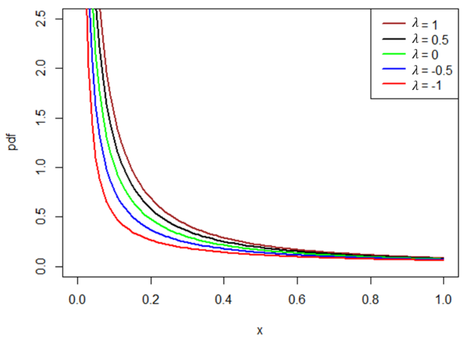


Figure 2: *pdf plots of CTAIW distribution for $\alpha = 0.2$, $\theta = 0.5$ and for varying values of λ*

Figures 1 and 2 present some possible shapes of the pdf of CTAIW distribution. It can be observed that the pdf of proposed distribution can generate a skewed unimodal and exponential shapes which are important in modelling lifetime data.

3. Statistical Properties

In this section, some statistical properties of the Cubic Transmuted Ailamujia Inverted Weibull distribution (CTAIW) such as moments, moment generating function, mean, variance, survival function and hazard function are derived.

3.1. Moments

Theorem 1. The *rth* moment of CTAIW distribution with density (4) is

$$\mu_r' = \lambda(2\theta)^{\frac{r}{\alpha}}[(1 - 2^{\frac{r}{\alpha}-1}3 + 3^{\frac{r}{\alpha}-1}2 - \frac{1}{\lambda})\Gamma(-\frac{r}{\alpha}+2) - (2^{\frac{r}{\alpha}-2}3 - 3^{\frac{r}{\alpha}-2}4)\Gamma(\frac{r}{\alpha}+3) + 3^{\frac{r}{\alpha}-3}2\Gamma(\frac{r}{\alpha}+4)] \tag{5}$$

The mean and variance are respectively, given by

$$\mu = \lambda(2\theta)^{\frac{1}{\alpha}}[(1 - 2^{\frac{1}{\alpha}3-1} + 3^{\frac{1}{\alpha}2-1} - \tfrac{1}{\lambda})\Gamma(-\tfrac{1}{\alpha}+2) - (2^{\frac{1}{\alpha}-2}3 - 3^{\frac{1}{\alpha}-2}4)\Gamma(\tfrac{1}{\alpha}+3) + 3^{\frac{1}{\alpha}2-3}\Gamma(\tfrac{1}{\alpha}+4)]$$

$$\begin{aligned}\sigma^2 =& \lambda(2\theta)^{\frac{2}{\alpha}}[(1 - 2^{\frac{2}{\alpha}-1}3 + 3^{\frac{2}{\alpha}-1}2 - \frac{1}{\lambda})\Gamma(-\frac{2}{\alpha}+2) - (2^{\frac{2}{\alpha}-2}3 - 3^{\frac{2}{\alpha}-2}4)\Gamma(\frac{2}{\alpha}+3) + 3^{\frac{2}{\alpha}-3}2\Gamma(\frac{2}{\alpha}+4)] \\ &- (\lambda(2\theta)^{\frac{1}{\alpha}}[(1 - 2^{\frac{1}{\alpha}-1}3 + 3^{\frac{1}{\alpha}-1}2 - \frac{1}{\lambda})\Gamma(-\frac{1}{\alpha}+2) - (2^{\frac{1}{\alpha}-2}3 - 3^{\frac{1}{\alpha}-2}4)\Gamma(\frac{1}{\alpha}+3) \\ &+ 3^{\frac{1}{\alpha}-3}2\Gamma(\frac{1}{\alpha}+4)])^2.\end{aligned}$$

Proof. The rth moment is defined by

$$\begin{aligned}\mu_r' &= E[X^r] \\ &= \int_{-\infty}^{\infty} x^r f(x)dx \\ &= \int_0^{\infty} x^r 4\alpha\theta^2 x^{-2\alpha-1} e^{-2\theta x^{-\alpha}}(1 - \lambda + 6\lambda(1 + 2\theta x^{-\alpha})e^{-2\theta x^{-\alpha}} - 6\lambda(1 + 2\theta x^{-\alpha})^2 e^{-4\theta x^{-\alpha}})dx\end{aligned}$$

$$= 4(1-\lambda)\theta^2\alpha\int_0^\infty x^r x^{-2\alpha-1}e^{-2\theta x^{-\alpha}}dx + 24\lambda\theta^2\alpha\int_0^\infty x^r x^{-2\alpha-1}(1+2\theta x^{-\alpha})e^{-4\theta x^{-\alpha}}dx$$
$$- 24\lambda\theta^2\alpha\int_0^\infty x^r x^{-2\alpha-1}(1+2\theta x^{-\alpha})^2 e^{-6\theta x^{-\alpha}}dx$$
$$= (1-\lambda)[-(2\theta)^{\frac{r}{\alpha}}\Gamma(-\frac{r}{\alpha}+2)] - 6\lambda[\frac{(4\theta)^{\frac{r}{\alpha}}}{4}(\Gamma(-\frac{r}{\alpha}+2)+\frac{1}{2}\Gamma(-\frac{r}{\alpha}+3))]$$
$$+ 6\lambda[-\frac{1}{9}(6\theta)^{\frac{r}{\alpha}}(\Gamma(-\frac{r}{\alpha}+2)+\frac{2}{3}\Gamma(-\frac{r}{\alpha}+3)+\frac{1}{9}\Gamma(-\frac{r}{\alpha}+4))]$$
$$= \lambda(2\theta)^{\frac{1}{\alpha}}[(1-2^{\frac{1}{\alpha}3-1}+3^{\frac{1}{\alpha}2-1}-\frac{1}{\lambda})\Gamma(-\frac{1}{\alpha}+2)-(2^{\frac{1}{\alpha}-2}3-3^{\frac{1}{\alpha}-2}4)\Gamma(\frac{1}{\alpha}+3)+3^{\frac{1}{\alpha}2-3}\Gamma(\frac{1}{\alpha}+4)]$$

The mean of CTAIW distribution is obtained by using $r = 1$ in (5) and is

$$\mu = \lambda(2\theta)^{\frac{1}{\alpha}}[(1-2^{\frac{1}{\alpha}3-1}+3^{\frac{1}{\alpha}2-1}-\tfrac{1}{\lambda})\Gamma(-\tfrac{1}{\alpha}+2)-(2^{\frac{1}{\alpha}-2}3-3^{\frac{1}{\alpha}-2}4)\Gamma(\tfrac{1}{\alpha}+3)+3^{\frac{1}{\alpha}2-3}\Gamma(\tfrac{1}{\alpha}+4)]$$

The $2nd$ raw moment of CTAIW distribution is obtained by using $r = 2$ in (5) and is

$$\mu_2' = \lambda(2\theta)^{\frac{2}{\alpha}}[(1-2^{\frac{2}{\alpha}-1}3+3^{\frac{2}{\alpha}-1}2-\tfrac{1}{\lambda})\Gamma(-\tfrac{2}{\alpha}+2)-(2^{\frac{2}{\alpha}-2}3-3^{\frac{2}{\alpha}-2}4)\Gamma(\tfrac{2}{\alpha}+3)+3^{\frac{2}{\alpha}-3}2\Gamma(\tfrac{2}{\alpha}+4)]$$

The variance σ^2 of CTAIW distribution obtained as

$$\sigma^2 = \mu_2' - (\mu_1')^2$$
$$= \lambda(2\theta)^{\frac{2}{\alpha}}[(1-2^{\frac{2}{\alpha}-1}3+3^{\frac{2}{\alpha}-1}2-\frac{1}{\lambda})\Gamma(-\frac{2}{\alpha}+2)-(2^{\frac{2}{\alpha}-2}3-3^{\frac{2}{\alpha}-2}4)\Gamma(\frac{2}{\alpha}+3)+3^{\frac{2}{\alpha}-3}2\Gamma(\frac{2}{\alpha}+4)]$$
$$-(\lambda(2\theta)^{\frac{1}{\alpha}}[(1-2^{\frac{1}{\alpha}-1}3+3^{\frac{1}{\alpha}-1}2-\frac{1}{\lambda})\Gamma(-\frac{1}{\alpha}+2)-(2^{\frac{1}{\alpha}-2}3-3^{\frac{1}{\alpha}-2}4)\Gamma(\frac{1}{\alpha}+3)$$
$$+3^{\frac{1}{\alpha}-3}2\Gamma(\frac{1}{\alpha}+4)])^2.$$

■

3.2. Moment Generating Function

Theorem 2. Let X follows the CTAIW distribution then the moment generating function $M_X(t)$ is given by

$$M_x(t) = \sum_{r=0}^{\infty}\frac{t^r\lambda(2\theta)^{\frac{r}{\alpha}}}{r!}[a\Gamma(-\frac{r}{\alpha}+2)-b\Gamma(\frac{r}{\alpha}+3)+c\Gamma(\frac{r}{\alpha}+4)]$$

where $a = 1-2^{\frac{r}{\alpha}-1}3+3^{\frac{r}{\alpha}-1}2-\frac{1}{\lambda}$, $b = 2^{\frac{r}{\alpha}-2}3-3^{\frac{r}{\alpha}-2}4$, $c = 3^{\frac{r}{\alpha}-3}2$ and $t \in R$.

Proof. By definition of moment generating function and equation (5), we have

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_0^\infty e^{tx}f(x,\theta,\alpha,\lambda)dx.$$

Recall that $e^{tx} = \sum_{r=0}^{\infty}\frac{t^r}{r!}x^r$ then we have

$$M_X(t) = \int_0^\infty\sum_{r=0}^{\infty}\frac{t^r}{r!}x^r f(x,\theta,\alpha,\lambda)dx = \sum_{r=0}^{\infty}\frac{t^r}{r!}x^r\int_0^\infty f(x,\theta,\alpha,\lambda)dx = \sum_{r=0}^{\infty}\frac{t^r}{r!}\mu_r'.$$

Thus,

$$M_x(t) = \sum_{r=0}^{\infty}\frac{t^r\lambda(2\theta)^{\frac{r}{\alpha}}}{r!}[a\Gamma(-\frac{r}{\alpha}+2)-b\Gamma(\frac{r}{\alpha}+3)+c\Gamma(\frac{r}{\alpha}+4)]$$

where $a = 1-2^{\frac{r}{\alpha}-1}3+3^{\frac{r}{\alpha}-1}2-\frac{1}{\lambda}$, $b = 2^{\frac{r}{\alpha}-2}3-3^{\frac{r}{\alpha}-2}4$ and $c = 3^{\frac{r}{\alpha}-3}2$. ■

3.3. Reliability Analysis

Let X be a random variable with cdf (3) and pdf (4) then the survival $S(x)$ and hazard $h(x)$ functions of CTAIW distribution are respectively, given as follows:

$$S(x,\theta,\alpha,\lambda) = 1-(1-\lambda)ze^{-2\theta x^{-\alpha}} - 3\lambda z^2 e^{-4\theta x^{-\alpha}} + 2\lambda z^3 e^{-6\theta x^{-\alpha}};$$

$$h(x,\theta,\alpha,\lambda) = \frac{4\alpha\theta^2 x^{-2\alpha-1} e^{-2\theta x^{-\alpha}}(1-\lambda+6\lambda z e^{-2\theta x^{-\alpha}} - 6\lambda z^2 e^{-4\theta x^{-\alpha}})}{1-(1-\lambda)ze^{-2\theta x^{-\alpha}} - 3\lambda z^2 e^{-4\theta x^{-\alpha}} + 2\lambda z^3 e^{-6\theta x^{-\alpha}}},$$

where $z = 1+2\theta x^{-\alpha}$, $x > 0$,$\alpha, \theta > 0$ and $\lambda \in [-1,1]$.

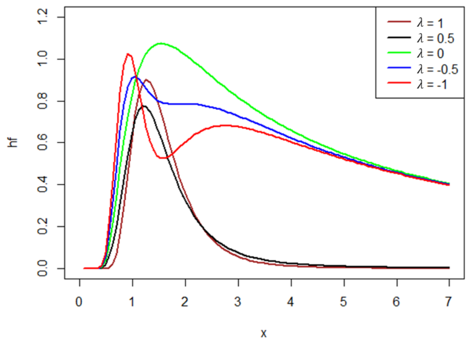


Figure 3: *hf plots of CTAIW distribution for* $\alpha = 1.5$, $\theta = 1.5$ *and for varying values of* λ

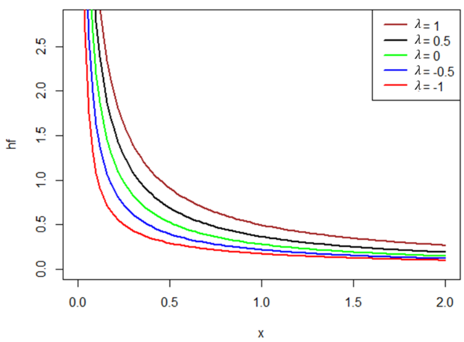


Figure 4: *hf plots of CTAIW distribution for* $\alpha = 0.2$, $\theta = 0.5$ *and for varying values of* λ

Figures 3 and 4 present some possible shapes of hazard rate function of CTAIW distribution. It reveals that the hazard rate function of the proposed distribution can model both monotonic and non-monotonic decreasing failure rate behavior of the data which are common in the reliability data.

4. Maximum Likelihood Estimation

In this section, the maximum likelihood approach is used to estimate the CTAIW distribution parameters.

Let X_1, X_2,..., X_n be a random sample of size n from CTAIW distribution, then the likelihood function is

$$\mathbf{L} = \prod_{i=1}^{n}[4\alpha\theta^2 x_i^{-2\alpha-1} e^{-2\theta x_i^{-\alpha}}(1-\lambda+6\lambda z e^{-2\theta x_i^{-\alpha}} - 6\lambda z^2 e^{-4\theta x_i^{-\alpha}})]$$

where $z = 1 + 2\theta x_i^{-\alpha}$. The log-likelihood function is

$$l = \sum_{i=1}^{n} log[4\alpha\theta^2 x_i^{-2\alpha-1} e^{-2\theta x_i^{-\alpha}} (1 - \lambda + 6\lambda z e^{-2\theta x_i^{-\alpha}} - 6\lambda z^2 e^{-4\theta x_i^{-\alpha}})]. \tag{6}$$

Taking the deivative of (6) with respect to the parameters α, θ and λ then, we have the following:

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} - 2\sum_{i=1}^{n} log(x_i) + 2\theta \sum_{i=1}^{n} x_i^{-\alpha} log(x_i) + 24\lambda\theta^2 \sum_{i=1}^{n} \frac{log(x_i) x^{-2\alpha} e^{-2\theta x_i^{-\alpha}} (1 - 2ze^{-2\theta x_i^{-\alpha}})}{1 - \lambda - 6\lambda z e^{-2\theta x_i^{-\alpha}} - 6\lambda z^2 e^{-4\theta x_i^{-\alpha}}} \tag{7}$$

$$\frac{\partial l}{\partial \theta} = \frac{2n}{\theta} - 2\sum_{i=1}^{n} x_i^{-\alpha} - 24\lambda\theta \sum_{i=1}^{n} \frac{x_i^{-2\alpha} e^{-2\theta x_i^{-\alpha}} (1 - 2ze^{-2\theta x_i^{-\alpha}})}{1 - \lambda - 6\lambda z e^{-2\theta x_i^{-\alpha}} - 6\lambda z^2 e^{-4\theta x_i^{-\alpha}}} \tag{8}$$

$$\frac{\partial l}{\partial \lambda} = -\sum_{i=1}^{n} \frac{1 - 6ze^{-2\theta x_i^{-\alpha}} + 6z^2 e^{-4\theta x_i^{-\alpha}}}{1 - \lambda - 6\lambda z e^{-2\theta x_i^{-\alpha}} - 6\lambda z^2 e^{-4\theta x_i^{-\alpha}}} \tag{9}$$

Equating equations (7), (8) and (9) to 0, respectively, one can get numerical maximum likelihood estimates of the CTAIW distribution parameters.

5. Application

In this section, the proposed distribution is applied to two real data sets and compared with the following distributions:

- Quasi Suja (QS) distribution [9]

$$f(x, \alpha, \theta) = \frac{\theta^4}{\alpha\theta^3 + 24}(\alpha + \theta x^4)e^{-\theta x}, x > 0, \theta > 0, \alpha > 0.$$

- Weibull-Lindley (WL) distribution [2]

$$f(x, \alpha, \beta) = \alpha(\beta x^{\beta-1}(1 + x)e^{x^\beta})e^{-\alpha((1+x)e^{x^\beta}-1)}, x > 0, \alpha, \beta > 0.$$

A data analysis is performed using a package "fitdistrplus" in R. Moreover, AKaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Kolmogorov-Smirnov (K-S), Anderson-Darling (A) and Cramer-von Mises (W*) statistics are used for comparison in this analysis. In addition, the two sets of data used in the analysis are given as follows:

Data Set 1. This data set is from Murthy [6]. It is a set of failure times of 50 electronic components. The data set is given as follows: 0.036, 0.058, 0.061, 0.074, 0.078, 0.086, 0.102, 1.103, 0.114, 0.116, 0.148, 0.183, 0.192, 0.254, 0.262, 0.379, 0.381, 0.538, 0.570, 0.574, 0.590, 0.618, 0.645, 0.961, 1.228, 1.600, 2.006, 2.054, 2.804, 3.058, 3.076, 3.147, 3.625, 3.704, 3.931, 4.073, 4.393, 4.534, 4.893, 6.274, 6.816, 7.896, 7.904, 8.022, 9.337, 10.940, 11.020, 13.880, 14.730 and 15.080.

Data Set 2. This dataset is from Khan [4] and it is an ICU data set which assess the intensive care unit (ICU) patients agitation-sedation (A-S) status. The data set is given as follows: 9, 3, 27, 8, 4, 3, 4, 3, 23, 3, 3, 4, 28, 18, 19, 6, 3, 26, 3, 12, 6, 9, 43, 4, 4, 3, 5, 12, 4, 36, 6, 8, 6, 5, 3, 3 and 33.

Tables 1 and 3 list the MLEs of the CTAIW, QS and WL distributions fitted to first and second sets of data. Tables 2 and 4 indicate that the proposed distribution has a lower values of the AIC, BIC, K-S, A and W* compared to the QS and WL distributions for both sets of data. So, the propsed distribution fit well the said datasets. Furthermore, the estimated pdf of the fitted models to both sets of data are presented in figure 5 and 6, respectively.

Table 1: *MLEs of the fitted models for first set of data.*

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$
CTAIW	0.5116727		0.6782907	-0.9999998
WL	0.07258343	0.12084431		
QS	211.8090902		0.6493223	

Table 2: *Numerical values of AIC, BIC, K-S, A and W* of the fitted models for first set of data.*

Distribution	AIC	BIC	$K-S$	A	$W*$
CTAIW	209.0186	214.7547	0.1531139	1.3078260	0.2494361
WL	228.7096	232.5337	0.2500551	3.3786131	0.5503281
QS	213.2069	217.031	0.2179699	3.6483026	0.4827523

Table 3: *MLEs of the fitted models for second set of data.*

Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$
CTAIW	1.1988356		8.5119091	-0.9999997
WL	0.02157018	0.11712539		
QS	20023.91		0.1678021	

Table 4: *Numerical values of AIC, BIC, K-S, A and W* of the fitted models for second set of data.*

Distribution	AIC	BIC	$K-S$	A	$W*$
CTAIW	232.7614	237.5941	0.1724946	1.4724752	0.2293943
WL	256.3175	259.5393	0.2192403	2.3094837	0.4026443
QS	253.2686	256.4904	0.3155350	2.5730059	0.4309077

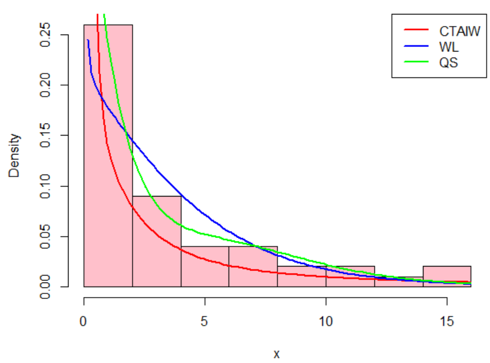


Figure 5: *Estimated pdf of the fitted models for the first set of data.*

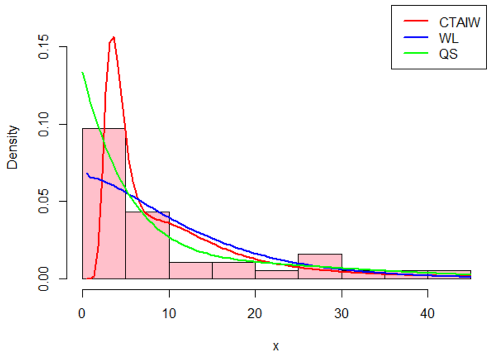


Figure 6: *Estimated pdf of the fitted models for the second set of data.*

6. Concluding Remarks

In this study, an extended version of the Ailamujia Inverted Weibull distribution called Cubic Transmuted Ailamujia Inverted Weibull distribution has been introduced. Some properties of the proposed distribution such as moment, moment generating function, mean and variance were derived. The maximum likelihood approach was used to estimate the proposed distribution parameters. Two real datasets were used to examine the flexibility of proposed distribution and compared to Quasi Suja and Weibull - Lindley distributions. It was found that the proposed distribution fit well the given data sets compared to said distributions.

References

[1] Aijaz, A., Ahmad, A. and Tripathi, R. (2020). Inverse analogue of Ailamujia distribution with statistical properties and applications. *Asian Research Journal of Mathematics*, 16:36–46.
[2] Chacko, V. M., S, D. K, Thomas, B. and C, R. (2018) Weibull-Lindley Distribution: A Bathtub Shaped Failure Rate Model. *Reliability: Theory and Application*, 13:9–20.
[3] Jayakumar B. and Elangovan, R. A. (2019). New generalization of Ailamujia distribution with applications in bladder cancer data. *IJSRMSS*, 6:61–68.
[4] Khan, M. S., King, R. and Hudson, I. L. (2017). Transmuted generalized exponential distribution: A generalization of the exponential distribution with applications to survival data. *Communication Statistics : Simulation and Computation*, 46:4377–4398.
[5] Kotz, S. L. and Pensky, M. The stress strength model and its generalizations, Theory and Applications, World Scientific, Singapore, 2003.
[6] Murthy, D. N. P., Xie, M., Jiang, R. Weibull Models, John Wiley and sons Inc, Hoboken, 2004.
[7] Rahman, M. M., Al-Zahrani, B. and Shahbaz, S. H. (2019). Cubic Transmuted Uniform Distribution: An Alternative to Beta and Kumaraswamy Distributions. *European Journal of Pure and Applied Mathematics*, 12:1106–1121.
[8] Rather, A., Subramanian, C., Shafi, S., Malik, K. A., Ahmad, P. J., Para, B. A. and Jan, T. A. (2018). New size biased distribution with Applications in Engineering and Medical Science. *IJSRMSS*, 5:75–85.
[9] Shanker, R., Upadhyay, R. and Shukla, K. K. (2022). A Quasi Suja Distribution. *Reliablilty: Theory and Application*,17:162–178.
[10] Smadi, M. M. and Ansari, S. I. (2022). Ailamujia Inverted Weibull Distribution with Application to Lifetime Data. *Pakistan Journal of Statistics*, 38:341–358.
[11] Uzma, J., Kawsar, F. and Ahmad, S. P. (2017). On weighted Ailamujia distribution and its applications to lifetime data. *Journal of Statistics Applications and Probability an International Journal*, 6:619–633.

Reliability Measures of a 2-out-of-3: G System with Priority and Failure of Service Facility during Repair

Anuradha
•
Department of Statistics, Maharshi Dayanand University Rohtak, India anudixit083@gmail.com

S.C. Malik
•
Department of Statistics, Maharshi Dayanand University Rohtak, India sc_malik@rediffmail.com

Abstract

Aim. The objective of this paper is to describe a particular case of the k-out-of-n: G system for k=2 and n=3 with different repair policies and to discuss the application of the proposed in a toxic waste incinerator. Methods. The system has all the three units identical in nature. The system model is developed using semi-Markov process and regenerative point technique. The preventive maintenance and repair activities of the units are carried out immediately by a single service facility whenever desires. The service facility is subjected to failure during repair of the units while it does preventive maintenance of the units without any problem. The failed service facility undergoes for treatment to restore its efficiency to perform the remaining jobs with full capacity. The provision of priority to preventive maintenance of the units has been made over the repair in order to avoid the earlier failure of the system. Findings. The measures that can affect and enhance the performance of the system have been discussed for arbitrary values of the rates which follow some arbitrary distributions including the negative exponential. The system is analysed in steady state and the graphs have been drawn to see the effect of different transition rates such as failure rate, preventive maintenance rate, treatment rate, and repair rate of the units on reliability measures and the profit. The study reveals that there is a decline in these measures with the increase of the rate by which unit undergoes for preventive maintenance, failure rates of the units and service facility. However, the values of reliability measures MTSF, availability and profit function keep on increasing with the increase of treatment rate, repair rate of the unit and preventive maintenance completion rate. The profit increases if the rate with which a unit completes its preventive maintenance. Hence, implementing the preventive maintenance repair policy for a 2-out-of-3 system is beneficial as it increases the availability and hence the profit of the system.

Keywords: 2-out-of-3 System, Preventive Maintenance, Treatment Rate, Priority, Failure of Service Facility and Reliability Measures

1. Introduction

Now a day, complex systems are used in almost all the areas of science and technology specially in the field of industries. Complex systems are made of multiple dependent and independent components. This kind of formulation may affect the efficient and accurate evaluation of reliability measures. So, for the accurate analysis of the system performance, it is obligatory to understand the nature of components, their failure rates, preventive maintenance rates and their interactions. The k -out-of-n: G system is one such type of complex system defined as the system which has 'n' identical units, all are working initially and at least k are required to be good to make the system work. k-out-of-n system is one of the most useful partial redundancy types in complex systems, which is often used in various areas including software and hardware engineering for the purpose of providing a proper level of redundancy during the operation of a system to increase the availability and hence the profit. There are many systems in our day-to-day life which are based upon k-out-of-n configuration for example multi engine system of an airplane, multi pump system in a hydraulic control system, networking and communication systems, toxic waste incinerator etc.

Several studies have been conducted so far for arbitrary k-out-of-n systems by assuming that system can perform nonstop work without requiring any maintenance. But in real life this assumption seems hard to believe as continuous operation deteriorates the performance and hence reliability of the system. To avoid such problems preventive maintenance can be considered as an alternative solution conducted after a specific operation time. Preventive maintenance also helps in slow down the deterioration rate of the process. Malik and Bhardwaj [1] have analysed 2-out-of-3 redundant system with general repair and waiting time distribution. They have also studied the same system with feasibility to repair. Malik and Singh [2] have studied a 2-out-of-3 redundant system with a single service facility for inspection and repair of the units by using the concept of degradation of unit after repair. Kishan and Kumar [3] studied a parallel system with preventive maintenance. Jain et al. [4] have analysed a Load Sharing m-out-of-n: G System with Non -Identical Components Subject to Common Cause Failure. Further, the concept of priority in repair policies has also been suggested by the researchers to make the system more profitable. Yang et al. [5] have discussed the reliability analysis of load-sharing k-out-of-n system considering component degradation. Poonia et al. [6] has performed the cost analysis of a repairable warm standby k-out-of-n: G and 2-out-of-4: G systems in series configuration under catastrophic failure using copula repair. Anuradha et al. [7] have analysed the profit of a 1-out-of-2: G system with the concept of server failure and priority to repair. Singh et al. [8] have done the performance assessment of the complex repairable system with n-identical units under (k-out-of-n: G) scheme and copula linguistic repair approach. Abdullahi et. al [9] have performed the cost analysis of 2-out-of-4 system connected to two units parallel supporting device for operation However, the idea of priority to preventive maintenance over repair along with the failure of service facility (all the three repair policies together) has not been introduced while analysing system reliability models of three or more identical units.

So, here a particular case of k-out-of-n: G system for k=2 and n=3 has been considered. Initially all the three units are operative out of which two units are required to be good for the proper functioning of the system. The system undergoes for preventive maintenance with an arbitrary rate in order to restore the efficiency of the working units. A single service facility performs the preventive maintenance and repair of the units whenever desires. The service facility may fail and therefore, undergoes for treatment to restore the efficiency to carry out the remaining repair activities. Preventive maintenance of the units is kept as priority over repair in order to avoid the earlier failure of the system. The well-known semi-Markov process and regenerative point technique are used to develop the system model under certain assumptions. The measures that can affect and enhance the performance of the system have been discussed for arbitrary values of the rates which follow some arbitrary distributions including the negative exponential. The system is analysed in steady state and the graphs have been drawn to see the effect of different transition rates such as failure rate, preventive maintenance rate, treatment rate, and repair rate of the units on reliability measures and the profit. Thus, the focus of the present paper is to analyse reliability measures and profit of a 2-out-of-3 system.

2. System description and Notations

2.1) Notations

Table 1: *Symbol Description*

Symbol	Description
O	Unit is operative
λ	Failure Rate of the Unit
μ	Failure Rate of the Server
WO	Unit is Waiting for Operation
p(t)/P(t)	pdf/cdf of the preventive maintenance time of the unit
$p_m(t)/P_m(t)$	pdf/cdf of the preventive maintenance completion time of the unit
$r_s(t)/R_s(t)$	pdf/cdf of the treatment time of the server
r(t)/R(t)	pdf/cdf of the repair time of the unit
Fur/FUR	Unit is failed and under repair/under repair continuously from previous state
FWr /FWR	Unit is failed and waiting for repair/waiting for repair continuously from previous state
UPm /UPM	Unit is failed and under preventive maintenance/under preventive maintenance continuously from previous state
SFut	Server failed under treatment
$q_{ij}(t)/Q_{ij}(t)$	pdf/cdf of first passage time from regenerative state S_i to a regenerative state S_i or to a failed state S_j without visiting anyother regenerative state in (0, t]
$q_{ij.k}(t)/Q_{ij.k}(t)$	pdf/cdf of direct transition time from regenerative state S_i to a regenerative state S_j or to a failed state S_j visiting state S_k once in (0, t]
$M_i(t)$	Probability that the system up initially in state S_i and is up at time t without visiting any regenerative state
$W_i(t)$	Probability that the server is busy in the state S_i up to time 't' without making any transition to any other regenerative state or returning to the same state via one or more non-regenerative states
μ_i	The mean sojourn time in state S_i given by: $\mu_i=E(t)=\int_0^\infty P(T>t)dt=\sum_j m_{ij}$, where T is the time to system failure
m_{ij}	Contribution to mean sojourn time(μ_i) in state S_i when the system directly transits to state S_j so that $\mu_i=\sum_j m_{ij}$ and $m_{ij}=\int tdQ_{ij}(t)dt=-q_{ij}^{*'}(0)$
Ⓢ/©	Symbols for Laplace Stieltjes convolution/Laplace convolution
*/**	Symbols for Laplace transformation/Laplace Stieltjes transformation

2.2) The state transition diagram of the system model

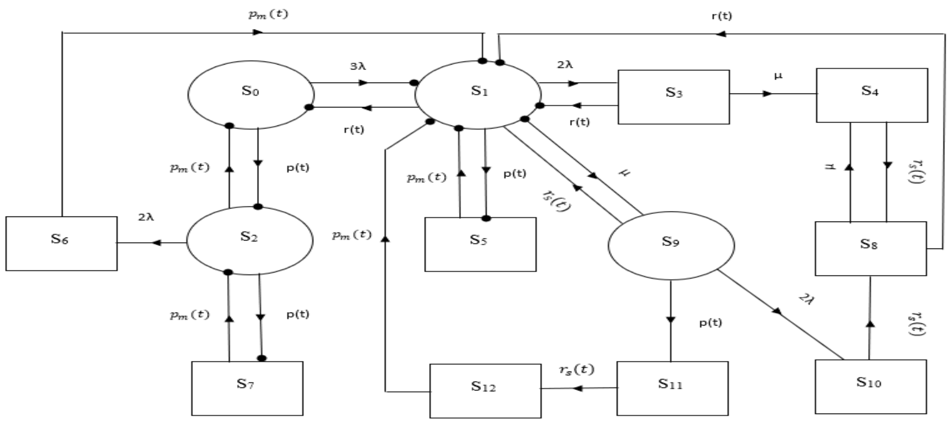


Figure 1: *State Transition Diagram*

3. State Transition Probabilities and Mean Sojourn Time

Table 1: *Transition State Description*

Transition state	Description
S_0(O,O,O)	All three units are in operation
S_1 (Fur ,O,O)	Two units are in operation and one unit is under repair
S_2 (O ,O ,UPm)	Two units are in operation and one unit is under preventive maintenance
S_3 (WO ,FUR ,Fwr)	One unit is waiting for operation. One unit is waiting for repair and one failed unit is under repair continuously from previous state.
S_4(WO ,Fwr ,FWR ,SFut)	One unit is waiting for operation. One unit is waiting for repair and third unit is waiting for repair from previous state. Server is under treatment.
S_5(WO ,Fwr , UPm)	One unit is waiting for operation, one is waiting for repair and one unit is under preventive maintenance.
S_6(WO ,Fwr ,UPM)	One unit is waiting for operation, one is waiting for repair and one unit is under preventive maintenance from previous state.
S_7 (WO, UPM, WPm)	One unit is waiting for operation, one unit is under preventive maintenance from its previous state and one unit is waiting for preventive maintenance for first time.
S_8 (WO, Fur, FWR)	One unit is waiting for operation, failed unit is under repair and one unit is waiting for repair from previous state.
S_9 (O,O, Fwr, SFut)	Two units are in operation. Failed unit is waiting for repair and failed server is under rectification.
S_{10}(WO, FWR ,Fwr ,SFUT)	Failed unit is waiting for repair from previous state another one is waiting for repair for the first time. Server is under rectification from previous state and one unit is waiting for operation.
S_{11}(WO ,WPm ,FWR,SFUT)	Failed unit is waiting for repair from previous state. One unit is waiting for preventive maintenance. Server is under treatment.
S_{12} (WO, FWR, UPm)	One unit is waiting for operation, failed unit is waiting for repair and one unit is under preventive maintenance.

4. Reliability Measures

4.1) Transition Probabilities

Using the formula given below transition probabilities from any state i to j are obtained as follows:

$p_{ij}= Q_{ij}(\infty)= \int_0^\infty q_{ij}(t)dt$

$p_{01}=\frac{\lambda}{\lambda+\alpha}$, $p_{02}=\frac{\alpha}{\lambda+\alpha}$,$p_{10}=\frac{b}{2(\lambda+\alpha)+\mu+b}$, $p_{15}=\frac{2\alpha}{2(\lambda+\alpha)+\mu+b}$, $p_{13}=\frac{2\lambda}{2(\lambda+\alpha)+\mu+b}$, $p_{19}=\frac{\mu}{2(\lambda+\alpha)+\mu+b}$

$p_{20}=\frac{\eta}{2(\lambda+\alpha)+\eta}$, $p_{27}=\frac{2\alpha}{2(\lambda+\alpha)+\eta}$, $p_{26}=\frac{2\lambda}{2(\lambda+\alpha)+\eta}$,$p_{31}=p_{81}=\frac{b}{\mu+b}$, $p_{3,8}=p_{84}=\frac{\mu}{\mu+b}$

$p_{4,8}=p_{51}=p_{61}= p_{72}= p_{10,12}=p_{11.12}=p_{12,1}=1$,$p_{91}=\frac{a}{2(\lambda+\alpha)+a}$, $p_{9,10}=\frac{2\lambda}{2(\lambda+\alpha)+a}$, $p_{9,11}=\frac{2\alpha}{2(\lambda+\alpha)+a}$

It is verified that:

$p_{01}+p_{02}=1$,$p_{10}+p_{15}+p_{19}+p_{13}=1$,$p_{20}+p_{22.7}+p_{21.6}=1$,$p_{31}+p_{3,8}=1$,$p_{81}+p_{84}=1$,$p_{91}+p_{9,10}+p_{9,11}=1$,$p_{4,8}=p_{51}=p_{61}=p_{72}=p_{10,12}=p_{11.12}=p_{12,1}=1$

4.2) Mean Sojourn Time

μ_i in state S_i are given as:

$\mu_0=\frac{1}{3(\lambda+\alpha)}$, $\mu_1=\frac{1}{(2\lambda+2\alpha+\mu+b)}$, $\mu_2=\frac{1}{(2\lambda+2\alpha+\eta)}$, $\mu_3=\mu_8=\frac{1}{\mu+b}$, $\mu_9=\frac{1}{2\lambda+2\alpha+a}$, $\mu_4=\mu_{10}=\mu_{11}=\frac{1}{a}$, $\mu_5=\mu_6=\mu_7=\mu_{12}=\frac{1}{\eta}$

$\mu_1'=m_{10}+m_{15}+m_{19}+m_{11.3}+m_{11.3(4,8)^n}$, $\mu_2'=m_{20}+m_{22.7}+m_{21.6}$, $\mu_9'=m_{91}+m_{91.(11,12)}+m_{91.(10,8)}+m_{91.10(8,4)^n}$

4.3) Mean Time to System Failure

Let $\Phi_i(t)$- cdf of first passage time from regenerative state S_i to a failed state.

$\Phi_0(t)=Q_{01}(t)Ⓢ\Phi_1(t)+Q_{02}(t)Ⓢ\Phi_2(t)$ (1)

$\Phi_1(t)=Q_{10}(t)Ⓢ\Phi_0(t)+Q_{19}(t)Ⓢ\Phi_9(t)+Q_{15}(t)+Q_{13}(t)$ (2)

$\Phi_2(t)=Q_{20}(t)Ⓢ\Phi_0(t)+Q_{26}(t)+Q_{27}(t)$ (3)

$\Phi_9(t)=Q_{91}(t)Ⓢ\Phi_1(t)+Q_{9,10}(t)+Q_{9,11}(t)$ (4)

Solving the above equations for $\Phi_0^{**}(s)$ by taking Laplace Stieltjes Transformation, We get,

$R^*(s)=\frac{1-\Phi_0^{**}(s)}{s}$

MTSF= $\lim_{n\to\infty}\frac{1-\Phi_0^{**}(s)}{s}=\frac{N_1}{D_1}$

Where $N_1=(1-p_{19}p_{91})(\mu_0+p_{02}\mu_2)+p_{01}(\mu_1+p_{19}\mu_9)$ and $D_1=(1-p_{02}p_{20})(1-p_{19}p_{91})-p_{01}p_{10}$

4.4) Long Run Availability of the System

Define, $A_i(t)$- Probability that the system is available at any instant time t given that it has entered regenerative state S_I at time t =0

$M_i(t)$- Probability that the system is up initially as well as at time t in state S_i without making any transition to regenerative state

The following expressions are obtained:

$A_0(t)=M_0(t)+q_{01}(t)©A_1(t)+q_{02}(t)©A_2(t)$ (6)

$A_1(t)=M_1(t)+q_{10}(t)©A_0(t)+(q_{11.3}+q_{11.3(4,8)^n})©A_1(t)+q_{15}(t)©A_5(t)+q_{19}©A_9(t)$ (7)

$A_2(t)=M_2(t)+q_{20}(t)©A_0(t)+q_{22.7}©A_2(t)+q_{21.6}©A_1(t)$ (8)

$A_9(t)=M_9(t)+(q_{91}(t)+q_{91.(11,12)}+q_{91.10(8,4)^n}+q_{91.(10,8)})©A_1(t)$ (9)

$A_5(t)=q_{51}©A_1(t)$ (10)

Where, $M_0(t)=e^{-3\lambda t}\overline{P(t)}$, $M_1(t)=e^{-(2\lambda+\mu)t}\overline{P(t)R(t)}$, $M_2(t)=e^{-2\lambda t}\overline{P_m(t)P(t)}$, $M_9(t)=e^{-2\lambda t}\overline{P(t)R_s(t)}$

Solving for $A_0^{**}(s)$ using LST

We have,

$A_0(\infty)=\lim_{n\to\infty}sA_0^{**}(s)=\frac{N_2}{D_2}$ (11)

$N_2=p_{10}(1-p_{27})\mu_0+\mu_2p_{02}p_{10}+(p_{01}p_{20}+p_{26})(p_{19}\mu_9+\mu_1)$

$D_2=p_{10}(1-p_{27})\mu_0+\mu'_2p_{02}p_{10}+(p_{01}p_{20}+p_{26})(\mu_1'+\mu_5p_{15}+p_{19}\mu_9')$

4.4) Busy Period of the Server Due to Repair

$BP_i^r(t)$- Probability that the server is busy in repairing at an instant 't'.

The recursive expressions for $BP_i^r(t)$ are given below:

$BP_0^r(t)=q_{01}(t)©BP_1^r(t)+q_{02}(t)©BP_2^r(t)$ (12)

$BP_1^r(t)=W_1^r(t)+q_{10}(t)©BP_0^r(t)+(q_{11.3}+q_{11.3(4,8)^n})©BP_1^r(t)+q_{15}©BP_5^r(t)+q_{19}©BP_9^r(t)$ (13)

$BP_2^r(t)=q_{20}(t)©BP_0^r(t)+q_{22.7}©BP_2^r(t)+q_{21.6}©BP_1^r(t)$ (14)

$BP_9^r(t)=(q_{91}(t)+q_{91.(11,12)}+q_{91.10(8,4)^n}+q_{91.(10,8)})©BP_1^r(t)$ (15)

$BP_5^r(t)=q_{51}©BP_1^r(t)$ (16)

$W_1^r(t)=e^{-(2\lambda+\mu)t}\overline{P(t)R(t)}$

Solving for $BP_0^{r**}(s)$, We have:

$BP_0^r(\infty)=\lim_{n\to\infty}sBP_0^{r**}(s)=\frac{N_5}{D_2}$, $N_5=W_1^*(0)(p_{20}p_{01}+p_{26})$ and D_2 is already specified.

4.5) Busy Period of the Server due to Preventive Maintenance

$BP_i^{pm}(t)$- Probability that the server is busy in preventive maintenance at an instant 't' Expressions for $BP_i^{pm}(t)$ are given below:

$BP_0^{pm}(t)= q_{01}(t)©BP_1^{pm}(t)+ q_{02}(t)©BP_2^{pm}(t)$ (17)

$BP_1^{pm}(t)=q_{10}(t)©BP_0^{pm}(t) + (q_{11.3}+q_{11.3(4,8)^n})©BP_1^{pm}(t)+q_{15}©BP_5^{pm}(t)+q_{19}©BP_9^{pm}(t)$ (18)

$BP_2^{pm}(t)= W_2^{pm}(t)+q_{20}(t)©BP_0^{r}(t) + q_{22.7}©BP_2^{r}(t)+q_{21.6}©BP_1^{r}(t)$ (19)

$BP_9^{pm}(t)= (q_{91}(t) + q_{91.(11,12)}+q_{91.10(8,4)^n}+q_{91.(10,8)})©BP_1^{pm}(t)$ (20)

$BP_5^{pm}(t)=W_5^{pm}(t)+q_{51}©BP_2^{pm}(t)$ (21)

Where, $W_2^{pm}(t)=e^{-2\lambda t}\overline{P(t)P_m(t)}, W_2^{pm}(t)=\overline{P_m(t)}$

Solving for $BP_0^{pm^{**}}(s)$

We get,

$BP_0^{pm}(\infty)= \lim_{n\to\infty} sBP_0^{pm^{**}}(s)=\frac{N_4}{D_2}$

$N_4=W_2^{pm^*}(0)p_{02}p_{10}+W_5^{pm^*}(0)p_{02}p_{26}p_{15}$ and D_2 is already specified.

4.6) Expected Number of Repairs (ENR) of the Unit

Let $Rp_i(t)$be the expected number of repairs by the server in (0, t].

The recursive relations for $Rp_i(t)$are given as:

$Rp_0(t)=Q_{01}(t)ⓈRp_1(t) + Q_{02}(t)ⓈRp_2(t)$ (22)

$Rp_1(t)=Q_{10}(t)Ⓢ(1+Rp_0(t)) + (Q_{11.3}+Q_{11.3(4,8)^n})Ⓢ(1+Rp_1(t))+Q_{11.5}ⓈRp_5(t) +Q_{19}ⓈRp_9(t)$ (23)

$Rp_2(t)=Q_{20}(t)ⓈRp_0(t) + Q_{22.7}ⓈRp_2(t) +Q_{21.6}ⓈRp_1(t)$ (24)

$Rp_9(t)=(Q_{91}(t) + Q_{91.(11,12)})ⓈRp_1(t)+(Q_{91.(10,8)}+ Q_{91.10(8,4)^n}))Ⓢ(1+Rp_1(t))$ (25)

$Rp_5(t)= Q_{51}ⓈRp_1(t)$ (26)

Solving for $Rp_0^{**}(s)$ We have,

$Rp_0(\infty)= \lim_{n\to\infty} sRp_0^{**}(s)=\frac{N_5}{D_2}$, $N_5= \left((p_{20}p_{01}+p_{26})(p_{10}+p_{13}+p_{19}p_{9,10})\right)$

4.7) Expected Number of Preventive Maintenance (PM) of the Unit

Let $Pm_i(t)$be the expected number of repairs by the server in (0, t].

The expressions for $Pm_i(t)$are given as:

$Pm_0(t)=Q_{01}(t)ⓈPm_1(t) + Q_{02}(t)ⓈPm_2(t)$ (27)

$Pm_1(t)=Q_{10}(t)ⓈPm_0(t) + (Q_{11.3}+Q_{11.3(4,8)^n})ⓈPm_1(t)+Q_{15}ⓈPm_5(t)) +Q_{19}ⓈPm_9(t)$ (28)

$Pm_2(t)=Q_{20}(t)Ⓢ(1+Pm_0(t)) + Q_{22.7}Ⓢ(1+Pm_2(t)) +Q_{21.6}Ⓢ(1+Pm_1(t)$ (29)

$Pm_9(t)=(Q_{91}(t)+Q_{91.(10,8)}+Q_{91.10(8,4)^n})ⓈPm_1(t)+(Q_{32}+Q_{91.(11,12)})Ⓢ(1+Pm_2(t))$ (30)

$Pm_5(t)= Q_{51}Ⓢ(1+Pm_1(t))$ (31)

Solving for $Pm_0^{**}(s)$(by taking LST)

$Pm_0(\infty)= \lim_{n\to\infty} sPm_0^{**}(s)=\frac{N_6}{D_2}$, $N_6=p_{02}p_{10}+(p_{26}+p_{01}p_{20})\left(p_{15}+p_{9,11}p_{19}\right)$

4.8) Expected Number of Visits of the Server

Let $V_i(t)$be the expected number of visits by the server.

The equations for $V_i(t)$ are as follow:

$V_0(t)=Q_{01}(t)Ⓢ(1+V_1(t)) + Q_{02}(t)Ⓢ(1+V_2(t))$ (32)

$V_1(t)=Q_{10}(t)ⓈV_0(t) + (Q_{11.3}+Q_{11.3(4,8)^n})ⓈV_1(t)+Q_{15}ⓈV_5(t)) +Q_{19}ⓈV_9(t)$ (33)

$V_2(t)=Q_{20}(t)ⓈV_0(t) + Q_{22.7}ⓈV_2(t) +Q_{21.6}ⓈV_1(t)$ (34)

$V_9(t)=(Q_{31}(t) + Q_{91.(11,12)}+Q_{91.10(8,4)^n}++Q_{91.10,8})\{V_1 t\}$ (35)

$V_5(t)= Q_{51}ⓈV_1(t)$ (36)

Solving for $V_0^{**}(s)$

$V_0(\infty)= \lim_{n\to\infty} sV_0^{**}(s)=\frac{N_7}{D_2}$,$N_7=p_{10}\left(p_{20}+p_{26}\right)$

5. Profit Analysis

The profit analysis of the model can be represented as : $P_C = C_0 A_0 - C_1 BP_0^r - C_2 BP_0^{pm} - C_3 V_0$

Here, C_0, C_1, C_2 & C_3 are respectively the revenue per unit time, cost per unit time the service facility is busy for repair, busy for preventive maintenance and costs per unit time visit by the service facility The particular case $p(t)=\alpha e^{-\alpha t}$, $p_m(t)=\eta e^{-\eta t}$, $r_s(t)=ae^{-at}$ and $r(t)=be^{-bt}$, $C_0=10000$, $C_1=1000$, $C_2=500$, $C_3=100$ has been considered to obtain the reliability measures MTSF, availability and profit function. The values of these measures have been evaluated for arbitrary values of the parameters since there is no reliable source of information which tells about the actual values of the parameters.

6. Results and Graphical Representation of Reliability Measures

Reliability measures such as MTSF, Availability and Profit have been studied for different values of parameters. The graphs have been plotted for a range (1.1-1.8) of values of preventive maintenance completion rate(η) and corresponding effects have been explained.

6.1) MTSF Vs Rate of Preventive Maintenance (η)

a: Treatment rate of the server
b: Repair rate of the unit
α: Rate by which unit undergoes for preventive maintenance

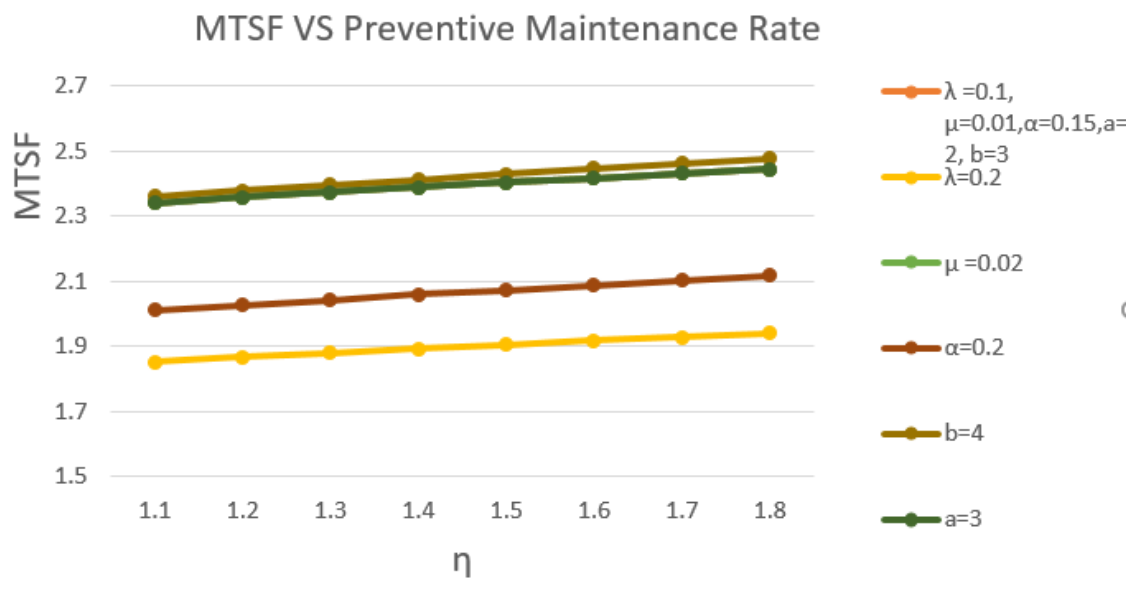


Figure 2: *MTSF vs Preventive Maintenance Rate*

From fig.2 it is quite evident that the MTSF increases with the increase in preventive maintenance rate, repair rate(b) of unit and server(a) as well. However, if the failure rate of the units is increasing (λ) from 0.1 to 0.2, MTSF decreases.

6.2) Availability vs Preventive maintenance Rate

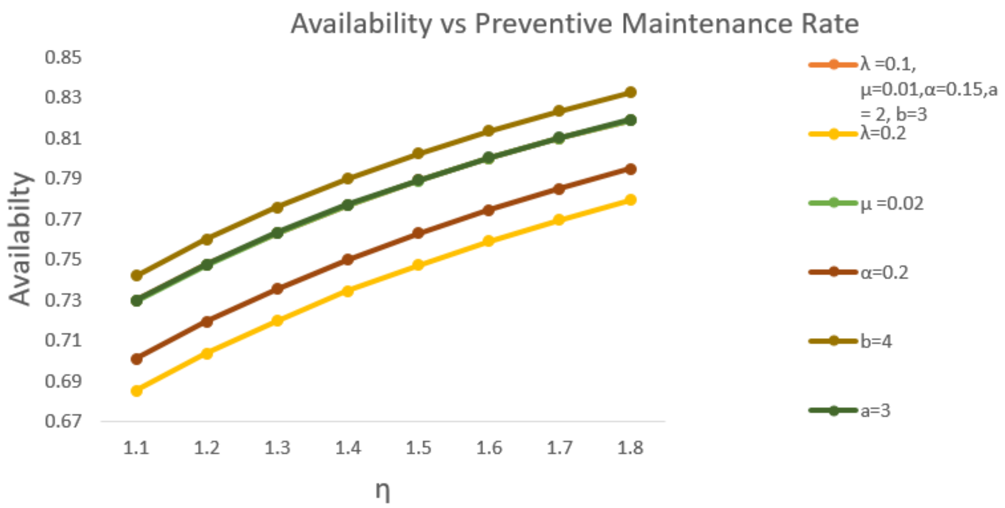


Figure 3: *Availability vs Preventive Maintenance Rate*

Fig.3 shows that availability of the system keeps on increasing if preventive maintenance rate (η) of the units is increased (from 1.1 to 1.8) and decreases with the increase in failure rate of units.

6.3) Profit vs Preventive Maintenance rate (η)

Figure 4: *Profit vs Preventive Maintenance Rate*

Fig.4 shows profit increases steadily with the increase in the preventive maintenance rate (1.1 to 1.8), repair rate (increased from 3 to 4) of the units however it sharply decreases with the increase in failure rates of the units and server.

7. Conclusion

A 2-out-of-3 system has been developed using the ideas of priority for preventive maintenance and conditional failure of the service facility. The MTSF, availability and profit function of this system have been obtained for particular values of the parameters. The study reveals that there is a decline in these measures with the increase of the rate by which unit undergoes for preventive maintenance, failure rates of the units and service facility. However, the values of MTSF, availability and profit function keep on increasing with the increase of treatment rate, repair rate of the unit and preventive maintenance completion rate.

The profit increases if the rate with which a unit completes its preventive maintenance. Hence, implementing the preventive maintenance repair policy for a 2-out-of-3 system is beneficial as it increases the availability and hence the profit of the system.

8.Application of the Proposed Model

Burning of toxic waste, especially waste produced in chemical factories, medical laboratories produce a lot of harmful and lethal gases. To neutralize the toxicity of these kind of gases huge amount of fire is required. Toxic waste incinerator is one such kind of combustion technique which helps in neutralising these harmful gases. It has a circuit named as flame detection circuit which works as a 2 -out-of-3system. The principal behind the working of the incinerator: As long as a sufficient amount of heat is maintained in the incinerator, it is safe to put the waste inside the chamber to neutralize the toxic gases produced due to combustion of waste. If the flame is not sufficient enough to neutralize the toxic gases produced, it would not be safe to keep inserting the waste inside the chambers because the gases produced would exit without being neutralized and may cause a severe health issue to anyone nearby. Therefore, our main focus is on designing the system in such a way that the system detects the sufficient amount of flame and permits waste insertion only when there is sufficient flame to neutralize the exhaust. Due to high risk involved in passing out the waste un neutralized it would be beneficial to make the flame detection system redundant. So that if one sensor fails to detect the flame, other may sense and cover up the risk involved. Hence it is very much necessary to design a system which opens the waste valve only if there is enough flame signalled by the sensors. The best designed system for this kind of incinerator is a 2-Out-of-3 system. There are three sensors to detect the flame which are identical in nature and the valve for injection of waste will open only if two out of three will signal that there is sufficient amount of flame inside the chamber to neutralize the toxic gases produced. If any of the two sensors fails to detect the flame it would lead to hazardous condition therefore it is highly recommended to keep an eye on the working condition of the sensors. Preventive maintenance is one such precautionary measure which ensures the proper functioning of the sensors hence it is kept as priority over repair while developing the system model.

A pictorial presentation of Toxic Waste Incinerator:

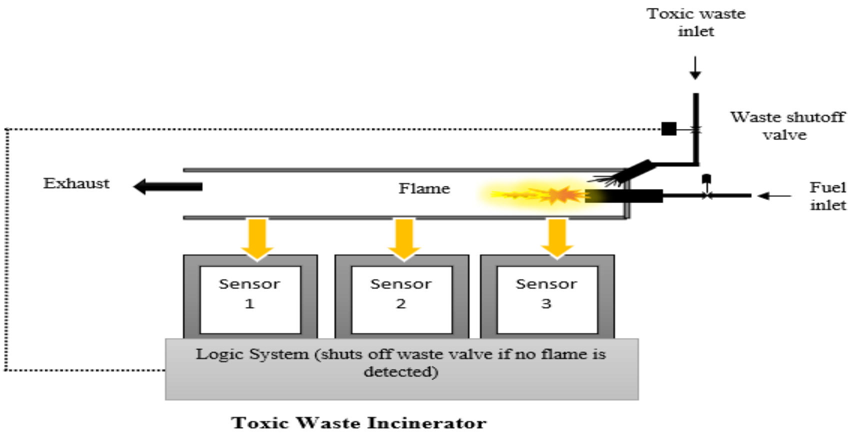


Figure 5: *Toxic Waste Incinerator*

References

[1] Malik, S.C. and Bhardwaj, R.K. (2007). Reliability and Cost Benefit Analysis of 2-out-of-3 Redundant System with General Distribution of Repair and Waiting Time, International Journal for Business and IT (DTR), Vol. 4 (1), pp. 28 – 35.

[2] Malik, S.C. and Singh, M. (2009). Reliability Modelling of 2-out-of 3 Redundant System Subject to Degradation After Repair, Journal of Reliability and Statistical Studies, Vol. 2(2), pp.91-104

[3] Kishan, R., and Kumar, M. (2009). Stochastic analysis of a two-unit parallel system with preventive maintenance, Journal of Reliability and Statistical Studies, vol. 2(2), pp. 31- 38.

[4] Jain, M. and Gupta, R. (2012). Load Sharing m-out-of-n: G System with Non -Identical Components Subject to Common Cause Failure. International Journal of Mathematics in Operational Research, Vol.4(5), pp. 586-605.

[5] Yang, Chunbo, Zeng, Shengkui and Guo, Jianbin (2015). Reliability Analysis of Load-Sharing K-out-of-N System considering Component Degradation, Mathematical Problems in Engineering, ID 72853, pp.1-10.

[6] Poonia, P.K., Sirohi, A. and Kumar, A. (2021). Cost analysis of a repairable warm standby *k*-out-of-*n*: G and 2-out-of-4: G systems in series configuration under catastrophic failure using copula repair,Life Cycle Reliability and Safety Engineering, Vol. 10, No. 2, pp. 121-133.

[7] Malik,S.C. Jha P.C., Anuradha (2021).Profit Analysis of a 1-out-of-2:G System with Priority to Repair and Conditional Failure of Service Facility, Journal of Statistics and Reliability Engineering,Vol.8(2),pp.264-271

[8] Singh, V.V., Mohammad, A.I., Ibrahim, K.H., and Yusuf, I. (2021). Performance assessment of the complex repairable system with n-identical units under (k-out-of-n: G) scheme and copula linguistic repair approach. International Journal of Quality Reliability and Management. Vol. 39(2), pp. 367-386.

[9] Abdullahi, S. and Yusuf, I. (2021). Cost Analysis of 2-out-of-4 System Connected to Two Unit Parallel Supporting Device for Operation. Life Cycle Reliability and Safety Engineering, Vol.10, pp. 113-119.

RT&A, No 1 (72)
Volume 18, March 2023

E-Bayesian estimations for Chen distribution under Type II censoring with medical application

Athirakrishnan R. B. and E. I. Abdul Sathar

•

Department of Statistics,University of Kerala
athirakrishnanrb91@gmail.com & sathare@gmail.com

Abstract

The study focuses on the E-Bayesian estimation of a Type-II censored sample from the Chen distribution. Three distinct prior distributions for the hyper-parameters and three different loss functions are considered here for deriving the E-Bayes estimators of the scale parameter and hazard rate of above said distribution under Type-II censoring. Also derived analytical expressions for the E-MSE of the proposed estimators. Additionally, several features of the E-Bayesian estimators and E-MSEs are derived. This paper compares E-Bayesian estimation with traditional estimation methods like MLE and Bayesian. The applicability of the proposed estimators is demonstrated using a real data application. Furthermore, the credible intervals of the scale parameter estimators are also provided. The numerical analysis demonstrates that the proposed method is simpler and more feasible than traditional techniques.

Keywords: Chen distribution; Type II censoring; Bayesian estimation; E-Bayesian estimation; E-MSE.

1 Introduction

Experiments in reliability and life-testing are done to learn more about the time of a significant event of interest. Examples of situations where the time of occurrence is significant are when a component fails, a disease abrogates, or a biological unit dies. For some reason, most investigations might not have complete information on the lifetimes or failure times of the experimental units. For example, in a medical trial, patients may withdraw from treatment, or the funding is only available for a specific period. In industrial trials, it is planned to remove accidentally damaged units before they fail to reduce testing time and costs. Censored data are those obtained from such experiments. The censoring schemes that appear the most frequently in the literature are Type-I and Type-II censoring. The experiment's endpoint is fixed, while the amount of failures reported is random in Type-I censoring. In contrast, the experiment's endpoint is random, and the number of failures is fixed in Type-II censoring. Type-II censoring is more cost-effective than Type-I censoring when comparing two censoring schemes. Inference under Type-II censoring for different parametric family distributions has been thoroughly studied in the literature. For more details, one can refer to [23], [7], [4], [9], [10].

A number of distributions with hazard rate functions that are constant, increasing, or decreasing in nature are discussed in the reliability literature. These distributions include generalized exponential, gamma, Weibull, and lognormal. These are the most often used models, and we use them to investigate various phenomena that occur in real life. However, these models do not work well with data sets showing bathtub-shaped hazard rates. In order to analyze real data with bathtub-shaped failure rates, several authors introduced probability models like modified Weibull by [18] and extended Weibull by [19], but these models are still unsuitable for producing accurate bathtub shape failure rates. Chen [8] showed a two-parameter lifespan distribution with a bathtub-shaped or increasing failure rate function. The hazard rate for this distribution initially declines, keeps the same, and increases. Chen distribution is a useful model for examining the lifespan of mechanical and electronic devices and humans. In addition to well-known models like lognormal and gamma, it can also be used to model positively skewed data. Due to the two parameters, closed-form confidence intervals for the shape parameter and joint confidence regions, this distribution is adaptable.

The MLEs of the Chen distribution parameter using samples that have been progressively Type-II censored are calculated by [24]. The MLE and Bayes estimators for the parameters of the Chen distribution using complete and censored samples are derived by [22], [2], [17], [16]. Kayal et al. [15] developed point and interval estimates of the multicomponent stress-strength reliability model of an s-out-of-j system using both classical and Bayesian techniques under the presumption that both the stress and the strength variables follow a Chen distribution. The literature on estimations of Chen distributions mentioned above mainly focuses on MLE or Bayesian techniques.

In addition to the Bayesian approach, the E-Bayesian estimation method, was developed in the literature. Originally, Han [13] addressed the definition of E-Bayesian estimation. Since the prior distribution of the hyperparameters is taken into account, the E-Bayesian approach is more reliable than Bayesian. The term "E-Bayesian estimation" refers to the expectation of the parameter's Bayesian estimate for all hyperparameters. In recent days so many works related to E-Bayesian inference of parameters and reliability functions of different distributions using complete and censored samples are discussed in the literature. For more details, one can refer to [1], [12],[3], [21], [20], [5]. The works mentioned above were a source of inspiration for further research on E-Bayes estimators for the scale parameter and hazard rate of Chen distributions under Type-II censoring schemes. The present work aims to develop E-Bayes estimators for the scale parameter and hazard rate of the Chen distribution using a Type-II censoring scheme and to calculate E-MSEs for the proposed estimators.

The organization of the remaining part of the work is as follows. In part 2, we go through the MLE of the scale parameter and hazard rate of the Chen distribution under the Type-II censoring scheme. Section 3 discusses the estimators' MSE as well as the Bayesian estimation of the scale parameter and hazard rate. Section 4 developed how to obtain E-Bayesian estimators of the scale parameter, hazard rate, and their associated E-MSEs. Section 5 of the article discusses the features that all of these estimators possess. In Section 6, it is discussed how well the estimators work with real data set. The final findings of the proposed study are provided in Section 7.

2 Maximum Likelihood Estimation

In this section, using a Type-II censoring technique, we derive the MLE of the scale parameter and hazard rate of the Chen distribution. The pdf, cdf and hazard function of the Chen distribution are respectively given by

$$f(x;\theta,\lambda) = \theta\lambda x^{\lambda-1} e^{x^{\lambda}+\theta(1-e^{x^{\lambda}})}, \quad x>0, \quad \lambda>0, \quad \theta>0, \tag{1}$$

$$F(x;\theta,\lambda) = 1-e^{\theta(1-e^{x^{\lambda}})}, \quad x>0, \quad \lambda>0, \quad \theta>0 \tag{2}$$

and

$$h(t) = \theta\lambda t^{\lambda-1} e^{t^{\lambda}}, \quad t>0. \tag{3}$$

With pdf and cdf defined in (1) and (2), respectively, assume that n distinct units selected from a population are put to the test and that the associated lifetimes are distributed identically. Let X = $(X_{(1)}, X_{(2)}, ..., X_{(r)})$ be the Type-II censored sample taken from (1) with r failure times. The likelihood function for Type-II censored sample is given by

$$\begin{aligned} L(\lambda,\theta|\underline{x}) &= \frac{n!}{(n-r)!}\pi_{i=1}^{r} f(x_{(i)})[1-F(x_{(r)})]^{n-r} \\ &= \frac{n!}{(n-r)!}\theta^{r}\nu(\lambda,\underline{x})e^{-\theta T}. \end{aligned} \tag{4}$$

where $\nu(\lambda, \underline{x}) = \lambda^r \pi_{i=1}^r x_i^{\lambda-1} e^{\sum_{i=1}^r x_i^\lambda}$, $\underline{x} = (x_{(1)}, x_{(2)}, ..., x_{(r)})$ and $T = \sum_{i=1}^r e^{x_i^\lambda} + (n-r)e^{x_r^\lambda} - n$.
The log-likelihood function is provided by Chen distribution when λ is known is given by

$$\ln(L) = \ln \theta^r + \ln e^{-\theta T}.$$

The normal equation is obtained by differentiating the log-likelihood with respect to the scale parameter θ and equating them to zero.

$$\frac{\partial \ln(L)}{\partial \theta} = 0 \implies \frac{r}{\theta} - T = 0.$$

By solving the above equation we can obtain MLE of the parameter θ as

$$\hat{\theta}_{ML} = \frac{r}{T}. \tag{5}$$

3 Bayesian Estimation

The Bayes estimators of the parameter θ are obtained in this section based on the squared error loss function (SELF), entropy loss function (ELF), and precautionary loss function (PLF). For developing the Bayesian estimation, we assume the gamma distribution as conjugate prior with probability density function

$$\pi(\theta|a,b) = \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta}, \quad \theta > 0, \quad a, b > 0, \tag{6}$$

where a and b are the hyper parameters. The posterior density of θ can be expressed as the following using the prior density (6) and likelihood function (4) as

$$q(\theta|\underline{x}) = \frac{(b+T)^{r+a}}{\Gamma(r+a)} \theta^{r+a-1} e^{-\theta(b+T)}, \quad \theta > 0. \tag{7}$$

We arrived at the Bayes estimators of θ and the hazard rate of (1) under three distinct loss functions in the subsequent theorem.

Theorem 1. For the Type-II censored sample $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(r)})$ from (1) under SELF, ELF, and PLF together with the likelihood function (4) and prior distribution (6), we obtain the Bayes estimators of θ and hazard rate, respectively, provided as

i) Under SELF

$$\hat{\theta}_{B1} = \frac{r+a}{b+T}, \tag{8}$$

$$\hat{h}(t)_{B1} = \frac{r+a}{b+T} \lambda t^{\lambda-1} e^{t^\lambda}. \tag{9}$$

ii) Under ELF

$$\hat{\theta}_{B2} = \frac{r+a-1}{b+T}, \tag{10}$$

$$\hat{h}(t)_{B2} = \frac{r+a-1}{b+T} \lambda t^{\lambda-1} e^{t^\lambda}. \tag{11}$$

iii) Under PLF

$$\hat{\theta}_{B3} = \sqrt{\frac{(r+a+1)(r+a)}{(b+T)^2}}, \tag{12}$$

$$\hat{h}(t)_{B3} = \sqrt{\frac{(r+a+1)(r+a)}{(b+T)^2}} \lambda t^{\lambda-1} e^{t^\lambda}. \tag{13}$$

Proof.

i) The mean of (7) serves as the Bayes estimator of θ under SELF and is written as

$$\hat{\theta}_{B1} = E(\theta|\underline{x}) = \frac{r+a}{b+T}.$$

Likewise, the Bayes estimator of hazard rate is

$$\hat{h}(t)_{B1} = E\left(\theta\lambda t^{\lambda-1}e^{t^{\lambda}}|\underline{x}\right) = \frac{r+a}{b+T}\lambda t^{\lambda-1}e^{t^{\lambda}}.$$

ii) The following is the expression of a Bayes estimator of θ using ELF:

$$\hat{\theta}_{B2} = \left[E\left(\frac{1}{\theta}|\underline{x}\right)\right]^{-1} = \frac{r+a-1}{b+T}.$$

Likewise, the Bayes estimator of hazard rate is

$$\hat{h}(t)_{B2} = \left[E\left(\left(\theta\lambda t^{\lambda-1}e^{t^{\lambda}}\right)^{-1}|\underline{x}\right)\right]^{-1} = \frac{r+a-1}{b+T}\lambda t^{\lambda-1}e^{t^{\lambda}}.$$

iii) The following is the expression of a Bayes estimator of θ using PLF:

$$\hat{\theta}_{B3} = \sqrt{E(\theta^2|\underline{x})} = \sqrt{\frac{(r+a)(r+a+1)}{(b+T)^2}}.$$

Likewise, the Bayes estimator of hazard rate is

$$\hat{h}(t)_{B3} = \sqrt{E\left(\left(\theta\lambda t^{\lambda-1}e^{t^{\lambda}}\right)^2|\underline{x}\right)} = \sqrt{\frac{(r+a+1)(r+a)}{(b+T)^2}}\lambda t^{\lambda-1}e^{t^{\lambda}}.$$

■

We determined the MSE of the Bayes estimators of θ and the hazard rate of (1) for three distinct loss functions in the subsequent theorem.

Theorem 2. For the Type-II censored sample $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(r)})$ from (1), the MSE of the Bayes estimators of θ and the hazard rate under SELF, ELF, and PLF, respectively as

i) Under SELF

$$MSE(\hat{\theta}_{B1}) = \frac{r+a}{(b+T)^2}, \tag{14}$$

$$MSE(\hat{h}(t)_{B1}) = (\lambda t^{\lambda-1}e^{t^{\lambda}})^2\frac{r+a}{(b+T)^2}. \tag{15}$$

ii) Under ELF

$$MSE(\hat{\theta}_{B2}) = \frac{r+a+1}{(b+T)^2}, \tag{16}$$

$$MSE(\hat{h}(t)_{B2}) = (\lambda t^{\lambda-1}e^{t^{\lambda}})^2\frac{r+a-1}{(b+T)^2}. \tag{17}$$

iii) Under PLF

$$MSE(\hat{\theta}_{B3}) = \frac{2(r+a)}{(b+T)^2}[(r+a+1) - \sqrt{(r+a+1)(r+a)}], \tag{18}$$

$$MSE(\hat{h}(t)_{B3}) = (\lambda t^{\lambda-1}e^{t^{\lambda}})^2\frac{2(r+a)}{(b+T)^2}[(r+a+1) - \sqrt{(r+a+1)(r+a)}]. \tag{19}$$

Proof.

i) MSE of the Bayes estimator of θ under SELF is defined as

$$\begin{aligned} MSE(\hat{\theta}_{B1}(a,b)) &= E(\theta^2|x) - 2\hat{\theta}_{B1}(a,b)E(\theta|x) + [\hat{\theta}_{B1}(a,b)]^2 \\ &= \frac{r+a}{(b+T)^2}. \end{aligned}$$

Likewise, the MSE of the Bayes estimator of hazard rate under SELF is as follows:

$$\begin{aligned} MSE(\hat{h}(t)_{B1}) &= E[h(t)^2|x] - 2\hat{h}(t)_{B1}E[h(t)|x] + [\hat{h}(t)_{B1}]^2 \\ &= (\lambda t^{\lambda-1}e^{t^\lambda})^2 \frac{r+a}{(b+T)^2}. \end{aligned}$$

ii) MSE of Bayes estimator of θ using ELF is defined as

$$\begin{aligned} MSE(\hat{\theta}_{B2}(a,b)) &= E(\theta^2|x) - 2\hat{\theta}_{B2}(a,b)E(\theta|x) + [\hat{\theta}_{B2}(a,b)]^2 \\ &= \frac{r+a-1}{(b+T)^2}. \end{aligned}$$

Likewise, the MSE of the Bayes estimator of hazard rate under ELF is as follows:

$$\begin{aligned} MSE(\hat{h}(t)_{B2}) &= E[h(t)^2|x] - 2\hat{h}(t)_{B2}E[h(t)|x] + [\hat{h}(t)_{B2}]^2 \\ &= \frac{r+a-1}{b+T}\lambda t^{\lambda-1}e^{t^\lambda}. \end{aligned}$$

iii) MSE of the Bayes estimator of θ using PLF is defined as

$$\begin{aligned} \hat{\theta}_{B3} &= E(\theta^2|x) - 2\hat{\theta}_{B3}(a,b)E(\theta|x) + [\hat{\theta}_{B3}(a,b)]^2 \\ &= \frac{2(r+a)}{(b+T)^2}[(r+a+1) - \sqrt{(r+a+1)(r+a)}]. \end{aligned}$$

Likewise, the MSE of the Bayes estimator of hazard rate under PLF is as follows:

$$\begin{aligned} \hat{h}(t)_{B3} &= E[h(t)^2|x] - 2\hat{h}(t)_{B3}E[h(t)|x] + [\hat{h}(t)_{B3}]^2 \\ &= (\lambda t^{\lambda-1}e^{t^\lambda})^2 \frac{2(r+a)}{(b+T)^2}[(r+a+1) - \sqrt{(r+a+1)(r+a)}]. \end{aligned}$$

■

4 E-Bayesian Estimation and its E-MSE

Han [13] is the author who first introduced E-Bayesian estimation in literature. Here we will obtain the E-Bayes estimator of the scale parameter and hazard rate of the Chen distribution under Type II censoring based on SELF, ELF and PLF and derive the properties exhibited by these estimators. Three different prior distributions of the hyper-parameters are considered to examine the impact of various prior distributions on the E-Bayesian estimate of θ. According to [13], it is important to establish that the prior distribution of a and b, indicated by $\pi(\theta|a,b)$, is a decreasing function in θ. Finding the first derivative of $\pi(\theta|a,b)$ with respect to θ and obtaining the result as

$$\frac{\partial \pi(\theta|a,b)}{\partial \theta} = \frac{b^a \theta^{a-2} e^{-b\theta}}{\Gamma a}[(a-1) - b\theta].$$

As a result, the function $\frac{\partial \pi(\theta|a,b)}{\partial \theta}$ <0 and $\pi(\theta|a,b)$ is a decreasing function of θ for 0< a <1 and b>0. Given 0< a <1, the gamma density function's tail will be thinner the larger b. The thinner-tailed prior distribution frequently affects the robustness of the Bayesian estimate, according to [6], which took the robustness of the Bayesian estimate into account. As a result, b must not exceed a specified upper bound c, where $c > 0$ is an unknown constant. As a result, the restriction of 0< a <1 and 0< $b < c$ should be used when choosing the hyper-parameters a and b.

Definition 4.1. $\hat{\theta}_{EB} = \int\int_D \hat{\theta}_B(a,b)\pi(a,b)dadb = E[\hat{\theta}(a,b)]$ is referred to as the E-Bayesian estimate of θ when $\hat{\theta}_B(a,b)$ is continuous and is considered finite. D is the domain of a and b, $\hat{\theta}_B(a,b)$ is the Bayesian estimation of θ with hyper-parameters a and b, and $\pi(a,b)$ is the prior density of a and b over D.

The expectation of the Bayesian estimation of θ for the hyperparameters is what definition 4.1 defines as the E-Bayesian estimation of θ. The definition of E-MSE of the E-Bayes estimators of θ presented by [11] is provided below.

Definition 4.2. $E - MSE(\hat{\theta}_{EB}) = \int\int_D MSE(\hat{\theta}_B(a,b))\pi(a,b)dadb = E[MSE(\hat{\theta}_B(a,b))]$ is referred to as the E-MSE of E-Bayes estimation of θ when $MSE(\hat{\theta}_B(a,b))$ is continuous and is considered finite. D is the domain of a and b, $MSE(\hat{\theta}(a,b))$ is the MSE of the Bayesian estimation of θ with hyper-parameters a and b, and $\pi(a,b)$ is the prior density of a and b over D.

The E-Bayesian estimators of the parameter θ are obtained in this section using three different prior distributions for the hyper-parameters a and b. These prior distributions were chosen to demonstrate how the various prior distributions affected the E-Bayesian estimation of the parameter θ. The prior distributions we used are given by

$$\pi_1(a,b) = \frac{2(c-b)}{c^2}, \quad 0 < a < 1 \quad 0 < b < c. \tag{20}$$

$$\pi_2(a,b) = \frac{1}{c}, \quad 0 < a < 1 \quad 0 < b < c. \tag{21}$$

$$\pi_3(a,b) = \frac{2b}{c^2}, \quad 0 < a < 1 \quad 0 < b < c. \tag{22}$$

These prior distributions are used to ensure that $\pi(\theta|a,b)$ is a decreasing function in θ. The E-Bayes estimators of the parameter θ and the hazard rate function using $\pi_1(a,b)$ are derived in the subsequent theorem under various loss functions.

Theorem 3. We have the E-Bayes estimators of θ and hazard rate, which are provided, for the censored sample $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(r)})$ from (1) using the prior distribution (20) under SELF, ELF, and PLF, respectively as

i) Under SELF

$$\hat{\theta}_{ES1} = \frac{2r+1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}, \tag{23}$$

$$\hat{h}(t)_{ES1} = \lambda t^{\lambda-1}e^{t^\lambda}\frac{2r+1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}. \tag{24}$$

ii) Under ELF

$$\hat{\theta}_{EE1} = \frac{2r-1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}, \tag{25}$$

$$\hat{h}(t)_{EE1} = \lambda t^{\lambda-1}e^{t^\lambda}\frac{2r-1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}. \tag{26}$$

iii) Under PLF

$$\hat{\theta}_{EP1} = \frac{2}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}\int_0^1 \sqrt{(r+a)(r+a+1)}da, \tag{27}$$

$$\hat{h}(t)_{EP1} = \lambda t^{\lambda-1}e^{t^\lambda}\frac{2}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}\int_0^1 \sqrt{(r+a)(r+a+1)}da. \tag{28}$$

Proof.

i) The following is the expression of the E-Bayes estimator of θ using SELF:

$$\hat{\theta}_{ES1} = \int_0^1 \int_0^c \hat{\theta}_{B1}(a,b)\pi_1(a,b)dadb$$

Using (8) and (20), the above equation simplifies to

$$\hat{\theta}_{ES1} = \frac{2r+1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}$$

Likewise, the E-Bayes estimator of hazard rate using SELF using (9) and (20) is as follows:

$$\begin{aligned}\hat{h}(t)_{ES1} &= \int\int_D \hat{h}_{B1}(a,b)\pi_1(a,b)dadb \\ &= \lambda t^{\lambda-1}e^{t^\lambda}\frac{2r+1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}.\end{aligned}$$

ii) The following is the expression of the E-Bayes estimator of θ using ELF:

$$\hat{\theta}_{EE1} = \int_0^1 \int_0^c \hat{\theta}_{B2}(a,b)\pi_1(a,b)dadb.$$

Using (10) and (20), the above equation simplifies to

$$\hat{\theta}_{EE1} = \frac{2r-1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}.$$

Likewise, the E-Bayes estimator of hazard rate using ELF using (11) and (20) is as follows:

$$\begin{aligned}\hat{h}(t)_{EE1} &= \int_0^1 \int_0^c \hat{h}_{B2}(a,b)\pi_1(a,b)dadb \\ &= \lambda t^{\lambda-1}e^{t^\lambda}\frac{2r-1}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}.\end{aligned}$$

iii) The following is the expression of the E-Bayes estimator of θ using PLF:

$$\hat{\theta}_{EP1} = \int_0^1 \int_0^c \hat{\theta}_{B3}\pi_1(a,b)dadb.$$

Using (12) and (20), the above equation simplifies to

$$\hat{\theta}_{EP1} = \frac{2}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}\int_0^1 \sqrt{(r+a)(r+a+1)}da.$$

Likewise, the E-Bayes estimator of hazard rate using PLF using (13) and (20) is as follows:

$$\hat{h}(t)_{EP1} = \lambda t^{\lambda-1}e^{t^\lambda}\frac{2}{c^2}\left\{(T+c)ln\left(\frac{T+c}{T}\right) - c\right\}\int_0^1 \sqrt{(r+a)(r+a+1)}da.$$

■

The E-Bayes estimators of the parameter θ and the hazard rate function using $\pi_2(a,b)$ are derived in the subsequent theorem under various loss functions.

Theorem 4. We have the E-Bayes estimators of θ and hazard rate, which are provided, for the censored sample $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(r)})$ from (1) using the prior distribution (21) under SELF, ELF, and PLF, respectively as

i) Under SELF

$$\hat{\theta}_{ES2} = \frac{2r+1}{2c}ln\left(\frac{T+c}{T}\right), \tag{29}$$

$$\hat{h}(t)_{ES2} = \lambda t^{\lambda-1}e^{t^\lambda}\frac{2r+1}{2c}ln\left(\frac{T+c}{T}\right). \tag{30}$$

ii) Under ELF

$$\hat{\theta}_{EE2} = \frac{2r-1}{2c} ln\left(\frac{T+c}{T}\right), \tag{31}$$

$$\hat{h}(t)_{EE2} = \lambda t^{\lambda-1} e^{t^\lambda} \frac{2r-1}{2c} ln\left(\frac{T+c}{T}\right). \tag{32}$$

iii) Under PLF

$$\hat{\theta}_{EP2} = \frac{1}{c} ln\left(\frac{T+c}{T}\right) \int_0^1 \sqrt{(r+a)(r+a+1)}da, \tag{33}$$

$$\hat{h}(t)_{EP2} = \lambda t^{\lambda-1} e^{t^\lambda} \frac{1}{c} ln\left(\frac{T+c}{T}\right) \int_0^1 \sqrt{(r+a)(r+a+1)}da. \tag{34}$$

Proof. The proof is excluded since it is similar to that of Theorem 3. ■

The E-Bayes estimators of the parameter θ and the hazard rate function using $\pi_3(a,b)$ are derived in the subsequent theorem under various loss functions.

Theorem 5. We have the E-Bayes estimators of θ and hazard rate, which are provided, for the censored sample $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(r)})$ from (1) using the prior distribution (22) under SELF, ELF, and PLF, respectively as

i) Under SELF

$$\hat{\theta}_{ES3} = \frac{2r+1}{c^2}\left\{c - Tln\left(\frac{T+c}{T}\right)\right\}, \tag{35}$$

$$\hat{h}(t)_{ES3} = \lambda t^{\lambda-1} e^{t^\lambda} \frac{2r+1}{c^2}\left\{c - Tln\left(\frac{T+c}{T}\right)\right\}. \tag{36}$$

ii) Under ELF

$$\hat{\theta}_{EE3} = \frac{2r-1}{c^2}\left\{c - Tln\left(\frac{T+c}{T}\right)\right\}, \tag{37}$$

$$\hat{h}(t)_{EE3} = \lambda t^{\lambda-1} e^{t^\lambda} \frac{2r-1}{c^2}\left\{c - Tln\left(\frac{T+c}{T}\right)\right\}. \tag{38}$$

iii) Under PLF

$$\hat{\theta}_{EP3} = \frac{2}{c^2}\left\{c - Tln\left(\frac{T+c}{T}\right)\right\} \int_0^1 \sqrt{(r+a)(r+a+1)}da, \tag{39}$$

$$\hat{h}(t)_{EP3} = \lambda t^{\lambda-1} e^{t^\lambda} \frac{2}{c^2}\left\{c - Tln\left(\frac{T+c}{T}\right)\right\} \int_0^1 \sqrt{(r+a)(r+a+1)}da. \tag{40}$$

Proof. The proof is excluded since it is similar to that of Theorem 3. ■

The E-MSE of the E-Bayes estimators of the parameter θ using different priors are derived in the subsequent theorem under various loss functions.

Theorem 6. The E-MSE of the E-Bayes estimators of θ using the priors $\pi_1(a,b)$, $\pi_2(a,b)$, and $\pi_3(a,b)$ under SELF, ELF, and PLF are presented, respectively, for the E-Bayes estimators of θ of Type-II censored sample $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(r)})$ from (1) as

i) Under SELF

$$E - MSE(\hat{\theta}_{ES1}) = \frac{2r+1}{c^2}\left\{ln\left(\frac{T}{T+c}\right) + \frac{c}{T}\right\}, \tag{41}$$

$$E - MSE(\hat{\theta}_{ES2}) = \frac{2r+1}{2}\left\{\frac{1}{T(c+T)}\right\}, \tag{42}$$

$$E - MSE(\hat{\theta}_{ES3}) = \frac{2r+1}{c^2}\left\{ln\left(\frac{T+c}{T}\right) - \frac{c}{c+T}\right\}. \tag{43}$$

ii) Under ELF

$$E - MSE(\hat{\theta}_{EE1}) = \frac{2r+3}{c^2}\left\{ln\left(\frac{T}{T+c}\right) + \frac{c}{T}\right\}, \quad (44)$$

$$E - MSE(\hat{\theta}_{EE2}) = \frac{2r+3}{2}\left\{\frac{1}{T(c+T)}\right\}, \quad (45)$$

$$E - MSE(\hat{\theta}_{EE3}) = \frac{2r+3}{c^2}\left\{ln\left(\frac{T+c}{T}\right) - \frac{c}{c+T}\right\}. \quad (46)$$

iii) Under PLF

$$\begin{aligned} E - MSE(\hat{\theta}_{EP1}) &= \frac{4}{c^2}\left\{ln\left(\frac{T}{T+c}\right) + \frac{c}{T}\right\}\int_0^1 (r+a) \\ &\quad [(r+a+1) - \sqrt{(r+a)(r+a+1)}]da, \end{aligned} \quad (47)$$

$$E - MSE(\hat{\theta}_{EP2}) = \frac{2}{T(c+T)}\int_0^1 (r+a)[(r+a+1) - \sqrt{(r+a)(r+a+1)}]da, \quad (48)$$

$$\begin{aligned} E - MSE(\hat{\theta}_{EP3}) &= \frac{4}{c^2}\left[ln\left(\frac{T+c}{T}\right) - \frac{c}{c+T}\right]\int_0^1 (r+a) \\ &\quad [(r+a+1) - \sqrt{(r+a)(r+a+1)}]da. \end{aligned} \quad (49)$$

Proof.

i) Under SELF, the E-MSE of the estimator, $\hat{\theta}_{ES1}$ can be obtained from (14) and (20) by using the definition (4.2) and is given by

$$\begin{aligned} E - MSE(\hat{\theta}_{ES1}) &= \int\int_D MSE(\hat{\theta}_{B1}(a,b))\pi_1(a,b)dadb \\ &= \frac{2r+1}{c^2}\left\{ln\left(\frac{T}{T+c}\right) + \frac{c}{T}\right\}. \end{aligned}$$

Similarly, the E-MSE of $\hat{\theta}_{ES2}$ and $\hat{\theta}_{ES3}$ can be obtained from (14), (21) and (22) and by using the definition (4.2) and are given, respectively, by

$$\begin{aligned} E - MSE(\hat{\theta}_{ES2}) &= \int\int_D MSE(\hat{\theta}_{B1}(a,b))\pi_2(a,b)dadb \\ &= \frac{2r+1}{2}\left\{\frac{1}{T(c+T)}\right\}, \end{aligned}$$

and

$$\begin{aligned} E - MSE(\hat{\theta}_{ES3}) &= \int\int_D MSE(\hat{\theta}_{B1}(a,b))\pi_3(a,b)dadb \\ &= \frac{2r+1}{c^2}\left\{ln\left(\frac{T+c}{T}\right) - \frac{c}{c+T}\right\}. \end{aligned}$$

ii) Under ELF, the E-MSE of the estimator, $\hat{\theta}_{EE1}$ can be obtained from (15) and (20), and using the definition (4.2) and is given by

$$\begin{aligned} E - MSE(\hat{\theta}_{EE1}) &= \int\int_D MSE(\hat{\theta}_{B2}(a,b))\pi_1(a,b)dadb \\ &= \frac{2r+3}{c^2}\left\{ln\left(\frac{T}{T+c}\right) + \frac{c}{T}\right\}. \end{aligned}$$

Similarly, the E-MSE of $\hat{\theta}_{EE2}$ and $\hat{\theta}_{EE3}$ can be obtained from (15), (21) and (22) and by using the definition (4.2) and are given, respectively, by

$$E - MSE(\hat{\theta}_{EE2}) = \int\int_D MSE(\hat{\theta}_{B2}(a,b))\pi_2(a,b)dadb$$

$$= \frac{2r+3}{2}\left\{\frac{1}{T(c+T)}\right\},$$

and

$$\begin{aligned} E-MSE(\hat{\theta}_{EE3}) &= \int\int_D MSE(\hat{\theta}_{B2}(a,b))\pi_3(a,b)dadb \\ &= \frac{2r+3}{c^2}\left\{ln\left(\frac{T+c}{T}\right)-\frac{c}{c+T}\right\}. \end{aligned}$$

iii) Under PLF, the E-MSE of the estimator, $\hat{\theta}_{EP1}$ can be obtained from (16) and (20) by using the definition (4.2) and is given by

$$\begin{aligned} E-MSE(\hat{\theta}_{EP1}) &= \int\int_D MSE(\hat{\theta}_{B3}(a,b))\pi_1(a,b)dadb \\ &= \frac{4}{c^2}\left\{ln\left(\frac{T}{T+c}\right)+\frac{c}{T}\right\}\int_0^1(r+a) \\ &\quad [(r+a+1)-\sqrt{(r+a)(r+a+1)}]da. \end{aligned}$$

Similarly, the E-MSE of $\hat{\theta}_{EP2}$ and $\hat{\theta}_{EP3}$ can be obtained from (16), (20) and (21) and by using the definition (4.2) and are given, respectively, by

$$\begin{aligned} E-MSE(\hat{\theta}_{EP2}) &= \int\int_D MSE(\hat{\theta}_{B3}(a,b))\pi_2(a,b)dadb \\ &= \frac{2}{T(c+T)}\int_0^1(r+a)[(r+a+1)-\sqrt{(r+a)(r+a+1)}]da, \end{aligned}$$

and

$$\begin{aligned} E-MSE(\hat{\theta}_{EP3}) &= \int\int_D MSE(\hat{\theta}_{B3}(a,b))\pi_2(a,b)dadb \\ &= \frac{4}{c^2}\left[ln\left(\frac{T+c}{T}\right)-\frac{c}{c+T}\right] \\ &\quad \int_0^1(r+a)[(r+a+1)-\sqrt{(r+a)(r+a+1)}]da. \end{aligned}$$

■

The E-MSE of the E-Bayes estimators of the hazard rate $h(t)$ using different priors are derived in the subsequent theorem under various loss functions.

Theorem 7. The E-MSE of the E-Bayes estimators of $h(t)$ using the priors $\pi_1(a,b)$, $\pi_2(a,b)$, and $\pi_3(a,b)$ under SELF, ELF, and PLF are presented, respectively, for the E-Bayes estimators of $h(t)$ of Type-II censored sample $\underline{X}=(X_{(1)},X_{(2)},...,X_{(r)})$ from (1) as

i) Under SELF

$$E-MSE(\hat{h}(t)_{ES1}) = (\lambda t^{\lambda-1}e^{t^\lambda})^2\frac{2r+1}{c^2}\left\{ln\left(\frac{T}{T+c}\right)+\frac{c}{T}\right\}, \tag{50}$$

$$E-MSE(\hat{h}(t)_{ES2}) = (\lambda t^{\lambda-1}e^{t^\lambda})^2\frac{2r+1}{2}\left\{\frac{1}{T(c+T)}\right\}, \tag{51}$$

$$E-MSE(\hat{h}(t)_{ES3}) = (\lambda t^{\lambda-1}e^{t^\lambda})^2\frac{2r+1}{c^2}\left\{ln\left(\frac{T+c}{T}\right)-\frac{c}{c+T}\right\}. \tag{52}$$

ii) Under ELF

$$E-MSE(\hat{h}(t)_{EE1}) = (\lambda t^{\lambda-1}e^{t^\lambda})^2\frac{2r+3}{c^2}\left\{ln\left(\frac{T}{T+c}\right)+\frac{c}{T}\right\}, \tag{53}$$

$$E-MSE(\hat{h}(t)_{EE2}) = (\lambda t^{\lambda-1}e^{t^\lambda})^2\frac{2r+3}{2}\left\{\frac{1}{T(c+T)}\right\}, \tag{54}$$

$$E-MSE(\hat{h}(t)_{EE3}) = (\lambda t^{\lambda-1}e^{t^\lambda})^2\frac{2r+3}{c^2}\left\{ln\left(\frac{T+c}{T}\right)-\frac{c}{c+T}\right\}. \tag{55}$$

iii) Under PLF

$$E-MSE(\hat{h}(t)_{EP1}) = (\lambda t^{\lambda-1}e^{t^{\lambda}})^2\frac{4}{c^2}\left\{ln\left(\frac{T}{T+c}\right)+\frac{c}{T}\right\}\int_0^1(r+a) \\ [(r+a+1)-\sqrt{(r+a)(r+a+1)}]da, \tag{56}$$

$$E-MSE(\hat{h}(t)_{EP2}) = (\lambda t^{\lambda-1}e^{t^{\lambda}})^2\frac{2}{T(c+T)}\int_0^1(r+a) \\ [(r+a+1)-\sqrt{(r+a)(r+a+1)}]da, \tag{57}$$

$$E-MSE(\hat{h}(t)_{EP3}) = (\lambda t^{\lambda-1}e^{t^{\lambda}})^2\frac{4}{c^2}\left[ln\left(\frac{T+c}{T}\right)-\frac{c}{c+T}\right]\int_0^1(r+a) \\ [(r+a+1)-\sqrt{(r+a)(r+a+1)}]da. \tag{58}$$

Proof. The proof is excluded since it is similar to that of Theorem 6. ∎

5 Properties of E-Bayesian estimation

We now go over some important features of E-Bayesian estimators and their E-MSE. The relationship between E-Bayes estimators of θ under various loss functions is given in the subsequent theorem.

Theorem 8. Using the priors $\pi_1(a,b), \pi_2(a,b)$ and $\pi_3(a,b)$ under various loss functions, the relationship between E-Bayes estimators of θ when 0< c <T is given as

a) under SELF
   i) $\hat{\theta}_{ES3} < \hat{\theta}_{ES1} < \hat{\theta}_{ES2}$,
   ii) $\lim_{T\to\infty}\hat{\theta}_{ES1} = \lim_{T\to\infty}\hat{\theta}_{ES2} = \lim_{T\to\infty}\hat{\theta}_{ES3}$,

b) under ELF
   i) $\hat{\theta}_{EE3} < \hat{\theta}_{EE1} < \hat{\theta}_{EE2}$,
   ii) $\lim_{T\to\infty}\hat{\theta}_{EE1} = \lim_{T\to\infty}\hat{\theta}_{EE2} = \lim_{T\to\infty}\hat{\theta}_{EE3}$,

c) under PLF
   i) $\hat{\theta}_{EP3} < \hat{\theta}_{EP1} < \hat{\theta}_{EP2}$,
   ii) $\lim_{T\to\infty}\hat{\theta}_{EP1} = \lim_{T\to\infty}\hat{\theta}_{EP2} = \lim_{T\to\infty}\hat{\theta}_{EP3}$.

Proof.

a) Under SELF
   i) From (23) and (29), we have

$$\hat{\theta}_{ES1}-\hat{\theta}_{ES2} = \frac{2r+1}{c}\left[\left(\frac{1}{2}+\frac{c}{T}\right)\ln\left(\frac{T+c}{T}\right)-1\right]. \tag{59}$$

For $-1 < x < +1$, we have, $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - ... = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k}$.
Let $x = \frac{c}{T}$, when $0 < c < T, 0 < \frac{c}{T} < 1$, we get

$$\begin{aligned}
&\left[\left(\frac{1}{2}+\frac{c}{T}\right)\ln\left(\frac{T+c}{T}\right)\right]-1\\
&\quad=\left(\frac{1}{2}+\frac{c}{T}\right)\left[\frac{c}{T}-\frac{1}{2}\left(\frac{c}{T}\right)^2+\frac{1}{3}\left(\frac{c}{T}\right)^3-\frac{1}{4}\left(\frac{c}{T}\right)^4+...\right]-1\\
&\quad=\left[\frac{1}{2}\left(\frac{c}{T}\right)-\frac{1}{4}\left(\frac{c}{T}\right)^2+\frac{1}{6}\left(\frac{c}{T}\right)^3-\frac{1}{8}\left(\frac{c}{T}\right)^4+...\right]+\left[\left(\frac{c}{T}\right)^2-\frac{1}{2}\left(\frac{c}{T}\right)^3\right.\\
&\quad\quad\left.+\frac{1}{3}\left(\frac{c}{T}\right)^4-\frac{1}{4}\left(\frac{c}{T}\right)^5+...\right]-1\\
&\quad=\frac{1}{2}\left(\frac{c}{T}\right)+\frac{3}{4}\left(\frac{c}{T}\right)^2\left[1-\frac{4}{9}\left(\frac{c}{T}\right)\right]+\frac{5}{24}\left(\frac{c}{T}\right)^4\left[1-\frac{18}{25}\left(\frac{c}{T}\right)\right]+...-1\\
&\quad<0.
\end{aligned} \tag{60}$$

So, we can say that,

$$\hat{\theta}_{ES1} < \hat{\theta}_{ES2}. \tag{61}$$

Now from (23) and (35), we have

$$\begin{aligned}
&\hat{\theta}_{ES1}-\hat{\theta}_{ES3}\\
&\quad=\frac{2r+1}{c^2}\left[(T+c)\ln\left(\frac{T+c}{T}\right)-c\right]-\frac{2r+1}{c^2}\left[c-T\ln\left(\frac{T+c}{T}\right)\right]\\
&\quad=\frac{2r+1}{c^2}\left[(2T+c)\ln\left(\frac{T+c}{T}\right)-2c\right]\\
&\quad=\frac{2r+1}{c}\left[\frac{(2T+c)}{c}\ln\left(\frac{T+c}{T}\right)-2\right].
\end{aligned} \tag{62}$$

$$\begin{aligned}
&\left[\left(\frac{2T}{c}+1\right)\ln\left(\frac{T+c}{T}\right)\right]-2\\
&\quad=\left[2-\left(\frac{c}{T}\right)+\left(\frac{2}{3}\right)\left(\frac{c}{T}\right)^2-\left(\frac{1}{2}\right)\left(\frac{c}{T}\right)^3+...\right]+\left[\left(\frac{c}{T}\right)-\frac{1}{2}\left(\frac{c}{T}\right)^2\right.\\
&\quad\quad\left.+\frac{1}{3}\left(\frac{c}{T}\right)^3-\frac{1}{4}\left(\frac{c}{T}\right)^4+...\right]-2\\
&\quad=\left(\frac{1}{6}\right)\left(\frac{c}{T}\right)^2\left[1-\left(\frac{c}{T}\right)\right]+\frac{3}{20}\left(\frac{c}{T}\right)^4\left[1-\frac{8}{9}\left(\frac{c}{T}\right)\right]+...\\
&\quad>0.
\end{aligned} \tag{63}$$

So we can say that,

$$\hat{\theta}_{ES1} > \hat{\theta}_{ES3}. \tag{64}$$

From (61) and (64),

$$\hat{\theta}_{ES3} < \hat{\theta}_{ES1} < \hat{\theta}_{ES2}. \tag{65}$$

ii) From (59) and (60) and by applying limit $T \to \infty$

$$\begin{aligned}
&\lim_{T\to\infty}(\hat{\theta}_{ES1}-\hat{\theta}_{ES2})\\
&=\left(\frac{2r+1}{c}\right)\lim_{T\to\infty}\{\frac{1}{2}\left(\frac{c}{T}\right)+\frac{3}{4}\left(\frac{c}{T}\right)^2\left[1-\frac{4}{9}\left(\frac{c}{T}\right)\right]+\frac{5}{24}\left(\frac{c}{T}\right)^4\left[1-\frac{18}{25}\left(\frac{c}{T}\right)\right]\\
&\quad+...-1\}\\
&=0.
\end{aligned} \tag{66}$$

From (62) and (63) and by applying limit $T \to \infty$

$$\begin{aligned} &\lim_{T\to\infty}(\hat{\theta}_{ES1}-\hat{\theta}_{ES3}) \\ &=\left(\frac{2r+1}{c}\right)\lim_{T\to\infty}\{\left(\frac{1}{6}\right)\left(\frac{c}{T}\right)^2\left[1-\left(\frac{c}{T}\right)\right]+\frac{3}{20}\left(\frac{c}{T}\right)^4\left[1-\frac{8}{9}\left(\frac{c}{T}\right)\right]+...\} \\ &=0. \end{aligned} \tag{67}$$

Hence from (66) and (67),

$$\lim_{T\to\infty}\hat{\theta}_{ES1}=\lim_{T\to\infty}\hat{\theta}_{ES2}=\lim_{T\to\infty}\hat{\theta}_{ES3}. \tag{68}$$

The remaining part of the proof is removed since it is comparable to that presented above. ■
The relationship between E-Bayes estimators of $h(t)$ under various loss functions is given in the subsequent theorem.

Theorem 9. Using the priors $\pi_1(a,b),\pi_2(a,b)$ and $\pi_3(a,b)$ under various loss functions, the relationship between E-Bayes estimators of $h(t)$ when 0< c <T is given as

a) under SELF

i) $\hat{h}(t)_{ES3}<\hat{h}(t)_{ES1}<\hat{h}(t)_{ES2}$,

ii) $\lim_{T\to\infty}\hat{h}(t)_{ES1}=\lim_{T\to\infty}\hat{h}(t)_{ES2}=\lim_{T\to\infty}\hat{h}(t)_{ES3}$,

b) under ELF

i) $\hat{h}(t)_{EE3}<\hat{h}(t)_{EE1}<\hat{h}(t)_{EE2}$,

ii) $\lim_{T\to\infty}\hat{h}(t)_{EE1}=\lim_{T\to\infty}\hat{h}(t)_{EE2}=\lim_{T\to\infty}\hat{h}(t)_{EE3}$,

c) under PLF

i) $\hat{h}(t)_{EP3}<\hat{h}(t)_{EP1}<\hat{h}(t)_{EP2}$,

ii) $\lim_{T\to\infty}\hat{h}(t)_{EP1}=\lim_{T\to\infty}\hat{h}(t)_{EP2}=\lim_{T\to\infty}\hat{h}(t)_{EP3}$.

The proof is excluded since it is similar to that of Theorem 8.

The relationship between E-MSE of the E-Bayes estimators of θ under various loss functions is given in the subsequent theorem.

Theorem 10. Using the priors $\pi_1(a,b),\pi_2(a,b)$ and $\pi_3(a,b)$ under various loss functions, the relationship between E-MSE of the E-Bayes estimators of θ when 0< c <T is given as

a) under SELF

i) $E-MSE(\hat{\theta}_{ES3})<E-MSE(\hat{\theta}_{ES1})<E-MSE(\hat{\theta}_{ES2})$,

ii) $\lim_{T\to\infty}E-MSE(\hat{\theta}_{ES1})=\lim_{T\to\infty}E-MSE(\hat{\theta}_{ES2})=\lim_{T\to\infty}E-MSE(\hat{\theta}_{ES3})$.

b) under ELF

i) $E-MSE(\hat{\theta}_{EE3})<E-MSE(\hat{\theta}_{EE1})<E-MSE(\hat{\theta}_{EE2})$,

ii) $\lim_{T\to\infty}E-MSE(\hat{\theta}_{EE1})=\lim_{T\to\infty}E-MSE(\hat{\theta}_{EE2})=\lim_{T\to\infty}E-MSE(\hat{\theta}_{EE3})$.

c) under PLF

i) $E-MSE(\hat{\theta}_{EP3})<E-MSE(\hat{\theta}_{EP1})<E-MSE(\hat{\theta}_{EP2})$,

ii) $\lim_{T\to\infty}E-MSE(\hat{\theta}_{EP1})=\lim_{T\to\infty}E-MSE(\hat{\theta}_{EP2})=\lim_{T\to\infty}E-MSE(\hat{\theta}_{EP3})$.

Proof.

a)

i) From (41) and (43), we have

$$
\begin{aligned}
E - MSE(\hat{\theta}_{ES3}) \quad - \quad & E - MSE(\hat{\theta}_{ES1}) \\
= \quad & \frac{1}{c}\left[\ln\left(\frac{T+c}{T}\right) - \left(\frac{c}{c+T}\right) - \ln\left(\frac{T}{T+c}\right) - \frac{c}{T}\right] \\
= \quad & \frac{1}{c}\left[2\ln\left(\frac{T+c}{T}\right) - \frac{c(2T+c)}{T(c+T)}\right] \\
= \quad & \frac{2}{c}\ln\left(\frac{T+c}{T}\right) - \frac{(2T+c)}{T(c+T)} \\
< \quad & 0.
\end{aligned} \tag{69}
$$

So, we can say that,

$$
E - MSE(\hat{\theta}_{ES3}) < E - MSE(\hat{\theta}_{ES1}). \tag{70}
$$

Now, from (41) and (42), we have

$$
\begin{aligned}
E - MSE(\hat{\theta}_{ES2}) \quad - \quad & E - MSE(\hat{\theta}_{ES1}) \\
= \quad & \frac{c}{2}\frac{1}{T(c+T)} - \frac{1}{c}\left\{\ln\left(\frac{T+c}{T}\right) - \frac{c}{c+T}\right\} \\
= \quad & \frac{c}{2T(c+T)} + \frac{1}{c+T} - \frac{1}{c}\ln\left(\frac{T+c}{T}\right) \\
= \quad & \frac{c+2T}{(c+T)2T} - \frac{1}{c}\left[\frac{c}{T} - \frac{1}{2}\left(\frac{c}{T}\right)^2 + \frac{1}{3}\left(\frac{c}{T}\right)^3 - ...\right] \\
= \quad & \frac{c+2T}{(c+T)2T} - \frac{1}{T} + \frac{1}{2T}\left(\frac{c}{T}\right)\left[1 - \frac{2}{3}\left(\frac{c}{T}\right)\right] \\
& + \frac{1}{4T}\left(\frac{c}{T}\right)^3\left[1 - \frac{4}{5}\left(\frac{c}{T}\right)\right] + ... \\
> \quad & 0.
\end{aligned} \tag{71}
$$

So, we can say that,

$$
E - MSE(\hat{\theta}_{ES2}) > E - MSE(\hat{\theta}_{ES1}). \tag{72}
$$

From (70) and (72),

$$
E - MSE(\hat{\theta}_{ES3}) < E - MSE(\hat{\theta}_{ES1}) < E - MSE(\hat{\theta}_{ES2}). \tag{73}
$$

ii) From (69) and by applying limit $T \to \infty$

$$
\begin{aligned}
\lim_{T\to\infty}(E \quad - \quad & MSE(\hat{\theta}_{ES3}) - E - MSE(\hat{\theta}_{ES1})) \\
= \quad & \lim_{T\to\infty}\left[\frac{2}{c}\ln\left(\frac{T+c}{T}\right) - \frac{(2T+c)}{T(c+T)}\right] \\
= \quad & \lim_{T\to\infty}\left[\frac{2}{T} - \frac{c}{T^2} + \frac{2c^2}{3T^3} - \frac{c^3}{2T^4} + ...\right] \\
& - \lim_{T\to\infty}\frac{2+\frac{c}{T}}{T(\frac{c}{T}+1)} \\
= \quad & 0.
\end{aligned} \tag{74}
$$

From (71) and by applying limit $T \to \infty$

$$
\begin{aligned}
\lim_{T\to\infty}(E \quad - \quad & MSE(\hat{\theta}_{ES2}) - E - MSE(\hat{\theta}_{ES1})) \\
= \quad & \lim_{T\to\infty}\frac{\frac{c}{T}+2}{2T(\frac{c}{T}+1)} - \frac{1}{T} + \frac{c}{2T^2}\left[1 - \frac{2}{3}\left(\frac{c}{T}\right)\right] + \\
& \frac{c^3}{4T^4}\left[1 - \frac{4}{5}\left(\frac{c}{T}\right)\right] + ... \\
= \quad & 0.
\end{aligned} \tag{75}
$$

Hence from (73) and (74),

$$
\lim_{T\to\infty} E - MSE(\hat{\theta}_{ES1}) = \lim_{T\to\infty} E - MSE(\hat{\theta}_{ES2}) = \lim_{T\to\infty} E - MSE(\hat{\theta}_{ES3}). \tag{76}
$$

The remaining part of the proof is removed since it is comparable to that presented above. ■
The relationship between E-MSE of the E-Bayes estimators of $h(t)$ under various loss functions is given in the subsequent theorem.

Theorem 11. Using the priors $\pi_1(a,b), \pi_2(a,b)$ and $\pi_3(a,b)$ under various loss functions, the relationship between E-MSE of the E-Bayes estimators of $h(t)$ when 0< c <T is given as

a) under SELF

i) $E - MSE(\hat{h}(t)_{ES3}) < E - MSE(\hat{h}(t)_{ES1}) < E - MSE(\hat{h}(t)_{ES2})$,

ii) $\lim_{T\to\infty} E - MSE(\hat{h}(t)_{ES1}) = \lim_{T\to\infty} E - MSE(\hat{h}(t)_{ES2}) = \lim_{T\to\infty} E - MSE(\hat{h}(t)_{ES3})$.

b) under ELF

i) $E - MSE(\hat{h}(t)_{EE3}) < E - MSE(\hat{h}(t)_{EE1}) < E - MSE(\hat{h}(t)_{EE2})$,

ii) $\lim_{T\to\infty} E - MSE(\hat{h}(t)_{EE1}) = \lim_{T\to\infty} E - MSE(\hat{h}(t)_{EE2}) = \lim_{T\to\infty} E - MSE(\hat{h}(t)_{EE3})$.

c) under PLF

i) $E - MSE(\hat{h}(t)_{EP3}) < E - MSE(\hat{h}(t)_{EP1}) < E - MSE(\hat{h}(t)_{EP2})$,

ii) $\lim_{T\to\infty} E - MSE(\hat{h}(t)_{EP1}) = \lim_{T\to\infty} E - MSE(\hat{h}(t)_{EP2}) = \lim_{T\to\infty} E - MSE(\hat{h}(t)_{EP3})$.

Table 1: *The AE (first row), MSE (second row) and ACI for MLE, Bayesian and E-Bayesian estimates of θ for real data.*

	n=148				
	r=30	r=60	r=90	r=120	ACI
$\hat{\theta}_{MLE}$	0.0140806 $9.58331 * 10^{-4}$	0.0213817 $5.65431 * 10^{-4}$	0.0302629 $2.45881 * 10^{-4}$	0.0346765 $1.07767 * 10^{-4}$	
$\hat{\theta}_{B1}$	0.014312 $1.0999 * 10^{-5}$	0.021557 $9.3838 * 10^{-6}$	0.0304276 $1.11832 * 10^{-5}$	0.0348179 $1.03929 * 10^{-5}$	(0.0113113, 0.0473139)
$\hat{\theta}_{B2}$	0.0138428 $1.13596 * 10^{-5}$	0.0212007 $9.5389 * 10^{-6}$	0.0300914 $1.13068 * 10^{-5}$	0.0345289 $1.04791 * 10^{-5}$	(0.0111863, 0.0467911)
$\hat{\theta}_{B3}$	0.0145447 $1.10877 * 10^{-5}$	0.0217345 $9.42226 * 10^{-6}$	0.0305953 $1.12139 * 10^{-5}$	0.0349621 $1.04144 * 10^{-5}$	(0.0113736, 0.0475745)
$\hat{\theta}_{ES1}$	0.0143109 $1.09967 * 10^{-5}$	0.0215561 $9.38285 * 10^{-6}$	0.0304265 $1.11823 * 10^{-5}$	0.0348168 $1.03922 * 10^{-5}$	(0.0113114, 0.0473116)
$\hat{\theta}_{ES2}$	0.0143087 $1.09922 * 10^{-5}$	0.0215542 $9.38095 * 10^{-6}$	0.0304243 $1.11805 * 10^{-5}$	0.0348148 $1.0391 * 10^{-5}$	(0.0113117, 0.0473072)
$\hat{\theta}_{ES3}$	0.0143065 $1.09877 * 10^{-5}$	0.0215524 $9.37904 * 10^{-6}$	0.030422 $1.11788 * 10^{-5}$	0.0348127 $1.03897 * 10^{-5}$	(0.011312, 0.0473028)
$\hat{\theta}_{EE1}$	0.0138417 $1.13573 * 10^{-5}$	0.0211998 $9.53794 * 10^{-6}$	0.0300903 $1.13059 * 10^{-5}$	0.0345279 $1.04785 * 10^{-5}$	(0.0111864, 0.0467889)
$\hat{\theta}_{EE2}$	0.0138396 $1.13526 * 10^{-5}$	0.021198 $9.536 * 10^{-6}$	0.0300881 $1.13041 * 10^{-5}$	0.0345258 $1.04772 * 10^{-5}$	(0.0111867, 0.0467845)
$\hat{\theta}_{EE3}$	0.0138374 $1.1348 * 10^{-5}$	0.0211961 $9.53407 * 10^{-6}$	0.0300859 $1.13023 * 10^{-5}$	0.0345238 $1.04759 * 10^{-5}$	(0.011187, 0.0467801)
$\hat{\theta}_{EP1}$	0.0300255 $1.10854 * 10^{-5}$	0.0353436 $9.4213 * 10^{-6}$	0.0443292 $1.1213 * 10^{-5}$	0.0472391 $1.04137 * 10^{-5}$	(0.0113737, 0.0475723)
$\hat{\theta}_{EP2}$	0.0145414 $1.10809 * 10^{-5}$	0.0217316 $9.41939 * 10^{-6}$	0.0305919 $1.12113 * 10^{-5}$	0.0349589 $1.04124 * 10^{-5}$	(0.011374, 0.0475678)
$\hat{\theta}_{EP3}$	0.0145391 $1.10763 * 10^{-5}$	0.0217297 $9.41748 * 10^{-6}$	0.0305896 $1.12095 * 10^{-5}$	0.0349568 $1.04112 * 10^{-5}$	(0.0113743, 0.0475634)

6 Results

In this section, we examine how well the estimators developed in this article performed.

6.1 *Real Data Analysis*

We used the real data set given by [14] for real-life situations indicating the graft survival periods in months of 148 renal transplant patients to examine the performance of the estimators developed in this research. The data were fitted using the Chen distribution, and the p-value and test statistic values for the Kolmogorov-Smirnov test are 0.5844 and 0.0626, respectively. The MLEs for the unknown Chen distribution parameters are calculated to be $\hat{\theta}$ =0.0429 and $\hat{\lambda}$ =0.3863. We generate Type-II censored samples by choosing different values for r (30, 60, 90 and 120). We presume that the shape parameter is always known and equal to its MLE, i.e., $\hat{\lambda}$ = 0.3863. Using the bootstrapping concept, we computed the AE, MSE, E-MSE and 95% average credible interval (ACI) of the estimators and are given in Tables 1 and 2.

Table 2: *The AE (first row), MSE (second row) and ACI for MLE, Bayesian and E-Bayesian estimates of $h(t)$ for real data.*

	n=148				
	r=30	r=60	r=90	r=120	ACI
$\hat{\lambda}_{MLE}$	0.0205438	0.0334508	0.0501238	0.0557942	
	$1.82571 * 10^{-3}$	$1.28701 * 10^{-3}$	$7.69232 * 10^{-4}$	$7.28922 * 10^{-4}$	
$\hat{h}_{B1}$	0.0198053	0.0325953	0.0490964	0.0556378	(0.00534318, 0.0855648)
	$2.10708 * 10^{-5}$	$2.20851 * 10^{-5}$	$3.14922 * 10^{-5}$	$3.08391 * 10^{-5}$	
$\hat{h}_{B2}$	0.019156	0.0320565	0.0485539	0.0551761	(0.00528414, 0.0846193)
	$2.17617 * 10^{-5}$	$2.24501 * 10^{-5}$	$3.18402 * 10^{-5}$	$3.1095 * 10^{-5}$	
$\hat{h}_{B3}$	0.0201274	0.0328636	0.0493669	0.0558682	(0.00537262, 0.0860362)
	$2.12408 * 10^{-5}$	$2.21756 * 10^{-5}$	$3.15787 * 10^{-5}$	$3.09028 * 10^{-5}$	
$\hat{h}_{ES1}$	0.0198039	0.0325939	0.0490946	0.0556361	(0.00534354, 0.0855612)
	$2.10668 * 10^{-5}$	$2.2083 * 10^{-5}$	$3.14898 * 10^{-5}$	$3.08372 * 10^{-5}$	
$\hat{h}_{ES2}$	0.019801	0.0325912	0.0490911	0.0556328	(0.00534425, 0.0855542)
	$2.10586 * 10^{-5}$	$2.2079 * 10^{-5}$	$3.14851 * 10^{-5}$	$3.08335 * 10^{-5}$	
$\hat{h}_{ES3}$	0.0197981	0.0325885	0.0490876	0.0556295	(0.00534496, 0.0855471)
	$2.10505 * 10^{-5}$	$2.20749 * 10^{-5}$	$3.14804 * 10^{-5}$	$3.08298 * 10^{-5}$	
$\hat{h}_{EE1}$	0.0191546	0.0320552	0.0485521	0.0551744	(0.00528449, 0.0846158)
	$2.17575 * 10^{-5}$	$2.2448 * 10^{-5}$	$3.18378 * 10^{-5}$	$3.10931 * 10^{-5}$	
$\hat{h}_{EE2}$	0.0191518	0.0320525	0.0485487	0.0551712	(0.0052852, 0.0846088)
	$2.17491 * 10^{-5}$	$2.24439 * 10^{-5}$	$3.1833 * 10^{-5}$	$3.10894 * 10^{-5}$	
$\hat{h}_{EE3}$	0.0191489	0.0320499	0.0485452	0.0551679	(0.0052859, 0.0846018)
	$2.17407 * 10^{-5}$	$2.24398 * 10^{-5}$	$3.18282 * 10^{-5}$	$3.10856 * 10^{-5}$	
$\hat{h}_{EP1}$	0.0201259	0.0328622	0.0493651	0.0558665	(0.00537298, 0.0860327)
	$2.12367 * 10^{-5}$	$2.21735 * 10^{-5}$	$3.15763 * 10^{-5}$	$3.09009 * 10^{-5}$	
$\hat{h}_{EP2}$	0.020123	0.0328595	0.0493616	0.0558632	(0.0053737, 0.0860256)
	$2.12285 * 10^{-5}$	$2.21695 * 10^{-5}$	$3.15716 * 10^{-5}$	$3.08972 * 10^{-5}$	
$\hat{h}_{EP3}$	0.02012	0.0328567	0.049358	0.0558599	(0.00537441, 0.0860184)
	$2.12203 * 10^{-5}$	$2.21654 * 10^{-5}$	$3.15669 * 10^{-5}$	$3.08935 * 10^{-5}$	

From Tables 1 and 2, it is to be noted that the approximated MSEs decrease as r increases. We can deduce from Tables that E-Bayesian estimators outperform MLE and Bayesian estimators in terms of MSE.

7 Discussion and Concluding Remark

The parameter and hazard rate of the Chen distribution based on Type II censoring are estimated using the MLE, Bayesian, and E-Bayesian approaches. The estimates are computed using real data, and various estimation techniques are compared. One of the study's key findings is the superiority of the proposed estimators versus existing estimators. The impact of various prior distributions and loss functions is also something we theoretically investigate. Important concluding remarks from our study are listed below:

- The lowest MSE of all the estimates is seen in the E-Bayesian estimations of θ.
- The lowest E-MSE among all estimates is found in the E-Bayesian estimations of θ based on ELF with prior distribution $\pi_3(a, b)$.
- For a fixed value of n and r the E-MSE is less for E-Bayesian estimators as compared to Bayesian and MLE.
- Compared to Bayesian and MLE, the proposed estimators perform better in terms of minimum MSE.

Combining the findings mentioned above, we recommended the E-Bayesian technique, which outperforms previous estimates in terms of minimum MSE, to estimate the scale parameter and hazard rate functions of the Chen distribution based on the type-II censoring scheme using prior distribution $\pi_3(a, b)$.

References

[1] Abdul Sathar E.I., AthiraKrishnan R.B., E-Bayesian and hierarchical Bayesian estimation for the shape parameter and reversed hazard rate of power function distribution under different loss functions, *Journal of the Indian Society for Probability and Statistics*, 20(2):227–253, 2019.

[2] Essam A.Ahmed., Bayesian estimation based on progressive Type-II censoring from two-parameter bathtub-shaped lifetime model: an Markov chain Monte Carlo approach, Journal of Applied Statistics, 41(4):752–768, 2014.

[3] Ali Algarni, Abdullah M Almarashi, Hassan Okasha, and Hon Ke-ung Tony Ng. E-bayesian estimation of chen distribution based on Type-I censoring scheme, Entropy, 22(6):636, 2020.

[4] Abdullah M Almarashi and Gamal A Abd-Elmougod, Accelerated competing risks model from Gompertz lifetime distributions with Type-II censoring scheme, Thermal Science, 24(Suppl. 1):165–175,2020.

[5] Athirakrishnan R.B. and Abdul Sathar E.I., E-Bayesian and hierarchical Bayesian estimation of inverse Rayleigh distribution, American Journal of Mathematical and Management Sciences, 41(1):70–87, 2022.

[6] James O Berger, Statistical decision theory and Bayesian analysis Springer Series in Statistics, New York., 1985.

[7] Ajit Chaturvedi and Taruna Kumari, Estimation and comparison of the stress-strength models with more than two states under Weibull distribution and Type-II censoring scheme, Communications in Statistics-Theory and Methods, 48(3):537–548, 2019.

[8] Zhenmin Chen, A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function, Statistics and Probability Letters, 49(2):155–161, 2000.

[9] Rajni Goel and Hare Krishna, Likelihood and Bayesian inference for k lindley populations under joint Type-II censoring scheme, Communications in Statistics-Simulation and Computation, pages 1–16, 2021.

[10] Rajni Goel and Hare Krishna, Statistical inference for two lindley populations under balanced joint progressive Type-II censoring scheme, Computational Statistics, pages 1–24, 2021.

[11] Ming 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.

[12] Ming Han, The E-Bayesian estimation and its E-MSE of Pareto distribution parameter under different loss functions, Journal of Statistical Computation and Simulation, 90(10):1834–1848, 2020.

[13] Ming Han and Yuanyao Ding, Synthesized expected Bayesian method of parametric estimate, Journal of Systems Science and Systems Engineering, 13(1):98–111, 2004.

[14] DJ Hand, F Daly, AD Lunn, KJ McConway, and E Ostrowski, Small data sets, 1994.

[15] Tanmay Kayal, Yogesh Mani Tripathi, Sanku Dey, and Shuo-Jye Wu, On estimating the reliability in a multicomponent stress-strength model based on Chen distribution, Communications in Statistics-Theory and Methods, 49(10):2429–2447, 2020.

[16] Tanmay Kayal, Yogesh Mani Tripathi, Debasis Kundu, and Manoj Kumar Rastogi, Statistical inference of Chen distribution based on Type-I progressive hybrid censored samples, Statistics, Optimization and Information Computing, 10(2):627–642, 2022.

[17] Tanmay Kayal, Yogesh Mani Tripathi, Devendra Pratap Singh, and Manoj Kumar Rastogi, Estimation and prediction for Chen distribution with bathtub shape under progressive censoring, Journal of Statistical Computation and Simulation, 87(2):348–366, 2017.

[18] CD Lai, Min Xie, and DNP Murthy, A modified Weibull distribution, IEEE Transactions on reliability, 52(1):33–37, 2003.

[19] Albert W Marshall and Ingram Olkin, A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, Biometrika, 84(3):641–652, 1997.

[20] Hassan Okasha, Mazen Nassar, and Saeed A Dobbah, E-Bayesian estimation of burr Type-XII model based on adaptive Type-II progressive hybrid censored data, AIMS Mathematics, 6(4):4173–96, 2021.

[21] Sowbhagya S Prabhu, E-Bayesian estimation of the shape parameter of Lomax model under symmetric and asymmetric loss functions, International Journal of Statistics and Applied Mathematics, 5(6):142–146, 2020.

[22] Manoj Kumar Rastogi, Yogesh Mani Tripathi, and Shuo Jye Wu, Estimating the parameters of a bathtub-shaped distribution under progressive Type-II censoring, Journal of Applied Statistics, 39(11):2389–2411, 2012.

[23] KR Renjini, EI Abdul Sathar, and G Rajesh, Bayes estimation of dynamic cumulative residual entropy for Pareto distribution under Type-II right censored data, Applied Mathematical Modelling, 40(19-20):8424–8434, 2016.

[24] Shuo-Jye Wu, Estimation of the two-parameter bathtub-shaped lifetime distribution with progressive censoring, Journal of Applied Statistics, 35(10):1139–1150, 2008.

NEW MEDIAN BASED ALMOST UNBIASED EXPONENTIAL TYPE RATIO ESTIMATORS IN THE ABSENCE OF AUXILIARY VARIABLE

Sajad Hussain*, Vilayat Ali Bhat[1]

•

*University School of Business, Chandigrah University, Mohali, Gharuan - 140413, INDIA
[1]Department of Statistics, Pondicherry University, Puducherry - 605014, INDIA
sajad.stat321@gmail.com, vilayat.stat@gmail.com

Abstract

The problem of biasness and availability of auxiliary variable for the estimating population mean is a big concern, both can be handled by proposing unbiased estimators in the absence of auxiliary variable. So in this paper unbiased exponential type estimators of population mean have been proposed. The estimators are proposed in the absence of the instrumental variable called the auxiliary variable by taking the advantage of the population and the sample median of the study variable. To about the first order approximation, the theoretical formulations of the bias and mean square error (MSE) *are obtained. The circumstances in which the suggested estimators have the lowest mean squared error values when compared to the existing estimators were also deduced. In comparison to the currently used estimators, it was discovered that the suggested estimators of population mean had the lowest MSE, hence highest efficiency. Also least influence from the data's influential observations when it came to accurately calculating the population mean for skewed data. The theoretical findings of the paper are validated by the numerical study.*

Keywords: Dual estimator, Median, Study variable, Unbiased Estimator, Mean square error

1. Introduction

In sampling, a representative part of the population called the sample is studied to determine the parameters or the characteristics of the population like the mean, variance, median, correlation coefficient etc. The sample mean estimator is a good device to find the approximate value of the population mean. The population mean can be estimated more precisely by using the auxiliary information as the pioneer work of Cochran [2] proposed a ratio estimator which is more precise than the sample mean estimator when the study and auxiliary variable are positively correlated, though this estimator is biased. Using auxiliary information which is often obtained at an extra survey cost, the authors such as Sisodia and Dwivedi [9], Yadav and Kadilar [10], Ekpenyong and Enang [15] etc. proposed modified ratio estimators which are more efficient than the sample mean and the classical ratio estimator. The pioneering work of Subramani [6] offered a precise ratio estimator to estimate the mean of skewed population without the usage of auxiliary variable by leveraging the median of main variable as the aforementioned estimators are inapplicable in the absence of the auxiliary variable. The median of the study variable may easily be available without having exact information on every data point (See Subramani [6] and these median based estimators are robust in nature since outliers have the least impact on them. The exponential ratio estimator of population mean introduced by Bahl and Tuteja [1] can be applied even when correlations are weak. Later, to be more accurate than the traditional exponential ratio estimator, Singh *et.al* [12], Yasmeen *et al.* [8], Zaman and Kadilar [4], Hussain *et al.* [3] and several others suggested various modified exponential ratio type estimators.

The estimators discussed above are all biassed, which could cause the population mean to be overestimated or underestimated. As a result, the authors Singh *et al.* [13], Yadav *et al.* [11], Singh *et al.* [14] etc. proposed almost unbiased estimators of population mean in presence of auxiliary variable. The auxiliary variable may not be always avaliable, therefore in the absence of auxiliary variable present study is carried to propose high precision exponential type estimators of population mean which may be unbiased in nature and also able to handle the voluminous data influenced by outliers.

2. Material and Methods

Assume that a simple random sampling without replacement (SRSWOR) sampling strategy is used to select a random sample of size n from a finite population of N number of units. The goal of the study is to estimate the population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ without the use of auxiliary variable information. Assume that data on the study variable Y's correlation with the auxiliary variable X is accessible for every member of the population. The notations and formula used in the paper are as follows

Study Variable

"$C_y = \frac{S_y}{\bar{Y}}$ is the coefficient of variation .
$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ is the population mean square.
$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ is the sample mean square.

Auxiliary Variable

: $C_x = \frac{S_x}{\bar{X}}$ is the coefficient of variation.
: $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$ is the population mean square.
: $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is the sample mean square.

Further,
M (m) are the population (sample) median of the main variable.
$\bar{M} = \frac{1}{^NC_n}\sum_{i=1}^{^NC_n} m_i$ is the average of sample medians of the main variable.
m_i is the sample median of i^{th} sample $(i = 1, 2, ..., ^N C_n)$.
^NC_n is the number of samples of size n from N.
The study variable dataset is skewed, M $\neq \bar{Y}$.
$C_m = \frac{S_m}{M}$, $C_{ym} = \frac{S_{ym}}{\bar{Y}M}$, $S_{ym} = \frac{1}{^NC_n}\sum_{i=1}^{^NC_n}(\bar{y}_i - \bar{Y})(m_i - M)$, $S_m^2 = \frac{1}{^NC_n}\sum_{i=1}^{^NC_n}(m_i - M)^2$.
$\rho = \frac{cov(x,y)}{S_x S_y}$, is the population correlation coefficient between X and Y.
$\gamma = \frac{1-f}{n}$, where the sampling fraction $f = \frac{n}{N}$.
$\theta = \frac{r\bar{X}}{2(r\bar{X}+s)}$".

3. Examining Current Ratio Type Estimators

The sample mean estimator is the fundamental estimator of the population mean without the usage of an auxiliary variable as

$$t_1 = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

Bias and MSE of the estimator t_1 up to $O(n)^{-1}$ are as

$$Bias(t_1) = 0. \tag{1}$$

$$MSE(t_1) = \gamma \bar{Y}^2 C_y^2. \tag{2}$$

Making the use of auxiliary variable Cochran [2] proposed a ratio estimator which is more efficient than the estimator t_1, if $\frac{C_x}{2C_y} < \rho \leq +1$ as

$$t_2 = \bar{y}\frac{\bar{X}}{\bar{x}}.$$

The Bias and MSE expressions of the estimator t_2 up to $O(n)^{-1}$ are as

$$Bias(t_2) = \gamma\bar{Y}(C_x^2 - C_{yx}). \tag{3}$$

$$MSE(t_2) = \gamma\bar{Y}^2(C_y^2 + C_x^2 - 2C_{yx}). \tag{4}$$

According to Bahl and Tuteja [1] proposal, an effective exponential ratio type estimator may be applied even when X and Y have weak correlations.

$$t_3 = \bar{y}\exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right).$$

The estimator t_3 is more efficient than the sample mean estimator t_1, if $\frac{1}{4} < \rho_{xy}\frac{C_y}{C_x} < \frac{3}{4}$. The Bias and MSE up to $O(n)^{-1}$ are as

$$Bias(t_3) = \gamma\bar{Y}\left(\frac{3}{8}C_x^2 - \frac{1}{2}C_{yx}\right). \tag{5}$$

$$MSE(t_3) = \gamma\bar{Y}^2\left(C_y^2 + \frac{C_x^2}{4} - C_{yx}\right). \tag{6}$$

A family of modified exponential ratio estimators was presented by Singh et al. [12] employing various well-known auxiliary variable characteristics, such as correlation coefficient, coefficient of variation, skewness, etc. as

$$t_4 = \bar{y}\exp\left[\frac{(r\bar{X} + s) - (r\bar{x} - s)}{(r\bar{X} + s) + (r\bar{x} - s)}\right].$$

The Bias and MSE expressions of up to $O(n)^{-1}$ for the estimator t_4 are as

$$Bias(t_4) = \gamma\bar{Y}(\theta^2 C_x^2 - \theta C_{yx}). \tag{7}$$

$$MSE(t_4) = \gamma\bar{Y}^2(C_y^2 + \theta^2 C_x^2 - 2\theta C_{yx}). \tag{8}$$

A median-based ratio type estimator without the usage of an auxiliary variable was proposed by Subramani [6] as

$$t_5 = \bar{y}\left[\frac{M}{m}\right].$$

The estimator t_5 is biased with the expressions of Bias and MSE up to $O(n)^{-1}$ as

$$Bias(t_5) = \gamma\bar{Y}\left(C_m^2 - C_{ym} - \frac{Bias(m)}{\gamma M}\right). \tag{9}$$

$$MSE(t_5) = \gamma\bar{Y}^2\left(C_y^2 + \frac{\bar{Y}^2}{M^2}C_m^2 - 2\frac{\bar{Y}}{M}C_{ym}\right). \tag{10}$$

The sample mean estimators is biased while as all other estimators viz. Cochran [2], Bahl and Tuteja [1], Singh *et al.* [12] and Subramani [6] discussed above are biased.

4. Proposed Exponential Ratio Estimators

In the absence of auxiliary variable, the proposed exponential ratio type estimators of the study are as

$$t_{ue1} = \bar{y}\left[\alpha\exp\left(\frac{M - m}{aM}\right) + (1-\alpha)\exp\left(\frac{m - M}{aM}\right)\right].$$

$$t_{ue2} = \bar{y}\left[\beta\exp\left(\frac{M - m}{bm}\right) + (1-\beta)\exp\left(\frac{m - M}{bm}\right)\right].$$

The value of non zero constants a and b is chosen such that the estimators t_{ue1} and t_{ue2} are unbiased and the value of constants α and β are chosen such that the MSE of t_{ue1} and t_{ue2} should be minimum.
Consider,

$$\bar{y} = \bar{Y}(1+e_0) \quad \text{and} \quad m = M(1+e_1)$$

Therefore,

$$E(e_0) = 0,\;\; E(e_1) = \frac{Bias(m)}{M}$$

$$E(e_0^2) = \gamma C_y^2; \quad E(e_1^2) = \gamma C_m^2; \quad E(e_0e_1) = \gamma C_{ym}$$

Transforming the estimator t_{ue1} and t_{ue2} in terms of $e_i(i = 0, 1)$, the equations obtained are as

$$t_{ue1} = \bar{Y}(1+e_0)\left[\alpha \exp\left(\frac{-e_1}{a}\right) + (1-\alpha)\exp\left(\frac{e_1}{a}\right)\right]. \tag{11}$$

$$t_{ue2} = \bar{Y}(1+e_0)\left[\beta \exp\left(\frac{-e_1}{b(1+e_1)}\right) + (1-\beta)\exp\left(\frac{e_1}{b(1+e_1)}\right)\right]. \tag{12}$$

Solving the equations (11) & (12) and retaining the terms only up to up 2^{nd} degree, the reduced equations are as

$$t_{ue1} = \bar{Y}\left[1 + e_0 + (1-2\alpha)\frac{e_1}{a} + \frac{e_1^2}{2a^2} + (1-2\alpha)\frac{e_0e_1}{a}\right].$$

$$\Rightarrow t_{ue1} - \bar{Y} = \bar{Y}\left[e_0 + (1-2\alpha)\frac{e_1}{a} + \frac{e_1^2}{2a^2} + (1-2\alpha)\frac{e_0e_1}{a}\right] \tag{13}$$

$$t_{ue2} = \bar{Y}\left[1 + e_0 + (1-2\beta)\frac{e_1}{b} + \left(\frac{1}{2b} + 2\beta - 1\right)\frac{e_1^2}{b} + (1-2\beta)\frac{e_0e_1}{b}\right].$$

$$\Rightarrow t_{ue2} - \bar{Y} = \bar{Y}\left[e_0 + (1-2\beta)\frac{e_1}{b} + \left(\frac{1}{2b} + 2\beta - 1\right)\frac{e_1^2}{b} + (1-2\beta)\frac{e_0e_1}{b}\right] \tag{14}$$

The bias of the estimators t_{ue1} and t_{ue2} is obtained by taking expectation on both sides of (13) and (14) as

$$Bias(t_{ue1}) = \gamma\bar{Y}\left[\frac{1}{2a^2}C_m^2 + \frac{1}{a}(1-2\alpha)\left(C_{ym} + \frac{Bias(m)}{\gamma M}\right)\right]. \tag{15}$$

$$Bias(t_{ue2}) = \gamma\bar{Y}\frac{1}{b}\left[\left(\frac{1}{2b} + 2\beta - 1\right)C_m^2 + (1-2\beta)\left(C_{ym} + \frac{Bias(m)}{\gamma M}\right)\right]. \tag{16}$$

Taking the expectation of the square of the equations (13) & (14), the mean square error of t_{de1} and t_{de2} is obtained as

$$MSE(t_{ue1}) = \gamma\bar{Y}^2\left[C_y^2 + (1-2\alpha)^2\frac{C_m^2}{a^2} + \frac{2}{a}(1-2\alpha)C_{ym}\right]. \tag{17}$$

$$MSE(t_{ue2}) = \gamma\bar{Y}^2\left[C_y^2 + (1-2\beta)^2\frac{C_m^2}{b^2} + \frac{2}{b}(1-2\beta)C_{ym}\right]. \tag{18}$$

The estimator t_{ue1} is unbiased, if

$$a = \frac{C_m^2}{2(2\alpha-1)\left(C_{ym} + \frac{Bias(m)}{\gamma M}\right)}. \tag{19}$$

Whereas the estimator t_{ue2} is unbiased, if

$$b = \frac{C_m^2}{2(2\beta-1)\left(\frac{Bias(m)}{\gamma M} + C_{ym} - C_m^2\right)}. \tag{20}$$

Substituting the values of (19) and (20) in equations (7) and (8) respectively, the following equations are obtained as

$$MSE(t_{ue1}) = \gamma\bar{Y}^2\left[C_y^2 + 4(1-2\alpha)^4\frac{\left(C_{ym}+\frac{Bias(m)}{\gamma M}\right)^2}{C_m^2} - 4(1-2\alpha)^2\frac{C_{ym}\left(C_{ym}+\frac{Bias(m)}{\gamma M}\right)}{C_m^2}\right]. \quad (21)$$

$$MSE(t_{ue2}) = \gamma\bar{Y}^2\left[C_y^2 + 4(1-2\beta)^4\frac{\left(C_{ym}+\frac{Bias(m)}{\gamma M}-C_m^2\right)^2}{C_m^2} - 4(1-2\beta)^2\frac{\left(C_{ym}+\frac{Bias(m)}{\gamma M}-C_m^2\right)}{C_m^2}C_{ym}\right]. \quad (22)$$

Differentiating equations (21) and (22) with respect to α and β respectively and equating to zero, the optimum value of α and β is obtained as

$$\alpha = \frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{C_{ym}}{2\left(C_{ym}+\frac{Bias(m)}{\gamma M}\right)}} \quad \text{and} \quad \beta = \frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{C_{ym}}{2\left(C_{ym}+\frac{Bias(m)}{\gamma M}-C_m^2\right)}}.$$

The constants a, b, α and β contain the unknowns C_m and C_{ym} whose value is considered to be known well in advance and if unknown, they can be determined from past surveys, experience carried by the researcher in the due course of time or from the pilot survey (See Srivenkataramana & Tracy [17], Singh & Kumar [16] and the references cited therein). Now on using the value of α and β in equation (21) and (22) respectively, the minimum value of MSE of the estimator t_{ue1} and t_{ue2} up to $O(n)^{-1}$ is obtained as

$$MSE_{min}(t_{uei}) = \gamma\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right]. \quad (23)$$

5. Theoretical Efficiency Comparisons

From equations (2), (4), (6), (8), (10) and (23), the circumstances and conditions in which the suggested estimators outperform the sample mean estimator, the existing estimators of Cochran [2], Bahl and Tuteja [1], Singh *et al.* [12] and Subramani [6] are obtained as

$$MSE_{min}(t_{uei}) < MSE(t_1)$$
$$\Rightarrow \gamma\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right] < \gamma\bar{Y}^2C_y^2, \text{ if } C_{ym}^2 > 0. \quad (24)$$

$$MSE_{min}(t_{uei}) < MSE(t_2)$$
$$\Rightarrow \gamma\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right] < \gamma\bar{Y}^2(C_y^2 + C_x^2 - 2C_{yx}), \text{ if } C_{ym}^2 > C_m^2(2C_{yx} - C_x^2). \quad (25)$$

$$MSE_{min}(t_{uei}) < MSE(t_3)$$
$$\Rightarrow \gamma\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right] < \gamma\bar{Y}^2\left(C_y^2 + \frac{C_x^2}{4} - C_{yx}\right), \text{ if } C_{ym}^2 > \frac{C_m^2}{4}(4C_{yx} - C_x^2). \quad (26)$$

$$MSE_{min}(t_{uei}) < MSE(t_4)$$
$$\Rightarrow \gamma\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right] < \gamma\bar{Y}^2(C_y^2 + \theta^2C_x^2 - 2\theta C_{yx}), \text{ if } C_{ym}^2 > C_m^2(2\theta C_{yx} - \theta^2C_x^2). \quad (27)$$

$$MSE_{min}(t_{uei}) < MSE(t_5)$$
$$\Rightarrow \gamma\bar{Y}^2\left[C_y^2 - \frac{C_{ym}^2}{C_m^2}\right] < \gamma\bar{Y}^2\left(C_y^2 + \frac{\bar{Y}^2}{M^2}C_m^2 - 2\frac{\bar{Y}}{M}C_{ym}\right), \text{ if } C_{ym}^2 > C_m^2\left(2\frac{\bar{Y}}{M}C_{ym} - \frac{\bar{Y}^2}{M^2}C_m^2\right). \quad (28)$$

The proposed median based unbiased exponential ratio type estimators t_{uei} are more precise than the estimators t_1 to t_6 under the conditions (24) to (28).

6. Numerical Study Comparisons

For the numerical study, data of populations P1 and P2 containing influential observations have been considered as given in Table-1. The data set of population P1 is sourced from Singh and Chaudhary [7] where the study variable is to estimate area of wheat under cultivation in the year 1974 and the auxiliary variable is the cultivated area under wheat in the year 1971. The data set of population P2 is taken from Mukhopadhyay [5] where the study variable is to estimate the ammount of raw materials for 20 jute mills and the auxiliary variable is the number of workers.

Table 1: ***Summary statistics of the population data sets.***

Data constants	Populations	
	P1	P2
N	34	20
n	5	5
${}^{N}C_n$	278256	15504
$\bar{Y}$	856.4118	41.50000
$\bar{M}$	736.9811	40.05520
M	767.5000	40.50000
$\bar{X}$	208.8824	441.9500
$\frac{\bar{Y}}{\bar{M}}$	1.115800	1.024700
C_y^2	0.125014	0.008338
C_x^2	0.088563	0.007845
C_m^2	0.100833	0.006606
C_{ym}	0.073140	0.005394
C_{yx}	0.047257	0.005275
ρ_{yx}	0.449100	0.652200

It can be observed from Table-1 that a sample of 5 units has been drawn from two populations P1 and P2 having size 34 and 20 respectively. The values of different parameters like population mean, population median, coefficient of variation etc. of the study and auxiliary variable are obtained and can be seen from the table.

Table 2: ***MSE, Bias and PRE of the estimators t_1, t_2, t_3, t_4, t_5 and t_{uei}.***

Estimator	**Population**					
	P1			**P2**		
	MSE	**| Bias |**	**PRE**	**MSE**	**| Bias |**	**PRE**
t_1	15641.306	0.000	100.000	2.154	0.000	100.000
t_2	14896.738	6.035	104.998	1.455	0.016	148.041
t_3	12498.850	1.399	125.142	1.297	0.002	166.076
t_4	12499.093	0.227	125.139	1.298	0.004	165.948
t_5	10926.773	38.100	143.147	1.090	0.463	197.615
t_{uei}	9003.545	0.000	173.724	1.016	0.000	212.007

The suggested median based exponential ratio estimators t_{uei} $(i = 1, 2)$ have the lowest MSE values for both population data sets P1 and P2, as can be seen from Table 2. When comparing with the sample mean estimators, estimators of Cochran [2], Bahl and Tuteja [1], Singh *et al.* [12], and Subramani [6], the PRE of the suggested estimators is found highest. Furthermore, as the estimators suggested in the paper are unbiased, they may be used to address the issue of under or overestimating the population mean.

7. Discussion

The paper presents two estimators of population mean as t_{ue1} and t_{ue2} based on population median. Since median is a type of parameter which is least influenced by the effect of outliers, so the proposed estimators may work efficiently for skewed data as evident from numerical study. It can be observed from empirical study that the MSE value for t_{uei} is 9003.545 and 1.016 for populations P1 and P2 respectively which can be observed as minimum value among all other estimators considered for comparision. The minimu value of MSE highlights that the estimators t_{uei} are most efficient. Further looking at the bias values, t_{uei} can be found as unbiased so will take care of under or over estimation problem . The population and sample median used in the construction of estimators t_{uei} are of study variable only, so do have a good advantage as the auxiliary variable may not be always available

8. Conclusion

- In the absence of an auxiliary variable, the suggested median-based almost-unbiased exponential ratio estimators of population mean are as follows

$$t_{ue1} = \bar{y}\left[\alpha \exp\left(\frac{M-m}{aM}\right) + (1-\alpha)\exp\left(\frac{m-M}{aM}\right)\right].$$

$$t_{ue2} = \bar{y}\left[\beta \exp\left(\frac{M-m}{bm}\right) + (1-\beta)\exp\left(\frac{m-M}{bm}\right)\right].$$

- The proposed estimators t_{ue1} and t_{ue2} are median based and therefore have least influence of outliers present in the data set.
- The estimators proposed in the study are more precise for a skewed data set than the estimators considered.

References

[1] Bahl, S. and Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Information and Optimization Sciences*, 12(1), 159–163.
[2] Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*, 30, 262–275.
[3] Hussain, S., Sharma, M. and Bhat, M. I. J. (2021). Optimum exponential ratio type estimators for estimating the population mean. *Journal of Statistics Applications and Probability Letters*, 8(2), 73–82.
[4] Zaman, T. and Kadilar, C. (2019). Novel family of exponential estimators using information of auxiliary attribute. *Journal of Statistics and Management Systems*, 22 (8),1499–1509.
[5] Mukhopadhyay, P. (2005). *Theory and methods of survey sampling*, PHI Learning, 2nd edition, New Delhi.
[6] Subramani, J. (2016). A new median based ratio estimator for estimation of the finite population mean, *Statistics in Transition New Series*, 17 (4): 1–14.
[7] Singh, D and Chaudhary, F.S. (1986). *Theory and analysis of sample survey designs*. New Age International Publisher, New Delhi.

[8] Yasmeen, U., Noor ul Amin, M., Hanif, M., (2016). Exponential ratio and product type estimators of population mean, *Journal of Statistics and Management System*, 19(1), pp. 55–71.

[9] Sisodia BVS, Dwivedi VK. A modified ratio estimator using coefficient of variation of auxiliary variable. *Journal of the Indian Society Agricultural Statistics*. 1981; 33(2): 13–18.

[10] Yadav SK, Kadilar C. Improved Class of Ratio and Product Estimators. *Applied Mathematics and Computation*. 2013; 219 (22): 10726–10731.

[11] Yadav R, Upadhyaya LN, Singh HP, Chatterjee S. Almost unbiased ratio and product type exponential estimators. *Statistics in Transition new series*. 2012; 13(3): 537–550.

[12] R. Singh, P. Chauhan, N. Sawan and F. Smarandache, Improvement in Estimating the Population Mean Using Exponential Estimators in Simple Random Sampling. *Bulletin of Statistics and Economics*. 3 (A09), 13–18 (2009).

[13] Singh, R.,Kumar, M., and Smarandache,F. (2008): Almost unbiased estimator for estimating population mean using known value of some population parameter(s). *Pak.j.stat.oper.res.* Vol.IV No.2 2008 pp63–76.

[14] Singh, R., Gupta, S.B. and Malik, S., 2016. Almost Unbiased Estimator Using Known Value of Population Parameter (s) in Sample Surveys. *Journal of Modern Applied Statistical Methods*, 15(1), p.30.

[15] Ekpenyong, E.J and Enang, E.I. 2015. A modified class of ratio and product estimators of population mean in simple random sampling using information on auxiliary variable. *Journal of statistics*, 22: 1–8.

[16] Singh HP, Kumar S. A general family of estimators of finite population ratio, product and mean using two phase sampling scheme in the presence of non-response. *Journal of Statistical Theory and Practice*. 2008; 2(4):677–692.

[17] Srivenkataramana. T and D.S Tracy (1980), An alternative to ratio method in sample surveys. *Annals of the Institute of Statistical Mathematics*. 32:111–120.

PERFORMABILITY OPTIMISATION OF MULTISTATE COAL HANDLING SYSTEM OF A THERMAL POWER PLANT HAVING SUBSYSTEMS DEPENDENCIES USING PSO AND COMPARATIVE STUDY BY PETRI NETS

Er. Sudhir Kumar, Dr. P.C. Tewari

•

Department of Production & Industrial Engineering
National Institute of Technology, Kurukshetra, India
sudhirtamak@gmail.com, pctewari1@gmail.com

Abstract

This paper deals with an analysis methodology for evaluating the performance of a coal handling system utilized in a coal based thermal power plant. To simulate the interactions between the subsystems, a stochastic Petri nets technique is used. A licensed software package named Petri module of GRIF were used for computations. This work addresses the performability and cost multi-objective optimization problem for a series-parallel coal handling system of a thermal power plant having subsystem failure dependencies. Performability of subsystems has been examined in relation to variations in failure and repair rates. The Particle Swarm Optimization Technique, which is based on an algorithm discussed, has been used to optimize the results. Based upon the observation and criticality of failure, the subsystems of the coal handling system were given maintenance order priority. A decision support provided at last which will the maintenance personnel™s to take better and informed decision while forming the maintenance policies. It has been observed that the Crusher and Tippler are crucial components that demand the full attention of plant manager.

Keywords: Petri Nets, Particle Swarm Optimization, Performability, Decision Support System

1. Introduction

Electricity consumption has significantly increased in India as a result of the country's fast development. The Thermal Power Plant (TPP) is one of the primary sources of power generating. The high availability of key components of equipment is necessary for continuous electricity generation. Reliability and maintainability of the used equipment, subsystem, and system are factors that affect thermal power plant availability to the utmost extent[1]. Keeping the systems of thermal power plant in operational state is a big challenge for the maintenance personnels. Unfortunately, system failure cannot be completely avoided, but it can be reduced to the absolute minimum[2]. To increase plant efficiency, the performability of subsystems TPP's must be evaluated. In this study, Stochastic Petri Nets (SPN) approach is used to examine the performance evaluation of TPP. The design of any system in the context of performability presents numerous challenges, and the objectives may be conflicting at a time. System reliability, availability, maintainability, safety, and cost are all prime aspects of system performability[3, 4, 5, 6]. To keep systems in a working state over the course of several decades, maintenance scheduling was crucial. The best maintenance plan takes care of TPP's maintenance requirements for the least cost[7]. The maintenance cost is influenced by both scheduled and unscheduled maintenance tasks. Due to this, research have been done in the past to decrease system downtime, which has improved plant availability. The field has slowly expanded to include more systems that schedule system maintenance using condition-based maintenance strategies. Various attempts have been undertaken in the past to identify and prioritise the TPP's important subsystems[8]. Furthermore, the field of power generation has widely embraced condition monitoring tools. It makes it easier to identify and fix major equipment flaws so that working equipment can be restored as soon as possible[9, 10, 11]. For a coal-fired thermal power plant, earlier scholars attempted a number of different attempts at RAM analysis. It's important to assess TPP's performance under actual circumstances. He [12] has proposed a method using the Petri Net approach for reliability evaluation and safety assessment of an industrial manufacturing system. Using Petri nets, Sachdeva et al. [13, 14, 15, 16, 17] assessed the pulping system's availability in the paper sector. The study's main aim was to determine the overall coast of operation and repair. In their investigations, it was explored how different parameters affected the system's availability. Based on the Petri Nets model's Mante Carlo simulation, the various reliability parameters were calculated. The uses of the Petri nets model and its capacity to simulate the real world were presented by Schneeweiss [18]. It has also been demonstrated how PN models can be used to analyse non-repairable systems. .

2. System Description

Coal handling system is a crucial element of a coal-based thermal power plant. The continuous operation of the thermal power plant and consequently the production of power will be ensured by the system's smooth operation. Tipplers are used to unload the coal and send it to the storage yard after it has been transported to the plant via a variety of routes, including rail, road, and water. The following description applies to the numerous subsystems of the coal handling system also represented by fig. 1, which is configured in a hybrid mode.

- Wagon Tippler: Tipplers are used to tip laden waggons to empty them of their contents. Tippler secures the waggon from the side and the top using gripping devices that are built into the waggon. In addition, waggon tippler features include wheel grippers, limit switches of various types, and track stops.
- Crusher: The bigger coal rocks are broken down into smaller coal rocks, gravel, and rock dust using a crusher. The coal is crushed or compressed in coal crushers using metallic surfaces.
- Bunker: Bunkers are utilised as a depot to guarantee a steady and smooth supply of coal to the mill.
- Feeder: Feeders are used to feed coal to coal mills, primarily ensuring a smooth and constant supply of coal in accordance with boiler demand.
- Coal Mills: For further combustion in the furnace, the coarse coal is ground into a fine powder in a coal mill

The above-mentioned subsystems are conned in the hybrid mode of configuration i.e., the combination of series and parallel configuration. The systems become more complex as it has the dependencies on the performance of all subsystems mentioned. The overall performance thus depends upon the failure and repair rate of all subsystems. Thus, it is necessary to optimize these failure and rapier rates so that the overall performance can be enhanced in terms of performability. Based upon the severity of failure and performance, the critical subsystems can be identified, which will be further provided with better maintenance policies and procedures. It will the maintenance engineers to make better strategies for repairs for the overall system.

3. Petri Nets Modelling

Petri Nets Modelling of Coal Handling System: The performability and long-term availability, of the various subsystems of the Coal Handling System of a coal-operated thermal power plant, has been assessed in this section. The numerous input values, including FRR for different subsystems, were taken from repair and maintenance manuals with the help of maintenance

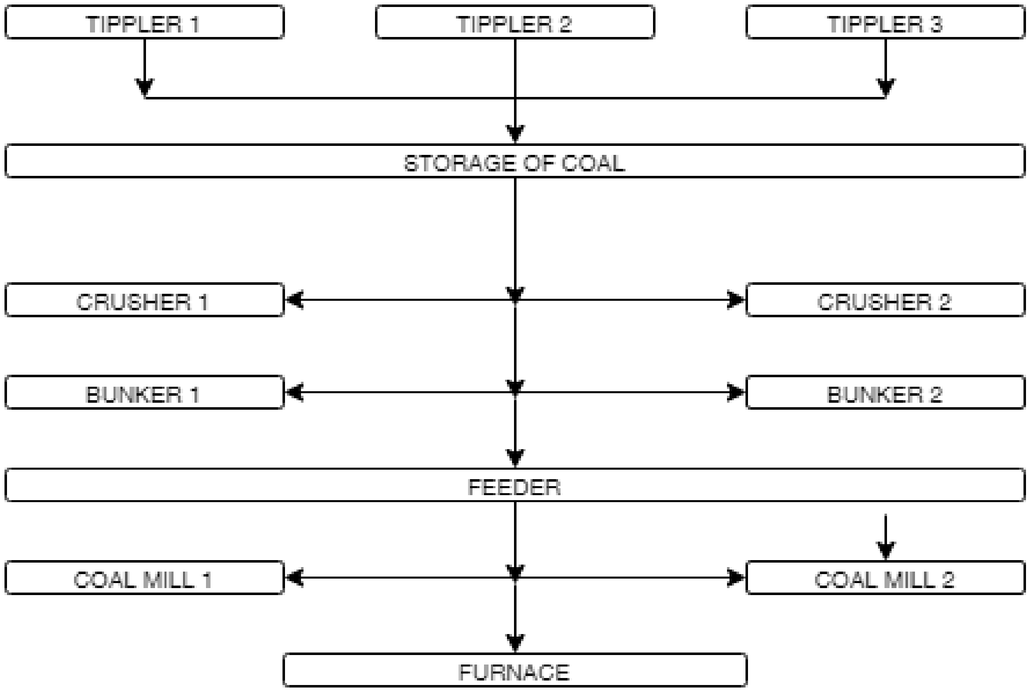


Figure 1: *Illustrative Diagram of Coal Handling System of Thermal Power Plant*

staff and supervisors. The Weibull distribution pattern was intended to be followed by the Failure and Repair Rates. The performance modelling of the coal handling system, which utilised the MOCA-Computation engine based on the Monte Carlo Simulation Approach, was done using the stochastic Petri nets technique (SPN) shown in fig.2. Using simulation for 10000 hours with 21000 replications and a 95 confidence level, the characteristics of plant behaviour were determined. By adjusting Failure and Repair Rates (FRR) within acceptable ranges while doing performance modelling of the plant, it is possible to determine the long-term availability of the various subsystems. the MATLAB program-generated charts for the performability matrices of the Coal Handling System's subsystems. Figures 3 to 7 display the performability w.r.t. FRR for different Coal Handling System subsystems, and tables 1 to 5 display the performability matrices.

Assumptions Notations: The Petri Nets were used for performability analysis, with the following notations and presumptions:

- The failure and repair rates of different subsystems of thermal power plant are exponentially dispersed.
- Single subsystem failures happened one at a time.
- Repaired devices work just as well as brand-new ones.
- Repair includes both component replacement and repair.
- There won't be any instances of two or more subsystems failing simultaneously.
- Standby systems are similar in nature comparable to active systems.
- Over time, the failure rate and repair rate patterns will be stable and statistically independent.

Places: in the petri nets are represented by the circles P = {P1 P2 P3 P4 P5. . . . . . Pn} is a non-empty finite set of places. Each place could be vacant or only store a certain number of tokens. A Petri net's state can be determined via the number of tokens it hold, often known as marking the net.

sys_available: indicates the upstate of the entire system, meaning that it is ready for use.

sys.works_full cap.: depicts the state of the entire system when it is operating at maximum efficiency.

sys.works_red.cap: indicates a system that is operating at a decreased capacity.

sys_failed: depicts the system's downstate

rep.facilities_available: indicates a facility for quick repairs.

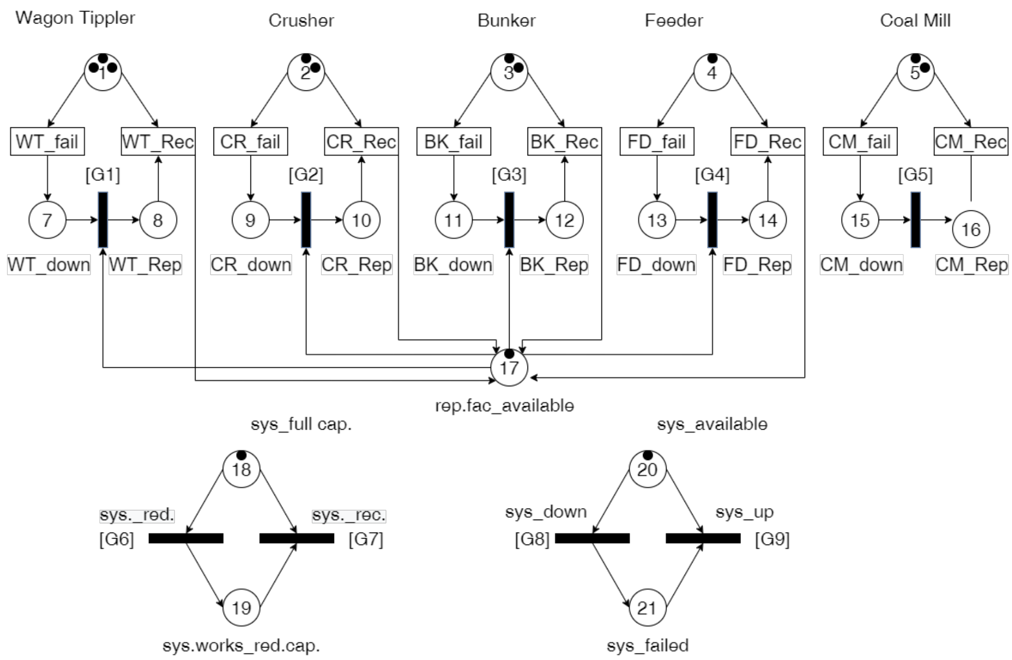


Figure 2: *Modelling of Coal Handling System of Thermal Power Plant Using Petri Nets*

WT_up, CR_up, BK_up, FD_up, CM_up: reflect the operational condition i.e., working state of Wagon Tipplers, Crushers, Bunkers, Feeders and Coal Mills.

WT_down, CR_down, BK_down, FD_down, CM_down: symbolises a state of inefficiency i.e., non-working state of Wagon Tipplers, Crushers, Bunkers, Feeders and Coal Mills.

WT_Rep, CR_Rep, BK_Rep, FD_Rep, CM_Rep: indicates restored conditions of Wagon Tipplers, Crushers, Bunkers, Feeders and Coal Mills.

Transitions: Firing of transitions means occurring of events. Transition fired only if it has at least one token in every location connected to it as an input. These are represented by black or white bars transition T = {T1 T2 T3 T4 T5. . . . Tn,} is a non-empty finite set of transitions. These transitions fired according some predefined sets of rules are known as guard functions. when a transition fires it removes one token from each of its input places and produces a single token on each of its output places. These are of two types timed and direct transitions. Timid transitions fired with some predefined delay. Similarly, the direct transitions fired at once without any kind of time delay. The various transitions used during the performance modelling of coal handling system are as follows:

WT_fail, CR_fail, BK_fail, FD_fail, CM_fail: depict timid transitions linked to failure patterns of Wagon Tipplers, Crushers, Bunkers, Feeders and Coal Mills.

WT_OK, CR_OK, BK_OK, FD_OK, CM_OK: indicates timid transitions that are linked to the failure pattern of waggon tippers., Crushers, Bunkers, Feeders and Coal Mills.

rep. avail_ WT, rep. avail_ CR, rep. avail_ BK rep. avail_ FD, rep. avail_CM: are the immediate transitions indicative of the presence of facilities for repair Wagon Tipplers, Crushers, Bunkers, Feeders and Coal Mills respectively.

sys_ red, sys_ recovered, sys_ fail, and sys_ ok: are immediate transitions which fired immediately without any delay.

Guard Functions: The guard functions, often referred to as enabling functions, are Boolean expressions built using PN primitives (places, transitions, tokens). In addition to the usual requirements, the enabling function must evaluate to true, which modifies the enabling rule. Below is a description of the guard functions related to various transitions.:

[G1]: = (#7>0 and #17>0) **rep. avail_ WT** transition was started via this guard function.

[G2]: = (#9>0 and #17>0) **rep. avail_ CR** transition was started via this guard function.

[G3]: = (#11>0 and #17>0) **rep. avail_ BK** transition was started via this guard function.
[G4]: = (#13>0 and #17>0) **rep. avail_ FD** transition was started via this guard function.
[G5]: = (#15>0 and #17>0) **rep. avail_ CM** transition was started via this guard function.
[G6]: = #2<3 and #2>0 or #3<3 and #3>0 or #5<3 and #5>0) **sys_ red** transition was started via this guard function.
[G7]: =(#2>2and#3>2and#5>2) blocks transition from firing **sys_ recovered.**
[G8]: = (#1>0 or #2>0, or #2>3, or #4>0, or #5>0) **sys_ fail** transition was started via this guard function.
[G9]: = (#1>0 and #2>0, and #2>3, and #4>0, and #5>0) blocks transition from firing **sys_ ok.**

4. Results and Discussion

Table 1: *Performability-Matrix for Wagon Tippler of Coal Handling System*

	Performability Matrix						
$\Phi 1$ $\rho 1$	0.02	0.03	0.04	0.05	0.06	Const. Parameters	
0.0045	0.8133	0.8613	0.8848	0.8932	0.8999		
0.0065	0.7534	0.8342	0.8662	0.8794	0.8839	$\rho 2 = 0.0075$	$\Phi 2 = 0.12$
0.0085	0.6854	0.7968	0.8532	0.8669	0.8794	$\rho 3 = 0.0125$	$\Phi 3 = 0.26$
0.0105	0.6320	0.7623	0.8301	0.8526	0.8734	$\rho 4 = 0.00025$	$\Phi 4 = 0.0075$
0.0125	0.5573	0.7285	0.8030	0.8412	0.8634	$\rho 5 = 0.0004$	$\Phi 5 = 0.005$

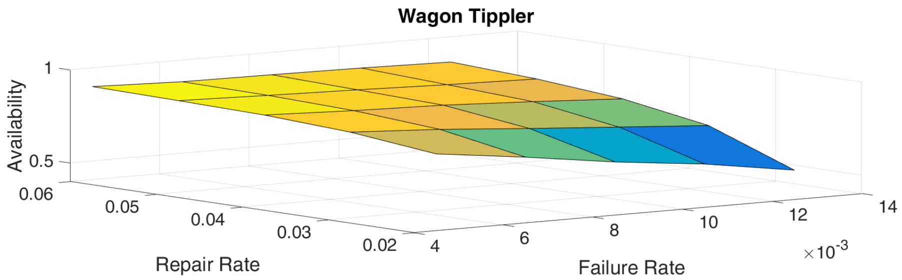


Figure 3: *Impact of varying FRR of Wagon Tippler on the Performability of Coal Handling System*

The effects of variations in the FRR on the performance levels of the waggon tipper (WT) of the coal handling system are shown in Table 1 and Figure 3. Maintaining the repair rate at 0.02, as the failure rate ($\phi 1$) of the WT grows from 0.0045 to 0.0125, and the system's performance severely declines from 0.8133 to 0.5573, or 25.06 %. Similar to this, by maintaining the Failure Rate at 0.0125 while the Repair Rate ($\rho 1$) rises from 0.02 to 0.06, the System's Performability increases immediately from 0.5573 to 0.8634, or by 30.61 %. With variations in FRR combinations, the overall availability of a subsystem can vary by up to 34.26 %.

Table 2: *Performability-Matrix for Crushers of Coal Handling System*

	Performability Matrix						
Φ2 ρ2	0.02	0.07	0.12	0.17	0.22	Const. Parameters	
0.0035	0.7626	0.8532	0.8643	0.8673	0.8707		
0.0055	0.7050	0.8375	0.8555	0.8556	0.8581	$\rho1 = 0.0085$	$\Phi1 = 0.040$
0.0075	0.6461	0.8270	0.8532	0.8489	0.8565	$\rho3 = 0.0125$	$\Phi3 = 0.26$
0.0095	0.5661	0.8229	0.8388	0.8449	0.8527	$\rho4 = 0.00025$	$\Phi4 = 0.0075$
0.0115	0.5140	0.7993	0.8310	0.8377	0.8456	$\rho5 = 0.0004$	$\Phi5 = 0.005$

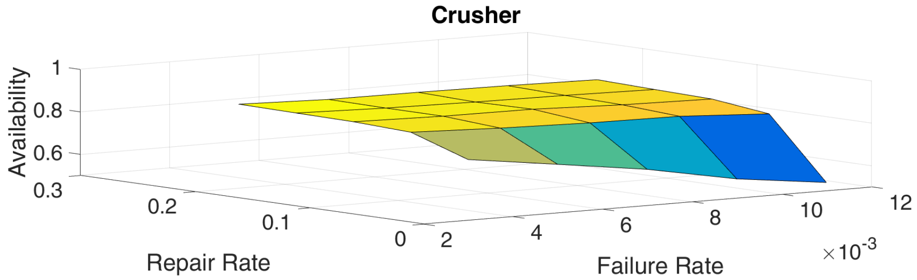


Figure 4: *Impact of varying FRR of Crushers on the Performability of Coal Handling System*

The performance of the Crusher (subsystem) of the coal handling system is affected by variations in the FRR, as shown in Table 2 and Figure 4. The performability levels obtained show that there is a significant impact of FRR fluctuation on the subsystem's performability. The Crusher's failure rate ($\phi2$) ranges from 0.0035 to 0.0115, keeping Repair Rate of 0.02 causes the Crusher's performability levels to drop significantly, from 0.7626 to 0.5140, or 24.86 %. Additionally, the performability levels increase abruptly from 0.5140 to 0.8456, or 33.16 %, when the repair rate ($\rho2$) varies from 0.02 to 0.22. By using various combinations of FRR, it is possible to identify the total variation of 35.67 % in the subsystem's performability.

Table 3: *Performability-Matrix for Bunkers of Coal Handling System*

	Performability Matrix						
Φ3 ρ3	0.16	0.21	0.26	0.31	0.36	Const. Parameters	
0.0085	0.8591	0.8595	0.8610	0.8630	0.8642		
0.0105	0.8480	0.8545	0.8573	0.8523	0.8594	$\rho1 = 0.0085$	$\Phi1 = 0.040$
0.0125	0.8354	0.8467	0.8532	0.8546	0.8517	$\rho2 = 0.0075$	$\Phi2 = 0.12$
0.0145	0.8350	0.8390	0.8416	0.8453	0.8431	$\rho4 = 0.00025$	$\Phi4 = 0.0075$
0.0165	0.8274	0.8332	0.8413	0.8336	0.8401	$\rho5 = 0.0004$	$\Phi5 = 0.005$

Table 3 and Figure 5 indicate the variations in the performability levels of the bunkers (a component of the coal handling system) at various combinations of FRR. According to the performability matrix and plot, the performability dramatically declines from 0.8591 to 0.8274, or 3.17 %, when the failure rate ($\phi3$) of the Bunker rises from 0.0085 to 0.0165. The performability only rises from 0.8274 to 0.8401, or 1.27 %, with an increase in repair rate ($\rho3$) from 0.16 to 0.25. Performability with combinations of FRR has been reported to vary by an average of 3.68 %.

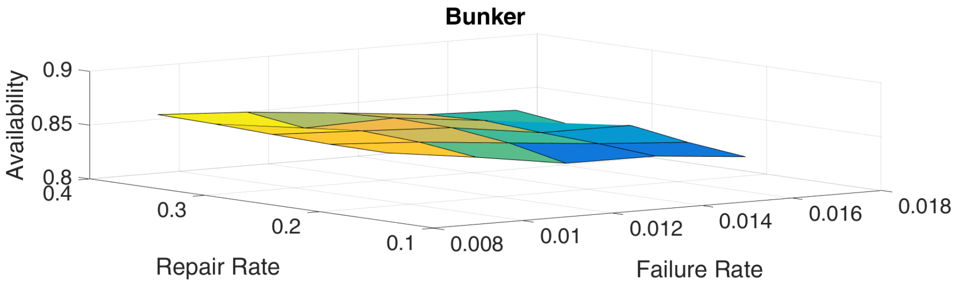


Figure 5: *Impact of varying FRR of Bunkers on the Performability of Coal Handling System*

Table 4: *Performability-Matrix for Feeders of Coal Handling System*

	Performability Matrix					
$\Phi4$ $\rho4$	0.0015	0.0045	0.0075	0.0105	0.0135	Const. Parameters
0.00005	0.8528	0.8744	0.8808	0.8835	0.8819	
0.00015	0.7971	0.8548	0.8664	0.8740	0.8805	$\rho1 = 0.0085$ $\Phi1 = 0.040$
0.00025	0.7332	0.8243	0.8532	0.8626	0.8686	$\rho2 = 0.0075$ $\Phi2 = 0.12$
0.00035	0.6804	0.8067	0.8332	0.8493	0.8529	$\rho3 = 0.0125$ $\Phi3 = 0.26$
0.00045	0.6418	0.7861	0.8188	0.8445	0.8495	$\rho5 = 0.0004$ $\Phi5 = 0.005$

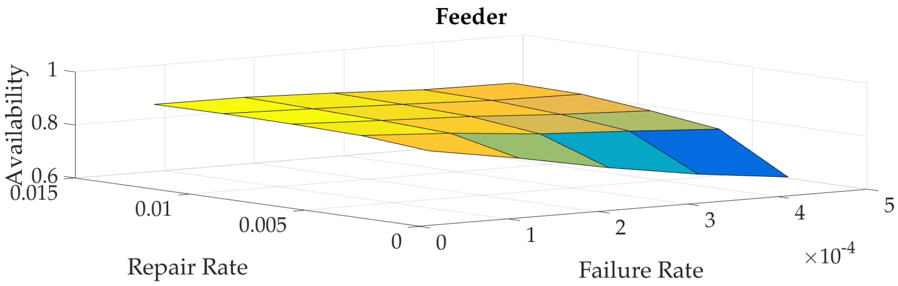


Figure 6: *Impact of varying FRR of Feeders on the Performability of Coal Handling System*

The effects of variations in FRR on the Feeder (subsystem) of the coal handling system are shown in Table 4 and Figure 6. It shows that the performability drops from 0.8528 to 0.6418, or 21.10 %, when the failure rate ($\phi4$) rises from 0.00005 to 0.00045. Similar to this, the performability increases from 0.6418 to 0.8495, or 20.77 %, with variation in repair rate ($\rho4$) from 0.0015 to 0.0135. The entire subsystem performability changes by an outrageous 24.01 %, as observed.

Table 5: *Performability-Matrix for Coal Mills of Coal Handling System*

	Performability Matrix					
$\Phi5$ $\rho5$	0.0040	0.0045	0.0050	0.0055	0.0060	Const. Parameters
0.00030	0.8464	0.8578	0.8623	0.8665	0.8774	
0.00035	0.8418	0.8490	0.8522	0.8639	0.8635	$\rho1 = 0.010$ $\Phi1 = 0.24$
0.00040	0.8255	0.8409	0.8532	0.8602	0.8640	$\rho2 = 0.00020$ $\Phi2 = 0.0035$
0.00045	0.8150	0.8317	0.8345	0.8493	0.8524	$\rho3 = 0.0011$ $\Phi3 = 0.5$
0.00050	0.8097	0.8208	0.8294	0.8409	0.8427	$\rho4 = 0.075$ $\Phi4 = 0.10$

According to Table 5 and Figure 7, the performability levels of the coal mill (subsystem) of the coal handling system vary significantly. The performability levels considerably drop from 0.8464

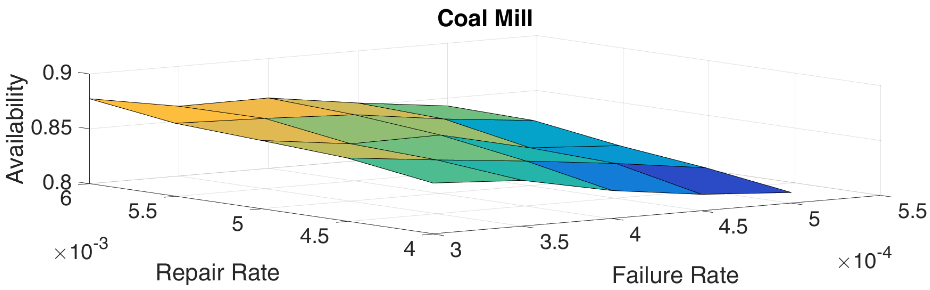


Figure 7: *Impact of varying FRR of Coal Mill on the Performability of Coal Handling System*

to 0.8097, or 3.67 %, when the failure rate of the coal mill ($\phi 5$) rises from 0.00030 to 0.00050. The performability also rises from 0.8097 to 0.8427, or 3.00 %, with an increase in repair rate ($\rho 5$) from 0.0040 to 0.0060. The performability of coal mills has changed by a total of 6.77%.

Table 6: *Variation in the Overall Performability with increase in Repair Facilities*

Availability Matrix					
No. of Repair Facilities	1	2	3	4	5
Availability	0.8532	0.9383	0.9515	0.9446	0.9441

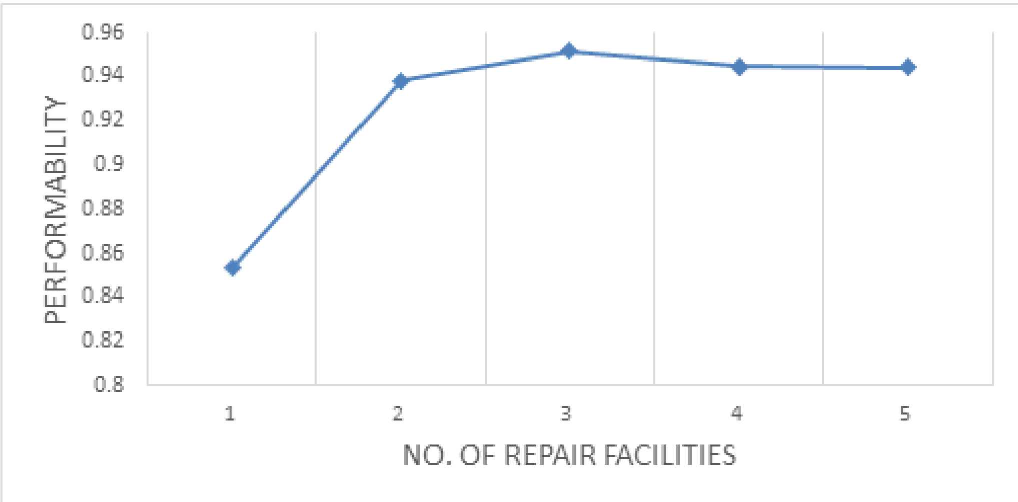


Figure 8: *Impact of variation in Repair Facilities on the Performability of System*

The effect of more repair facilities on the system's overall availability is seen in Table 6 and Fig. 8. The total performability of the coal handling system is seen to dramatically enhance from 0.8532 to 0.9383 as the repair facility grows from 1 to 2. The performability increases noticeably from 0.9383 to 0.9515 when the repair facility is increased from 2 to 3. As the availability goes from three to four and beyond, the performance becomes practically consistent.

5. Optimization Using PSO

The goal of optimization is to identify the best possible solution to a given issue while taking into consideration all of its constraints. In the current study, a population-based global optimisation method known as PSO was used to evaluate the performance of the coal handling system. A stochastic global optimization technique called particle swarm optimization (PSO), proposed by

Eberhart and Kennedy in 1995, was motivated by the social behaviour of fish schools and flocks of birds. The PSO is motivated by swarm behaviour in nature[19, 20, 21, 22]. The movement patterns of several fish and birds served as inspiration for this optimization process. A particle in the swarm is referred to as a bird or a fish. PSO consists of several particles, each of which consists of its current objective value, velocity, and position. Its personal best position, which is the position at which the personal best value has been attained, contrasts with its personal best value, which is the best objective value the particle has ever experienced. The following relations are used by the global best PSO, a common variant of classical PSO, to calculate the ith particle's velocity and position[24].

$$Vi(n+1) = w * Vi(n) + C1(n) * R1i(n) * p - besti - Xi(n) + C2(n) * R2i(n) * g - best - Xi(n) \tag{1}$$

N=0,1,....., N-1

$$Xi(n+1) = Xi(n) + Vi(n+1); n = 0, 1, ..., N-1 \tag{2}$$

where Xi is the position of the ith particle, Vi is its velocity, and n in parentheses is the number of iterations; n = 0 refers to the initialization. N stands for the overall number of iterations, C1 and C2 stand for the personal weight and the global weight, respectively, and they range from 0 to 2 (ideally, C1 = C2 = 2). The random numbers, which are dispersed between 0 and 1, are R1i and R2i. The inertia weight, or w, is a number between 0.4 and 1.4.

Algorithm 1 provides the major steps of the PSO that was implemented also represented in Fig. 9
Steps of the implemented PSO [23]
Repeat
Evaluate z(n, r);
For all particles y Do
Update velocities;
Move to the new position;
If z(n, r) < z(pbesty) then pbesty= (n, r);
If z(n, r) < z(gbest) then gbest= z(n, r);
Update position and velocity;

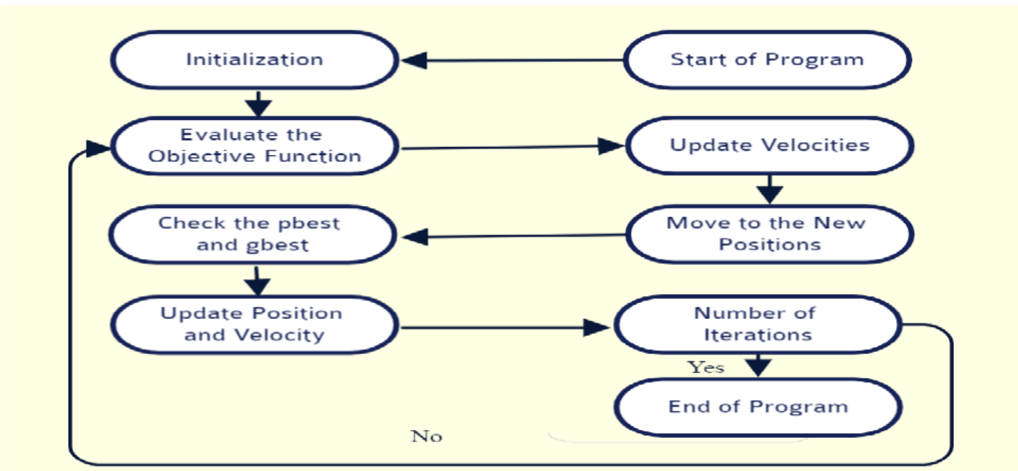


Figure 9: *Flowchart of PSO Implemented*

The following subsection gives more specific results on performance optimization.

Table 7: *Numerous PSO Parameters for Various System*

PSO Parameters			
Sr. No.	Parameter	Range/ Value	Remarks
1	Particle or Population Size	10-100	for optimum Performability
2	Number of Generations	5-50	optimum Performability
3	Inertia Weight (w)	0-1	Its value lies between 0-1
4	Cognitive Factor (c1)	1.49	Selected Arbitrarily
5	Social Factor (c2)	1.49	Selected Arbitrarily
6	Random Number (R1)	0-1	Selected Arbitrarily
7	Random Number (R2)	0-1	Selected Arbitrarily

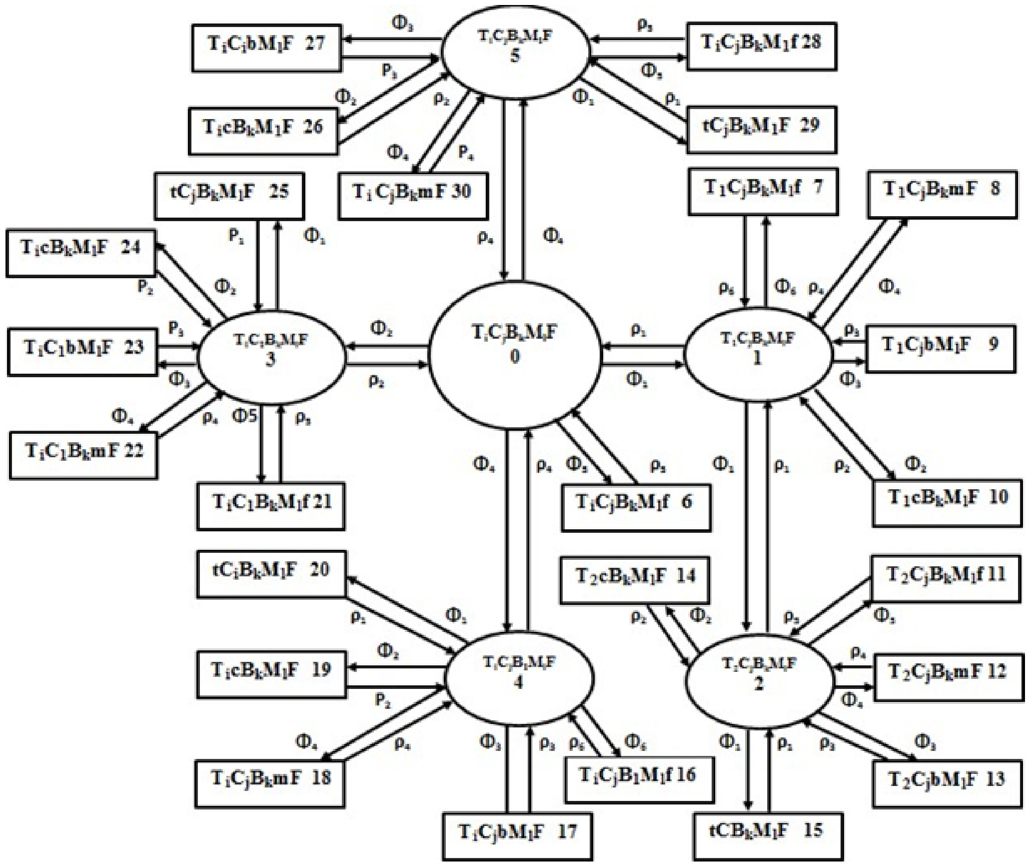


Figure 10: *Transition Diagram for Coal Handling System*

The transition diagram of the coal handling system were obtained as shown in fig. 10. After solving the transition diagram of coal handling system using the Markovian approach the following equation has been obtained for the performance measurement.

$$P0 = 1/[1 + (K1 + K2 + K3 + K4 + K5)(1 + K1 + K1K1 + K2 + K3 + K4)] \tag{3}$$

Where,

Ki = ϕi/ρi

i = 1, 2, 3, 4, 5

By adjusting two factors, PS and the number of generations, the performance was optimised using the aforementioned approach. The following chart illustrates the designed ranges for the failure (ϕ) and repair rate (ρ) parameters of the various subsystems of :

$\phi 1$ (0.0045 - 0.0125), $\rho 1$ (0.02 - 0.06)- Tippler
$\phi 2$ (0.0035-0.0115), $\rho 2$(0.02 - 0.22) Crushers
$\phi 3$ (0.0085-0.0165), $\rho 3$ (0.16-0.36) Bunkers
$\phi 4$ (0.00005-0.00045), $\rho 4$ (0.0015-0.0135) Feeders
$\phi 5$ (0.00030-0.00050), $\rho 5$ (0.0040-0.0060) Coal Mill

The optimum performability of the coal handling system achieved is 93.34 % by using the PSO algorithm at a PS of 40 and by taking constant GS i.e. 100. Table 8 gives the appropriate combinations of FRR as $\phi 1 = 0.006$, $\phi 2 = 0.007$, $\phi 3 = 0.011$, $\phi 4 = 0.0002$, $\phi 5 = 0.0003$, $\rho 1 = 0.07$, $\rho 2 = 0.21$, $\rho 3 = 0.36$, $\rho 4 = 0.006$ and $\rho 5 = 0.005$. The effect of common parameters like PS at constant GS on the performability of the system is described in Figure 11. The performability levels for coal handling system at PS varied from 5 to 50 in a step of 5 taking constant GS are mentioned in Table 8.

Table 8: *Effect of PS on System Performability at constant GS (100)*

Performability Matrix										
FRR	PS5	PS10	PS15	PS20	PS25	PS30	PS35	PS40	PS45	PS50
$\phi 1$ 1	0.008	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
$\phi 1$ 2	0.007	0.006	0.006	0.006	0.005	0.006	0.007	0.007	0.007	0.007
$\phi 1$ 3	0.012	0.009	0.009	0.009	0.010	0.010	0.011	0.011	0.011	0.011
$\phi 1$ 4	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
$\phi 1$ 5	0.0004	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
$\rho 1$	0.06	0.07	0.07	0.07	0.06	0.06	0.07	0.07	0.07	0.07
$\rho 2$	0.10	0.15	0.15	0.15	0.15	0.15	0.20	0.21	0.20	0.20
$\rho 3$	0.33	0.25	0.25	0.25	0.25	0.26	0.35	0.36	0.36	0.36
$\rho 4$	0.007	0.009	0.009	0.009	0.009	0.009	0.006	0.006	0.006	0.006
$\rho 5$	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
PA	90.19	92.40	92.41	92.40	92.42	92.46	93.10	93.34	93.22	93.21

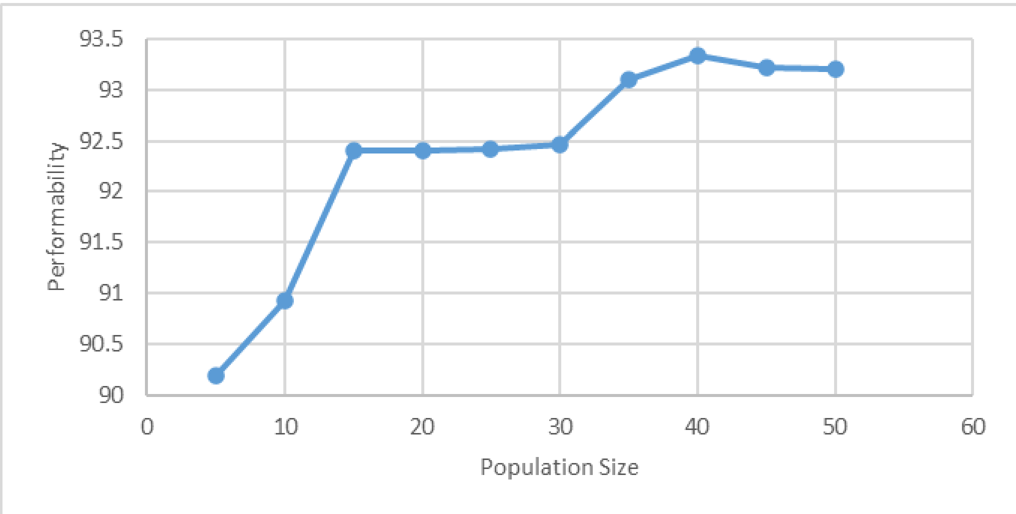


Figure 11: *Effect of PS on System Performability*

The optimum performability of coal ash handling attained is 93.32 % by using the PSO algorithm at a GS of 70 and by taking constant PS i.e. 40. Table 9 offers the appropriate combinations of FRR as ϕ 1 = 0.006, $\phi 2 = 0.007$, $\phi 3 = 0.011$, $\phi 4 = 0.0002$, $\phi 5 = 0.0003$, $\rho 1 = 0.07$, $\rho 2 = 0.21$, $\rho 3 = 0.36$, $\rho 4 = 0.006$ and $\rho 5 = 0.005$. The effect of common parameters like PS at constant

GS is indicated in Figure 12. The performability levels for power generation system at GS varied from 10 to 100 in a step of 10 taking constant PS are given in Table 9.

Table 9: *Effect of GS on System Performability at constant PS (40)*

Performability Matrix										
FRR	GS10	GS20	GS30	GS40	GS50	GS60	GS70	GS80	GS90	GS100
ϕ1 1	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
ϕ1 2	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
ϕ1 3	0.012	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
ϕ1 4	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
ϕ1 5	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
ρ1	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
ρ2	0.20	0.20	0.21	0.20	0.21	0.21	0.21	0.21	0.21	0.21
ρ3	0.34	0.35	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36
ρ4	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
ρ5	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
PA	91.70	93.03	93.09	93.16	93.20	93.26	93.32	93.23	93.21	93.13

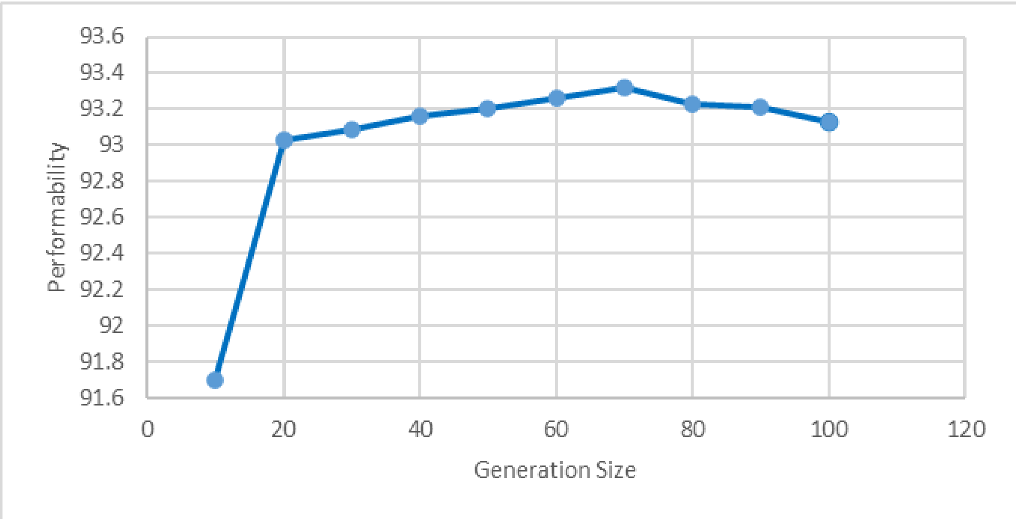


Figure 12: *Effect of GS on System Performability*

Table 10: *Effect of PS on System Performability at constant GS (100)*

Decision Support System				
Subsystem	Variation in FR (ϕ)	Variation in RR (ρ)	Change in Availability	Priority
Tippler	0.0045- 0.0125	0.02- 0.06	55.7 to 89.99	2
Crusher	0.0035- 0.0115	0.02 -0.22	51.40 to 87.07	1
Bunker	0.0085- 0.0165	0.16 -0.36	82.7 to 86.42	5
Feeder	0.00005- 0.00045	0.0015 -0.0135	64.1 to 88.19	3
Coal Mill	0.00030- 0.00050	0.0040 -0.0060	80.97 to 87.74	4

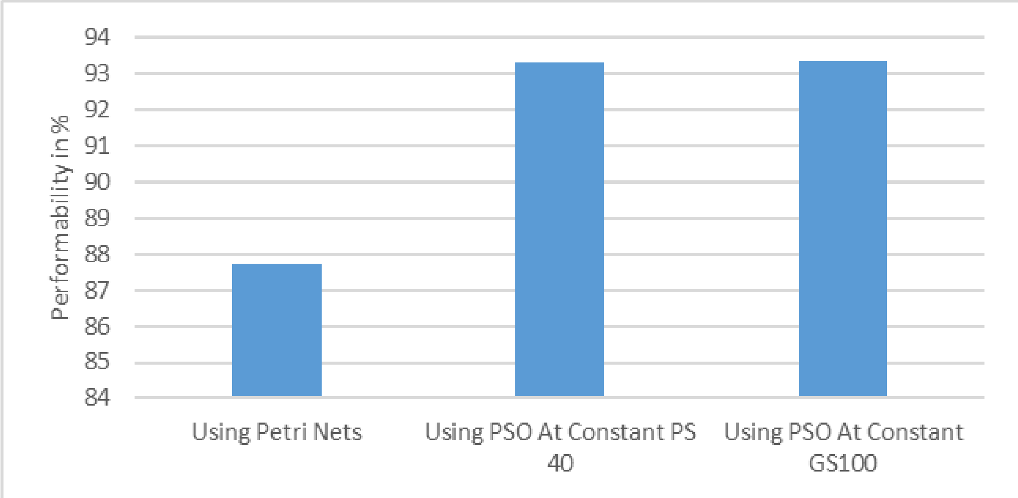


Figure 13: *Comparative Analysis of Performability Levels using Petri Nets and PSO*

6. Conclusion

The analytical identification of the most important subsystems is one of the key conclusions of present case study's . Due to the high risk of failure, these subsystems should be given top priority during maintenance. It is observed that the Crusher and Tippler are crucial components that demand the plant manager's full attention. As a result, the concerned plant authorities can develop and implement suitable maintenance plans to enhance system functionality. A DSS has been suggested (in Table 10) to help the plant managers, based on a thorough analysis, and is likely to increase efficiency even more. Due to the resources available and the maintenance planning techniques being used, managing the maintenance of industrial systems that can be repaired is a very difficult task. This approach also establishes a trade-off between financial investments and benefits earned in terms of revenue, reputation, safety, etc. The best value of the system performability obtained by PSO as shown in fig. 13 reveal that the PSO has outperformed the Petri Nets.

References

[1] Carazas, F.J. G. and De Souza,G.F. M. (2009). Availability analysis of gas turbines used in power plants *Int. J. Thermodyn.* , 12 (1):28–37.
[2] Yang, S.K., 2004. A condition-based preventive maintenance arrangement for thermal power plants. Electr. Power Syst. Res. 72 (1), 49"62.
[3] D Cotroneo, AK Iannillo, R Natella, S. Rosiello, Dependability Assessment of the Android OS Through Fault Injection, IEEE Trans Reliab (2019) 1"16, https://doi.org/10.1109/tr.2019.2954384.
[4] P Zhou, D Zuo, K Hou, Z Zhang, J. Dong, Improving the Dependability of Self- Adaptive Cyber Physical System With Formal Compositional Contract, IEEE Trans Reliab (2019), https://doi.org/10.1109/TR.2019.2930009.
[5] L. Li, Software reliability growth fault correction model based on machine learning and neural network algorithm, Microprocess Microsyst 80 (2021), 103538, https://doi.org/10.1016/j.micpro.2020.103538.
[6] Y Tang, Y Yuan, Y. Liu, Cost-aware reliability task scheduling of automotive cyberphysical systems, Microprocess Microsyst (2020), 103507, https://doi.org/ 10.1016/j.micpro.2020.103507.
[7] Eti, M.C., Ogaji, S.O.T., Probert, S.D., 2007. Integrating reliability availability, maintainability and supportability with risk analysis for improved operation of the afam thermal power-station. Appl. Energy 84 (2), 202"221.

[8] Melani, A.H.A., Murad, C.A., Caminada Netto, A., de Souza, G.F.M., Nabeta, S.I., 2018. Criticality-based maintenance of a coal-fired power plant. Energy 147 (C), 767"781.

[9] Wang, L., Chu, J., Wu, J., 2007. Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process. Int. J. Prod. Econ. 107 (1), 151"163.

[10] Singh, R.K., Kulkarni, M.S., 2013. Criticality analysis of power-plant equipments using the analytic hierarchy process. Int. J. Ind. Eng. Technol. 3 (4), 1"14.

[11] Jagtap, H.P., Bewoor, A.K., 2017. Use of analytic hierarchy process methodology for criticality analysis of thermal power plant equipments. Mater. Today Proc. 4 (2), 1927"1936.

[12] Adamyan, A. and He, D. (2002), Analysis of sequential failures for assessment of reliability and safety of manufacturing systems´, Reliability Engineering and System Safety, Vol. 76, No. 3, pp. 227-23 6.

[13] Sachdeva, A., Kumar, D. and Kumar, P. (2008), Reliability analysis of pulping system using petri nets´, International Journal of Quality Reliability Management, Vol. 25, No. 8, pp. 860-877.

[14] Sachdeva, A., Kumar, D., and Kumar, P. (2008), Availability modeling of screening system of a paper plant using GSPN´, Journal of Modelling in Management´, Vol. 3, No. 1, pp. 26-39.

[15] Sachdeva, A., Kumar, D., and Kumar, P. (2008), Planning and optimization the maintenance of paper production systems in a paper plant´, Computers Industrial Engineering´, Vol. 55, No. 4, pp. 817-829.

[16] Sachdeva, A., Kumar, P., and Kumar, D. (2009), Behavioral and performance analysis of feeding system using stochastic reward nets´, International Journal of Advanced Manufacturing Technology´, Vol. 45, No. 1-2, pp. 156-169.

[17] Sachdeva, A. (2009), RAM analysis of industrial systems using Petri Nets-Ph.D Thesis´, Department of Mechanical and Industrial Engineering IIT-Roorkee, pp. 1-307. 6.

[18] Schneeweiss, W.G. (2001), Tutorial: Petri Nets as a graphical description medium for many reliability scenarios´, IEEE Transactions on Reliability, Vol. 50, No. 2, pp. 159-164.

[19] J Kennedy, R. Eberhart, Particle swarm optimization. Neural Networks, in: 1995 Proceedings, IEEE Int Conf, 1995, pp. 1942"1948, https://doi.org/10.1109/ ICNN.1995.488968, 4vol.4.

[20] Mellal MA, Williams EJ. A survey on ant colony optimization, particle swarm optimization, and cuckoo algorithms. Handb. Res. Emergent Appl. Optim. Algorithms. IGI Global, USA: 2018.

[21] A Khare, S. Rangnekar, A review of particle swarm optimization and its applications in Solar Photovoltaic system, Appl Soft Comput 13 (2013) 2997"3006, https://doi.org/10.1016/J.ASOC.2012.11.033.

[22] Jyoti Saharia B, H Brahma, N Sarmah, A review of algorithms for control and optimization for energy management of hybrid renewable energy systems, J Renew Sustain Energy 10 (2018), 053502, https://doi.org/10.1063/1.5032146.

[23] [Mellal MA, Williams EJ. A survey on ant colony optimization, particle swarm optimization, and cuckoo algorithms. Handb. Res. Emergent Appl. Optim. Algorithms. IGI Global, USA: 2018.

[24] Schoene, T. (2011) Step-Optimized Particle Swarm Optimization, MS thesis, University of Saskatchewan, Saskatoon, Canada.

Bayesian and Non-Bayesian Inference of Exponentiated Moment Exponential Distribution with Progressive Censored Samples

Amal S. Hassan [1], Samah A. Atia [2], Hiba Z. Muhammed [3]

1. Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research, Egypt, Email: amal52_soliman@cu.edu.eg
2 Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research, Egypt, Email: 12422019446259@pg.cu.edu.eg
3. Department of Mathematical Statistics, Cairo University, Faculty of Graduate Studies for Statistical Research, Egypt, Email: hiba_stat@cu.edu.eg

Abstract

In this paper, a progressive type-II censoring strategy is used to estimate the parameters, reliability and hazard rate functions of the exponentiated moment exponential distribution. The maximum likelihood and Bayesian techniques have been used to estimate the proposed estimators. Gamma (informative) and uniform (non-informative) priors are taken into account under the squared error loss function to produce the Bayesian estimators. The highest posterior density interval estimations and the 95% approximate confidence intervals along with coverage probability are calculated. In order to evaluate the effectiveness of estimates produced by the Metropolis-Hastings sampling algorithms, we provide a numerical research. According to the study's findings, the Bayes estimates under informative priors are typically more accurate than other estimates.

Key Words: Exponentiated moment exponential, gamma prior, credible interval, Metropolis-Hastings, Progressive censorning

1. Introduction

Censoring is widely used in reliability data analysis and other practical life-testing investigations. It becomes apparent when precise failure times for a subset of the test units used in an experiment are observed. The experimenter frequently runs into incomplete data in this scenario. Typical censoring systems include type I censoring (T1C) and type II censoring (T2C). The units can only be expelled after the conclusion of the experiment, which is a major drawback of TIC and T2C methods. In a more open-ended censoring technique known as progressive censoring (PC), units are designated to be discarded from the test at times other than the eventual termination time point. The remaining units are then tested again while being observed. To learn more, visit Balakrishnan [1].

Progressive T2C (PT2C) is the major topic of this research project. Let's assume that n identical items are used in the experiment, and that the PC scheme R is pre-fixed so that, after the first failure, R_1 surviving items are ejected from remaining live $(n-1)$ items, R_2 surviving items are ejected from remaining live $(n-R_1-2)$ items, and so on. After m^{th} failure, this procedure is maintained until all $R_m = n-m-R_1-...R_{m-1}$ remaining objects are expelled (see Hofmann et al [2]). Therefore, a PT2C procedure consists of m and $R_1, R_2, ..., R_m$ such that $\sum_{i=1}^{m} R_i + m = n$. Note that, if $R_1 = R_2 = \ldots = R_m = 0$ then the PT2C provides complete sampling and if $R_1 = R_2 = \ldots = R_{m-1} = 0$ and $R_m = n-m$ then PT2C yields T2C scheme (see Krishna and Kumar [3]).

According to PT2C samples, the likelihood function of random variable X (Balakrishnan and Aggarwala [4]) is supplied as follows .

$$L(\theta) = C \prod_{i=1}^{m} f(x_{(i)}) \left[1 - F(x_{(i)})\right]^{R_i}, \tag{1}$$

where $C = n(n - R_1 - 1) \ldots (n - R_1 - R_2 - \cdots - R_{m-1} - m + 1)$. Some important literature regarding the estimation studies under PT2C scheme can be found in Wu [5] , Ng [6], Dey et al. [7], Hassan et al. [8], EL-Sagheer [9], Noor et al. [10], Alshenawy et al. [11], and Shrahili et al. [12].

Moment distributions are essential in probability theory and several economic, reliability, and biological studies, as well as other areas of mathematics and statistics. Some of the fundamental features of the moment exponential (ME) distribution were studied and suggested by Dara and Ahmad [13]. The version of the ME distribution that includes an additional shape parameter is known as the exponentiated ME (EME) distribution, and it is frequently employed in reliability research. Hasnain et al. [14] suggested several EME distribution features, including conditional-based characterisation, explored maximum likelihood (ML) estimators, and fitted it to actual data sets. Compared to the ME distribution and exponentiated exponential (EE) distribution, the EME distribution is more adaptable when fitting data. As described by Hasnain et al. [14] the EME distribution's cumulative distribution function (CDF), is

$$F(x;\alpha,\beta) = [1-\psi]^{\alpha}; \quad x,\beta,\alpha > 0, \tag{2}$$

where $\psi = (1 + x\beta^{-1}) e^{-x/\beta}$, β is scale parameter and α is shape parameter. The probability density function (PDF) of the EME distribution is

$$f(x;\alpha,\beta) = \alpha\beta^{-2} [1-\psi]^{\alpha-1} x e^{-x/\beta}; \qquad x,\beta,\alpha > 0. \tag{3}$$

For $\beta = 1$, the CDF (2) gives the CDF of one parameter EE distribution (Gupta and Kundu [15]). Also, for $\alpha = 1$, the CDF (2) gives the CDF of ME distribution. The reliability function (RF) and hazard rate function (HRF) related to (3) are defined as:

$$R(x) = 1 - [1-\psi]^{\alpha}, \qquad h(x) = \frac{\alpha x e^{-x/\beta} [1-\psi]^{\alpha-1}}{\beta^2 \left(1 - [1-\psi]^{\alpha}\right)}.$$

Plots of the PDF and HRF of the EME distribution are displayed in Figure 1. It is evident that different parameter values result in varied forms for the PDF for the EME distribution. The distribution may alternatively be characterised as favourably skewed to right and uni-modal. It is clear that the EME distribution's HRF has an increasing trend.

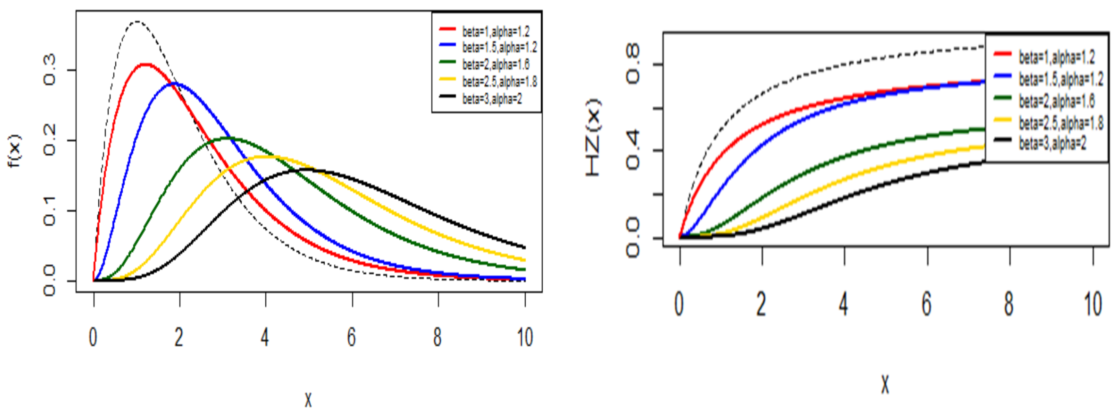


Figure 1: PDF and HRF plots of the EME distribution

Different approaches to estimating the PDF and CDF of the EME distribution were provided by Tripathi et al. [16]. The ML and Bayesian techniques developed by Fatima and Ahmad [17] have been taken into consideration when discussing the parameter estimators of the EME distribution. Akhter et al. [18] provided explicit algebraic equations that are generated from the EME distribution for both single and product moments of order statistics. Additionally, they used a full sample as well as a T2C sample to identify the best linear unbiased estimators based on these moments. Some generalizations of EME distribution may be found in Iqbal et al. [19], Ahmadini et al. [20] and Shrahili et al. [21].

The RF, HRF, and parameter estimators of the EME utilising ML and Bayesian techniques are addressed in the current study. Both the Bayesian credible intervals (BCIs) and the approximate confidence intervals (ACIs) are built using the PT2C data. This document can be constructed as shown below. Section 2 deals with ML estimators and the ACIs of parameters, RF and HRF. Sections 3 explore Bayesian estimate under informative (IF) and non-informative (NIF) priors. Sections 4 and 5, respectively, provide numerical illustrative studies and a conclusion.

2. Maximum Likelihood Procedure

Here, using PT2C data, we obtain the ML estimators of the parameters, RF, and HRF of the EME distribution. In addition, the ACIs for the RF, HRF and the parameters β and α are built.

Let $x_{(1)}, x_{(2)},..., x_{(m)}$ be the observed PT2C random samples extracted from the EME distribution. Based on (1), then the likelihood function of the EME distribution takes the following form:

$$L(\underline{x}|\beta,\alpha) \propto \prod_{i=1}^{m} \alpha\beta^{-2}\left(1-\psi_i\right)^{\alpha-1} x_i e^{-x_i/\beta}\left[1-\left(1-\psi_i\right)^{\alpha}\right]^{R_i}, \tag{4}$$

where $\psi_i=(1+x_i\beta^{-1})e^{-x_i/\beta}$ and we write $x_{(i)}=x_i$ for simplified form. The logarithm of (4), say $\ell=\log L(\underline{x}|\beta,\alpha)$ becomes:

$$\ell \propto m\ln\alpha - 2m\ln\beta + \sum_{i=1}^{m}\ln x_i - \frac{1}{\beta}\sum_{i=1}^{m} x_i + (\alpha-1)\sum_{i=1}^{m}\ln\left(1-\psi_i\right) + \sum_{i=1}^{m} R_i \ln\left[1-\left(1-\psi_i\right)^{\alpha}\right]. \tag{5}$$

The first derivative of (5) via β and α are given as:

$$\frac{\partial\ell}{\partial\beta} = \frac{-2m}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{m} x_i - (\alpha-1)\sum_{i=1}^{m}\frac{\varpi_i}{(1-\psi_i)} + \alpha\sum_{i=1}^{m} R_i \frac{\left(1-\psi_i\right)^{\alpha-1} x_i\left(\psi_i - e^{-x_i/\beta}\right)}{\beta^2\left[1-\left(1-\psi_i\right)^{\alpha}\right]},$$

$$\frac{\partial \ell}{\partial \alpha}=\frac{m}{\alpha}+\sum_{i=1}^{m}\ln\left(1-\psi_i\right)-\sum_{i=1}^{m}R_i\,\frac{\left(1-\psi_i\right)^{\alpha}\ln\left(1-\psi_i\right)}{1-\left(1-\psi_i\right)^{\alpha}},$$

where $\varpi_i = x_i\beta^{-2}(\psi_i - e^{-x_i/\beta})$. The estimator of β and α is the solution of the first derivative of $\partial\ell/\partial\beta\big|_{\beta=\hat{\beta}}=0$ and $\partial\ell/\partial\alpha\big|_{\alpha=\hat{\alpha}}=0$. Numerical iterative approach may be used to calculate the estimator of β and α for the specified values of ($m,\underline{R},\underline{x}$). Additionally, the invariance feature of the ML method is used to evaluate $R(x)$ and $h(x)$ as below

$$\hat{R}(x)=1-\left[1-\psi\right]^{\hat{\alpha}},\qquad \hat{h}(x)=\alpha x\hat{\beta}^{-2}e^{-x_i/\hat{\beta}}\left[1-\psi\ \right]^{\hat{\alpha}-1}\left(1-\left[1-\psi\right]^{\hat{\alpha}}\right)^{-1}.$$

In addition, we get the observed information matrix, say $I(\hat{\alpha},\hat{\beta})$, to build ACIs. The multivariate normal distribution $N_2(0,I^{-1}(\hat{\alpha},\hat{\beta}))$ is used to create ACIs for the parameters β and α with the usual regularity requirements. Based on the asymptotic normality criteria of the ML, the two-sided $100(1-\varepsilon)\%$ACI for parameter β and α is

$$\text{AsyCI\_Upper}=\hat{\beta}\ +Z_{\varepsilon/2}\sqrt{\text{var}(\hat{\beta})}\ \text{and}\ \hat{\alpha}+Z_{\varepsilon/2}\sqrt{\text{var}(\hat{\alpha})}\ ,$$
$$\text{AsyCI\_Lower}=\hat{\beta}\ -Z_{\varepsilon/2}\sqrt{\text{var}(\hat{\beta})}\ \text{and}\ \hat{\alpha}-Z_{\varepsilon/2}\sqrt{\text{var}(\hat{\alpha})}\ ,$$
$$\text{AIL= AsyCI\_Upper}-\text{AsyCI\_Lower},$$

where $Z_{\varepsilon/2}$ is the right tail probability's percentile $\varepsilon/2$ for the standard normal distribution. Once more, an R-based numerical method is offered to get the variance-covariance matrix. Also, the $100(1-\varepsilon)\%$ACI for $R(x)$ and $h(x)$ are given by $\hat{R}(x)\ \pm Z_{\varepsilon/2}\sqrt{\tilde{\text{var}}(\hat{R}(x))}$, $\hat{h}(x)\ \pm Z_{\varepsilon/2}\sqrt{\tilde{\text{var}}(\hat{h}(x))}$.

3. Bayesian Estimators

Here, Bayesian estimator of the parameters, RF and HRF of the EME distribution in case of IF and NIF priors under squared error (SE) loss function. Firstly, consider β and α have a gamma distribution with parameters (*a, b*) and (*c, d*) respectively. Assuming that β and α are independently distributed, the joint prior distribution of β and α is given by:

$$h_{1,2}(\beta,\alpha|\underline{x})=\frac{b^a d^c}{\Gamma(a)\Gamma(c)}\beta^{a-1}\alpha^{c-1}e^{-b\beta-d\alpha}\,,$$

where *a, b, c* and *d* are chosen to reflect the prior knowledge about the unknown parameters (the criteria to select the hyper-parameter values is discussed in Section 3.1). The joint posterior distribution of parameters β and α is defined as:

$$\pi_2(\beta,\alpha|\underline{x})=\frac{\beta^{a-3}\alpha^{c}e^{-b\beta-d\alpha}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1}x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i}}{\int_0^\infty\int_0^\infty \beta^{a-3}\alpha^{c}e^{-b\beta-d\alpha}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1}x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i}\,d\beta\,d\alpha}.$$

Hence, the marginal posterior distributions of β and α take the following forms:

$$h_1(\beta|x) = k_1^{-1}\ \beta^{a-3}e^{-b\beta}\int_0^\infty \alpha^c e^{-d\alpha}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ d\alpha\,,$$

$$h_2(\alpha|\underline{x}) = k_1^{-1}\ \alpha^c e^{-d\alpha}\int_0^\infty \beta^{a-3}e^{-b\beta}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ d\beta\,,$$

where $$k_1 = \int_0^\infty\int_0^\infty \beta^{a-3}\alpha^c e^{-b\beta-d\alpha}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ d\beta\, d\alpha\,.$$

The Bayesian estimator of β and α, expressed by $\tilde{\beta}$ and $\tilde{\alpha}$, are obtained as follows:

$$\tilde{\beta} = k_1^{-1}\int_0^\infty\int_0^\infty \beta^{a-2}\alpha^c e^{-b\beta-d\alpha}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ d\beta\, d\alpha\ ,$$

$$\tilde{\alpha} = k_1^{-1}\int_0^\infty\int_0^\infty \beta^{a-3}\alpha^{c+1} e^{-b\beta-d\alpha}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ d\beta\, d\alpha\,.$$

The Bayesian estimators of $R(x)$ and $h(x)$ are given by:

$$\tilde{R}(x) = k_1^{-1}\int_0^\infty\int_0^\infty\left(1-\left[1-(1+x\beta^{-1})e^{-x/\beta}\right]^\alpha\right)\beta^{a-3}\alpha^c e^{-b\beta-d\alpha}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i} d\alpha\, d\beta, \quad (6)$$

$$\tilde{h}(x) = k_1^{-1}\int_0^\infty\int_0^\infty\frac{x e^{-x/\beta}\left[1-(1+x\beta^{-1})\,e^{-x/\beta}\right]^{\alpha-1}}{\left(1-\left[1-(1+x\beta^{-1})e^{-x/\beta}\right]^\alpha\right)}\ \beta^{a-5}\alpha^{c+1} e^{-b\beta-d\alpha}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i} d\alpha\, d\beta. \quad (7)$$

The above Bayesian estimators $\tilde{\beta},\tilde{\alpha},\tilde{R}(x)$ and $\tilde{h}(x)$ are not in closed forms but can be evaluated numerically for the given values of $a,b,c,d,n,m,\underline{x}$ and $\underline{R}$.
Secondly, assuming the prior of parameters β and α, denoted by $g_1(\beta)$ and $g_2(\alpha)$ has the uniform (NIF) prior distribution. The joint prior for parameters β and α, represented by $g_{1,2}(\alpha,\beta)$, assuming independent of priors, is

$$g_{1,2}(\alpha,\beta|\underline{x}) = (\alpha\beta)^{-1},\quad 0<\beta,\ \alpha<1.$$

The joint posterior density of β and α given the data $\underline{x}$ is given by:

$$\pi_1(\beta,\alpha|\underline{x}) = k_2^{-1}\beta^{-3}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}.$$

where $$k_2 = \int_0^\infty\int_0^\infty \beta^{-3}\prod_{i=1}^m (1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ d\beta\, d\alpha\,.$$

Hence, the marginal posterior distributions of β and α take the following forms:

$$g_1(\beta|\underline{x}) = \int_\alpha \pi_1(\beta,\alpha|\underline{x})\ d\alpha = k_2^{-1}\beta^{-3}\int_0^\infty\prod_{i=1}^m (1-\psi_i)^{\alpha-1}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ x_i e^{-x_i/\beta} d\alpha\,,$$

$$g_2(\alpha|\underline{x}) = \int_\beta \pi_1(\beta,\alpha|\underline{x})\, d\beta = k_2^{-1}\int_0^\infty \beta^{-3}\prod_{i=1}^m (1-\psi_i)^{\alpha-1}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\ x_i e^{-x_i/\beta}\ d\beta\,,$$

The Bayesian estimator of β and α, denoted by $\ddot{\beta}$ and $\ddot{\alpha}$, are obtained as follows:

$$\ddot{\beta} = E(\beta|\underline{x}) = k_2^{-1}\int_0^\infty\int_0^\infty \beta^{-2}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\beta\, d\alpha, \tag{8}$$

$$\ddot{\alpha} = E(\alpha|\underline{x}) = k_2^{-1}\int_0^\infty\int_0^\infty \alpha\beta^{-3}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\beta\, d\alpha. \tag{9}$$

The Bayesian estimator of $R(x)$ and $h(x)$ are given by:

$$\ddot{R}(x) = \int_0^\infty\int_0^\infty\left(1-\left[1-(1+x\beta^{-1})\,e^{-x_i/\beta}\right]^{\alpha}\right)\beta^{-3}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\alpha\, d\beta, \tag{10}$$

$$\ddot{h}(x) = k_2^{-1}\int_0^\infty\int_0^\infty \frac{\alpha x e^{-x/\beta}\left[1-(1+x\beta^{-1})\,e^{-x_i/\beta}\right]^{\alpha-1}}{\beta^2\left(1-\left[1-(1+x\beta^{-1})e^{-x_i/\beta}\right]^{\alpha}\right)}\beta^{-3}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\alpha\, d\beta. \tag{11}$$

The above Bayes estimates $\ddot{\beta}, \ddot{\alpha}, \ddot{R}(x)$ and $\ddot{h}(x)$ are assessed numerically for the given values of $n, m, \underline{x}$ and $\underline{R}$. Integrals (8)–(11) are very hard to be solved analytically, so the Metropolis-Hastings (MH) algorithm will be used to solve these integrals.

3.1 Hyper-Parameter Elicitation

This sub-section handled the elicitation of the hyper-parameter values in case of IP. These hyper-parameters of IP are obtained from ML estimators for β and α, by equating the mean and variance of $\hat{\beta}^i$ and $\hat{\alpha}^i$ with the mean and variance of the gamma distributions, where i=1,2,….,N and N is the number of samples available from the EME distribution. Thus,

$$\frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^i = \frac{a}{b},\quad \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\beta}^i - \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^i\right)^2 = \frac{a}{b^2},\ \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}^i = \frac{c}{d},\quad \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\alpha}^i - \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}^i\right)^2 = \frac{c}{d^2}.$$

Hence, the estimated hyper-parameters are obtained as follows
:

$$a = \frac{\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^i\right)^2}{\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\beta}^i - \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^i\right)^2}, b = \frac{\frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^i}{\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\beta}^i - \frac{1}{n}\sum_{i=1}^{N}\hat{\beta}^i\right)^2}, c = \frac{\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}^i\right)^2}{\frac{1}{N-1}\sum_{i=1}^{n}\left(\hat{\alpha}^i - \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}^i\right)^2}, d = \frac{\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}^i}{\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\alpha}^i - \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}^i\right)^2}.$$

For more information (see Dey and Pradhan, 2014).

3.2 Bayesian Credible Intervals

Furthermore, the BCI of α and β denoted by $\tilde{\alpha}_{BCIF}$ and $\tilde{\beta}_{BCIF}$ is obtained under IF and NIF priors as follows:

$$\tilde{\beta}_{BCIF} = k_1^{-1}\int_L^U\int_0^\infty k_2^{-1}\beta^{a-2}\alpha^{c}e^{-b\beta-d\alpha}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\beta\, d\alpha = 0.95, \tag{12}$$

$$\tilde{\alpha}_{BCIF} = k_1^{-1}\int_L^U\int_0^\infty \beta^{a-3}\alpha^{c+1}e^{-b\beta-d\alpha}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\beta\, d\alpha = 0.95. \tag{13}$$

$$\ddot{\beta}_{BCNIF} = \int_L^U\int_0^\infty k_2^{-1}\beta^{-2}\prod_{i=1}^{m}(1-\psi_i)^{\alpha-1} x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^{\alpha}\right\}^{R_i} d\beta\, d\alpha = 0.95, \tag{14}$$

$$\ddot{\alpha}_{BCNIF}=\int_L^U\int_0^\infty k_2^{-1}\alpha\beta^{-3}\prod_{i=1}^m(1-\psi_i)^{\alpha-1}\,x_i e^{-x_i/\beta}\left\{1-(1-\psi_i)^\alpha\right\}^{R_i}\,d\beta\,d\alpha=0.95. \tag{15}$$

Integrals (12)–(15) are very hard to be solved analytically, so the MH algorithm will be used to solve these integrals Similary, the BCI of $R(x)$ and $h(x)$ provided in (6), (7) under IP and the BCI of $R(x)$ and $h(x)$ provided in (10) and (11) under NIP are obtained using the above procedure.

4. Numerical Illustration

To determine ML estimates (MLEs) and Bayesian estimates (BEs) for parameters, RF and HRF under the PT2C scheme, a simulation study was conducted. Different sample sizes (n), effective failure sizes (m), and picking parameter values are taken into consideration. The R 3.6.1 software is used to complete the following stages.

1. Using the same technique as that provided by Balakrishnan and Sandhu [22], which includes the following, random samples $X_1, X_2,..., X_n$ are produced from the EME distribution under PT2C samples:

i. Generate m independent and identically (iid) random numbers $W_1, W_2,..., W_m$ from uniform distribution U(0,1).

ii. Set $V_i = W_i^{(1/i+R_m+R_{m-1}+...+R_{m-i+1})}$ for $i = 1,2,..., m$.

iii. Set $U_i = 1-V_m\,V_{m-1}\,...\,V_{m-i+1}$ and for $i = 1,2,...,m$ .Then $U_1,U_2,...,U_m$ is the PT2C sample from U(0, 1) distribution.

iv. Finally, set $X_i = F^{-1}(U_i)$ for $i = 1,2,..., m$, where $F^{-1}(.)$ is the inverse CDF of EME distribution consideration, then $X_1, X_2,..., X_m$ are the required PT2C samples from EME distribution with censoring scheme $\underset{\sim}{R} = (R_1, R_2,..., R_m)$.

2. Three different sampling schemes are considered as follows:

Scheme I: $R_1 = R_2 =\ldots= R_{m-1}$ and $R_m = n-m$ (T2C),

Scheme II: $R_m = n-m$, $R_2 = R_3 =\ldots= R_m$=0 and

Scheme III: $R_1 = R_2 = (n-m)/2$, $R_3 = R_4 =\ldots= R_m$=0.

3. The parameters β and α are chosen with values; Case 1: $\beta = 0.5$, $\alpha = 1.5$ and Case 2: $\beta = 0.5$, $\alpha = 3$

4. With the mission time x = 0.8, the number of stages m, and the censoring strategy $\underset{\sim}{R} = (R_1, R_2,..., R_m)$, various sample sizes of n=50, 100, and 150 are chosen. The method described by Dey et al. [23] is used to choose the hyper-parameters for gamma priors

5. To create samples from the posterior distributions, the MH approach is applied.

6. The biases, mean squared errors (MSEs), average lengths (AILs), and CPs for MLEs and BEs are computed for various sample sizes, with the number of repeated samples being 1000 samples

7. A portion of the results, which are lengthy numerically, are shown Tables 1–3 for MLEs and BEs under IF.

Figures 2–8 provide examples from the investigation.

Regarding the behaviour of various estimations, the following findings are found.

- All the precision measures for MLEs and BEs tend to decrease with sample sizes n and number of stages m, in majority of the cases. The sample size n and number of stages m both enhance the CPs of the HRF estimates.
- Figure 2 shows that the MSEs of $\tilde{\alpha}$ and $\tilde{\beta}$, obtain the least values across all schemes, and the MSEs of $\hat{\alpha}$ and $\hat{\beta}$ in Case 1 get the biggest values across all schemes.

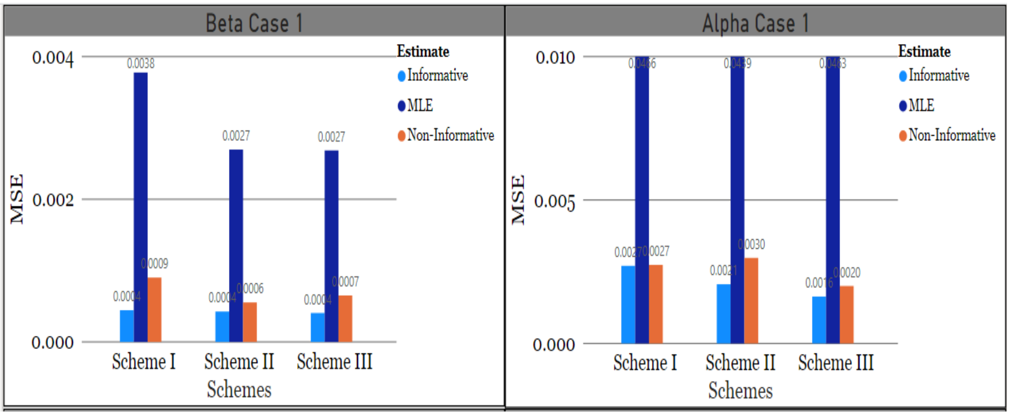


Figure 2: *MSEs for* α and β *estimates in Case 1 for all values of m*

- Figure 3 demonstrates that the MSEs of $\tilde{\alpha}$ and $\tilde{\beta}$ in Case 2 obtain the lowest values among all schemes, whereas the MSEs of $\hat{\alpha}$ and $\hat{\beta}$ obtain the highest values within all schemes

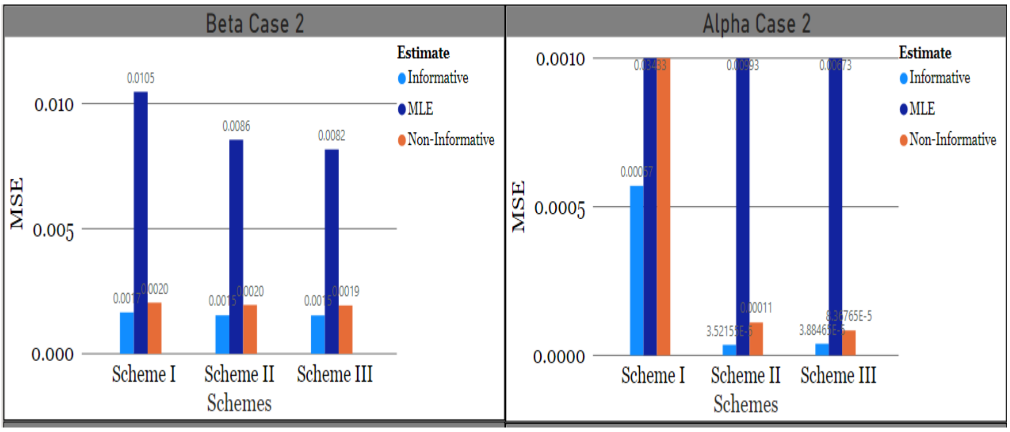


Figure 3: *MSEs for* α and β *estimates in Case 2 for all values of m*

- Regarding Case 1, in Figure 4, the MSEs of $\tilde{R}(x)$ and $\tilde{h}(x)$ in all schemes take the least value, while the MSEs of $\tilde{R}(x)$ and $\tilde{h}(x)$ receive the biggest value

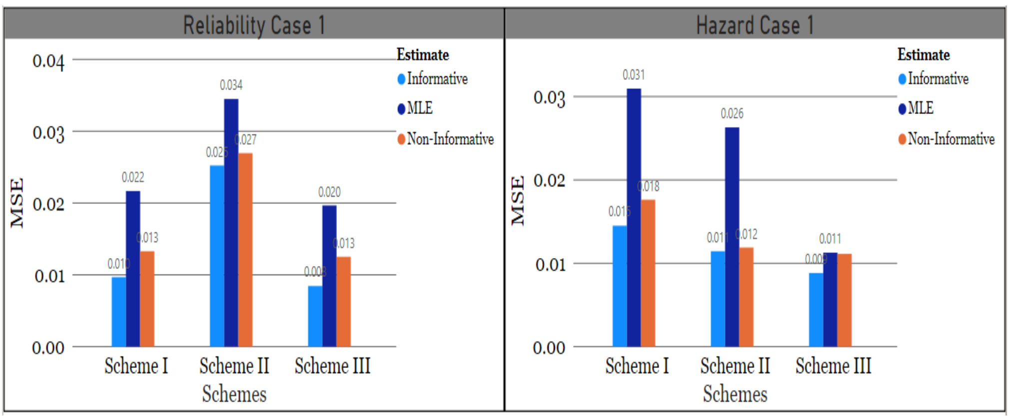


Figure 4: *MSEs for RF and HRF estimates in Case 1 for all values of m*

❖ The MSEs of $\tilde{R}(x)$ and $\tilde{h}(x)$, in all schemes, obtain the least values, as shown in Figure 5.

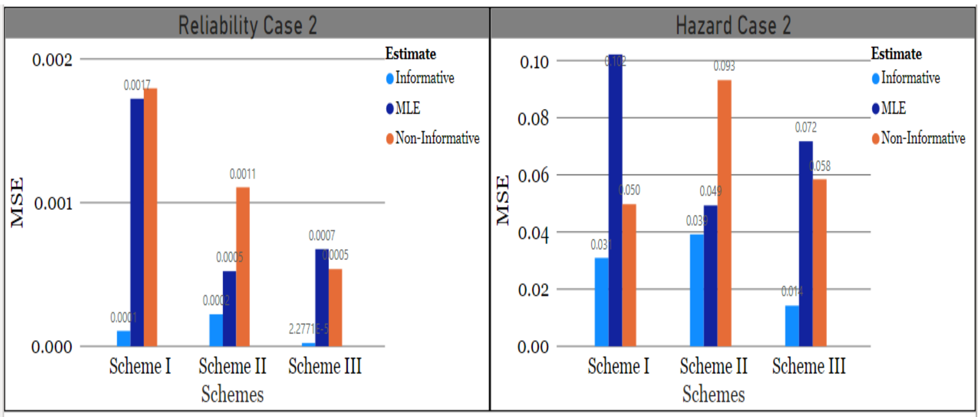


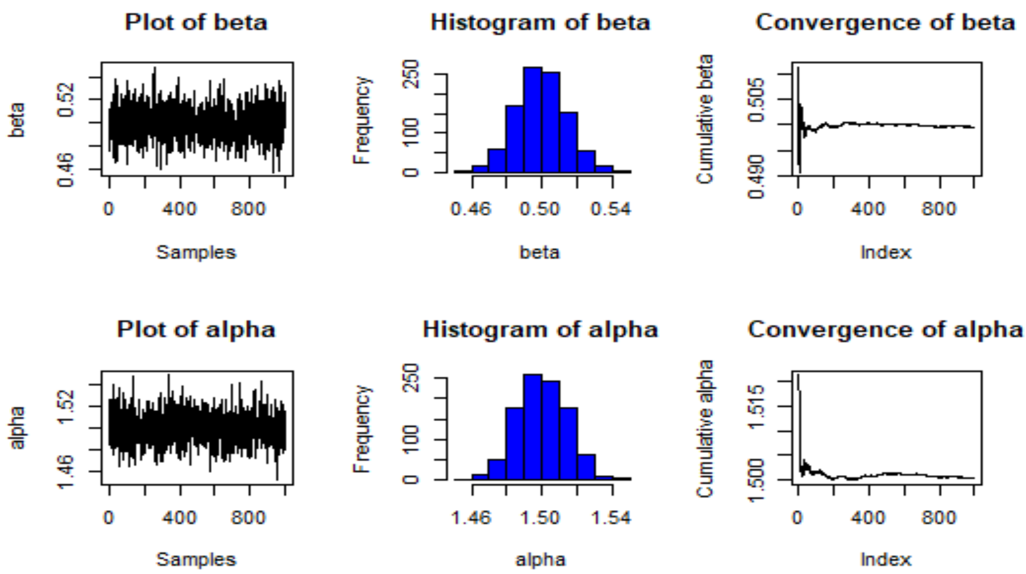
Figure 5: *MSEs for RF and HRF in Case 2 for all values of m*

❖ In most cases, it is possible to draw the conclusion that the MSEs of population parameters employing IF priors take the lowest values.

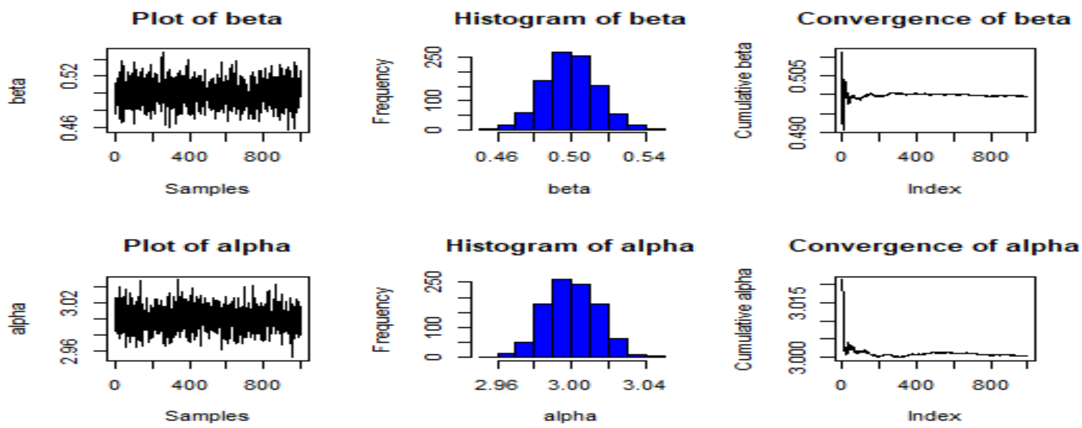
❖ The widths of the BCIs via IF priors are shorter than those of the MLEs and BEs under NIP priors in Case 1 (β =0.5, α =1.5).

❖ The CPs for BEs under IF priors are higher than the equivalent for MLEs and BEs under NIF priors.

❖ In Figure 6, for NIF prior, history graphs for various estimates of β and α are demonstrated. The plots of the parameter chains resemble a horizontal band without any discernible lengthy upward or downward trends, which are evidence of convergence.



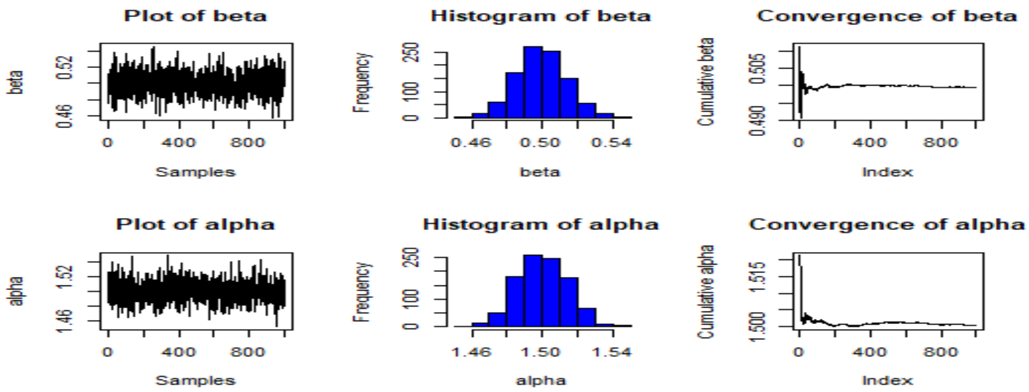
(a) $\ddot{\beta}$ and $\ddot{\alpha}$ at *n*=100, *m*= 50 for β =0.5, α =1.5



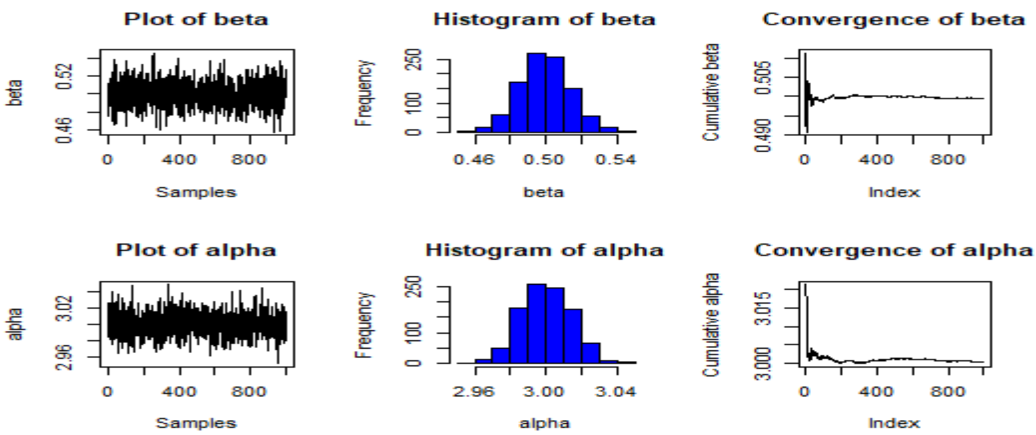
(b) *$\ddot{\beta}$ and $\ddot{\alpha}$ at n=100, m= 50 for β =0.5, $\alpha = 3$*

Figure 6: *Different BEs for β and α under NIF priors*

- In Figure 7, for IF priors, history graphs for various estimations of β and α are shown. The plots of the chains for the parameters resemble a horizontal band without any significant long-term rising or downward trends, which are signs of convergence



(a) $\tilde{\beta}$ and $\tilde{\alpha}$ at *n*=100, *m*= 50 for β =0.5, $\alpha = 1.5$



(b) *$\tilde{\beta}$ and $\tilde{\alpha}$ at n=100, m= 50 for β =0.5, $\alpha = 3$*

Figure 7: *Different Bayesian estimates for β and α under gamma priors*

Table 1: *MLEs and associated measures for* $\alpha, \beta, R(x)$ and $h(x)$ *in case 1*

Scheme I							
n	m	Estimate	Mean	Bias	MSE	AIL	CP
50	20	$\hat{\beta}$	0.490	0.010	0.017	0.511	95.6
		$\hat{\alpha}$	1.780	0.280	0.562	2.728	95.8
		$\hat{R}(x)$	0.842	0.169	0.029	0.299	95.0
		$\hat{h}(x)$	0.627	0.366	0.134	0.758	95.0
	30	$\hat{\beta}$	0.485	0.015	0.010	0.383	96.6
		$\hat{\alpha}$	1.734	0.234	0.471	2.531	96.2
		$\hat{R}(x)$	0.790	0.117	0.014	0.483	96.7
		$\hat{h}(x)$	0.729	0.264	0.070	1.118	96.7
100	20	$\hat{\beta}$	0.477	0.023	0.020	0.550	95.2
		$\hat{\alpha}$	1.810	0.310	0.560	2.670	95.6
		$\hat{R}(x)$	0.909	0.237	0.056	0.184	95.0
		$\hat{h}(x)$	0.470	0.523	0.273	0.737	95.0
	50	$\hat{\beta}$	0.492	0.008	0.006	0.297	96.9
		$\hat{\alpha}$	1.626	0.126	0.159	1.484	95.5
		$\hat{R}(x)$	0.772	0.100	0.010	0.469	96.0
		$\hat{h}(x)$	0.765	0.228	0.052	1.031	96.0
	70	$\hat{\beta}$	0.495	0.005	0.004	0.237	96.5
		$\hat{\alpha}$	1.582	0.082	0.098	1.184	96.7
		$\hat{R}(x)$	0.638	0.035	0.001	0.640	97.1
		$\hat{h}(x)$	0.978	0.015	0.000	1.108	97.1
150	50	$\hat{\beta}$	0.490	0.010	0.007	0.324	96.1
		$\hat{\alpha}$	1.615	0.115	0.138	1.384	95.8
		$\hat{R}(x)$	0.835	0.162	0.026	0.306	96.0
		$\hat{h}(x)$	0.676	0.317	0.100	0.767	96.0
	70	$\hat{\beta}$	0.491	0.009	0.004	0.258	96.8
		$\hat{\alpha}$	1.599	0.099	0.103	1.198	95.7
		$\hat{R}(x)$	0.826	0.153	0.023	0.393	97.1
		$\hat{h}(x)$	0.686	0.308	0.095	0.964	97.1
	100	$\hat{\beta}$	0.495	0.005	0.003	0.202	96.7
		$\hat{\alpha}$	1.562	0.062	0.070	1.011	95.3
		$\hat{R}(x)$	0.685	0.012	0.000	0.592	97.0
		$\hat{h}(x)$	0.915	0.078	0.006	1.103	97.0
	130	$\hat{\beta}$	0.498	0.002	0.002	0.165	96.8
		$\hat{\alpha}$	1.530	0.030	0.047	0.838	96.3
		$\hat{R}(x)$	0.583	0.089	0.008	0.756	96.9
		$\hat{h}(x)$	1.046	0.053	0.003	1.173	96.9

Continued Table 1

Scheme II							
n	m	Estimate	Mean	Bias	MSE	AIL	CP
50	20	$\hat{\beta}$	0.488	0.012	0.010	0.385	96.5
		$\hat{\alpha}$	1.673	0.173	0.314	2.090	96.8
		$\hat{R}(x)$	0.533	0.139	0.019	0.906	95.0
		$\hat{h}(x)$	1.092	0.099	0.010	1.491	95.0
	30	$\hat{\beta}$	0.494	0.006	0.007	0.335	95.8
		$\hat{\alpha}$	1.652	0.152	0.263	1.921	95.2
		$\hat{R}(x)$	0.577	0.096	0.009	0.898	96.7
		$\hat{h}(x)$	1.026	0.033	0.001	1.380	96.7
100	20	$\hat{\beta}$	0.498	0.002	0.010	0.385	95.9
		$\hat{\alpha}$	1.615	0.115	0.214	1.759	96.5
		$\hat{R}(x)$	0.491	0.181	0.033	0.893	95.0
		$\hat{h}(x)$	1.133	0.140	0.020	1.393	95.0
	50	$\hat{\beta}$	0.497	0.003	0.004	0.257	96.2
		$\hat{\alpha}$	1.585	0.085	0.123	1.336	94.8
		$\hat{R}(x)$	0.512	0.160	0.026	0.967	96.0
		$\hat{h}(x)$	1.110	0.117	0.014	1.573	96.0
	70	$\hat{\beta}$	0.497	0.003	0.000	0.085	97.0
		$\hat{\alpha}$	1.559	0.059	0.004	0.084	96.8
		$\hat{R}(x)$	0.549	0.123	0.015	0.938	100.0
		$\hat{h}(x)$	1.057	0.064	0.004	1.462	100.0
150	50	$\hat{\beta}$	0.490	0.010	0.004	0.251	96.5
		$\hat{\alpha}$	1.600	0.100	0.111	1.247	96.3
		$\hat{R}(x)$	0.499	0.173	0.030	0.859	96.0
		$\hat{h}(x)$	1.169	0.176	0.031	1.156	96.0
	70	$\hat{\beta}$	0.497	0.003	0.003	0.210	96.2
		$\hat{\alpha}$	1.557	0.057	0.079	1.081	95.8
		$\hat{R}(x)$	0.569	0.104	0.011	0.913	97.1
		$\hat{h}(x)$	1.043	0.050	0.002	1.312	97.1
	100	$\hat{\beta}$	0.500	0.000	0.002	0.183	97.0
		$\hat{\alpha}$	1.538	0.038	0.059	0.939	96.5
		$\hat{R}(x)$	0.472	0.201	0.040	0.949	97.0
		$\hat{h}(x)$	1.171	0.178	0.032	1.425	97.0
	130	$\hat{\beta}$	0.498	0.002	0.002	0.160	97.1
		$\hat{\alpha}$	1.536	0.036	0.044	0.809	96.0
		$\hat{R}(x)$	0.465	0.208	0.043	0.912	96.9
		$\hat{h}(x)$	1.193	0.200	0.040	1.280	96.9

Continued Table 1

Scheme III							
n	m	Estimate	Mean	Bias	MSE	AIL	CP
50	20	$\hat{\beta}$	0.483	0.017	0.010	0.392	96.1
		$\hat{\alpha}$	1.693	0.193	0.319	2.081	95.9
		$\hat{R}(x)$	0.566	0.107	0.011	0.938	95.0
		$\hat{h}(x)$	1.057	0.064	0.004	1.502	95.0
	30	$\hat{\beta}$	0.495	0.005	0.007	0.332	96.3
		$\hat{\alpha}$	1.645	0.145	0.254	1.892	95.9
		$\hat{R}(x)$	0.580	0.093	0.009	0.958	96.7
		$\hat{h}(x)$	0.988	0.006	0.000	1.626	96.7
100	20	$\hat{\beta}$	0.490	0.010	0.010	0.381	96.3
		$\hat{\alpha}$	1.617	0.117	0.191	1.653	95.9
		$\hat{R}(x)$	0.603	0.069	0.005	0.952	95.0
		$\hat{h}(x)$	0.958	0.035	0.001	1.614	95.0
	50	$\hat{\beta}$	0.496	0.004	0.004	0.253	96.4
		$\hat{\alpha}$	1.588	0.088	0.123	1.333	95.4
		$\hat{R}(x)$	0.559	0.114	0.013	0.919	96.0
		$\hat{h}(x)$	1.069	0.075	0.006	1.371	96.0
	70	$\hat{\beta}$	0.494	0.006	0.003	0.209	96.8
		$\hat{\alpha}$	1.581	0.081	0.093	1.151	95.6
		$\hat{R}(x)$	0.538	0.134	0.018	0.869	97.1
		$\hat{h}(x)$	1.100	0.106	0.011	1.227	97.1
150	50	$\hat{\beta}$	0.498	0.002	0.004	0.247	96.6
		$\hat{\alpha}$	1.560	0.060	0.092	1.166	96.8
		$\hat{R}(x)$	0.462	0.210	0.044	0.932	96.0
		$\hat{h}(x)$	1.173	0.180	0.032	1.444	96.0
	70	$\hat{\beta}$	0.500	0.000	0.003	0.216	96.8
		$\hat{\alpha}$	1.546	0.046	0.071	1.030	96.3
		$\hat{R}(x)$	0.536	0.136	0.019	0.947	97.1
		$\hat{h}(x)$	1.090	0.097	0.009	1.471	97.1
	100	$\hat{\beta}$	0.496	0.004	0.002	0.185	97.2
		$\hat{\alpha}$	1.556	0.056	0.064	0.967	95.4
		$\hat{R}(x)$	0.544	0.129	0.017	0.937	97.0
		$\hat{h}(x)$	1.081	0.088	0.008	1.423	97.0
	130	$\hat{\beta}$	0.497	0.003	0.002	0.160	96.3
		$\hat{\alpha}$	1.542	0.042	0.046	0.828	95.7
		$\hat{R}(x)$	0.485	0.188	0.035	0.960	96.9
		$\hat{h}(x)$	1.153	0.160	0.025	1.560	96.9

Table 2: *Bayes estimates and associated measures for* $\alpha, \beta, R(x)$ and $h(x)$ *in case 1 using IF priors*

Scheme I							
N	*m*	Estimate	Mean	Bias	MSE	CIL	CP
50	20	$\tilde{\beta}$	0.489	0.011	0.001	0.082	97.4
		$\tilde{\alpha}$	1.779	0.279	0.078	0.082	97.0
		$\tilde{R}(x)$	0.841	0.169	0.028	0.308	100.0
		$\tilde{h}(x)$	0.629	0.364	0.132	0.787	100.0
	30	$\tilde{\beta}$	0.484	0.016	0.001	0.080	97.0
		$\tilde{\alpha}$	1.734	0.234	0.055	0.082	98.5
		$\tilde{R}(x)$	0.789	0.116	0.014	0.483	96.7
		$\tilde{h}(x)$	0.732	0.261	0.068	1.100	100.0
100	20	$\tilde{\beta}$	0.477	0.023	0.001	0.084	96.7
		$\tilde{\alpha}$	1.809	0.309	0.096	0.080	96.7
		$\tilde{R}(x)$	0.909	0.237	0.056	0.186	100.0
		$\tilde{h}(x)$	0.471	0.522	0.273	0.745	100.0
	50	$\tilde{\beta}$	0.492	0.008	0.001	0.080	97.2
		$\tilde{\alpha}$	1.624	0.124	0.016	0.077	98.1
		$\tilde{R}(x)$	0.714	0.041	0.002	0.452	96.0
		$\tilde{h}(x)$	0.895	0.098	0.010	0.922	100.0
	70	$\tilde{\beta}$	0.495	0.005	0.000	0.077	97.6
		$\tilde{\alpha}$	1.580	0.080	0.007	0.084	97.4
		$\tilde{R}(x)$	0.638	0.035	0.001	0.640	97.1
		$\tilde{h}(x)$	0.977	0.016	0.000	1.099	98.6
150	50	$\tilde{\beta}$	0.491	0.009	0.001	0.081	98.9
		$\tilde{\alpha}$	1.619	0.119	0.015	0.082	96.4
		$\tilde{R}(x)$	0.793	0.120	0.014	0.281	96.0
		$\tilde{h}(x)$	0.779	0.214	0.046	0.598	100.0
	70	$\tilde{\beta}$	0.490	0.010	0.001	0.081	98.0
		$\tilde{\alpha}$	1.599	0.099	0.010	0.085	98.2
		$\tilde{R}(x)$	0.825	0.153	0.023	0.374	100.0
		$\tilde{h}(x)$	0.688	0.305	0.093	0.884	100.0
	100	$\tilde{\beta}$	0.496	0.004	0.000	0.078	96.9
		$\tilde{\alpha}$	1.549	0.049	0.003	0.084	96.8
		$\tilde{R}(x)$	0.697	0.024	0.001	0.571	99.0
		$\tilde{h}(x)$	0.899	0.094	0.009	1.114	100.0
	130	$\tilde{\beta}$	0.496	0.004	0.000	0.076	97.5
		$\tilde{\alpha}$	1.547	0.047	0.003	0.084	97.3
		$\tilde{R}(x)$	0.567	0.105	0.011	0.799	96.2
		$\tilde{h}(x)$	1.067	0.074	0.006	1.196	99.2

Continued Table 2

Scheme II							
n	m	Estimate	Mean	Bias	MSE	CIL	CP
50	20	$\tilde{\beta}$	0.486	0.014	0.001	0.080	97.4
		$\tilde{\alpha}$	1.673	0.173	0.030	0.085	97.5
		$\tilde{R}(x)$	0.532	0.140	0.020	0.965	100.0
		$\tilde{h}(x)$	1.097	0.104	0.011	1.595	100.0
	30	$\tilde{\beta}$	0.494	0.006	0.000	0.080	96.5
		$\tilde{\alpha}$	1.651	0.151	0.023	0.090	97.5
		$\tilde{R}(x)$	0.577	0.095	0.009	0.891	96.7
		$\tilde{h}(x)$	1.024	0.031	0.001	1.308	100.0
100	20	$\tilde{\beta}$	0.496	0.004	0.000	0.082	97.1
		$\tilde{\alpha}$	1.615	0.115	0.014	0.083	97.1
		$\tilde{R}(x)$	0.490	0.182	0.033	0.991	100.0
		$\tilde{h}(x)$	1.137	0.144	0.021	1.704	100.0
	50	$\tilde{\beta}$	0.493	0.007	0.000	0.083	98.8
		$\tilde{\alpha}$	1.590	0.090	0.009	0.080	97.8
		$\tilde{R}(x)$	0.522	0.151	0.023	0.922	96.0
		$\tilde{h}(x)$	1.118	0.125	0.016	1.333	100.0
	70	$\tilde{\beta}$	0.497	0.003	0.000	0.078	97.2
		$\tilde{\alpha}$	1.557	0.057	0.004	0.083	98.1
		$\tilde{R}(x)$	0.549	0.123	0.015	0.938	100.0
		$\tilde{h}(x)$	1.057	0.064	0.004	1.461	100.0
150	50	$\tilde{\beta}$	0.495	0.005	0.000	0.079	96.8
		$\tilde{\alpha}$	1.574	0.074	0.006	0.082	96.6
		$\tilde{R}(x)$	0.505	0.167	0.028	0.958	100.0
		$\tilde{h}(x)$	1.117	0.124	0.015	1.555	98.0
	70	$\tilde{\beta}$	0.496	0.004	0.000	0.082	97.7
		$\tilde{\alpha}$	1.555	0.055	0.004	0.085	96.9
		$\tilde{R}(x)$	0.567	0.105	0.011	0.911	98.6
		$\tilde{h}(x)$	1.046	0.053	0.003	1.319	97.1
	100	$\tilde{\beta}$	0.498	0.002	0.000	0.077	98.2
		$\tilde{\alpha}$	1.547	0.047	0.003	0.085	97.2
		$\tilde{R}(x)$	0.511	0.162	0.026	0.940	96.0
		$\tilde{h}(x)$	1.125	0.132	0.018	1.370	100.0
	130	$\tilde{\beta}$	0.497	0.003	0.000	0.075	97.7
		$\tilde{\alpha}$	1.540	0.040	0.002	0.082	97.2
		$\tilde{R}(x)$	0.486	0.186	0.035	0.925	98.5
		$\tilde{h}(x)$	1.164	0.170	0.029	1.330	96.9

Continued Table 2

Scheme III							
n	m	Estimate	Mean	Bias	MSE	CIL	CP
50	20	$\tilde{\beta}$	0.483	0.017	0.001	0.081	97.1
		$\tilde{\alpha}$	1.693	0.193	0.038	0.088	98.2
		$\tilde{R}(x)$	0.565	0.107	0.012	0.968	100.0
		$\tilde{h}(x)$	1.059	0.066	0.004	1.590	100.0
	30	$\tilde{\beta}$	0.494	0.006	0.000	0.083	98.9
		$\tilde{\alpha}$	1.644	0.144	0.021	0.080	97.4
		$\tilde{R}(x)$	0.579	0.094	0.009	0.959	100.0
		$\tilde{h}(x)$	0.991	0.003	0.000	1.626	100.0
100	20	$\tilde{\beta}$	0.488	0.012	0.001	0.079	96.4
		$\tilde{\alpha}$	1.617	0.117	0.014	0.084	97.1
		$\tilde{R}(x)$	0.602	0.070	0.005	0.973	100.0
		$\tilde{h}(x)$	0.962	0.031	0.001	1.672	100.0
	50	$\tilde{\beta}$	0.498	0.002	0.000	0.079	96.8
		$\tilde{\alpha}$	1.562	0.062	0.004	0.084	96.9
		$\tilde{R}(x)$	0.457	0.215	0.046	0.927	98.0
		$\tilde{h}(x)$	1.189	0.196	0.038	1.427	98.0
	70	$\tilde{\beta}$	0.494	0.006	0.000	0.077	97.1
		$\tilde{\alpha}$	1.580	0.080	0.007	0.089	97.3
		$\tilde{R}(x)$	0.538	0.135	0.018	0.865	100.0
		$\tilde{h}(x)$	1.102	0.109	0.012	1.254	97.1
150	50	$\tilde{\beta}$	0.497	0.003	0.000	0.079	96.7
		$\tilde{\alpha}$	1.559	0.059	0.004	0.082	97.3
		$\tilde{R}(x)$	0.472	0.200	0.040	0.912	96.0
		$\tilde{h}(x)$	1.172	0.179	0.032	1.271	100.0
	70	$\tilde{\beta}$	0.499	0.001	0.000	0.079	98.0
		$\tilde{\alpha}$	1.547	0.047	0.003	0.084	96.8
		$\tilde{R}(x)$	0.535	0.137	0.019	0.928	100.0
		$\tilde{h}(x)$	1.094	0.101	0.010	1.487	97.1
	100	$\tilde{\beta}$	0.495	0.005	0.000	0.076	97.6
		$\tilde{\alpha}$	1.552	0.052	0.003	0.082	98.0
		$\tilde{R}(x)$	0.515	0.158	0.025	0.887	99.0
		$\tilde{h}(x)$	1.127	0.134	0.018	1.416	97.0
	130	$\tilde{\beta}$	0.497	0.003	0.000	0.076	98.3
		$\tilde{\alpha}$	1.534	0.034	0.002	0.085	97.1
		$\tilde{R}(x)$	0.489	0.183	0.034	0.950	100.0
		$\tilde{h}(x)$	1.148	0.154	0.024	1.438	100.0

5. Discussion and Summary

This study uses maximum likelihood and Bayesian techniques to analyse parameter estimators, reliability function estimator, and hazard rate function estimator for EME distributions under PT2Cschemes. Gamma and uniform priors are taken into account under the squared error loss function to construct the Bayesian estimators. On the basis of IF and NIF priors, it is possible to derive approximate confidence intervals as well as Bayesian credible intervals. A simulation study is conducted to compare the effectiveness of every estimate. The Bayesian estimates using the gamma prior are, roughly speaking, generally more accurate than the MLEs, according to a numerical illustration. When compared to other schemes, Scheme I's MSEs have the highest value. Additionally, the MSEs for each estimate use the value for Scheme III that is the lowest. Comparatively speaking, the Bayesian estimates using gamma priors have the highest coverage probability.

References

[1] Balakrishnan, N. (2007). Progressive censoring methodology: an appraisal (with discussion). Test 16, 211–296.

[2] Hofmann, G., Cramer, E., Balakrishnan, N. and Kunert, G. (2005). An asymptotic approach to progressive censoring. Journal of Statistical Planning and Inference, 130(1), 207–227.

[3] Krishna, H., and Kumar, K. (2011). Reliability estimation in Lindley distribution with progressively type II right censored sample. Mathematics and Computers in Simulation, 82(2), 281–294.

[4] Balakrishnan, N., and Aggarwala, R. (2000). Progressive Censoring Theory, Methods and Applications. Birkhauser Boston, MA.

[5] Wu, S. J. (2002). Estimation of the parameters of the Weibull distribution with progressively c-ensored data. Journal of the Japan Statistical Society, 32(2), 155–163.

[6] Ng, H.K.T. (2005). Parameter estimation for a modified Weibull distribution, for progressively type-II censored samples. IEEE Transactions on Reliability, 54(3), 374–380.

[7] Dey, S., Singh, S., Tripathi, Y.M., and Asgharzadeh, A. (2016). Estimation and prediction for a progressively censored generalized inverted exponential distribution. Statistical Methodology, 32, 185–202.

[8] Hassan, A.S. and Abd-Alla, M. and El-Elaa, H.G.A. (2017). Estimation in step stress partially accelerated life test for exponentiated Pareto distribution under progressive censoring with random removal. Journal of Advances in Mathematics and Computer Science, 25(1), 1–16.

[9] EL-Sagheer, R.M. (2019). Estimating the parameters of Kumaraswamy distribution using progressively censored data. Journal of Testing and Evaluation, 47(2). https://doi.org/10.1520/JTE20150393

[10] Noor, F., Sajid, A., Ghazal, M., Khan, I, Zaman, M., and Baig, I. (2020). Bayesian estimation of Rayleigh distribution in the presence of outliers using progressive censoring. Hacettepe Journal of Mathematics & Statistics, 49 (6), 2119 – 2133.

[11] Alshenawy, R., Ali Al-Alwan, Ehab M. Almetwally, Ahmed Z. Afify, and Hisham M. Almongy. (2020). Progressive Type-II censoring schemes of extended odd Weibull exponential distribution with applications in medicine and engineering. Mathematics, 8(10)1679. https://doi. org/10.3390 /math 8101679

[12] Shrahili, M., El-Saeed, A.R., Hassan, A.S., Elbatal, I., and Elgarhy, M. (2022). Estimation of entropy for log-logistic distribution under progressive Type II censoring. Journal of Nanomaterials, Doi.org/10.1155/2022/2739606

[13] Dara, S.T., Ahmad, M. (2012). Recent Advances in Moment Distributions and their Hazard Rate", Ph.D. Thesis. National College of Business Administration and Economics, Lahore, Pakistan.

[14] Hasnain, S.A., Iqbal, Z., and Ahmad, M., (2015).On exponentiated moment exponential distribution. Pakistan Journal of Statistics, 31(2), 267–280.

[15] Gupta, R. D., and Kundu, D., (1999). Generalized exponential distribution. Australian and New Zealand Journal Statistical, 41(2), 173–188.

[16] Tripathi, Y.M., Kayal, T., and Dey, S. (2017). Estimation of the PDF and the CDF of exponentiated moment exponential distribution. International Journal of System Assurance Engineering and Management, 8(2), 1282–1296.

[17] Fatima, K. and Ahmad, S.P. (2018). Bayesian approach in estimation of shape parameter of the exponentiated moment exponential distribution. Journal of Statistical Theory and Applications, 17(2), 359–374.

[18] Akhter, Z., MirMostafaee, S. M.T.K. and Ormoz, E. (2022). On the order statistics of exponentiated moment exponential distribution and associated inference. Journal of Statistical Computation and Simulation, 92(6), 1322-1346, DOI: 10.1080/00949655.2021.1991927.

[19] Iqbal, Z. and Hasnain, S.A., Salman, M., Ahmad, M. and Hamedani, G. (2014). Generalized exponentiated moment exponential distribution. Pakistan Journal of Statistics, 30(4), 537–554.

[20] Ahmadini, A.A.H., Hassan, A.S., Mohamed, R.E., Alshqaq, S.S. and Nagy, H.F. (2021). A new four-parameter moment exponential model with applications to lifetime data. Intelligent Automation & Soft Computing,29 (1), 131−146.

[21] Shrahili, M., Hassan, A.S., Almetwally E.M., Ghorbal, A.B. and Elbatal, I. (2022). Alpha power moment exponential model with application to biomedical science. Scientific Programming. https:// doi.org/ 10.1155/ 2022 /6897405

[22] Balakrishnan, N., Sandhu, R.A.(1995). A simple simulation algorithm for generating progressively type II censored samples. American Statistical Association, 49(2), 229–230.

[23] Dey, S., Pradhan, B. (2014). Generalized inverted exponential distribution under hybrid censoring. Statistical Methodology. 18, 101–114.

RT&A, No 1 (72)
Volume 18, March 2023

Ratio Transformation Lomax Distribution with Applications

SHAMSHAD UR RASOOL *

•

Department of Statistics, University of Kashmir, Srinagar, India
srasool92@gmail.com

S.P.AHMAD

•

Department of Statistics, University of Kashmir, Srinagar, India
sprvz@yahoo.com

Abstract

It has been noted in the literature on probability theory that the classical probability distributions do not adequately fit real-world data and do not exhibit non-monotonic hazard rate behavior. To overcome this limitation, researchers are focusing on the improvement of these distributions. In this manuscript, we have introduced a new probability model called Ratio Transformation Lomax Distribution (RTLD) as a new generalization of Lomax distribution. A thorough mathematical analysis of the new distribution is provided in closed form such as density function, distribution function, the r-th moment, survival function, hazard function, moment generating function, generalized entropy and also the order statistics. The new model's parameters are calculated using the method of maximum likelihood estimation. The proposed distribution's performance and adaptability is backed by three sets of real lifetime data as well as simulated data.

Keywords: Ratio Transformation Lomax distribution, hazard rate function, moments, maximum likelihood estimation

1. Introduction

In several literary contexts, the Lomax distribution has been employed. It has been frequently utilised for reliability modelling and life testing. But it does not provide an acceptable fit for several applications, particularly when the risk rates include bimodal or bathtub-shaped hazards. To overcome these limitations, researchers have created a variety of extensions and changes to the Lomax distribution to model various sorts of data. In the statistical literature, a variety of probability models are available to simulate various real-life random processes. Every year more distinct models with high degrees of flexibility are developed because no single distribution can fully represent all phenomena. As a result, researchers are focusing on creating new families of distribution and releasing a new variety of families of distribution in order to more thoroughly analyse real-world data in a variety of applications. Among these, some of the extensions of the Lomax distribution found in the literature are exponentiated Weibull-Lomax distribution proposed by [7], power Lomax distribution introduced by [8], a new extension of Lomax distribution formulated by [5], Marshall-0lkin alpha power Lomax distribution presented by [4]. The generalization of probability models has been very popular in recent years. There are variety of approaches for generalizing probability distributions, such as Alpha Power Transformation (APT) proposed by [12], exponentiation, mixture and Weighted Technique ,Power Transformation, and several others. Recently, [11] proposed a new method for generating distributions known as Ratio Transformation(RT) method. In this manuscript, our motive is the generalization of Lomax distribution to develop the new probability model called as Ratio Transformation Lomax Distribution (RTLD) by using Ratio Transformation (RT) method. The primary justification for making this generalization is that RTLD's hazard rate displays a variety of complex shapes, such as constant, increasing-decreasing, decreasing-increasing, etc., which overcomes Lomax distribution's drawbacks. Additionally, when considering a real-world data sets, the new distribution performs better than the baseline distribution and certain well-known competitive models. The remaining portions of the manuscript are structured as follows: In section 2, the Ratio Transformation (RT) method is discussed. In section 3, the RTLD's pdf and cdf are defined, and its sub-cases are covered. In section 4, the reliability analysis of the RTLD is presented. In sections 5, 6, 7, and 8 the statistical properties ,generating functions, order statistics, and information measure of the RTLD are respectively discussed. A very effective method is used to carry out the parameter estimation in section 9. Sections 10, 11 and 12, respectively, provide information on the simulation study, applicability of RTLD and its conclusion.

2. Ratio Transformation (RT) Method

The Ratio Transformation (RT) family of probability distributions, as proposed by [11] is highlighted in this section. Suppose the continuous random variable X has cdf $F(x)$. Therefore, the RT of $F(x)$ denoted by $F_{RT}(x)$ for $x \in \mathbb{R}$ and is defined by

$$F_{RT}(X) = \frac{F(x)}{1+\eta-\eta^{F(x)}}\ ; \qquad \eta > 0 \tag{1}$$

The pdf of the Ratio Transformation(RT) distribution is defined as follows

$$f_{RT}(X) = f(x)\frac{\left(1+\eta-\eta^{F(x)}\left(1-F(x)log\eta\right)\right)}{\left(1+\eta-\eta^{F(x)}\right)^2}\ ; \qquad \eta > 0 \tag{2}$$

3. Ratio Transformation Lomax Distribution (RTLD)

Suppose the random variable X has the Lomax distribution with shape parameter β and scale parameter θ respectively, then its probability density function(pdf) and Cumulative distribution function (cdf) are respectively given by

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

$$F(x;\beta,\theta) = 1-\left(1+\frac{x}{\theta}\right)^{-\beta}\ ;\ x>0,\beta>0,\theta>0 \tag{4}$$

The RTLD is constructed from the Lomax distribution by using the (3) and (4)into (2) and(1) respectively. Therefore, the cdf of the RTLD is obtained as;

$$F_{RTLD}(x;\beta,\theta,\eta) = \frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}\ ; \qquad ;\ x>0,\eta>0,\beta>0,\theta>0 \tag{5}$$

and the corresponding pdf is

$$f_{RTLD}(x,\beta,\theta,\eta) = \frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)}\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta})log\eta\right)\right)}{\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^2}\ ;\ x>0,\eta>0,\beta>0,\theta>0 \tag{6}$$

Table 1: *Sub-Cases of RTLD*

η	θ	β	**Reduced Model**
-	1	-	Two parameter RTLD
1	-	-	Two parameter Lomax distribution
1	1	-	Beta Prime distribution
1	1	1	$F(2,2)$
1	$\frac{1}{\theta_q(q-1)}$	$\frac{(2-q)}{(q-1)}$	q-exponential distribution

Figure 1 and 2 have been displayed to provide a visual representation of the potential shapes of pdf and cdf of RTLD. Figure 3 represents the hazard rate plots of the RTLD for different parameter values.

Remark: For $\eta = 1$ in 6, RTLD becomes the two parametric Lomax distrbution. The important sub-cases of RTLD are presented in Table 1

4. Reliability analysis of the RTLD

This section primarily focuses on calculating the reliability (survival function), hazard rate (failure rate), reverse hazard function, cumulative hazard function, and mills ratio expressions for RTLD respectively.

4.1. Survival function

The survival function/reliability function is the complement of the cumulative distribution function and it is defined as the probability that a system will survive beyond a specified time. For the RTLD, the survival function denoted as $R_{RTLD}(x)$ is given by

$$R_{RTLD}(x) = 1 - F_{RTLD}(x;\beta,\theta,\eta) = \frac{\eta\left(1-\eta^{-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right)+\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \tag{7}$$

4.2. Hazard Rate

Hazard rate also known as hazard function , force of mortality or failure rate. The expression for the hazard rate of RTLD is expressed as

$$h(x;\eta,\beta,\theta) = \frac{f_{RTLD}(x,\beta,\theta,\eta)}{R_{RTLD}(x,\beta,\theta,\eta)}$$

$$h(x;\eta,\beta,\lambda) = \frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)}\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta})log\eta\right)\right)\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^{-1}}{\eta\left(1-\eta^{-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right)+\left(1+\frac{x}{\theta}\right)^{-\beta}} \tag{8}$$

4.3. Reverse Hazard function

The reverse hazard function for the RTLD is expressed as

$$h_r(x;\eta,\beta,\theta) = \frac{f_{RTLD}(x,\beta,\theta,\eta)}{F_{RTLD}(x;\beta,\theta,\eta)}$$

Using equation (6) and(5) , the reverse hazard function for the RTLD is obtained as

$$h_r(x;\eta,\beta,\theta)=\frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)}\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta})log\eta\right)\right)\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^{-1}}{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} \tag{9}$$

4.4. Cumulative Hazard function

The Cumulative hazard function for the RTLD is obtained as

$$\Lambda_{RTLD}(x;\eta,\beta,\theta)=-\log R_{RTLD}(x)$$

$$\Lambda_{RTLD}(x;\eta,\beta,\theta)=\log\left\{\frac{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{\eta\left(1-\eta^{-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right)+\left(1+\frac{x}{\theta}\right)^{-\beta}}\right\} \tag{10}$$

4.5. Mills Ratio

The Mills ratio for the RTLD is obtained as

$$M.R=\frac{F_{RTLD}(x;\beta,\theta,\eta)}{R_{RTLD}(x)}=\left\{\frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{\eta\left(1-\eta^{-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right)+\left(1+\frac{x}{\theta}\right)^{-\beta}}\right\} \tag{11}$$

5. STATISTICAL PROPERTIES OF RTLD

This part focuses on discussing the related measures that are connected to the formulated model, including the raw moments, central moments, pearson's coefficients, coefficient of variation, and index of dispersion.

5.1. Raw Moments

The r^{th}moment of the RTLD about origin μ_r' is given by

$$\mu_r'=E(x^r)=\int_0^\infty x^r f_{RTLD}(x,\beta,\theta,\eta)dx$$

$$\mu_r'=\int_0^\infty x^r\frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)}\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta})log\eta\right)\right)}{\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^2}dx \tag{12}$$

Here, r^{th} moment of the RTLD is obtained by using the following series representations.

$$\eta^{-x}=\sum_{k=0}^{\infty}\frac{(-log\eta)^k x^k}{k!} \tag{13}$$

$$(1-x)^{-2}=\sum_{k=0}^{\infty}(k+1)x^k\ ;\qquad |x|<1, \tag{14}$$

$$(1-x)^{-1}=\sum_{k=0}^{\infty}x^k\ ;\qquad |x|<1, \tag{15}$$

By substituting $y = \left(1+\frac{x}{\theta}\right)^{-\beta}$ in (12) and solving the integral further,we obtain r^{th}moment of the RTLD about origin $\mu_r^{'}$ as

$$\mu_r^{'} = \sum_{k,m=0}^{\infty} \frac{\beta\theta^r \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A + (1+k)^{m+1} log\eta \left[A - C\right] \right\} \tag{16}$$

where

$A = B(r+1, \beta(m+1) - r)$ and $C = B(r+1, \beta(m+2) - r)$ represents the beta functions of second type. Using equation (16) and substituting $r = 1, 2, 3, 4$, the first four moments about origin of the RTLD are obtained as

$$\mu_1^{'} = \sum_{k,m=0}^{\infty} \frac{\beta\theta \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A_1^{'} + (1+k)^{m+1} log\eta \left[A_1^{'} - C_1^{'}\right] \right\} \tag{17}$$

where

$$A_1^{'} = B(2, \beta(m+1) - 1)$$

$$C_1^{'} = B(2, \beta(m+2) - 1)$$

The equation (17) represents the mean of the RTLD.

$$\mu_2^{'} = \sum_{k,m=0}^{\infty} \frac{\beta\theta^2 \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A_2^{'} + (1+k)^{m+1} log\eta \left[A_2^{'} - C_2^{'}\right] \right\} \tag{18}$$

where

$$A_2^{'} = B(3, \beta(m+1) - 2)$$

$$C_2^{'} = B(3, \beta(m+2) - 2)$$

$$\mu_3^{'} = \sum_{k,m=0}^{\infty} \frac{\beta\theta^3 \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A_3^{'} + (1+k)^{m+1} log\eta \left[A_3^{'} - C_3^{'}\right] \right\} \tag{19}$$

where

$$A_3^{'} = B(4, \beta(m+1) - 3)$$

$$C_3^{'} = B(4, \beta(m+2) - 3)$$

$$\mu_4^{'} = \sum_{k,m=0}^{\infty} \frac{\beta\theta^4 \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A_4^{'} + (1+k)^{m+1} log\eta \left[A_4^{'} - C_4^{'}\right] \right\} \tag{20}$$

where

$$A_4^{'} = B(5, \beta(m+1) - 4)$$

$$C_4^{'} = B(5, \beta(m+2) - 4)$$

5.2. Moments about Mean (Central Moments)

The moments about the mean also known as central moments of RTLD are obtained as

$$\mu_2 = \mu_2^{'} - (\mu_1^{'})^2$$

$$\mu_2 = \sum_{k,m=0}^{\infty} \frac{\beta\theta^2\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_2^{'} + (1+k)^{m+1}log\eta \left[A_2^{'} - C_2^{'} \right] \right\}$$
$$- \left\{ \sum_{k,m=0}^{\infty} \frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_1^{'} + (1+k)^{m+1}log\eta \left[A_1^{'} - C_1^{'} \right] \right\} \right\}^2 \quad (21)$$

The equation (21) represents the variance of RTLD.

$$\mu_3 = \sum_{k,m=0}^{\infty} \frac{\beta\theta^3\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_3^{'} + (1+k)^{m+1}log\eta \left[A_3^{'} - C_3^{'} \right] \right\}$$
$$-3 \left(\sum_{k,m=0}^{\infty} \frac{\beta\theta^2\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_2^{'} + (1+k)^{m+1}log\eta \left[A_2^{'} - C_2^{'} \right] \right\} \right)$$
$$\left(\sum_{k,m=0}^{\infty} \frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_1^{'} + (1+k)^{m+1}log\eta \left[A_1^{'} - C_1^{'} \right] \right\} \right)$$
$$+2 \left\{ \sum_{k,m=0}^{\infty} \frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_1^{'} + (1+k)^{m+1}log\eta \left[A_1^{'} - C_1^{'} \right] \right\} \right\}^3 \quad (22)$$

$$\mu_4 = \sum_{k,m=0}^{\infty} \frac{\beta\theta^4\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_4^{'} + (1+k)^{m+1}log\eta \left[A_4^{'} - C_4^{'} \right] \right\}$$
$$-4 \left(\sum_{k,m=0}^{\infty} \frac{\beta\theta^3\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_3^{'} + (1+k)^{m+1}log\eta \left[A_3^{'} - C_3^{'} \right] \right\} \right)$$
$$\left(\sum_{k,m=0}^{\infty} \frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_1^{'} + (1+k)^{m+1}log\eta \left[A_1^{'} - C_1^{'} \right] \right\} \right)$$
$$+6 \left(\sum_{k,m=0}^{\infty} \frac{\beta\theta^2\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_2^{'} + (1+k)^{m+1}log\eta \left[A_2^{'} - C_2^{'} \right] \right\} \right)$$
$$\left(\sum_{k,m=0}^{\infty} \frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_1^{'} + (1+k)^{m+1}log\eta \left[A_1^{'} - C_1^{'} \right] \right\} \right)$$
$$-3 \left\{ \sum_{k,m=0}^{\infty} \frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!} \left\{ k^m(1+\eta)A_1^{'} + (1+k)^{m+1}log\eta \left[A_1^{'} - C_1^{'} \right] \right\} \right\}^4 \quad (23)$$

As a result, these equations may be used to calculate the skewness measure, kurtosis, coefficient of variation and index of dispersion for the RTLD.

5.3. Pearson's Coefficients

The following four coefficients can be obtained for the RTLD based upon the first four moments about the mean using the above section as:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

$$\gamma_1 = \sqrt{\beta_1}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$\gamma_2 = \beta_2 - 3$$

5.4. Coefficient of Variation

$$CV = \frac{\sqrt{\mu_2}}{\mu_1^{'}}$$

On using the equations (17) and(21), the coefficient of variation can be obtained for RTLD.

5.5. Index of Dispersion

The index of dispersion is defined as :

$$D = \frac{\mu_2}{\mu_1^{'}}$$

On using the equations (17) and(21), the index of dispersion can be obtained for RTLD.

6. Generating Functions RTLD

6.1. Moment Generating Function

Moment generating function (MGF) is used to represent all the moments of a distribution. The MGF for RTLD distribution is given in the following theorem.

Theorem 1. Let X follows the RTLD distribution, then the moment generating function, $M_X(t)$ is

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \sum_{k,m=0}^{\infty} \frac{\beta\theta^r \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A + (1+k)^{m+1} log\eta \left[A - C\right] \right\} \tag{24}$$

Proof: The moment generating function of RTLD distribution is defined as

$$M_x(t) = \int_0^{\infty} e^{tx} f(x) dx$$

Using the series representation of e^{tx}, we have

$$\sum_{r=0}^{\infty} \frac{t^r}{r!} \int_0^{\infty} x^r f(x; \eta, \beta, \theta) dx$$

Using equation (16) we obtain the moment generating function for RTLD as

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \sum_{k,m=0}^{\infty} \frac{\beta\theta^r \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A + (1+k)^{m+1} log\eta \left[A - C\right] \right\} \tag{25}$$

6.2. Characteristic Function

The characteristic function for RTLD distribution is given in the following theorem.

Theorem 2. Let X follows the RTLD distribution, then the characteristic function, $\phi_X(t)$is

$$\phi_X(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \sum_{k,m=0}^{\infty} \frac{\beta\theta^r \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A + (1+k)^{m+1} log\eta \left[A - C\right] \right\} \tag{26}$$

Proof: The characteristic function for the RTLD can be obtained using the relation $\phi_X(t) = M_x(it)$

$$\phi_X(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \sum_{k,m=0}^{\infty} \frac{\beta\theta^r \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A + (1+k)^{m+1} log\eta \left[A - C\right] \right\} \tag{27}$$

6.3. Cumulant Function

The cumulant function for the RTLD can be obtained using the relation $k_x(t) = \log M_x(t)$

$$k_v(t) = \log \sum_{r=0}^{\infty} \frac{t^r}{r!} \sum_{k,m=0}^{\infty} \frac{\beta\theta^r \eta^k (-log\eta)^m}{(1+\eta)^{k+2} m!} \left\{ k^m (1+\eta) A + (1+k)^{m+1} log\eta \left[A - C \right] \right\} \tag{28}$$

7. Order Statistics of RTLD

The order statistics connected to the RTLD is devoted in this section. Let $X_{(t;n)}$ be the t^{th} order statistics with the random sample $x_{(1)}, x_{(2)}, x_{(3)}, ...x_{(m)}$ derived from the RTLD having the probability density function (pdf) $f(x;\eta,\beta,\theta)$ and cumulative distribution function (cdf) $F(x;\eta,\beta,\theta)$. Therefore, the probability density function (pdf) and cumulative distribution function (cdf) of $x_{(t;n)}$ say $f_{(t;n)}(x)$ and $F_{(t;n)}(x)$ respectively is defined as

$$f_{(t;n)}(x) = \frac{n!}{(t-1)!(n-t)!} \left[F(x;\eta,\beta,\theta) \right]^{t-1} \left[1 - F(x;\eta,\beta,\theta) \right]^{n-t} f(x;\eta,\beta,\theta) \tag{29}$$

$$F_{(t;n)}(x) = \sum_{j=t}^{n} \binom{n}{j} \left[F(x;\eta,\beta,\theta) \right]^{j} \left[1 - F(x;\eta,\beta,\theta) \right]^{n-j} \tag{30}$$

Using equation(5) and equation(6) in equation(29) and equation(30), the pdf and cdf of t^{th} ordered statistics for the RTLD is derived and is expressed as

$$f_{(t;n)}(x) = \frac{n!}{(t-1)!(n-t)!} \left[\frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{t-1} \left[\frac{\left(1+\frac{x}{\theta}\right)^{-\beta}+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{n-t}$$
$$\left\{ \frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)} \left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} \left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta}) log\eta\right)\right)}{\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^2} \right\} \tag{31}$$

$$F_{(t;n)}(x) = \sum_{j=t}^{n} \binom{n}{j} \left[\frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{j} \left[\frac{\left(1+\frac{x}{\theta}\right)^{-\beta}+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{n-j} \tag{32}$$

In order to obtain the expression for pdf of smallest(minimum) order statistics $x_{(1)}$and the largest (maximum) order statistics $x_{(m)}$ of RTLD , we assume $t = 1$ and n respectively and is expressed in the form as

$$f_{(1;n)}(x) = n \left[\frac{\left(1+\frac{x}{\theta}\right)^{-\beta}+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{n-1} \left\{ \frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)} \left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} \left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta}) log\eta\right)\right)}{\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^2} \right\} \tag{33}$$

$$f_{(n;n)}(v) = n \left[\frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{n-1} \left\{ \frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)} \left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} \left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta}) log\eta\right)\right)}{\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^2} \right\} \tag{34}$$

7.1. Median order statistics

Theorem 3. The Pdf of median order statistics for the RTLD is given as

$$f_{(n+1;n)}(x) = \frac{(2n+1)!}{(n)!(n)!} \left[\frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{n} \left[\frac{\left(1+\frac{x}{\theta}\right)^{-\beta}+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \right]^{n}$$
$$\left\{ \frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)} \left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} \left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta}) log\eta\right)\right)}{\left(1+\eta-\eta^{1-(1+\frac{x}{\theta})^{-\beta}}\right)^2} \right\} \tag{35}$$

Proof The pdf of median order statistics, $x_{(n+1)}$ is defined as

$$f_{(n+1;n)}(x) = \frac{(2n+1)!}{n!n!}\left[F(x;\eta,\beta,\theta)\right]^n \left[1 - F(v;\eta,\beta,\theta)\right]^n f(v;\eta,\beta,\theta)$$

$$f_{(n+1;n)}(x) = \frac{(2n+1)!}{(n)!(n)!}\left[\frac{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}\right]^n \left[\frac{\left(1+\frac{x}{\theta}\right)^{-\beta}+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}}\right]^n \left\{\frac{\frac{\beta}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\beta+1)}\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta})log\eta\right)\right)}{\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right)^2}\right\} \tag{36}$$

8. Information measure of RTLD

Entropy is a quantitative measures of the amount of uncertainty in a random variable. In this section we derive the expression for generalized entropy of RTLD.

Theorem 4. The generalized entropy for the RTLD is expressed as

$$I(\alpha) = \frac{1}{\alpha(\alpha-1)}\left\{\frac{\sum_{k,m=0}^{\infty}\frac{\beta\theta^{\alpha}\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!}\left\{k^m(1+\eta)A+(1+k)^{m+1}log\eta\left[A-C\right]\right\}}{\left\{\sum_{k,m=0}^{\infty}\frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!}\left\{k^m(1+\eta)A_1^{'}+(1+k)^{m+1}log\eta\left[A_1^{'}-C_1^{'}\right]\right\}\right\}^{\alpha}}-1\right\} \tag{37}$$

Proof: The generalized entropy is defined as

$$I(\alpha) = \frac{x_\alpha \mu^{-\alpha}-1}{\alpha(\alpha-1)}$$

where

$$x_\alpha = \int_{-\infty}^{\infty} x^\alpha f(x)dx$$

and μ represents mean. For RTLD, we have

$$x_\alpha = \sum_{k,m=0}^{\infty}\frac{\beta\theta^{\alpha}\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!}\left\{k^m(1+\eta)A+(1+k)^{m+1}log\eta\left[A-C\right]\right\} \tag{38}$$

where

$A = B(\alpha+1, \beta(m+1)-\alpha)$ and $C = B(\alpha+1, \beta(m+2)-\alpha)$ represents the beta functions of second type.

$$\mu^{-\alpha} = \left\{\sum_{k,m=0}^{\infty}\frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!}\left\{k^m(1+\eta)A_1^{'}+(1+k)^{m+1}log\eta\left[A_1^{'}-C_1^{'}\right]\right\}\right\}^{-\alpha} \tag{39}$$

where

$$A_1^{'} = B(2, \beta(m+1)-1)$$

$$C_1^{'} = B(2, \beta(m+2)-1)$$

Therefore, the expression for the generalized entropy of RTLD is obtained as

$$I(\alpha) = \frac{1}{\alpha(\alpha-1)}\left\{\frac{\sum_{k,m=0}^{\infty}\frac{\beta\theta^{\alpha}\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!}\left\{k^m(1+\eta)A+(1+k)^{m+1}log\eta\left[A-C\right]\right\}}{\left\{\sum_{k,m=0}^{\infty}\frac{\beta\theta\eta^k(-log\eta)^m}{(1+\eta)^{k+2}m!}\left\{k^m(1+\eta)A_1^{'}+(1+k)^{m+1}log\eta\left[A_1^{'}-C_1^{'}\right]\right\}\right\}^{\alpha}}-1\right\} \tag{40}$$

9. Estimation of Parameters

This section is devoted to maximum likelihood estimation procedure for estimating unknown parameters η, β, θ of RTLD.

9.1. Maximum Likelihood Estimation(MLE)

Suppose $x_1, x_2, x_3, ...x_n$ be the random sample derived from the RTLD having the probability density function (pdf) $f(x; \eta, \beta, \theta)$. Therefore, for n observations , the logarithm of the likelihood function of RTLD is obtained as

$$l = nlog\beta + n\beta log\theta - (\beta+1)\sum_{i=1}^{n} log(x_i+\theta) - 2\sum_{i=1}^{n} log\left(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right) + \sum_{i=1}^{n} log\left[1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-log\eta\ (1-\left(1+\frac{x}{\theta}\right)^{-\beta})\right)\right] \quad (41)$$

The MLEs of η, θ and β are obtained by partially differentiating (41) with respect to the corresponding parameters and equating to zero, we have

$$\frac{\partial l}{\partial \eta} = \sum_{i=1}^{n} \frac{1+(1-\left(1+\frac{x}{\theta}\right)^{-\beta})^2\eta^{-\left(1+\frac{x}{\theta}\right)^{-\beta}} log\eta}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-log\eta(1-\left(1+\frac{x}{\theta}\right)^{-\beta})\right)} - 2\sum_{i=1}^{n} \frac{1-(1-\left(1+\frac{x}{\theta}\right)^{-\beta})\eta^{-\left(1+\frac{x}{\theta}\right)^{-\beta}}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} \quad (42)$$

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} + nlog\theta - \sum_{i=1}^{n} log(x_i+\theta) + 2\sum_{i=1}^{n} \frac{\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} log\eta\ \left(1+\frac{x}{\theta}\right)^{-\beta} log\left(1+\frac{x}{\theta}\right)}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} - \sum_{i=1}^{n} \frac{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}(1+\frac{x}{\theta})^{-\beta} log\eta(1+\frac{x}{\theta}) + \left(1-log\eta\ (1-\left(1+\frac{x}{\theta}\right)^{-\beta})\right)\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} log\eta(1+\frac{x}{\theta})^{-\beta} log(1+\frac{x}{\theta})}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-log\eta(1-\left(1+\frac{x}{\theta}\right)^{-\beta})\right)} \quad (43)$$

$$\frac{\partial l}{\partial \theta} = \frac{n\beta}{\theta} - (\beta+1)\sum_{i=1}^{n}\frac{1}{(x_i+\theta)} - 2\sum_{i=1}^{n} \frac{\beta\theta^{-2}x\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}} log\eta\ \left(1+\frac{x}{\theta}\right)^{-(\beta+1)}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}} - \sum_{i=1}^{n} \frac{\left[\beta\theta^{-2}xlog\eta\ \left(1+\frac{x}{\theta}\right)^{-(\beta+1)}\right]\left\{(1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}) + \left(1-log\eta\ (1-\left(1+\frac{x}{\theta}\right)^{-\beta})\right)\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\right\}}{1+\eta-\eta^{1-\left(1+\frac{x}{\theta}\right)^{-\beta}}\left(1-log\eta(1-\left(1+\frac{x}{\theta}\right)^{-\beta})\right)} \quad (44)$$

The above three non-linear equations (42),(43) and (44) are not in closed form.Therefore,we shall solve these equations with the help of R software.

10. SIMULATION ILLUSTRATION

In this section, the effectiveness of the (MLEs) of RTLD is explored. To demonstrate the behavior of MLEs in terms of random generating sample sizes of n= 50, 100, 200, 300, 400 a simulation research was conducted using R Software. The procedure is repeated 500 times. Different sets of parameter combinations are selected as (1,0.5,1)and (0.5,1,1) with reference to the usual order (θ, η, β). The average MLE values, bias, and related empirical mean squared errors (MSEs) were determined for each scenario. From Table 2 and Table 3 the simulation findings are shown. The estimates are stable and near to the genuine parameter values, as presented in Tables 2 and 3. In all circumstances, the MSE drops as the sample size increases.

Table 2: *Results of the simulation study for the RTLD model at parameter combination set as* $(\theta = 1, \eta = 0.5, \beta = 1)$.

Sample	MLE			BIAS			MSE		
n	$\hat{\theta}$	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\eta}$	$\hat{\beta}$
50	1.88737	0.81308	1.48624	0.88737	0.31308	0.48624	5.31104	4.05814	1.24
100	1.36960	0.77404	1.23155	0.369604	0.27404	0.23155	1.26655	3.01730	0.32662
200	1.17796	0.61684	1.12613	0.177968	0.11684	0.126133	0.456912	0.24202	0.085288
300	1.09238	0.59553	1.07560	0.09238	0.0955	0.07560	0.381143	0.08623	0.06454
400	1.02375	0.59153	1.02596	0.02375	0.09153	0.02596	0.23470	0.06871	0.033937

Table 3: *Results of the simulation study for the RTLD model at parameter combination set as* $(\theta = 0.5, \eta = 1, \beta = 1)$.

Sample	MLE			BIAS			MSE		
n	$\hat{\theta}$	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\eta}$	$\hat{\beta}$
50	0.82485	1.9607	1.36407	0.3248	0.9607	0.36407	1.07362	6.2609	0.59748
100	0.78055	1.64962	1.20104	0.28055	0.6496	0.2010	1.02573	4.14133	0.25056
200	0.60470	1.16408	1.07572	0.1047	0.1640	0.07572	0.16791	1.0627	0.04460
300	0.59327	1.14369	1.06681	0.0932	0.14369	0.06681	0.122244	0.50144	0.033350
400	0.56298	1.13550	1.0422	0.0629	0.135501	0.04229	0.0884	0.5013	0.01751

11. Application

This section concentrates on application of the proposed model to real life data sets. The significance and superiority of RTLD are highlighted in this part by the use of three real-life data sets. The MLEs of the model parameters are computed along with the corresponding Standard Error (SE) and goodness-of-fit statistics for these models are compared with other competing models. We compare the fits of the RTLD distribution with some competitive models which are listed in Table 4. To choose the best model among the compared models, performance comparing tools such as Akaike Information Criteria (AIC),Bayesian Information Criteria (BIC) and Akaike Information Criteria Corrected (AICc) are exploited. These Criterions choose the superior distribution as the one which is having the smallest value of AIC,BIC,and AICc. Furthermore, the Kolmogorov -Smirnov (KS)-distance and associated p- value is obtained to assess the goodness of fit. The superior probability model is considered the one which is having the least value of KS and maximum value of p- value.
The performance comparing tools are mentioned below:

- Akaike Information Criterion(AIC) is calculated as

$$\text{AIC} = -2\hat{l} + 2m$$

- Bayesian Information Criterion (BIC) is defined as

$$\text{BIC} = -2\hat{l} + m\ln(n)$$

- Akaike Information Criterion Corrected(AICC) is defined as.

$$\text{AICC} = \text{AIC} + \frac{2m(m+1)}{n-m-1}$$

where

$\hat{l}$ is the log-likelihood function of the model given the data.
m represents the number of parameters involved in the given model.
n is the sample size.

Table 5 and Table 6 displays the MLE's with corresponding Standard Error (SE) and the comparison of performance of RTLD with compared distributions for data set 1 ,which represents the COVID-19 vaccination rate from different countries. Table 7 and Table 8 presents the MLE's along with corresponding Standard

Table 4: *Competitive models of the RTLD model.*

Competitive models of the RTLD model	
Distribution(s)	**Author(s)**
(1) Sine Power Lomax (SPL)	[14]
(2) Length Biased Weighted Lomax Distribution(LBWLD)	[2]
(3) Topp-Leone Lomax (TLLo)	[15]
(4) Power Lomax (PL)	[16]
(5) Exponentiated Lomax (EL)	[1]
(6) Weibull Lomax(WL)	[17]
(7) Lomax (L)	[10]

Table 5: *MLE's of RTLD and compared distributions with corresponding standard error (given in parenthesis) for Covid -19 vaccination rate data set .*

Model	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\lambda}$
RTLD	0.3446	2.7564	25.9981		
	(0.1854)	(2.3906)	(31.3471)	-	-
PL		0.9390		1.6917	6.7054
	-	(0.2478)	-	(1.6295)	(6.9138)
SPL		0.9402		0.7755	0.2037
	-	(0.2714)	-	(0.7075)	(0.2041)
EL		0.1952		1.3525	1.0322
	-	(0.2336)	-	(0.6882)	(0.3999)
TLLo		0.6762		0.1952	1.0322
	-	(0.3441)	-	(0.2336)	(0.3999)
L		5.5782	1.3924		
	-	(3.3975)	(0.5346)	-	-

Table 6: *Comparison of RTLD and compared distributions for Covid -19 vaccination rate data set*

Model	-2$\hat{l}$	AIC	AICC	BIC	K-S	p-value
RTLD	**284.6826**	**290.6826**	**291.2540**	**296.1685**	**0.0845**	**0.8697**
PL	285.6881	291.6881	292.2595	297.1740	0.1013	0.6938
SPL	285.8228	291.8229	292.3943	297.3088	0.1030	0.6751
EL	285.8365	291.8365	292.4079	297.3224	0.10699	0.6294
TLLo	285.8365	291.8365	292.4079	297.3224	0.10699	0.6294
L	288.7432	292.7432	293.0222	296.4004	0.108	0.6124

Error (SE) and comparison of performance of RTLD with compared distributions for data set 2. In addition to these, the result findings of the RTLD for the data set 3 , that represents the organic carbon content percentage in the soil of the district Ganderbal and are discussed in Table 9 and Table 10. The results shown in Table 6, Table 8 and 10 reveals that RTLD is having a minimum value of AIC ,BIC and AICC, and thus

Table 7: *MLE's of RTLD and compared distributions with corresponding standard error (given in parenthesis) for the dataset of the life of fatigue fracture of Kevlar 373/epoxy data .*

Model	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\lambda}$
RTLD	20.9773	3.31514	0.635077		
	(17.388)	(0.892)	(0.550)	-	-
SPL		1.543		1.768	0.127
	-	(0.246)	-	(1.385)	(0.106)
PL		1.591		3.629	9.746
	-	(0.243)	-	(3.024)	(8.519)
TLLo		15.937		0.023	1.772
	-	(17.019)	-	(0.026)	(0.303)
EL		0.026		27.846	1.794
	-	(0.027)	-	(26.984)	(0.309)
WL	9288.5276	41544.7577	1.32698	19.90441	
	(17578.45506)	(548.674)	(0.11376)	(30.64844)	-
L		112,212.8	219,815.9		
	-	(11,863.8471)	(231.3384)	-	-

Table 8: *Comparison of RTLD and compared distributions for the data set of the life of fatigue fracture of Kevlar 373/epoxy data*

Model	-2$\hat{l}$	AIC	AICC	BIC	K-S	p-value
RTLD	**239.2362**	**245.2362**	**245.5695**	**252.2284**	**0.0669**	**0.863**
SPL	242.6924	248.6924	249.0257	255.6846	0.0829	0.6416
PL	243.0583	249.0583	249.3916	256.0505	0.0844	0.6204
TLLo	244.5824	250.5824	250.9157	257.5746	0.0907	0.53
EL	244.6087	250.6087	250.9421	257.6009	0.0906	0.52
WL	245.0592	253.0592	253.6226	262.3822	0.11003	0.2943
L	254.2288	258.2289	258.3931	262.8902	0.16631	0.02635

outperforms the base model of Lomax and a few well-known competitive models as shown in Table 4 for the provided data set 1, data set 2 and data set 3 (Given in Appendix A). The claim is further supported by Figures 4 and 5. Also the P-P plots of the RTLD model for all the given 3 data sets are shown in Figure 6 ,supports the results presented in Table 6, Table 8 and 10.
***Note:** The findings of the TLLO and EL models for data set 1 and data set 2 are nearly equal,because of their similar nature but slight numerical variations are seen without rounding.

12. Conclusion

In this manuscript, the main contribution is to propose a flexible generalization of Lomax distribution that can acts as a potential substitute for the base model in various situations. In this regard , we use the Ratio Transformation (RT) method and introduced a new model called as RTLD. Some of its key characteristics are discussed, and parameters are determined using a fairly potent estimation technique. The application of

Table 9: *MLE's of RTLD and compared distributions with corresponding standard error (given in parenthesis) for the dataset of the Organic carbon content percentage in the soil of district Ganderbal .*

Model	$\hat{\eta}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\lambda}$
RTLD	18.0874	112511.8	24858.69		
	(5.0829)	(6957.718)	(241.6045)	-	-
SPL		2.2218		48.3872	0.01316
	-	(0.2901)	-	(60.3950)	(0.0164)
LBWLD	82006.2083				34926.7946
	(8393.8383)	-	-	-	(141.5032)
L		58010.4142	49436.1672		
	-	(10617.0983)	(227.7487)	-	-

Table 10: *Comparison of RTLD and compared distributions for the data set of the Organic carbon content percentage in the soil of district Ganderbal*

Model	**-2$\hat{l}$**	AIC	AICC	BIC	K-S	p-value
RTLD	**36.7457**	**42.7457**	**43.3457**	**48.0982**	**0.0787**	**0.948**
SPL	42.9822	48.9822	49.5822	54.3347	0.137	0.3808
LBWLD	56.3150	60.3150	60.60774	63.8834	0.22998	0.01904
L	73.91046	77.91046	78.20314	81.47884	0.30589	0.01

RTLD from a practical perspective is demonstrated through the incorporation of a three real life data sets. The goodness of fit measure is used to assess the effectiveness of the proposed model to other existing known models. The acquired findings are quite encouraging and demonstrate that the RTLD model outperforms the competing models for the provided data sets.

Acknowledgements

We would like to appreciate the referees inputs, valuable comments, and suggestions on the manuscript.

Appendix A

Data set 1: The first data represents the COVID-19 vaccination rate from 46 different countries in southern Africa . The data has been previously analyzed by [3]. The data is as follows: 0.042, 0.205, 0.285, 0.319, 0.464, 0.550, 0.889, 0.895, 0.939, 0.986, 1.000, 1.088, 1.212, 1.244, 1.450, 1.593, 1.844, 2.039, 2.157, 2.167, 2.334, 2.440, 2.657, 3.685, 3.879, 4.493, 4.800, 4.944, 5.155, 5.674, 7.602, 10.004, 12.238, 12.520, 12.553, 13.063, 15.105, 15.229, 15.629, 15.848, 18.641, 18.940, 29.885, 58.162, 61.838, 72.286.

Data set 2 The data set represents the life of fatigue fracture of Kevlar 373/epoxy that are subject to constant pressure at the 90 % stress level until all had failed. For previous studies on the data sets, see, [9] and [6] The data are: 0.0251, 0.0886, 0.0891, 0.2501, 0.3113, 0.3451, 0.4763, 0.5650, 0.5671, 0.6566, 0.6748, 0.6751, 0.6753, 0.7696, 0.8375, 0.8391, 0.8425, 0.8645, 0.8851, 0.9113, 0.9120, 0.9836, 1.0483, 1.0596, 1.0773, 1.1733, 1.2570, 1.2766, 1.2985, 1.3211, 1.3503, 1.3551, 1.4595, 1.4880, 1.5728, 1.5733, 1.7083, 1.7263, 1.7460, 1.7630, 1.7746, 1.8275, 1.8375, 1.8503, 1.8808, 1.8878, 1.8881, 1.9316, 1.9558, 2.0048, 2.0408, 2.0903, 2.1093, 2.1330, 2.2100, 2.2460, 2.2878, 2.3203, 2.3470, 2.3513, 2.4951, 2.5260, 2.9911, 3.0256, 3.2678, 3.4045, 3.4846, 3.7433, 3.7455, 3.9143, 4.8073, 5.4005, 5.4435, 5.5295, 6.5541, 9.0960.

Data set 3 The data set represents the organic carbon(%) content in the soil of district Ganderbal. For previous studies on the data set, see, [13].The data set are: 0.99, 0.81, 0.57, 1.11, 0.97, 0.78, 0.85, 0.85, 0.91, 0.79, 0.66, 0.99, 0.94, 1.17, 1.06, 0.99, 0.84, 1.47, 1.14, 1.41, 0.2, 0.6, 0.03,0.12, 1.11, 0.25, 1.14, 0.63, 0.45, 0.76, 1.2, 1.08, 1.26, 1.08, 0.27, 0.15, 0.75, 0.33, 0.75, 0.63, 1.47, 1.21, 1.24, 1.48.

References

[1] I. Abdul-Moniem and H. Abdel-Hameed. On exponentiated lomax distribution. *International Journal of Mathematical Archive*, 22:2144–2150, 2012.

[2] A. Ahmad, S. Ahmad, and A. Ahmed. Length-biased weighted lomax distribution: statistical properties and application. *Pakistan Journal of Statistics and Operation Research*, pages 245–255, 2016.

[3] H. M. Almongy, E. M. Almetwally, H. Haj Ahmad, and A. H. Al-nefaie. Modeling of covid-19 vaccination rate using odd lomax inverted nadarajah-haghighi distribution. *Plos one*, 17(10):e0276181, 2022.

[4] H. M. Almongy, E. M. Almetwally, and A. E. Mubarak. Marshall–olkin alpha power lomax distribution: Estimation methods, applications on physics and economics. *Pakistan Journal of Statistics and Operation Research*, pages 137–153, 2021.

[5] M. M. Elbiely and H. M. Yousof. A new extension of the lomax distribution and its applications. *Journal of Statistics and Applications*, 2(1):18–34, 2018.

[6] J. Gillariose and L. Tomy. The marshall-olkin extended power lomax distribution with applications. *Pakistan Journal of Statistics and Operation Research*, pages 331–341, 2020.

[7] A. S. Hassan and M. Abd-Allah. Exponentiated weibull-lomax distribution:properties and estimation. *Journal of Data Science*, 16(2):277–298, 2004.

[8] A. S. Hassan and S. Nassr. Power lomax poisson distribution:properties and estimation. *Journal of Data Science*, 18(1):105–128, 2018.

[9] F. Jamal and C. Chesneau. A new family of polyno-expo-trigonometric distributions with applications. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 22(04):1950027, 2019.

[10] K. S. Lomax. Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association*, 49(268):847–852, 1954.

[11] M. A. Lone, I. H. Dar, and T. R. Jan. A new method for generating distributions with an application to weibull distribution. *Reliability Theory and Applications*, 17(1):223–239, 2022.

[12] A. Mahdavi and D. Kundu. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557, 2017.

[13] A.S. Malik. Ph.d thesis. Extension of Rayleigh and Inverse Rayleigh Distributions. *University Of Kashmir*, 2020.

[14] B. V. Nagarjuna, Vasili, R. V. Vardhan, and C. Chesneau. On the accuracy of the sine power lomax model for data fitting. *Modelling*, 2(1):78–104, 2021.

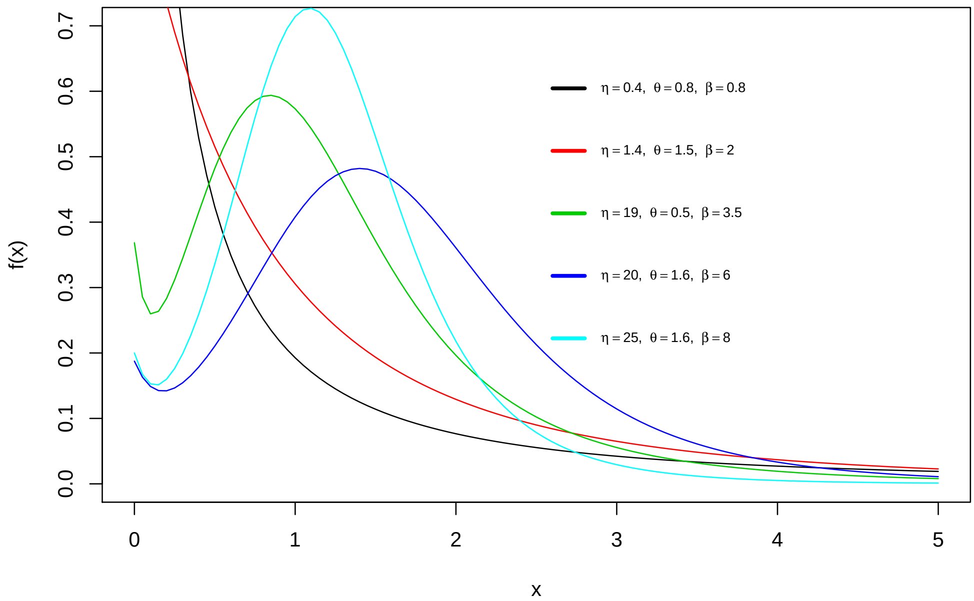
[15] P. E. Oguntunde, M. Khaleel, H. Okagbuei, and O. Okagbue. The topp-leone lomax (tllo) distribution with applications to airbone communication transceiver dataset. *Wirel. Pers. Commun*, 109:349–360, 2019.

[16] E.-H. A. Rady, W. A. Hassanein, and T. A. Elhaddad. The power lomax distribution with an application to bladder cancer data. *SpringerPlus*, 5(1):1–22, 2016.

[17] M. H. Tahir, G. M. Cordeiro, M. Mansoor, and M. Zubair. The weibull-lomax distribution: properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 44(2):455–474, 2015.

13. figures

Here in this section the figures related to the probability density functions, distribution functions, hazard rate and different plots for the three given data sets



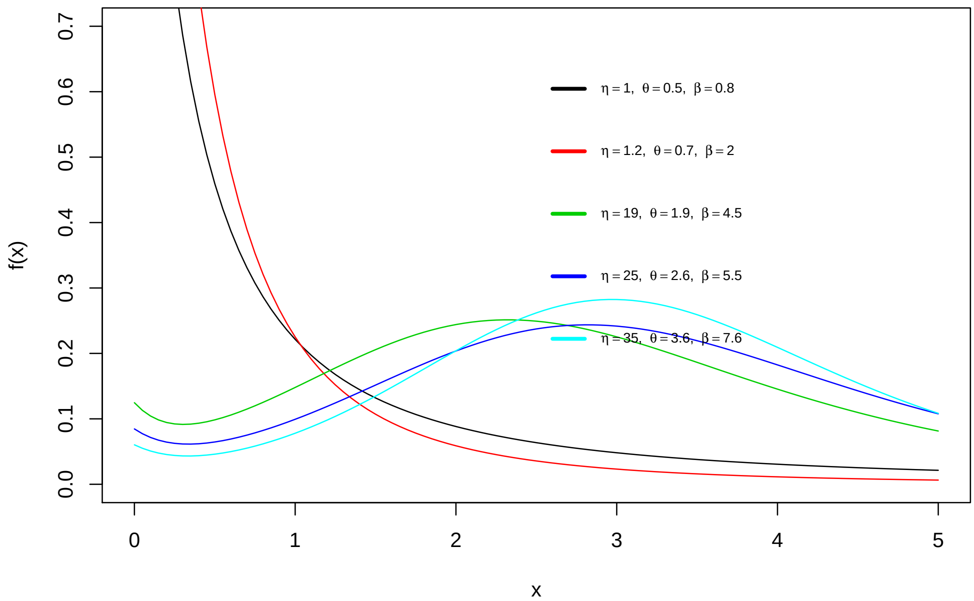
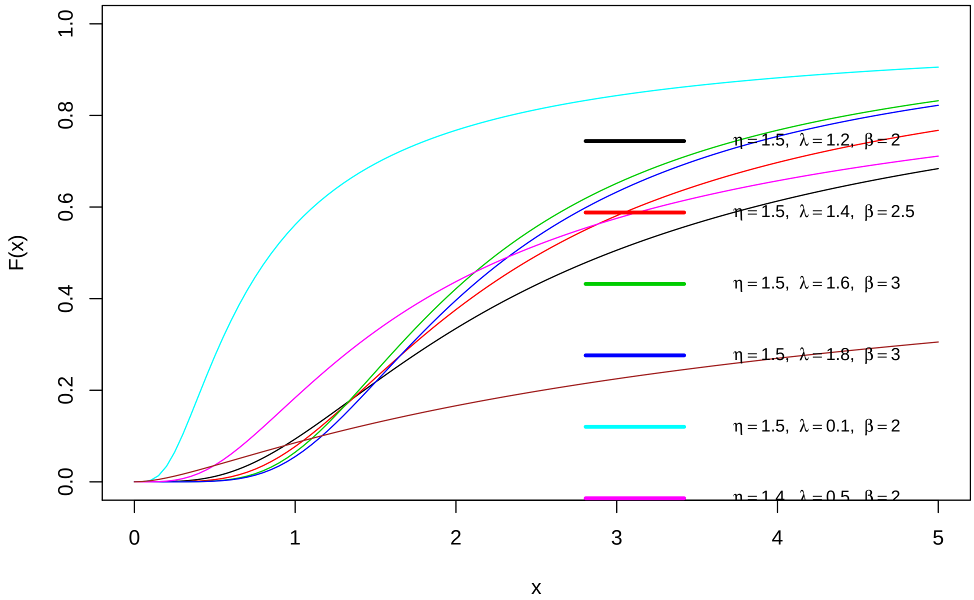


Figure 1: *Probability density plots of the RTLD for various values of* $\eta, \beta, \theta > 0$



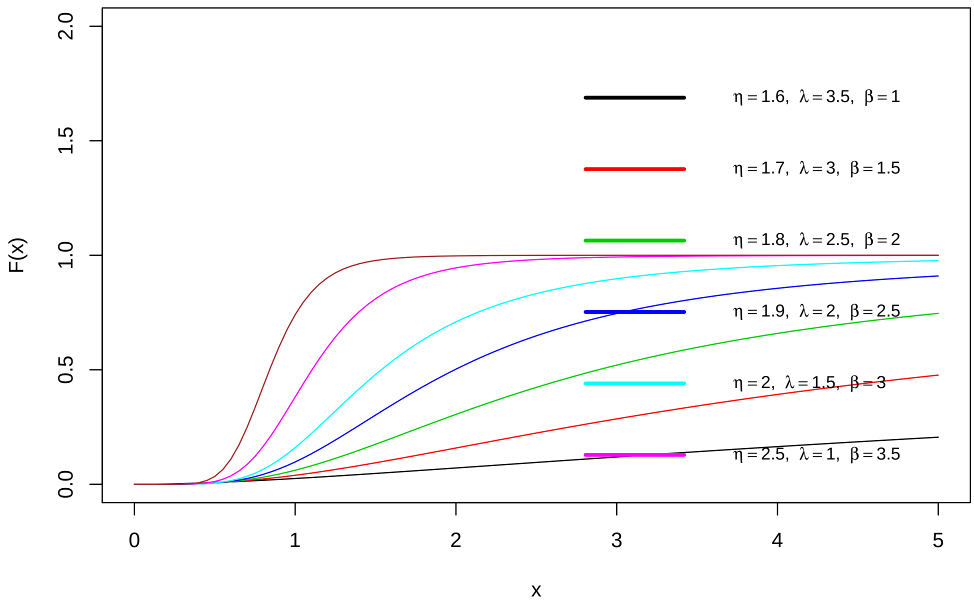


Figure 2: *Distribution function plots of the RTLD for various values of* $\eta, \beta, \theta > 0$

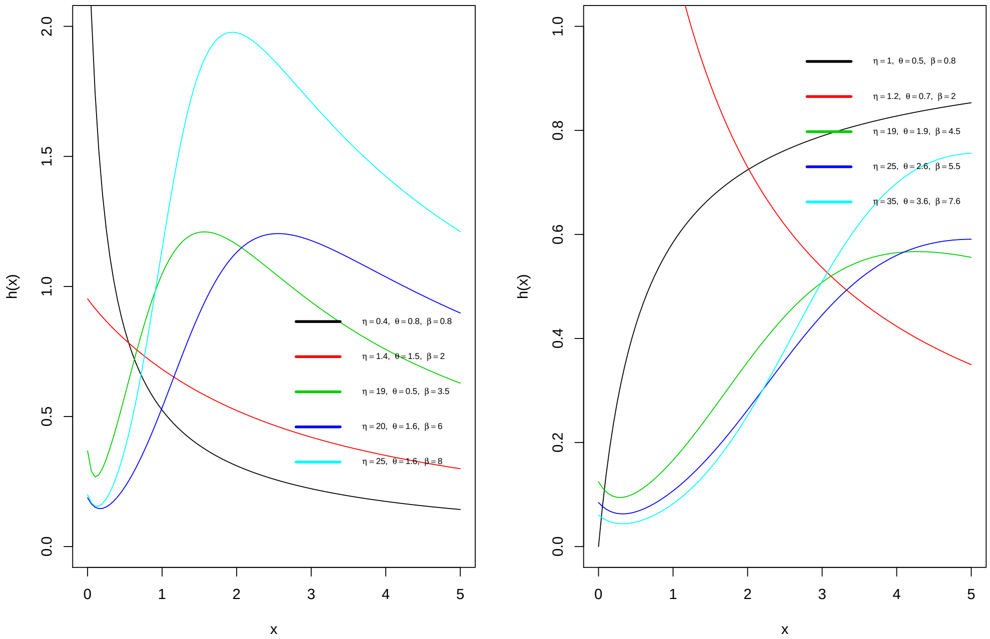


Figure 3: *Hazard rate plots of the RTLD for various values of* $\eta, \beta, \theta > 0$

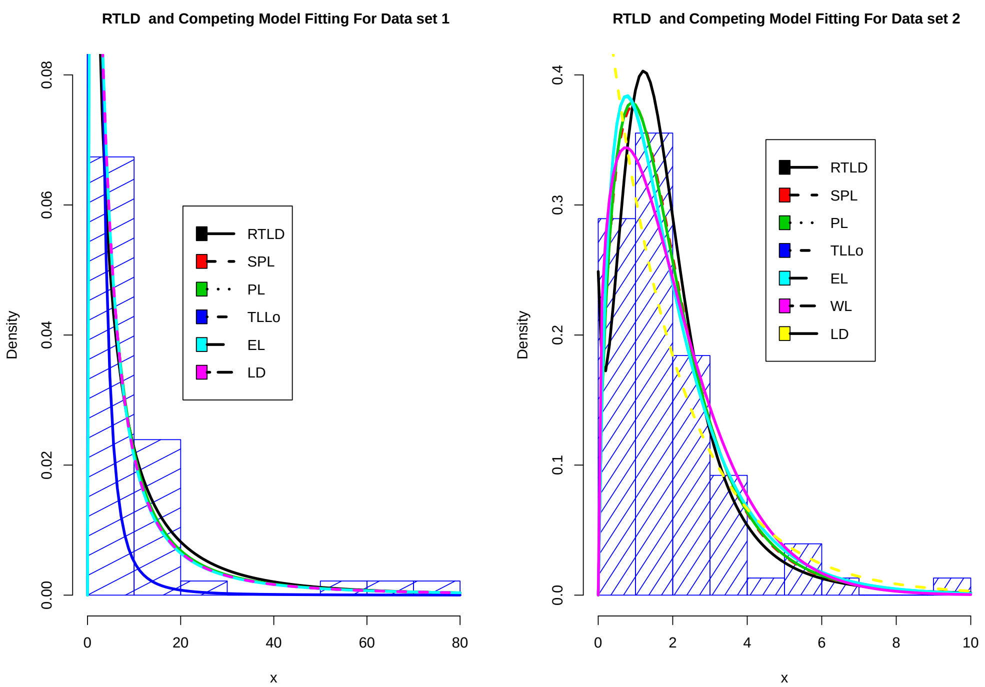


Figure 4: *Plot of the Fitted densities.*

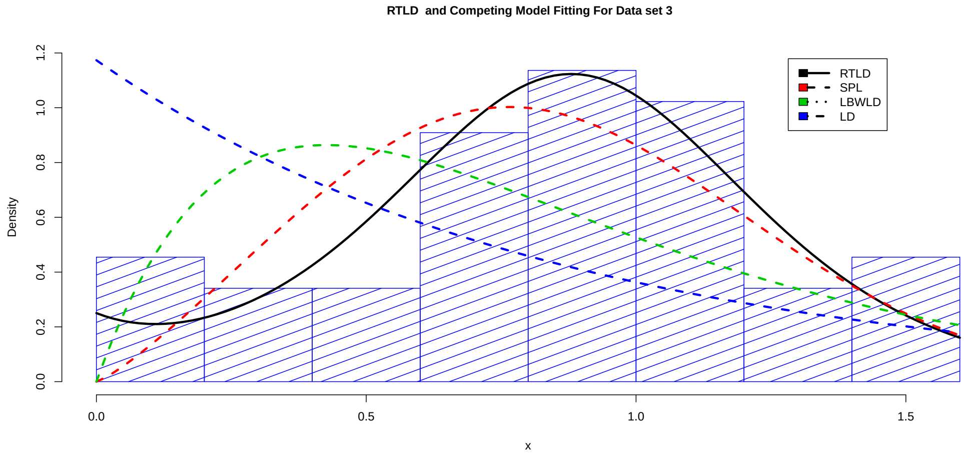


Figure 5: *Plot of the Fitted densities.*

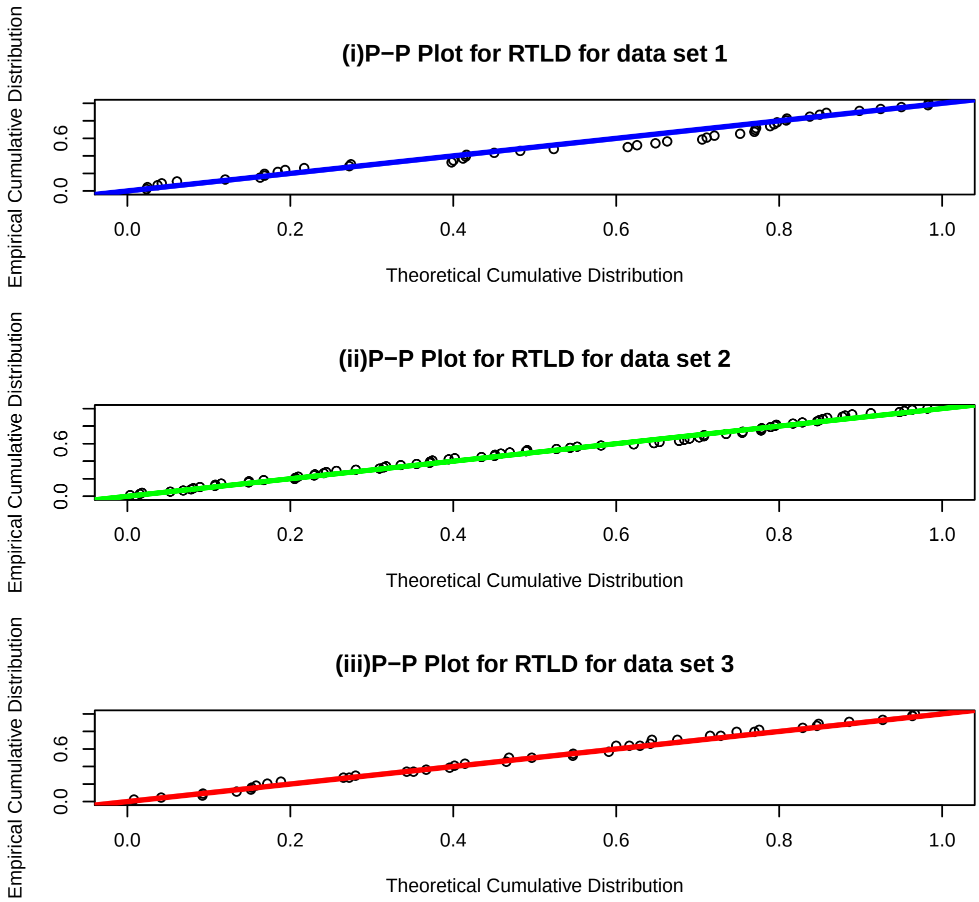


Figure 6: *P-P plot of the RTLD model for data set 1,data set 2 and data set 3 respectively.*

RT&A, No 1 (72)
Volume 18, March 2023

Applications and Some Characteristics of Inverse Power Cauchy Distribution

Laxmi Prasad Sapkota [1] and Vijay Kumar [2]

•

[1,2] Department of Mathematics & Statistics, DDU Gorakhpur University, Gorakhpur, India
[1] laxmisapkota75@gmail.com, [2] vkgkp@rediffmail.com

Abstract

Based on the power Cauchy distribution, we have purposed a new distribution called inverse power Cauchy distribution that offers greater modeling flexibility for lifetime data. Real-world data can be efficiently analyzed using the suggested model because it is analytically sound. Its density function can take on a number of different shapes, including reversed-J, symmetrical, and right-skewed. Depending on the different values of the parameters, it can adapt to different hazard forms, such as an upside-down bathtub, a monotonically increasing or decreasing curve, and others. Its moments, quantile, reliability, hazard, order statistics with density function, moment generating function, and entropy are all given with various explicit forms. The observed information matrix is created when the new model's parameters are calculated through maximum likelihood technique. A simulation study is conducted to investigate the behaviour of maximum likelihood estimators. The proposed model gets a superior fit compared to certain well-known distributions, according to the test of goodness-of-fit we conducted. The significance of the purposed distribution is demonstrated empirically using two real-world data sets.

Keywords: Entropy, Maximum likelihood estimation, Moment, order statistics, Power Cauchy distribution

1. Introduction

Many families of probability models have been developed in the recent decades. New distributions are frequently produced from modifications of a random variable X of parent distribution by: power (e.g., Weibull is obtained from the exponential); linear; non-linear (e.g., log-logistic from logistic); log (e.g., log gamma, log-normal, logistic) and inverse transformation (e.g., inverse Lindley, inverse Exponential models); T-X family framework proposed by Alzaatreh et al. [3]; the compounding of some important lifetime and discrete distributions (e.g. the Poisson-X family distribution) by Tahir et al. [28]. In most cases, a given mixture of baseline models or linear combination defines a class of probability distributions where baseline model is a particular case.

In contrast to the Gaussian distribution, the Cauchy distribution has a significantly thicker tail and is symmetric, unimodal, and bell-shaped. It can be used to analyze data that contains outliers. The ratio of two independent normal variates can be used to generate the Cauchy distribution. It is a widely used distribution that has applications in a variety of disciplines, including biology, applied mathematics, engineering, econometrics, physics, clinical trials, stochastic modeling of decreasing failure rate survival data, queuing theory, and reliability. With location parameter $\theta > 0$ and non-negative scale parameter $\beta > 0$, the cumulative distribution function (CDF) of the Cauchy distribution is,

$$F_X(x;\beta,\theta) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x-\theta}{\beta}\right) \ ; x \in \Re,\ \beta ¿\ 0. \tag{1}$$

and probability density function (PDF) of Equation (1) is

$$f_X(x;\beta,\theta) = \left\{\beta\pi\left[1+\left(\frac{x-\theta}{\beta}\right)^2\right]\right\}^{-1} \ ; x \in \Re,\ \beta ¿\ 0 \tag{2}$$

The centre limit theorem (CLT) is invalid because there is no finite moment generating function for the Cauchy distribution. Due to the lack of a closed-form solution, the MLEs of its parameters are also not the good. These factors make it unlikely or unreasonable to use this distribution to model real-life data. Therefore, the Cauchy distribution must be modified to address the aforementioned shortcomings, some of which are listed below. The generalized form of Cauchy distribution has introduced by Rider [25] whose PDF is given by

$$f_x(x) = \frac{\Gamma(m)}{\beta\,\Gamma(1/2)\,\Gamma(m-1/2)}\left[1+\left(\frac{x-\theta}{\beta}\right)^2\right]^{-m};\ \mathrm{m} \geqslant 1,\ \beta > 0 -\infty < x < \infty. \tag{3}$$

A truncated Cauchy distribution was suggested by Nadarajah and Kotz [21] to address the issue of the lack of MLEs and moments of Cauchy distribution and having PDF,

$$f(x) = \frac{1}{\pi\beta} + \left[1+\left(\frac{x-\theta}{\beta}\right)^2\right]^{-1}\left[\tan^{-1}\left(\frac{B-\theta}{\beta}\right) - \tan^{-1}\left(\frac{A-\theta}{\beta}\right)\right]^{-1};\quad -\infty < A \leqslant x \leqslant B, \theta \in \Re, \beta > 0 \tag{4}$$

Additionally, a relationship between the Cauchy and the hyperbolic secant distribution has been established by Manoukian and Nadeau [17] and Kravchuk [15]. A Modified form of Cauchy distribution has introduced by Ohakwe and Osu [22] and another generalization of the Cauchy distribution is presented by Eugene et al. [11] and Alshawarbeh et al. [[1], [2]]. Further, the Half- Cauchy (HC) distribution has used by using Marshall and Olkin [18] and defined a generators, Beta-G by Eugene et al. [11], and Kumaraswamy-G by Cordeiro and de Castro [9]. Similar to this, some half-Cauchy families have been proposed, including the Marshall-Olkin- HC, beta- HC, and Kumaraswamy- HC families by Jacob and Jayakumar [14], Cordeiro and Lemonte [10], and Ghosh [12] respectively, and the truncated form of Cauchy power-exponential model by Chaudhary et al. [8] and exponentiated form of PC distribution has presented by Sapkota [27].

Extensive study has recently been conducted to develop models that fit survival data, which can be negatively or positively skewed and can have the unimodal hazard function. Power Cauchy (PC) distribution, a two-parameter model that is a sub-model of the modified Beta family that does well with the survival data, was introduced by Rooks et al. [26]. Positively skewed data can be employed with the PC distribution's PDF, which has a somewhat larger right tail than the other recognized humped-shaped two-parameter sub-model of the transformed beta family Rooks et al. [26]. The PC distribution's CDF and PDF are,

$$F(x) = 2\pi^{-1}\tan^{-1}(\lambda x)^{\alpha};\ x > 0,\ \alpha, \lambda > 0. \tag{5}$$

and

$$f(x) = 2\pi^{-1}(\lambda x)(\lambda x)^{\alpha-1}\left[1+(\lambda x)^{2\alpha}\right]^{-1}; x > 0, \alpha, \lambda > 0. \tag{6}$$

respectively.
The hazard function of PC distribution is

$$h(x) = \frac{2\pi^{-1}(\lambda x)(\lambda x)^{\alpha-1}\left[1+(\lambda x)^{2\alpha}\right]^{-1}}{1-2\pi^{-1}\tan^{-1}(\lambda x)^{\alpha}};\ x > 0,\ \alpha, \lambda > 0 \tag{7}$$

Also using PC distribution Weibull PC has been defined by Tahir et al. [29] and Burr XII-power Cauchy distribution by Bhatti et al. [4] using the T-X family technique. The fundamental goal of this work is to present a more adaptable model, demonstrate its applicability, and improve the fitting to the real-life data. In this study, a novel model known as the inverse power Cauchy (IPC) distribution was created using the inversion method. It is the inverse of PC distribution. We have also demonstrated some of the intended model's mathematical and statistical characteristics. The structure of this paper's contents is as follows. In section 2, we presented the IPC distribution and a few distributional features, including the graph of the density, the survival and hazard rate function, the quantile function, random number generation, and skewness and kurtosis. Section 3 presents some of the IPC distribution's crucial characteristics. We go over the procedure for estimating the model parameters in section 4. In section 5, a simulation experiment is done to examine how maximum likelihood estimators behave. We conduct a goodness-of-fit test and a model adequacy test in section 6 using two real data sets. We present some conclusions in section 7.

2. Inverse Power Cauchy Distribution

We have introduced new IPC distribution and visualized some graphs of its PDF and HRF in this section. The CDF of IPC distribution with shape parameter α and scale parameter λ can be expressed by using Equation (5) as

$$F(x) = 1 - 2\pi^{-1}\tan^{-1}\left[\left(\frac{\lambda}{x}\right)^{\alpha}\right];\ x>0,\ \alpha,\lambda>0. \tag{8}$$

And its corresponding PDF is obtained as,

$$f(x) = 2\pi^{-1}\alpha\lambda^{\alpha}x^{-(\alpha+1)}\left[1+\left(\frac{\lambda}{x}\right)^{2\alpha}\right]^{-1};\ x>0,\ \alpha,\lambda>0. \tag{9}$$

2.1. Survival and hazard rate function (HRF)

The survival and HRF of $X \sim IPC(\alpha,\lambda)$ are

$$S(x) = 2\pi^{-1}\tan^{-1}\left[\left(\frac{\lambda}{x}\right)^{\alpha}\right];\ x>0,\ \alpha,\lambda>0 \tag{10}$$

and

$$h(x) = \frac{\alpha\lambda^{\alpha}x^{-(\alpha+1)}\left[1+(\lambda/x)^{2\alpha}\right]^{-1}}{\tan^{-1}\{(\lambda/x)^{\alpha}\}};\ x>0,\ \alpha,\lambda>0. \tag{11}$$

respectively.
A special case of the IPC distribution: If $\alpha = 1$ and $\lambda = 1$ in Equation (9) the IPC distribution tends to two times the standard Cauchy distribution.

2.2. Cumulative hazard function and Failure rate average (FRA)

The cumulative hazard function of IPC distribution can be expressed as

$$\begin{aligned} H(x) &= \int_{-\infty}^{x} h(x)dx \\ &= -\log\left[1-F(x)\right] \\ &= -\log\left[2\pi^{-1}\tan^{-1}(\lambda/x)^{\alpha}\right] \end{aligned} \tag{12}$$

and FRA function is

$$FRA\,(x) = \frac{H(x)}{x} = -\frac{1}{x}\log\left[2\pi^{-1}\tan^{-1}(\lambda/x)^{\alpha}\right];x>0 \tag{13}$$

here H(x) is the cumulative hazard rate function. The IPC distribution exhibits unimodal, decreasing, right skewed and symmetrical shapes of PDF and decreasing, increasing, and inverted bathtub hazard rate shapes in Figure 1.

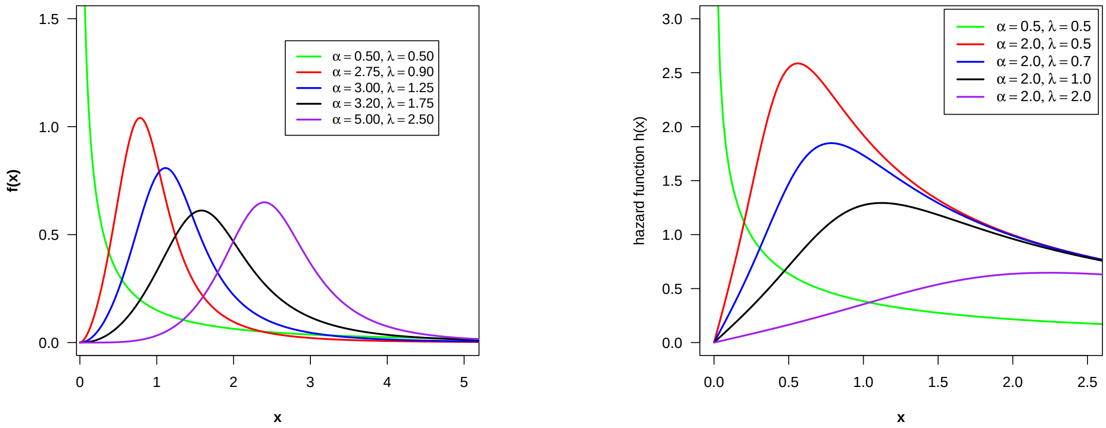


Figure 1: *Various shapes of PDF and HRF function of IPC distribution.*

3. SOME PROPERTIES OF IPC DISTRIBUTION

3.1. The Quantile Function

The quantile function can be expressed as

$$Q(p) = \lambda \left[\tan\left\{\frac{(1-p)\,\pi}{2}\right\}\right]^{-\frac{1}{\alpha}} ; 0 < p < 1. \tag{14}$$

The median, lower quartile, and upper quartile can be calculated by using Equation (14) as follows

$$median = \lambda \left[\tan\left\{\frac{\pi}{4}\right\}\right]^{-\frac{1}{\alpha}}$$

lower quartile

$$Q_1 = \lambda \left[\tan\left\{\frac{3\pi}{8}\right\}\right]^{-\frac{1}{\alpha}}$$

and upper quartile

$$Q_3 = \lambda \left[\tan\left\{\frac{\pi}{8}\right\}\right]^{-\frac{1}{\alpha}}$$

Random deviate generation
The random deviate can be generated from $IPC(\alpha, \lambda)$ by

$$x = \lambda \left[\tan\left\{\frac{(1-u)\,\pi}{2}\right\}\right]^{-\frac{1}{\alpha}} ; 0 < u < 1 \tag{15}$$

Where u has the $U(0,1)$ uniform distribution.

3.2. Mode of IPC distribution

The probability distribution of the provided PDF's mode is its most frequent value. For computing the mode, the following requirements must be met: $\frac{df(x)}{dx} = 0$ and $\frac{d^2 f(x)}{dx^2} < 0$ respectively. Since $f(x) > 0$, by resolving the equation, the model value of the suggested distribution is determined as

$$2\alpha\lambda^{2\alpha} - (1+\alpha)\left(x^{2\alpha} + \lambda^{2\alpha}\right) = 0$$

Hence we obtained the mode of the IPC distribution is

$$x = \left\{\frac{2\alpha\lambda^{2\alpha}}{(1+\alpha)} - \lambda^{2\alpha}\right\}^{1/2\alpha}$$

3.3. Skewness and Kurtosis

The quantile based Skewness and Kurtosis can be computed using the following expressions,

- Coefficient of Bowley's skewness can be computed by using

$$Skewness\,(B) = \frac{Q\,(0.75) - 2Q\,(0.5) + Q\,(0.25)}{Q\,(0.75) - Q\,(0.25)}$$

and

- The coefficient of kurtosis based on octiles which was defined by Moors [20] is

$$M_K = \frac{Q\,(0.875) + Q\,(0.375) - Q\,(0.625) - Q\,(0.125)}{Q\,(3/4) - Q\,(1/4)}$$

In table 1 we have displayed the distributional nature of the IPC distribution. Nine sets of random samples of equal size 100 are generated from Equation (15) for different values of parameters α *and* λ We have computed the values of central tendencies and dispersions like mean, mode, median, skewness, and kurtosis of the proposed distribution. From the table, we observed that the median is increased as the values of the parameters are increases while the mean and mode increases first and then decrease. Similarly, skewness is positive at first and becomes negative when the values of the parameters increase. The proposed distribution is leptokurtic initially, changing progressively to platykurtic when the parameter values are increased, according to the measure of kurtosis.

Table 1: *The mean, median, mode, skewness, and kurtosis for different values of the parameters*

Parameters		**Mean**	**Median**	**Mode**	**Skewness**	**Kurtosis**
alpha	**lambda**					
1	0.1	0.2907	0.095	0.2041	9.2321	89.1813
1	1.0	2.9069	0.9504	2.0408	9.2321	89.1813
2	1.1	1.2948	1.0723	1.5714	5.2338	38.0526
5	1.2	1.1874	1.1878	1.384	1.2998	5.4178
10	1.3	1.2765	1.2934	1.3961	0.317	2.1062
15	1.4	1.3791	1.3953	1.4682	0.0238	1.6876
20	1.5	1.4815	1.4962	1.5545	-0.1187	1.5904
25	1.6	1.5834	1.5967	1.6463	-0.2032	1.5666
30	1.8	1.6848	1.6979	1.7049	-0.2592	1.5649

4. Expansion and Properties of IPC distribution

The PDF and CDF of IPC distribution can be transform into linear form by using the binomial expansion as

$$(1+a)^{-c} = \sum_{k=0}^{\infty} (-1)^k \binom{c+k-1}{k} \mathrm{a}^k, \text{ for } |a| < 1, \text{ c} > 0. \tag{16}$$

Applying Equation (16) the PDF of IPC distribution can be expressed as

$$f\,(x) = \frac{2\alpha}{\pi} \sum_{k=0}^{\infty} \eta_k \; x^{-\alpha(1+2k)-1} \tag{17}$$

where $\eta_k = (-1)^k \lambda^{\alpha(1+2k)}$ also the CDF

$$\begin{aligned}
[F(x)]^h &= \left[1 - 2\pi^{-1}\tan^{-1}\left\{\left(\frac{\lambda}{x}\right)^{\alpha}\right\}\right]^h \\
&= \sum_{j=0}^{\infty}(-1)^j \begin{pmatrix} h \\ j \end{pmatrix}\left[2\pi^{-1}\tan^{-1}\left\{\left(\frac{\lambda}{x}\right)^{\alpha}\right\}\right]^j \\
&= \sum_{j=0}^{\infty}(-2\pi^{-1})^j \begin{pmatrix} h \\ j \end{pmatrix}\left[\tan^{-1}\left\{\left(\frac{\lambda}{x}\right)^{\alpha}\right\}\right]^j
\end{aligned} \tag{18}$$

4.1. Moments of IPC distribution

Let X be a random variable that follows $IPC(\alpha,\lambda)$ then raw moments about the origin can be calculated as

$$\begin{aligned}
\mu_r' = E(X^r) &= \int_0^{\infty} x^r 2\pi^{-1}\alpha\lambda^{\alpha}x^{-(\alpha+1)}\left[1+\left(\frac{\lambda}{x}\right)^{2\alpha}\right]^{-1}dx \\
&= 2\pi^{-1}\alpha\lambda^{\alpha}\int_0^{\infty} x^{r+\alpha-1}\frac{1}{x^{2\alpha}+\lambda^{2\alpha}}dx \\
&= \frac{\lambda^{\alpha}}{\pi}\sum_{k=0}^{\frac{r-\alpha}{2\alpha}}\begin{pmatrix} \frac{r-\alpha}{2\alpha} \\ k \end{pmatrix}\left(-\lambda^{2\alpha}\right)^k\int_{\lambda^{2\alpha}}^{\infty} z^{\frac{r-\alpha}{2\alpha}-k-1}dz
\end{aligned} \tag{19}$$

Where $z = x^{2\alpha} + \lambda^{2\alpha}$

Case I: if $\frac{r-\alpha}{2\alpha} - k = 0$, then $\mu_r' = \frac{\lambda^{\alpha}}{\pi}\sum_{k=0}^{\frac{r-\alpha}{2\alpha}}\begin{pmatrix} \frac{r-\alpha}{2\alpha} \\ k \end{pmatrix}\left(-\lambda^{2\alpha}\right)\left(1-\lambda^{r-\alpha-2k\alpha}\right)$.

Case II: if $\frac{r-\alpha}{2\alpha} - k < 0$, then $\mu_r' = \frac{\lambda^{\alpha}}{\pi}\sum_{k=0}^{\frac{r-\alpha}{2\alpha}}\begin{pmatrix} \frac{r-\alpha}{2\alpha} \\ k \end{pmatrix}\left(-\lambda^{2\alpha}\right)\left(-\lambda^{r-\alpha-2k\alpha}\right)$.

Case III: if $\frac{r-\alpha}{2\alpha} - k > 0$ then μ_r' is undefined.

4.2. Moment Generating Function

A moment-generating function is an important tool for studying random variables. Let X be a random variable that follows IPC distribution and the moment generating function can be defined as $M_X(\delta) = E(e^{\delta x}) = \sum_{r=0}^{\infty}\frac{\delta^r}{r!}E(X^r)$. By using equation (19) we can write

$$M_X(\delta) = \sum_{r=0}^{\infty}\sum_{k=0}^{\frac{r-\alpha}{2\alpha}}\frac{\delta^r}{r!}\frac{\lambda^{\alpha}}{\pi}\begin{pmatrix} \frac{r-\alpha}{2\alpha} \\ k \end{pmatrix}\left(-\lambda^{2\alpha}\right)^k\int_{\lambda^{2\alpha}}^{\infty} z^{\frac{r-\alpha}{2\alpha}-k-1}dz \tag{20}$$

Where $z = x^{2\alpha} + \lambda^{2\alpha}$, hence the moment generating function of IPC distribution can be expressed as

$$M_X(\delta) = \sum_{r=0}^{\infty}\sum_{k=0}^{\frac{r-\alpha}{2\alpha}}\frac{\delta^r}{r!}\frac{\lambda^{\alpha}}{\pi}\begin{pmatrix} \frac{r-\alpha}{2\alpha} \\ k \end{pmatrix}\left(-\lambda^{r-\alpha}\right) \quad for\ \frac{r-\alpha}{2\alpha} - k < 1. \tag{21}$$

4.3. Residual life function

In reliability studies, the additional lifetime given that an event or a component or a system has survived until time t is called the residual life function (RLF) of the event or component, or a system. The r^{th} moment of the residual life of random variable X of IPC distribution can be defined as

$$D_r(t) = \frac{1}{R(t)}\int_t^{\infty}(x-t)^r f(x)dx. \tag{22}$$

Using the binomial expansion $(x-t)^k = \sum_{r=0}^{k} (-1)^r \binom{k}{r} x^{k-r} t^r$ in Equation (22) and using Equation (17) we get

$$D_r(t) = \frac{1}{R(t)} \sum_{j=0}^{\infty} \sum_{r=0}^{k} (-1)^r \binom{k}{r} t^r \ \eta_j \int_t^{\infty} x^{-\alpha(1+2j)-k+r-1} dx \tag{23}$$

Where $\eta_j = \frac{2}{\pi}(-1)^k \alpha \lambda^{\alpha(1+2k)}$. Hence RLF for the time t is

$$D_r(t) = \frac{1}{R(t)} \sum_{j=0}^{\infty} \sum_{r=0}^{k} \theta_{jr} t^{r-(k+\alpha+2\alpha j)} \ for \ r < k+\alpha+2\alpha j \tag{24}$$

here $\theta_{jr} = (-1)^r \binom{k}{r} t^r \eta_j$.

4.4. Entropy

The entropy of a random variable T is a measure of the variance of an uncertainty and has density function $f(t)$. The Renyi entropy is defined as,

$$\begin{aligned} I_R(\rho) &= (1-\rho)^{-1} \log \left[\int f(t)^\rho dt \right]; \quad \text{where } \rho > 0, \rho \neq 1 \\ &= (1-\rho)^{-1} \log \left[\int_0^{\infty} \left\{ 2\pi^{-1} \alpha \lambda^{\alpha} t^{-(\alpha+1)} \left[1 + \left(\frac{\lambda}{t} \right)^{2\alpha} \right]^{-1} \right\}^{\rho} dt \right] \\ &= (1-\rho)^{-1} \log \left[(2\alpha)^{\rho-1} \pi^{-1} \sum_{k=1}^{\frac{\alpha\rho-2\alpha-\rho+1}{2\alpha}} \lambda^{\alpha\rho+2\alpha k} (-1)^k \int_{\lambda^{2\alpha}} z^{\frac{1-\alpha\rho-2\alpha-\rho}{2\alpha}-k-\rho} dz \right] \end{aligned} \tag{25}$$

Where $z = t^{2\alpha} + \lambda^{2\alpha}$

Case I: when $\frac{1-\alpha\rho+2\alpha-\rho}{2\alpha} < k+\rho$ then

$I_R(\rho) = (1-\rho)^{-1} \log \left[(2\alpha)^{\rho-1} \pi^{-1} \lambda^{\alpha\rho} \sum_{k=1}^{\frac{\alpha\rho-2\alpha-\rho+1}{2\alpha}} (-1)^k \left\{ -\lambda^{1-\alpha(\rho+2k+2\rho+2)-\rho} \right\} \right]$

Case II: when $\frac{1-\alpha\rho+2\alpha-\rho}{2\alpha} > k+\rho$, then the integral is divergent.

4.5. Order Statistics (OS)

Numerous applications of probability theory and applied statistics can make use of OS. So, for the suggested distribution, we have shown some OS features. Let $X_1, ..., X_n$ be n iid random variates, each with $F(x)$. Suppose represents the r^{th} OS and denote PDF of r^{th} OS for $X_1, ..., X_n$ be n iid random variables from CDF $F(x)$ and can be defined as

$$\begin{aligned} f_{r:n}(x) &= \frac{n!}{(r-1)!\,(n-r)!} f(x) \left[F(x) \right]^{r-1} \left[1 - F(x) \right]^{n-r} \\ &= \frac{n!}{(r-1)!\,(n-r)!} f(x) \sum_{j=1}^{n-r} \binom{n-r}{j} \left[F(x) \right]^{j+r-1} \\ &= \frac{n!}{(r-1)!\,(n-r)!} 2\pi^{-1} \alpha \lambda^{\alpha} x^{-(\alpha+1)} \left[1 + \left(\frac{\lambda}{x} \right)^{2\alpha} \right]^{-1} \sum_{j=1}^{n-r} \binom{n-r}{j} \times \\ & \left[1 - 2\pi^{-1} \tan^{-1} \left[\left(\frac{\lambda}{x} \right)^{\alpha} \right] \right]^{j+r-1} \end{aligned}$$

Hence the PDF of r^{th} order statistic can be expressed as

$$f_{r:n}(x) = C\sum_{k=0}^{\infty}\sum_{j=1}^{n-r}\omega_{jk}\left[\tan^{-1}\left(\frac{\lambda}{x}\right)\right]^{\alpha k} \tag{26}$$

where $C = \frac{n!}{(r-1)!(n-r)!}2\pi^{-1}\alpha\lambda^{\alpha}x^{-(\alpha+1)}\left[1+\left(\frac{\lambda}{x}\right)^{2\alpha}\right]^{-1}$ and $\omega_{jk} = (-1)^k\left(2\pi^{-1}\right)^k\begin{pmatrix} n-r \\ j\end{pmatrix}\begin{pmatrix} j+r-1 \\ k\end{pmatrix}$.

5. PARAMETER ESTIMATION OF IPC DISTRIBUTION

In this section, we talk about the maximum likelihood estimation (MLE) method and their asymptotic properties to get approximate confidence intervals based on MLEs. Let $\underset{\sim}{x} = (x_1, \ldots, x_n)$ be a random sample drawn from $IPC(\alpha,\ \lambda)$ distribution, and log-likelihood function $l(\alpha, \lambda/\underset{\sim}{x})$ can be written as,

$$l(\alpha, \lambda/\underset{\sim}{x}) = n\ln(2/\pi) + n\ln(\alpha) + n\alpha\ln(\lambda) - (\alpha+1)\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\ln\left\{1+\left(\frac{\lambda}{x_i}\right)^{2\alpha}\right\} \tag{27}$$

Differentiating Equation (27) with respect to α and λ we get,

$$\frac{\partial l}{\partial\alpha} = \frac{n}{\alpha} + n\ln(\lambda) + \sum_{i=1}^{n}\ln(x_i) - 2\sum_{i=1}^{n}\frac{\ln(\lambda/x_i)(\lambda/x_i)^{2\alpha}}{1+(\lambda/x_i)^{2\alpha}} \tag{28}$$

$$\frac{\partial\ell}{\partial\lambda} = \frac{\alpha}{\lambda} - \frac{2\alpha(\lambda/x)^{2x}}{\lambda\left\{1+(\lambda/x)^{2x}\right\}} \tag{29}$$

By setting Equations (28) and (29) to zero and solving them simultaneously we get the maximum likelihood estimate $\hat{\alpha}$ *and* $\hat{\lambda}$ of the model parameters. For the parameters α *and* λ the $100(1-\tau)\%$ confidence intervals can be calculated as the customary asymptotic normality of the maximum likelihood estimators $\text{var}(\hat{\alpha})$ *and* $\text{var}(\hat{\lambda})$ estimated from the inverse of the matrix of second derivatives of the log-likelihood function Casella & Berger [5] locally at $\hat{\alpha}$ *and* $\hat{\lambda}$.

$$\frac{\partial^2\ell}{\partial\alpha^2} = -\frac{n}{\alpha^2} - 4\sum_{i=1}^{n}\frac{\left\{\ln(\lambda/x_i)\right\}^2(\lambda/x_i)^{2\alpha}}{\left[1+(\lambda/x_i)^{2\alpha}\right]^2}$$

$$\frac{\partial^2\ell}{\partial\lambda^2} = -\frac{n\alpha}{\lambda^2} + \frac{2\alpha}{\lambda^2}\sum_{i=1}^{n}\frac{(\lambda/x_i)^{2\alpha}\left\{1-2\alpha+(\lambda/x_i)^{2\alpha}\right\}}{\left[1+(\lambda/x_i)^{2\alpha}\right]^2}$$

$$\frac{\partial^2\ell}{\partial\alpha\partial\lambda} = \frac{n}{\lambda} + \frac{2}{\lambda}\sum_{i=1}^{n}\frac{(\lambda/x_i)^{2\alpha}\left\{1+2\alpha\ln(\lambda/x_i)+(\lambda/x_i)^{2\alpha}\right\}}{\left[1+(\lambda/x_i)^{2\alpha}\right]^2}$$

$$\frac{\partial^2\ell}{\partial\lambda\partial\alpha} = \frac{n}{\lambda} + \frac{2}{\lambda}\sum_{i=1}^{n}\frac{(\lambda/x_i)^{2\alpha}\left\{1+2\alpha\ln(\lambda/x_i)+(\lambda/x_i)^{2\alpha}\right\}}{\left[1+(\lambda/x_i)^{2\alpha}\right]^2}$$

Let $\underset{\sim}{\delta} = (\alpha, \lambda)$ represent the parameter vector and the corresponding MLE of $\underset{\sim}{\delta}$ as $\hat{\underset{\sim}{\delta}} = (\hat{\alpha}, \hat{\lambda})$, then, $\hat{\underset{\sim}{\delta}} - \underset{\sim}{\delta} \to N_2\left[0, \left(I\left(\underset{\sim}{\delta}\right)\right)^{-1}\right]$ where $I\left(\underset{\sim}{\delta}\right)$ is the matrix of Fisher's information (FIM) given by,

$$I\left(\underset{\sim}{\delta}\right) = -\begin{pmatrix} E\left(\frac{\partial^2 l}{\partial\alpha^2}\right) & E\left(\frac{\partial^2 l}{\partial\alpha\partial\lambda}\right) \\ E\left(\frac{\partial^2 l}{\partial\lambda\partial\alpha}\right) & E\left(\frac{\partial^2 l}{\partial\lambda^2}\right)\end{pmatrix}$$

Because we don't know $\underline{\delta}$, it is pointless for practical purposes that the MLEs have asymptotic variance $\left(I\left(\underline{\delta}\right)\right)^{-1}$. So, using the estimated values of the parameters, we approximate the asymptotic variance. The general procedure is to use the observed FIM $O\left(\hat{\underline{\delta}}\right)$ as an estimate of $I\left(\underline{\delta}\right)$ given by

$$O\left(\hat{\underline{\delta}}\right) = -\begin{pmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 l}{\partial \lambda \partial \alpha} & \frac{\partial^2 l}{\partial \lambda^2} \end{pmatrix}_{|(\hat{\alpha},\hat{\lambda})} = -H\left(\underline{\delta}\right)_{|\left(\underline{\delta}=\hat{\underline{\delta}}\right)}$$

here H is the Hessian matrix. The observed information matrix is given maximum likelihood using the Newton-Raphson algorithm. Consequently, the variance-covariance matrix is calculated as follows:

$$\left[-H\left(\underline{\delta}\right)_{|\left(\underline{\delta}=\hat{\underline{\delta}}\right)}\right]^{-1} = \begin{pmatrix} \mathrm{var}(\hat{\alpha}) & \mathrm{cov}(\hat{\alpha},\hat{\lambda}) \\ \mathrm{cov}(\hat{\lambda},\hat{\alpha}) & \mathrm{var}(\hat{\lambda}) \end{pmatrix}$$

Confidence intervals for α and λ can be constructed using the asymptotic normality of MLEs, $\hat{\alpha} \pm Z_{\tau/2}\sqrt{\mathrm{var}(\hat{\alpha})}$ and $\hat{\lambda} \pm Z_{\tau/2}\sqrt{\mathrm{var}(\hat{\lambda})}$, where $Z_{\tau/2}$ is the upper percentile of standard normal variate.

6. SIMULATION STUDY

A simulation study is carried out to investigate the capability of the ML estimators for estimating the parameters of the model $IPC\left(\alpha,\ \lambda\right)$. Using the random deviate function defined in Equation (15) we have generated $N = 10000$ independent samples of different sizes $n = (50, 100, 200, 250, 500)$ for three different sets of the parameter values. The estimated value of the parameters $\hat{\alpha}$ *and* $\hat{\lambda}$, absolute average bias (Bias) and mean square errors (MSEs) of the ML estimators are reported in Table 2. From Table 2 we have observed that the ML estimators tend to the actual values, Biases are comes close to zero, and MSEs are decreased as we expected under asymptotic theory when the size of the samples is increased.

Table 2: *The estimated values, Biases, and MSEs are based on 10000 simulations of IPC distribution.*

	Actual values		MLEs		Bias		MSEs	
n	α	λ	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$
50	0.5	0.25	0.5142	0.2731	0.0142	0.0231	0.0048	0.014
	0.75	0.5	0.7705	0.5185	0.0205	0.0185	0.0106	0.0203
	1	0.75	1.0274	0.7659	0.0274	0.0159	0.0189	0.0248
100	0.5	0.25	0.5066	0.2604	0.0066	0.0104	0.0022	0.0057
	0.75	0.5	0.7597	0.5086	0.0097	0.0086	0.0049	0.0095
	1	0.75	1.0122	0.7567	0.0122	0.0067	0.0088	0.0119
200	0.5	0.25	0.5033	0.2557	0.0033	0.0057	0.0011	0.0028
	0.75	0.5	0.7548	0.505	0.0048	0.005	0.0024	0.0045
	1	0.75	1.0065	0.7544	0.0065	0.0044	0.0042	0.0058
250	0.5	0.25	0.5031	0.2537	0.0031	0.0037	0.00083	0.0021
	0.75	0.5	0.7535	0.5037	0.0035	0.0037	0.0019	0.0037
	1	0.75	1.0053	0.7539	0.0053	0.0039	0.0033	0.0045
500	0.5	0.25	0.5016	0.2519	0.0016	0.0019	0.0004	0.001
	0.75	0.5	0.752	0.5018	0.002	0.0018	0.0009	0.0018
	1	0.75	1.0029	0.7515	0.0029	0.0015	0.0016	0.0023

7. APPLICATIONS TO REAL DATASET

Using two real datasets, we have examined the applicability of IPC distribution in this section.

Data set 1

The data set below is from an accelerated life test of 59 conductors Lawless [16] where failure times are measured in hours and there are no censored observations.

"6.545, 9.289, 7.543, 6.956, 6.492, 5.459, 8.120, 4.706, 8.687, 2.997, 8.591, 6.129, 11.038, 5.381, 6.958, 4.288, 6.522, 4.137, 7.459, 7.495, 6.573, 6.538, 5.589, 6.087, 5.807, 6.725, 8.532, 9.663, 6.369, 7.024, 8.336, 9.218, 7.945, 6.869, 6.352, 4.700, 6.948, 9.254, 5.009, 7.489, 7.398, 6.033, 10.092, 7.496, 4.531, 7.974, 8.799, 7.683, 7.224, 7.365, 6.923, 5.640, 5.434, 7.937, 6.515, 6.476, 6.071, 10.491, 5.923".

Data set 2

The second data set of 72 observations represents the coating weight (gm/m2) of Iron Sheets by a chemical method on the top center side (TCS) Rao and Mbwambo [24].

"36.8, 47.2, 35.6, 36.7, 55.8, 58.7, 42.3, 37.8, 55.4, 45.2, 31.8, 48.3, 45.3, 48.5, 52.8, 45.4, 49.8, 48.2, 54.5 ,50.1 ,48.4, 44.2, 41.2, 47.2, 39.1, 40.7, 40.3, 41.2,30.4, 42.8, 38.9, 34.0, 33.2, 56.8, 52.6, 40.5, 40.6, 45.8, 58.9, 28.7, 37.3, 36.8, 40.2, 58.2, 59.2, 42.8, 46.3, 61.2, 58.4, 38.5, 34.2, 41.3, 42.6, 43.1 ,42.3, 54.2,44.9, 42.8, 47.1, 38.9, 42.8, 29.4, 32.7, 40.1, 33.2, 31.6, 36.2, 33.6, 32.9, 34.5, 33.7, 39.9"

7.1. Estimation of the Parameters of IPC distribution

By using the **maxLik()** function Henningsen and Toomet [13] in R programming software R Core Team [23], we have computed the MLEs (see McElreath [19]) by maximizing the likelihood function for IPC distribution along with some models with their standard errors (SE) and presented in Tables 3 and 4.

Table 3: *MLEs and SE (parenthesis) for the distributions under study (data set 1)*

Distribution	Estimated parameters				
	$\hat{k}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$
IPC	-	6.3604(0.7692)	-	6.8773(0.1997)	-
WPC	-	2.7753(2.2584)	2.1901(1.7000)	6.3793(0.7396)	-
BurrXII-PC	1.480(5.005)	4.547(4.182)	4.016(15.330)	-	7.361(8.356)
LHC	-	-	-	4.495(1.117)	18.374(4.324)
NLHC	-	-	-	4.448(2.530)	1.787 (1.062)
HC	-	6.8049(0.9046)	-	-	-
Cauchy	-	-	-	6.8049(0.9046)	-

Table 4: *MLEs and SE (parenthesis) for the distributions under study (data set 2)*

Distribution	Estimated parameters				
	$\hat{k}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\theta}$
IPC	-	7.5515(0.8957)	-	42.1573(1.1228)	-
WPC	-	3.368(1.901)	2.129(1.152)	40.00(2.521)	-
BurrIX-PC	0.865(0.3027)	1.4109(0.8936)	11.658(3.871)	-	27.3261(4.1769)
LHC	-	-	-	4.2178(0.4471)	106.0269(4.4036)
NLHC	-	-	-	114.9316(4.42782)	0.5731(0.071)
HC	-	42.323(4.194)	-	-	-
Cauchy	-	-	-	42.323(4.194)	-

7.2. Adequacy test of the IPC model

To evaluate the goodness-of-fit and adequacy of the proposed model we have considered some well-known models like three-parameter Weibull power Cauchy (WPC) by Tahir et al. [29], four-parameter BurrXII-power Cauchy (BurrXII-PC) by Bhatti et al. [4], two parameters Lindley half Cauchy (LHC) by Chaudhary and Kumar [6], new Lindley half Cauchy (NLHC) by Chaudhary and Kumar [7], single parameter half Cauchy (HC) and Cauchy distributions. For the adequacy test of the model we have compared the IPC model with underlying models using the criterions Akaike information criterion (AIC), Corrected Akaike Information criterion (CAIC), Bayesian information criterion (BIC), and Hannan-Quinn information criterion (HQIC) respectively, and calculate as

$$AIC = -2l(\hat{\alpha}, \hat{\lambda}) + 2d$$

$$BIC = -2l(\hat{\alpha}, \hat{\lambda}) + d\log(n)$$

$$CAIC = AIC + \frac{2d\,(d+1)}{n-d-1}$$

$$HQIC = -2l(\hat{\alpha}, \hat{\lambda}) + 2d\log[\log(n)]$$

where d is the number of parameters associated with the concern model and n is the sample size. The results are reported in Tables 5 and 6 and noticed that a two-parameter simple IPC model can perform as complex models with three and four parameters WPC and BurrXII-PC respectively and our model is better than LHC, NLHC, HC, and Cauchy models. To evaluate the fit attained by the IPC model we have also presented the quantile-quantile (Q-Q) plot and Kolmogorov-Smirnov (KS) plot (Figures 2 and 3) and also verified that the IPC model fits the real data under study very nicely.

Table 5: *Model selection statistics (data set 1)*

Model	AIC	BIC	CAIC	HQIC	LL
IPC	228.1826	232.3376	228.3968	229.8045	-112.091
WPC	228.3944	234.627	228.8308	230.8274	-111.197
BurrXII-PC	230.6105	238.9206	231.3512	233.8544	-111.305
LHC	344.665	348.8201	344.8793	346.287	-170.333
NLHC	345.139	349.2941	345.3533	346.761	-170.57
HC	366.3982	368.4757	366.4684	367.2092	-182.199
Cauchy	448.1896	450.2671	448.2597	449.0005	-223.095

Table 6: *Model selection statistics (data set 2)*

Model	AIC	BIC	CAIC	HQIC	LL
IPC	513.6304	518.1838	513.8043	515.4431	-254.815
WPC	513.1188	519.9488	513.4718	515.8379	-253.559
BurrIX-PC	515.0942	524.2009	515.6913	518.7196	-253.547
LHC	681.5617	686.115	681.7356	683.3744	-338.781
NLHC	689.2971	693.8504	689.471	691.1098	-342.649
HC	627.2789	629.5556	627.336	628.1852	-312.639
Cauchy	808.5337	810.8104	808.5909	809.4401	-403.267

7.3. Test of goodness of fit of the proposed model

Further, we have compared the proposed distribution with competitive distributions by computing the Kolmogorov-Smirnov (KS), the Cramer-Von Mises (A^2) and the Anderson-Darling (W)

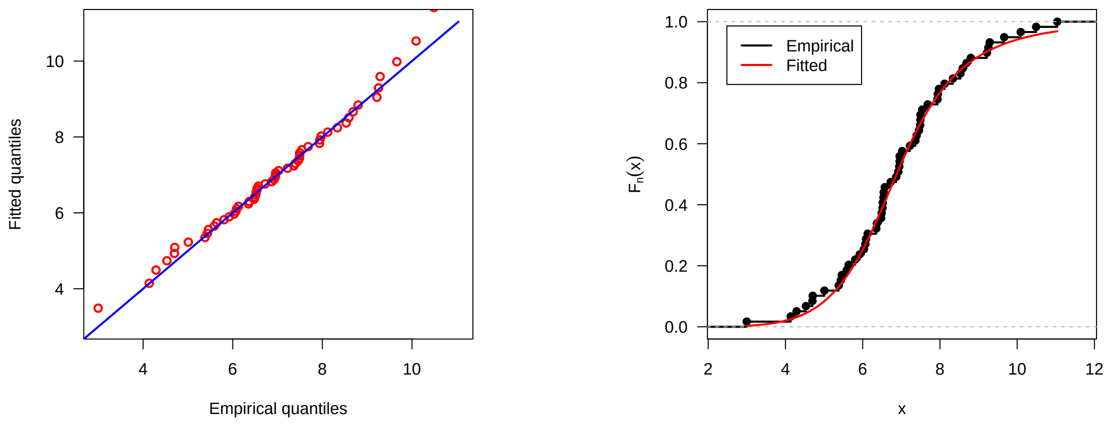


Figure 2: *Q-Q and KS plot (data set 1)*

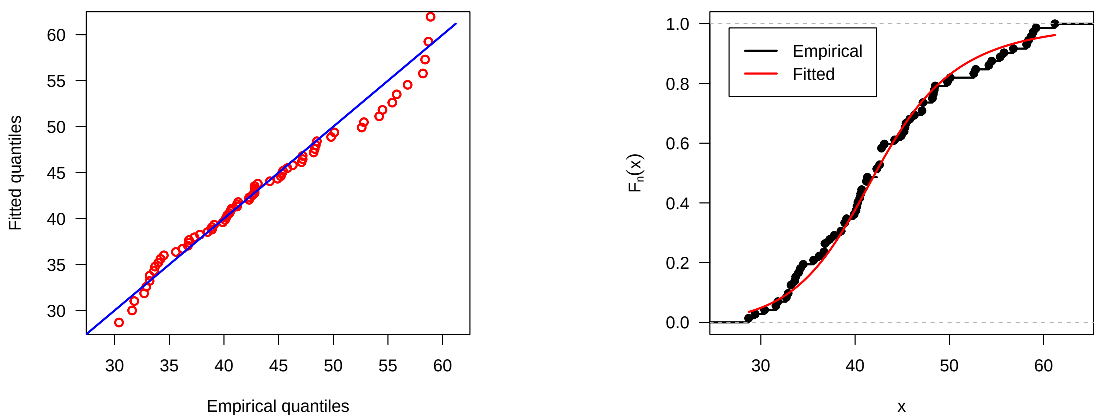


Figure 3: *Q-Q and KS plot (data set 2)*

statistics. These measures are frequently used to contrast non-nested models and show how well a certain CDF matches the empirical distribution of a particular data and computed as

$$KS = \max_{1 \leqslant j \leqslant n} \left(c_j - \frac{j-1}{n}, \frac{j}{n} - c_j \right)$$

$$W = -n - \frac{1}{n} \sum_{j=1}^{n} (2j-1) \left[\log c_j + \log \left(1 - c_{n+1-j} \right) \right]$$

$$A^2 = \frac{1}{12n} + \sum_{j=1}^{n} \left[\frac{(2j-1)}{2n} - c_j \right]^2$$

where $c_j = CDF(x_j)$; the x_j's are the ordered observations. The results are reported in Tables 7 and 8 and found that IPC model gets small test statistics values (except the value of AD for WPC and Burr XII-PC) and the highest p-value, hence the proposed model fits the data under study very well. This result is also verified by figures 4 and 5.

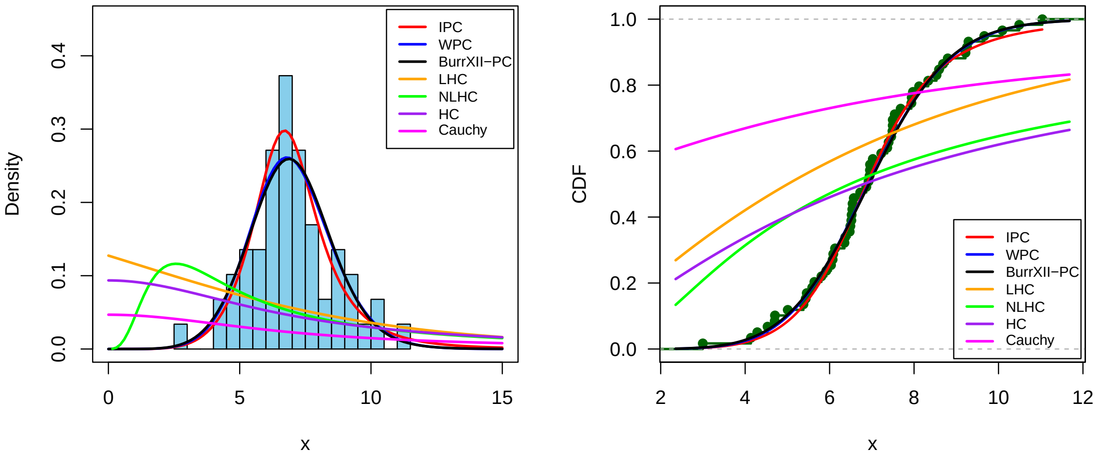


Figure 4: *PDF and CDF fit attained by competing models (data set 1)*

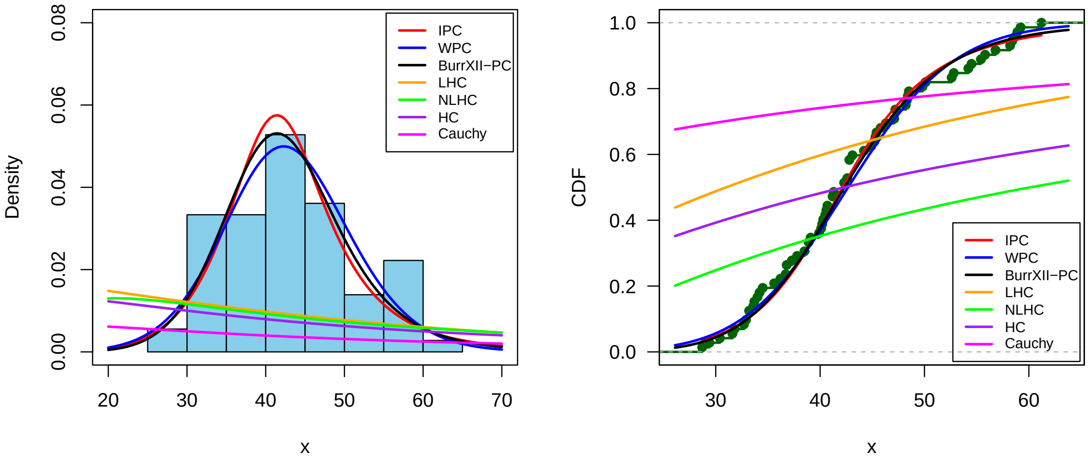


Figure 5: *PDF and CDF fit attained by competing models (data set 2)*

8. CONCLUSION

In this paper, we present the inverse power Cauchy (IPC) distribution, a novel class of continuous distributions. The plots of the PDF, and HRF, explicit expressions for CDF, quantile function, survival function, reverse HRF, skewness, kurtosis, moments, order statistics, and entropy have all been covered. The parameters of the suggested model were estimated using the maximum likelihood estimation (MLE) approach. For the purpose of illustrating the IPC distribution, two real data sets are taken into account. We come to the conclusion that the IPC distribution outperforms by comparing it to several other lifetime models the WPC, BurrXI-PC, LHC, NLHC, HC, and Cauchy distributions that are taken into account. We anticipate that the fields of applied statistics, probability theory, and survival analysis may choose to use this distribution.

Conflict of interest: The authors declare that there is no conflict of interest.

Table 7: *The goodness-of-fit test statistic with p-value (data set 1)*

Model	*KS(p-value)*	*W(p-value)*	A^2*(p-value)*
IPC	0.0480(0.9982)	0.1780(0.9953)	0.0199(0.9973)
WPC	0.05602(0.9876)	0.1541(0.9983)	0.0257(0.9887)
BurrXII-PC	0.0589(0.979)	0.1659(0.9971)	0.0282(0.9825)
LHC	0.4152(1.004e-09)	15.561(1.017e-05)	3.2415(7.108e-09)
NLHC	0.4348(1.156e-10)	15.321(1.017e-05)	3.214(8.475e-09)
HC	0.3517(4.956e-07)	14.226(1.017e-05)	2.8148(9.604e-08)
Cauchy	0.6569(7.772e-16)	33.963(1.017e-05)	7.4666(¡ 2.2e-16)

Table 8: *The goodness-of-fit test statistic with p-value (data set 2)*

Model	*KS(p-value)*	*W(p-value)*	A^2*(p-value)*
IPC	0.0623(0.9429)	0.4434(0.8044)	0.0453(0.9055)
WPC	0.0776(0.7792)	0.5364(0.7094)	0.0595(0.8189)
BurrIX-PC	0.0599(0.9581)	0.3916(0.8566)	0.0406(0.9315)
LHC	0.4719(2.387e-14)	20.943(8.333e-06)	4.4141(2.2e-16)
NLHC	0.4940(1.11e-15)	22.016(8.333e-06)	4.6737(¡ 2.2e-16)
HC	0.3852(1.054e-09)	18.986(8.333e-06)	3.8139(2.697e-10)
Cauchy	0.6897(¡ 2.2e-16)	42.534(8.333e-06)	9.3442(¡ 2.2e-16)

References

[1] Alshawarbeh, E., Famoye, F. and Lee, C. (2013). Beta-Cauchy distribution: Some properties and applications. *Journal of Statistical Theory and Applications*, 12, 378-391.

[2] Alshawarbeh, E., Lee, C. and Famoye, F. (2012). The beta-Cauchy distribution. *Journal of Probability and Statistical Science*, 10, 41-57.

[3] Alzaatreh, A., Famoye, F. and Lee, C. (2014). T-normal family of distributions: A new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications* 1, Article 16.

[4] Bhatti, F. A., Hamedani, G., Al Sobhi, M. M., and Korkmaz, M. ?. (2021). On the Burr XII-Power Cauchy distribution: Properties and applications. *AIMS Mathematics.* 6(7), 7070–7092.

[5] Casella, G., and Berger, R. L. (2002). *Statistical inference second edition.* Duxbur y advanced series.

[6] Chaudhary, A. K. and Kumar, V. (2020a). Lindley Half Cauchy Distribution: Properties and Applications. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 8(9), 1233-1242.

[7] Chaudhary, A.K. and Kumar, V. (2020b). New Lindley Half Cauchy Distribution: Theory and Applications. *International Journal of Recent Technology and Engineering (IJRTE)*, 9(4), 1-7.

[8] Chaudhary, A. K., Sapkota, L. P. and Kumar, V. (2020). Truncated Cauchy power– exponential distribution: Theory and Applications. *IOSR Journal of Mathematics (IOSR-JM)*, 16(6), Ser. I, 44-52.

[9] Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81, 883-893.

[10] Cordeiro, G. M. and Lemonte, A. J. (2011). The beta-half Cauchy distribution. *Journal of Probability and Statistics*, Article ID 904705, 18 pages.

[11] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, 31, 497-512.

[12] Ghosh, I. (2014).The Kumaraswamy-half Cauchy distribution: Properties and applications. *Journal of Statistical Theory and Applications*, 13, 122-134.

[13] Henningsen, A., and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26(3), 443-458.

[14] Jacob, E. and Jayakumar, K. (2012). On half-Cauchy distribution and process. International *Journal of Statistika and Mathematika*, 3, 77-81.

[15] Kravchuk, O.Y. (2005). Rank test for location optimal for hyperbolic secant distribution. *Communications in Statistics-Theory and Methods*, 34, 1617-1630.

[16] Lawless, J. F. (2003). *Statistical models and methods for lifetime data* (Vol. 362). John Wiley & Sons.

[17] Manoukian, E. B. and Nadeau, P. (1988). A note on hyperbolic secant distribution. *The American Statistician*, 42, 77-79.

[18] Marshall, A. N. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika*, 84, 641-652.

[19] McElreath, R. (2020). *Statistical rethinking*: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC.

[20] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 37(1), 25-32.

[21] Nadarajah, S. and Kotz, S. A. (2006). Truncated Cauchy distribution, *International Journal of Mathematical Education in Science and Technology*, 37, 605-608.

[22] Ohakwe, J. and Osu, B. (2011). The existence of the moments of the Cauchy distribution under a simple transformation of dividing with a constant. *Journal of Theoretical Mathematics and Application*, 1, 27-35.

[23] R Development Core Team. (2022). R- A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Austria, Vienna).

[24] Rao, G. S., and Mbwambo, S. (2019). Exponentiated inverse Rayleigh distribution and an application to coating weights of iron sheets data. *Journal of probability and statistics*, 2019.

[25] Rider, P.R. (1957). Generalized Cauchy distribution. *Annals of Mathematical Statistics*, 9, 215-223.

[26] Rooks, B., Schumacher, A. and Cooray, K. (2010). The power Cauchy distribution: derivation, description, and composite models. NSF-REU Program Reports.

[27] Sapkota, L. P. (2022). Statistical Properties and Applications of Exponentiated Inverse Power Cauchy Distribution. *Journal of Institute of Science and Technology*, 27(1), 113-123.

[28] Tahir, M. H., Zubair, M., Cordeiro, G. M., Alzaatreh, A., and Mansoor, M. (2016). The Poisson-X family of distributions. *Journal of Statistical Computation and Simulation*, 86(14), 2901-2921.

[29] Tahir, M. H., Zubair, M., Cordeiro, G. M., Alzaatreh, A., and Mansoor, M. (2017). The Weibull-Power Cauchy distribution: model, properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 46(4), 767-789.

PERFORMANCE AND BEHAVIOR ANALYSIS OF WATER CIRCULATION SYSTEM OF A THERMAL POWER PLANT

Seema Sharma[1], Sushma[2,*]

•

[1]Professor, Department of Mathematics and Statistics, Gurukul Kangri (Deemed to be University) Haridwar, Uttarakhand, India-249404
[2] Research Scholar, Department of Mathematics and Statistics, Gurukul Kangri (Deemed to be University) Haridwar, Uttarakhand, India-249404

[1]e-mail: seema@gkv.ac.in
[2]e-mail: sushmanirwal1885@gmail.com
*Corresponding Author

Abstract

This paper analyses the performance and behavior of water circulation system (WCS) of a thermal power plant in fuzzy environment. For this purpose, fuzzy λ–τ technique coupled with petrinet modelling has been used. To address the vagueness in data, trapezoidal fuzzy numbers have been employed in fuzzy λ–τ technique. Various reliability indicators of WCS viz. failure rate, repair time, expected number of failures, mean time between failures, reliability and availability have been computed at ±15%, ±25% and ±40% spreads using fuzzy λ–τ technique. Further, fuzzy values of reliability indicators have been defuzzified employing COA method and the failure dynamics of WCS have been studied on account of decreasing / increasing trends of reliability indicators. The outcomes of this study are of great importance for plant personnel / management to enhance the availability of WCS.

Keywords: thermal power plant, water circulation system, fuzzy λ–τ technique, trapezoidal fuzzy number

1. Introduction

Thermal power plant is a prominent source of electricity generation. It is a complex arrangement of various systems / subsystems / components. The highest availability of a thermal power plant depends upon reliability and maintainability of each of its systems / subsystems. Unfortunately, system failure cannot be avoided entirely, but it can be reduced to minimum possible level by proper planning and following a suitable maintenance strategy [1]. Therefore, the failure prediction of each system in a thermal power plant is necessary for its successful and perpetual operation. Further, Water is an essential fluid used in a thermal power plant for ash removal, condenser cooling and steam generation etc. WCS of a thermal power plant regulates the flow of water and plays a vital role in its proper functioning. Therefore, it should remain operative with full capacity for longer duration.

Several techniques namely Markov technique, fault tree analysis and petrinet among others, are available in the literature for assessment of the performance / behavior of a system [2, 3]. The performance of many real-world industrial systems such as urea plants, sugar mills, paper mills and thermal power plants have been improved by using the probabilistic Markov technique, which considers a constant repair time and failure rate. Gupta et al. [4] calculated reliability indicators for plastic-pipe production plants, while, Garg et al. [5] analysed performance of a cattle feed plant employing Markov technique. Sikos and Klemes [6] optimised the availability, reliability and maintainability of a heat exchanger in a thermal power plant employing Markov technique. Modgil et al. [7] employed Markov technique to discuss the performance of a shoe upper manufacturing unit. Sharma and Vishwakarma [8] analysed performance of feeding system in a sugar industry employing Markov technique. Malik and Tewari [9] modeled the performance of water flow system in a thermal power plant using Markov technique.

To handle the uncertainty / imprecision of data due to system complexity as well as ambiguity in human verdicts, several academicians have employed fuzzy methodology (FM) for reliability analysis of systems in different sectors like healthcare, tunnel boring machines, power distribution systems and process industries. In fuzzy methodology, fuzzy numbers can be represented by triangular, trapezoidal, normal, gamma, gaussian types of membership functions. In literature, most of the studies using fuzzy methodology have been conducted using triangular membership functions due to their simplicity.

Fuzzy λ–τ technique using petrinet was proposed by Knezevic and Odoom [10] to analyse the behavior of a general production plant. Verma et al. [11] employed vague λ–τ technique with triangular fuzzy numbers and petrinet model to assess the reliability of combustion system of a gas turbine plant. The reliability of coal handling unit of a thermal power plant has been investigated by Kumar and Panchal [12] using fuzzy λ–τ technique with triangular fuzzy numbers. Srivastava et al. [13] analysed the reliability of a CNG dispensing unit using fuzzy λ–τ technique with triangular fuzzy numbers. Gopal and Panchal [14] evaluated the risk and reliability of milk process industry implementing fuzzy λ–τ technique with triangular fuzzy numbers. The performance of juice clarification unit has been analysed by Kushwaha et al. [15] employing triangular fuzzy numbers in intuitionistic fuzzy λ–τ technique.

In this paper, the performance of WCS in a thermal power plant has been analysed employing trapezoidal fuzzy numbers in fuzzy λ–τ technique coupled with petrinet modelling. The interval expressions for OR / AND transitions of petrinet model of WCS have been evaluated using trapezoidal fuzzy numbers. Here, trapezoidal fuzzy numbers have been chosen in the light of a comparison study presented by Princy and Dhenakaran [16]. They compared trapezoidal and triangular fuzzy membership functions and revealed the fact that "although trapezoidal membership functions make the procedure more complex but their performance is still better than the triangular membership functions". To counter the vagueness of input data, some reliability indicators of WCS have been evaluated at ±15%, ± 25% and ± 40% spreads. Then the obtained fuzzy values of reliability indicators have been defuzzified using COA method for quantitative analysis of performance of WCS.

2. Proposed Methodology

Fuzzy λ–τ technique [10] is a quantitative technique for analysing the performance of repairable systems in an ambiguous environment. In order to study the behavior of complex repairable systems, failure and repair rates are assumed to be constant in this technique. Moreover, this technique uses fuzzified values of repair time and failure rate data of each of its components / subsystems. This technique is more powerful than Markov technique since it takes care of ambiguity in repair time and failure rate data. Numerous researchers have implemented it in various areas like sugar mills, thermal power plants and chemical sector, among others. The various steps of methodology used in this paper are given below:

Step 1: Collect the information regarding various components and subsystems of WCS to construct its fault tree model and then its analogous petrinet model.

Step 2: Collect the data for failure rate and repair time of various components and subsystems of WCS.

Step 3: Fuzzify failure and repair data of each components and subsystems of WCS using trapezoidal membership function.

Step 4: Compute various reliability indicators of WCS at different spreads.

Step 5: Defuzzify the fuzzy values of reliability indicators employing COA method.

Step 6: Analyse the behavior of WCS.

3. Preliminaries

This section discusses some fuzzy set theory concepts to be used in this paper [17-19].

3.1. Fuzzy set

A fuzzy set $\widetilde{A}$ is defined as

$$\widetilde{A} = \left\{ \left(x, \mu_{\widetilde{A}}(x) \right) : x \in X \right\}, \tag{1}$$

where, $\mu_{\widetilde{A}} : X \rightarrow [0,1]$ is the membership function.

3.2. Fuzzy number

A normal convex fuzzy set $\widetilde{A}$ on real line R is called fuzzy number if its support is bounded.

3.3. Trapezoidal fuzzy number

A trapezoidal fuzzy number $\widetilde{A} = (l_1, l_2, u_1, u_2)$ has its membership function $\mu_{\widetilde{A}}(x)$ given by

$$\mu_{\widetilde{A}}(x) = \begin{cases} \dfrac{x - l_1}{l_2 - l_1}, & l_1 \le x \le l_2 \\ 1, & l_2 \le x \le u_1 \\ \dfrac{u_2 - x}{u_2 - u_1}, & u_1 \le x \le u_2 \\ 0, & otherwise \end{cases} . \tag{2}$$

3.4. α–cut of a trapezoidal fuzzy number

The α-cut of the trapezoidal fuzzy number $\widetilde{A}$ is given by

$$\left[l^\alpha, u^\alpha\right] = \left[l_1 + (l_2 - l_1)\alpha,\ u_2 - (u_2 - u_1)\alpha\right], \qquad \text{where, } \alpha \in [0,1]$$

and is shown in the figure 1.

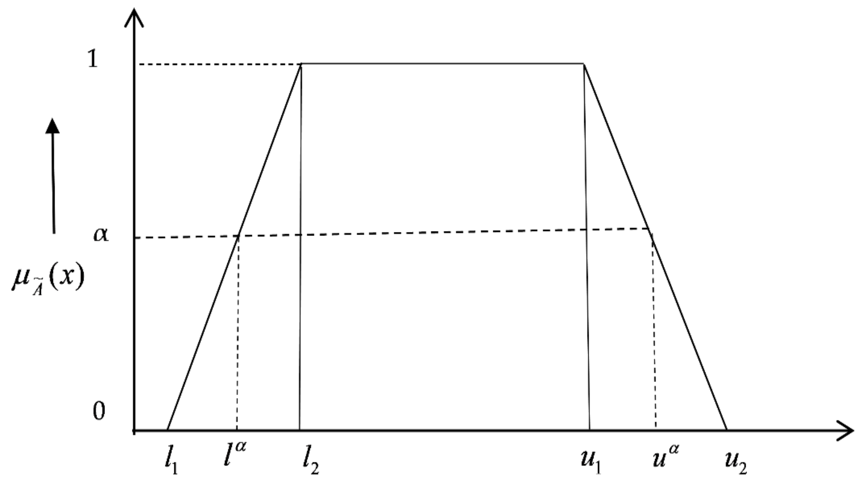


Figure 1: *Trapezoidal fuzzy number with α-cut*

4. An application

Water circulation system has a significant role for proper functioning of a thermal power plant and therefore should remain operative for longer duration. The hot water collected from the condenser enters into the condensate extraction pump and then passes to low pressure (LP) heater where its temperature is further increased to improve the system efficiency. The hot water further passes through the deaerator where the dissolved gases are removed from the hot water. The ash content free water then enters into the high pressure (HP) heater with the help of the boiler feed pump and further enters into the economiser. In economiser, the hot flue gases exiting the boiler are used to heat the feed water. This heated water enters the boiler drum, where super-heated steam is generated by super heaters of boiler. This steam is expanded in turbine to produce mechanical energy by the turbine shaft. Then the exhaust steam is condensed in condenser. Figure 2 depicts schematic flow diagram of WCS. It is made up of five subsystems as described below:

i. **Subsystem 1 (SS_1):** It consists a condensate extraction pump (CEP) which is used to extract the condensate. The system will stop working if CEP fails.

ii. **Subsystem 2 (SS_2):** It is made up of three LP heaters which are linked in a parallel arrangement and are used to raise the condensate's temperature. If one of them fails, the system will run at a lower efficiency.

iii. **Subsystem 3 (SS_3):** It consists a deaerator to extract dissolved gases from hot water obtained from LP heater. The efficiency of the system reduces if deaerator fails.

iv. **Subsystem 4 (SS_4):** Three boiler feed pumps (BFP) are connected in a parallel configuration in this subsystem. They are responsible for supplying hot water to the HP heater. The system efficiency is reduced if any of these pumps fails.

v. **Subsystem 5 (SS_5):** It is made up of two HP heaters that are connected in a parallel manner. They are used to further heat the water received from the boiler feed pump. The efficiency of the system reduces if anyone HP heater fails.

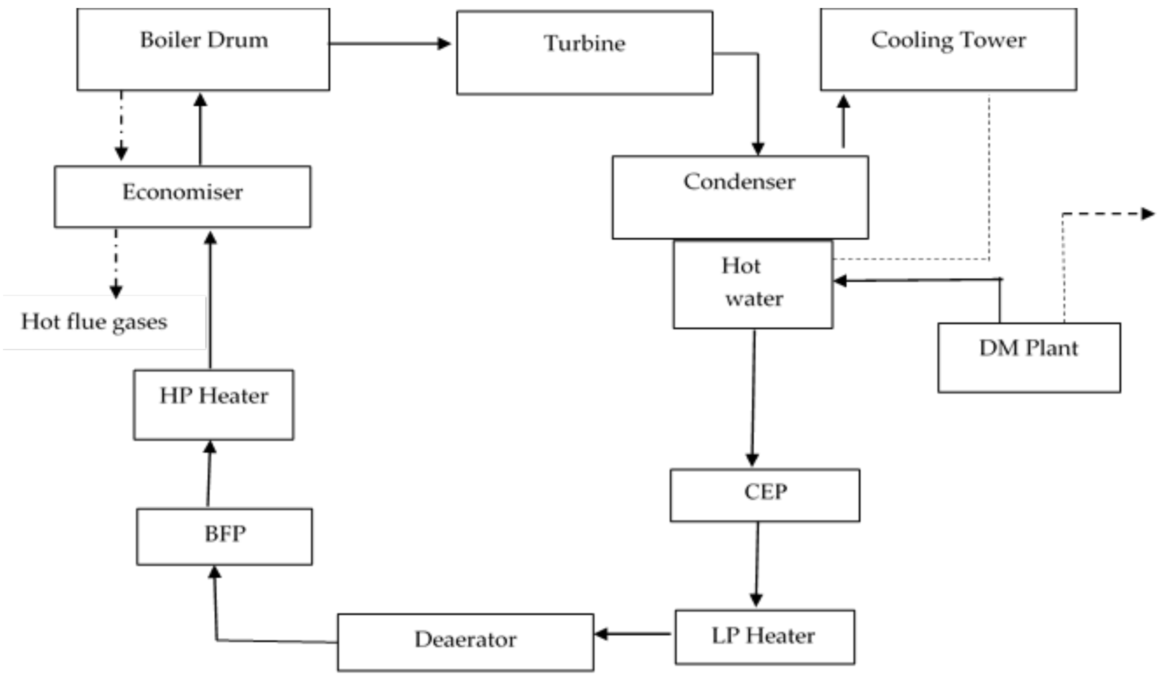


Figure 2: *Flow diagram of water circulation system*

5. Reliability Analysis of WCS

The reliability of WCS has been calculated using following steps:

Step 1. Fault tree and petrinet models construction: First a fault tree model of the complicated parallel and series arrangement of the WCS has been prepared (figure 3) and then its equivalent petrinet model has been developed (figure 4). The AND gate represents the parallel arrangement of components, while, the OR gate represents the series arrangement of components in these models.

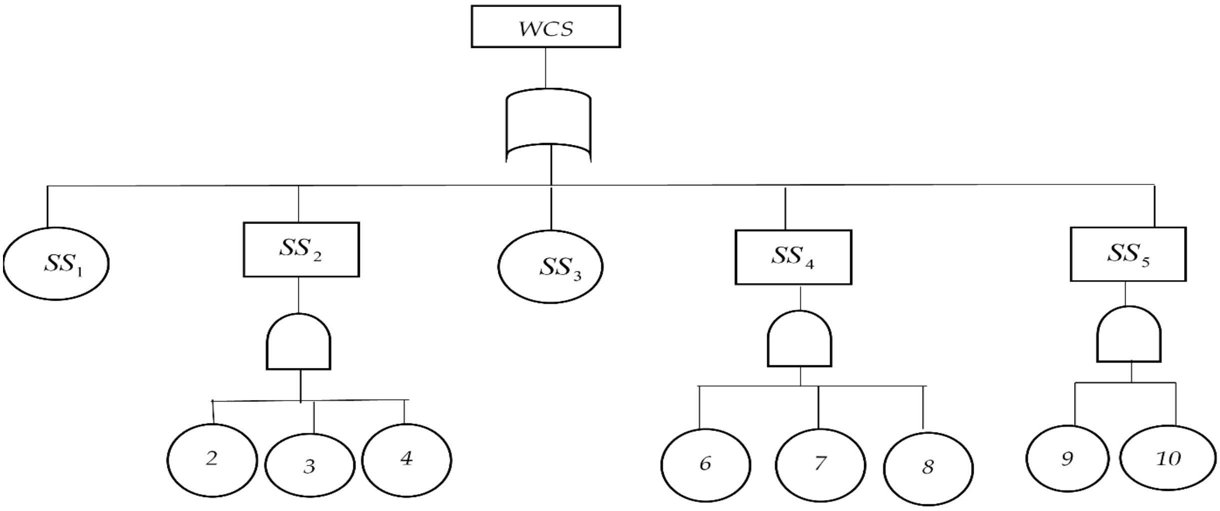


Figure 3: *Fault tree model of WCS*

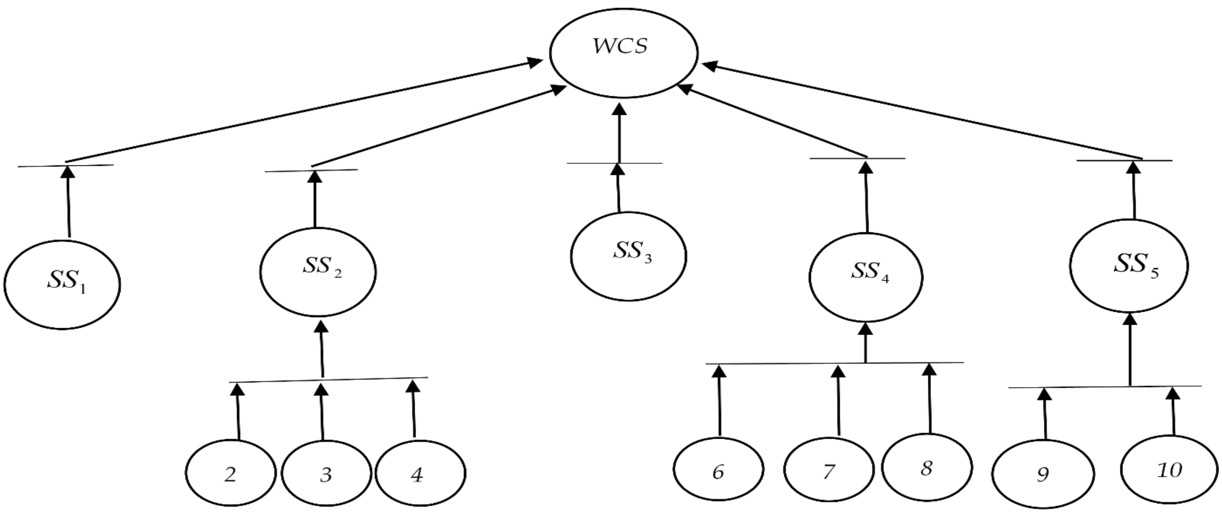


Figure 4: *Petrinet model of WCS*

Step 2. Data extraction: The failure and repair data of all ten components of WCS acquired from service logbook and validated by a maintenance specialist are given in Table 1 [20].

Table 1: *Failure and repair data of different components of WCS*

Components	Failure rate λ_k (failures / h)	Repair time τ_k (h)
Condensate extraction pump (SS_1) (k = 1)	5.78×10^{-5}	11
Low pressure heater (SS_2) (k = 2, 3, 4)	3.85×10^{-5}	16
Deaerator (SS_3) (k = 5)	1.15×10^{-4}	9
Boiler feed pump (SS_4) (k = 6, 7, 8)	1.15×10^{-4}	11
High pressure heater (SS_5) (k = 9, 10)	3.85×10^{-5}	11

Step 3. Fuzzification of crisp data: The acquired crisp data for repair time and failure rate has been fuzzified into trapezoidal fuzzy numbers using trapezoidal membership functions to reduce ambiguity in the collected data. The values of $\lambda_{k2}, \lambda_{k3}, \tau_{k2}$ and τ_{k3} have been fixed at ±10% of acquired crisp data whereas the values of $\lambda_{k1}, \lambda_{k4}, \tau_{k1}$ and τ_{k4} have been changed at ±15%, ± 25% and ± 40% spreads for each component of the system. Figure 5 depicts trapezoidal fuzzy numbers for failure rate (λ_k) and repair time (τ_k) for k^{th} component.

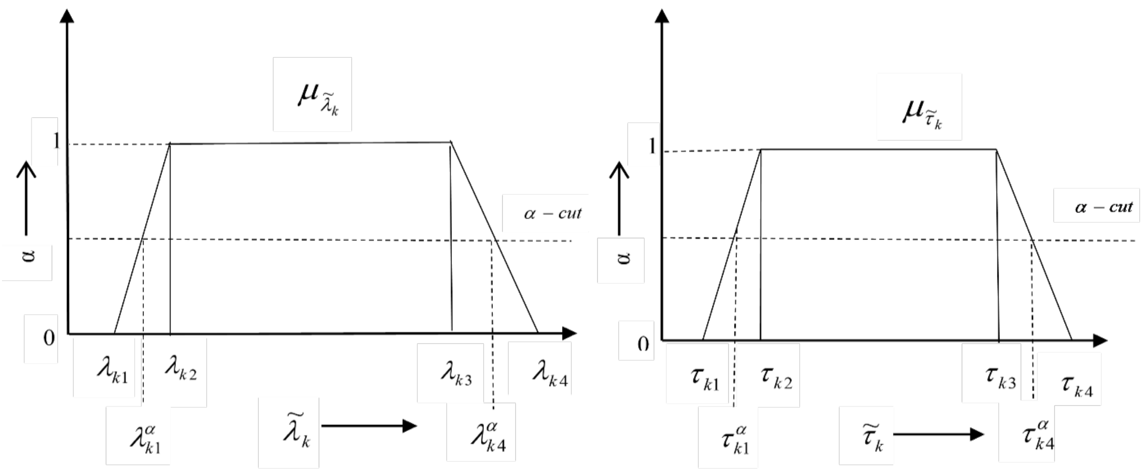


Figure 5: *Trapezoidal fuzzy numbers for failure rate* (λ_k) *and repair time* (τ_k) *for* k^{th} *component*

Step 4. Computation of fuzzy reliability indicators: After fuzzifying the failure rate and repair time of all the components of WCS shown in petrinet model (figure 4) as trapezoidal fuzzy numbers, the fuzzy failure rate and repair time of the top most position of petrinet model of WCS are evaluated using interval expressions for OR / AND transitions presented in equations (3-6). These interval expressions are obtained utilising the extension principle and α-cut along with interval arithmetic operations on the basic λ-τ expressions for OR / AND gates given in Table 2.

Table 2: *Basic expressions for λ-τ technique*

Gate	λ_{OR}	τ_{OR}	λ_{AND}	τ_{AND}
Expressions (n-inputs)	$\sum_{k=1}^{n}\lambda_k$	$\frac{\sum_{k=1}^{n}\lambda_k\tau_k}{\sum_{k=1}^{n}\lambda_k}$	$\prod_{l=1}^{n}\lambda_l\left[\sum_{l=1}^{n}\prod_{\substack{l=1\\k\neq l}}^{n}\tau_l\right]$	$\frac{\prod_{k=1}^{n}\tau_k}{\sum_{l=1}^{n}\left[\prod_{\substack{k=1\\k\neq l}}^{n}\tau_k\right]}$

Interval expressions for OR transition

$$\tau^{\alpha}=\left[\frac{\sum_{k=1}^{n}\left[\{\lambda_{k1}+(\lambda_{k2}-\lambda_{k1})\alpha\}.\{\tau_{k1}+(\tau_{k2}-\tau_{k1})\alpha\}\right]}{\sum_{k=1}^{n}\{\lambda_{k4}-(\lambda_{k4}-\lambda_{k3})\alpha\}},\ \frac{\sum_{k=1}^{n}\left[\{\lambda_{k4}-(\lambda_{k4}-\lambda_{k3})\alpha\}.\{\tau_{k4}-(\tau_{k4}-\tau_{k3})\alpha\}\right]}{\sum_{k=1}^{n}\{\lambda_{k1}+(\lambda_{k2}-\lambda_{k1})\alpha\}}\right], \tag{3}$$

$$\lambda^{\alpha}=\left[\sum_{k=1}^{n}\{\lambda_{k1}+(\lambda_{k2}-\lambda_{k1})\alpha\},\ \sum_{k=1}^{n}\{\lambda_{k4}-(\lambda_{k4}-\lambda_{k3})\alpha\}\right]. \tag{4}$$

Interval expressions for AND transition

$$\tau^{\alpha}=\left[\frac{\prod_{k=1}^{n}\{\tau_{k1}+(\tau_{k2}-\tau_{k1})\alpha\}}{\sum_{l=1}^{n}\left[\prod_{\substack{k=1\\k\neq l}}^{n}\{\tau_{k4}-(\tau_{k4}-\tau_{k3})\alpha\}\right]},\ \frac{\prod_{k=1}^{n}\{\tau_{k4}-(\tau_{k4}-\tau_{k3})\alpha\}}{\sum_{l=1}^{n}\left[\prod_{\substack{k=1\\k\neq l}}^{n}.\{\tau_{k1}+(\tau_{k2}-\tau_{k1})\alpha\}\right]}\right], \tag{5}$$

$$\lambda^{\alpha}=\left[\Pi_{k=1}^{n}\{\lambda_{k1}+(\lambda_{k2}-\lambda_{k1})\alpha\}.\sum_{l=1}^{n}\left[\Pi_{k=1,k\neq l}^{n}\{\tau_{k1}+(\tau_{k2}-\tau_{k1})\alpha\}\right],\right.$$
$$\left.\Pi_{k=1}^{n}\{\lambda_{k4}-(\lambda_{k4}-\lambda_{k3})\alpha\}.\sum_{l=1}^{n}\left[\Pi_{k=1,k\neq l}^{n}\{\tau_{k4}-(\tau_{k4}-\tau_{k3})\alpha\}\right]\right]. \tag{6}$$

To analyse the performance of WCS quantitatively the reliability indicators viz., failure rate, repair time, expected number of failures, mean time to failure, mean time to repair, mean time between failures, reliability and availability are calculated using expressions given in Table 3 at ±15%, ±25% and ± 40% spreads for α = 0.0 (0.1) 1.0.

Table 3: *Expressions for reliability indicators*

Reliability indicator	Expression
Expected number of failures	$ENOF = \frac{\lambda \gamma t}{\gamma + \lambda} + \frac{\lambda^2}{(\gamma + \lambda)^2}\left[1 - e^{-(\gamma + \lambda)t}\right]$
Mean time to failure	$MTTF = \frac{1}{\lambda}$
Mean time to repair	$MTTR = \frac{1}{\gamma} = \tau$
Mean time between failures	$MTBF = MTTF + MTTR$
Reliability	$R = e^{-\lambda t}$
Availability	$A = \frac{\gamma}{\gamma + \lambda} + \frac{\lambda}{\gamma + \lambda} e^{-(\gamma + \lambda)t}$

The left and right spread fuzzy values of the reliability indicators of WCS at 15% spread are presented in Table 4 and Table 5, respectively. The fuzzy reliability indicators at ±15%, ± 25% and ± 40% spreads for α = 0.0 (0.1) 1.0 are depicted in figure 6.

Table 4: *The left spread fuzzy values of WCS at 15% spread*

α	Failure Rate $\times 10^{-4}$ (/h)	Repair Time (h)	ENOF $\times 10^{-2}$	MTBF $\times 10^{3}$ (h)	Reliability	Availability
1.0	1.555441	7.118736	2.610373	5.266834	0.968564	0.997535
0.9	1.546797	7.007978	2.595923	5.242912	0.968423	0.997487
0.8	1.538153	6.898654	2.581470	5.219206	0.968282	0.997438
0.7	1.529509	6.790745	2.567017	5.195713	0.968142	0.997389
0.6	1.520865	6.684231	2.552621	5.172431	0.968001	0.997339
0.5	1.512222	6.579095	2.538106	5.149355	0.967861	0.997288
0.4	1.503578	6.475318	2.523648	5.126484	0.967720	0.997236
0.3	1.494934	6.372881	2.509189	5.103816	0.967579	0.997183
0.2	1.486290	6.271767	2.494729	5.081346	0.967439	0.997129
0.1	1.477647	6.171960	2.480267	5.059073	0.967298	0.997074
0	1.469003	6.073441	2.465804	5.036994	0.967158	0.997019

Table 5: *The right spread fuzzy values of WCS at 15% spread*

α	Failure Rate $\times 10^{-4}$ (/h)	Repair Time (h)	ENOF $\times 10^{-2}$	MTBF $\times 10^{3}$ (h)	Reliability	Availability
1.0	1.901244	12.999494	3.186822	6.442043	0.974207	0.998894
0.9	1.909890	13.191290	3.201183	6.478163	0.974349	0.998917
0.8	1.918536	13.385785	3.215540	6.514689	0.974490	0.998940
0.7	1.927182	13.583026	3.229894	6.551628	0.974632	0.998962
0.6	1.935829	13.783060	3.244246	6.588987	0.974773	0.998984
0.5	1.944475	13.985935	3.258594	6.626774	0.974915	0.999006
0.4	1.953122	14.191699	3.272938	6.664995	0.975056	0.999027
0.3	1.961768	14.400403	3.287279	6.703659	0.975198	0.999048
0.2	1.970415	14.612099	3.301617	6.742773	0.975339	0.999069
0.1	1.979061	14.826838	3.315951	6.782345	0.975481	0.999089
0	1.987708	15.044674	3.330282	6.822383	0.975623	0.999109

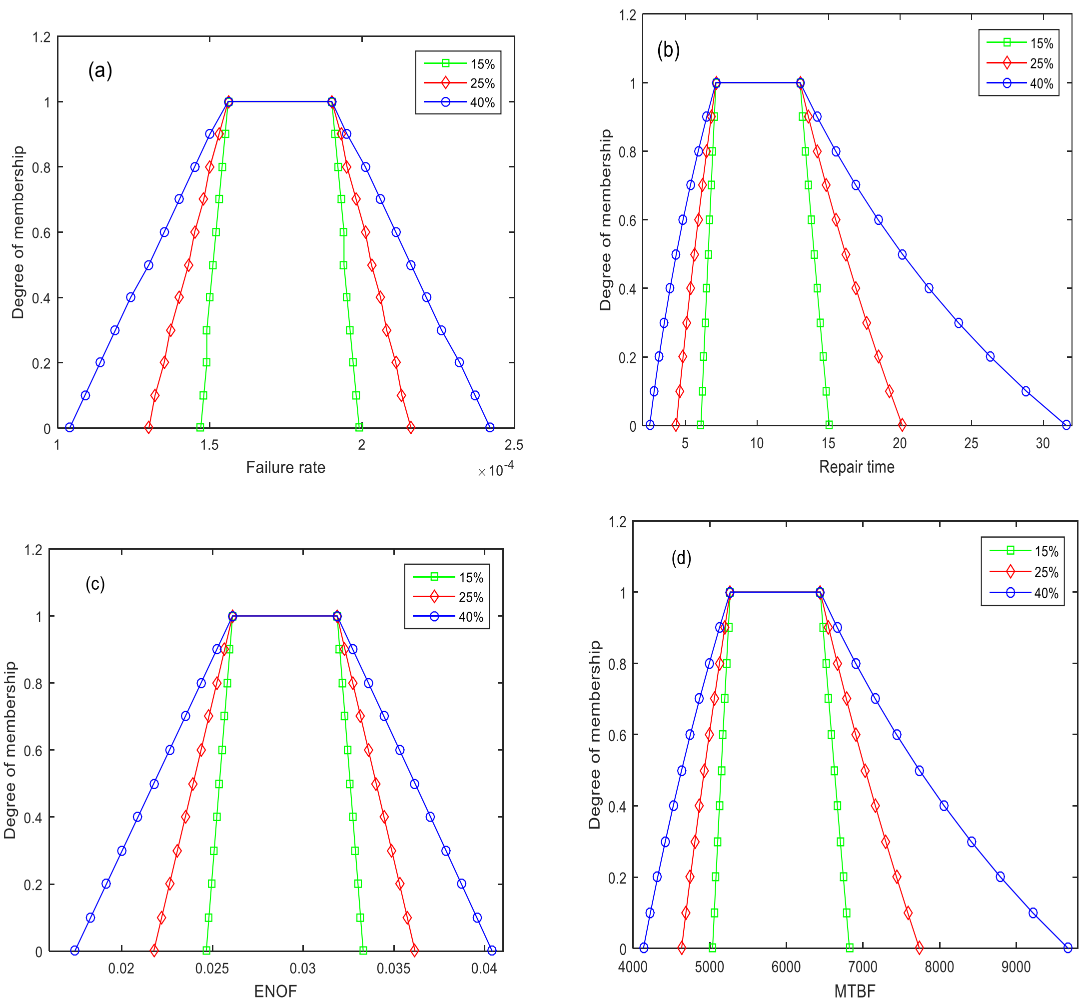


Figure 6: *Graph representation of fuzzy reliability indicators*

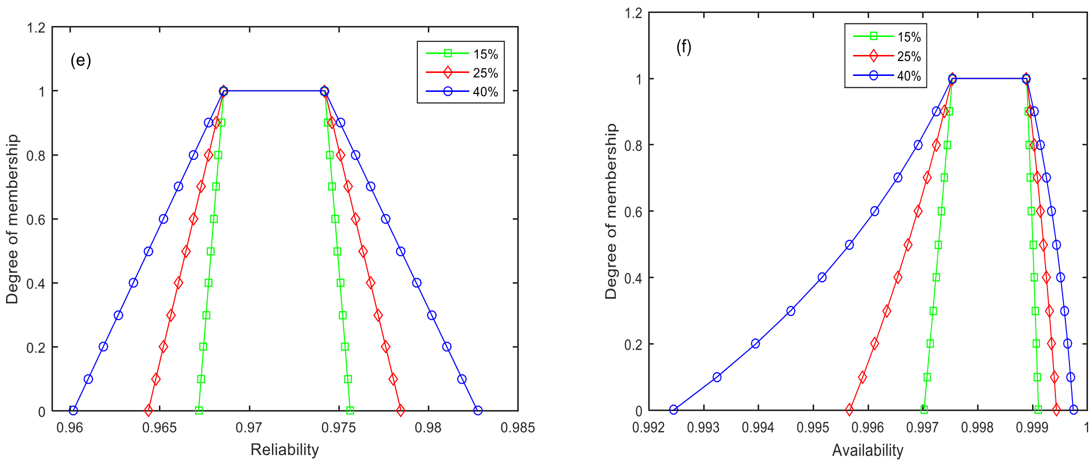


Figure 6: *Graph representation of fuzzy reliability indicators (continued)*

Step 5. Defuzzification of fuzzy reliability indicators: The fuzzy results of reliability indicators are defuzzified by employing COA method for defuzzification. The reliability indicators have been defuzzified for ±15%, ±25% and ±40% spreads using equation (7) and are presented in Table 6 together with their crisp values.

$$\bar{x} = \frac{\int_{x_1}^{x_2} x.\mu_{\tilde{A}}(x)dx}{\int_{x_1}^{x_2} \mu_{\tilde{A}}(x)dx} \tag{7}$$

Table 6: *Values of reliability indicators of WCS*

Reliability Indicator	Crisp Value	Defuzzified Values		
		Spreads		
		± 15 %	± 25 %	± 40 %
Failure rate $\times 10^{-4}$ (/h)	1.728332	1.730000	1.730000	1.731031
Repair time (h)	9.668175	10.327083	11.320939	14.323654
ENOF $\times 10^{-2}$	2.898757	2.898380	2.897209	2.894103
MTBF $\times 10^{3}$ (h)	5.795593	5.894658	6.043965	6.493699
Reliability	0.971382	0.971388	0.971397	0.971421
Availability	0.998332	0.998134	0.997831	0.996912

6. Results and Discussion

It is observed from the behavioral graphs given in figure 6 that the membership curves of many reliability indicators are deformed trapezoids, whose non-parallel sides are parabolic. Further, it is demonstrated that the defuzzified values vary with changes in spread. For instance, for the spread expansion from ±15% to ±25%, the failure rate does not change, the reliability increases marginally, the repair time and MTBF increase by 9.62% and 2.53%, respectively, the ENOF and availability decrease by 0.04% and 0.03%, respectively. Further, for spread expansion from ±25% to ±40%, failure rate, repair time, MTBF and reliability increase by 0.06%, 26.52%, 7.44% and 0.002%, respectively, while, ENOF and availability decrease by 0.10% and 0.09%, respectively. The effect of spread change is most significant on repair time in comparison to other reliability indicators, resulting in system availability loss. These findings will assist plant managers in planning and adapting best maintenance policy to improve performance of WCS and minimise operational and maintenance expenses.

7. Conclusion

In this study, the performance of WCS of a thermal power plant has been analysed using fuzzy λ–τ technique. Trapezoidal fuzzy numbers have been employed to remove ambiguity in acquired data. To eliminate unexpected failure of the plant, the trend of different reliability indicators viz. repair time, failure rate, reliability, availability, ENOF and MTBF with respect to spread changes has been analysed. The maintenance personnels of the considered system should focus at system's availability because it diminishes by increasing spread, which is unfavorable. The reason behind this is that repair time varies more rapidly in comparison to other reliability indicators. A structured framework has been established in this research to assist maintenance engineers for improving availability of the considered plant.

References

[1] Yang, S. K. (2004). A condition – based preventive maintenance arrangement for thermal power plant. *Electric Power Systems Research*, 72(1):49-62.

[2] Jin, Q. and Sugasawa, Y. (1995). Representation and analysis of behavior for multiprocessor systems by using stochastic petri nets. *Mathematical* and *Computer Modelling*, 22(10–12):109–118.

[3] Adamyan, A. and He, D. (2004). System failure analysis through counters of petri net models. *Quality and Reliability Engineering International*, 20(4):317–335.

[4] Gupta, P., Lal, A. K., Sharma, R. K. and Singh, J. (2007). Analysis of reliability and availability of serial processes of plastic-pipe manufacturing plant: a case study. *International Journal of Quality and Reliability Management*, 24(4):404–419.

[5] Garg, D., Singh, J. and Kumar, K. (2009). Performance analysis of a cattle feed plant. *ICFAI University Journal of Science and Technology,* 5(2):83-94.

[6] Sikos, L. and Klemes, J. (2010). Reliability, availability and maintainability optimisation of heat exchanger networks. *Applied Thermal* Engineering, 30(1): 63-69.

[7] Modgil, V., Sharma, S. K. and Singh, J. (2013). Performance modeling and availability analysis of shoe upper manufacturing unit. *International Journal of Quality and Reliability Management*, 30(8):816-831.

[8] Sharma, S. P. and Vishwakarma, Y. (2014). Application of Markov process in performance analysis of feeding system of sugar industry. *Journal of Industrial Mathematics*, 2014:1-9.

[9] Malik, S. and Tewari, P. C. (2018). Performance modeling and maintenance priorities decision for the water flow system of a coal-based thermal power plant. *International Journal of Quality and Reliability Management*, 35(4):996–1010.

[10] Knezevic, J. and Odoom, E. R. (2001). Reliability modelling of repairable systems using Petri nets and fuzzy lambda-tau methodology. *Reliability Engineering & System Safety*, 73:1-17.

[11] Verma, M., Kumar, A. and Singh, Y. (2013). Vague reliability assessment of combustion system using Petri nets and vague lambda-tau methodology. *Engineering Computations*, 30(5): 665-681.

[12] Panchal, D. and Kumar, D. (2017). Stochastic behaviour analysis of real industrial system. *International Journal of System Assurance Engineering and Management*, 8(2):1126-1142.

[13] Srivastava, P., Khanduja, D., Narayanan, G. A., Agarwal, M. and Tulsian, M. (2019). Reliability analysis of CNG dispensing unit by lambda-tau Approach. *Operations Management and Systems Engineering*, Lecture Notes on Multidisciplinary Industrial Engineering. 153-168.

[14] Gopal, N. and Panchal, D. (2021). A structured framework for reliability and risk evaluation in the milk process industry under fuzzy environment. *Facta Universitatis. Series: Mechanical Engineering*, 19(2):307-333.

[15] Kushwaha, D. K., Panchal, D. and Sachdeva, A. (2022). Intuitionistic fuzzy modelling-based integrated framework for performance analysis of juice clarification unit. *Applied Soft Computing*, 124, Article ID 109056.

[16] Princy, S. and Dhenakaran, S. S. (2016). Comparison of triangular and trapezoidal fuzzy membership function. *Journal of Computer Science and Engineering*, 2:46-51.

[17] Chen, C. H. and Mon, D. L. (1993). Fuzzy system reliability analysis by interval of confidence. *Fuzzy Set System*, 56(1): 29-35.

[18] Zadeh, L. A. Fuzzy Sets, Fuzzy Logic, Fuzzy Systems: Selected Papers, World Scientific, Singapore, 1996.

[19] Zimmermann, H. Fuzzy Set Theory and its Applications, third ed., Kluwer Academic Publishers, New York, 2001.

[20] Panchal, D. (2015). Performance analysis and maintenance planning for the real operating systems. Ph.D. thesis, Indian Institute of Technology Roorkee, India.

RELIABILITY ANALYSIS OF PARALLEL SYSTEM USING PRIORITY TO PM OVER INSPECTION

Neetu Dabas
•
Department of Statistics, Amity University, Noida, Uttar Pradesh-201313
Email: neetudabas2@gmail.com

Reetu Rathee
•
School of Business, Galgtias University, Greater Noida, Uttar Pradesh-203201
Email: dr.reetustat@gmail.com

Abhishek Sheoran
•
Department of Statistics, Ramanujan College, University of Delhi, Delhi-110019
Email: stat.abhi@gmail.com

Abstract

Reliability optimization of a system is an extant problem. By solving these problems, new methodologies are obtained that have invent new engineering technology and changes the management perspective. Aim of the reliability analysis is to study the failure mechanisms of a system and and outcomes of the analysis serve to identify design solutions and maintenance actions for preventing the failures from occurring. So, it is used to evaluate and improve the quality of products, processes, and systems. Measurement, planning, and improvement in the reliability are the things which are well do in any business but only when efforts are focused on important problems which are highlighted by monetary values, improve reliability, reduce unreliability costs, generate more profit, and get more business. To serve this purpose, present study investigates a parallel system of two identical units which is based on several assumptions like, the system is served by one serviceman who is immediately available for service when it will call. System failure rate is fix and the failure type (repairable or replaceable) is known by inspection. The failure and repair activities are stochastically independent, and their rates are exponentially distributed. Priority to PM over inspection is given in the system. Several measures of reliability effectiveness like MTSF, availability and cost-benefit analysis of the system are obtained by semi-Markov and regenerative point approach. Reliability characteristics parameters are random variables, and results are obtained in the form of graphs and tables by changing the values of these variables one by one, while keeping other variables constant at that point. From the results we conclude how to make the given system more profitable. Findings of present system model shows that when the failure rate is low then the system obtained more profit by increasing preventive maintenance rate. On the other hand, when failure rate going high then we make the system more profitable by increasing inspection rate. These insights of modelling and analysis helps the system developers and managers to make good choices of action against specified criteria that managing engineered products and industrial plants safely and reliably. This leads to more profit and making a business more growing.

Keywords: Parallel system, priority, inspection, semi-Markov, regenerative points, repair activities, profit.

1. Introduction

The objectives of any equipment or system manufactured is to design it in such a way that it achieves its goals in terms of production. A reliable equipment is the one which works satisfactorily for a given time period under given environmental conditions without any interruptions. Hence the reliability is the key point in the manufacturing/ production industries. Although ever-increasing urge of the society is making the design of system more complicated and to control the strength and effectiveness of failure of such system reliability experts frame a model which is more productive and profitable. Various researcher has done a lot of research work in the reliability theory to improve the repair techniques. Several authors such as Gupta and Agarwal [1], Dhillon and Yang [2] extensively discussed complex systems by considering various failure and repair disciplines. Ram [3] gave the summery of various reliability approaches. Parallel redundancy is highly used by the researcher to enhance the system reliability. Hitomi [4] investigates the reliability of a manufacturing system in which two machines are arranged in parallel. Termoto et al. [5] studied an optimal inspection policy for an n-unit parallel system which is checked at successive times and PM is carried out after the failure of a certain number of units at each inspection. Gupta et al. [6] analyzed a two non-identical unit parallel system. They have taken the joint distribution of lifetimes of both units as bivariate negative exponential. Malik et al. [7] considered two reliability models with two parallel units in which one is original and other is duplicate. Priority to repair of the original unit is given in model I and no priority is given in model II. Yu and Khambadkone [8] derived a parallel inverter system to analyses the reliability and cost optimization using sensitivity analysis. Rathee and Malik [9] observed a parallel system under the aspect of priority to PM over repair and replacement. Sivakumar and Jayanth [10] studied the reliability of a Mobile network system during communication. Kakkar et.al. [11] worked on the reliability and profit analysis of a parallel industrial system. Bhardwaj and Parasher [12] analyzed a cold standby system with geometric failure and repair rates. Pundir et.al. [13] studied a two non-identical units' parallel system with priority in repair to first unit. Chopra and Ram [14] investigate the reliability measures of the system, which has two dissimilar units in the parallel network under copula. Li et al. [15] developed a two similar component parallel degradation repairable system. They considered that when the repairman is on vacation then the failed component is not repaired as good as new. Dabas and Rathee [16,17] derived a parallel system with the idea of priority to preventive maintenance over replacement and inspection for repair activities. Sherbeny [18] studied the impact of some system parameters on an industrial system consisting parallel units with one repairer and optional vacations under Poisson shocks.

In this paper we consider a two identical units parallel system using priority to PM over inspection. All repair activities and inspection is done by a single serviceman. Units' failure rate is taken to be constant, and inspection is done to find the type of failure. Time taken in repair activities is distributed arbitrarily and their rates are exponentially distributed. The failure and repair activities are stochastically independent. To meet the objective of reliability evolution we derive some suitable measure like MTSF, availability and cost-benefit analysis of the system using semi-Markov and regenerative point approach. Graphs are plotted to observe the change in the behaviour of these measures with failure rate for particular cases of various rates included in the system.

2. Notations for System Model

λ	:	Constant Failure Rate
$a/b/\alpha_0$	:	Rate by which system goes for Repair / Replacement / Preventive Maintenance respectively
$\alpha/\beta/\gamma/\theta$	:	Repair / Replacement / Inspection / Preventive Maintenance rate respectively done by the server
h(t)/f(t)/r(t)/g(t)	:	pdf of the Inspection / Repair / Replacement / Preventive Maintenance time respectively
H(t)/F(t)/R(t)/G(t)	:	cdf of the Inspection / Repair / Replacement / Preventive Maintenance time respectively
p_{ij}	:	Transition probability from state S_i to state S_j
$p_{ij.kr}$	:	Transition probability from state S_i to state S_j via state S_k, S_r
$Q_{ij}(t)/q_{ij}(t)$	:	Cdf/pdf of passage time from regenerative state S_i to a regenerative state S_j or to a failed state Sj without visiting any other regenerative state in (0, t]
$Q_{ij.kr}(t)/q_{ij.kr}(t)$	:	pdf/cdf of direct transition time from regenerative state S_i to a regenerative state S_j or to a failed state S_j visiting state S_k, S_r once in (0, t]
μ_i	:	Mean sojourn time in state S_i
m_{ij}	:	Contribution to mean sojourn time in state S_i when the system transits directly to state S_j
*/**	:	Symbol for Laplace transformation/ Laplace Stieltjes Transformation
©/Ⓢ	:	Symbol for Laplace transformation/Laplace Stieltjes convolution

3. System Description and Assumptions

To evaluate the effectiveness of the present research, the numerical data is examined in order to establish essential assumptions to the system model. The Semi-markov and re-generative point process are used to provide formulations of system dependability measures such as reliability, mean time to system failure (MTSF), availability, and profit function. Numerical examples are provided to demonstrate the acquired conclusions. Results are obtained in tabular and graphical form to investigate the influence of various system features.

Assumptions

- Initially both the units are in working mode
- Units are failed with constant rate
- System is served by single serviceman
- All the times associated with all events are random and independent.
- All the repair activities follow exponential distribution

Table 1: *Description of the states*

States	Description
S_0	Both the units are operative
S_1	One unit is operative and other is failed under inspection
S_2	Resume for PM
S_3	One is working and other is failed under replacement
S_4	One unit is continuously under inspection from previous state and other is waiting for inspection
S_5	One is working and other is failed under repair
S_6	One unit is waiting for inspection from previous state and other is under PM
S_7	One unit is working and other under PM
S_8	One unit is continuously under replacement from previous state and other is waiting

	for inspection
S_9	One unit is under repair and other is continuously waiting for inspection from previous state
S_{10}	One unit is under replacement and other is continuously waiting for inspection from previous state
S_{11}	One unit is continuously under repair from previous state and other is waiting for inspection
S_{12}	One unit is continuously under repair from previous state and other is waiting for PM
S_{13}	One unit is continuously under replacement from previous state and other is waiting for PM
S_{14}	One unit is continuously under PM from previous state and other is waiting for inspection

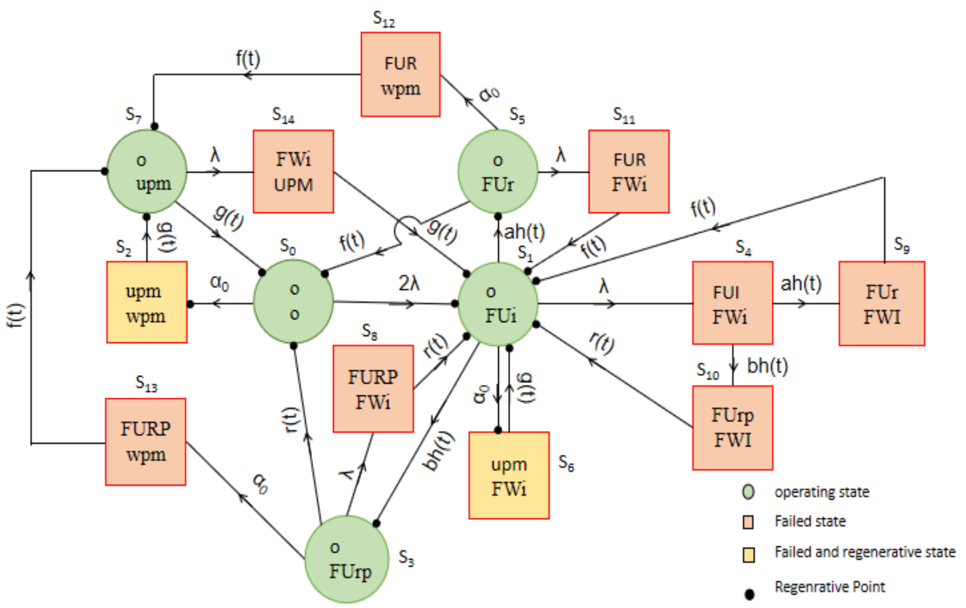


Figure 1: *Transition State Diagram*

4. Formulation and Stochastic Analysis of the Model

4.1. Transition Probabilities & Mean Sojourn Times (μ_i)

Steady- state transition probabilities from regenerative state i to state j are given by the formula

$$p_{ij} = Q_{ij}(\infty) = \int_0^\infty q_{ij}(t)dt$$

$p_{01} = \frac{2\lambda}{2\lambda+\alpha_0}$, $p_{02} = \frac{\alpha_0}{2\lambda+\alpha_0}$, $P_{13} = bh^*(\lambda+\propto_0)$, $p_{14} = \frac{\lambda}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$,

$p_{15} = ah^*(\lambda+\propto_0)$, $p_{16} = \frac{\propto_0}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$, $p_{30} = r^*(\lambda+\propto_0)$, $p_{38} = p_{31.8} = \frac{\lambda}{\lambda+\propto_0}(1-r^*(\lambda+\propto_0))$,

$p_{3,13} = p_{37.13} = \frac{\propto_0}{\lambda+\propto_0}(1-r^*(\lambda+\propto_0))$,

$p_{49} = a$, $p_{4,10} = b$, $p_{50} = f^*(\lambda+\propto_0)$, $p_{5,11} = p_{51.11} = \frac{\lambda}{\lambda+\propto_0}(1-f^*(\lambda+\propto_0))$,

$p_{5,12} = p_{57.12} = \frac{\propto_0}{\lambda+\propto_0}(1-f^*(\lambda+\propto_0))$, $p_{70} = g^*(\lambda)$, $p_{7,14} = p_{71.14} = 1-g^*(\lambda)$,

$p_{11.49} = \frac{\lambda a}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$, $p_{11.4,10} = \frac{\lambda b}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$,

$p_{1,14.6} = \frac{\propto_0 b}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$, $p_{17.6,13} = \frac{\propto_0 a}{\lambda+\propto_0}(1-h^*(\lambda+\propto_0))$,

$p_{27} = p_{61} = p_{81} = p_{91} = p_{10,1} = p_{11,1} = p_{12,7} = p_{13,7} = p_{14,1} = 1$

It is noticed that the $\sum_j p_{ij} = 1$ for all possible values of 'i'.
Further mean sojourn times (μ_i) is the expected time taken by the system in a particular state before transiting to any other state. If T_i is the sojourn time in state 'i' ,then
$\mu_i = E(t) = \int_0^\infty P(T_i > t)\,dt = \sum_j m_{ij}$ and $m_{ij} = \frac{d[Q_{ij}^{**}(s)]}{ds}|s = 0.$
Expressions for μ_i are given as

$$\mu_0 = \frac{1}{2\lambda + \alpha_0},\ \mu_1 = \frac{1}{\lambda + \alpha_0}(1 - h^*(\lambda + \alpha_0)), \mu_3 = \frac{1}{\lambda + \alpha_0}(1 - r^*(\lambda + \alpha_0))$$

$$\mu_5 = \frac{1}{\lambda+\alpha_0}(1 - f^*(\lambda + \alpha_0)), \mu_7 = \frac{1}{\lambda}(1 - g^*(\lambda))$$

$$\mu_1' = [\frac{1}{\lambda + \alpha_0} + \frac{\lambda}{\lambda + \alpha_0}(\frac{b}{\beta} + \frac{1}{\gamma} + \frac{a}{\alpha})](1 - h^*(\lambda + \alpha_0))$$

$$\mu_3' = \frac{1}{\beta},\ \mu_5' = \frac{1}{\propto}, \mu_7' = \frac{1}{\theta}$$

4.2. Reliability & Mean Time to System Failure (MTSF)

Let cdf of first transition time from the state S_i to the state in which failure occur is represented by $\Phi_i(t)$. We take absorbing state as the failed state. So, the expressions for $\Phi_i(t)$ from which MTSF of discussed system is obtained as

$$\Phi_i(t) = \sum_{i,j} Q_{ij}(t) \circledS \Phi_j(t) + \sum_{i,k} Q_{ik}(t) \quad (1)$$

Where i is the operating state from which transition takes place to j (operating & regenerative state) and k (failed state).
If we take LST of above relation (1) and solved them for $\Phi_0^{**}(s)$, we have

$$R^*(s) = \frac{1 - \Phi^{**}(s)}{s} \quad (2)$$

By taking Inverse Laplace transform of (2), we get system reliability.
And MTSF is obtained as: $\text{MTSF} = \lim_{s\to 0} \frac{1 - \Phi^{**}(s)}{s} = \frac{N}{D}$
Where, $N = \mu_0 + \mu_1 p_{01} + \mu_3 p_{01} p_{13} + \mu_5 p_{01} p_{15}$ and $D = 1 - p_{01}p_{13}p_{30} - p_{01}p_{15}p_{50}$

4.3. Analysis of Availability

Let $A_i(t)$ be the probability of the system availability at an instant 't' given that system goes to regenerative state S_i at t = 0. So the expressions for $A_i(t)$ as

$$A_i(t) = M_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © A_j(t) \quad (3)$$

Where i is regenerative state from which transition takes place to j (successive regenerative state) through n transitions.
$M_i(t)$ be the probability of the system in up state Si up to the time 't' without visiting to any other regenerative state.

$$M_0(t) = e^{-(2\lambda + \alpha_0)t},\ M_1(t) = e^{-(\lambda+\alpha_0)t}\overline{H(t)},$$

$$M_3(t) = e^{-(\lambda+\alpha_0)t}\overline{R(t)}, M_5(t) = e^{-(\lambda+\alpha_0)t}\overline{F(t)},\ M_7(t) = e^{-(\lambda t)}\overline{G(t)}$$

Now, if we use LT of (3) and solved it for $A_0^*(s)$.We get the result for steady state availability as

$$A_0(\infty) = \lim_{s\to 0} sA_0^*(s) = \frac{N_1}{D_1} \quad (4)$$

Where
$N_1 = \mu_0 X + (\mu_1 + \mu_3 p_{13} + \mu_5 p_{15})Y + \mu_7 Z$ and
$D_1 = (\mu_0 + \mu_2 p_{02})X + (\mu_1' + \mu_3' p_{13} + \mu_5' p_{15} + \mu_6 p_{16})Y + \mu_7' Z$

4.4. Busy Period Analysis for Server

Let $B_i^I(t), B_i^R(t), B_i^{Rp}(t), B_i^P(t)$be the probability of busy period of server during inspection, repair, replacement and PM at instant 't' with the given condition that the system go to regenerative state Si at t=0. The expressions for $B_i^I(t), B_i^R(t), B_i^{Rp}(t), B_i^P(t)$ are as follows:

$$B_i^I(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © B_j^I(t) \ , \quad B_i^R(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © B_j^R(t)$$
$$B_i^{Rp}(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © B_j^{Rp}(t) \ , B_i^P(t) = W_i(t) + \sum_{i,j} q_{ij}^{(n)}(t) © B_j^P(t) \tag{5}$$

Where i is regenerative state from which transition takes place to j (successive regenerative state) through n transitions.

Wi(t) be the probability of server busyness at state Si due to repair activities at time t without making any transition to any other regenerative state or returning to the same via one or more non regenerative state.

Here,

$$W_1(t) = e^{-(\lambda+\alpha_0)t}\overline{H(t)} + (\lambda e^{-(\lambda+\alpha_0)t} © 1)\overline{H(t)}$$
$$W_5(t) = e^{-(\lambda+\alpha_0)t}\overline{F(t)} + (\lambda e^{-(\lambda+\alpha_0)t} © 1)\overline{F(t)} + (\alpha_0 e^{-(\lambda+\alpha_0)t} © 1)\overline{F(t)}$$
$$W_3(t) = e^{-(\lambda+\alpha_0)t}\overline{R(t)} + (\lambda e^{-(\lambda+\alpha_0)t} © 1)\overline{R(t)} + (\alpha_0 e^{-(\lambda+\alpha_0)t} © 1)\overline{R(t)}$$
$$W_2(t) = \overline{G(t)} = W_6(t), W_7(t) = e^{-(\lambda)t}\overline{G(t)} + (\lambda e^{-(\lambda)t} © 1)\overline{G(t)}$$

Take LT of (5) and solving it for $B_0^{I^*}(s), B_0^{R^*}(s), B_0^{Rp^*}(s), B_0^{P^*}(s)$ .The busy time in inspection, repair, replacement and PM for server is given by

$$B_0^I(\infty) = \lim_{s\to 0} sB_0^{I^*}(s) = \frac{N_2}{D_1}, B_0^R(\infty) = \lim_{s\to 0} sB_0^{R^*}(s) = \frac{N_3}{D_1}$$
$$B_0^{Rp}(\infty) = \lim_{s\to 0} sB_0^{Rp^*}(s) = \frac{N_4}{D_1}, \ B_0^P(\infty) = \lim_{s\to 0} sB_0^{P^*}(s) = \frac{N_5}{D_1}$$

Here,

$N_2 = W_1^*(0)Y, N_3 = W_5^*(0)p_{15}Y, N_4 = W_3^*(0)p_{13}Y$ and

$N_5 = W_2^*(0)p_{02}X + W_6^*(0)p_{16}Y + W_7^*(0)Z$ and D1 is mentioned above.

4.5. Expected Number of Visits by The Server

Consider $I_i(t), R_i(t), Rp_i(t), Pm_i(t)$ as the expected number of visits make by the server for inspection, repair, replacement and PM in (0, t] .We have the following recursive relations for $I_i(t), R_i(t), Rp_i(t), Pm_i(t)$ are

$$I_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + I_j(t)) \ , \quad R_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + R_j(t))$$
$$Rp_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + Rp_j(t)) \ , \ Pm_i(t) = \sum_{i,j} Q_{ij}^{(n)}(t) Ⓢ (C + Pm_j(t)) \tag{6}$$

Where i is regenerative state from which transition takes place to j (successive regenerative state) through n transitions and and C = 1 server does the job afresh at j, otherwise C = 0.

Take LST of (6) and solving it for $I_0^{**}(s), R_0^{**}(s), Rp_0^{**}(s), Pm_0^{**}(s)$.The expected number of inspections, repairs, replacements and PM by the server is given by (per unit time)

$$I_0(\infty) = \lim_{s\to 0} sI_0^{**}(s) = \frac{N_6}{D_1} \ , \quad R_0(\infty) = \lim_{s\to 0} sR_0^{**}(s) = \frac{N_7}{D_1}$$
$$Rp_0(\infty) = \lim_{s\to 0} sRp_0^{**}(s) = \frac{N_8}{D_1} \ , Pm_0(\infty) = \lim_{s\to 0} sPm_0^{**}(s) = \frac{N_9}{D_1}$$

Where,

$N_6 = (1 - p_{16})Y, N_7 = (p_{11.49} + p_{15})Y, N_8 = (p_{11.4,10} + p_{13})Y,$

$N_9 = p_{02}X + p_{16}Y + Z$ and D_1 is already mentioned.

Here X, Y & Z are

$X = p_{13}p_{30} + p_{15}p_{50} + p_{70}(p_{13}p_{37.13} + p_{15}p_{57.12})$, $Y = (1 - p_{02}p_{70})$ and

$Z = p_{13} + p_{15} - p_{13}p_{31.8} - p_{15}p_{51.11} - p_{01}p_{13}p_{30} - p_{01}p_{15}p_{50}$

4.6. Profit Analysis

In steady state the profit function of the system model can be obtained as

$$P = k_0A_0 - k_1B_0^I - k_2B_0^R - k_3B_0^{Rp} - k_4B_0^{Pm} - k_5I_0 - k_6R_0 - k_7Rp_0 - k_8Pm_0$$

Here,

P = Profit function of system model

k_0 = Revenue per unit up − time of the system

k_1, k_2, k_3, k_4 = Cost per unit time of the server when it is busy in inspection, repair, replacement, preventive maintenance

k_5, k_6, k_7, k_8 = Cost per unit time for inspection, repair, replacement, preventive maintenance

5. Analytical Study of the Model

To make the study more practical we draw the results in the form of tables and graphs. Tables 1, 2, 3 and figures 2, 3, 4 show the behaviour of MTSF, availability and profit with respect to failure rate for different values of the given parameters by assuming that the rate of repair activities follow exponential distribution. Table 1 and figure 2 talks about the values of MTSF goes decreasing when failure rate increases. If we increase the values of repair rate (α=4.1), replacement rate (β=5) and inspection rate (γ=3) one by one and keep all other parameters fix we find that the MTSF is increases. But if we increase α_0=3.1 (rate by which system goes for PM) the MTSF is decreases.

Table 2 and figure 3 shows the effect of various parameters on availability with respect the failure rate and we observe that the availability of the system decreases with increase in the failure rate. Similarly if we increase α, β, γ and θ (rate by which system do preventive maintenance) then the availability is also increases but availability decreases when we increase in α_0.

From table 3 and figure 4 we conclude about the profit of the system. We see that if values of α, β, γ and θ increases then the profit is also increases but if we increase α_0 then the profit goes decline.

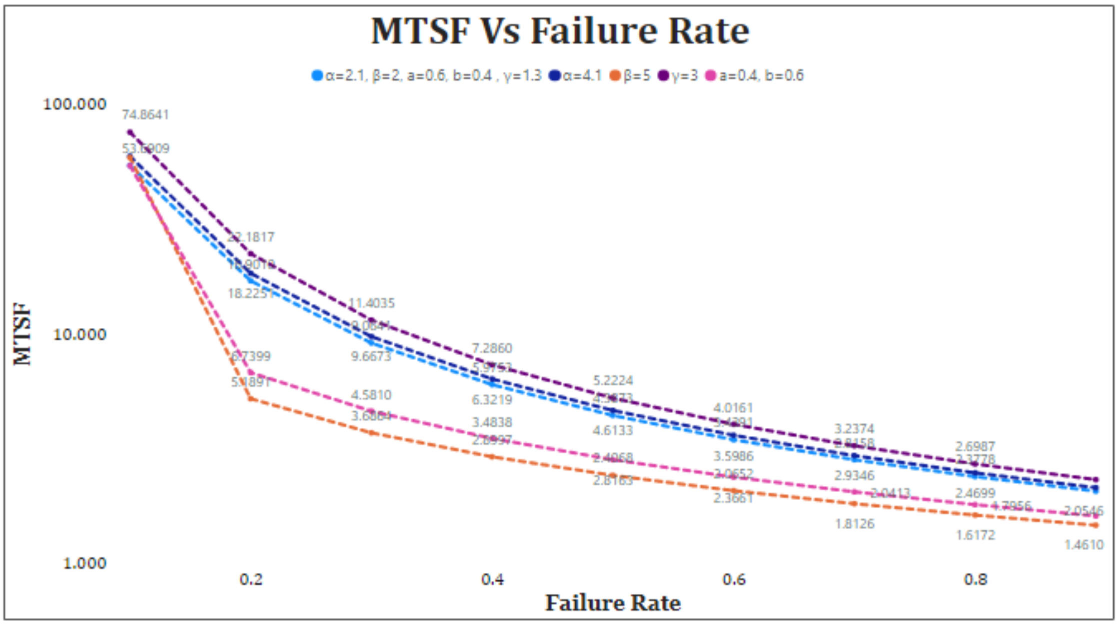


Figure 2: *MTSF VS Failure Rate*

Table 2: *MTSF w.r.t various parameters*

Failure rate	α=2.1,β=2,a=0.6,b=0.4,γ =1.3,α0=3	α= 4.1	β=5	γ= 3	α0= 3.1	a= 0.4, b= 0.6
0.1	0.33274	0.33275	0.33275	0.33279	0.32204	0.33274
0.2	0.33113	0.33119	0.33119	0.33130	0.32056	0.33113
0.3	0.32875	0.32887	0.32886	0.32908	0.31838	0.32875
0.4	0.32577	0.32596	0.32594	0.32627	0.31563	0.32577
0.5	0.32235	0.32260	0.32258	0.32302	0.31247	0.32234
0.6	0.31859	0.31890	0.31887	0.31943	0.30899	0.31858
0.7	0.31458	0.31495	0.31492	0.31558	0.30527	0.31457
0.8	0.31040	0.31082	0.31079	0.31153	0.30138	0.31039
0.9	0.30610	0.30657	0.30653	0.30736	0.29737	0.30609

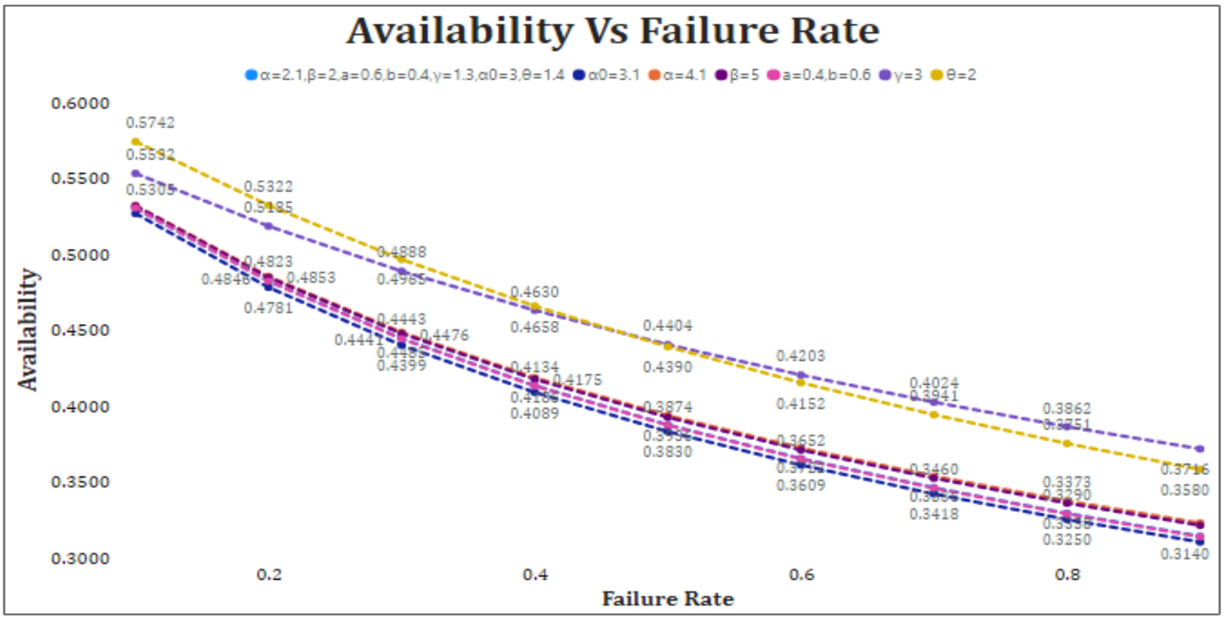


Figure 3: *Availability VS Failure Rate*

Table 3: *Availability w.r.t various parameters*

Failure Rate	α=2.1,β=2,a=0.6,b=0.4,γ=1.3,α0=3,θ=1.4	α=4.1	β=5	a=0.4,b=0.6	γ=3	α0=3.1	θ=2
0.1	0.53052	0.53208	0.53170	0.53043	0.55323	0.52682	0.57425
0.2	0.48233	0.48528	0.48462	0.48218	0.51847	0.47809	0.53218
0.3	0.44435	0.44852	0.44764	0.44414	0.48875	0.43990	0.49653
0.4	0.44435	0.41859	0.41754	0.41312	0.46298	0.40888	0.46580
0.5	0.38741	0.41859	0.39237	0.38714	0.44036	0.38300	0.43895
0.6	0.36524	0.37219	0.37087	0.36494	0.42031	0.36094	0.41524
0.7	0.34599	0.35362	0.35221	0.34567	0.40238	0.34183	0.39411
0.8	0.34599	0.33727	0.33578	0.32871	0.38622	0.32505	0.37514
0.9	0.34599	0.33727	0.32115	0.31363	0.37156	0.31014	0.35799

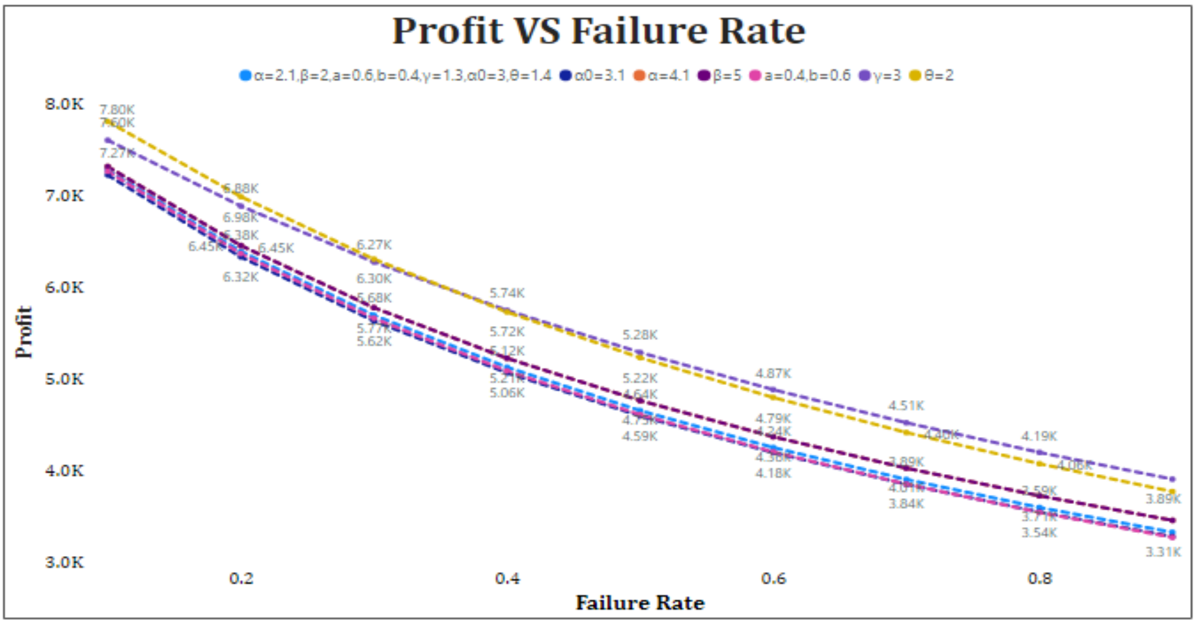


Figure 4: *Profit VS Failure Rate*

Table 4: *Profit w.r.t various parameters*

Failure Rate	α=2.1,β=2,a=0.6,b=0.4,γ=1.3,α0=3,θ=1.4	α=4.1	β=5	a=0.4,b=0.6	γ=3	α0=3.1	θ=2
0.1	7271.45	7307.88	7308.10	7256.59	7596.16	7216.20	7800.53
0.2	6382.77	6445.78	6446.11	6357.37	6877.03	6321.97	6980.11
0.3	5684.64	5767.44	5767.80	5651.40	6265.67	5622.79	6297.90
0.4	5117.12	5214.77	5215.12	5077.84	5738.43	5056.42	5719.43
0.5	4643.64	4752.40	4752.75	4599.59	5278.18	4585.18	5221.28
0.6	4240.52	4357.50	4357.87	4192.61	4872.23	4184.81	4786.85
0.7	3891.68	4014.63	4015.04	3840.60	4510.94	3838.93	4404.02
0.8	3585.81	3712.90	3713.39	3532.09	4186.87	3536.04	4063.65
0.9	3314.65	3444.43	3445.03	3258.70	3894.18	3267.80	3758.75

6. Practical Implication

Redundancy is a useful method of increasing reliability and optimizing the balance between operation effectiveness and expenditure. Arranging elements of the system in parallel provide alternative paths of operation. The parallel structure in the reliability engineering is widely used in many industrial systems such as power generation systems, pump systems, production systems and computing systems. One of the examples is in the commercial boiler market. Packaged boilers installed in multiple boiler cascade systems offer long-term energy savings and reliability. L. Vorsteveld [19] gives that the preferred control scheme is parallel cascading, and it leads to high turn down ratio.

Figure 5: *Parallel Boiler Cascade System*

References

[1] Gupta, P.P., and Agarwal, S.C. (1984). A parallel redundant complex system with two types of failure under preemptive-repeat repair discipline. *Microelectronics Reliability,* 24(3): 395-399.

[2] Dhillon, B.S. and Yang, N. (1992). Stochastic analysis of standby systems with common-cause failures and human errors. *Microelectronics Reliability,* 32(12):1699-1712.

[3] Ram, M. (2013). On System Reliability Approaches: A Brief Survey. *International Journal of System Assurance Engineering and Management,* 4(2):101-117.

[4] Hitomi, K., Nakajima, M. and Takahashi, N. (1979). Reliability analysis of parallel manufacturing systems with two machines. *Journal of Manufacturing Science and Engineering,* 101(3):250-255.

[5] Teramoto, K., Nakagawa, T. and Motoori, M. (1990). Optimal inspection policy for a parallel redundant system. *Microelectronics Reliability,* 30(1):151-155.

[6] Gupta, R., Kishan, R. and Kumar, P. (1999). A two-non-identical-unit parallel system with correlated lifetimes. *International Journal of Systems Science,* 30(10):1123-1129.

[7] Malik, S.C., Bhardwaj, R.K. and Grewal, A.S. (2010). Probabilistic analysis of a system of two nonidentical parallel units with priority to repair subject to inspection. *Journal of Reliability and Statistical Studies, 3(1):1-11*.

[8] Yu, X. and Khambadkone, A.M. (2012). Reliability analysis and cost optimization of parallel-inverter system. *IEEE Transaction on Industrial Electronics,* 59(10):3881-3889.

[9] Rathee, R. and Malik,S.C. (2013). A parallel system with priority to preventive maintenance subject to maximum operation and repair times. *American Journal of Mathematics and Statistics,* 3(6):436-444.

[10] Shivkumar, M. and Jayanth, M.K. (2014). Reliability analysis of link stability in secured routing protocols for MANETs. *Engineering Journal,* 18(1):65-76.

[11] Kakkar, M.K., Chitkara, A.K. and Bhatti, J. (2015). Reliability analysis of two-unit parallel repairable industrial system. *Decision Science Letters,* 4(4):525-536.

[12] Bhardwaj, N. and Parasher, B. (2017). Stochastic analysis of two identical cold standby units with geometric failure and repair rates. *Mathematical Journal of Interdisciplinary Sciences,* 5:131-142.

[13] Pundir, P.S., Patawa, R. and Gupta, P.K. (2018). Stochastic outlook of two non-identical unit parallel system with priority in repair. *Cogent Mathematics & Statistics,5(1):1-18.*

[14] Chopra, G. and Ram, M. (2019). Reliability measures of two dissimilar unit parallel system using gumbel-hougaard family copula. *International Journal of Mathematical, Engineering and Management Sciences,* 4(1):116–130.

[15] YangLing, L., GenQi, X. and Hao, C. (2021). Analysis of two components parallel repairable degenerate system with vacation. *AIMS Mathematics,* 10:10602-10619.

[16] Dabas, N. and Rathee, R. (2022). Parallel system reliability analysis with priority and inspection using semi-Markov approach. *Reliability: Theory and Applications,* 17(2):56-66.

[17] Dabas, N. and Rathee, R. (2022). Reliability analysis of parallel system using inspection for repair activities. *International Journal of Agricultural Statistical Sciences,* 18(2):899-906.

[18] El-Sherbeny, M.S. and Hussien, Z.M. (2022). Reliability analysis of a two nonidentical unit parallel system with optional vacations under poison shocks. *Mathematical Problems in Engineering.* https://doi.org/10.1155/2022/2488182.

[19] Lou,V. (2021). Higher turndown ratios, parallel cascading are keys to efficiency. https://www.pmmag.com/articles/103686,.

SOME USEFUL PATHWAY MODELS FOR RELIABILITY ANALYSIS

T. PRINCY

Department of Statistics
Cochin University of Science and Technology
Cochin-682022, Kerala, India
princyt@cusat.ac.in

Abstract

In this paper, we first discuss pathway model in general. Then a special case for the real scalar variable is considered. This special case is relevant in reliability problems. In the pathway model, an arbitrary function is introduced so that the hazard function resulting from this model is of a given shape such as a bathtub type hazard function. The model is also derived by using an entropy optimization procedure by introducing a new entropy measure. It is shown that a large number of densities in current use are connected to the pathway model. Certain combinations of pathway densities resulting in hazard functions of desired shapes, multi-component failure situation etc are examined from a reliability point of view. For further use of the proposed model, the unknown parameters are estimated using the method of maximum likelihood estimation. The behaviour of the reliability measure has been observed graphically for arbitrary values of the parameters related to the number of components and operating time.

Keywords: Pathway model, reliability analysis, hazard function, entropy optimization

1. Introduction

The common knowledge used in the literature is that an approximate model for the data at hand can be found, regardless of whether the data comes from the biological, physical, engineering, social, or other fields. The information at hand can be described in the area around the stable condition or along a path that leads there. In (2005) Mathai [1] introduced a pathway model to cover the stable as well as the transitional stages, which describes transitions of rectangular matrix-variate distributions in the real case. It is a mathematical or stochastic model that switches from one functional form to another through pathway parameter so that intermediate steps can be represented. Also the parameters of pathway model connect many families of functions, and as a result, it is possible to identify a suitable member either a given family or between stages of two families. The idea was extended to cover the complex rectangular matrix-variate case in Mathai and Provost [2]. The real scalar version of the pathway model can be stated as the following:

$$f_1(x) = c_1|x|^{\gamma}[1 - a(1-q)|x|^{\delta}]^{\frac{\eta}{1-q}} \tag{1}$$

for $a > 0, q < 1, \gamma > -1, \delta > 0, \eta > 0, 1 - a(1-q)|x|^{\delta} > 0$, and zero elsewhere, where c_1 can act as the normalizing constant if we wish to create a statistical density out of $f_1(x)$. The support of (1) is on $-[a(1-q)]^{-\frac{1}{\delta}} < x < [a(1-q)]^{-\frac{1}{\delta}}$. It is a finite range model. The functional part in a basic type-1 beta model is of the form $x^{\alpha-1}(1-x)^{\beta-1}, \alpha > 0, \beta > 0, 0 \leq x \leq 1$ and zero elsewhere, and hence the model in (1) can be looked upon as a generalized, extended and power-transformed type-1 beta model for $q < 1$. When $q \to 1$ the range of x will go from $-\infty$ to ∞. Note that (1) is a symmetric model for $x < 0$ and $x > 0$. If an asymmetric model is required then different weights can be used for $x < 0$ and $x > 0$ situations. These differing weights can be introduced either by multiplying with constants or through one of the parameters. For $q > 1$, write $1 - q = -(q-1), q > 1$, then (1) switches into the following model:

$$f_2(x) = c_2|x|^{\gamma}[1 + a(q-1)|x|^{\delta}]^{-\frac{\eta}{q-1}} \tag{2}$$

for $-\infty < x < \infty, a > 0, q > 1, \delta > 0, \gamma > -1, \eta > 0$. The model in (2) can be looked upon as a generalized, power-transformed and extended type-2 beta family of functions. The functional part of the basic type-2 beta model is $x^{\alpha-1}(1+x)^{-(\alpha+\beta)}, 0 \le x < \infty, \alpha > 0, \beta > 0$. Thus, (2) is the generalized and extended version of this basic type-2 beta function. When $q \to 1_-$ in (1) and $q \to 1_+$ in (2) the models in (1) and (2) go to the generalized and extended gamma family of functions as follows:

$$f_3(x) = c_3|x|^{\gamma}\mathrm{e}^{-a\eta|x|^{\delta}} \tag{3}$$

for $a > 0, \eta > 0, \delta > 0, \gamma > -1, -\infty < x < \infty$. This is the generalized gamma family of functions. Thus, the basic pathway model is (1), and (2) and (3) are available from (1). For $q < 1$ the family of functions is the generalized and extended type-1 beta family of functions. When q goes to 1, then one goes into the generalized and extended gamma family of functions. When q moves to $q > 1$ then we go into the generalized and extended type-2 beta family of functions. The parameter q enables us to go to three different families of functions and hence q is called the pathway parameter. The pathway idea in model building situation is to capture the stable situation as well as the unstable neighborhoods by the same model. If the gamma family is the stable situation in a physical problem then the paths leading to this stable situation through the generalized type-1 beta model and generalized type-2 beta model and the transitional stages are captured by the pathway parameter q. If (1) to (3) are to be treated as statistical densities then the following are the normalizing constants:

$$c_1 = \frac{\delta[a(1-q)]^{\frac{\gamma+1}{\delta}}\Gamma(\frac{\eta}{1-q} + 1 + \frac{\gamma+1}{\delta})}{2\Gamma(\frac{\gamma+1}{\delta})\Gamma(\frac{\eta}{1-q}+1)}, \text{ for } \gamma > -1, a > 0, q < 1, \eta > 0, \delta > 0 \tag{4}$$

$$c_2 = \frac{\delta[a(q-1)]^{\frac{\gamma+1}{\delta}}\Gamma(\frac{\eta}{q-1})}{2\Gamma(\frac{\gamma+1}{\delta})\Gamma(\frac{\eta}{q-1} - \frac{\gamma+1}{\delta})}, \text{ for } \eta > 0, \delta > 0, a > 0, q > 1, \gamma > -1, \frac{\eta}{q-1} - \frac{\gamma+1}{\delta} > 0 \tag{5}$$

$$c_3 = \frac{\delta(a\eta)^{\frac{\gamma+1}{\delta}}}{2\Gamma(\frac{\gamma+1}{\delta})}, a > 0, \eta > 0, \delta > 0, \gamma > -1. \tag{6}$$

Our interest is to consider a special case of the pathway model for $x \ge 0$ and when $\gamma = \delta - 1$. In this case the normalizing constants reduce to very simple forms. These forms of the pathway model are the most relevant in reliability analysis. Our main focus will be on this special case. Before we concentrate on the special case, let us see some other special cases of the general models. Note that (1) for $\gamma = 0, a = 1, \delta = 1, \eta = 1, q < 1, q > 1, q \to 1$ is Tsallis statistics in non-extensive statistical mechanics. It is claimed that between 1990 and 2010, over 3000 articles are published on this topic of Tsallis statistics. Details of the development may be seen from his book Tsallis [3]. Equation (2) for $q > 1, q \to 1, a = 1, \delta = 1, \eta = 1$ is superstatisitcs in statistical mechanics. Dozens of articles are also written in this area since 2003. The basic article in this area is Beck and Cohen [4].

If location and scale parameters are to be incorporated into the models in (1) to (3) then replace $|x|$ by $|\frac{x-\mu}{\sigma}|$ for some constant μ and some constant $\sigma > 0$ in (1) to (3). Note that (3) for $\gamma = 0, \delta = 2$ is the normal or Gaussian density. For $x > 0$, (3) produces the generalized gamma density, Weibull density, gamma density, chisquare density, exponential density, Rayleigh density, Maxwell-Boltzmann density etc. Exponentiation in (2), that is put $x = \mathrm{e}^{-cy}, c > 0$, produces the generalized logistic density of Mathai and Provost [5], logistic density and other special cases of the generalized logistic density, which are relevant in reliability analysis. (2) can produce Cauchy density, Student-t density, F-density etc. For $\gamma = \delta - 1$, (1) and (2) produce many models in reliability analysis. As a limiting form of (2) one can obtain Fermi-Dirac density from (2) and Bose-Einstein density from (1), after exponentiation.

2. Construction of the Pathway Model through Entropy Optimization

Model building in physical situations is often done by optimizing an appropriate entropy measure and then deriving the model from therein. Consider Mathai's entropy, for an earlier version see Mathai and Haubold [6],

$$M_q(f) = \frac{\int_X [f(X)]^{\frac{1-q+\eta}{\eta}} \mathrm{d}X - 1}{q-1}, q \neq 1, q < 2 \tag{7}$$

where $f(X)$ is a statistical density where X could be real or complex scalar or matrix variable, and $\int_X$ denotes the integral over the support of f. A corresponding version for the discrete situation can be constructed. One can look upon (7) as an expected value of $f^{\frac{1-q}{\eta}}$ which then corresponds to Kerridge's measure of inaccuracy for $\eta = 1$, see Mathai and Rathie [7]. For the real scalar variable case, (7) for $\eta = 1$ can be looked upon as a modified Havrda-Charvat entropy, see Mathai and Rathie [7]. In the following discussion we consider the real scalar variable case first. Let us optimize (7) subject to the following moment-type restrictions for real scalar variables:

$$\int_x x^{\gamma \frac{(1-q)}{\eta}} f(x)\mathrm{d}x = \text{ fixed } \& \int_x x^{\gamma \frac{(1-q)}{\eta} + \delta} f(x)\mathrm{d}x = \text{ fixed} \tag{8}$$

for $\gamma > -1, q < 1, \delta > 0$, and it is assumed that the integrals in (8) exist. Note that for $\gamma = 0$, (8) states that the total integral is unity and that the first moment is given. This is equivalent to the physical law of conservation of energy when we consider energy distribution. It is convenient to use calculus of variation for optimizing (7) subject to the conditions in (8). Then the Euler equation is the following:

$$\frac{\partial}{\partial f}[f^{\frac{1-q+\eta}{\eta}} - \lambda_1 x^{\gamma \frac{(1-q)}{\eta}} f + \lambda_2 x^{\gamma \frac{(1-q)}{\eta} + \delta} f] = 0 \tag{9}$$

where λ_1 and λ_2 are Lagrangian multipliers. Then (9) gives

$$f^{\frac{(1-q)}{\eta}} = \mu_1 x^{\gamma \frac{(1-q)}{\eta}} [1 - \mu_2 x^\delta]$$

for some constants μ_1 and μ_2, which then gives

$$f_1 = \nu\, x^\gamma [1 - \mu_2 x^\delta]^{\frac{\eta}{1-q}} \tag{10}$$

for some ν and μ_2. Take $\mu_2 = a(1-q)$ and ν as the normalizing constant to obtain the model (1). In (7) if X is a $p \times 1$ vector random variable and if (8) is replaced by the following conditions

$$\int_X [(X-\mu)' V^{-1} (X-\mu)]^{\gamma \frac{(q-1)}{\eta}} f(X)\mathrm{d}X = \text{ fixed} \tag{11}$$

and

$$\int_X [(X-\mu)' V^{-1} (X-\mu)]^{\gamma \frac{(q-1)}{\eta} + \delta} f(X)\mathrm{d}X = \text{ fixed} \tag{12}$$

where V is $p \times p$ real symmetric and positive definite constant matrix, μ is a $p \times 1$ constant vector, a prime denotes transpose, $\gamma > 0, q > 1, \eta > 0$, then from (2.3) and (10) we have the following density:

$$f_4(X) = c_4[(X-\mu)' V^{-1} (X-\mu)]^\gamma [1 + a(q-1)\{(X-\mu)' V^{-1} (X-\mu)\}^\delta]^{-\frac{\eta}{q-1}}, q > 1. \tag{13}$$

Then, when $q \to 1$, $f_4(X)$ goes to $f_5(X)$ given by

$$f_5(X) = c_5[(X-\mu)' V^{-1} (X-\mu)]^\gamma \mathrm{e}^{-a\eta[(X-\mu)' V^{-1} (X-\mu)]^\delta}. \tag{14}$$

Note that (13) and (14) are also associated with type-2 and gamma distributed random points in Euclidean n-space, $p \leq n$, see Mathai [8]. Also, (14) for $\gamma = 0, \delta = 1$ is the p-variate Gaussian density with mean value vector μ and covariance matrix V. The quantity

$$(X-\mu)'V^{-1}(X-\mu) = c > 0 \tag{15}$$

is known as the ellipsoid of concentration. Hence the constraints can be explained in terms of ellipsoid of concentration. In (7) if X is a $p \times q$, $q \geq p$ rectangular matrix-variate random variable and if the conditions in (8) are replaced by the following:

$$\int_X [\text{tr}(AXBX')]^{\gamma \frac{(q-1)}{\eta}} f(X)\text{d}X = \text{ fixed} \tag{16}$$

and

$$\int_X [\text{tr}(AXBX')]^{\gamma \frac{(q-1)}{\eta}+\delta} f(X)\text{d}X = \text{ fixed} \tag{17}$$

where A is $p \times p$ and B is $q \times q$ constant real positive definite matrices, $q > 1, \eta > 0, \delta > 0$ then the steps in (9) and (10) give the density

$$f_6(X) = c_6[\text{tr}(AXBX')]^{\gamma}[1 + a(q-1)\{\text{tr}(AXBX')\}^{\delta}]^{-\frac{\eta}{q-1}}, q > 1 \tag{18}$$

which when $q \to 1$ gives

$$f_7(X) = c_7[\text{tr}(AXBX')]^{\gamma}\text{e}^{-a\eta[\text{tr}(AXBX')]^{\delta}} \tag{19}$$

for $a > 0, \eta > 0, \delta > 0, \gamma > -1$. Note that (19) for $\gamma = 0, \delta = 1$ is the real matrix-variate Gaussian density. By replacing X by $X - M$, where M is a $p \times q$ constant matrix, one can also incorporate a location parameter matrix in the models in (18) and (19). Mathai and Princy [9] illustrate the significance of models (18) and (19) to the real multivariate reliability analysis.

3. A Special Case of Pathway Model and Reliability Analysis

A special case of the pathway model (1),(2),(3) is the case where $\gamma = \delta - 1$, for $x \geq 0$. This will then correspond to a power-transformed basic model. Take the basic model in (1) as the one with $\gamma = 0$ and $\delta = 1$ for $x \geq 0$ or $c[1 - a(1-q)x]^{\frac{\eta}{1-q}}$ where $a > 0, q < 1$ and c is a constant. Make the power transformation here or put $x = y^{\delta}$ for some $\delta > 0$. Then we have

$$g_1(y) = C_1 y^{\delta-1}[1 - a(1-q)y^{\delta}]^{\frac{\eta}{1-q}}, q < 1, \eta > 0, a > 0, \delta > 0 \tag{20}$$

and C_1 is a constant. The cases for $q > 1, q \to 1$ are available from (20) as

$$g_2(y) = C_2 y^{\delta-1}[1 + a(q-1)y^{\delta}]^{-\frac{\eta}{q-1}}, q > 1 \tag{21}$$

and

$$g_3(y) = C_3 y^{\delta-1}\text{e}^{-a\eta y^{\delta}}, a > 0, \eta > 0, \delta > 0. \tag{22}$$

If g_1, g_2, g_3 are to be treated as statistical densities then the normalizing constants are the following:

$$C_1 = \delta a(\eta + 1 - q), q < 1; C_2 = \delta a(\eta + 1 - q), q > 1; C_3 = \delta\eta a. \tag{23}$$

The following are the graphs showing the relative positions of g_1, g_2, g_3.

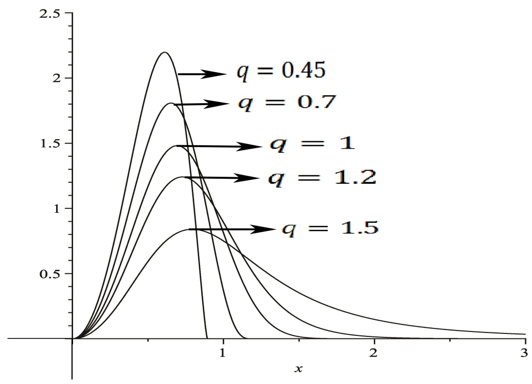


Figure 1: Pathway models for $\gamma = \delta - 1$

Note that in $g_2(y)$ and $g_3(y)$ the densities of $\frac{1}{y}$ also belong to the same families. Put $x = \frac{1}{y}$ then $g_2(x)$ and $g_3(x)$ go to the following for $0 \le x < \infty$:

$$g_{2*}(x) = a\delta(\eta + 1 - q)x^{-\delta-1}[1 + a(q-1)x^{-\delta}]^{\frac{\eta}{1-q}} \tag{24}$$

and

$$g_{3*}(x) = a\delta(\eta + 1 - q)x^{-\delta-1}e^{-a\eta x^{-\delta}}. \tag{25}$$

Thus, in $g_2(y)$ and $g_3(y)$ both δ and $-\delta$ with $\delta > 0$ are admissible with the corresponding change in $y^{\delta-1}$. Let us compute the survival functions. For $q < 1$

$$\begin{aligned} S_1(t) = Pr\{x \ge t\} &= \int_{x=t}^{[a(1-q)]^{-\frac{1}{\delta}}} a\delta(\eta + 1 - q)x^{\delta-1}[1 - a(1-q)x^{\delta}]^{\frac{\eta}{1-q}}dx \\ &= [1 - a(1-q)t^{\delta}]^{\frac{\eta}{1-q}+1}, q < 1, \eta > 0, a > 0, \delta > 0. \end{aligned} \tag{26}$$

For $q > 1$ the survival function $S_2(t)$ is given by the following:

$$S_2(x) = [1 + a(q-1)t^{\delta}]^{-\frac{\eta}{q-1}+1}, q > 1, \eta > 0, a > 0, \delta > 0 \tag{27}$$

and for $q \to 1$

$$S_3(t) = Pr\{x \ge t\} = e^{-a\eta t^{\delta}}, a > 0, \eta > 0, \delta > 0. \tag{28}$$

The hazard functions for $q < 1, q > 1, q \to 1$, denoted by $h_1(t), h_2(t), h_3(t)$ are the following:

$$h_1(t) = \frac{g_1(t)}{S_1(t)} = \frac{\delta a(\eta + 1 - q)t^{\delta-1}}{1 - a(1-q)t^{\delta}}, q < 1 \tag{29}$$

$$h_2(t) = \frac{\delta a(\eta + 1 - q)t^{\delta-1}}{1 + a(q-1)t^{\delta}}, q > 1 \tag{30}$$

$$h_3(t) = a\delta\eta t^{\delta-1}. \tag{31}$$

Here (29) to (31) do not show any interesting shapes. A useful shape is the bathtub shape. All different shapes are available from (29) to (31).

4. Exponentiation of the Pathway Model in the Special Case

In this paper our main objective is to construct some useful bathtub shaped models for reliability analysis. For this, let us make the transformation $y = e^{cx}, c > 0$ in (20). Then the condition

$$1 - a(1-q)y^{\delta} > 0 \Rightarrow 1 - a(1-q)e^{c\delta x} > 0 \text{ or } x < \frac{-1}{c\delta}\ln[a(1-q)]$$

for $q < 1$. But for $q > 1$ and $q \to 1$, $-\infty < x < \infty$ under exponentiation. Then the models reduce to the following:

$$g_4(x) = ca\delta(\eta + 1 - q)\mathrm{e}^{c\delta x}[1 - a(1-q)\mathrm{e}^{c\delta x}]^{\frac{\eta}{1-q}}, q < 1 \tag{32}$$

for $c > 0, a > 0, \delta > 0, \eta > 0, \eta + 1 - q > 0, -\infty < x < \frac{-1}{c\delta}\ln[a(1-q)]$ and zero elsewhere.

$$g_5(x) = ca\delta(\eta + 1 - q)\mathrm{e}^{c\delta x}[1 + a(q-1)\mathrm{e}^{c\delta x}]^{-\frac{\eta}{q-1}}, q > 1 \tag{33}$$

for $a > 0, \eta > 0, \delta > 0, c > 0, \eta + 1 - q > 0, -\infty < x < \infty$.

$$g_6(x) = ca\delta\eta\mathrm{e}^{c\delta x}\mathrm{e}^{-a\eta\mathrm{e}^{c\delta x}} \tag{34}$$

for $a > 0, \eta > 0, \delta > 0, c > 0, -\infty < x < \infty$. The hazard functions for (32) to (34) can be seen to the the following:

$$h_5(t) = \frac{ac\delta(\eta + 1 - q)\mathrm{e}^{c\delta t}}{[1 - a(1-q)\mathrm{e}^{c\delta t}]}, q < 1. \tag{35}$$

$$h_6(t) = \frac{ac\delta(\eta + 1 - q)\mathrm{e}^{c\delta t}}{[1 + a(q-1)\mathrm{e}^{c\delta t}]}, q > 1. \tag{36}$$

$$h_7(t) = ac\delta\eta\mathrm{e}^{c\delta t}, a > 0, c > 0, \delta > 0, \eta > 0. \tag{37}$$

If we make the transformation $y = \mathrm{e}^{-cx}, c > 0$ in (20), we can connect the transformed models to Bose-Einstein density, Logistic and Fermi-Dirac densities. The corresponding transformed models are

$$g_7(x) = ca\delta(\eta + 1 - q)\mathrm{e}^{-c\delta x}[1 - a(1-q)\mathrm{e}^{-c\delta x}]^{\frac{\eta}{1-q}}, q < 1 \tag{38}$$

for $c > 0, a > 0, \delta > 0, \eta > 0, \eta + 1 - q > 0, \frac{1}{c\delta}\ln[a(1-q)] < x < \infty$ and zero elsewhere.

$$g_8(x) = ca\delta(\eta + 1 - q)\mathrm{e}^{-c\delta x}[1 + a(q-1)\mathrm{e}^{-c\delta x}]^{-\frac{\eta}{q-1}}, q > 1 \tag{39}$$

for $a > 0, \eta > 0, \delta > 0, c > 0, \eta + 1 - q > 0, -\infty < x < \infty$.

$$g_9(x) = ca\delta\eta\mathrm{e}^{-c\delta x}\mathrm{e}^{-a\eta\mathrm{e}^{-c\delta x}} \tag{40}$$

for $a > 0, \eta > 0, \delta > 0, c > 0, -\infty < x < \infty$.

4.1. Bose-Einstein density

Consider the limiting case $\eta + 1 = q$. This is not admissible in the density (38). For $\eta + 1 = q$ can we re-normalize the function for $0 \le x < \infty$ and create a density out of it? That is, can

$$g_{10}(x) = c_{10}[\frac{1}{a(1-q)}\mathrm{e}^{c\delta x} - 1]^{-1} = C_7[\mathrm{e}^{\alpha + c\delta x} - 1]^{-1}, 0 \le x < \infty \tag{41}$$

be a density where $\frac{1}{a(1-q)} = \mathrm{e}^{\alpha}$?. Let us consider the integral

$$\begin{aligned}\int_0^\infty [\mathrm{e}^{\alpha + c\delta x} - 1]^{-1}\mathrm{d}x &= \frac{1}{\xi}\int_\alpha^\infty [\mathrm{e}^u - 1]^{-1}\mathrm{d}u, \xi = c\delta, \quad \text{put } \nu = \mathrm{e}^u \\ &= \frac{1}{\xi}\int_{\mathrm{e}^\alpha}^\infty \frac{1}{v}\frac{1}{v-1}\mathrm{d}v = \frac{1}{\xi}\int_{\mathrm{e}^\alpha}^\infty [\frac{1}{v-1} - \frac{1}{v}]\mathrm{d}v \\ &= \frac{1}{\xi}\ln\left(\frac{\mathrm{e}^\alpha}{\mathrm{e}^\alpha - 1}\right), \mathrm{e}^\alpha \neq 1.\end{aligned}$$

Hence for $c_{10} = \xi[\ln\left(\frac{\mathrm{e}^\alpha}{\mathrm{e}^\alpha - 1}\right)]^{-1}$ is a density for $0 \le x < \infty$. This density in (41) is the Bose-Einstein density in Physics.

4.2. Logistic and Fermi-Dirac densities

Consider (33) and (39) for $q = \frac{3}{2}$ and $\eta = 1$ and let $a = 2$. Then (33) and (39) become

$$g_{11}(x) = c\delta \frac{e^{-c\delta x}}{(1 + e^{-c\delta x})^2} = c\delta \frac{e^{c\delta x}}{(1 + e^{c\delta x})^2}, -\infty < x < \infty. \tag{42}$$

This is the logistic density. Hence (33) is a generalized form of the logistic density. Still more general forms can be obtained by deleting the condition $\gamma = \delta - 1$ in (2) and then exponentiating. Consider again the model (39). This can be written as follows:

$$\begin{aligned} g_8(y) &= ac\delta(\eta + 1 - q)e^{-c\delta x}[1 + a(q-1)e^{-c\delta x}]^{-\frac{\eta}{q-1}} \\ &= ac\delta(\eta + 1 - q)[a(q-1)]^{-\frac{\eta}{q-1}}[e^{-c\delta x}]^{-\frac{(\eta+1-q)}{q-1}} \qquad (43) \\ &\times [1 + \frac{1}{a(q-1)}e^{c\delta x}]^{-\frac{\eta}{q-1}}. \qquad (44) \end{aligned}$$

Note that $\frac{\eta}{q-1} = 1$ or $\eta + 1 - q = 0$ is not an admissible value in (44). This is the limiting situation. Can we re-normalize the function in (44) for $\eta + 1 = q$ and $0 \le y < \infty$ and create a density from (44)? Consider

$$g_{12}(x) = c_{12}[1 + e^{\alpha + c\delta x}]^{-1}, \frac{1}{a(q-1)} = e^{\alpha}. \tag{45}$$

Consider

$$\begin{aligned} \int_0^\infty g_{12}(x)dx &= c_{12}\int_0^\infty [1 + e^{\alpha + c\delta x}]^{-1}dx = \frac{c_{12}}{c\delta}\int_\alpha^\infty (1 + e^u)^{-1}du, \ \text{ put } v = e^u \\ &= \frac{c_{12}}{c\delta}\int_{e^\alpha}^\infty [\frac{1}{v}\frac{1}{1+v}]dv = \frac{c_{12}}{c\delta}\ln(\frac{e^\alpha}{1+e^\alpha}). \end{aligned}$$

Hence for $c_{12} = \frac{c\delta}{\ln(\frac{e^\alpha}{1+e^\alpha})}$, $g_{12}(x)$ is a density for $0 \le x < \infty$ and this density is known as Fermi-Dirac density in Physics.

4.3. Arbitrary function $P(x)$

Instead of power transformations and exponentiation let us take an arbitrary function $P(x)$ in our basic pathway models in (20) to (22). Then the models become the following:

$$g_{13}(x) = a\delta(\eta + 1 - q)P'(x)[P(x)]^{\delta - 1}[1 - a(1-q)(P(x))^\delta]^{\frac{\eta}{1-q}} \tag{46}$$

for $a > 0, \delta > 0, \eta > 0, q < 1, P'(x) = \frac{d}{dx}P(x) > 0, P(x) > 0$.

$$g_{14}(x) = a\delta(\eta + 1 - q)P'(x)[P(x)]^{\delta - 1}[1 + a(q-1)(P(x))^\delta]^{-\frac{\eta}{q-1}} \tag{47}$$

for $a > 0, \delta > 0, \eta > 0, q > 1, P'(x) > 0, P(x) > 0$.

$$g_{15}(x) = a\delta\eta P'(x)[P(x)]^{\delta-1}e^{-a\eta[P(x)]^\delta} \tag{48}$$

for $a > 0, \eta > 0, \delta > 0, P'(x) > 0, P(x) > 0$. Thus, $P(x)$ must be a positive and increasing function so that $P(x)$ and $P'(x)$ are positive. One such function is the distribution function for an arbitrary density. Let y be a real continuous random variable with density $g(y)$ and distribution function $F_y(x) = Pr\{y \le x\}$. Then take $P(x) = F_y(x)$ so that $P'(x) = g(x)$ is the density which is positive on the support of $g(x)$. Observe that when $F_y(x)$ is a distribution function then $[F_y(x)]^\delta, \delta > 0$ is again a distribution function for some other random variable. Our problem is the following: Can we select a function $P(x)$ so that the hazard function coming out of the pathway model is of the desired shape? From (29) and (30) note that the hazard function is of the following structure:

$$h_1(x) = -\frac{(\eta + 1 - q)}{1 - q}\frac{\partial}{\partial x}\ln[1 - a(1-q)(P(x))^\delta], q < 1 \tag{49}$$

for $1-a(1-q)[P(x)]^\delta > 0, q < 1, a > 0, \eta > 0, \delta > 0$, and

$$h_2(x) = \frac{(\eta+1-q)}{q-1}\frac{\partial}{\partial x}\ln[1+a(q-1)[P(x)]^\delta], q > 1, \tag{50}$$

for $a > 0, \delta > 0, \eta > 0$. Then for a pre selected hazard function of the desired shape one can solve for (49) and (50) and compute the corresponding $P(x)$. This is the aim. Suppose that the hazard function $h_1(x)$ is of the form

$$h_1(x) = \frac{1}{x-a+\epsilon_1} + \frac{1}{b-x+\epsilon_2}, \epsilon_1 > 0, \epsilon_2 > 0, a \le x \le b$$

and zero elsewhere. When $x = a$ one has $\frac{1}{\epsilon_1} + \frac{1}{(b-a)+\epsilon_2}$ and when $x = b$ it is $\frac{1}{\epsilon_2} + \frac{1}{(b-a)+\epsilon_1}$. It is bathtub shaped, take $b-a$ large. Then

$$\int_x h_1(x)\mathrm{d}x = \ln(x-a+\epsilon_1) - \ln(b-x+\epsilon_2) = \ln\left(\frac{x-a+\epsilon_1}{b-x+\epsilon_2}\right).$$

That is, for example, (49) yields

$$\int_x h_1(x)\mathrm{d}x = -\frac{(\eta+1-q)}{1-q}\ln[1-a(1-q)[P(x)]^\delta] = \ln\left(\frac{x-a+\epsilon_1}{b-x+\epsilon_2}\right).$$

That is,

$$\ln[1-a(1-q)[P(x)]^\delta] = \ln\left[\frac{b-x+\epsilon_2}{x-a+\epsilon_1}\right]^{\frac{1-q}{\eta+1-q}}$$

One solution is

$$[P(x)]^\delta = \frac{1}{a(1-q)}\left[1-\left[\frac{b-x+\epsilon_2}{x-a+\epsilon_1}\right]^{\frac{1-q}{\eta+1-q}}\right].$$

But when we impose the conditions $P(x) > 0, P'(x) > 0$ the bathtub shape cannot be maintained. The conditions $P(x) > 0, P'(x) > 0$ can be met only by taking $\epsilon_1 = \epsilon_2 + (b-a)$ or by shifting the minimum point to the left-end. Then we get a semi-bathtub shaped hazard function of the following form:

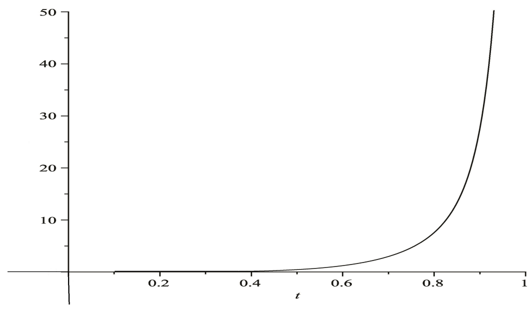


Figure 2: Semi-bathtub shaped hazard function

Consider the case of the arbitrary function $P(x)$ being the distribution function for some random variable. Let $P(x) = F_y(x)$ for a distribution function $F_y(t) = Pr\{y \le t\}$ for some continuous random variable y. Then $P'(t) = g_y(t)$ where $g_y(t)$ is the density of some random variable y, evaluated at t. Let $P(t) = [F_y(t)]$. Then the hazard function of (46)

$$h_{13}(t) = \frac{P'(t)}{Pr\{x \ge t\}} = \frac{a\delta(\eta+1-q)g_y(t)[F_y(t)]^{\delta-1}}{1-a(1-q)[F_y(t)]^\delta}.$$

Take $\delta = 1$. Then

$$h_{13}(t) = \frac{a(\eta+1-q)g_y(t)}{1-a(1-q)F_y(t)}, 0 \le F_y(t) \le 1.$$

Take a fast decreasing $g_y(t)$ with $g_y(0) \neq 0$ then such a $g_y(t)$ should produce a bathtub shaped hazard function curve. We shall examine a few such cases here. **Case (1):** Let

$$g_y(t) = \theta e^{-\theta t}, t \geq 0, \theta > 0$$

and zero elsewhere. In this case the hazard function

$$h_{14}(t) = \frac{a\theta(\eta + 1 - q)e^{-\theta t}}{(1 - a(1-q))[1 - e^{-\theta t}]}$$

for $a(1-q) < 1, > 0, q < 1, \eta > 0, \eta + 1 - q > 0$. Some of the plots are given below for the various values of the parameters of $h_{14}(t)$

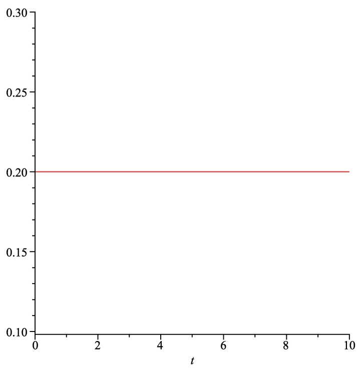


Figure 3: $a = 2, \theta = \frac{1}{10}, \eta = \frac{1}{2}, q = \frac{1}{2}$

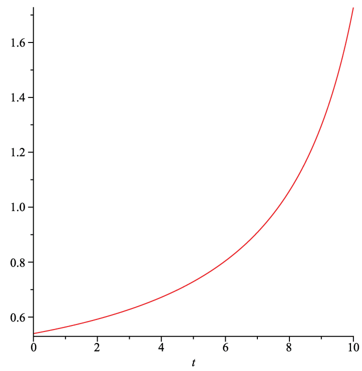


Figure 4: $a = 2, \theta = \frac{1}{10}, \eta = 2, q = \frac{3}{10}$

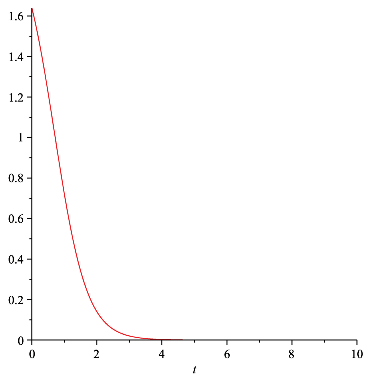


Figure 5: $a = 2, \theta = 2, \eta = \frac{1}{100}, q = \frac{6}{10}$

Case (2): Consider a type-1 beta model with the density $g_y(t) = \beta(b-t)^{\beta-1}, 0 \leq t \leq b, \beta > 0$. Then the distribution function is $F_y(t) = b^\beta - (b-t)^\beta$. Again, taking the pathway model for $q < 1$ and $\gamma = \delta - 1$ in (1.1), with $\delta = 1$, produces the hazard function

$$h_{15}(t) = \frac{a(\eta + 1 - q)\beta(b-t)^{\beta-1}}{1 - a(1-q)[b^\beta - (b-t)^\beta]}, 0 \leq t \leq b.$$

When $t = 0$, $h(0) = a(\eta + 1 - q)\beta b^{\beta-1}$. When t is nearing b the numerator nears zero and the denominator nears $1 - a(1-q)$. Observe that $a(1-q) < 1$. Select a and q such that $a(1-q)$ nears 1 then $h_{15}(t)$ will be a large quantity. This is plotted for various values of the parameters a, q, β, b.

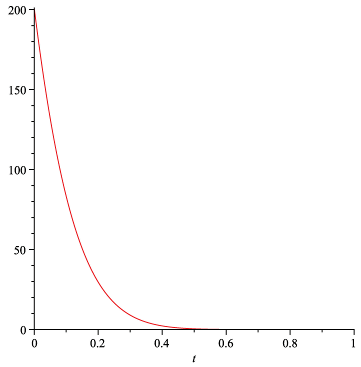


Figure 6: $a = 1, \beta = 10, b = 1, \eta = 20, q = \frac{9}{10}$

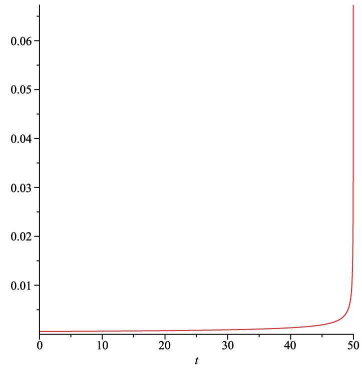


Figure 7: $a = 0.0002, \beta = \frac{2}{10}, b = 50, \eta = 5, q = \frac{9}{10}$

Case (3): Consider a Pareto type density for $g_y(t)$. Let

$$g_y(t) = C\alpha(1+t)^{-(\alpha+1)}, 0 \leq t \leq b, \alpha > 0, C = \frac{(1+b)^\alpha}{(1+b)^\alpha - 1}.$$

When $b = \infty$ then $C = 1$ and in this case $g_y(t)$ is a type-2 beta density. The hazard function in this case of pathway model of (1.1) for $q < 1$ with $\gamma = \delta - 1$ and $\delta = 1$ is as follows:

$$h_{16}(t) = \frac{a(\eta + 1 - q)C\alpha(1+t)^{-(\alpha+1)}}{1 - a(1-q)C[1 - (1+t)^{-\alpha}]}$$

for $\alpha > 0, q < 1, a > 0, 0 \leq t \leq b$. Note that when $t = 0$ the denominator is 1 and the numerator is $a(\eta + 1 - q)C\alpha$ and when t is large then the numerator is very small and the denominator nears $1 - a(1-q)$. Selecting a and q such that $a(1-q) < 1$ but close to 1 we can make a bathtub shaped curve. The following are the curves for some selected values of the parameters.

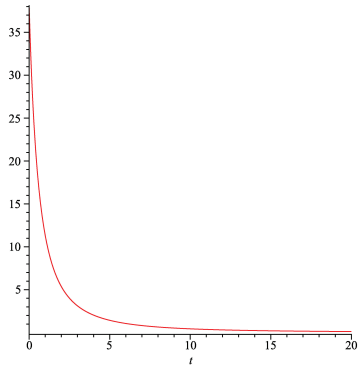


Figure 8: $a = 3, \alpha = 1, \eta = 12, b = 50, q = \frac{9}{10}$

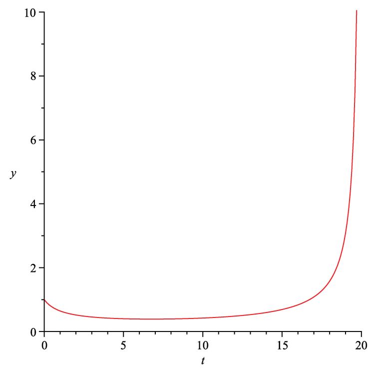


Figure 9: $a = 2, \alpha = \frac{1}{100}, \eta = 1, b = 20, q = \frac{5}{10}$

Hence this approach does not produce a satisfactory hazard function.

5. Combinations of Pathway Models with Other Models

Consider an exponentiated pathway model of (38) for $c = 1$, of the form

$$f_{10}(x) = a\delta(\eta + 1 - q)e^{-\delta x}[1 + a(q-1)e^{-\delta x}]^{-\frac{\eta}{q-1}}, q > 1, \tag{51}$$

for $\eta > 0, \delta > 0, a > 0$. Take another function of the power function type.

$$f_{11} = \frac{\gamma}{e^{b\gamma} - 1}e^{\gamma x}, 0 \leq x \leq b, \gamma > 0 \tag{52}$$

and zero elsewhere. Let $f(x)$ be a convex combination of $f_{10}(x)$ and $f_{11}(x)$. Let

$$f(x) = \frac{a_1}{a_1+a_2} f_{10}(x) + \frac{a_2}{a_1+a_2} f_{11}(x), a_1 > 0, a_2 > 0.$$

Then the survival function $S(t)$ is the following:

$$\begin{aligned} S(t) = Pr\{x \geq t\} &= \frac{a_1}{a_1+a_2}\int_t^{\infty} f_{10}(x)\mathrm{d}x + \frac{a_2}{a_1+a_2}\int_t^{b} f_{11}(x)\mathrm{d}x \\ &= \frac{a_1}{a_1+a_2}[1+a(q-1)\mathrm{e}^{-\delta t}]^{-\frac{\eta}{q-1}+1} + \frac{a_2}{a_1+a_2}\frac{1}{\mathrm{e}^{\gamma b}-1}[\mathrm{e}^{b\gamma} - \mathrm{e}^{\gamma t}]. \end{aligned}$$

Therefore the hazard function is the following:

$$h(x) = \frac{f(x)}{S(x)} = \frac{a_1 a\delta(\eta+1-q)\mathrm{e}^{-\delta x}[1+a(q-1)\mathrm{e}^{-\delta x}]^{-\frac{\eta}{q-1}} + a_2\frac{\gamma}{\mathrm{e}^{b\gamma}-1}\mathrm{e}^{\gamma x}}{a_1[1+a(q-1)\mathrm{e}^{-\delta x}]^{-\frac{\eta}{q-1}+1} + a_2\frac{1}{\mathrm{e}^{\gamma b}-1}[\mathrm{e}^{b\gamma} - \mathrm{e}^{\gamma x}]}, 0 \leq x \leq b. \quad (53)$$

Note that $a_1 + a_2$ will be canceled. Hence we may take any positive linear combinations of $f_{10}(x)$ and $f_{11}(x)$ to get the numerator of $h(x)$ and the same linear combination to get the denominator. This (53) is plotted for various parameter values. The graphs are for the following combinations of the parameters:

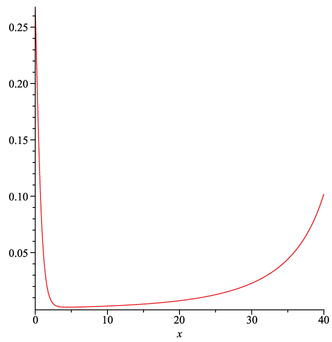


Figure 10

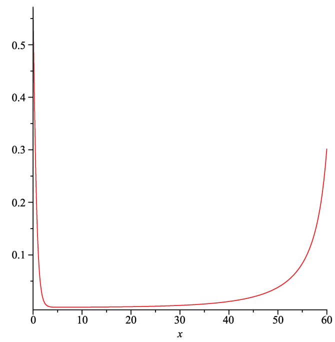


Figure 11

Figure 10: $a_1 = 1, a_2 = 1, b = 40, a = 1, \delta = 2, q = 1.5, \eta = 1, \gamma = 0.1$
Figure 11: $a_1 = 9, a_2 = 1, b = 40, a = 1, \delta = 2, q = 1.5, \eta = 1, \gamma = 0.1$

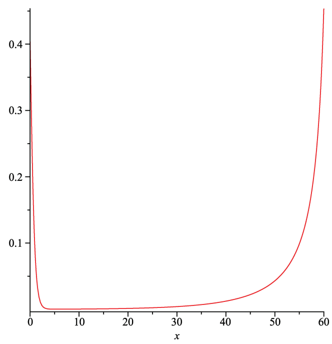


Figure 12

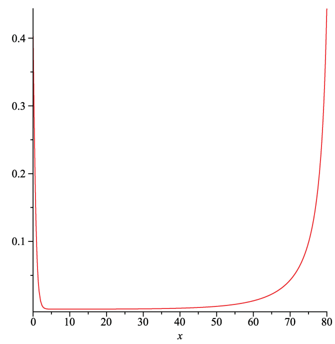


Figure 13

Figure 12: $a_1 = 7, a_2 = 3, b = 50, a = 1.\delta = 2, q = 1.5, \eta = 1, \gamma = 0.1$
Figure 13: $a_1 = 7, a_2 = 3, b = 70, a = 1, \delta = 2, q = 1.5, \eta = 1, \gamma = 0.1$

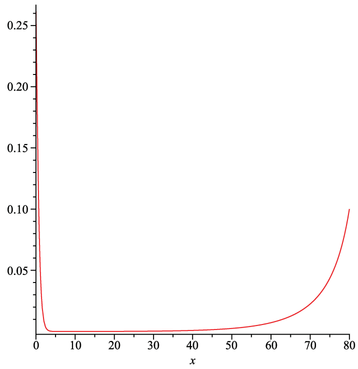


Figure 14

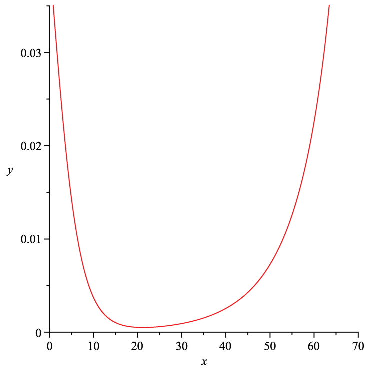


Figure 15

Figure 14: $a_1 = 7, a_2 = 3, b = 80, a = 1, \delta = 2, q = 1.5, \eta = 1, \gamma = 0.1$
Figure 15: $a_1 = 1, a_2 = 1, b = 70, \delta = 2.2, q = 1.9, \eta = 1, \gamma = 0.1$

This approach can give hazard functions of the desired shapes by selecting the parameters appropriately.

We shall try combination of $f_{10}(x)$ with a power function of the type

$$f_{12}(x) = \frac{\ln a}{a^b - 1} a^x, 0 \leq x \leq b, a > 0 \tag{54}$$

and zero elsewhere. We would like to have a slow rising $f_{12}(x)$. This can be achieved by taking a near 1 such as $a = 1.01$. The hazard function in this case is the following, denoting $f^*(x) = a_1 f_{10}(x) + a_2 f_{12}(x)$ and $S^*(t) = Pr\{x \geq t\}$ in $f^*(x)$:

$$h^*(x) = \frac{f^*(x)}{S^*(x)} = \frac{a_1 a\delta(\eta + 1 - q)\mathrm{e}^{-\delta x}[1 + a(q-1)\mathrm{e}^{-\delta x}]^{-\frac{\eta}{q-1}} + a_2 \frac{\ln a}{a^b - 1} a^x}{a_1[1 - [1 + a(q-1)\mathrm{e}^{-\delta x}]^{-\frac{\eta}{q-1}+1}] + a_2[\frac{a^b - a^x}{a^b - 1}]}. \tag{55}$$

For convenience, consider the pathway model $f_{10}(x) = \frac{\alpha(\rho-1)}{(1+\alpha x)^\rho}, 0 \leq x < \infty, \alpha > 0, \rho > 1$ and zero elsewhere. Then

$$h^*(x) = \frac{a_1[\alpha(\rho-1)(a^b-1)] + a_2[(\ln a)a^x(1+\alpha x)^\rho]}{a_1(1+\alpha x)(a^b-1) + a_2(a^b - a^x)(1+\alpha x)^\rho}. \tag{56}$$

From here we can get bathtub shaped curves for the hazard function. This is plotted for the following sets of parameters.

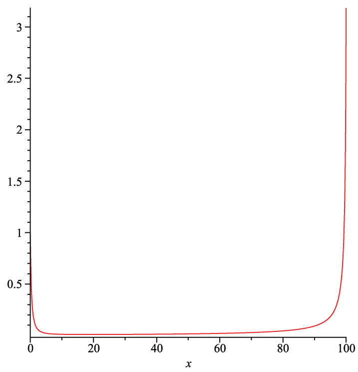


Figure 16

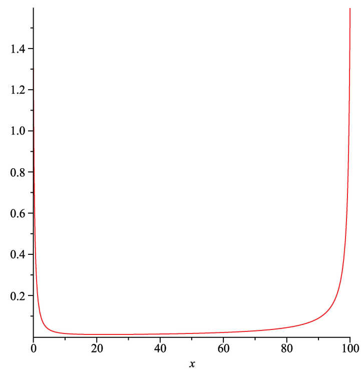


Figure 17

Figure 16: $a_1 = 1, a_2 = 1, \alpha = 1, \rho = 2, b = 100, a = 1.01$
Figure 17: $a_1 = 2, a_2 = 1, \alpha = 1, \rho = 2, b = 100, a = 1.01$

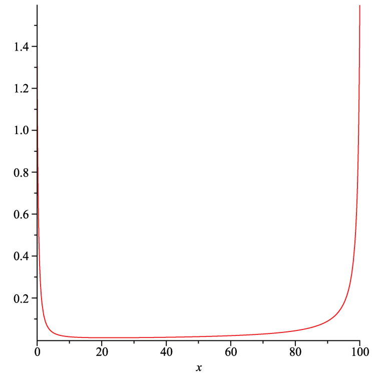


Figure 18: $a_1 = 2, a_2 = 1, \alpha = 2, \rho = 2, b = 100, a = 1.01$

6. Moments and Laplace Transforms

Let us consider the special case where $\gamma = \delta - 1$ in (1),(2),(3) with $x \geq 0$. This situation is more relevant to reliability analysis. In this case the pathway models are the following:

$$K_1(x) = a\delta(\eta + 1 - q)x^{\delta-1}[1 - a(1-q)x^{\delta}]^{\frac{\eta}{1-q}} \tag{57}$$

for $q < 1, a > 0, \delta > 0, \eta > 0.\eta + 1 - q > 0, 1 - a(1-q)x^{\delta} > 0$.

$$K_2(x) = a\delta(\eta + 1 - q)x^{\delta-1}[1 + a(q-1)x^{\delta}]^{-\frac{\eta}{q-1}} \tag{58}$$

for $q > 1, a > 0, \delta > 0, \eta > 0, \eta + 1 - q > 0, x \geq 0$.

$$K_3(x) = a\delta\eta x^{\delta-1}\mathrm{e}^{-a\eta x^{\delta}}, a > 0, \eta > 0, \delta > 0. \tag{59}$$

As illustrated before, in (58) and (59) both x and $\frac{1}{x}$ belong to the same family of distributions, namely type-2 pathway model and generalized gamma model respectively. Consider arbitrary moments $E(x^h)$ for a complex number h. This is available from the type-1 beta integral, type-2 beta integral and gamma integral respectively, and they are the following, denoted by $E^{(j)}(x^h), j = 1, 2, 3$:

$$E^{(1)}(x^h) = \frac{a(\eta + 1 - q)}{[a(1-q)]^{\frac{h}{\delta}+1}} \frac{\Gamma(\frac{h}{\delta}+1)\Gamma(\frac{\eta}{1-q}+1)}{\Gamma(\frac{h}{\delta}+\frac{\eta}{1-q}+2)}, \tag{60}$$

for $a < 1, \Re(\frac{h}{\delta} + 1) > 0$.

$$E^{(2)}(x^h) = \frac{a(\eta + 1 - q)}{[a(q-1)]^{\frac{h}{\delta}+1}} \frac{\Gamma(\frac{h}{\delta}+1)\Gamma(\frac{\eta}{q-1}-\frac{h}{\delta}-1)}{\Gamma(\frac{\eta}{q-1})} \tag{61}$$

for $q > 1, \Re(\frac{h}{\delta} + 1) > 0, \Re(\frac{\eta}{q-1} - \frac{h}{\delta} - 1) > 0$ or $-\delta < \Re(h) < \frac{\delta\eta}{q-1} + \delta$. Thus, only a few moments will exist here. But the above strip of analyticity is sufficient to compute the density via the inverse Mellin transform.

$$E^{(3)}(x^h) = \frac{\Gamma(\frac{h}{\delta}+1)}{(a\eta)^{\frac{h}{\delta}}}, a > 0, \delta > 0, \eta > 0, \Re(\frac{h}{\delta} + 1) > 0. \tag{62}$$

Let the Laplace transforms with Laplace parameter t be denoted by $L_{g_1}(t), L_{g_2}(t), L_{g_3}(t)$ respectively.

$$\begin{aligned} L_{g_1}(t) &= a\delta(\eta + 1 - q)\int_0^{[a(1-q)]^{-\frac{1}{\delta}}} \mathrm{e}^{-tx}x^{\delta-1}[1 - a(1-q)x^{\delta}]^{\frac{\eta}{1-q}}\mathrm{d}x \\ &= \frac{(\eta + 1 - q)}{(1-q)}\int_0^1 \mathrm{e}^{-\left[\frac{t^{\delta}u}{a(1-q)}\right]^{\frac{1}{\delta}}}(1-u)^{\frac{\eta}{1-q}}\mathrm{d}u. \end{aligned}$$

Expanding the exponential part and integrating term by term we have the following:

$$L_{g_1}(t) = \frac{(\eta+1-q)}{(1-q)} \sum_{k=0}^{\infty} \frac{(-t)^k}{k!\ [a(1-q)]^{\frac{k}{\delta}}} \int_0^1 u^{\frac{k}{\delta}} (1-u)^{\frac{\eta}{1-q}} \mathrm{d}u$$
$$= \frac{(\eta+1-q)\Gamma(\frac{\eta}{1-q}+1)}{(1-q)} \sum_{k=0}^{\infty} \frac{\Gamma(\frac{k}{\delta}+1)}{k!\Gamma(\frac{k}{\delta}+\frac{\eta}{1-q}+2)} \left[\frac{-t}{[a(1-q)]^{\frac{1}{\delta}}}\right]^k. \quad (63)$$

If $\frac{1}{\delta} = m, m = 2, 3, ...$ then we can expand the gammas by using the multiplication formula for gamma functions, namely,

$$\Gamma(mz) = \sqrt{2\pi}\ m^{mz-\frac{1}{2}} \Gamma(z)\Gamma(z+\frac{1}{m})...\Gamma(z+\frac{m-1}{m}) \quad (64)$$

and then (63) can be written as a hypergeometric series of the type ${}_mF_m$.
Now, consider $L_{g_2}(t)$. Since only a few moments exist the Laplace transform or moment generating function does not exist here. But for truncated case the Laplace transform will exist. Let the right tail be truncated out from $x = b$ onward where $b < [a(q-1)]^{-\frac{1}{\delta}}$ so that $0 < a(q-1)x^{\delta} < 1$. In this case the truncated density is the following:

$$g_2^*(x) = \frac{(\eta+1-q)a\delta}{Pr\{x \le b\}} x^{\delta-1} [1 + a(q-1)x^{\delta}]^{-\frac{\eta}{q-1}}, 0 \le x \le b$$

and zero elsewhere. Then the Laplace transform in this truncated case is the following:

$$L_{g_2^*}(t) = \frac{(\eta+1-q)a\delta}{Pr\{x \le b\}} \int_0^b \mathrm{e}^{-tx} x^{\delta-1} [1 + a(q-1)x^{\delta}]^{-\frac{\eta}{q-1}} \mathrm{d}x$$
$$r = \frac{(\eta+1-q)a\delta}{Pr\{x \le b\}} \sum_{k=0}^{\infty} (\frac{\eta}{q-1})_k [\frac{(-1)^k}{k!}] [a(q-1)]^k \int_0^b x^{\delta k+\delta-1} \mathrm{e}^{-tx} \mathrm{d}x$$
$$= \frac{(\eta+1-q)a\delta}{t^{\delta} Pr\{x \le b\}} \sum_{k=0}^{\infty} (\frac{\eta}{q-1})_k \frac{(-1)^k}{k!} [\frac{a(q-1)}{t^{\delta}}]^k \gamma(k\delta+\delta, bt) \quad (65)$$

for $|\frac{a(q-1)}{t^{\delta}}| < 1$ where $\gamma(\xi, c) = \int_0^c x^{\xi-1} \mathrm{e}^{-x} \mathrm{d}x$ is the incomplete gamma function. In the truncated case of $g_2^*(x)$ we can also derive arbitrary moments. Let $c = Pr\{x \le b\}$. Then

$$E(x^h) = \frac{(\eta+1-q)a\delta}{c} \int_0^b x^h x^{\delta-1} [1 + a(q-1)x^{\delta}]^{-\frac{\eta}{q-1}} \mathrm{d}x$$
$$= \frac{(\eta+1-q)}{c(q-1} \frac{1}{[a(q-1)]^{\frac{h}{\delta}}} \int_0^{a(q-1)b^{\delta}} u^{\frac{h}{\delta}} [1+u]^{-\frac{\eta}{q-1}} \mathrm{d}u.$$

Since $0 < u < 1$ we can expand the binomial part. $(1+u)^{-\frac{\eta}{q-1}} = \sum_{k=0}^{\infty} (\frac{\eta}{q-1})_k \frac{(-1)^k}{k!} u^k$. Then

$$E(x^h) = \frac{(\eta+1-q)}{c(q-1)} \frac{1}{[a(q-1)]^{\frac{h}{\delta}}} \sum_{k=0}^{\infty} (\frac{\eta}{q-1})_k \frac{(-1)^k}{k!} \int_0^{a(q-1)b^{\delta}} u^{\frac{h}{\delta}+k} \mathrm{d}u$$
$$= \frac{(\eta+1-q)}{c(q-1)} b^h \frac{\delta[a(q-1)b^{\delta}]}{(h+\delta)} {}_2F_1(\frac{\eta}{q-1}, \frac{h}{\delta}+1; \frac{h}{\delta}+2; -a(q-1)b^{\delta}).$$

We already have $0 < a(q-1)b^{\delta} < 1$ and hence the ${}_2F_1$ series is convergent. Note that for $h = 0$ the ${}_2F_1$ reduces to a ${}_1F_0$ multiplied by a constant which is equal to $\frac{(q-1)c}{\eta+1-q}$ which when multiplied by the remaining factor gives 1.
The Laplace transform for g_3 is given by the following:

$$L_{g_3}(t) = a\delta\eta \int_0^{\infty} \mathrm{e}^{-tx} x^{\delta-1} \mathrm{e}^{-a\eta x^{\delta}} \mathrm{d}x$$
$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left[\frac{t}{(a\eta)^{\frac{1}{\delta}}}\right]^k \Gamma(\frac{k}{\delta}+1), \text{ for } |\frac{t}{(a\eta)^{\frac{1}{\delta}}}| < 1, \delta > 1. \quad (66)$$

7. Component Failures

In engineering fields, a system or network is described as a collection of parts or components. Usually a system is represented by as a network in which the system components are connected together either in series, parallel or a combination of these. Consider a system consisting of k components, these components acting independently. Let the life times of these k components be denoted by $x_1, .., x_k$. Suppose that the system fails if any component fails. Let the system failure time be denoted by x. Then $x = \min\{x_1, ..., x_k\}$. Then $Pr\{x \geq t\}$ is given by the product of the probabilities $Pr\{x_j \geq t\}, j = 1, ..., k$ because if the smallest is $\geq t$ then all are $\geq t$. The corresponding distribution function of x is

$$P\{x \leq t\} = 1 - P\{x > t\} = 1 - \prod_{i=1}^{k} P\{x_j > t\} = 1 - [P\{x_j > t\}^k]$$
$$= 1 - [1 - P\{x_1 \leq t\}]^k, \tag{67}$$

when $x_1, \ldots, x_n$ are independently and identically distributed. Then the survival function for x, denoted by $S_k(x)$, is the following:

$$S_k(t) = Pr\{x \geq t\} = \prod_{j=1}^{k} Pr\{x_j \geq t\} = \prod_{j=1}^{k} S_{x_j}(t). \tag{68}$$

Several generalized statistical models are developed by using the formula given in (67) (i.e, the concept of series system) for details see Pascoa et al. [10]. In (1980), Kumaraswamy [11] proposed a two-parameter distribution on $(0, 1)$, so called Kumaraswamy distribution. This is contained in the pathway model for $q < 1$. This type of generalizations contains distributions with unimodal and bathtub shaped hazard functions, see Cordeiro and de Castro [12] and Jones [13]. These generalized models include Kumaraswamy-Weibull distribution by Cordeiro et al. [14], Kumaraswamy-Gumbel distribution by Cordeiro et al. [15], Kumaraswamy-generalized gamma distribution by Pascoa et al. [10], Kumaraswamy-log-logistic distribution by Tiago et al. [16], Kumaraswamy-modified Weibull distribution by Cordeiro et al. [17], Kumaraswamy-half Cauchy distribution by Gosh [18], Kw-generalized Rayleigh distribution by Antonio et al. [19] and Kumaraswamy-Gompertz distribution by Rocha et al. [20].

Let the life times be pathway distributed. For convenience let us take the case where the pathway parameter $q > 1$ or $1 < q < \eta + 1$. Then from (27)

$$S_{x_j}(t) = Pr\{x_j \geq t\} = [1 + a_j(q_j - 1)t^{\delta_j}]^{b_j+1} \tag{69}$$

where $b_j = -\frac{\eta_j}{q_j-1}, b_j + 1 = -\frac{\eta_j+1-q_j}{q_j-1}$. The density of x, denoted by $f_x(t)$ is available from $S_x(t)$ by differentiation. That is,

$$f_x(t) = -\frac{\mathrm{d}}{\mathrm{d}t} S_x(t) = -\frac{\mathrm{d}}{\mathrm{d}t} \prod_{j=1}^{k} [1 + a_j(q_j - 1)t^{\delta_j}]^{b_j+1}$$
$$= \sum_{j=1}^{k} [-(b_j + 1)][1 + a_j(q_j - 1)t^{\delta_j}]^{b_j} [\frac{\mathrm{d}}{\mathrm{d}t}(1 + a_j(q_j - 1)t^{\delta_j})] \prod_{i \neq j=1}^{k} [1 + a_i(q_i - 1)t^{\delta_i}]^{b_i+1}$$
$$= \sum_{j=1}^{k} [\prod_{i \neq j=1}^{k} (1 + a_i(q_i - 1)t^{\delta_i})^{b_i+1}][1 + a_j(q_j - 1)t^{\delta_j}][a_j(q_j - 1)\delta_j t^{\delta_j - 1}]. \tag{70}$$

Hence the hazard function for x is given by the following:

$$h_x(t) = \frac{\sum_{j=1}^{k} (\eta_j + 1 - q_j)[a_j \delta_j t^{\delta_j - 1}][\prod_{i \neq j=1}^{k} (1 + a_i(q_i - 1)t^{\delta_i})^{b_i+1}][1 + a_j(q_j - 1)t^{\delta_j}]}{\prod_{j=1}^{k} (1 + a_j(q_j - 1)t^{\delta_j})^{b_j+1}} \tag{71}$$

This is a very interesting form. Note that in (7.4) all the three families of functions, namely the generalized type-1 beta, type-2 beta and gamma families are involved. For different j, q_j can be $q_j < 1$ or $q_j > 1, 1 < q_j < \eta_j + 1$ or $q_j \to 1$. Hence we can consider many special cases of various types. For example, let $k = 2$ and $q_2 \to 1$. Then

$$h_x(t) = (\eta_1 + 1 - q_1)\frac{\delta_1 a_1 t^{\delta_1 - 1}}{[1 + a_1(q_1 - 1)t^{\delta_1}]^{b_1}} + \delta_2 a_2 \eta_2 t^{\delta_2 - 1}. \tag{72}$$

The graph for the following parameter values is given below.

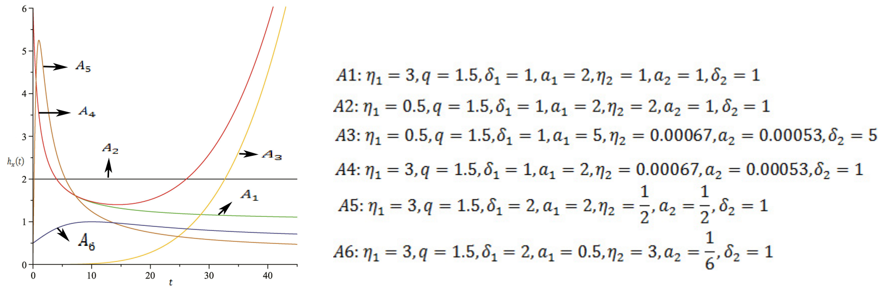


Figure 19: Plots of $h_x(t)$ for $b_1 = 1$ and different values of the other parameters

8. Estimation of Reliability Function

Let $x_1, \ldots, x_n$ denote a random sample of size n from the pathway model (21) with parameters a, q, δ, η. From (21), the logarithmic likelihood function is

$$L(a, \delta, \eta, q; y) = n \ln a + n \ln \delta + n \ln(\eta + 1 - q) + (\delta - 1)\sum_{i=1}^{n} \ln y_i - \frac{\eta}{q-1}\sum_{i=1}^{n} \ln(1 + a(q-1)y_i^{\delta}). \tag{73}$$

First of all we differentiate (73) with respect to all unknown parameters and equate these differential equations to zero. The MLEs of the unknown parameters are obtained on solving these differential equations simultaneously. Let $\hat{a}, \hat{q}, \hat{\eta}$ and $\hat{\delta}$ be the MLEs of the a, q, η, and δ respectively.

Theorem 1. The MLE of $S_2(t)$ is given by

$$\hat{S_2(t)} = [1 + \hat{a}(\hat{q} - 1)t^{\hat{\delta}}]^{-\frac{\hat{\eta}}{\hat{q}-1}+1}. \tag{74}$$

Using the invariance property of the MLEs, we can easily establish the above result.
The first derivative of the logarithmic likelihood function relative to the parameters is non-linear, and analytical solutions are difficult to obtain. A constrained optimization method can be used to solve such kinds of equations. This optimization problem can be carried out using constrOptim () or optim () function in R software.

8.1. Simulation Study

Random samples were generated from the pathway model using its distribution function. We considered a random sample of sizes $n = 200, 400, 600$ and the procedure was repeated 1000 times. The maximum likelihood estimate was computed using the optim () function in R. The results are given in the following tables, bias and MSE can be observed to decrease as the sample size increases. Similarly, the parameters of the distribution corresponding to $q < 1$ can also estimated by using this method. We are able to estimate the parameters of all the models proposed in this paper using the same procedure.

Table 1: ML estimate, Bias, MSE's for the parameter value $(q, a, \eta, \delta) = (1.2, 2, 3, 1)$

n	$\hat{q}$	Abs.Bias	MSE	$\hat{a}$	Abs.Bias	MSE
200	1.2256	0.0256	0.00065	2.1298	0.1298	0.0168
400	1.2134	0.0237	0.00018	2.0237	0.0237	0.0005
600	1.2109	0.0108	0.00011	2.0079	0.0079	0.00006
n	$\hat{\eta}$	Abs.Bias	MSE	$\hat{\delta}$	Abs.Bias	MSE
200	3.1186	0.1186	0.0140	1.015	0.0158	0.0002
400	3.1148	0.1148	0.0131	1.0063	0.0066	0.00004
600	3.0894	0.0894	0.0079	1.0049	0.0049	0.00002

Table 2: ML estimate, Bias, MSE's for the parameter value $(q, a, \eta, \delta) = (1.1, 2, 2, 5)$

n	$\hat{q}$	Abs.Bias	MSE	$\hat{a}$	Abs.Bias	MSE
200	1.1360	0.0360	0.0013	2.1013	0.1013	0.0102
400	1.1190	0.0190	0.0003	2.0496	0.0496	0.0024
600	1.1113	0.0113	0.00012	2.0433	0.0433	0.0018
n	$\hat{\eta}$	Abs.Bias	MSE	$\hat{\delta}$	Abs.Bias	MSE
200	2.0831	0.0831	0.0069	5.0953	0.0953	0.00908
400	2.0397	0.0397	0.0015	5.0466	0.0466	0.00217
600	2.0148	0.0148	0.0002	5.0344	0.0344	0.00118

Table 3: ML estimate, Bias, MSE's for the parameter value $(q, a, \eta, \delta) = (1.2, 4, 3, 5)$

n	$\hat{q}$	Abs.Bias	MSE	$\hat{a}$	Abs.Bias	MSE
200	1.2423	0.0423	0.0017	4.3068	0.3068	0.0941
400	1.2243	0.0243	0.0005	4.0791	0.0791	0.0062
600	1.2155	0.0155	0.0002	4.0746	0.0746	0.0055
n	$\hat{\eta}$	Abs.Bias	MSE	$\hat{\delta}$	Abs.Bias	MSE
200	3.1272	0.1272	0.0161	5.0793	0.0793	0.0062
400	3.1246	0.1246	0.0155	5.0331	0.0331	0.0010
600	3.0505	0.0505	0.0025	5.0247	0.0247	0.00061

9. Reliability for Arbitrary Values of the Parameters

In this section we discussed some simple numerical results for illustrating the theory developed in this paper. The reliability measurement of the system was obtained for arbitrary values of parameters related to the number of components, running time, etc. The reliability behavior of the system was graphically observed to identify the best possible configuration of the components with enhanced reliability of the system.

Table 4: Reliability measure for $a = 0.5, \delta = 0.1, \eta = 2$ and various values of q at time $t = 10$

No of components	$q = 1.2$	$q = 1.3$	$q = 1.4$	$q = 1.5$	$q = 1.6$
1	0.3439755	0.3752348	0.4072686	0.4400367	0.4735025
2	0.1183192	0.1408012	0.1658677	0.1936323	0.2242046
3	0.0406989	0.0528335	0.0675527	0.0852053	0.1061614
4	0.0139994	0.019825	0.0275121	0.0374935	0.0502677
5	0.0048155	0.007439	0.0112048	0.0164985	0.0238019
6	0.0016564	0.0027914	0.0045634	0.00726	0.0112702
7	0.0005698	0.0010474	0.0018585	0.0031946	0.0053365
8	0.000196	0.000393	0.0007569	0.0014058	0.0025268
9	0.0000674	0.0001475	0.0003083	0.0006186	0.0011965
10	0.0000232	0.0000553	0.0001255	0.0002722	0.0005665

Table 5: Reliability measure for $q = 1.6, \delta = 0.1, \eta = 2$ and various values of a at time $t = 10$

No. Of Components (n)	$a = 0.02$	$a = 0.03$	$a = 0.04$	$a = 0.05$	$a = 0.06$
1	0.9656186	0.9490587	0.9329014	0.9171343	0.9017451
2	0.9324194	0.9007123	0.870305	0.8411353	0.8131442
3	0.9003615	0.8548288	0.8119088	0.771434	0.7332488
4	0.8694059	0.8112827	0.7574309	0.7075086	0.6612035
5	0.8395145	0.7699549	0.7066083	0.6488803	0.596237
6	0.8106509	0.7307323	0.6591959	0.5951104	0.5376538
7	0.7827796	0.6935078	0.6149648	0.5457961	0.4848267
8	0.7558666	0.6581796	0.5737015	0.5005684	0.4371901
9	0.7298789	0.6246511	0.535207	0.4590884	0.394234
10	0.7047846	0.5928305	0.4992953	0.4210457	0.3554986

Table 6: Reliability measure for $q = 1.6, a = 0.06, \eta = 2$ and various values of δ at time $t = 10$

No of Components	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$	$\delta = 0.5$	$\delta = 0.6$
1	0.8785595	0.8505635	0.8170973	0.7775798	0.7316045
2	0.7718669	0.7234583	0.667648	0.6046303	0.5352452
3	0.678131	0.6153472	0.5455333	0.4701483	0.3915878
4	0.5957784	0.5233919	0.4457538	0.3655778	0.2864874
5	0.5234268	0.445178	0.3642242	0.2842659	0.2095955
6	0.4598616	0.3786522	0.2976066	0.2210394	0.153341
7	0.4040158	0.3220677	0.2431735	0.1718758	0.112185
8	0.3549519	0.273939	0.1986964	0.1336471	0.082075
9	0.3118464	0.2330025	0.1623543	0.1039213	0.0600465
10	0.2739756	0.1981835	0.1326593	0.0808071	0.0439303

Table 7: Reliability measure for $q = 1.6, a = 0.06, \delta = 0.6$ and various values of η at time $t = 10$

No of Components	$\eta = 3$	$\eta = 4$	$\eta = 5$	$\eta = 6$	$\eta = 7$
1	0.5852359	0.4681505	0.3744899	0.2995675	0.2396344
2	0.342501	0.2191649	0.1402427	0.0897407	0.0574247
3	0.2004439	0.1026022	0.0525195	0.0268834	0.0137609
4	0.117307	0.0480333	0.019668	0.0080534	0.0032976
5	0.0686522	0.0224868	0.0073655	0.0024125	0.0007902
6	0.0401778	0.0105272	0.0027583	0.0007227	0.0001894
7	0.0235135	0.0049283	0.001033	0.0002165	0.0000454
8	0.0137609	0.0023072	0.0003868	0.000064	0.0000109
9	0.0080534	0.0010801	0.0001449	0.0000194	0.0000026
10	0.0047131	0.0005057	0.0000543	0.0000058	0.0000006

The following are the graphical representations of reliability relative to the number of components n.

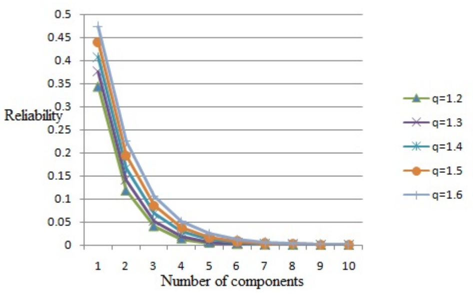


Figure 20

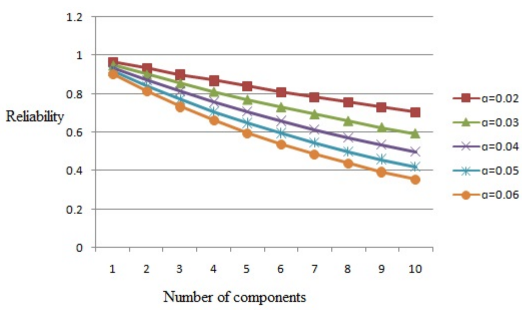


Figure 21

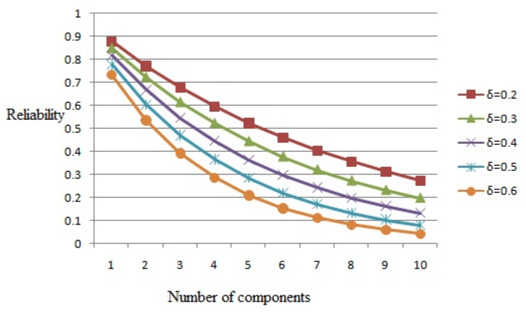


Figure 22

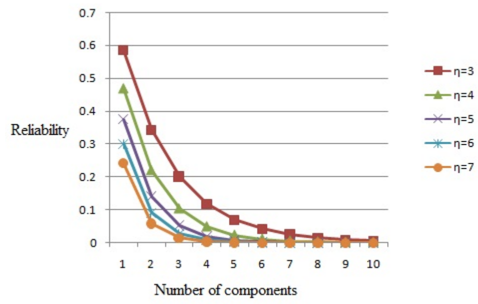


Figure 23

The results obtained for arbitrary values of the parameters indicate that reliability of a series system of 10 identical components keeps on decreasing with the increasing number of components. Several distributions are obtained from our paper for particular parameter values and various arbitrary functions $P(x)$. As $q \to 1$ the models (20) and (21) become the Weibull distribution, so it can be considered as an extended version of the Weibull distribution.

10. Conclusions

An arbitrary function is introduced in the pathway model, for constructing the hazard functions of desired shapes. In the present study, we conclude that reliability continues to decline as the number of components increases. It is recommended to use the smallest number of components in a series system for better performance. However, the performance of these systems can be enhanced by using components that follow the extended form of Weibull failure laws.

References

[1] Mathai, A. M. (2005): A pathway to matrix-variate gamma and normal densities, *Linear Algebra and its Applications*, 396: 317–328.

[2] Mathai, A. M. and Provost, S. B. (2006). Some complex matrix variate statistical distributions in rectangular matrices. *Linear Algebra and its Applications*, 410: 198–216.

[3] Tsallis, C. Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World, Springer, New York, 2009.

[4] Beck, C. and Cohen, E. G. D. (2003). Superstatistics. *Physica A*, 322: 267–275.

[5] Mathai, A. M. and Provost, S. B. (2006a). On q-logistic and related distributions. *IEEE Transactions on Reliability*, 55: 237–244.

[6] Mathai, A. M. and Haubold, H. J. (2007). Pathway model, superstatistics, Tsallis statistics and generalized measure of entropy. *Physica A*, 375: 110–122.

[7] Mathai, A. M. and Rathie, P. N. Basic Concepts in Information Theory and Statistics: Axiomatic Foundations and Applications, Wiley Halsted, New York and Wiley Eastern, New Delhi, 1975.

[8] Mathai, A. M. An Introduction to Geometrical Probability: Distributional Aspects with Applications, Gordon and Breach, Newark, USA, 1999.

[9] Mathai, A. M. and Princy, T. (2017). Analogues of reliability analysis for matrix-variate cases. *Linear Algebra and its Applications*, 532: 287–311.

[10] Pascoa, A. R. M., Ortega, E. M. M. and Cordeiro, G. M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, 8: 411-"433.

[11] Kumaraswamy, P. (1980). Generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 462: 79–88.

[12] Cordeiro, G. M., Castro, M. de. (2011). A new family of generalized distributions, *Journal of Statistical Computation and Simulation*. 81: 883–898.

[13] Jones, M. C. (2009). Kumaraswamy's distribution: a beta-type distribution with some tractability advantages, *Statistical Methodology*. 6: 70–81.

[14] Cordeiro, G. M., Ortega, E. M. M. and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*, 347: 1399–1429.

[15] Cordeiro, G. M., Nadarajah, S. and Ortega, E. M. M. (2011). The Kumaraswamy Gumbel distribution. *Statistical Methods and Applications*. 21: 139–168.

[16] Tiago Viana Flor de Santana, Edwin M. M. Ortega, Gauss M. Cordeiro and Giovana O. Silva, (2012). The Kumaraswamy-log-logistic distribution. *Journal of Statistical Theory and Applications*. 11: 265–291.

[17] Cordeiro, G. M., Ortega, E. M. M. and Silva, G. O. (2012). The Kumaraswamy modified Weibull distribution: theory and applications. *Journal of Statistical Computation and Simulation*. 84: 1387–1411.

[18] Ghosh, I. (2014). The Kumaraswamy-half-Cauchy distribution: properties and applications. *Journal of Statistical Theory and Applications*. 13: 122–134.

[19] Antonio E. Gomes, Cibele Q. da-Silva, Gauss M. Cordeiro and Edwin M.M. Ortega (2014). A new lifetime model: the Kumaraswamy generalized Rayleigh distribution. *Journal of Statistical Computation and Simulation*, 84: 290–309.

[20] Rocha,R., Nadarajah, S., Tomazella, V., Louzada, F. and Eudes, A. (2015). New defective models based on the Kumaraswamy family of distributions with application to cancer data sets. *Statistical methods in medical research*, 0962280215587976.

REPRESENTATION OF CERTAIN LIFETIME MODELS VIA SEQUENCES OF SPECIAL NUMBERS

Christian Genest and Nikolai Kolev

•

McGill University and Universidade de São Paulo
christian.genest@mcgill.ca and kolev.ime@gmail.com

Abstract

Stimulated by the work of Rządkowski et al. (2015, J. Nonlinear Math. Phys., 22 (2015), 374–380), the authors derive representations for certain classes of univariate and bivariate lifetime distributions in terms of sequences of Bell, Bernoulli, and Stirling numbers of the second kind, their generalizations, and associated polynomials. Gould–Hopper polynomials are used in the bivariate case, leading to representations for large classes of distributions satisfying a law of uniform seniority for dependent lives formulated by Genest and Kolev (Scan. Act. J., 2021-8 (2021), 726–743).

Keywords: Bell numbers, Bernoulli numbers, Gould–Hopper polynomials, law of uniform seniority, lifetime distributions, Stirling numbers of the second kind, survival functions.

1. Introduction

A random variable X is said to have a Gumbel distribution with parameters $a \in \mathbb{R}$ and $b \in (0, \infty)$ if and only if, for every real $x \in \mathbb{R}$, one has

$$\Pr(X \le x) = \exp\{-e^{-(x-a)/b}\}. \tag{1}$$

Also known as the log-Weibull and double exponential distribution, this model belongs to the class of generalized extreme-value or Fisher–Tippett distributions. It was identified by Fisher and Tippett [6] as one of the three possible limit distributions of properly normalized maxima of a sequence of mutually independent and identically distributed random variables.

In a note published in 2015, Rządkowski et al. [18] pointed out that Gumbel distributions are related to Stirling numbers of the second kind and Bernoulli numbers. Limiting the discussion to the case $a = 0$ and $b = 1$ for simplicity, these authors showed that the cumulative distribution function of the standard Gumbel distribution, defined for every real $x \in \mathbb{R}$, by $G(x) = \exp(-e^{-x})$, is such that, for every integer $n \in \mathbb{N} = \{1, 2, \ldots\}$,

$$\int \{G^{(n)}(x)\}^2 dx = \frac{(-1)^n}{2n} B_{2n}(1 - 2^{2n}),$$

where $G^{(n)}$ denotes the nth derivative of G and B_n is the nth Bell polynomial defined in terms of the Stirling numbers $S(n, k)$ of the second kind by setting, for every real $x \in \mathbb{R}$,

$$B_n(x) = \sum_{k=1}^{n} S(n, k)\, x^k.$$

Recall that for arbitrary integers $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$, $S(n, k)$ represents the number of ways in which one can partition a set of n elements into k non-empty and non-overlapping subsets. The nth Bell number $B_n = B_n(1) = S(n, 1) + \cdots + S(n, n)$ is then the number of partitions of a set with n elements. For convenience, one also sets $B_0 = S(0,0) = 1$ and $S(n,0) = 0$ for every integer $n \in \mathbb{N}$. For additional information, the reader is referred to the book by Comtet [3].

The purpose of this note is to extend the observation of Rządkowski et al. [18] to a large class of lifetime distributions routinely used in actuarial practice, and to show how the approach can be extended to the bivariate case. The starting point is the fact that the exponential generating function of the Bell numbers is such that, for every real $x \in \mathbb{R}$,

$$B(x) = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!} = \exp(e^x - 1), \tag{2}$$

and hence both $e^x = 1 + \ln\{B(x)\}$ and $x = \ln[1 + \ln\{B(x)\}]$.

As shown in Section 2, these relations can be used to express some of the most common univariate lifetime distributions in terms of Bell numbers. A bivariate extension is then considered in Section 3 using a special case of the Appell polynomials introduced by Gould and Hopper [10], and recently discussed in the multivariate case by Ricci et al. [17].

Specifically, for any integers $m, n \in \mathbb{N}$ with $m \geq 2$, the Gould–Hopper generalization of the classical Hermite polynomial of order $m-1$ is defined, for every pair $(x, y) \in \mathbb{R}^2$, by

$$H_n^{[m-1]}(x,y) = \sum_{k=0}^{\lfloor \frac{n}{m-1} \rfloor} \frac{n!}{k!(n-mk+k)!} x^{n-mk+k} y^k,$$

where in general, $\lfloor x \rfloor$ refers to the integer part of x. As shown by Gould and Hopper [10], one has, for every pair $(x, y) \in \mathbb{R}^2$ and $\gamma \in (0, \infty)$,

$$\sum_{n=0}^{\infty} H_n^{[m-1]}(x,y) \frac{\gamma^n}{n!} = \exp(\gamma x + \gamma^{m-1} y). \tag{3}$$

Relation (3) will be used in Section 3 to derive the Gould–Hopper polynomial expansion of continuous bivariate models satisfying a law of uniform seniority for dependent lives recently introduced and characterized in [7]. A few concluding comments are given in Section 4.

2. Expansions for univariate lifetime distributions

Representations in terms of Bell, Bernoulli, and Stirling numbers of the second kind are given below for various classes of univariate lifetime distributions. The classical Gompertz law of mortality is considered in Section 2.1, and its extension to the Gompertz–Makeham model is the object of Section 2.2. The little known Teissier and Chiang–Conforti distributions are treated in Section 2.3.

2.1. Bell-number expansion of Gompertz's law

Gompertz's law is possibly the oldest documented demographic model. Proposed by the British actuary Benjamin Gompertz (1779–1865) in the early 19th century [9], this model states that for any fixed scale parameter $\lambda \in (0, \infty)$ and shape parameter $\theta \in [0, \infty)$, the survival probability of a random lifetime T is given, at any age $t \in [0, \infty)$, by

$$\Pr(T > t) = \exp\{-\theta(e^{\lambda t} - 1)\}. \tag{4}$$

As is readily seen, Gompertz's law is the same as the Gumbel distribution for the negative of age, restricted to negative values. More specifically, if $T = -X$ in Eq. (1) and $t = -x$ is restricted to positive values for age, then $\Pr(T > t)$ is of the form (4) with $\lambda = 1/b$, $\theta = e^{a/b}$, save for the normalizing constant e^θ which ensures that $\Pr(T > 0) = 1$.

Given the result of Rządkowski et al. [18], it may thus be suspected that a similar result holds for Gompertz's law. The following simple result confirms this suspicion.

Proposition 1. *The Bell number representation of Gompertz's survival function* (4) *is given, for every real* $t \in [0, \infty)$, *by* $\Pr(T > t) = \{B(\lambda t)\}^{-\theta}$.

Indeed, it follows at once from Eq. (2) that, for every real $t \in [0, \infty)$, one has

$$\{B(\lambda t)\}^{-\theta} = \exp\{-\theta(e^{\lambda t} - 1)\}, \tag{5}$$

which is precisely the survival function (4) of Gompertz's law. The hazard rate (or force of mortality) implied by Gompertz's law is given, at any age $t \in (0, \infty)$, by $-\{\Pr(T > t)\}^{-1} \, d \Pr(T > t)/dt = \theta\lambda \exp(\lambda t)$. It corresponds to an exponential increase in death rate with age which is echoed in the exponential increase of the Bell numbers: the first ten are 1, 1, 2, 5, 15, 52, 203, 877, 4140 and 21147. Similarly, the convexity of this hazard rate is mirrored by the convexity of the sequence of Bell numbers, namely the fact that for every integer $n \in \mathbb{N}$, one has $B_n \leq (B_{n-1} + B_{n+1})/2$; see Exercise 1 on p. 291 in Comtet [3].

If desired, connections between Gompertz's law and other famous sequences of numbers could be found just as easily through their corresponding exponential generating function. For instance, a representation of the survival function (4) could be established in terms of

(i) Stirling numbers of the second kind $S_2(k, n)$, given that for every integer $k \in \mathbb{N}$ and real $x \in \mathbb{R}$, one has
$$(e^x - 1)^k / k! = \sum_{n=k}^{\infty} S(k, n) \, \frac{x^n}{n!};$$

(ii) Bernoulli numbers, $\mathcal{B}_n$, given that $\mathcal{B}_0 = 1$ and that for every real $x \in \mathbb{R}$, one has
$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \mathcal{B}_n \, \frac{x^n}{n!}.$$
In particular, $\mathcal{B}_1 = -0.5$, $\mathcal{B}_2 = 1/6$, and $\mathcal{B}_{2k+1} = 0$ for every integer $k \in \mathbb{N}$.

Further note that if λ and θ are replaced by $-\lambda$ and $-\theta$, respectively, Eq. (5) transforms into the survival function of the negative Gompertz distribution defined, at every real $t \in [0, \infty)$, by $\exp\{\theta(e^{-\lambda t} - 1)\}$. When $\lambda = 1$ in the latter expression, one gets the Laplace transform of a Poisson distribution with parameter $\theta \in (0, \infty)$; see Chapter 10 of the book by Marshall and Olkin [15].

2.2. r-Bell number expansion of the Gompertz–Makeham law

The Gompertz–Makeham law is an extension of the Gompertz model which was proposed by another British actuary, William Matthew Makeham (1826–1891). The latter proposed to add an age-independent term to the exponentially age-dependent term of Gompertz's model. The resulting construction is one of the most effective theories to describe human mortality.

Stated differently, let X and Y be independent random variables, where X follows Gompertz's law with survival function (4) and Y has an exponential distribution with mean $\xi\lambda\theta$ for some $\xi \in (0, \infty)$. Then the random variable $Z = \min(X, Y)$ has the Gompertz–Makeham distribution with survival function given, at any age $t \in [0, \infty)$, by

$$\Pr(Z > t) = \exp\{-\xi\lambda\theta t - \theta(e^{\lambda t} - 1)\}. \tag{6}$$

Refer to Chapter 10 of the book by Marshall and Olkin [15] for details and historical notes. This model is still used today to describe the age dynamics of human mortality and construct life tables, with remarkable accuracy, for individuals between 30 and 80 years of age.

Now consider the r-Stirling number of the second kind with integer-valued parameters $n \geq k \geq r$, denoted $S(n, k, r)$. This number is a count of the partitions of the set $\{1, \ldots, n\}$ into k non-empty non-overlapping subsets, such that the integers $1, \ldots, r$ are in distinct subsets. By convention, one sets $S(n,k,0) = S(n,k)$. See Broder [1] for basic facts about r-Stirling numbers of the first and second kind.

Mezo [16] defined the nth r-Bell number, denoted $B_{n,r}$, by $B_{n,r} = S(n,0,r) + \cdots + S(n,n,r)$ so that by convention, one has $B_{n,0} = B_n$. More generally, the nth r-Bell number is a count of the partitions of a set with $n + r$ elements whose first r elements are in distinct subsets of the partition. The first few r-Bell numbers are given in Table A134980 of Sloane [19]. For example, the first seven 6-Bell numbers are 1, 7, 50, 365, 2727, 20878, and 163967. Recent results about r-Bell numbers are reported by Corcino et al. [4].

Mezo [16] showed that the exponential generating function of the nth r-Bell number $B_{n,r}$ is given, at every real $x \in \mathbb{R}$, by

$$B_r(x) = \sum_{n=0}^{\infty} B_{n,r} \frac{x^n}{n!} = \exp(e^x - 1 + rx).$$

The following result is then immediate.

Proposition 2. *Suppose that the random variable Z has a Gompertz–Makeham distribution with scale parameter $\lambda \in (0,\infty)$ and shape parameter $\theta \in (0,\infty)$. Further suppose that the parameter ξ is integer-valued with $\xi \in \{r,\ldots,n\}$. Then the survival function of Z given in Eq.* (6) *is such that, for every real $t \in (0,\infty)$, one has $\Pr(Z > t) = \{B_\xi(\lambda t)\}^{-\theta}$.*

A connection between the Gompertz–Makeham survival function and r-Stirling numbers of the second kind can also be deduced from the fact that, for every real $x \in (0,\infty)$,

$$\frac{1}{k!}\left(e^x - \sum_{\ell=0}^{r-1} \frac{x^\ell}{\ell!}\right)^k = \sum_{n=kr}^{\infty} S(n,k,r)\frac{x^n}{n!}.$$

2.3. Two other Bell-number representations of distributions

In 1934, the French biologist Georges Teissier (1900–1972) introduced a model to describe the mortality of several domestic animal species protected from accidents and disease, i.e., dying from "pure aging" [20]. Based on data collected on several species, this author found that animal mortality does not follow Gompertz's law, as it does for humans.

The survival function of Teissier's distribution is defined, at any age $t \in [0,\infty)$, by $\exp(t + 1 - e^t)$. This model was later rediscovered by Laurent [13], who considered a one-parameter extension. A two-parameter version called the scaled Teissier distribution with parameters $\lambda \in (0,\infty)$ and $\theta \in (0,\infty)$ such that $\lambda\theta \le 1$ has survival function is given, for every real $t \in (0,\infty)$, by

$$S_{\mathrm{T}}(t) = \exp\{\lambda t - (e^{\lambda t} - 1)/(\lambda\theta)\}. \tag{7}$$

See Kolev et al. [12] for more details on the history of the Teissier distribution.

Taking into account Eq. (5) and the identity $\ln\{B(\lambda t)\} + 1 = e^{\lambda t}$, valid for all real $t \in [0,\infty)$, one can get the following Bell-number representation of Teissier's scaled model.

Proposition 3. *The Bell-number representations of the Teissier survival functions* (7) *is given, for every real $t \in [0,\infty)$, by $S_{\mathrm{T}}(t) = [\ln\{B(\lambda t)\} + 1]/\{B(\lambda t)\}^{1/(\lambda\theta)}$.*

As a final example, Chiang and Conforti [2] designed a stochastic model assuming that mortality intensity is a function of the accumulated effect of a person's continuous exposure to toxic material in the environment (absorbing coefficient) and their biological reaction to the toxin absorbed (discharging coefficient). Given parameters $\alpha \in (0,\infty)$ and $\theta \in (1,\infty)$, the survival function of the Chiang–Conforti model is given, for every real $t \in [0,\infty)$, by

$$S_{\mathrm{CC}}(x) = \exp\{-\lambda x - \theta(e^{-\lambda x} - 1)\}. \tag{8}$$

Note in passing that an equivalent form of this model had appeared earlier, with a different parametrization, in the work of Ghurye [8].

From Eq. (2), one has, for every real $t \in [0,\infty)$, $\ln\{B(-\lambda t)\} + 1 = e^{-\lambda t}$ and $\{B(-\lambda t)\}^{-\theta} = \exp\{-\theta(e^{-\lambda t} - 1)\}$, leading to the following representation of the Chiang–Conforti model.

Proposition 4. *The Bell-number representation of the Chang–Conforti survival function* (8) *is given, for every real* $t \in [0, \infty)$*, by* $S_{\mathrm{CC}}(t) = [\ln\{B(-\lambda t)\} + 1]/\{B(-\lambda t)\}^{\theta}$.

3. A bivariate extension

A common characteristic of the univariate models considered in Section 2 is that they have what Laurent [13] termed an exponential structure. Namely, in all cases considered, a survival function S could be expressed, for every real $t \in [0, \infty)$, in the form

$$S(t) = \exp\{-\Lambda(t)\} \tag{9}$$

for a map Λ which involved a linear and an exponential term. This map is called the cumulative hazard rate in survival analysis.

A similar approach can actually be used in higher dimensions. To make this point, consider the b-BLUS class of bivariate survival functions introduced in [7] as a model for dependent lives based on an extension of the actuarial law of uniform seniority [5].

Specifically, a pair (X, Y) of continuous random variables is said to belong to the b-BLUS class if there exist real-valued parameters $\alpha, \beta \in (0, \infty)$ and a continuous, strictly decreasing univariate survival function $\psi : [0, \infty) \to [0, 1]$ such that, for every pair $(x, y) \in (0, \infty)^2$, $\Pr(X > x, Y > y) = \psi(\alpha x + \beta y)$.

It will be shown below how various models from the b-BLUS class can be expressed in terms of Bell numbers, Bernoulli numbers, and Stirling numbers of the second kind with the help of the Gould–Hopper polynomials specified by Eq. (3).

3.1. Gould–Hopper expansion of Gompertz's bivariate law

Recall that a random pair (X, Y) has a bivariate Gompertz distribution with parameters $\alpha, \beta \in (0, \infty)$ and $\theta \in [1, \infty)$ whenever, for every pair $(x, y) \in [0, \infty)^2$, one has

$$S_{\mathrm{BG}}(x, y) = \Pr(X > x, Y > y) = \exp\{-\theta(e^{\alpha x + \beta y} - 1)\}. \tag{10}$$

The proof of the following result is straightforward.

Proposition 5. *The Gould–Hopper polynomial expansion of the bivariate Gompertz joint survival function* (10) *with real-valued parameters* $\theta \in (0, \infty)$*,* $\alpha = \gamma \in (0, \infty)$*, and* $\beta = \gamma^{m-1}$ *for some integer* $m \in \mathbb{N}$ *is given, for every pair* $(x, y) \in [0, \infty)^2$*, by*

$$S_{\mathrm{BG}}(x, y) = \exp[-\theta\{\mathcal{H}^{[m-1]}(x, y; \gamma) - 1\}],$$

where $\mathcal{H}^{[m-1]}$ *stands for the parametric function implicitly defined in Eq.* (3).

3.2. Gould–Hopper expansion of two other bivariate models

In [7], bivariate versions of the univariate Teissier and Chang–Conforti distributions (7) and (8) are defined as follows.

Definition 1. *A random pair* (X, Y) *is said to have a bivariate Teissier distribution with real-valued parameters* $\alpha \in (0, \infty)$*,* $\beta \in (0, \infty)$*, and* $\theta \in [(3 + \sqrt{5})/2, \infty)$ *if and only if, for every pair* $(x, y) \in (0, \infty)^2$*, one has*

$$S_{\mathrm{BT}}(x, y) = \Pr(X > x, Y > y) = \exp\{\alpha x + \beta y - \theta(e^{\alpha x + \beta y} - 1)\}. \tag{11}$$

Definition 2. *A random pair* (X, Y) *is said to have a bivariate Chang–Conforti distribution with real-valued parameters* $\alpha \in (0, \infty)$*,* $\beta \in (0, \infty)$*, and* $\theta \in [0, (3 - \sqrt{5})/2]$ *if and only if, for every pair* $(x, y) \in (0, \infty)^2$*, one has*

$$S_{\mathrm{BCC}}(x, y) = \Pr(X > x, Y > y) = \exp\{-\alpha x - \beta y - \theta(e^{-\alpha x - \beta y} - 1)\}. \tag{12}$$

The following results are then immediate from relation (3).

Proposition 6. *The Gould–Hopper polynomial expansion of the bivariate Teissier distribution with survival function* (11) *and real-valued parameters* $\alpha = \gamma \in (0, \infty)$, $\theta \in [(3 + \sqrt{5})/2, \infty)$, *and* $\beta = \gamma^{m-1}$ *for some integer* $m \in \mathbb{N}$, *is given, for every pair* $(x, y) \in [0, \infty)^2$, *by*

$$S_{\mathrm{BT}}(x, y) = \exp\{\ln[\mathcal{H}^{[m-1]}(x, y; \gamma)] - \theta[\mathcal{H}^{[m-1]}(x, y; \gamma) - 1]\}$$

where $\mathcal{H}^{[m-1]}$ *stands for the parametric function implicitly defined in Eq.* (3).

Proposition 7. *The Gould–Hopper polynomial expansion of the bivariate Chang–Conforti distribution with survival function* (12) *and real-valued parameters* $\alpha = \gamma \in (0, \infty)$, $\theta \in [0, (3 - \sqrt{5})/2]$, *and* $\beta = \gamma^{m-1}$ *for some integer* $m \in \mathbb{N}$ *is given, for every pair* $(x, y) \in [0, \infty)^2$, *by*

$$S_{\mathrm{BCC}}(x, y) = \exp\{-\ln[\mathcal{H}^{[m-1]}(x, y; \gamma)] - \theta[\mathcal{H}^{[m-1]}(-x, -y; \gamma) - 1]\},$$

where $\mathcal{H}^{[m-1]}$ *stands for the parametric function implicitly defined in Eq.* (3).

3.3. Two b-BLUS models with positive or negative dependence

The b-BLUS property discussed in [7] can model lives that are either positively or negatively dependent. The following two examples illustrate each one of these two cases. Given real-valued parameters $\alpha \in (0, \infty)$, $\beta \in (0, \infty)$, $\delta \in (0, \infty)$ and $\rho \in [1, \infty)$, consider the joint survival functions of random pairs (X, Y) defined, for every pair $(x, y) \in [0, \infty)^2$, by

$$S_1(x, y) = \Pr(X > x, Y > y) = (1 + \alpha x + \beta y)^{-1/\delta}$$

and

$$S_2(x, y) = \Pr(X > x, Y > y) = (1 - \alpha x - \beta y)^{\rho} \mathbf{1}_{(0,1)}(\alpha x + \beta y),$$

where in general, $\mathbf{1}_A$ refers to the indicator function of the set A.

In these two cases, a direct application of identity (3) with $\alpha = \gamma$ and $\beta = \gamma^{m-1}$ for some integer $m \in \mathbb{N}$ leads to the following Gould–Hopper polynomial expansions:

$$S_1(x, y) = \left\{1 + \ln \mathcal{H}^{[m-1]}(x, y; \gamma)\right\}^{-1/\delta} \quad \text{and} \quad S_2(x, y) = \left\{1 + \ln \mathcal{H}^{[m-1]}(x, y; \gamma)\right\}^{\rho}.$$

Here again, $\mathcal{H}^{[m-1]}$ stands for the parametric function implicitly defined in Eq. (3).

4. Final remarks

The initial motivation for this note was the intriguing observation of Rządkowski et al. [18] to the effect that Gumbel's distribution is related to Stirling numbers of the second kind and Bernoulli numbers. It was shown that upon expressing a survival function S through its cumulative hazard rate Λ as in Eq. (9), new relations involving Bell numbers, Bernoulli numbers, and Stirling numbers of the second kind can be obtained via Eq. (2) for a large class of survival models, at the cost of suitable restrictions on the parameters. Such representations may or may not be elegant, insightful or useful for computational purposes.

In the bivariate case, the survival function of any continuous random pair (X, Y) can be written, for every pair $(x, y) \in [0, \infty)^2$, in the form

$$\Pr(X > x, Y > y) = \exp\left\{-\int_{\mathcal{C}} R(z)\, dz\right\}, \tag{13}$$

where R is the hazard gradient vector and $\mathcal{C}$ is any sufficiently smooth continuous path beginning at $(0, 0)$ and terminating at (x, y). Relation (13) holds so long as S is absolutely continuous along the path of integration; see Marshall [14]. Therefore, the expressions derived in Section 3 for distributions in the b-BLUS class in terms of the Gould–Hopper polynomial are just as natural as those presented by Rządkowski et al. [18] and in Section 2.

Of course, the Gould–Hopper polynomials are not always the right tool. For instance, consider a random pair (X, Y) with Gumbel's bivariate exponential distribution [11] with real-valued parameters $\alpha \in (0, \infty)$, $\beta \in (0, \infty)$, and $\delta \in [0, \alpha\beta]$, whose survival function is defined, for every pair $(x, y) \in [0, \infty)^2$, by

$$\Pr(X > x, Y > y) = \exp(-\alpha x - \beta y - \delta xy) = \exp(-\alpha x - \beta y) \exp(xy)^{-\delta}.$$

The first factor can be expressed via the Gould–Hopper polynomial via Eq. (3), but for the second factor one must resort to the map defined, for every pair $(x, y) \in [0, \infty)^2$, by $W(x, y) = ye^{xy}/(e^y - 1)$, i.e., the generating function of the classical Bernoulli polynomials; see Comtet [3] for details. Letting B denote the Bell polynomial, one then finds, for every pair $(x, y) \in [0, \infty)^2$,

$$\exp(xy) = W(x, y)(e^y - 1)/y = W(x, y) \ln\{B(y)\} / \ln[1 + \ln\{B(y)\}].$$

Representations of multivariate continuous survival functions might be derived from recent results for Appell polynomials due to Ricci et al. [17]. This may be the object of future work.

Acknowledgments. Grants in support of this work were provided by the Canada Research Chairs Program (950-231937), the Natural Sciences and Engineering Research Council of Canada (RGPIN-2016-04720), and the Fundação de Amparo à Pesquisa do Estado de São Paulo (2013/07375-0).

References

[1] Broder, A. Z. (1984). The r-Stirling numbers. *Discrete Math.*, 49(3):241–259.
[2] Chiang, C. L. and Conforti, P. M. (1989). A survival model and estimation of time to tumor. *Math. Biosci.*, 94(1):1–29.
[3] Comtet, L. (1974). *Advanced Combinatorics: The Art of Finite and Infinite Expansions.* Revised and enlarged edition. D. Reidel Publishing Co., Dordrecht, The Netherlands.
[4] Corcino, R. B., Corcino, C. B., Malusay, J. T., and Bercero, G. I. M. R. (2019). The r-Bell numbers and matrices containing non-central Stirling and Lah numbers. *J. Math. Comp. Sci.*, 19:181–191.
[5] de Morgan, A. (1839). On the rule for finding the value of an annuity on three lives. *London, Edinburgh, and Dublin Phil. Mag. J. Sci.*, 15(94):337–339.
[6] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Camb. Phil. Soc.*, 24(2):18–190.
[7] Genest, C. and Kolev, N. (2021). A law of uniform seniority for dependent lives. *Scan. Act. J.*, 2021(8):726–743.
[8] Ghurye, S. G. (1987). Some multivariate lifetime distributions. *Adv. in Appl. Probab.*, 19(1):138–155.
[9] Gompertz, B. (1825). On the nature of function expressive of the law of human mortality and a new mode of determining the value of life contingencies. *Phil. Trans. Roy. Soc. London*, 115:513–583.
[10] Gould, H. W. and Hopper, A. T. (1962). Operational formulas connected with two generalizations of Hermite polynomials. *Duke Math. J.*, 29:51–63.
[11] Gumbel, E. J. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.*, 55(292):698–707.
[12] Kolev, N., Ngoc, N., and Ju, Y. T. (2017). Bivariate Teissier distributions. In: *Analytical and Computational Methods in Probability Theory* (Rykov, V., Singpurwalla, N., and Zubkov, A., eds), ACMPT 2017. Lecture Notes in Computer Science, vol. 10684. Springer, Cham.
[13] Laurent, A. G. (1975). Failure and mortality from wear and aging: The Teissier model. In: *Statistical Distributions in Scientific Work — Model Building and Model Selection*, vol. 2 (Patil, G., Kotz, S., and Ord, H., eds.), pp. 301–320.

[14] Marshall, A. W. (1975). Some comments on the hazard gradient. *Stoch. Proc. Appl.*, 3(3):293–300.

[15] Marshall, A. W. and Olkin, I. (2007). *Life Distributions*. Springer, New York.

[16] Mezo, I. (2011). The *r*-Bell numbers. *J. Integer Seq. 14*, paper 11.1.1, 14 pp.

[17] Ricci, P. E., Srivastava, R., and Natalini, P. (2021). A family of the *r*-associated Stirling numbers of the second kind and generalized Bernoulli polynomials. *Axioms*, 2021(10):219.

[18] Rządkowski, G., Rządkowski, W., and Wójcicki, P. (2015). On some connections between the Gompertz function and special numbers. *J. Nonlinear Math. Phys.*, 22(3):374–380.

[19] Sloane, N. (2016). *The On-Line Encyclopedia of Integer Sequences*. Published electronically at `http://oeis.org/`.

[20] Teissier, G. (1934). Recherches sur le vieillissement et sur les lois de la mortalité. *Ann. Physiol. Physicoch. Biol.*, 10(2):237–284.

POWER GENERALIZED DUS TRANSFORMATION IN WEIBULL AND LOMAX DISTRIBUTIONS

Beenu Thomas and V. M. Chacko

•

St.Thomas College (Autonomous), Thrissur
beenuneel.18@gmail.com, chackovm@gmail.com

Abstract

A strong need for an appropriate lifetime model arises in reliability analysis. A large number of lifetime distributions are available in the literature. To analyze reliability data, a more suitable lifetime distribution is plausible. Power Generalized DUS (PGDUS) transformation of the lifetime model gives a solution to fit the data with more precision. PGDUS transformation of the exponential distribution is the first attempt in this regard. This new class of distributions can be used for model series systems in which the components are distributed as DUS transformations of some lifetime model. This paper introduces two novel classes of distributions using PGDUS transformation, which is a generalization of DUS transformation, with Weibull and Lomax distributions as the baseline distributions. Some analytical properties like moments, moment generating function, characteristic function, cumulant generating function, quantile function, distribution of order statistics, and Rényi entropy are derived. The maximum likelihood estimation procedure is employed to estimate the unknown parameters. Moreover, a simulation study has been conducted, and data has been analyzed for each of the proposed distributions to demonstrate how well the distributions would perform in a real-life situation. In comparison with some other recent new models, the proposed distribution is found to be a better model.

Keywords: PGDUS transformation, Weibull Distribution, Moments, Lomax Distribution, Maximum likelihood estimator, Simulation

1. Introduction

A large number of distributions are available in the literature to model monotone failure rate data. The Weibull distribution is confined to model data that exhibit monotone failure rate behavior. Due to the inability to handle non-monotone failure rate behavior, various modifications and generalizations are made to the existing Weibull distribution. The generalized Weibull distribution is widely applied in survival analysis and reliability engineering due to its simplicity and relative flexibility. Xie and Lai [24] introduced an additive Weibull model by adding two Weibull survival functions having a bathtub-shaped failure rate function. Theoretical investigations of the exponentiated Weibull family were carried out by Mudholkar and Srivastava ([18], [19]). Bagdonavicius and Nikulin [3] proposed a power-generalized Weibull distribution as an extension of the Weibull distribution. Xie et al. [25] proposed a modified Weibull bathtub-shaped failure rate distribution.

The Lomax distribution (also called Pareto-II distribution) is a heavily skewed probability distribution that plays an imperative role in the analysis of lifetime data sets in business, actuarial science, computer science, queueing theory, Internet traffic modeling, economics, income and wealth inequality, and reliability modeling. A few generalizations and extensions of the Lomax distribution can be seen in the literature, such as the Marshall-Olkin extended Lomax distribution (Ghitany et al. [9], Gupta et al. [10]), exponentiated Lomax distribution (Abdul-Moniem and Abdel-Hameed [1]), Beta-Lomax distribution (BL), Kumaraswamy Lomax distribution and McDonald-Lomax distribution (Lemonte and Cordeiro [15]), Gamma-Lomax distribution (Cordeiro et al. [6]), transmuted Weibull Lomax distribution (Afify et al. [2]) and the generalized transmuted Lomax distribution (Nofal et al. [20]).

With application to survival data analysis, Kumar et al. (2015) proposed a method, called DUS transformation, for getting a new distribution based on exponential baseline distributions. In terms of computation and interpretation, this transformation produces a parsimonious result since it does not include any new parameters other than those involved in the baseline distribution. In the case where F(x) is the CDF of the baseline distribution, then the CDF of the new DUS transformed distribution is as follows:

$$G(x) = \frac{1}{e-1}[e^{F(x)} - 1]$$

Maurya et al. [17] introduced the DUS transformation of the Lindley distribution. Tripathi et al. [23] studied the DUS transformation of an exponential distribution and its inference based on the upper record values. Recent studies using the DUS transformation can be seen in the works of Deepthi and Chacko [7], Kavya and Manoharan [11], and Gauthami and Chacko [8]. Recently, Thomas and Chacko [22] introduced an exponentiated generalization of the DUS transformation called the power generalized DUS transformation. When researchers deal with series systems with components distributed as DUS-transformed lifetime distributions, the PGDUS transformation is highly useful. So the investigation of the PGDUS transformation of various lifetime distributions is relevant in the sense of the selection of appropriate lifetime models for series systems.

The main goal of this study is to introduce two novel distributions using the power generalized DUS (PGDUS) transformation. Let X be a random variable with baseline cumulative distribution function (CDF) G(x) and corresponding probability density function (PDF) g(x). Then the CDF of the proposed PGDUS distribution is defined as:

$$F(x) = \left(\frac{e^{G(x)} - 1}{e-1}\right)^{\theta}, \theta > 0, x > 0. \quad (1)$$

and the PDF is,

$$f(x) = \frac{\theta}{(e-1)^{\theta}}(e^{G(x)} - 1)^{\theta-1} e^{G(x)} g(x), \theta > 0, x > 0. \quad (2)$$

The survival function is,

$$R(x) = 1 - \left(\frac{e^{G(x)} - 1}{e-1}\right)^{\theta}, \theta > 0, x > 0.$$

The failure rate function is,

$$h(x) = \frac{\theta g(x) e^{G(x)} (e^{G(x)} - 1)^{\theta-1}}{(e-1)^{\theta} - (e^{G(x)} - 1)^{\theta}}, \theta > 0, x > 0.$$

The paper is organized as follows. In Section 2, the distribution based on PGDUS transformation with Weibull distribution as baseline distribution is proposed. Moments, moment generating function, characteristic function, cumulant generating function, quantile function, distribution of order statistics, and Rényi entropy are derived. Parameter estimation based on the maximum likelihood method, simulation study, and real data application are also discussed. In section 3, a different distribution using the Lomax distribution as the baseline distribution in the PGDUS transformation is proposed. As in section 2, properties of PGDUS transformation of Lomax distribution are derived. Parameter estimation using the maximum likelihood method, simulation study, and real data application are also discussed. Finally, concluding remarks are given in Section 4.

2. PGDUS Weibull Distribution

In this section, the Weibull distribution is used as the baseline distribution for PGDUS transformation and investigated the distributional properties. The CDF of Weibull distribution with parameters α and β is

$$G(x) = 1 - e^{-(x\beta)^{\alpha}}, \alpha, \beta > 0, x > 0. \tag{3}$$

and corresponding PDF is

$$g(x) = \alpha\beta(x\beta)^{\alpha-1}e^{-(x\beta)^{\alpha}}, \alpha, \beta > 0, x > 0 \tag{4}$$

Using Eq.3 in Eq.1, the CDF of PGDUS transformation of Weibull distribution is obtained as

$$F(x) = \left(\frac{e^{1-e^{-(x\beta)^{\alpha}}} - 1}{e - 1}\right)^{\theta}, \alpha, \beta > 0, \theta > 0, x > 0. \tag{5}$$

and the corresponding PDF is given as

$$f(x) = \frac{\theta\alpha\beta^{\alpha}}{(e-1)^{\theta}} x^{\alpha-1}(e^{1-e^{-(x\beta)^{\alpha}}} - 1)^{\theta-1} e^{1-(x\beta)^{\alpha} - e^{-(x\beta)^{\alpha}}}, \alpha, \beta, \theta > 0, x > 0. \tag{6}$$

Then, the failure rate function associated to Eq.5 and Eq.6 is,

$$h(x) = \frac{\theta\alpha\beta^{\alpha}x^{\alpha-1}(e^{1-e^{-(x\beta)^{\alpha}}} - 1)^{\theta-1} e^{1-(x\beta)^{\alpha} - e^{-(x\beta)^{\alpha}}}}{(e-1)^{\theta} - (e^{1-e^{-(x\beta)^{\alpha}}} - 1)^{\theta}}, \alpha, \beta, \theta > 0, x > 0. \tag{7}$$

The distribution with CDF Eq.5 and PDF Eq.6 is referred to as Power generalized DUS Weibull distribution with parameters α, β and θ and is denoted as $PGDUSW(\alpha, \beta, \theta)$. Figures 1 and 2 provide the graphical representation of the PDF and failure rate function respectively for various parameter values

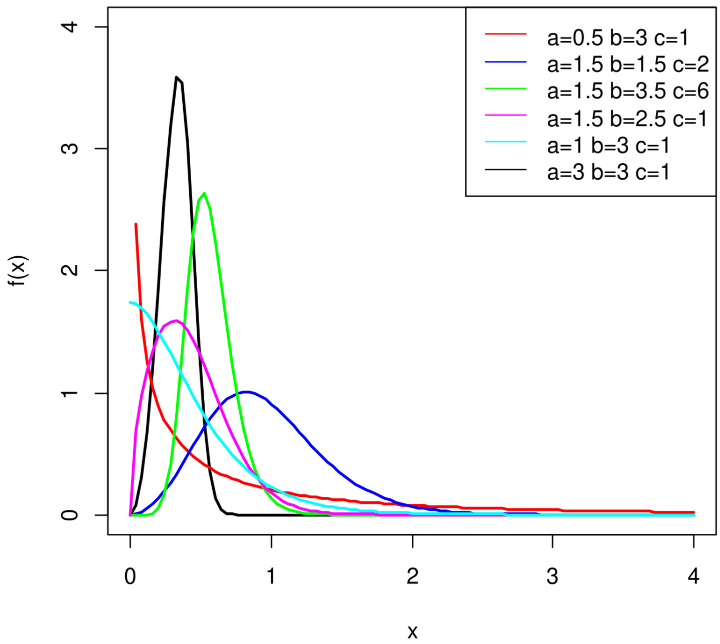


Figure 1: *PGDUSW distribution density plot for various parameter values.*

2.1. Analytical Properties

Moments, moment generating function (MGF), characteristic function (CF), cumulant generating function (CGF), quantile function, distribution of order statistics, and Rényi entropy of the proposed $PGDUSW(\alpha, \beta, \theta)$ distribution are derived.

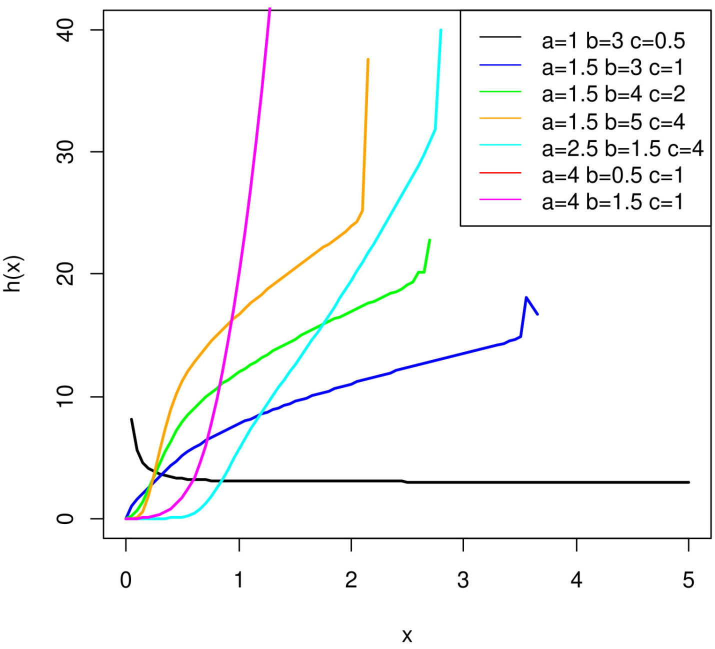


Figure 2: *PGDUSW distribution failure rate plot for various parameter values.*

2.1.1 Moments

The r^{th} raw moment of the $PGDUSW(\alpha,\beta,\theta)$ distribution is given by

$$\mu'_r = \int_0^\infty x^r \frac{\theta\alpha\beta^\alpha x^{\alpha-1}}{(e-1)^\theta}(e^{1-e^{-(x\beta)^\alpha}}-1)^{\theta-1}e^{1-(x\beta)^\alpha-e^{-(x\beta)^\alpha}}dx$$

$$= \frac{\theta\alpha\beta^\alpha e}{(e-1)^\theta}\int_0^\infty x^{r+\alpha-1}e^{-(x\beta)^\alpha}e^{-e^{-(x\beta)^\alpha}}\sum_{k=0}^\infty \binom{\theta-1}{k}(e^{1-e^{-(x\beta)^\alpha}})^{\theta-k-1}(-1)^k dx$$

$$= \frac{\theta\alpha\beta^\alpha e}{(e-1)^\theta}\sum_{k=0}^\infty \binom{\theta-1}{k}(-1)^k e^{\theta-k-1}\int_0^\infty x^{r+\alpha-1}e^{-(x\beta)^\alpha}e^{-(\theta-k)e^{-(x\beta)^\alpha}}dx$$

$$= \frac{\theta\alpha\beta^\alpha e}{(e-1)^\theta}\sum_{k=0}^\infty \binom{\theta-1}{k}(-1)^k e^{\theta-k-1}\sum_{m=0}^\infty \frac{(-1)^m}{m!}(\theta-k)^m\int_0^\infty x^{r+\alpha-1}e^{-(1+m)(x\beta)^\alpha}dx$$

$$= \frac{\theta\beta^{-r}e}{(e-1)^\theta}\sum_{k=0}^\infty\sum_{m=0}^\infty \frac{(-1)^{m+k}}{m!}e^{\theta-k-1}\binom{\theta-1}{k}(\theta-k)^m\frac{\Gamma(\frac{r}{\alpha}+1)}{(1+m)^{\frac{r}{\alpha}+1}}$$

2.1.2 Moment Generating Function

The MGF of $PGDUSW(\alpha,\beta,\theta)$ distribution is

$$M_X(t) = \frac{\theta\alpha\beta^\alpha e}{(e-1)^\theta}\int_0^\infty x^{\alpha-1}e^{tx}e^{-(x\beta)^\alpha}e^{-e^{-(x\beta)^\alpha}}(e^{1-e^{-(x\beta)^\alpha}}-1)^{\theta-1}dx$$

$$= \frac{\theta\alpha\beta^\alpha e}{(e-1)^\theta} \int_0^\infty x^{\alpha-1} e^{tx} e^{-(x\beta)^\alpha} e^{-e^{-(x\beta)^\alpha}} \sum_{k=0}^{\infty} \binom{\theta-1}{k} (-1)^k (e^{1-e^{-(x\beta)^\alpha}})^{\theta-k-1} dx$$

$$= \frac{\theta\alpha\beta^\alpha e}{(e-1)^\theta} \sum_{k=0}^{\infty} \binom{\theta-1}{k} (-1)^k e^{\theta-k-1} \int_0^\infty x^{\alpha-1} e^{tx} e^{-(x\beta)^\alpha} e^{-(\theta-k)e^{-(x\beta)^\alpha}} dx$$

$$= \frac{\theta\alpha\beta^\alpha}{(e-1)^\theta} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{k+m}}{m!} \binom{\theta-1}{k} e^{\theta-k} (\theta-k)^m \int_0^\infty x^{\alpha-1} e^{tx} e^{-(1+m)(x\beta)^\alpha} dx$$

$$= \frac{\theta\alpha\beta^\alpha}{(e-1)^\theta} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^{k+m+n}}{m!n!} \binom{\theta-1}{k} e^{\theta-k} (\theta-k)^m (1+m)^n \beta^{\alpha n} \int_0^\infty x^{\alpha+\alpha n-1} e^{tx} dx$$

$$= \frac{\theta\alpha}{(e-1)^\theta} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^{k+m+n}}{m!n!} \binom{\theta-1}{k} e^{\theta-k} (\theta-k)^m (1+m)^n \beta^{\alpha+\alpha n} \frac{\Gamma(\alpha+\alpha n)}{t^{\alpha+\alpha n}}$$

2.1.3 Characteristic Function and Cumulant Generating Function

The CF of $PGDUSW(\alpha, \beta, \theta)$ is given by

$$\phi_X(t) = \frac{\theta\alpha}{(e-1)^\theta} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^{k+m+n}}{m!n!} \binom{\theta-1}{k} e^{\theta-k} (\theta-k)^m (1+m)^n \beta^{\alpha+\alpha n} \frac{\Gamma(\alpha+\alpha n)}{(it)^{\alpha+\alpha n}},$$

and the CGF of $PGDUSW(\alpha, \beta, \theta)$ is given by

$$\begin{aligned} K_X(t) &= \log \phi_X(t) \\ &= \log \left[\frac{\theta\alpha}{(e-1)^\theta} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^{k+m+n}}{m!n!} \binom{\theta-1}{k} e^{\theta-k} (\theta-k)^m (1+m)^n \beta^{\alpha+\alpha n} \frac{\Gamma(\alpha+\alpha n)}{(it)^{\alpha+\alpha n}} \right] \end{aligned}$$

where $i = \sqrt{-1}$ is the unit imaginary number.

2.1.4 Quantile Function

The p^{th} quantile $Q(p)$ of the $PGDUSW(\alpha, \beta, \theta)$ is the real solution of the following equation

$$((e^{1-e^{-(\beta Q(p))^\alpha}} - 1)/(e-1))^\theta = p$$

where $p \sim Uniform(0, 1)$. Solving the above equation for $Q(p)$, we have

$$Q(p) = \frac{-1}{\beta^\alpha} \log[1 - \log{(e-1)} p^{\frac{1}{\theta}} + 1]^{\frac{1}{\alpha}}. \tag{8}$$

The median is obtained by setting $p = 0.5$ in the Eq.8. Thus,

$$Median = \frac{-1}{\beta^\alpha} \log[1 - \log{(e-1)} 0.5^{\frac{1}{\theta}} + 1]^{\frac{1}{\alpha}}.$$

Similarly, the quartiles Q_1 and Q_3 are obtained respectively by setting $p = 0.25$ and $p = 0.75$ in Eq.8.

2.1.5 Distribution of Order Statistic

Let $X_1, X_2, \ldots, X_m$ be m independent random variables from the $PGDUSW(\alpha, \beta, \theta)$ distribution with CDF Eq.5 and PDF Eq.6. Then the PDF of r^{th} order statistics $X_{(r)}$ of the $PGDUSW(\alpha, \beta, \theta)$ distribution is given by

$$f_{X_{(r)}} = \frac{m!}{(r-1)!(m-r)!} \frac{\theta\alpha\beta^\alpha x^{\alpha-1}}{(e-1)^{\theta m}} \left(e^{1-e^{-(x\beta)^\alpha}} - 1\right)^{\theta r-1} e^{1-(x\beta)^\alpha - e^{-(x\beta)^\alpha}}$$
$$\left[(e-1)^\theta - (e^{1-e^{-(x\beta)^\alpha}})^\theta\right]^{m-r}, r = 1, 2, \ldots, m. \tag{9}$$

Then, the PDF of $X_{(1)}$ and $X_{(m)}$ are obtained by setting $r = 1$ and $r = m$ respectively in Eq.9. This can be used to analyze the reliability of serial and parallel systems.

2.1.6 Rényi Entropy

Rényi entropy introduced by Rényi [21] is defined as

$$\daleth_R(\nu) = \frac{1}{1-\nu} \log \left(\int f^\nu(x) dx \right)$$

where $\nu > 0$ and $\nu \neq 1$.

$$\int_0^\infty f^\nu(x) dx = \frac{(\theta\alpha\beta^\alpha e)^\nu}{(e-1)^{\theta\nu}} \int_0^\infty x^{\nu\alpha-\nu} e^{-\nu(x\beta)^\alpha} e^{-\nu e^{-(x\beta)^\alpha}} \sum_{k=0}^\infty \binom{\nu\theta-\nu}{k} (-1)^k (e^{1-e^{-(x\beta)^\alpha}})^{\nu\theta-\nu-k} dx$$

$$= \frac{(\theta\alpha\beta^\alpha e)^\nu}{(e-1)^{\theta\nu}} \sum_{k=0}^\infty \sum_{m=0}^\infty \frac{(-1)^{k+m}}{m!} \binom{\nu\theta-\nu}{k} (\nu\theta-k)^m e^{\nu\theta-\nu-k} \int_0^\infty x^{\nu\alpha-\nu} e^{-(\nu+m)(x\beta)^\alpha} dx$$

$$= \frac{(\theta\alpha)^\nu}{(e-1)^{\theta\nu}} \sum_{k=0}^\infty \sum_{m=0}^\infty \frac{(-1)^{k+m}}{m!} \binom{\nu\theta-\nu}{k} (\nu\theta-k)^m e^{\nu\theta-k} \frac{\Gamma(\nu - \frac{\nu}{\alpha} + 1)}{(\nu+m)^{\nu-\frac{\nu}{\alpha}+1} \beta^{\alpha-\nu}}$$

Then the Rényi entropy becomes

$$\daleth_R(\nu) = \frac{1}{1-\nu} \log \left[\frac{(\theta\alpha)^\nu}{(e-1)^{\theta\nu}} \sum_{k=0}^\infty \sum_{m=0}^\infty \frac{(-1)^{k+m}}{m!} \binom{\nu\theta-\nu}{k} (\nu\theta-k)^m e^{\nu\theta-k} \frac{\Gamma(\nu - \frac{\nu}{\alpha} + 1)}{(\nu+m)^{\nu-\frac{\nu}{\alpha}+1} \beta^{\alpha-\nu}} \right]$$

2.2. Estimation

The method of Maximum likelihood estimation is used to estimate the unknown parameters of the $PGDUSW(\alpha, \beta, \theta)$. For this, consider a random sample of size n from $PGDUSW(\alpha, \beta, \theta)$ distribution. Therefore, the likelihood function is given by,

$$L(\alpha, \beta, \theta | x) = \prod_{i=1}^n f(x) = \prod_{i=1}^n \frac{\theta\alpha\beta^\alpha}{(e-1)^\theta} x^{\alpha-1} e^{1-(x_i\beta)^\alpha - e^{-(x_i\beta)^\alpha}} (e^{1-e^{-(x_i\beta)^\alpha}} - 1)^{\theta-1} \tag{10}$$

Applying the natural logarithm to Eq.10, the log-likelihood function becomes

$$\log L = n\log(\theta) + n\log(\alpha) + \alpha n \log(\beta) - \theta n \log(e-1) + n + \sum_{i=0}^n (\alpha-1)\log(x_i)$$
$$- \sum_{i=0}^n (x_i\beta)^\alpha - \sum_{i=0}^n e^{-(x_i\beta)^\alpha} + (\theta-1) \sum_{i=0}^n \log(e^{1-e^{-(x_i\beta)^\alpha}} - 1).$$

Computing the first order partial derivatives, we get

$$\begin{aligned}\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=0}^{n}(x_i\beta)^{\alpha}\log(x_i\beta) + \sum_{i=0}^{n}\log(x_i) + \sum_{i=0}^{n}(x_i\beta)^{\alpha}e^{-(x_i\beta)^{\alpha}}\log(x_i\beta) \\ +n\log(\beta) + \frac{(\theta-1)(x_i\beta)^{\alpha}}{(e^{1-e^{-(x_i\beta)^{\alpha}}}-1)}\log(x_i\beta)e^{1-(x_i\beta)^{\alpha}-e^{-(x_i\beta)^{\alpha}}},\end{aligned} \tag{11}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \beta} = \frac{n\alpha}{\beta} - \sum_{i=0}^{n}\frac{\alpha(x_i\beta)^{\alpha}}{\beta} + \sum_{i=0}^{n}\frac{\alpha(x_i\beta)^{\alpha}}{\beta}e^{-(x_i\beta)^{\alpha}} \\ +(\theta-1)\sum_{i=0}^{n}\frac{\alpha}{\beta}(x_i\beta)^{\alpha}\frac{e^{1-(x_i\beta)^{\alpha}-e^{-(x_i\beta)^{\alpha}}}}{(e^{1-e^{-(x_i\beta)^{\alpha}}}-1)},\end{aligned} \tag{12}$$

and

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} - n\log(e-1) + \sum_{i=0}^{n}\log(e^{1-e^{-(x_i\beta)^{\alpha}}}-1). \tag{13}$$

Equations 11, 12, and 13 are not in closed form. The solution of these explicit equations can be obtained analytically and can be solved numerically using R software by taking arbitrary initial values.

2.3. Simulation Study

In order to illustrate the performance of the maximum likelihood method for $PGDUSW(\alpha,\beta,\theta)$ distribution, the inversion transformation method is used. For different values of α,β, and θ, samples of sizes n = 100, 250, 500, 750, and 1000 are generated from the $PGDUSW(\alpha,\beta,\theta)$ model. For 1000 repetitions, the bias and mean square error (MSE) of the estimated parameters are computed. The selected parameter values are $\alpha = 0.5, \beta = 0.5$ and $\theta = 0.5$, $\alpha = 0.5, \beta = 1$ and $\theta = 0.5$ and $\alpha = 1, \beta = 1$ and $\theta = 0.5$. From Tables 1, 2, and 3, it is noted that bias and MSE decrease for the selected parameter values as sample size increases.

Table 1: *Estimate, Biases and MSEs for PGDUSW model at* $\alpha = 0.5, \beta = 0.5$ *and* $\theta = 0.5$

n	Estimated value of Parameters	Bias	MSE
	$\hat{\alpha}$=0.5668	0.0668	0.0473
100	$\hat{\beta}$=0.7541	0.2541	1.0617
	$\hat{\theta}$=0.5021	0.0031	0.0413
	$\hat{\alpha}$=0.5251	0.0251	0.0118
250	$\hat{\beta}$=0.5832	0.0832	0.1488
	$\hat{\theta}$=0.5032	0.0022	0.0165
	$\hat{\alpha}$=0.5297	0.0189	0.0057
500	$\hat{\beta}$=0.4929	0.0177	0.0318
	$\hat{\theta}$=0.4922	0.0007	0.0068
	$\hat{\alpha}$=0.5188	0.0188	0.0034
750	$\hat{\beta}$=0.4936	-0.0065	0.0223
	$\hat{\theta}$=0.5026	0.0003	0.0050
	$\hat{\alpha}$=0.5165	0.0165	0.0025
1000	$\hat{\beta}$=0.4795	-0.0205	0.0160
	$\hat{\theta}$=0.4922	-0.0078	0.0035

Table 2: *Estimate, Biases and MSEs for PGDUSW model at $\alpha = 0.5, \beta = 1$ and $\theta = 0.5$*

n	Estimated value of Parameters	Bias	MSE
	$\hat{\alpha}$=0.5730	0.0730	0.0460
100	$\hat{\beta}$=1.4827	0.4827	3.7354
	$\hat{\theta}$=0.5134	0.0434	0.0485
	$\hat{\alpha}$=0.5019	0.0019	0.0083
250	$\hat{\beta}$=1.2852	0.2852	0.6372
	$\hat{\theta}$=0.5333	0.0393	0.0169
	$\hat{\alpha}$=0.4943	-0.0057	0.0041
500	$\hat{\beta}$=1.2236	0.2236	0.2915
	$\hat{\theta}$=0.5399	0.0339	0.0102
	$\hat{\alpha}$=0.4886	-0.0109	0.0023
750	$\hat{\beta}$=1.1045	0.1814	0.1353
	$\hat{\theta}$=0.5244	0.0244	0.0050
	$\hat{\alpha}$=0.4822	-0.0178	0.0022
1000	$\hat{\beta}$=1.1814	0.1045	0.1195
	$\hat{\theta}$=0.5207	0.0207	0.0042

Table 3: *Estimate, Biases and MSEs for PGDUSW model at $\alpha = 1, \beta = 1$ and $\theta = 0.5$*

n	Estimated value of Parameters	Bias	MSE
	$\hat{\alpha}$=1.1273	0.1273	0.1628
100	$\hat{\beta}$=1.1460	0.1460	0.8851
	$\hat{\theta}$=0.5223	0.0223	0.0545
	$\hat{\alpha}$=1.0184	0.0184	0.0450
250	$\hat{\beta}$=1.0889	0.0889	0.1068
	$\hat{\theta}$=0.5205	0.0205	0.0177
	$\hat{\alpha}$=1.0109	0.0109	0.0185
500	$\hat{\beta}$=1.0490	0.0490	0.0447
	$\hat{\theta}$=0.5151	0.0151	0.0085
	$\hat{\alpha}$=1.0056	0.0056	0.0107
750	$\hat{\beta}$=1.0381	0.0381	0.0260
	$\hat{\theta}$=0.5095	0.0095	0.0049
	$\hat{\alpha}$=0.9851	-0.0149	0.0074
1000	$\hat{\beta}$=1.0239	0.0239	0.0167
	$\hat{\theta}$=1.0012	0.0012	0.0035

2.4. Application

A real data analysis is carried out to determine the performance of the proposed model. For this, the data on the number of million revolutions before the failure of 23 ball bearings put on test is considered (Lawless [13]), see Table 4.

Different distributions namely, Inverse Weibull (IW) distribution, DUS Exponential (DUSE) distribution, and Kavya-Manoharan Weibull (KMW) distribution are used to compare the performance with the proposed $PGDUSW(\alpha, \beta, \theta)$ distribution. In order to perform the necessary numerical evaluations, the software R is used.

Table 4: *Lawless Data*

17.88	28.92	33.00	41.52	42.12	45.60
48.80	51.84	51.96	54.12	55.56	67.80
68.64	68.64	68.88	84.12	93.12	98.64
105.12	105.84	127.92	128.04	173.40	

Table 5: *Findings for PGDUSW Distribution*

Model	**MLEs**	log **L**	**AIC**	**CAIC**	**KS**	**p-value**
IW	$\hat{\lambda}$ = 1.8341 $\hat{\theta}$ = 0.0206	-115.7887	235.5774	236.1774	0.1328	0.8118
DUSE	$\hat{a}$ = 0.0182	-127.4622	256.9244	257.1149	0.2774	0.0580
KMW	$\hat{\lambda}$ = 2.3169 $\hat{\kappa}$ = 0.0107	-113.4076	230.8152	231.4152	0.1421	0.7419
PGDUSW	$\hat{\alpha}$ = 0.9362 $\hat{\beta}$ = 0.0383 $\hat{\theta}$ = 4.4478	-113.0114	230.0228	230.6228	0.10921	0.9467

To check the acceptability of the $PGDUSW(\alpha, \beta, \theta)$ distribution for the given data set Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), log-likelihood value, and Kolmogorov-Smirnov goodness of fit test statistic (KS) with the p-value are used and the computed values are provided in Table 5. It is worth noting that in the goodness of fit test, the purpose is to determine whether the sets of data with the distribution function $F(y)$ and the hypothesised distribution $F_{PGDUSW}(y)$ are compatible. This problem can be formulated as $H_0 : F(y) = F_{PGDUSW}(y)$ versus the alternative $H_1 : F(y) = F_{PGDUSW}(y)$.

From Table 5, it is noted that the $PGDUSW(\alpha, \beta, \theta)$ distribution fits well for the given data set. To facilitate a better understanding of the results, the plot of the empirical CDF (ECDF) is shown in the figure along with other CDFs of the distributions for the Lawless dataset. Furthermore, our proposed distribution is found to fit better than those of the other distributions.

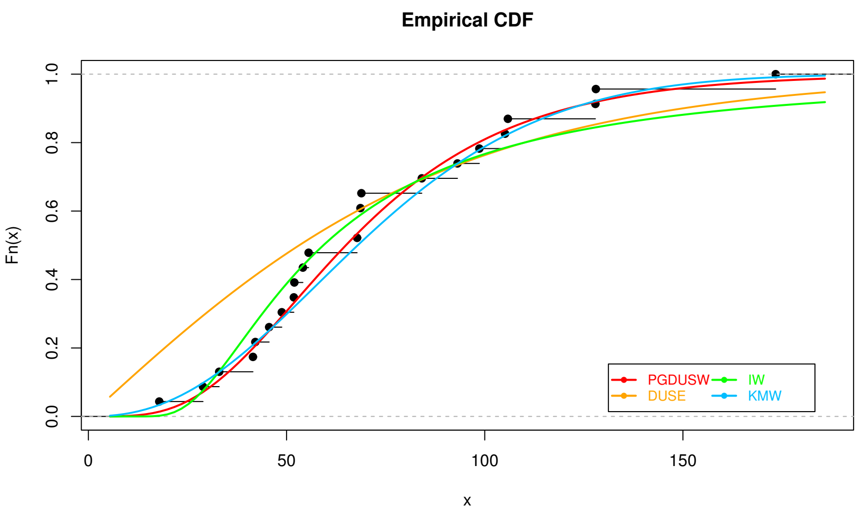


Figure 3: *ECDF plot for various distributions.*

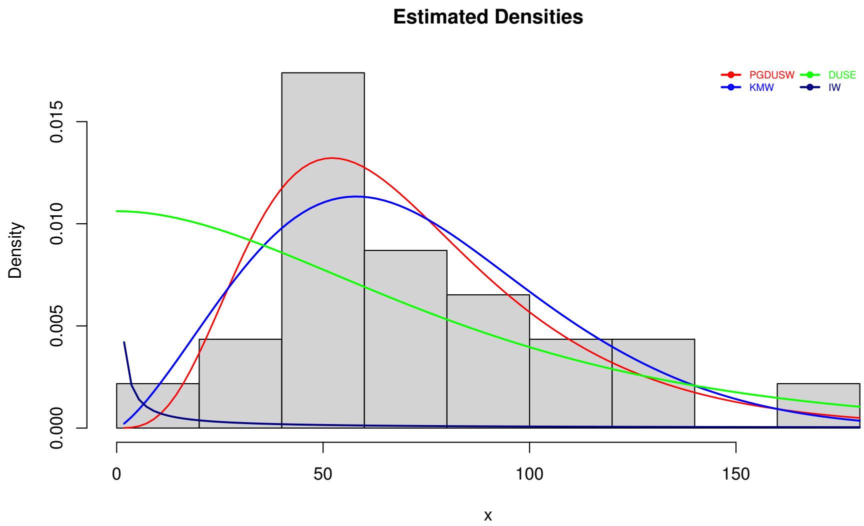


Figure 4: *ECDF plot for various distributions.*

3. PGDUS Lomax Distribution

Power Generalized DUS Lomax distribution denoted as $PGDUSL(\alpha, \beta, \theta)$, is obtained using PGDUS transformation with Lomax distribution as baseline distribution. Then the CDF of the $PGDUSL(\alpha, \beta, \theta)$ distribution using Eq.1 is given by

$$F(x) = \left(\frac{e^{1-(1+x\beta)^{-\alpha}} - 1}{e-1} \right)^{\theta}, \alpha, \beta > 0, \theta > 0, x > 0. \tag{14}$$

Then the PDF is

$$f(x) = \frac{\theta\alpha\beta}{(e-1)^{\theta}} (e^{1-(1+x\beta)^{-\alpha}} - 1)^{\theta-1} e^{1-(1+x\beta)^{-\alpha}} (1+x\beta)^{-(\alpha+1)}. \tag{15}$$

The failure rate function is

$$h(x) = \frac{\theta\alpha\beta(e^{1-(1+x\beta)^{-\alpha}} - 1)^{\theta-1} e^{1-(1+x\beta)^{-\alpha}} (1+x\beta)^{-(\alpha+1)}}{(e-1)^{\theta} - (e^{1-(1+x\beta)^{-\alpha}} - 1)^{\theta}}$$

3.1. Properties of PGDUSL Distribution

Here, we explore a few properties of the PGDUSL distribution.

3.1.1 Moments

The r^{th} raw moments of $PGDUSL(\alpha, \beta, \theta)$ is

$$\mu'_r = \frac{\theta\alpha}{(e-1)^{\theta}} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(-1)^{k+m+n}}{n!} \binom{\alpha+k}{k} \binom{\theta-1}{m} \beta^{k+1} e^{\theta-m} (\theta-m)^n$$
$$B(r+k+1, \alpha n - r - k - 1)$$

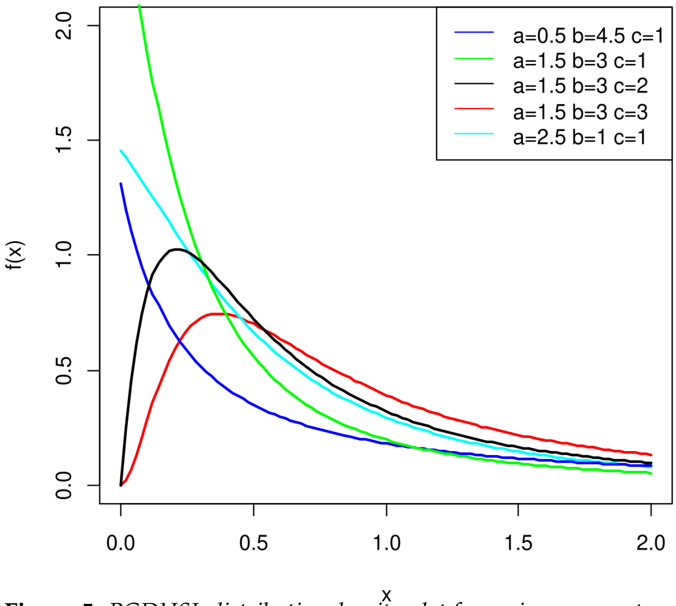


Figure 5: *PGDUSL distribution density plot for various parameter values.*

3.1.2 Quantile Function

The p^{th} quantile $Q(p)$ of the $PGDUSL(\alpha, \beta, \theta)$ is the real solution of the following equation

$$((e^{1-1+(\beta Q(p))^{\alpha}} - 1)/(e-1))^{\theta} = p$$

where $p \sim Uniform(0,1)$. Solving the above equation for $Q(p)$, we have

$$Q(p) = \frac{1}{\beta}\{\left[1 - \log\left[p^{\frac{1}{\theta}}(e-1)+1\right]\right]^{\frac{-1}{\alpha}} - 1\}.$$

The median is obtained by setting $p = 0.5$ in the above equation. Thus,

$$Median = \frac{1}{\beta}\{\left[1 - \log\left[0.5^{\frac{1}{\theta}}(e-1)+1\right]\right]^{\frac{-1}{\alpha}} - 1\}$$

3.2. Estimation of PGDUSL Distribution

The method of maximum likelihood estimation is used to estimate the unknown parameters of $PGDUSL(\alpha, \beta, \theta)$. For this, consider a random sample of size n from $PGDUSL(\alpha, \beta, \theta)$ distribution. Therefore, the likelihood function is given by,

$$L(\alpha, \beta, \theta | x) = \prod_{i=1}^{n} f(x) = \frac{(\theta\alpha\beta)^n}{(e-1)^{\theta n}} \prod_{i=1}^{n} (e^{1-(1+x_i\beta)^{-\alpha}} - 1)^{\theta-1} e^{1-(1+x_i\beta)^{-\alpha}} (1+x_i\beta)^{-\alpha+1} \quad (16)$$

The log-likelihood function becomes

$$\log L = n\log(\theta) + n\log(\alpha) + n\log(\beta) - \theta n \log(e-1) + n - \sum_{i=1}^{n}(1+x_i\beta)^{-\alpha}$$
$$- (\alpha+1)\sum_{i=1}^{n}\log(1+x_i\beta) + (\theta-1)\sum_{i=1}^{n}\log(e^{1-(1+x_i\beta)^{-\alpha}} - 1) \quad (17)$$

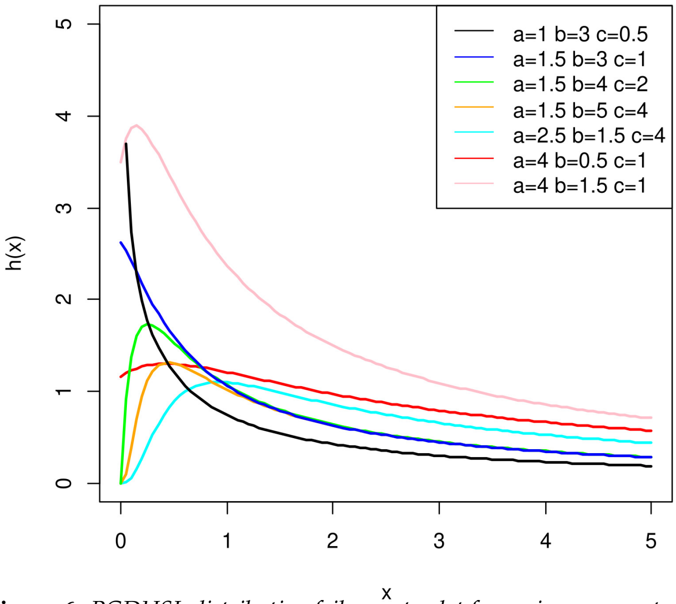


Figure 6: *PGDUSL distribution failure rate plot for various parameter values.*

Computing the first order partial derivatives of Eq.17, we get

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log(1+x_i\beta)(1+x_i\beta)^{-\alpha} - \sum_{i=1}^{n} \log(1+x_i\beta) + \sum_{i=1}^{n} \frac{(\theta-1)\log(1+x_i\beta)e^{1-(1+x_i\beta)^{-\alpha}}(1+x_i\beta)^{-\alpha}}{(e^{1-(1+x_i\beta)^{-\alpha}}-1)}, \tag{18}$$

$$\frac{\partial \log L}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \alpha x_i (1+x_i\beta)^{-(\alpha+1)} - (\alpha+1)\sum_{i=1}^{n} \frac{x_i}{1+x_i\beta} - \sum_{i=1}^{n} \frac{\alpha x_i(\theta-1)(1+x_i\beta)^{-(\alpha+1)}}{(e^{1-(1+x_i\beta)^{-\alpha}}-1)} \tag{19}$$

and

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} - n\log(e-1) + \sum_{i=1}^{n} \log(e^{1-(1+x_i\beta)^{-\alpha}}-1) \tag{20}$$

Equations 18, 19, and 20 are not in closed form. The solution to these explicit equations can be obtained analytically and can be solved numerically using R software by taking arbitrary initial values.

3.3. Simulation Study

In order to illustrate the performance of the maximum likelihood method for $PGDUSL(\alpha, \beta, \theta)$ distribution, the inverse transformation method is used. For different combinations of values of α, β, and θ, samples of sizes n = 250, 500, 750, and 1000 are generated from the $PGDUS-L(\alpha, \beta, \theta)$ model. For 1000 repetitions, the bias and mean square error (MSE) of the estimated parameters are computed. The selected parameter values are $\alpha = 0.5, \beta = 0.5$ and $\theta = 0.5$, $\alpha = 1, \beta = 1.5$ and $\theta = 0.5$ and $\alpha = 1, \beta = 1.5$ and $\theta = 1$. From Tables 6, 7, and 8, it is observed that bias and MSE decrease for the selected parameter values as sample size increases.

Table 6: *Estimate, Biases and MSEs for PGDUSL model at $\alpha = 0.5, \beta = 0.5$ and $\theta = 0.5$*

n	Estimated value of Parameters	Bias	MSE
250	$\hat{\alpha}$=0.5100	0.0100	0.0031
	$\hat{\beta}$=0.5520	0.0720	0.0665
	$\hat{\theta}$=0.5218	0.0218	0.0049
500	$\hat{\alpha}$=0.4921	-0.0039	0.0016
	$\hat{\beta}$=0.5926	0.0526	0.0422
	$\hat{\theta}$=0.5197	0.0197	0.0023
750	$\hat{\alpha}$=0.4960	-0.0079	0.0010
	$\hat{\beta}$=0.5313	0.0343	0.0181
	$\hat{\theta}$=0.5088	0.0088	0.0013
1000	$\hat{\alpha}$=0.4889	-0.0111	0.0008
	$\hat{\beta}$=0.5343	0.0313	0.0134
	$\hat{\theta}$=0.5046	0.0046	0.0009

Table 7: *Estimate, Biases and MSEs for PGDUSL model at $\alpha = 1, \beta = 1.5$ and $\theta = 0.5$*

n	Estimated value of Parameters	Bias	MSE
250	$\hat{\alpha}$=1.0268	0.0268	0.0314
	$\hat{\beta}$=1.6452	0.1800	0.4484
	$\hat{\theta}$=0.5217	0.0217	0.0037
500	$\hat{\alpha}$=1.0140	0.0140	0.0131
	$\hat{\beta}$=1.6800	0.1452	0.2215
	$\hat{\theta}$=0.5187	0.0187	0.0017
750	$\hat{\alpha}$=0.9838	-0.0070	0.0080
	$\hat{\beta}$=1.6374	0.1374	0.1404
	$\hat{\theta}$=0.5040	0.0050	0.0008
1000	$\hat{\alpha}$=0.9930	-0.0162	0.0059
	$\hat{\beta}$=1.6070	0.1070	0.0906
	$\hat{\theta}$=0.5050	0.0040	0.0006

Table 8: *Estimate, Biases and MSEs for PGDUSL model at $\alpha = 1, \beta = 1.5$ and $\theta = 1$*

n	Estimated value of Parameters	Bias	MSE
250	$\hat{\alpha}$=1.0284	0.0284	0.0194
	$\hat{\beta}$=1.69386	0.19386	0.71071
	$\hat{\theta}$=1.05298	0.05297	0.03426
500	$\hat{\alpha}$=1.0179	0.0179	0.0082
	$\hat{\beta}$=1.5999	0.0999	0.1999
	$\hat{\theta}$=1.0472	0.0472	0.0144
750	$\hat{\alpha}$=0.9917	-0.0083	0.0049
	$\hat{\beta}$=1.5596	0.0596	0.1101
	$\hat{\theta}$=1.0145	0.0145	0.0068
1000	$\hat{\alpha}$=0.9836	-0.0164	0.0033
	$\hat{\beta}$=1.5187	0.0187	0.0755
	$\hat{\theta}$=0.9967	-0.0033	0.0051

From Tables 6, 7, and 8, it is observed that bias and MSE are getting closer to zero, as the sample size increases. Therefore, it can be concluded that the proposed model is more consistent and the performance of MLE is highly adequate.

3.4. Real Data Application

Real data analysis is used to determine the applicability of the PGDUSL model. The data set shown in Table 9 is an uncensored data set. As reported by Lee and Wang [14], Table 9 shows the number of months in which 128 bladder cancer patients experienced remission. Different distributions namely, Lomax distribution (LD), DUS Exponential distribution (DUSE), and DUS Lomax distribution (DUSL) are used to compare the performance with the proposed $PGDUSL(\alpha, \beta, \theta)$ distribution. In order to perform the necessary numerical evaluations, the software R is used.

Table 9: ***Remission Times in Months of Blood Cancer Patients***

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23
0.52	4.98	6.97	9.02	13.29	0.40	2.26	3.57	5.06	7.09
0.22	13.80	25.74	0.50	2.46	3.64	5.09	7.26	9.47	14.24
0.82	0.51	2.54	3.70	5.17	7.28	9.74	14.76	26.31	0.81
0.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64	3.88	5.32
0.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
0.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01
0.19	2.75	4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33
0.66	11.25	17.14	79.05	1.35	2.87	5.62	7.87	11.64	17.36
0.40	3.02	4.34	5.71	7.93	11.79	18.10	1.46	4.40	5.85
0.26	11.98	19.13	1.76	3.25	4.50	6.25	8.37	12.02	2.02
0.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76	12.07
0.73	2.07	3.36	6.93	8.65	12.63	22.69	5.49		

To check the acceptability of the $PGDUSL(\alpha, \beta, \theta)$ distribution for the given data set AIC, Consistent AIC (CAIC), log-likelihood value, and KS statistic with the p-value are used and the computed values are provided in Table 10. It is worth noting that in the goodness of fit test, the purpose is to determine whether the sets of data with the distribution function $F(y)$ and the hypothesised distribution $F_{PGDUSL}(y)$ are compatible. This problem can be formulated as $H_{04} : F(y) = F_{PGDUSL}(y)$ versus the alternative $H_{14} : F(y) \neq F_{PGDUSL}(y)$.
Similarly the following hypotheses are tested.
$H_{01} : F(y) = F_{LD}(y)$ Vs $H_{11} : F(y) \neq F_{LD}(y)$
$H_{02} : F(y) = F_{DUSE}(y)$ Vs $H_{12} : F(y) \neq F_{DUSE}(y)$
$H_{03} : F(y) = F_{DUSL}(y)$ Vs $H_{13} : F(y) \neq F_{DUSL}(y)$
From Table 10, it is clear that $PGDUSL(\alpha, \beta, \theta)$ distribution fits well for the given data set. To facilitate a better understanding of the results, the plot of the empirical CDF (ECDF) is shown in the Fig.7 along with other CDFs of the distributions for the blood cancer patients dataset. Also, the plot of fitted densities for the blood cancer patients dataset are given. Furthermore, our proposed distribution is found to fit better than those of the other distributions.

Table 10: *Findings for PGDUSL distribution*

Model	MLEs	log L	AIC	CAIC	KS	p- value
LD	$\hat{\lambda} = 15.2817$ $\hat{\theta} = 0.0074$	-414.98	833.960	834.056	0.094	0.208
DUSE	$\hat{\mu} = 0.1342$	-433.139	868.278	868.309	0.081	0.366
DUSL	$\hat{\lambda} = 6.471$ $\hat{\theta} = 0.0253$	-413.077	830.153	830.249	0.075	0.463
PGDUSL	$\hat{\alpha} = 3.842$ $\hat{\beta} = 0.0605$ $\hat{\theta} = 1.3984$	-411.019	828.039	828.2324	0.035	0.998

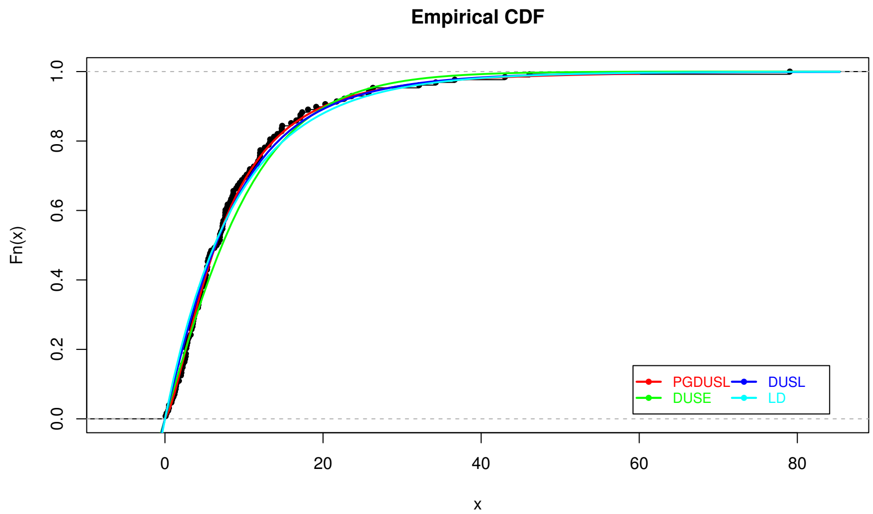


Figure 7: *ECDF plot of the models for blood cancer patients dataset.*

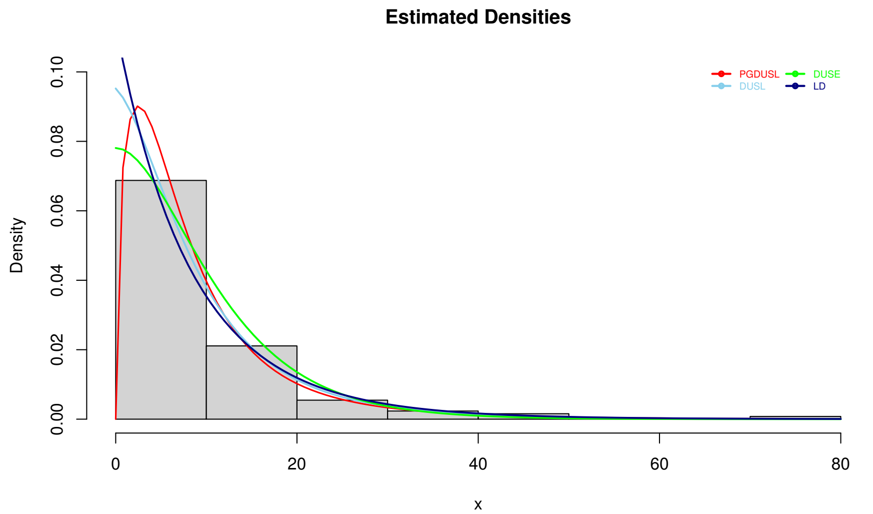


Figure 8: *Estimated densities of the models for the blood cancer patients dataset.*

4. Discussion

This paper proposes the power generalized DUS transformation of Weibull and Lomax distributions. Moments, MGF, CF, CGF, quantile function, distribution of order statistics, and Rényi entropy are derived. The parameter estimation has been done using the maximum likelihood method. By using a simulation study, it is observed that the estimates of the proposed distributions have smaller bias and mean square error when the sample size is larger. Real-world applications have been performed to determine the applicability of the proposed model. Furthermore, the newly developed models are compared with a few existing models, and it is found that the newly developed distributions perform better than the few existing models. When conducting reliability analysis with a series system where each of the components has a specific lifetime distribution, the PGDUS approach is highly useful.

Acknowledgement

The author(s) are thankful to the editor and referee(s) for providing anonymous comments for improving the contents and quality of the manuscript.

References

[1] Abdul-Moniem, I. B. and Abdel-Hameed H. F., (2012). On exponentiated Lomax distribution. *International Journal of Mathematical Archive*, 33(5):1-7.

[2] Afify, A. Z., Nofal Z. M., Yousof H. M., Gebaly Y. M., and Butt N. S. (2015). The transmuted Weibull Lomax distribution: properties and application. *Pakistan Journal of Statistics and Operation Research*, 11:135-152.

[3] Bagdonavicius, V., and M. Nikulin. (2001), Accelerated life models: modeling and statistical analysis: CRC press.

[4] Chen Z.(2002). A new two- parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics and Probability Letters*, 49:155-161.

[5] Cordeiro, G. M., Ortega, E. M. M., and da Cunha, D. C. C.(2013). The exponentiated generalized class of distributions. *Journal of Data Science*, 11:1-27.

[6] Cordeiro, G. M., Ortega, E. M., and Popović, B. V. (2015). The gamma-Lomax Distribution. *Journal of Statistical Computation and Simulation*, 85(2):305-319.

[7] Deepthi, K. S. and Chacko, V. M.(2020). An upside-down bathtub-shaped failure rate model using a DUS transformation of Lomax distribution, Lirong Cui, Ilia Frenkel, raAnatoly Lisnianski (Eds), Stochastic Models in Reliability Engineering, chapter 6:81-100. Taylor & Francis Group, Boca Raton, CRC Press.

[8] Gauthami, P. and Chacko, V. M. (2021). DUS transformation of Inverse Weibull distribution: An upside-down failure rate rate model.*Reliaility : Theory and Applications*, 16, 2(62):58-71.

[9] Ghitany M. E., Al-Awadhi F. A. and Alkhalfan L. A. (2007). Marshall - Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics - Theory and Methods*, 36(10):1855-1866.

[10] Gupta R. C., Ghitany M. E. and Al-Mutairi D. K.(2010). Estimation of reliability from Marshall- Olkin extended Lomax distributions. *Journal of Statistical Computation and Simulation*, 80(8):937-947.

[11] Kavya, P. and Manoharan, M.(2020). On a Generalized lifetime model using DUS transformation, Joshua, V., Varadhan, S., Vishnevsky, V. (Eds), Applied Probability and Stochastic Processes, 281-291. Infosys Science Foundation Series, Springer, Singapore.

[12] Kumar, D., Singh, U., and Singh, S. K.(2015). A method of proposing new distribution and its application to bladder cancer patients data, *Journal of Statistics Applications and Probability Letters*, 2(3):235-245.

[13] Lawless, J. F. Statistical models and methods for lifetime data, John Wiley and Sons, New York, 1982.

[14] Lee E.T. and Wang J.W. Statistical Methods for Survival Data Analysis, 3rd ed.,Wiley, NewYork, 2003.

[15] Lemonte A. J. and Cordeiro G. M. (2013). An extended Lomax distribution. *A Journal of Theoretical and Applied Statistics*, 47(4):800-816.

[16] Lomax K. S. (1987). Business failures: another example of the analysis of failure data. *Journal of the American Statistical Association*, 49:847-852.

[17] Maurya, S. K., Kaushik, A., Singh, S. K., and Singh, U.(2016). A new class of exponential transformed Lindley distribution and its application to Yarn data. *International Journal of Statistics and Economics*, 18(2).

[18] Mudholkar G. S. and Srivastava D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliabilty*, 42(2):299-302.

[19] Mudholkar G. S. and Srivastava D. K. (1995). The exponentiated Weibull family: A reanalysis of the Bus-Motor-failure data. *Technometrics*, 37(4):436-445.

[20] Nofal Z. M., Afify A. Z., Yousof H. M. and Cordeiro G. M. (2017). The generalized transmuted-G family of distributions. *Communication in Statistics- Theory and Methods*, 46(8):355-378.

[21] Rényi A. (1961). On measures of entropy and information. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, 1:547-561. Berkeley: University of California Press.

[22] Thomas, B. and Chacko V. M. (2022). Power generalized DUS transformation of exponential distribution.*International Journal of Statistics and Reliability Engineering*, 9(1):150-157.

[23] Tripathi, A., Singh, U. and Singh, S. K.(2019) Inferences for the DUS-Exponential Distribution Based on upper record values, *Annals of Data Science*, 8:387-403.

[24] Xie M. and Lai C. D. (1996). Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. *Reliability Engineering and System Safety*, 52(10):87-93.

[25] Xie M., Tang , Y. and Goh T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering and System Safety*, 76:279-285.

On an extension of the two-parameter Lindley distribution

Jiju Gillariose[1] and Lishamol Tomy[2]

•

[1] Department of Statistics, CHRIST (Deemed to be University), Bangalore, Karnataka– 560029, India
[2] Department of Statistics, Deva Matha College, Kuravilangad, Kerala, 686633, India
[1]jijugillariose@yahoo.com
[2]lishatomy@gmail.com

Abstract

AIM: Lindley distribution has been widely studied in statistical literature because it accommodates several interesting properties. In lifetime data analysis contexts, Lindley distribution gives a good description over exponential distribution. It has been used for analysing copious real data sets, specifically in applications of modeling stress-strength reliability. This paper proposes a new generalized two-parameter Lindley distribution and provides a comprehensive description of its statistical properties such as order statistics, limiting distributions of order statistics, Information theory measures, etc.

METHODS: We study shapes of the probability density and hazard rate functions, quantiles, moments, moment generating function, order statistic, limiting distributions of order statistics, information theory measures, and autoregressive models are among the key characteristics and properties discussed. The two-parameter Lindley distribution is then subjected to statistical analysis. The paper uses methods of maximum likelihood to estimate the parameters of the proposed distribution. The usefulness of the proposed distribution for modeling data is illustrated using a real data set by comparison with other generalizations of the exponential and Lindley distributions and is depicted graphically.

RESULTS/FINDINGS: This paper presents relevant characteristics of the proposed distribution and applications. Based on this study, we found that the proposed model can be used quite effectively to analyzing lifetime data.

CONCLUSIONS: In this article, we proffered a new customized Lindley distribution. The proposed distribution enfolds exponential and Lindley distributions as sub-models. Some properties of this distribution such as quantile function, moments, moment generating function, distributions of order statistics, limiting distributions of order statistics, entropy, and autoregressive time series models are studied. This distribution is found to be the most appropriate model to fit the carbon fibers data compared to other models. Consequently, we propose the MOTL distribution for sketching inscrutable lifetime data sets.

Keywords: Exponential distribution, Generalized family, Lindley distribution, Marshall-Olkin extended distribution, Maximum likelihood estimation

1. Introduction

Lindley distribution [16, 17] has been proposed to describe a difference between fiducial distribution and posterior distribution. The works on Lindley distribution; see, for example, [11], [14], [8], [2], [28], [33], etc. In the last decades, a lot of attempts have been made to define new probability distributions based on Lindley model, for example, three parameters-Lindley distribution [37], generalized Poisson-Lindley distribution [18], generalized Lindley distribution [23], Marshall-Olkin Lindley distribution [38], power Lindley distribution [10], two-parameter Lindley distribution [29], quasi Lindley distribution [30], transmuted Lindley distribution [20], transmuted Lindley-geometric distribution [21], beta-Lindley distribution [22], discrete Harris extended Lindley distribution [35], etc. Moreover, [36] has provided a detailed review study on

the generalizations of the Lindley distribution.
[31] introduced a new distribution, called two-parameter Lindley distribution. A random variable X is said to have the two-parameter Lindley distribution with parameters α and β if its survival function (sf) takes the form

$$\bar{F}(x,\alpha,\beta) = \frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta}e^{-\alpha x}\,,\; x>0,\alpha>0,\beta>-\alpha \tag{1}$$

and the corresponding probability density function (pdf) can be expressed as

$$f(x,\alpha,\beta) = \frac{\alpha^2(1+\beta x)e^{-\alpha x}}{\alpha+\beta}\,,\; x>0,\alpha>0,\beta>-\alpha. \tag{2}$$

It can easily be seen that at $\beta = 1$, the distribution in equation (2) reduces to the Lindley distribution and at $\beta = 0$, equation (2) reduces to the exponential distribution. [12] has also studied this distribution as a new flexible form of exponential distribution is called flexible exponential distribution. Some generalizations and extensions of this flexible exponential distribution are proposed in [34] and [25].

On the other hand, there is a vast amount of statistical literature on methods of introducing new family of distributions. Notable among them are Azzalinis skewed family of distributions [4], exponentiated family of distributions [13], gamma-generated family of distributions [39, 27], Kumaraswamy family of distributions [6], Weibull generalized family of distributions [5], logistic-generated family of distributions [Torabi and Montazari(2014)], Kumaraswamy Marshal-Olkin family of distributions [1] and Marshall-Olkin Kumaraswamy family of distributions [15]. Moreover, [19] has introduced a general method for adding parameter to a baseline distribution, the resulting distribution is called Marshall-Olkin family of distributions, its sf $\bar{G}(x)$ and pdf $g(x)$ are given by the following formulae,

$$\bar{G}(x,\alpha) = \frac{\gamma\bar{F}(x)}{1-\bar{\gamma}\bar{F}(x)}\,,\; x\in R,\gamma>0 \tag{3}$$

$$g(x,\alpha) = \frac{\gamma f(x)}{(1-\bar{\gamma}\bar{F}(x))^2}\,,\; x\in R,\gamma>0 \tag{4}$$

where $\bar{F}(x)$ is sf of the random variable X to be generated, $\bar{\gamma} = 1-\gamma$ and γ is a tilt parameter. If $F(x)$ has the hazard rate function (hrf) $r(x)$ then the hrf of MOE family is given by

$$h(x,\alpha) = \frac{r(x)}{1-\bar{\gamma}\bar{F}(x)}\,,\; x\in R,\gamma>0$$

The main object of this paper is to present an extension for the two-parameter Lindley distribution, that can be used as an alternative to the existing generalized exponential and Lindley distributions. The rest of this article is organized as follows: Section 2 introduces the Marshall-Olkin two-parameter Lindley distribution; its properties including quantile function, moments, moment generating function, distributions of order statistics, limiting distributions of order statistics, entropy and autoregressive time series models are presented in Section 3; Section 4 proposes parameter estimation of the proposed distribution by the method of maximum likelihood estimation; Section 5 deals with the application of the new distribution to a real data set; Section 6 presents the conclusion of the study.

2. Marshall-Olkin two-parameter Lindley distribution

If X is distributed according to equation (2), then the corresponding Marshall-Olkin (MO) generalized form of its sf and pdf using equations (3) and (4) is given by

$$\bar{G}(x,\alpha,\beta,\gamma) = \frac{\frac{\gamma(\alpha+\beta+\alpha\beta x)}{\alpha+\beta}e^{-\alpha x}}{\left[1-\bar{\gamma}\left(\frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta}e^{-\alpha x}\right)\right]}\,,\; x>0,\alpha,\gamma>0,\beta>-\alpha \tag{5}$$

and

$$g(x, \alpha, \beta, \gamma) = \frac{\gamma\alpha^2(1+\beta x)e^{-\alpha x}}{\alpha + \beta\left[1 - \bar{\gamma}\left(\frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta}e^{-\alpha x}\right)\right]^2}, \; x > 0, \alpha, \gamma > 0, \beta > -\alpha \tag{6}$$

respectively. The new distribution given by the pdf equation (6) is called the Marshall-Olkin two-parameter Lindley (MOTL) distribution. In addition, hrf of the MOTL distribution is given by following equation

$$h(x, \alpha, \beta, \gamma) = \frac{\alpha^2(1+\beta x)}{\left\{1 - \bar{\gamma}\left[\frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta}e^{-\alpha x}\right]\right\}(\beta + \alpha + \alpha\beta x)}, \; x > 0, \alpha, \gamma > 0, \beta > -\alpha. \tag{7}$$

Notably, the classical exponential and one-parameter Lindley distributions are special cases of the MOTL distribution. Some distributions that are special cases of MOTL distribution are:

Exponential distribution : when γ=1 and β=0 in equation (6) with pdf

$$g(x, \gamma) = \alpha e^{-\alpha x}$$

One-parameter Lindley distribution : when γ=1 and β=1 in equation (6) with pdf

$$g(x, \alpha, \gamma) = \frac{\alpha^2(1+x)e^{-\alpha x}}{\alpha + 1}$$

MO exponential distribution : when $\beta = 0$ in equation (6) with pdf

$$g(x, \alpha, \gamma) = \frac{\gamma\alpha e^{-\alpha x}}{[1 - \bar{\gamma}(e^{-\alpha x})]^2}$$

MO Lindley distribution: when $\beta = 1$ in equation (6) with pdf

$$g(x, \alpha, \gamma) = \frac{\gamma\alpha^2(1+x)e^{-\alpha x}}{\alpha + 1\left[1 - \bar{\gamma}\frac{(\alpha+1+\alpha x)}{\alpha+1}e^{-\alpha x}\right]^2}$$

The different shapes of the pdf and hrf of the MOTL distribution are displayed in Figure 1 and Figure 2 for selected parameter values. From figures it is clear that the pdf and hrf of MOTL distribution can be increasing, decreasing, upside-down bathtub (unimodal) depending on the values of its parameters.

3. Statistical Properties

In this section, we study the statistical properties for the MOTL distribution.

3.1. Quantiles

The quantile function of the MOTL distribution is given by

$$x = G^{-1}(u) = -\left(\frac{\alpha+\beta}{\alpha\beta}\right) - \frac{1}{\alpha}W_{-1}\left[-\frac{1}{\beta}\left(\frac{u-1}{1-u+u\gamma}\right)(\alpha+\beta)e^{\left(-\frac{\alpha+\beta}{\beta}\right)}\right] \tag{8}$$

where $G(x) = u$ and $0 \leq u \leq 1$. W_{-1} denotes the negative branch of the Lambert W function. Table 1 represents the quantiles for selected values of the parameters of the MOTL distribution using R programming language.

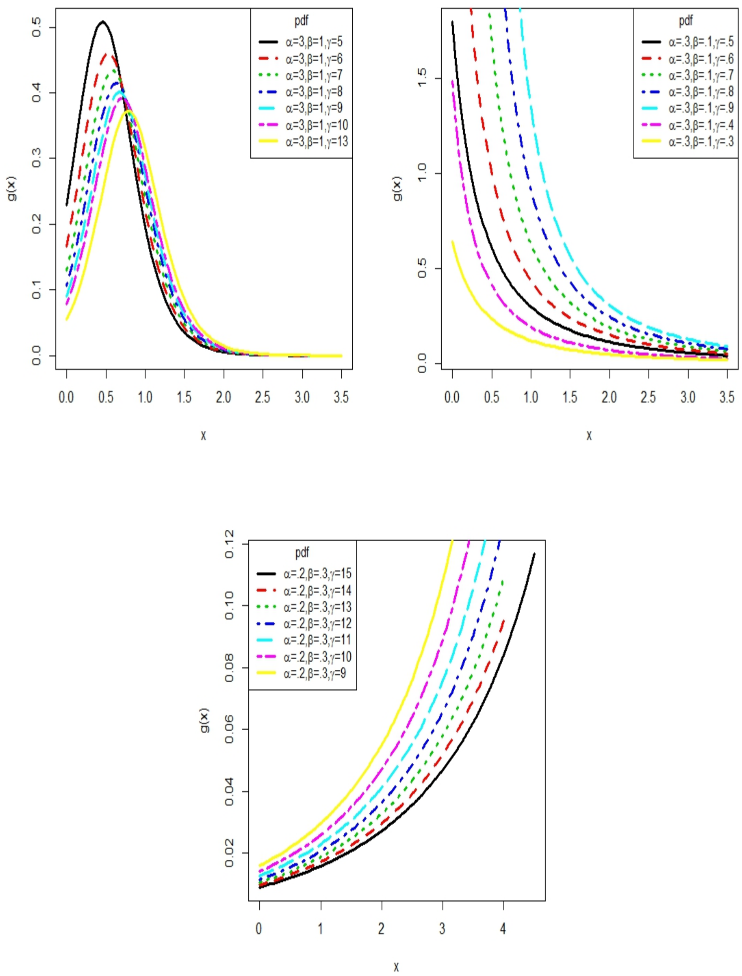


Figure 1: *Graphs of pdf of the MOTL distribution for different values of* α, β *and* γ.

Table 1: *The quantiles for different values of the parameters of the MOTL distribution*

u	(α,β,γ) (0.5,0.5,1.5)	(α,β,γ) (0.4,0.3,.05)	(α,β,γ) (2.1,2.2,0.5)	(α,β,γ) (3.5,4.2,2.5)	(α,β,γ) (5,6,7)
0.1	0.5784	0.2331	0.0513	0.1360	0.1515
0.2	1.1374	0.5001	0.1087	0.2503	0.2596
0.3	1.7018	0.8107	0.1741	0.3563	0.3515
0.4	2.2923	1.1796	0.2504	0.4604	0.4360
0.5	2.9324	1.6295	0.3422	0.5676	0.5202
0.6	3.6556	2.1986	0.4565	0.6838	0.6078
0.7	4.5205	2.9585	0.6074	0.8170	0.7061
0.8	5.6539	4.0696	0.8255	0.9879	0.8282
0.9	7.4520	6.0382	1.2080	1.2505	1.0135

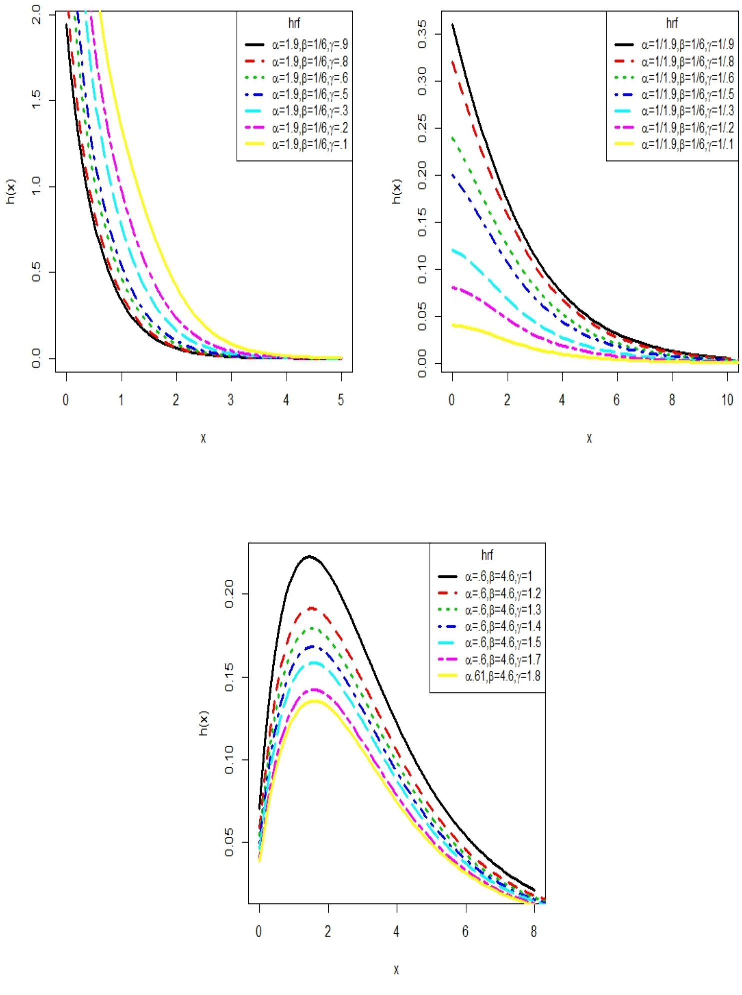


Figure 2: *Graphs of pdf of the MOTL distribution for different values of* α, β *and* γ.

3.2. Moments

In statistical analysis and its applications, moments have received important role. It can be used to study the most eminent features and characteristics such as tendency, dispersion, skewness and kurtosis of a distribution. We now give simple expansions for the pdf of the MOTL distribution. We have following expansion

$$(1-z)^{-r} = \sum_{i=0}^{\infty} \binom{r+i-1}{i} z^i, \; |z| < 1, r > 0$$

Put

$$S(x) = \frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}$$

When $\gamma \in (0,2)$

$$(1-(1-\gamma)(1-S(x))^{-2} = \sum_{i}^{\infty}\sum_{j=0}^{i}(i+1)(1-\gamma)^i \binom{i}{j} S(x)^j \quad (9)$$

Using the series expansion in equation (9) and the representation for the MOTL pdf in equation (6), we obtain

$$\begin{aligned} g(x) &= \gamma \sum_{i=0}^{\infty} \frac{\alpha^2}{\alpha+\beta}(i+1)(1-\gamma)^i \left\{1+\frac{\alpha\beta x}{\alpha+\beta}\right\}^i (1+\beta x)e^{-(i+1)\alpha x} \\ &= \gamma \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\alpha^2}{(\alpha+\beta)}(i+1)(1-\gamma)^i \binom{i}{j} \left\{\frac{\alpha\beta x}{\alpha+\beta}\right\}^j (1+\beta x)e^{-(i+1)\alpha x} \\ &= \gamma \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\alpha^2}{(\alpha+\beta)^{j+1}}(i+1)(1-\gamma)^i \binom{i}{j} (\alpha\beta)^j \left[x^j e^{-(i+1)\alpha x} + \beta x^{j+1} e^{-(i+1)\alpha x}\right] \quad (10) \end{aligned}$$

We have

$$E(X^r) = \int_0^{\infty} x^r g(x,\alpha,\beta,\gamma)dx. \quad (11)$$

Substituting equation (10) into the equation (11), we obtain the r^{th} moment of MOTL distribution in the form

$$\begin{aligned} E(X^r) &= \gamma \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\alpha^2}{(\alpha+\beta)^{j+1}}(i+1)(1-\gamma)^i \binom{i}{j} (\alpha\beta)^j \int_0^{\infty} x^r \left[x^j e^{-(i+1)\alpha x} + \beta x^{j+1} e^{-(i+1)\alpha x}\right] dx \\ &= \gamma \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{\alpha^2}{(\alpha+\beta)^{j+1}}(i+1)(1-\gamma)^i \binom{i}{j} (\alpha\beta)^j w_{ijr} \end{aligned}$$

$$w_{ijr} = \frac{\Gamma(r+j+1)}{[(i+1)\alpha]^{r+j+1}} + \beta \frac{\Gamma(r+j+2)}{[(i+1)\alpha]^{r+j+2}}$$

Similarly, when $\gamma > 1/2$

$$g(x) = \frac{\gamma\alpha^2(1+\beta x)e^{-\alpha x}}{\alpha+\beta}\left[1-\frac{\bar{\gamma}}{\gamma}(1-S(x))\right]^{-2}$$

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

The r^{th} moment of MOTL distribution is

$$E(X^r) = \gamma \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{i} \frac{\alpha^2}{(\alpha+\beta)^{j+1}} (-1)^{i+k} \left(\frac{\bar{\gamma}}{\gamma}\right)^i \binom{i-1}{k} \binom{k}{j} (\alpha\beta)^j w_{ijr}$$

Table 3.2 lists the moments, standard deviation (SD), coefficient of variation (CV), coefficient of skewness (CS) and coefficient of kurtosis (CK) of the MOTL distribution for selected values of the parameters.

Table 2: *Moments of the MOTL distribution for different values of the parameters α, β, and γ*

	(α,β,γ) (0.5,0.5,.5)	(α,β,γ) (1.5,0.5,1.5)	(α,β,γ) (2.5,0.005,3.5)	(α,β,γ) (3,5,3.5)	(α,β,γ) (5,6,7)
μ_1'	1.5273	1.0084	0.7029	0.8856	1.3452
μ_2'	3.0999	1.7819	0.7641	1.0953	2.5024
μ_3'	7.5709	4.3852	1.0910	1.6799	5.7702
μ_4'	21.4836	13.7554	1.9276	3.0565	15.7912
SD	0.8760	0.8746	0.5196	0.5577	0.8326
CV	0.5735	0.8676	0.7392	0.6297	0.6076
CS	3811.034	749.2536	211.0792	548.7451	2284.693
CK	3.8968	6.5565	5.3869	4.2850	4.3509

3.3. Moment Generating Function

Moment generating function is given by the following formula

$$M_X(t) = E(e^{tX}) = \sum_{r=0}^{\infty} \frac{t^r}{r!} E(X^r) \tag{12}$$

The moment generating function of MOTL distribution is obtained by using equation (12). When $\gamma \in (0,2)$, it has following form

$$M_X(t) = \gamma \sum_{r=0}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\alpha^2 t^r}{(\alpha+\beta)^{j+1} r!} (i+1)(1-\gamma)^i \binom{i}{j} (\alpha\beta)^j w_{ijr}$$

Similarly, when $\gamma > 1/2$

$$M_X(t) = \gamma \sum_{r=0}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{i} \frac{\alpha^2 t^r}{(\alpha+\beta)^{j+1} r!} (-1)^{i+j} \left(\frac{1-\gamma}{\gamma}\right)^i \binom{i-1}{j} \binom{k}{j} (\alpha\beta)^j w_{ijr}$$

where

3.4. Order Statistics

Let $X_1, X_2, ..., X_n$ be a random sample taken from the MOTL distribution and $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be the corresponding order statistics. The pdf $g_{r:n}(x, \alpha, \beta, \gamma)$ of the r^{th} order statistics $X_{r:n}$ is given by

$$g_{r:n}(x,\alpha,\beta,\gamma) = \frac{n!}{(r-1)(n-r)!} g(x,\alpha,\beta,\gamma) G(x,\alpha,\beta,\gamma)^{r-1} [1 - G(x,\alpha,\beta,\gamma)]^{n-r} \quad (13)$$

where g(x), G(x) are pdf and cdf of MOTL distribution by equations (5) and (6). We can use the binomial expansion of $[1-G(x)]^{n-i}$ given as follows

$$[1-G(x)]^{n-i} = \sum_{i=0}^{n-r} \binom{n-r}{i} (-1)^i G(x,\alpha,\beta,\gamma)^i \quad (14)$$

substitute equation (15) into the equation (14), we get

$$g_{r:n}(x,\alpha,\beta,\gamma) = \sum_{i=0}^{n-r} \binom{n-r}{i} \frac{n!(-1)^i}{(r-1)(n-r)!} g(x,\alpha,\beta,\gamma) G(\alpha,\beta,\gamma)^{i+r-1} \quad (15)$$

We get pdf of r^{th} order statistics for MOTL distribution from equation (15), by using above equations (5) and (6). We can express the k^{th} ordinary moment of the r^{th} order statistics $X_{r:n}$ ($\mathrm{E}(X_{r:n}^k)$) as a liner combination of the k^{th} moments of the MOTL distribution with different shape parameters. The r^{th} order statistic for MOTL distribution can be expressed as

$$g_{r:n}(x,\alpha,\beta,\gamma) = \sum_{i=0}^{n-r} \binom{n-r}{i} \frac{n!(-1)^i}{(r-1)(n-r)!} \frac{\gamma \frac{\alpha^2(1+\beta x)e^{-\alpha x}}{\alpha+\beta}}{\left[1 - \bar{\gamma} \frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right]^2} \left\{ \frac{\frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)}{\alpha+\beta}}{\left[\gamma + (1-\gamma) \frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right]} \right\}^{i+r-1} \quad (16)$$

When r=1 and r=n in equation (16), we get the equations of pdf of the smallest and largest order statistics, respectively.

3.5. Limiting distributions of order statistics

Theorem: 1. *If $X_{1:n}$ be the minimum and $X_{n:n}$ be the maximum of a random sample $x = (x_1, ..., x_n)$ from MOTL distribution, then*

(a) $\lim_{n\to\infty} P\left(\frac{X_{1:n} - p_n}{q_n} \le x\right) = 1 - e^{-x}$

(b) $\lim_{n\to\infty} P\left(\frac{X_{n:n} - p_n^*}{q_n^*} \le t\right) = \exp(-e^{-t})$

where $p_n = 0$, $q_n = G^{-1}(\frac{1}{n})$, $p_n^* = G^{-1}(1-\frac{1}{n})$, $q_n^* = 1$ and $G^{-1}(\cdot)$ is given in (8).

Proof

For MOTL distribution, by applying L′ Hospital rule, we obtain

$$\lim_{\epsilon\to 0+} \frac{G(\epsilon x)}{G(\epsilon)} = \lim_{\epsilon\to 0+} \frac{x g(\epsilon x)}{g(\epsilon)} = x$$

Hence, the minimal domain of attraction of the MOTL distribution can be the standard exponential distribution (see Theorem 8.3.6 [3]).

Subsequently, for the MOTL distribution, we can express

$$\lim_{x\to\infty+} \frac{d}{dx}\left(\frac{1}{h(x)}\right) = 0$$

Therefore, the maximal domain of attraction of MOLT distribution can be the standard Gumbel distribution (see Theorem 8.3.3 [3]).

3.6. Information Theory Measures

The concept of entropy has played a vital role in information theory. The entropy of a random variable can be defined in terms of its probability distribution and it has been used in distinct fields in science as a measure of variation of the uncertainty. Numerous measures of entropy have been mentioned and compared in the literature. The entropy and mutual information concepts have been formalized by Shannon (1948). [26] proposed a other measure of entropy namely, Rényi entropy as a generalization of Shannon entropy. The Rényi entropy is defined as $I_R(\gamma) = \frac{1}{1-\delta} \log \int_R g^{\delta}(x)dx$, $\delta > 0$ and $\delta \neq 1$. Rényi entropy of order 1 is Shannon entropy. We consider first $g^{\delta}(x)$ given by,

$$g^{\delta}(x) = \frac{\gamma^{\delta} \frac{\alpha^{2\delta}(1+\beta x)^{\delta} e^{-\alpha\delta x}}{\alpha+\beta}^{\delta}}{\left\{1 - \bar{\gamma}\frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right\}^{2\delta}}$$

Suppose that $\gamma > \frac{1}{2}$. Using the series expansion

$$\left[1 - \bar{\gamma}\frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right]^{-2\delta} = \gamma^{-2\delta} \sum_{k=0}^{\infty} \frac{\Gamma(2\delta+k)}{\Gamma(2\delta)k!} \left(1 - \frac{1}{\gamma}\right)^k \left[\frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right]^k$$

Thus we obtain that in the case $\gamma > 1/2$, the Rényi entropy is

$$\begin{aligned} I_R(\delta) &= \frac{1}{1-\delta} \log\{\frac{\gamma^{\delta}\alpha^{2\delta}}{(\alpha+\beta)^{\delta}} \sum_{i,j,k,l=0}^{\infty} \frac{\Gamma(2\delta+k)}{\Gamma(2\delta)k!(\alpha+\beta)^j} \left(1 - \frac{1}{\delta}\right)^k \\ &\quad \binom{j}{k}\binom{i+k}{i}\binom{l}{\delta}(-1)^i(\alpha\beta)^j\beta^{\delta} \int_R x^{(\delta+j)} e^{-\alpha(\delta+k)x} dx\} \\ &= \frac{1}{1-\delta} \log\{\frac{\gamma^{-\delta}\alpha^{2\delta}}{(\alpha+\beta)^{\delta}} \sum_{i,j,k,l=0}^{\infty} \frac{\Gamma(2\delta+k)}{\Gamma(2\delta)k!(\alpha+\beta)^j} \left(1 - \frac{1}{\delta}\right)^k \\ &\quad \binom{j}{k}\binom{i+k}{i}\binom{l}{\delta}(-1)^i(\alpha\beta)^j\beta^{\delta} \frac{\Gamma(\delta+j+1)}{[\alpha(\delta+k)]^{\delta+j+1}} \end{aligned}$$

Similarly, we can show that in the case $0 < \gamma < 2$ and by using the series expansion

$$\left\{1 - \bar{\gamma}\frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)e^{-\alpha x}}{\alpha+\beta}\right\}^{-2\delta} = \sum_{k=0}^{\infty} \frac{\Gamma(2\delta+k)}{\Gamma(2\delta)k!} (1-\gamma)^k \left[\frac{(\beta+\alpha+\alpha\beta x)e^{-\alpha x}}{\alpha+\beta}\right]^k$$

Corresponding Rényi entropy is

$$\begin{aligned} I_R(\delta) &= \frac{1}{1-\delta} \log\{\frac{\gamma^{-\delta}\alpha^{2\delta}}{(\alpha+\beta)^{\delta}} \sum_{i,j,k=0}^{\infty} \frac{\Gamma(2\delta+k)}{\Gamma(2\delta)k!(\alpha+\beta)^i} (1-\delta)^k \\ &\quad \binom{i}{k}\binom{j}{\delta}(\alpha\beta)^i\beta^j \int_R x^{(i+j)} e^{-\alpha(\delta+k)x} dx\} \\ &= \frac{1}{1-\delta} \log\{\frac{\gamma^{-\delta}\alpha^{2\delta}}{(\alpha+\beta)^{\delta}} \sum_{i,j,k=0}^{\infty} \frac{\Gamma(2\delta+k)}{\Gamma(2\delta)k!(\alpha+\beta)^i} (1-\delta)^k \\ &\quad \binom{i}{k}\binom{j}{\delta}(\alpha\beta)^i\beta^j \frac{\Gamma(i+j+1)}{[\alpha(\delta+k)]^{i+j+1}}\} \end{aligned}$$

3.7. Autoregressive Time Series Modeling

Autoregressive models are types of random process, has utilized to model and predict various type of natural phenomena. In other words, autoregressive models are group of linear prediction formulas which try to predict an output of a system based on the past observations. In the following Subsections we construct and explore different autoregressive models of order 1 (AR(1)), that is, MIN AR(1) model I, MIN AR(1) model II, MAX-MIN AR(1) model I and MAX-MIN AR(1) model II with MOTL as marginals.

3.7.1 MIN AR (1) Model-1 with MOTL Marginal Distribution

The first AR (1) structure is given by

$$X_n = \begin{cases} \mathfrak{e}_n, & \text{with probability } \varrho \\ min(X_{n-1}, \mathfrak{e}_n), & \text{with probability } 1-\varrho \end{cases} \tag{17}$$

where $\{\mathfrak{e}_n\}$ is a sequence of independently and identically distributed (iid) random variables independent of $\{X_n\}$ and $\varrho \in$ (0,1). Hence the process is stationary Markovian with MOTL distribution as marginal.

Theorem: 2. *$\{X_n\}$ is stationary Markovian with MOTL distribution with parameters $\varrho, \alpha, \beta \Longleftrightarrow \{\mathfrak{e}_n\}$ is distributed as two-parameter Lindley distribution, under in an AR (1) process defined in (17).*

Proof. Let $\mathfrak{e}_n$ follows two-parameter Lindley distribution with parameters α and β. Using equation (17), we can express

$$\bar{F}_{X_n}(x) = \varrho \bar{F}_{\mathfrak{e}_n}(x) + (1-\varrho)\bar{F}_{X_{n-1}}(x)\bar{F}_{\mathfrak{e}_n}(x)$$

While under stationary equilibrium,

$$\bar{F}_X(x) = \frac{\varrho \bar{F}_{\mathfrak{e}}(x)}{1-(1-\varrho)\bar{F}_{\mathfrak{e}}(x)}$$

and therefore

$$\bar{F}_{\mathfrak{e}}(x) = \frac{\bar{F}_X(x)}{\varrho + (1-\varrho)\bar{F}_X(x)}$$

When $\mathfrak{e}_n$, it follows two-parameter Lindley with parameters α and β

$$\bar{F}_{\mathfrak{e}}(x) = \frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}$$

Hence

$$\bar{F}_X(x) = \frac{\frac{\varrho(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}}{\left[1-\bar{\varrho}\left(\frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right)\right]}$$

It is easy to see that this is the sf of the MOTL(α, β, ϱ).

Conversily, if

$$\bar{F}_X(x) = \frac{\frac{\varrho(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}}{\left[1-\bar{\varrho}\left(\frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right)\right]}$$

then $\bar{F}_{\mathfrak{e}_n}(x)$ is distributed as two-parameter Lindley distribution with parameters α and β, and the process is stationary.

We have to prove its stationarity, so we take that $X_{n-1} \sim$ MOTL(α, β, ϱ) and $\mathfrak{e}_n$ follows two-parameter Lindley distribution with parameters α and β, then

$$\bar{F}_X(x) = \frac{\frac{\varrho(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}}{\left[1 - \bar{\varrho}\left(\frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}\right)\right]}$$

It is easy to see that X_n is distributed as MOTL(α, β, ϱ).

3.7.2 MIN AR (1) Model-II with MOTL Marginal Distribution

The second AR (1) structure is given by

$$X_n = \begin{cases} X_{n-1}, & \text{with probability } \varrho_1 \\ \mathfrak{e}_n, & \text{with probability } \varrho_2 \\ min(X_{n-1}, \mathfrak{e}_n), & \text{with probability } 1 - \varrho_1 - \varrho_2 \end{cases} \quad (18)$$

where ϱ_1, $\varrho_2 > 0$, $\varrho_1 + \varrho_2 < 1$ and $\{\mathfrak{e}_n\}$ is a sequence of iid random variables independent of $\{X_n\}$. This structure allows probabilistic selection of process values, innovations and combinations of both. Then the process is stationary with MOTL as marginal.

Theorem: 3. *$\{X_n\}$ is stationary Markovian with MOTL distribution with parameters γ, α and $\beta \Longleftrightarrow$ $\{\mathfrak{e}_n\}$ is distributed as two-parameter Lindley distribution with parameters α and β, where $\gamma = \frac{\varrho_2}{1-\varrho_1}$, under in an AR (1) process with structure defined in equation (18).*

Proof.

Let $\mathfrak{e}_n$ follows two-parameter Lindley with parameters α and β. By using equation (18),

$$\bar{F}_{X_n}(x) = \varrho_1[1 - (1 - \bar{F}_{X_{n-1}}(x))(1 - \bar{F}_{\mathfrak{e}_n}(x))] + \varrho_2 \bar{F}_{X_{n-1}}(x)\bar{F}_{\mathfrak{e}_n}(x) + (1 - \varrho_1 - \varrho_2)\bar{F}_{X_{n-1}}(x)$$

Under stationary equilibrium it becomes,

$$\bar{F}_X(x) = \frac{\gamma \bar{F}_{\mathfrak{e}}(x)}{1 - (1-\gamma)\bar{F}_{\mathfrak{e}}(x)}$$

where $\gamma = \frac{\varrho_2}{1-\varrho_1}$, it is evident that it has Marshall-Olkin form. Next, we assume that $\{X_n\} \sim$ MOTL(α, β, γ). By using equation (18), under stationarity, we can write

$$\bar{F}_{\mathfrak{e}}(x) = \frac{(1-\varrho_1)\bar{F}_X(x)}{\varrho_2 + (1 - \varrho_1 - \varrho_2)\bar{F}_X(x)}$$

Next by using X_n as MOTL(α, β, γ), it can be obtained as

$$\bar{F}_{\mathfrak{e}}(x) = \frac{(\alpha + \beta + \alpha\beta x)}{\alpha + \beta} e^{-\alpha x}$$

which is the sf of two-parameter Lindley distribution with parameters α and β.

3.7.3 MAX-MIN AR(1) model-I with MOTL Marginal Distribution

Consider now the third model with AR(1) structure

$$X_n = \begin{cases} max(X_{n-1}, \mathfrak{e}_n), & \text{with probability } \varrho_1 \\ min(X_{n-1}, \mathfrak{e}_n), & \text{with probability } \varrho_2 \\ X_{n-1}, & \text{with probability } 1 - \varrho_1 - \varrho_2 \end{cases} \quad (19)$$

where $0 < \varrho_1, \varrho_2 < 1$, $\varrho_2 < \varrho_1$, $\varrho_1 + \varrho_2 < 1$ and $\{\mathfrak{e}_n\}$ is a sequence of iid random variables independent of $\{X_n\}$. Then the process is stationary Markovian with MOTL distribution as marginal.

Theorem: 4. *$\{X_n\}$ is stationary Markovian AR (1) max-MIN process with MOTL distribution with parameters α, β and $\gamma \iff \{\mathfrak{e}_n\}$ is distributed as two-parameter Lindley distribution with parameters α and β, where $\gamma = \frac{\varrho_1}{\varrho_2}$, under an AR (1) MAX-MIN process with structure (19).*

Proof. Let $\mathfrak{e}_n$ follows two-parameter Lindley with parameters α and β. It is obvious from equation (19),

$$\bar{F}_{X_n}(x) = \varrho_1[1-(1-\bar{F}_{X_{n-1}}(x))(1-\bar{F}_{\mathfrak{e}_n}(x))] + \varrho_2 \bar{F}_{X_{n-1}}(x)\bar{F}_{\mathfrak{e}_n}(x) + (1-\varrho_1-\varrho_2)\bar{F}_{X_{n-1}}(x)$$

Under stationary equilibrium,

$$\bar{F}_{X_n}(x) = \frac{\gamma \bar{F}_{\mathfrak{e}}(x)}{1-(1-\gamma)\bar{F}_{\mathfrak{e}}(x)}$$

where $\gamma = \frac{\varrho_1}{\varrho_2}$ and $\bar{F}_{X_n}(x)$ has Marshall-Olkin form. Let $X_n \sim$ MOTL(α, β, γ).
Then by using (19), under stationarity,

$$\bar{F}_{\mathfrak{e}}(x) = \frac{\varrho_2 \bar{F}_{X_n}(x)}{\varrho_1 + (\varrho_2-\varrho_1)\bar{F}_{X_n}(x)}$$

Thus, after simplification it can be written as

$$\bar{F}_{\mathfrak{e}}(x) = \frac{(\alpha+\beta+\alpha\beta x)}{\alpha+\beta} e^{-\alpha x}$$

Consequently, which is the sf of two-parameter Lindley with parameters α and β.

3.7.4 MAX-MIN AR(1) model-II with MOTL Marginal Distribution

Finally, we consider the more general max-min process that includes minimum, maximum innovations and the process. The relating model with AR(1) structure being of the form

$$X_n = \begin{cases} max(X_{n-1}, \mathfrak{e}_n), & \text{with probability } \varrho_1 \\ min(X_{n-1}, \mathfrak{e}_n), & \text{with probability } \varrho_2 \\ \mathfrak{e}_n, & \text{with probability } \varrho_3 \\ X_{n-1}, & \text{with probability } 1-\varrho_1-\varrho_2-\varrho_3 \end{cases} \tag{20}$$

where $0 < \varrho_1, \varrho_2, \varrho_3 < 1$, $\varrho_1+\varrho_2+\varrho_3 < 1$ and $\{\mathfrak{e}_n\}$ is a sequence of iid random variables independent of $\{X_n\}$. Then the process is stationary Markovian with MOTL distribution as marginal.

Theorem: 5. *AR (1) MAX-MIN process $\{X_n\}$ with structure (20) is a stationary Markovian AR (1) MAX-MIN process with MOTL distribution (α, β, γ) $\iff \{\mathfrak{e}_n\}$ is distributed as two-parameter Lindley distribution with parameters α and β where $\gamma = \frac{\varrho_1+\varrho_3}{\varrho_2+\varrho_3}$*

Proof. Let $\mathfrak{e}_n$ follows two-parameter Lindley with parameters α and β. It is clear from equation (20),

$$\begin{aligned} \bar{F}_{X_n}(x) &= \varrho_1[1-(1-\bar{F}_{X_{n-1}}(x))(1-\bar{F}_{\mathfrak{e}_n}(x))] + \varrho_2\bar{F}_{X_{n-1}}(x)\bar{F}_{\mathfrak{e}_n}(x) + \varrho_3\bar{F}_{\mathfrak{e}_n}(x) \\ &+ (1-\varrho_1-\varrho_2-\varrho_3)\bar{F}_{X_{n-1}}(x) \end{aligned}$$

Under stationary equilibrium it gives,

$$\bar{F}_{X_n}(x) = \frac{\gamma \bar{F}_{\mathfrak{e}}(x)}{1-(1-\gamma)\bar{F}_{\mathfrak{e}}(x)}$$

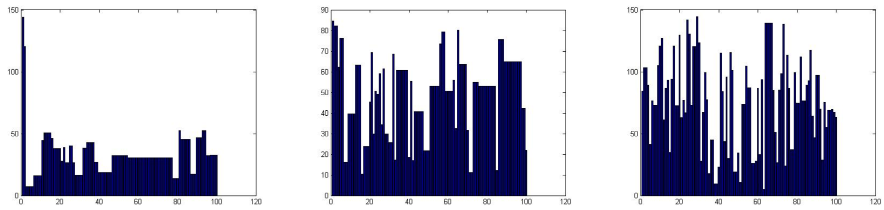


Figure 3: *Graphs of sample path of AR(1) Minification model I for different values of ϱ=0.3,0.5,0.8, α=30 and $\beta = 0.02$*

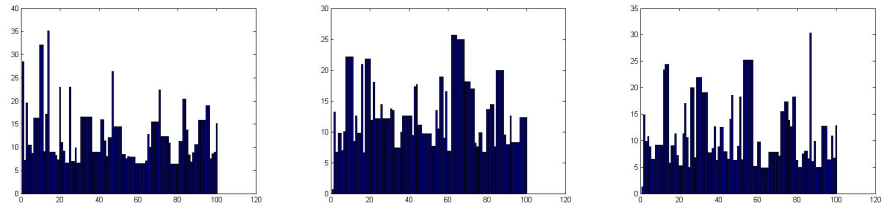


Figure 4: *Graphs of sample path of AR(1) Minification model II for different sets of (ϱ_1, ϱ_2)=(0.1,0.4), (0.4,0.1), (0.4,0.4), α=0.2 and β =0.3*

where $\gamma = \frac{\varrho_1+\varrho_3}{\varrho_2+\varrho_3}$, which has Marshall-Olkin form. Now let $X_n \sim$ MOTL(α, β, γ) and from (20), we obtain

$$\bar{F}_{\mathfrak{e}}(x) = \frac{(\varrho_2 + \varrho_3)\bar{F}_X(x)}{(\varrho_1 + \varrho_3) + (\varrho_2 - \varrho_1)\bar{F}_X(x)}$$

Thus, after simplification it reduces to

$$\bar{F}_{\mathfrak{e}}(x) = \frac{(\alpha + \beta + \alpha\beta x)}{\alpha + \beta} e^{-\alpha x}$$

which is the sf of two-parameter Lindley with parameters α and β. The sample path properties of the four AR(1) models developed in this section are displayed in Figure 3-6 and it shows how these measures vary with different values of parameters.

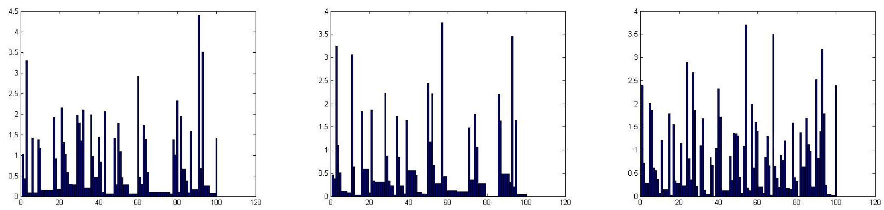


Figure 5: *Graphs of sample path of AR(1) Minification model II for different sets of (ϱ_1, ϱ_2)=(0.3,0.5), (0.5,0.3), (0.4,0.4), α=1.7 and β =1.4*

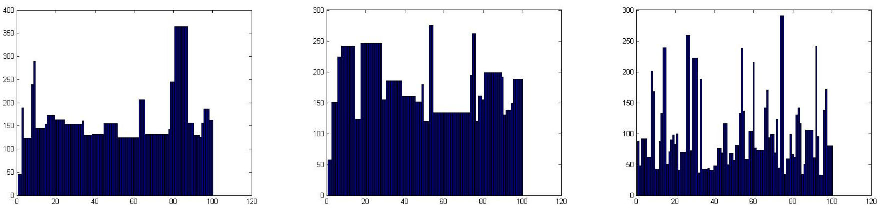


Figure 6: *Graphs of sample path of AR(1) Min-max model II for different sets of* $(\varrho_1, \varrho_2, \varrho_3)$*=(0.1,0.3,0.4), (0.2,0.2,0.4), (0.1,0.1,0.1),* α*=0.02 and* β *=0.02*

4. Estimation of Parameters

In this section, we consider maximum likelihood estimation (MLE) for a given sample of size $x_1, x_2, ..., x_n$ from MOTL(α, β, γ), then the log likelihood function is given by

$$\begin{aligned} l(\xi) &= n(\log\gamma + 2\log\alpha) + \sum_{i=1}^{n}\log(1+\beta x_i) - \alpha\sum_{i=1}^{n}x_i - n\log(\alpha+\beta) \\ &- 2\sum_{i=1}^{n}\log\left\{1-(1-\gamma)\frac{(\alpha+\beta+\alpha\beta x_i)e^{-\alpha x_i}}{\alpha+\beta}\right\} \end{aligned}$$

The partial derivative of the log likelihood functions with respect to the parameters are

$$\frac{\partial \text{ł}(\xi)}{\partial\alpha} = \frac{2n}{\alpha} - \sum_{i=1}^{n}x_i - \frac{n}{\alpha+\beta} - 2\sum_{i=1}^{n}(1-\gamma)\frac{\left((\alpha^2+(\alpha+\beta)\alpha\beta x_i)x_i e^{-\alpha x_i}\right)}{\left\{1-(1-\gamma)\frac{(\alpha+\beta+\alpha\beta x_i)e^{-\alpha x_i}}{\alpha+\beta}\right\}(\alpha+\beta)^2}$$

$$\frac{\partial l(\xi)}{\partial\beta} = \sum_{i=1}^{n}\frac{x_i}{1+\beta x_i} - \frac{n}{\alpha+\beta} + 2\sum_{i=1}^{n}\frac{(1-\gamma)\alpha^2 x_i e^{-\alpha x_i}}{\left\{1-(1-\gamma)\frac{(\alpha+\beta+\alpha\beta x_i)e^{-\alpha x_i}}{\alpha+\beta}\right\}(\alpha+\beta)^2}$$

$$\frac{\partial l(\xi)}{\partial\gamma} = \frac{n}{\gamma} - 2\sum_{i=1}^{n}\frac{\frac{(\alpha+\beta+\alpha\beta x_i)e^{-\alpha x_i}}{\alpha+\beta}}{\left\{1-(1-\gamma)\frac{(\alpha+\beta+\alpha\beta x_i)e^{-\alpha x_i}}{\alpha+\beta}\right\}}$$

The MLE $\hat{\xi}=(\hat{\alpha}, \hat{\beta}, \hat{\gamma})^T$ of $\xi=(\alpha, \beta, \gamma)^T$ can be numerically obtained by solving the equations $\frac{\partial l(\xi)}{\partial\alpha} = 0$, $\frac{\partial l(\xi)}{\partial\beta} = 0$, $\frac{\partial l(\xi)}{\partial\gamma} = 0$. For this purpose, we can use functions like *nlm*, *fitdist* or *optimize* from the programming language R.

5. Application

Now we use a real data set to show that the MOTL distribution can be a better model than the some other generalized exponential and Lindley distributions. The distributions are given below:

1. beta generalized exponential distribution (BGE) [9]
2. exponentiated exponential distribution (EE) [7]
3. gamma generalized exponential distribution (GGE) [39]
4. Kumaraswamy generalized exponential distribution (KGE) [6]

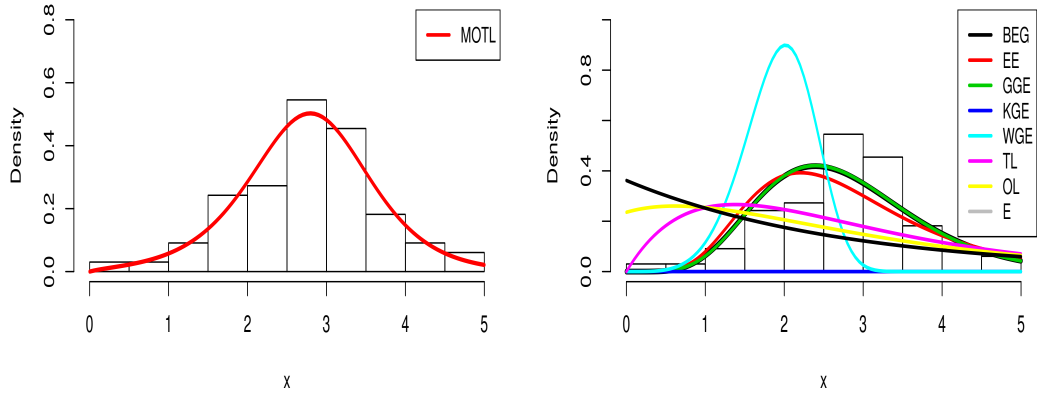


Figure 7: *pdf for fitted distributions of the breaking stress of carbon fibers data*

5. Weibull generalized exponential distribution (WGE) [1]
6. two-parameter Lindley distribution (TL) [31]
7. one-parameter Lindley distribution (OL) [16, 17]
8. classical exponential distribution (E)

The data set represents the breaking stress of carbon fibers of 50mm in length (n=66) and it has been given by [24]. The data set is given as:

3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11,3.56, 4.42, 2.41, 3.19, 3.22, 1.96, 3.28, 3.09, 1.87, 3.15, 4.90, 1.57, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20,3.33, 2.55, 3.31, 3.31, 2.85, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.89, 2.88,2.82, 2.05, 3.65, 3.75, 2.43, 2.95, 2.97, 3.39,2.96, 2.35, 2.55, 2.59, 2.03, 1.61, 2.12, 3.15, 1.08, 2.56, 1.80, 2.53

The distribution of this data set is unimodal and slightly left skewed (skewness = 0.131 and kurtosis = 3.223). For each distribution, we estimated the unknown parameters (by the maximum likelihood method), the values of the $-\log$-likelihood ($-\log L$), AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), the values of the Kolmogorov-Smirnov (K-S) statistic and the corresponding p-values.

All the computations were done through the use of the R programming language. The results for these data are listed in Table 3. From the table, we can observe that MOTL distribution provides smallest $-\log L$, AIC, BIC and K-S statistics values and highest p-value as compare to other distributions. This indicates that the MOTL distribution provides a better fit than other distributions. Plots of the histogram with fitted density functions and comparison of the cumulative distribution function for the each models with the empirical distribution function are displayed in Figure **??** and Figure **??**. From figures, one can easily identify the suitability behavior of the MOTL distribution. Therefore, the new model may be an interesting alternative to the other available generalized exponential models in the literature.

Table 3: *Estimated values, −logL, AIC, BIC, K-S statistics and p-value for data set.*

Distribution	Estimates	−logL	AIC	BIC	K-S	*p*-value
MOTL(α,β,γ)	$\hat{\alpha}$ = **2.2964** $\hat{\beta}$= **11012.6970** $\hat{\gamma}$= **79.8819**	**84.5821**	**175.1641**	**181.7331**	**0.0599**	**0.9717**
BEG(a,b,β)	$\hat{\beta}$= 4.323281e+06 $\hat{a}$=7.5387 $\hat{b}$=20.1066 $\hat{\beta}$=0.1176	90.9961	187.9922	194.5611	0.1309	0.2082
EE(a,β)	$\hat{a}$=9.2945 $\hat{\beta}$=1.0092	95.2187	194.4373	198.8166	0.1527	0.0921
GGE(a,β)	a=7.5710 $\hat{\beta}$=2.7395	90.9359	185.8719	190.2512	0.1303	0.2124
KGE(a,b,β)	$\hat{a}$ =4.3710 $\hat{b}$= 53.0724 $\hat{\beta}$= 0.1696	86.5579	179.1158	185.6848	0.0899	0.6603
WGE(a,b,β)	$\hat{a}$ = 3.4608 $\hat{b}$= 5.0392 $\hat{\beta}$= 1.6439 $\hat{\gamma}$=50.4204	85.7900	177.58	184.149	0.0799	0.7925
TL(α, β)	$\hat{\alpha}$= 2.2410 $\hat{\beta}$=174.011	112.0511	228.1022	232.4815	0.2510	-
OL(β)	$\hat{\beta}$=0.5895	181.7535	246.8953	249.085	0.3017	-
E(β)	$\hat{\beta}$= 0.3618	133.0921	268.1842	270.3739	0.35764	-

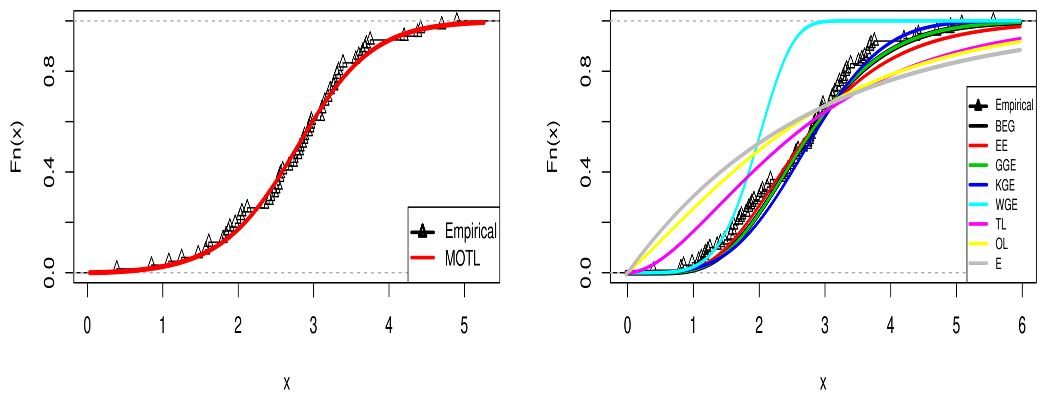


Figure 8: *Estimated cumulative distribution function for the data set*

6. Conclusions

The quality of the methods used in statistical analysis is eminently dependent on the underlying statistical distributions. In this article, we proffered a new customized Lindley distribution. The proposed distribution enfolds exponential and Lindley distributions as sub-models. Some properties of this distribution such as quantile function, moments, moment generating function, distributions of order statistics, limiting distributions of order statistics, entropy and autoregressive time series models are studied. This distribution is found to be the most appropriate model to fit the carbon fibers data compared to other models. Consequently, we propose the MOTL distribution for sketching inscrutable lifetime data sets.

References

[1] Alizadeh, M, Tahir, M.H., Cordeiro, G.M., Mansoor, M., Zubair, M , Hamedani, G.G. (2015). The Kumaraswamy Marshal-Olkin family of distributions. Journal of the Egyptian Mathematical Society, 23, 546–557.

[2] Al-Mutairi, D.K., Ghitany, M.E., Kundu, D., (2013). Inferences on stress-strength reliability from Lindley distributions, Communications in Statistics - Theory and Methods, 42, 1443–1463.

[3] Arnold B.C., Balakrishnan N., Nagaraja H.N., (1992), A First Course in Order Statistics, Wiley, New York.

[4] Azzalini, A. (1985). A class of distributions which includes the normal ones. Scand Journal of Statistics, 12, 171–178.

[5] Bourguignion, M., Silva, R.B., Cordeiro, G.M. (2014). The Weibull Generalized family of probability distributions. Journal of Data Science, 12: 53–68.

[6] Cordeiro, G.M, Castro, M. (2011). A New Family of Generalized Distributions. Journal of Statistical Computation and Simulation, 81, 883–898.

[7] Cordeiro, G.M, Ortega, E.M.M, Cunha, D.C.C. (2013). The Exponentiated Generalized Class of Distributions. Journal of Data Science, **11**, 1–27.

[8] Deniz, E. and Ojeda, E. (2011). The discrete lindley distribution: Properties and application. Journal of Statistical Computation and Simulation, 81, 11, 1405-1416.

[9] Eugene, N., Lee, C., Famoye, F. (2002). Beta-Normal Distribution and Its Applications. Communications in Statistics-Theory and Methods, 31, 497–512.

[10] Ghitany, M.E, Al-Mutairi, D.K, Balakrishnan, N., Al-Enezi, L.J. (2013). Power lindley distribution and associated inference. Computational Statistics and Data Analysis, 6, 20-33.

[11] Ghitany, M.E., Atieh, B. Nadarajah, S. (2008). Lindley distribution and its Applications, Mathematical Computation and Simulation, 78, 4, 493-506.

[12] Gómez, Y.M., Bolfarine, H., Gómez, H.W. (2014). A new extension of the exponential distribution. Revista Colombiana de Estadistica, 37, 25–34.

[13] Gupta, R.C., Gupta, P.L., Gupta, R.D., (1998). Modeling Failure Time Data by Lehman Alternatives. Communications in Statistics-Theory and Methods, 27, 887–904.

[14] Krishna, H., Kumar, K. (2011). Reliability estimation in Lindley distribution with Progressive type II right censored sample, Journal Mathematics and Computers in Simulation archive, 82, 2, 281-294.

[15] Laba, H., Subrata, C. (2017). The Marshall-Olkin-Kumaraswamy-G family of distributions. Journal of Statistical Theory and Applications, 16(4), 427–447.

[16] Lindley, D. V. (1958). Fiducial distributions and Bayes theorem, Journal of the Royal Statistical Society, A, 20, 102-107.

[17] Lindley, D. V., (1965). Introduction to Probability and Statistics from a Bayesian Viewpoint, Part II : inference, Combridge University Press, New Yourk.

[18] Mahmoudi, E., Zakerzadeh, H., (2010). Generalized Poisson-Lindley distribution, Communications in Statistics - Theory and Methods, 39, 1785–1798.

[19] Marshall, A.W., Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrica, 84, 641–652.

[20] Merovci, F. (2013). Transmuted Lindley distribution, International Journal of Open Problems in Computer Science and Mathematics, 6, 63–72.

[21] Merovci, F., Elbatal, I. (2014). Transmuted Lindley-geometric distribution and its applications, Journal of Statistics Applications & Probability, 3, 77–91.

[22] Merovci, F., Sharma, V.K. (2014). The beta Lindley distribution: Properties and applications, Journal of applied mathematics 2014, 1–10.

[23] Nadarajah, S., Bakouch, H., Tahmasbi, R., (2011). A generalized Lindley distribution. Sankhya B-Applied and Interdisciplinary Statistics, 73, 331–359.

[24] Nichols, M.D., Padgett, W.J. (2006). A bootstrap control chart for Weibull percentiles. Quality and Reliability Engineering, 22, 141–151.

[25] Rasekhi, M., Alizadeh, M., Altun, E., Hamedani, G.G., Afify, A.Z., Ahmad, M. (2017). The Modified Exponential Distribution with Applications. Pakistan Journal of Statistics, 33(5), 383–398.

[26] Rényi, A. (1960). On Measures of Entropy and Information. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1, 547–561.

[27] Ristić, M.M., Balakrishnan. (2012). The gamma-exponentiated exponential distribution. Journal of Statistical Computation and Simulation, 82, 1191–1206.

[28] Shanker, R., Hagos, F, Sujatha, S. (2015). On modeling of Lifetimes data using exponential and Lindley distributions. Biometrics & Biostatistics International Journal, 2, 5, 1-9.

[29] Shanker, R., Mishra, A.(2013a). A two-parameter Lindley distribution. Statistics in transition new series, 14, 1, 45-56.

[30] Shanker, R. Mishra A. (2013b). A quasi Lindley distribution. African Journal of Mathematics and Computer Science Research, 6, 4, 64-71.

[31] Shanker, R., Sharma, S., Shanker. R., (2013). A two-parameter Lindley distribution for modeling waiting and survival times data. Applied Mathematics, 4, 363-368.

[32] Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal, 27(3):379–423.

[33] Sharma, V,K, Singh, S., Singh, U. Agiwal, V. (2015). The inverse Lindley distribution: a stress-strength reliability model with applications to head and neck cancer data. Journal of Industrial and Production Engineering, 32, 3, 162-173.

[34] Thiago, A.N., Bourguignon, A.N.M., Cordeiro, G.M. (2016). The exponentiated generalized extended exponential distribution. Journal of Data Science, 14, 393–414.

[35] Thomas S.P., Jose K.K., Tomy L. (2019). Discrete Harris Extended Lindley Distribution and Applications, preprint.

[36] Tomy, L. (2018). A retrospective study on Lindley distribution, Biometrics and Biostatistics International Journal, 7, 3, 163-169.

[Torabi and Montazari(2014)] Torabi, H., Montazari, N.H. (2014). The logistic-uniform distribution and its application. Communications in Statistics-Theory and Methods, 43, 2551–2569.

[37] Zakerzadeh, H., Dolati, A. (2009). Generalized Lindley distribution, Journal of mathematical extension, 3, 2, 13-25.

[38] Zakerzadeh H, Mahmoudi E (2012). A New Two Parameter Lifetime Distribution: Model and Properties. arXiv:1204.4248 [stat.CO], URL http://arxiv.org/abs/1204.4248.

[39] Zografos, K., Balakrishnan. N (2009). On families of beta-and generalized gamma generated distributions and associated inference. Statistical Methodology, 6, 344–362.

Analysing Random Censored Data from Discrete Teissier Model

Abhishek Tyagi[1], Bhupendra Singh[1], Varun Agiwal[2], Amit Singh Nayal[1*]

•

[1] Department of Statistics, Chaudhary Charan Singh University, Meerut, India
[2]Indian Institute of Public Health, Hyderabad, Telangana, India.
abhishektyagi033@gmail.com
bhupendra.rana@gmail.com
varunagiwal.stats@gmail.com
*amitnayal009@gmail.com

Abstract

This paper deals with the classical and Bayesian estimation of the discrete Teissier distribution with randomly censored data. We have obtained the maximum likelihood point and interval estimator for the unknown parameter. Under the squared error loss function, a Bayes estimator is also computed utilising informative and non-informative priors. Furthermore, an algorithm to generate randomly right-censored data from the proposed model is presented. The performance of various estimation approaches is compared through comprehensive simulation studies. Finally, the applicability of the suggested discrete model has been demonstrated using two real datasets. The results show that the suggested discrete distribution fits censored data adequately and can be used to analyse randomly right-censored data generated from various domains.

Keywords: Bayesian estimation; Classical estimation; Discrete Teissier distribution; Random censoring.

1. Introduction

In many instances, the collection of data is constrained by time or budgetary limitations, making it difficult to obtain the whole data set. Such partial data is known as censored data. Various censoring schemes are available in the literature to examine this partial data. Conventional Type I and Type II censoring techniques are the most often used censoring schemes. In Type I censoring, the event is observed only if it occurs prior to some pre-specified time, whereas, in Type II censoring, the study continues until the predetermined number of individuals are observed to have failed. Random censoring is an another important censoring technique in the literature, this censoring scheme occurs when the subject under study is lost or removed from the experiment before its failure or event of intrest. This type of censoring commonly arises in medical time-to-event studies for example in clinical trials some patients do not complete the course of treatment and leave before the termination point. Therefore, the subject who leaves the study area before the event of interest occurs has a randomly censored value. The random censoring was introduced in literature by [1], he did so as part of his doctoral dissertation. For more details about the censoring schemes, their generalization, and analysis, one can refer to [2].

Randomly censored lifetime data frequently occur in many applications like medical science, biology, reliability studies, etc., which need to be analysed properly to make correct inferences and suitable research conclusions. These data are often right censored because it is not possible to observe the patients or the items under study until their death or patients may withdraw during the study period. In the existing literature, the random censoring scheme is widely studied under continuous models [see [3]].

In recent years, researchers in a variety of domains have acknowledged the distinctive significance played by discrete distributions. In certain circumstances, discrete distributions are more suitable than continuous distributions, even if the data is collected on a continuous scale. Also, discrete distributions are the only choice if the number of completed cycles of operation is used to measure the lifetime of different types of equipment [see [4],[5],[6] and references cited therein]. Most of the discrete models in the present literature were designed to fit the count data primarily, and in most cases, they fail to capture the diversity of the censored data. In literature, a few studies considered a random censoring scheme for discrete models, viz. [7] , [8], [9], and recently [10], discussed inferences of discrete inverted Nadarajah-Haghighi distribution with complete and random censored data. Due to the fact that most of the discrete distributions do not sufficiently portray the variety of real-world censored data, there is always a need for novel discrete distributions that can fit censored data adequately. One of such discrete distributions is the Discrete Teissier (DT) distribution proposed by [11], which provides the flexibility to fit the censored data with just a single parameter. Moreover, it can also model equi-, over-, and under-dispersed, positively skewed, negatively skewed, and increasing failure data. The probability mass function (PMF) of DT distribution is given by,

$$p_y = P[Y = y] = \exp(1)\exp(\alpha y)(\exp(-e^{\alpha y}) - \exp(\alpha - e^{\alpha(y+1)})); y = 0, 1, 2, ..., \alpha > 0. \quad (1)$$

Putting $\theta = \exp(\alpha)$, the PMF (1) can be written as

$$p_y = P[Y = y] = \exp(1)\theta^y(\exp(-\theta^y) - \theta\exp(-\theta^{(y+1)})); y = 0, 1, 2, ..., \theta > 1. \quad (2)$$

The cumulative distribution function (CDF) corresponding to PMF (2) is

$$F(y) = 1 - \theta^{y+1}\exp(1 - \theta^{(y+1)}); y = 0, 1, 2, ..., \theta > 1. \quad (3)$$

In this paper, we investigate the features of the DT distribution under randomly censored data. The article is organized as follows: The maximum likelihood estimator (MLE) for the model's parameter under randomly right-censored data is discussed in Section 2. Section 3 deals with the Bayesian estimation of the unknown parameter. The algorithm to generate censored data from the proposed model is given in Section 4. We use a Monte Carlo simulation analysis in Section 5 to investigate the characteristics of the different estimates established in the previous sections. Section 6 deals with the real data analysis to study the applications of random censoring in DT distribution. Finally, some concluding remarks are given in Section 7.

2. Method of maximum likelihood

2.1. Point Estimation

In this part, we compute the maximum likelihood point estimator for the DT distribution's parameter θ in the presence of random censored data. Let y_i be the i^{th} individual lifetime. In the presence of right-censored observations, the i^{th} individual contributes to the likelihood function (LF) based on a random sample (y_i, d_i) of size n as follows:

$$L_i = [p(y_i)]^{d_i}[S(y_i)]^{1-d_i},$$

where $S(y_i)$ is the survival function and d_i is a censoring indicator variable, that is, $d_i = 1$ for an observed lifetime and $d_i = 0$ for a censored lifetime $(i = 1, 2, ..., n)$. Then, for the DT distribution under random censoring, the LF of θ is given by

$$L(\underline{y}, \theta) = \exp(n)\theta^{\sum_{i=1}^{n}(y_i - d_i + 1)}\exp\left(-\sum_{i=1}^{n}\theta^{y_i+1}\right)\prod_{i=1}^{n}\left[\exp(\theta^{y_i+1} - \theta^{y_i}) - \theta\right]^{d_i}. \quad (4)$$

The log likelihood (LL) function corresponding to LF (4) is

$$LL = n + \left(\sum_{i=1}^{n} (y_i - d_i + 1)\right) \log\theta - \sum_{i=1}^{n} \theta^{y_i+1} + \sum_{i=1}^{n} d_i \log\left[\exp(\theta^{y_i+1} - \theta^{y_i}) - \theta\right]. \quad (5)$$

Taking the partial derivative of the LL function (5) with respect to the parameter θ, we get the following normal equation,

$$\frac{\partial LL}{\partial\theta} = \frac{1}{\theta}\left[\sum_{i=1}^{n} (y_i - d_i + 1) - \sum_{i=1}^{n} (E_2 + 1) - \sum_{i=1}^{n} \frac{d_i(\theta E_1 + E_3)}{1 - \theta E_1}\right] = 0. \quad (6)$$

where $E_1 = \exp(\theta^{y_i} - \theta^{y_i+1})$, $E_2 = (y_i + 1)\,\theta^{y_i+1} - 1$ and $E_3 = \theta^{y_i}\,(y_i - (y_i + 1)\,\theta)$.
The MLE of the parameter θ can be obtained by simplifying Equation (6); however, this equation does not provide an analytical solution. As a result, we employ an iterative method such as Newton-Raphson (NR) to compute the estimate computationally using built-in codes in R software.

2.2. Interval Estimation

The MLE of the unknown parameter θ is not found in closed form, hence exact distribution of MLE of θ cannot be derived. Therefore, it is infeasible to compute exact confidence interval for θ. Hence, we will construct the asymptotic confidence interval (ACI) for θ using the asymptotic distribution of MLE of θ. We know, the MLE $\hat{\theta}$ of θ, is consistent and asymptotic Gaussian distribution with $\sqrt{n}(\hat{\theta} - \theta)$ follows $N(0, I^{-1}(\theta))$, where $I(\theta) = E\left(-\frac{\partial^2 LL}{\partial\theta^2}\right)$. Therefore, the variance of the estimator $\hat{\theta}$ can be computed as $V(\hat{\theta}) \approx J^{-1}(\hat{\theta})$ where $J(\hat{\theta}) = -\left(\frac{\partial^2 \log L}{\partial\theta^2}\right)\Big|_{\theta=\hat{\theta}}$. The second-order partial derivative of LL function (5) is

$$\frac{\partial^2 LL}{\partial\theta^2} = -\frac{1}{\theta^2}\sum_{i=1}^{n} (y_i - d_i + 1) - \frac{1}{\theta^2}\sum_{i=1}^{n} y_i(1 + E_2) + \frac{1}{\theta^2}\sum_{i=1}^{n} \frac{d_i((1-\theta E_1)(E_3^2 - y_i(E_3 - \theta^{y_i})) - (\theta E_1 + E_3)^2)}{(1-\theta E_1)^2}.$$

Hence, the $100 \times (1 - \gamma)\%$ ACI for the parameter θ is

$$\hat{\theta} \mp Z_{\gamma/2}\sqrt{V(\hat{\theta})},$$

where $Z_{\gamma/2}$ is the upper $\gamma/2$ quantile of the standard Gaussian distribution.

3. Bayesian estimation

The Bayesian estimation blends prior and experimental information in terms of prior density and LF, respectively, to derive posterior inferences about the unknown quantities. The prior information is generally divided into two categories: informative priors and non-informative priors. Here, we will perform Bayesian estimation using both informative and non-informative priors to obtain Bayes estimators of the unknown parameter. Furthermore, the highest posterior density (HPD) interval for the parameter θ is also derived.

Case 1: When a probability distribution for the parameter θ provides adequate and full information, informative prior (IP) is used. In this scenario, we suppose θ has an exponential prior distribution with a density as

$$g(\theta) = \lambda e^{-\lambda(\theta-1)}; \theta > 1, \lambda > 0. \quad (7)$$

By combining prior distribution (7) with the LF (4) using the Bayes rule, the posterior distribution of θ given data is

$$P_1(\theta|\underline{y}) \propto \theta^{\sum_{i=1}^{n}(y_i - d_i + 1)} \exp\left(-\sum_{i=1}^{n} \theta^{y_i+1} - \lambda\theta\right) \prod_{i=1}^{n} \left(\exp(\theta^{(y_i+1)} - \theta^{y_i}) - \theta\right)^{d_i} \quad (8)$$

A loss function reflects the statistical risk (error) that arises while estimating parameters. It is a function of true and estimated parameters and is used to choose the best estimator with the lowest risk. The squared error loss function (SELF), which gives equal weight to overestimation and underestimation, is one of the most often used loss functions in literature. The Bayes estimator of a parameter under SELF is simply the expectation of that parameter with respect to its posterior distribution.

In the case of proposed distribution, the Bayes estimator of a function of the parameter θ under SELF, say $\psi(\theta)$ is

$$\hat{\psi}(\theta) = \int_1^\infty \psi(\theta) P_1(\theta|\underline{y}) d\theta. \tag{9}$$

The integral (9) cannot be given explicitly because the posterior distribution (8) is not in closed form. In this case, we may use a family of Markov Chain Monte Carlo (MCMC) algorithms to mimic draws from a posterior distribution. The Metropolis-Hastings (MH) algorithm [[12] and [13]] is a prominent approach in MCMC that generates a chain of random samples based on a given function, which may then be used to get Bayes estimates of interest.

To implement the MH algorithm for the proposed model, we go through the following steps:

Step 1. Set initial value of θ as $\theta^{(0)}$ and begin with $i = 1$.

Step 2. Propose a move $\theta^{*(i)}$, with candidate proposal density $g\left(\theta^{(i-1)}, \theta^{*(i)}\right)$.

Step 3. Calculate the Hastings ratio

$$\rho\left(\theta^{(i-1)}, \theta^{*(i)}\right) = \frac{P_1\left(\theta^{*(i)}|\underline{y}\right) g\left(\theta^{*(i)},\ \theta^{(i-1)}\right)}{P_1\left(\theta^{(i-1)}|\underline{y}\right) g\left(\ \theta^{(i-1)}, \theta^{*(i)}\right)}.$$

Step 4. Accept the proposed move $\theta^{*(i)}$ with probability $\tau = U(0,1) \leq min\left[1, \rho\left(\theta^{(i-1)}, \theta^{*(i)}\right)\right]$ and reject with probability $1 - \tau$.

Step 5. Set $i = i + 1$.

Step 6. Repeat steps 2-5 for all $i = 1, 2, 3, ..., M$ where M is large, and simulate the sequence of samples of $\theta^{(i)}$, $i = 1, 2, 3, ..., M$.

Step 7. The Bayes estimator of θ under SELF is calculated as

$$\hat{\theta} = \frac{1}{M - m} \sum_{i=(m+1)}^{M} \theta^{(i)},$$

where m is the burn-in iterations of the Markov Chain.

To compute the HPD interval for θ, let $\theta_{(m+1)} \leq \theta_{(m+2)} \leq ... \leq \theta_{(M)}$ denote the ordered values of θ_{m+1}, θ_{m+2}, ..., θ_M . Then, by [14] algorithm, the $100 \times (1-\gamma)\%$ HPD interval for θ is $(\theta_{(m+i^*)}, \theta_{(m.\,i^*+[(1-\gamma)(M-m)])})$, where i^* is chosen so that,

$$\theta_{(m+i^*+[(1-\gamma)(M-m)])} - \theta_{(m+i^*)} = \min_{m \leq i \leq (M-m)-[(1-\gamma)(M-m)]} (\theta_{(m+i+[(1-\gamma)(M-m)])} - \theta_{(m+i)}).$$

Case 2: In non-informative prior (NIP), least or no information is available about the unknown parameter. For the proposed model, we perform the Bayesian analysis, when θ has NIP of the following form,

$$g(\theta) \propto \frac{1}{\theta}; \theta > 1. \tag{10}$$

The un-normalized posterior distribution of θ given data is computed by combining prior distribution (10) with LF (4).

$$P_2(\theta|\underline{y}) \propto \theta^{\sum_{i=1}^n (y_i - d_i) + (n-1)} \exp\left(-\sum_{i=1}^n \theta^{y_i+1}\right) . \prod_{i=1}^n \left(\exp(\theta^{(y_i+1)} - \theta^{y_i}) - \theta\right)^{d_i}. \tag{11}$$

Since the posterior distribution $P_2(\theta|\underline{y})$ is again in non-closure form, so the Bayes estimator of θ is cannot be solved analytically. Therefore, using a similar algorithm as we have done in case 1, we can obtain the required point and interval estimates.

4. Algorithm to simulate random right censored data

In this section, we present a simple algorithm to generate the randomly right-censored data from the proposed model [15]. The algorithm consists of the following steps:

Step 1. Fix the value of the parameter θ.

Step 2. Draw n random pseudo from Uniform(0,1) i.e. $u_i \sim U(0,1); i = 1,2,...,n$.

Step 3. Obtain $y_i' = F^{-1}(u_i); i = 1,2,...,n$, where $F^{-1}(\bullet)$ is defined in Equation (6).

Step 4. Draw n random pseudo from $c_i \sim U(0, \max(y_i')); i = 1,2,...,n$. This is the distribution that controls the censorship mechanism.

Step 5. If $y_i' \leq c_i$, then $y_i = [y_i']$ and $d_i = 1, i = 1,2,...,n$, else, $y_i = [c_i]$ and $d_i = 0, i = 1,2,...,n$. Hence, pairs of values $(y_1, d_1), (y_2, d_2), ..., (y_n, d_n)$ are obtained as the random right-censored data.

5. Simulation study

The performance of the MLE and Bayes (IP and NIP under SELF) estimator under randomly right-censored data is investigated in this section via a simulation study. The whole study is based on random samples drawn from the DT distribution of sizes 20, 25,...,150. The parametric values of the parameter are taken as 1.05, 1.50, 2.0, and 3.0. To produce the needed random variable Y from the DT distribution, we employed the conventional strategy of first drawing the pseudo-random value X from the continuous Teissier distribution and then discretizing this value to store Y. A random variable X may be generated by using the following formula:

$$Q(u) = \frac{1}{\alpha} \log \left[-W_{-1} \left(\frac{u-1}{\exp(1)} \right) \right];\ 0 < u < 1,$$

where $\theta = \exp(\alpha)$ and W_{-1} denotes the Lambert function and its value can be easily obtained by the inbuilt R-function *lambertWm*1 available in the package *lamW*. The method described in Section 4 is utilized to produce the required random right-censored data. All simulation results are based on 2000 replicates for the different sample sizes considered for each parameter setting. We have calculated the mean-squared error (MSE) and the average absolute bias (AB) for MLE and Bayes point estimates and average width (AW) of 95% ACI and HPD intervals with their respective coverage probability (CP) based on these 2000 values, and the resulting findings are shown in Figures 1-3. Notably, when Bayesian estimation is used, an estimate for the parameter of the DT distribution is made using an exponential prior as an IP and a uniform prior as a NIP. Under exponential prior, the value of the hyper-parameter is calculated so that the expectation of the related prior density of the unknown parameter is equal to its actual parametric value. In this estimation scenario, we drew 51,000 MCMC samples for the parameter of the proposed distribution using the MH algorithm, excluding the first 11,000 samples as a burn-in phase to eliminate the effect of initial values. Additionally, to nullify the autocorrelation between successive draws, every tenth observation has been preserved. We have finally calculated the posterior quantities of interest by using generated posterior samples.

The following are some important inferences that are drawn from Figures 1-3:

- The MLE and Bayes estimator of the unknown parameters show the consistency property, i.e., the MSE reduces as the sample size rises.
- As n becomes larger, the average AB approaches zero.
- The Bayes estimator with IP performs better as compared to the MLE and Bayes estimator with NIP.
- The AW of HPD intervals under IP is lesser than those obtain under ACI and HPD with NIP.
- For large values of the parameter θ, all estimation methods produce nearly similar results. A similar trend is observed when sample size n becomes large.

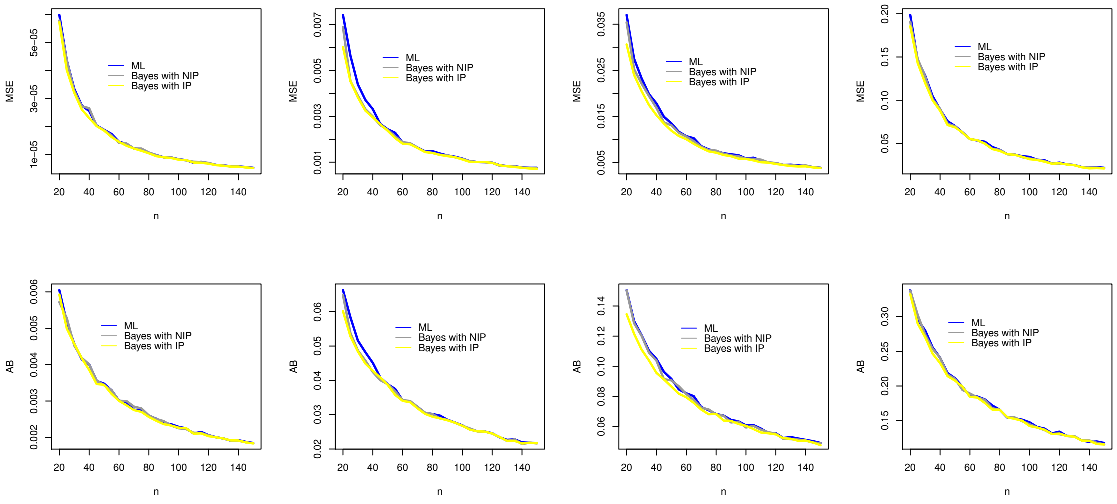


Figure 1: *The MLE and Bayes estimate for (i) θ =1.05 (ii) θ =1.50 (iii) θ =2.0 (iv) θ =3.0.*

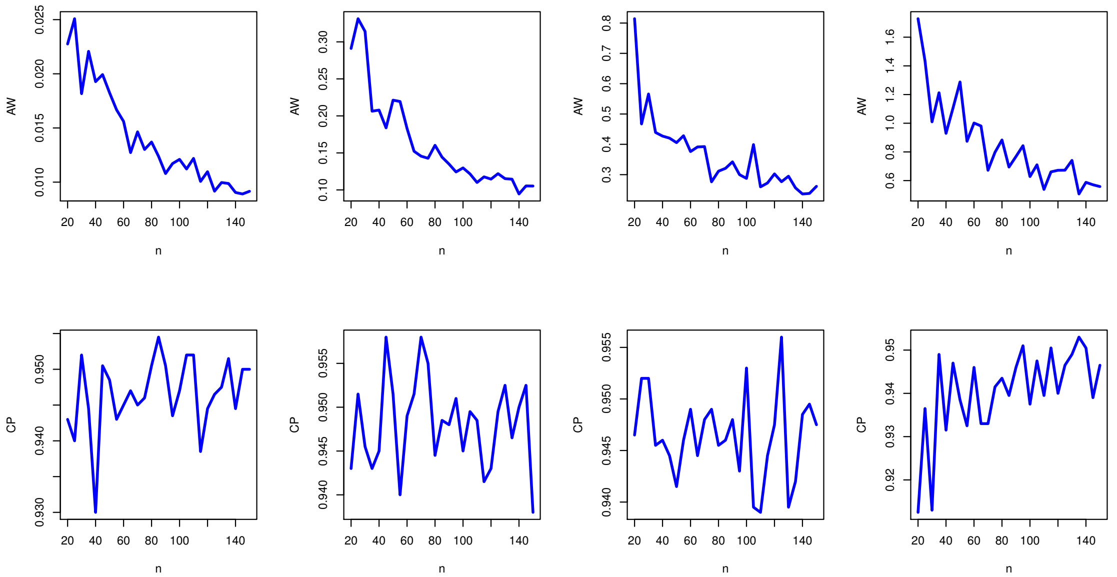


Figure 2: *The classical AW and CP for (i) θ =1.05 (ii) θ =1.50 (iii) θ =2.0 (iv) θ =3.0.*

6. Application to random censored data

Here, we examine two real datasets to demonstrate the applicability of the DT model to censored data. These data sets along with their fitting are described as follows:

The first data set (I): This data set consists of failure times for Epoxy Insulation Specimens at the voltage level 57.5 Kv [see [16], pp. 335]. The failure times, in minutes, for the insulation specimens are given below (censoring times are indicated with asterisks)
510, 1000*, 252, 408, 528, 690, 900*, 714, 348, 546, 174, 696, 294, 234, 288, 444, 390, 168, 558, 288.
Using Kolmogorov-Smirnov (K-S) statistics, we now evaluate the suitability of the DT distribution for modelling the above data. The K-S statistic and associated p-values of 0.17219 and 0.5936 indicate that the proposed model with MLE and associated standard error (SE) in parenthesis is 1.00195 (0.00018) sufficiently reflects the diversity of the data. Figure 4 (upper left panel)

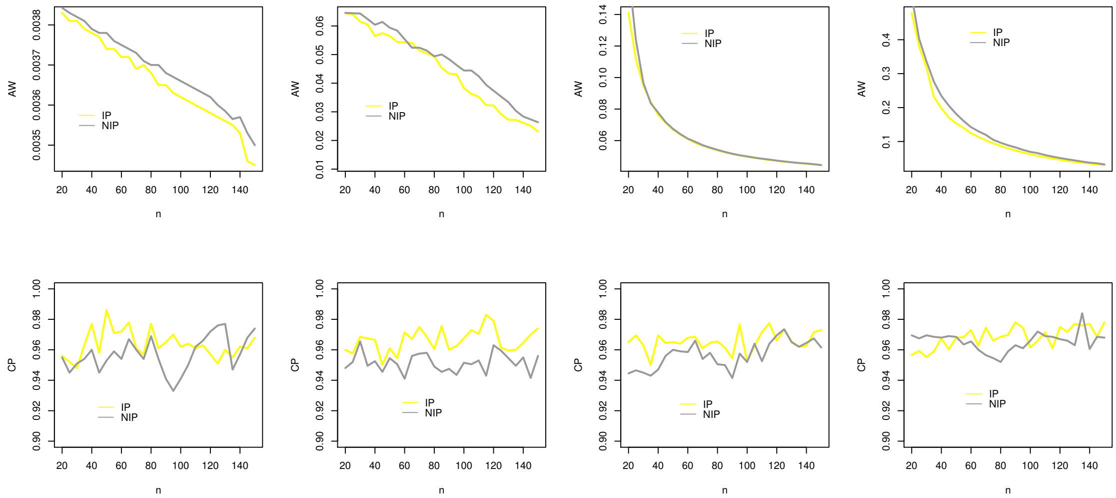


Figure 3: *The HPD AW and CP for (i) θ =1.05 (ii) θ =1.50 (iii) θ =2.0 (iv) θ =3.0.*

depicts the unique existence of MLE, whereas Figure 4 (upper right panel) demonstrates that the suggested model captures the data accurately. In addition, Table 1 displays the ACI, Bayes estimates, HPD interval with NIPs, and K-S statistics with its p-value.

The second data set (II): This data come from a nine-month study on the effect of known carcinogens DES and DMBA in the induction of mammary tumors in female rats [see [16], pp. 339]. After treatment, the times to tumor appearance for the animals were noted. The censored observations are indicated by asterisks. The data values are
57*, 67*, 88, 94, 100, 107, 113, 123, 123, 125, 129, 129, 129, 136, 136, 143, 144, 191, 191, 192, 211, 218, 266*, 266*.
For this data, the MLE (SE) of the parameter θ is 1.00658(0.00058). Now, using this estimate for the considered data, the K-S statistics and associated p-value are 0.2354 and 0.1396, respectively. This well-known goodness-of-fit measure indicates that the suggested discrete model is adequate for modelling the given censored data. The unique existence of the MLE can be verified by Figure 4 (lower left panel). Graphically, from Figure 4 (lower right panel), we can conclude that the DT model closely follows the pattern of this censored data. Also, we have obtained the ACI, Bayes estimates and HPD interval with NIPs, and the K-S statistics with its p-value, and they can be viewed in Table 1.

Table 1: *Classical and Bayesian estimates of censored data set I and II .*

Data set	Estimates		K-S	P-value
I	MLE (SE)	1.00193 (0.00018)	0.17219	0.5936
	ACI	[1.00158, 1.00228]		
	Bayes (SE)	1.00199 (0.00028)	0.1596	0.68790
	HPD	[1.00148, 1.00246]		
II	MLE (SE)	1.00658(0.00058)	0.23549	0.13960
	ACI	[1.00546, 1.00773]		
	Bayes (SE)	1.00660 (0.00061)	0.23465	0.14220
	HPD	[1.00549, 1.00766]		

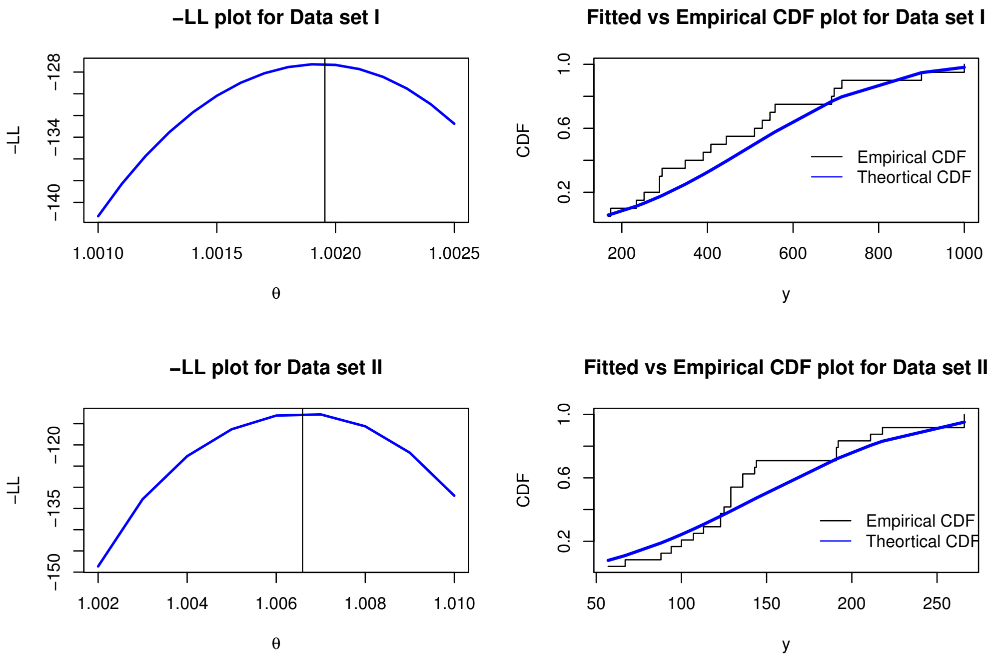


Figure 4: *The –LL and CDFs plots for data set I and II.*

7. Conclusion

In this article, the one-parameter DT distribution introduced by [11] was studied, taking into account the use of right-censored data. We use both classical and Bayesian methods to estimate the unknown parameter of the DT distribution. Furthermore, an algorithm to produce randomly right-censored data is also provided. An extensive simulation study is presented for the assessment of the various estimation procedures under censored data. Finally, the uselfuness of the proposed model is illustrated with two examples considering right censored real data sets. The study suggested that the proposed model can be used to analyse randomly right-censored data generated from various domains. Moreover, the DT distribution has the potential to attract more comprehensive applications in a variety of fields. A future plan of action regarding the current study might be an examination of the other types of censored data using the proposed model.

Conflict of interest

The authors declare no conflicts of interest.

References

[1] Gilbert, J. P. (1962). Random censorship. The University of Chicago.

[2] Klein, J. P., & Moeschberger, M. L. (2003). Survival analysis: techniques for censored and truncated data (Vol. 1230). New York: Springer.

[3] Garg, R., Dube, M.,& Krishna, H. (2020). Estimation of parameters and reliability characteristics in Lindley distribution using randomly censored data. Statistics, Optimization & Information Computing, 8(1), 80-97.

[4] El-Morshedy, M., Eliwa, M. S., & Tyagi, A. (2022). A discrete analogue of odd Weibull-G family of distributions: properties, classical and Bayesian estimation with applications to count data. Journal of Applied Statistics, 49(11), 2928-2952.

[5] Pandey, A., Singh, R. P., & Tyagi, A. (2022). AN INFERENTIAL STUDY OF DISCRETE BURR-HATKE EXPONENTIAL DISTRIBUTION UNDER COMPLETE AND CENSORED DATA. Reliability: Theory & Applications, 17(4 (71)), 109-122.

[6] Eliwa, M. S., Tyagi, A., Almohaimeed, B., & El-Morshedy, M. (2022). Modelling coronavirus and larvae Pyrausta data: A discrete binomial exponential II distribution with properties, classical and Bayesian estimation. Axioms, 11(11), 646.

[7] Krishna, H., & Goel, N. (2017). Maximum likelihood and Bayes estimation in randomly censored geometric distribution. Journal of Probability and Statistics, 2017.

[8] de Oliveira, R. P., de Oliveira Peres, M. V., Martinez, E. Z., & Achcar, J. A. (2019). Use of a discrete Sushila distribution in the analysis of right-censored lifetime data. Model Assisted Statistics and Applications, 14(3), 255-268.

[9] Achcar, J. A., Martinez, E. Z., de Freitas, B. C. L., & de Oliveira Peres, M. V. (2021). Classical and Bayesian inference approaches for the exponentiated discrete Weibull model with censored data and a cure fraction. Pakistan Journal of Statistics and Operation Research, 467-481.

[10] Singh, B., Singh, R. P., Nayal, A. S., & Tyagi, A. (2022). Discrete Inverted Nadarajah-Haghighi Distribution: Properties and Classical Estimation with Application to Complete and Censored data. Statistics, Optimization & Information Computing, 10(4), 1293-1313.

[11] Singh, B., Agiwal, V., Nayal, A. S., & Tyagi, A. (2022). A Discrete Analogue of Teissier Distribution: Properties and Classical Estimation with Application to Count Data. Reliability: Theory & Applications, 17(1 (67)), 340-355.

[12] Metropolis N, Ulam S (1949). The Monte Carlo method. J Am Stat Assoc 44:335–341.

[13] Hastings, WK. (1970). Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109.

[14] Chen M. H & Shao, Q., M. (1999). Monte Carlo estimation of Bayesian credible intervals and HPD intervals. J Comput Graph Stat., 8(1):69–92.

[15] Ramos, P. L., Guzman, D. C., Mota, A. L., Rodrigues, F. A., & Louzada, F. (2020). Sampling with censored data: a practical guide. arXiv preprint arXiv:2011.08417.

[16] Lawless, J. F. (2003). Statistical models and methods for lifetime data (Vol. 362). John Wiley & Sons.

ON GOMPERTZ EXPONENTIATED INVERSE RAYLEIGH DISRIBUTION

Sule Omeiza Bashiru[1] and Halid Omobolaji Yusuf[2]

•

[1]Prince Abubakar Audu University, Anyigba, Kogi State, Nigeria
[2]Ekiti State University, Ado-Ekiti, Ekiti State, Nigeria
Email: [1]bash0140@gmail.com, Email: [2]omobolajihalid01@gmail.com

Abstract

In this paper, we proposed a four parameter Gompertz Exponentiated Inverse Rayleigh Distribution. The proposed distribution is an extention of the Exponentiated Inverse Rayleigh Distribution which was compounded with the Gompertz generated family of distribution. Several of its statistical and mathematical properties including quantiles, median, moments, skewness and kurtosis are derived. Also, the reliability and hazard rate functions are derived. To estimate the new model parameters, the maximum likelihood technique is used. To evaluate the effectiveness of the estimators in this model, a simulation study was carried out and the result of the simulation study indicated that the model is consistent since the value of the mean square error decrease as sample size increases. Finally, the usefulness of the proposed distribution is illustrated with two datasets and it is discovered that this model is more adaptable when compared to well-known models..

Keywords: Maximum likelihood estimation; Skewness; Kurtosis; Probability density function; Cummulative probability distribution.

1. Introduction

The classical distributions frequently do not offer an appropriate match to some real data sets in real-world circumstances. In order to create novel distributions, researchers devised numerous generators by inserting one or more parameters. The newly generated distributions are more adaptable than the classical distributions.

Gompertz[10] introduced a continuous probability distribution known as the Gompertz probability distribution (GD). The GD is employed to explore nature of human mortality by determining the value of life's unexpected events. Several branches of statistics have used the Gompertz distribution where survival time is necessary such as in demography Vaupel [18], Preston et al [16] and in actuary Willemse and Koppelaar[19); in gerontology, medicine, biology, and related sciences Economos [7], Brown and Forbes [6]. In this article, the Gompertz family of distributions is used to create a novel model. Some authors that have employed the Gompertz Family of distributions include : Halid and Sule [12], Alizadeh et al. [4], and Abdal-Hameed et al. [1] . Halid and Sule [12] defined the cummulative density function (CDF) of the Gompertz family of distribution as:

$$F_X(x) = 1 - e^{\left(\frac{\varphi}{\eta}\right)\left[1-(1-G(x))^{-\eta}\right]} \tag{1}$$

and the corresponding PDF to (1) is given by

$$f_X(x) = \varphi g(x)\left[1 - G(x)\right]^{-\eta-1} e^{\left(\frac{\varphi}{\eta}\left\{1-\left[1-G(x)\right]^{-\eta}\right\}\right)} \tag{2}$$

where φ and η are the extra shape parameters

The article is broken down into the following sections: In Section 2, the new distribution GEIR's derivation is described. In Section 3, the mathematical characteristics of the new distribution are explained and the Maximum likelihood estimation of the distribution is used to estimate the parameters. In Section 4, we presented and explored the new distribution's practical applicability. Finally, Section 5 displays the concluding remarks.

2. Derivation of Gompertz Exponentiated Inverse Rayleigh Distribution

In this section, we derived the Gompertz Exponentiated Inverse Rayleigh (GEIR) distribution. Rao and Mbwambo [17], introduced the CDF and PDF of Exponentiated Inverse Rayleigh (EIR) Distribution as

$$G_X(x) = 1 - \left(1 - e^{\frac{\xi}{x}}\right)^{\alpha}; \qquad x \geq 0, \xi > 0, \alpha > 0 \tag{3}$$

The corresponding pdf is given as:

$$g_X(x) = \frac{2\alpha\xi^2}{x^3} e^{-\left(\frac{\xi}{x}\right)^2} \left(e^{-\left(\frac{\xi}{x}\right)^2}\right)^{\alpha-1}; \qquad x \geq 0, \xi > 0, \alpha > 0 \tag{4}$$

putting equation (3) into (1) we have the CDF of Gompertz Exponentiated Inverse Rayleigh (GEIR)

$$F_X(x) = 1 - e^{\left(\frac{\varphi}{\eta}\right)\left[1-\left(1-exp(-(\xi/x)^2)\right)^{-\eta\alpha}\right]} \qquad x \geq 0, \xi > 0, \alpha > 0, \eta > 0, \varphi > 0 \tag{5}$$

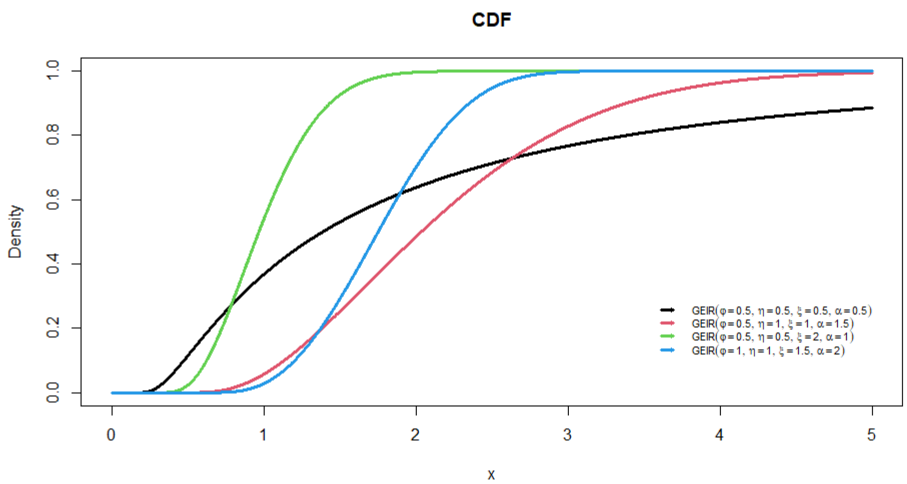


Figure 1: *CDF plot of GEIR distribution for different parameter values*

Now, putting (3)and (4) into (2), we now obtained the PDF of the proposed GEIR distribution given by

$$f_X(x) = 2\varphi\xi^2 x^{-3} e^{-\left(\frac{\xi}{x}\right)^2} \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta-1} e^{\frac{\varphi}{\eta}\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta}\right\}} \tag{6}$$

where η and α are shape parameters, φ and ξ are scale parameters.

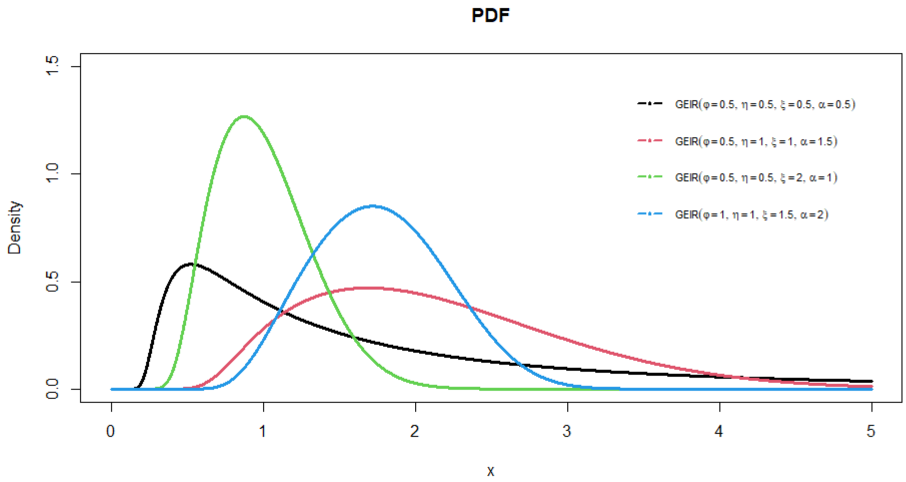


Figure 2: *PDF plot of GEIR distribution for different parameter values*

3. Properties of GEIR Distribution

3.1. Linear Mixture of GEIR

Given the CDF and PDF of GEIR distribution (5) and (6), the expressions

$$e^{\frac{\varphi}{\eta}\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}} = \sum_{k}^{\infty} \frac{(-1)^k}{m!}\left(\frac{\varphi}{\eta}\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}\right)^k$$

$$= \sum_{k}^{\infty} \frac{(-1)^k}{m!}\left(\frac{\varphi}{\eta}\right)^k\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}^k$$

$$\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}^k = 2\varphi\alpha\xi^2x^{-3}e^{-\left(\frac{\xi}{x}\right)^2}\sum_{k}^{\infty}\sum_{m}^{\infty}\frac{(-1)^{k+m}}{m!}\binom{k}{m}\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha(m+1)-1}$$

$$f(x) = \sum_{k}^{\infty}\sum_{m}^{\infty}\frac{(-1)^{k+m}}{m!}\binom{k}{m}2\varphi\alpha\xi^2x^{-3}e^{-\left(\frac{\xi}{x}\right)^2}\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha(m+1)-1}$$

$$\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha(m+1)-1} = \sum_{k}^{\infty}\sum_{m}^{k}\sum_{n}^{\infty}\frac{(-1)^{k+m+n}}{m!}\binom{k}{m}\binom{-\eta\alpha(m+1)-1}{n}e^{-n\left(\frac{\xi}{x}\right)^2}$$

So therefore, the PDF of GEIR distribution can be expressed as

$$f(x) = \omega_{k,m,n}2\,(n+2)\,\varphi\alpha\xi^2x^{-3}\left(e^{-\left(\frac{\xi}{x}\right)^2}\right)^{n+1}$$

where

$$\omega_{k,m} = \sum_{k}^{\infty}\sum_{m}^{k}\sum_{n}^{\infty}\frac{(-1)}{m!\,(n+2)}\binom{k}{m}\binom{-\eta\alpha(m+1)-1}{n}$$

and the CDF of GEIR distribution can be expressed as

$$F(x) = \omega_{k,m,n}\varphi\alpha\left(e^{-\left(\frac{\xi}{x}\right)^2}\right)^{n+2}$$

where $g_{n+1}(x)$ is the PDF of Gompertz Exponentiated Inverse Rayleigh distribution with shape parameter n+2

3.2. Survival Function

According to Ieren and Balogun [13], the survival function discribes the probability that a unit, or component, or individual will not fail at a given time . A survival function is generally expressed as

$$S(x) = 1 - F(x;\varphi,\alpha,\eta,\xi) \tag{7}$$

Therefore the survival function of GEIR distribution is derived by substituting (5) into (7) which resulted to

$$S_X(x) = e^{\frac{\varphi}{\eta}\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta}\right\}} \tag{8}$$

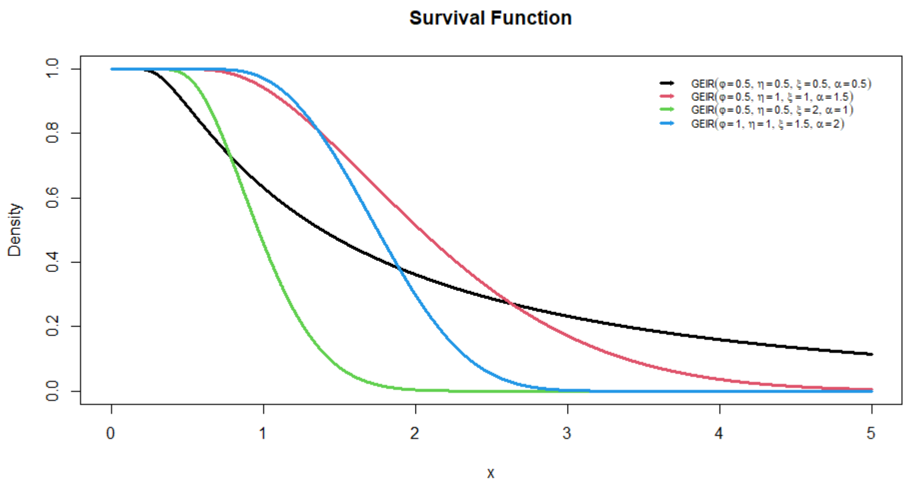


Figure 3: *GEIR distribution survival plot for various parameter values*

3.3. Hazard Function

The hazard function is given as

$$h(x) = \frac{f(x)}{1-F(x)} \tag{9}$$

The hazard function for GEIR distribution is derived by substituting (5) and (6) into (9) and it resulted into

$$h_X(x) = \frac{2\varphi\xi^2 x^{-3} e^{-\left(\frac{\xi}{x}\right)^2} \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta - 1} e^{\frac{\varphi}{\eta}\left\{1 - \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}}}{e^{\frac{\varphi}{\eta}\left\{1 - \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}}} \tag{10}$$

$$h_X(x) = 2\varphi\xi^2 x^{-3} e^{-\left(\frac{\xi}{x}\right)^2} \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta - 1} \tag{11}$$

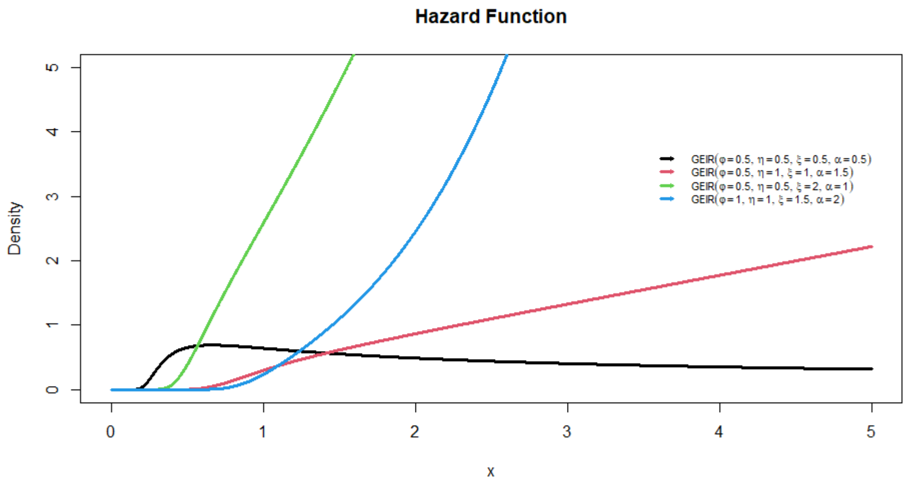


Figure 4: *Hazard plot of GEIR distribution for different parameter values*

3.4. Cumulative Hazard Function

From this definition, the cumulative hazard function, $H_X(x)$, of a continuous random variable, X, which follows the GEIR distribution is obtained.

$$H_X(x) = -log[S_X(x)] \tag{12}$$

substituting equation (8) into (12), we obtain

$$H_X(x) = -log\left[e^{\frac{\varphi}{\eta}\left\{1 - \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\}}\right] \tag{13}$$

$$H_X(x) = -\frac{\varphi}{\eta}\left\{1 - \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right\} \tag{14}$$

3.4.1 Reversed Hazard Function

The reversed hazard function can be obtained by applying the formula below:

$$\tau(x) = \frac{f(x)}{F(x)} \tag{15}$$

Hence, we obtain the reversed hazard function by substituting (5) and (6) in (15)

$$\tau(x) = \frac{2\varphi\xi^2 x^{-3} e^{-\left(\frac{\xi}{x}\right)^2} \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta-1} e^{\frac{\varphi}{\eta}\left\{1-\left[1-e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\alpha\eta}\right\}}}{1 - e^{\left(\frac{\varphi}{\eta}\right)\left[1-(1-exp(-(\xi/x)^2))^{-\eta\alpha}\right]}} \tag{16}$$

3.5. Quantile Function, Median, Skewness and Kurtosis

The p^{th} quantile of the GEIR distribution is derived as

$$Q_X(p) = \frac{\xi}{\sqrt{-\log\left(1 - \left[1 - \frac{\eta}{\varphi}\log(1-p)\right]^{-\frac{1}{\eta\alpha}}\right)}} \tag{17}$$

we have the first three , $Q_1 1 = Q(1/4)$ and $Q_3 = Q(3/4)$, that is by substituting value of p=0.25 and p=0.75 in X_p, respectively. Also Quantile is also used in finding the skewness and kurtosis of the distribution.

3.5.1 Median

Substitute p=0.5 in (17), we have

$$M_e = Q_X(0.5) = \frac{\xi}{\sqrt{-\log\left(1 - \left[1 - \frac{\eta}{\varphi}\log(0.5)\right]^{-\frac{1}{\eta\alpha}}\right)}} \tag{18}$$

3.5.2 Skewness and Kurtosis

According to Galton [9] and Moors[15] we can obtain the skewness (Sk) and kurtosis (Ku) measures, respectively for GEIR distribution using the following expression

$$Sk = \frac{Q(\frac{3}{4}) + Q(\frac{1}{4}) - 2Q(\frac{1}{2})}{Q(\frac{3}{4}) - Q(\frac{1}{4})} \tag{19}$$

and

$$Ku = \frac{Q(\frac{7}{8}) - Q(\frac{5}{8}) + Q(\frac{3}{8}) - Q(\frac{1}{8})}{Q(\frac{6}{8}) - Q(\frac{2}{8})} \tag{20}$$

3.5.3 Moment

In this section, we consider the moment of the GEIR distribution. Let X = $(x_1, x_2, \ldots, x_n)$ be a sample drawn from GEIR distribution with pdf, then the r^{th} moment μ_r' can be written as

$$\mu_r' = \int_0^\infty x^r f(x)dx$$

$$\mu_r' = \int_0^\infty x^r \omega_{k,m,n} 2\,(n+2)\,\varphi\alpha\xi^2 x^{-3}\left(e^{-\left(\frac{\xi}{x}\right)^2}\right)^{n+1} dx$$

after some mathematical derivations, we obtained the r^{th} moment as

$$= \omega_{k,m,n} 2\,(n+2)\,(n+1)^{(r/2-1)}\varphi\alpha\xi^2\Gamma(1-\frac{r}{2}) \qquad r < 2$$

3.6. Maximum Likelihood Estimation

Due to its consistency, asymptotic efficiency, and invariance property, the Maximum Likelihood Estimation (MLE) method is frequently used to estimate unknown parameter(s). Let $x_1, x_2, \ldots, x_n$ be random sample of size n drawn from GEIR distribution, then the likelihood can be expressed as :

$$L(\varphi, \xi, \eta, \alpha) = 2^n \varphi^2 \xi^{2n} \sum_{i=1}^{n} x^{-3} e^{-\sum_{i=1}^{n}\left(\frac{\xi^2}{x^2}\right)} \prod_{i=1}^{n} \left[1 - e^{-\frac{\xi^2}{x^2}}\right]^{-1-\eta\alpha} e^{-\sum_{i=1}^{n}\left(\frac{\varphi}{\eta}\left[1-\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\alpha\eta}\right]\right)} \tag{21}$$

and the log-likelihood of expression (21) can be expressed as

$$\begin{aligned} l = \log L \quad = \quad & n \ln\left(\varphi\alpha\xi^2\right) - \eta\alpha \sum_{i=1}^{n} \ln\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right) - \xi^2 \sum_{i=1}^{n} \left(\frac{1}{x^2}\right) + \frac{\varphi}{\eta} \sum_{i=1}^{n} \left(1 - \left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\right) \\ & -3 \sum_{i=1}^{n} \ln(x) - \sum_{i=1}^{n} \ln\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right) \end{aligned} \tag{22}$$

Differentiating (22) with respect to φ, ξ, α and η, if equated to zero, we obtain the following estimating equations

$$\frac{\partial l}{\partial \varphi} = \frac{n}{\varphi} + \frac{1}{\eta} \sum_{i=1}^{n} \left(1 - \left[1 - e^{-\left(\frac{\xi}{x}\right)^2}\right]^{-\eta\alpha}\right) \tag{23}$$

$$\begin{aligned} \frac{\partial l}{\partial \xi} \quad = \quad & \frac{2n}{\xi} - 2\eta\alpha\xi \sum_{i=1}^{n} \left(\frac{\mathrm{e}^{-\frac{\xi^2}{x^2}}}{x^2 \left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)} \right) - 2\xi \sum_{i=1}^{n} \left(\frac{2}{x^2}\right) + 2\varphi\alpha\xi \sum_{i=1}^{n} \left(\frac{\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha} \mathrm{e}^{-\frac{\xi^2}{x^2}}}{x^2 \left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)} \right) \\ & -2\xi \sum_{i=1}^{n} \left(-\frac{\mathrm{e}^{-\frac{\xi^2}{x^2}}}{x^2 \left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)} \right) \end{aligned} \tag{24}$$

$$\frac{\partial l}{\partial \eta} = -\alpha \sum_{i=1}^{n} \ln\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right) + \frac{\alpha\varphi}{\eta} \sum_{i=1}^{n} \left[\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha} \ln\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right) \right] - \frac{\varphi}{\eta^2} \sum_{i=1}^{n} \left(1 - \left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\right) \tag{25}$$

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} - \eta \sum_{i=1}^{n} \ln\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right) + \varphi \sum_{i=1}^{n} \left[\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha} \ln\left(1 - \mathrm{e}^{-\frac{\xi^2}{x^2}}\right) \right] \tag{26}$$

The maximum likelihood estimator $\hat{\theta} = (\hat{\varphi}, \hat{\xi}, \hat{\eta}, \hat{\alpha})$ of $\theta = (\varphi, \xi, \eta, \alpha)$ is obtained by solving the nonlinear system of equations (23) - (26). In this study, we used the Newton Raphson technique, a nonlinear optimization procedure, to numerically optimize the log-likelihood function shown in (22). The asymptotic distribution of the element of the 4×4 observed information matrix of GEIR distribution can be expressed as

$$\sqrt{n}(\hat{\theta} - \theta) \sim N_4(0, \Sigma^{-1}) \tag{27}$$

where Σ is the expected information matrix. Thus, the expected information matrix is expressed as

$$\Sigma^{-1} = -E \begin{bmatrix} \frac{\partial^2 l}{\partial \varphi^2} & \frac{\partial^2 l}{\partial \varphi \partial \xi} & \frac{\partial^2 l}{\partial \varphi \partial \eta} & \frac{\partial^2 l}{\partial \varphi \partial \alpha} \\ \frac{\partial^2 l}{\partial \varphi \partial \xi} & \frac{\partial^2 l}{\partial \xi^2} & \frac{\partial^2 l}{\partial \eta \partial \xi} & \frac{\partial^2 l}{\partial \alpha \partial \xi} \\ \frac{\partial^2 l}{\partial \varphi \partial \eta} & \frac{\partial^2 l}{\partial \eta \partial \xi} & \frac{\partial^2 l}{\partial \eta^2} & \frac{\partial^2 l}{\partial \eta \partial \alpha} \\ \frac{\partial^2 l}{\partial \varphi \partial \alpha} & \frac{\partial^2 l}{\partial \alpha \partial \xi} & \frac{\partial^2 l}{\partial \eta \partial \alpha} & \frac{\partial^2 l}{\partial \alpha^2} \end{bmatrix} \tag{28}$$

The solutions to be obtained by solving (28) will yield the asymptotic variance and covariances of the parameters $\hat{\varphi}, \hat{\xi}$, $\hat{\alpha}$ and $\hat{\eta}$. Using (28), the approximate $100(1-\lambda)\%$ confidence intervals for φ, ξ, α and η can be expressed as
$\hat{\varphi} \pm Z_{\frac{\lambda}{2}}\sqrt{\hat{\Sigma}_{11}}, \hat{\xi} \pm Z_{\frac{\lambda}{2}}\sqrt{\hat{\Sigma}_{22}}, \hat{\eta} \pm Z_{\frac{\lambda}{2}}\sqrt{\hat{\Sigma}_{33}}, \hat{\alpha} \pm Z_{\frac{\lambda}{2}}\sqrt{\hat{\Sigma}_{44}}$
where $Z_{\frac{\lambda}{2}}$ is the upper λ^{th} percentile of the standard normal distribution. where

$$\frac{d^2}{d\eta^2} = \sum_{i=1}^{n}\left[\frac{2\varphi\left(1-\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\right)}{\eta^3}\right] \tag{29}$$

$$\frac{d^2}{d\varphi^2} = -\frac{n}{\varphi^2} \tag{30}$$

$$\frac{d^2}{d\alpha^2} = -\frac{n}{\alpha^2} \tag{31}$$

$$\begin{aligned}\frac{d^2}{d\xi^2} = & -\frac{2n}{\xi^2} - \sum_{i=1}^{n}\left[\frac{2\eta\alpha e^{-\frac{\xi^2}{x^2}}}{x^2\left(1-e^{-\frac{\xi^2}{x^2}}\right)}\right] + \sum_{i=1}^{n}\left[\frac{4\eta\alpha\xi^2 e^{-\frac{\xi^2}{x^2}}}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)}\right] + \sum_{i=1}^{n}\left[\frac{4\eta\alpha\xi^2\left(e^{-\frac{\xi^2}{x^2}}\right)^2}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)^2}\right] \\ & -\frac{2n}{x^2} - \sum_{i=1}^{n}\left[\frac{4\varphi\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\eta\alpha^2\xi^2\left(e^{-\frac{\xi^2}{x^2}}\right)^2}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)^2\eta}\right] + \sum_{i=1}^{n}\left[\frac{2\varphi\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\eta\alpha e^{-\frac{\xi^2}{x^2}}}{x^2\left(1-e^{-\frac{\xi^2}{x^2}}\right)\eta}\right] \\ & -\sum_{i=1}^{n}\left[\frac{4\varphi\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\eta\alpha\xi^2 e^{-\frac{\xi^2}{x^2}}}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)\eta}\right] - \sum_{i=1}^{n}\left[\frac{4\varphi\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\eta\alpha\xi^2\left(e^{-\frac{\xi^2}{x^2}}\right)^2}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)^2\eta}\right] \\ & -\sum_{i=1}^{n}\left[\frac{2e^{-\frac{\xi^2}{x^2}}}{x^2\left(1-e^{-\frac{\xi^2}{x^2}}\right)}\right] + \sum_{i=1}^{n}\left[\frac{4\xi^2 e^{-\frac{\xi^2}{x^2}}}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)}\right] + \sum_{i=1}^{n}\left[\frac{4n\xi^2\left(e^{-\frac{\xi^2}{x^2}}\right)^2}{x^4\left(1-e^{-\frac{\xi^2}{x^2}}\right)^2}\right]\end{aligned} \tag{32}$$

$$\frac{\partial^2}{\partial\xi\partial\varphi} = \sum_{i=1}^{n}\left[\frac{2n\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\eta\alpha\xi e^{-\frac{\xi^2}{x^2}}}{x^2\left(1-e^{-\frac{\xi^2}{x^2}}\right)\eta}\right] \tag{33}$$

$$\frac{\partial^2}{\partial\alpha\partial\varphi} = 0 \tag{34}$$

$$\frac{\partial^2}{\partial\eta\partial\varphi} = -\sum_{i=1}^{n}\left[\frac{\left(1-\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\right)}{\eta^2}\right] \tag{35}$$

$$\frac{\partial^2}{\partial\eta\partial\alpha} = 0 \tag{36}$$

$$\frac{\partial^2}{\partial\eta\partial\xi} = -\sum_{i=1}^{n}\left[\frac{2n\varphi\left(1-e^{-\frac{\xi^2}{x^2}}\right)^{-\eta\alpha}\eta\alpha\xi e^{-\frac{\xi^2}{x^2}}}{x^2\left(1-e^{-\frac{\xi^2}{x^2}}\right)\eta^2}\right] \tag{37}$$

$$\frac{\partial^2}{\partial\alpha\partial\xi} = 0 \tag{38}$$

4. Data Analysis

4.1. Simulation Studies

In this section, we simulated data set for sizes n = 30,100,200 and 500 that follows Gompertz Exponentiated Inverse Rayleigh distribution using different parameter values for the four parameters φ, η, α and ξ using the quantile function (inverse transformation method of simulation). We considered the following combinations for the parameters $(\varphi, \eta, \alpha, \xi)$ = ((0.5,1,1,0.8), (1.5,0.6,1.2,1.5), (0.5,0.5,0.5,0.5) and (1,0.5,0.5,1))at different sample sizes n = 30, 100, 200, and 500. The results presented in Table 1 displayed the true values of $(\varphi, \eta, \alpha, \xi)$ and estimated values of $(\varphi, \eta, \alpha, \xi)$ with the standard errors. The results are replicated 10,000 times and the average result were presented in the Table 1.

Table 1: *The MLE estimates and their MSE for different parameter values*

		φ	η	α	ξ	MSE_φ	MSE_η	MSE_α	MSE_ξ
30	φ = 0.5	0.4164	0.6713	1.0835	0.6710	5.9423	9.5783	15.4494	0.1593
100	η = 1	0.5453	1.4859	0.6865	0.7407	2.7936	7.6042	3.5094	0.1109
200	α = 1	0.4947	0.9795	1.0074	0.7908	2.6984	5.3368	5.4865	0.0784
500	ξ = 0.8	0.4410	1.3019	0.8647	0.7593	1.2450	3.6662	2.4340	0.0556
30	φ = 1.5	1.1616	0.7094	0.8432	1.2341	16.8549	10.3031	12.2263	0.2242
100	η = 0.6	1.0910	1.0760	0.8447	1.3581	8.7396	8.6236	6.7624	0.1484
200	α = 1.2	1.0830	0.6560	1.2001	1.4536	5.4328	3.2944	6.0152	0.1018
500	ξ = 1.5	0.8998	0.7517	1.2599	1.4447	5.4193	4.5324	7.5899	0.0701
30	φ = 0.5	1.8939	1.3197	0.1331	0.4055	14.6817	10.2453	1.0298	0.0964
100	η = 0.5	0.4297	0.5878	0.4620	0.4533	3.1636	4.3259	3.3989	0.0672
200	α = 0.5	0.4996	0.5109	0.4939	0.4895	4.1287	4.2278	4.0834	0.0474
500	ξ = 0.5	0.5095	0.7248	0.4092	0.4760	1.5357	2.1832	1.2322	0.0317
30	φ = 1	1.3023	0.5704	0.3324	0.8097	27.7198	12.1465	7.0727	0.1621
100	η = 0.5	0.9354	0.7846	0.3940	0.9009	4.7474	3.9884	1.9979	0.1091
200	α = 0.5	0.9282	0.5241	0.5117	0.9699	4.5882	2.5960	2.5287	0.0763
500	ξ = 1	0.9232	0.7251	0.4500	0.9601	5.0199	3.9469	2.4472	0.0514

4.2. Data Description

The strength data was originally reported by Badar and Priest [5] where the strength is measured in GPA for single carbon fibers and impregnated 1000-carbon fiber tows at gauge lengths of 20 mm. These data set were fitted to GEIR distribution, the Half- Logistics Inverse Rayleigh (HLIR) distribution by Almarashi et al [3] and the Type II Topp-Leone Inverse Rayleigh (T2TLIR) distribution by Mohammed and Yahia [14]. Other distributions that have been fitted to these same data are the Transmuted Inverse Rayleigh distribution (TIR) by Ahmad et al [2], the Odd Frechet Inverse Rayleigh (OFIR) distribution by Elgarhy and Alrajhi [8], one parameter Inverse Rayleigh (IR) by Trayer [20].

Data I: carbon fibers Strength (20mm) Data set

1.312, 1.314, 1.479, 1.552, 1.700, 1.803, 1.861, 1.865, 1.944, 1.958, 1.966, 1.997, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.14, 2.179, 2.224, 2.240, 2.253, 2.270, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382, 2.382, 2.426, 2.434, 2.435, 2.478, 2.490, 2.511, 2.514, 2.535, 2.554, 2.566, 2.57, 2.586, 2.629, 2.633, 2.642, 2.648, 2.684, 2.697, 2.726, 2.770, 2.773, 2.800, 2.809, 2.818, 2.821, 2.848, 2.88, 2.954, 3.012, 3.067, 3.084, 3.090, 3.096, 3.128, 3.233, 3.433, 3.585, 3.585.

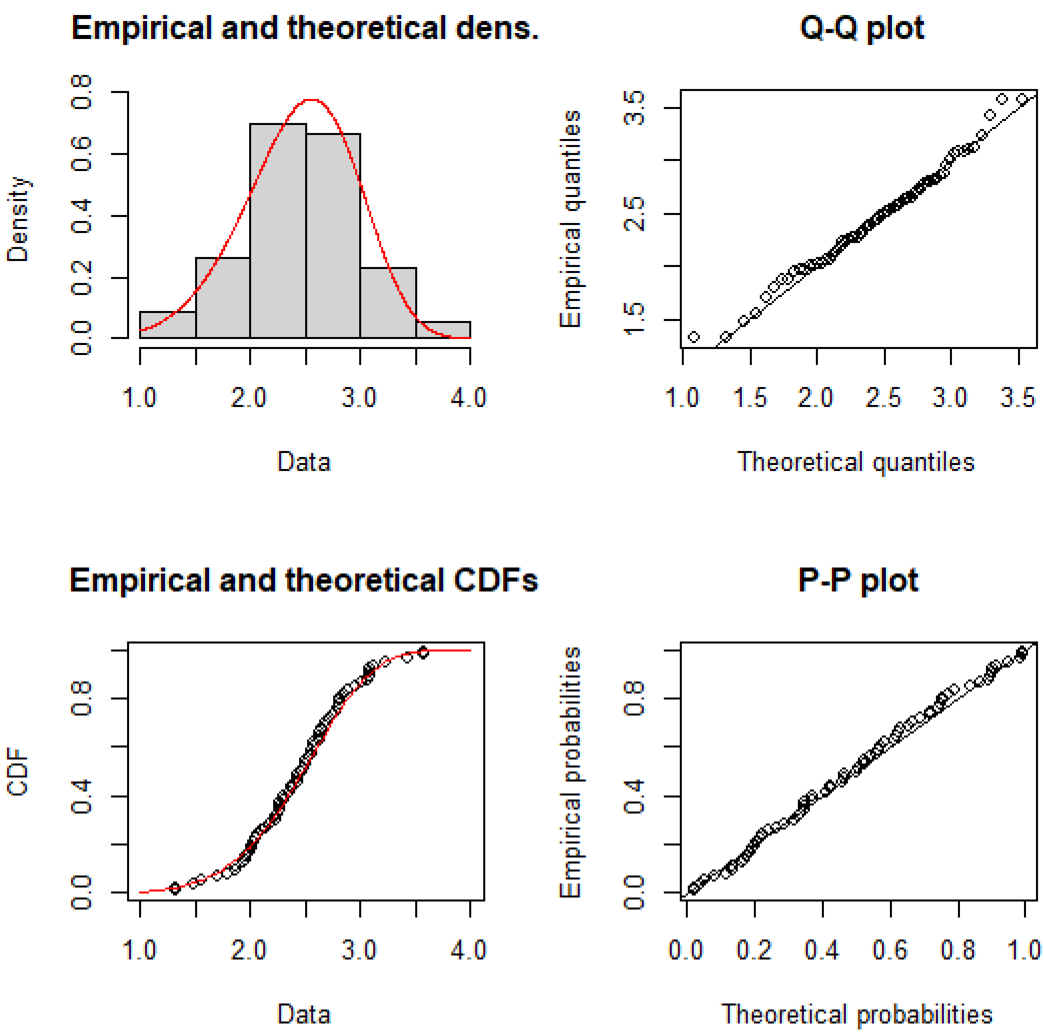


Figure 5: *Empirical and theoretical plot of Carbon*

Table 2 shows the summary statistics of the GEIR, HLIR, T2TLIR, TIR, OFIR, and IR distributions. These five distributions are fitted to data 1 using maximum likelihood estimation.

Table 2: Goodness-of-fit measures based on AIC, BIC, HQIC, K-S values for the Strength (20mm) data set (data 1)

Models	AIC	BIC	HQIC	K-S Value	P-value
GEIR($\varphi, \eta, \alpha, \xi$)	104.43	107.367	109.112	0.021	0.9321
HLIR(α, λ)	105.003	109.472	106.776	0.0596	0.9668
T2TLIR(α, θ)	108.137	112.605	109.91	0.0776	0.7993
TIR(θ, λ)	145.879	148.113	146.765	0.254	0.0002
OFIR(θ, α)	147.423	151.891	149.196	0.1801	0.0227
IR(α)	178.826	181.06	179.713	0.3549	0

Data II: Patients receiving an analgesic dataset

The data set is taken from Gross and Clark [11] which consists of 20 observations of patients receiving an analgesic 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0

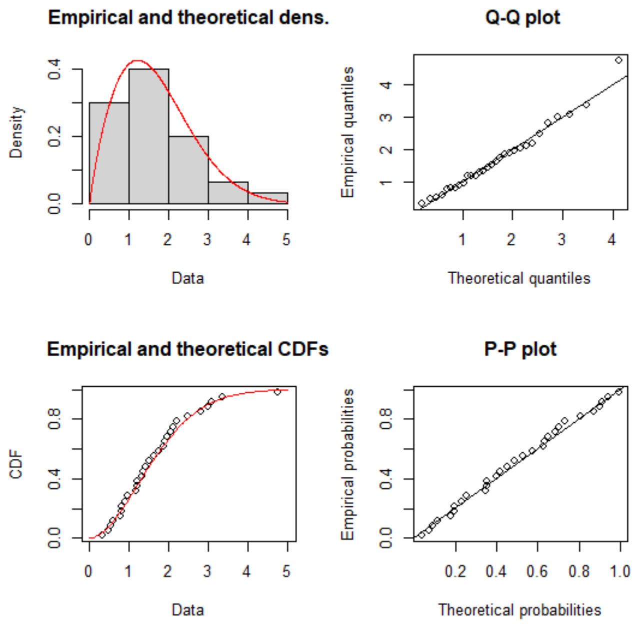


Figure 6: *Empirical and theoretical plot for Patients receiving an analgesic*

Table 3: Estimates and Goodness-of-fit measures based on AIC, BIC, HQIC, K-S values for Patients receiving an analgesic

Distribution	Parameters				AIC	CAIC	BIC	HQIC	Pvalue
GEIR	2.3362134	-0.3288	2.4818	2.2056	39.356	40.212	43.338	40.1332	0.4493
EIR	0.8714	3.1686			46.365	47.0709	48.3564	46.7537	0.1435
WR	11.8552	1.2364	0.0545		48.5149	50.0149	51.5021	49.098	0.4597
GR	3.2748	0.6926			40.805	41.5109	42.7965	41.1938	0.463

5. Conclusion

In this study, a proposed four parameter distributions are added to Gompterz family of distribution called Gompterz exponentiated Inverse Rayleigh (GEIR). Some structural mathematical properties; Moment, Order Statistic, Skewness and kurtosis of the derived model are obtained. A simulation study is carried out to estimate the behaviour of the shape and scale parameters, also maximum likelihood estimation method was employed to estimate the parameters of the distribution and simulation studies were performed to assess the flexibility of the proposed distribution. For the simulated dataset, the result presented in Table (1), from the result, we observed that the estimated values gotten are close to the predefined parameters and that as n increases the MSE reduces which confirms to the law of large numbers.

However, application of two real-life data set shows that the GEIR has strong and better fit than other competing models i.e., the data sets were fitted to the Half- Logistics Inverse Rayleigh (HLIR) distribution and the Type II Topp-Leone Inverse Rayleigh (T2TLIR). Other distributions that have been fitted to these same data are the Transmuted Inverse Rayleigh distribution (TIR), the Odd Frechet Inverse Rayleigh (OFIR) distribution, exponentiated inverse Rayleigh distribution (EIR), Weibul Rayleigh (WR), Gamma Rayleigh (GR) , one parameter Inverse Rayleigh (IR) distributions using goodness of fit and information criterion.

References

[1] Abdal-Hameed M.K., Khaleel2 M.A., Abdullah Z.M, Oguntade P.E., and Adejumo A.O(2018) Parameter Estimation and Reliability, Hazard Functions of Gompertz Burr Type XII Distribution. *Tikrit Journal of Administration and Ecomomics Sciences*, *14(1)*, 381 - 400

[2] Ahmad, A, Ahmad S.P.,and Ahmed A. (2014) Transmuted Inverse Rayleigh distribution: A generalization of the inverse rayleigh distribution. *Math. Theory Model*, *4(7)*, 111-131

[3] Almarashi A.M., Majdah M.B., Elgarhy M., Jamal F. and Chesneau C. (2020). Statistical Inference of the Half-Logistic Inverse Rayleigh Distribution. *Entropy*, *2020,22*, 449

[4] Alizadeh, M., Cordeiro, G. M., Pinho, L.G.B. and Ghosh, I. (2017). The Gompertz-G family of distributions. *Journal of Statistical Theory and Practice*, *11(1)*, 179-207.

[5] Bader, M. and Priest, A. (1982) Statistical Aspect 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

[6] Brown, K. and Forbes, W.A. (1974). Mathematical model of aging processes. *Journal of Gerontology*, *29(1)*, 46-51.

[7] Economos, A. (1982). Rate of aging, rate of dying and the mechanism of mortality. *Archives of Gerontology and Geriatrics*, *1*, 46-51.

[8] Elgarhy M. and Alrajhi, S (2019) The odd Frechet inverse Rayleigh distribution: Statistical Properties and applications *J. nonlinear Sci. Appl 12*, 291-299

[9] Galton, F. (1983). Enquiries into Human Faculty and its Development. *Macmillan and Company, London.*

[10] Gompertz B. (1825). On the nature of the function expressive of the law of human mortality and on a new model of determining life contigencies. *Phil Trans. R. Soc.*, *115*, 513-585

[11] Gross, A.J. and Clark, V.A. (1975). Survival Distributions: Reliability Application in the Biometrical Sciences, *John Wiley, New York*.

[12] Halid O.Y, and Sule, O.B. (2022). A Classical and Bayesian Estimation Techniques for Gompertz Invere Rayleigh Distribution: Properties and Application *Pakistan Journal of Statistics*, *38(1)*, 49-76

[13] Ieren T.G., and Balogun O.S(2021). Exponential-Lindley Distribution: Theory and Application to Bladder Cancer Data. *Journal of Applied Probability and Statistics*, *16(2)*, 129-146.

[14] Mohammed, H.F, and Yehia N.(2019) The type II Topp-Leone generalized inverse raleigh distribution *Int. J. Contemp. Math. Sci 14(3)*, 113-122

[15] Moors, J.J. (1988). A quantile alternative for kurtosis. *The Statistician*, *37*, 25-32.

[16] Preston, S., Heuveline, P. and Guillot, M. (2001). Demograpy. Measuring and modeling population processes. *Blackwell Publisher, Oxford.*

[17] Rao G.S and Mbwambo S. (2018) Exponentiated Inverse Rayleigh Distribution and Application to Coating Weights of Iron Sheets Data. *Journal of Probability and Statistics*, *2019* 1-13

[18] Vaupel, J. (1986). How Change in Age-specific Mortality Affects Life Expentancy. *Population Studies*, *40*, 147-157

[19] Willemse, W. and Koppelaar, H. (2000). Knowledge elicitation of Gompertz law of mortality. *Scandinavian Actuarial Journal*, *2*, 168-179.

[20] Trayer, V.N.(1964) Proceedings of the Academy of Science Belarus, USSR

MODELLING OF RELIABILITY INDICATORS BY MEANS OF THE INTERFERENCE METHOD

Alena Breznická[1] Pavol Mikuš[2]

•

Faculty of Special Technology, Alexander Dubček University of Trenčín,
Ku kyselke 469, 911 06, Trenčín, Slovakia[1,2]
alena.breznicka@tnuni.sk
pavol.mikus@tnuni.sk

Abstract

The article describes the application of the simulation modeling of tasks of the interference theory of reliability (SSI - stress - strength interference reliability model) in the MATLAB programming language. It points to the possibility of creating your own purpose-built models designed to predict the evolution of the reliability of the technical system during the user's interactive activity. Reliability simulation by changing load and resistance parameters makes it possible to find acceptable reliability parameters.

Keywords: reliability, MATLAB programming, failure

1. Introduction

Systems simulation is currently one of the most progressive means to investigate the operation of complex systems, which can be applied to most technical problems showing the nature of service processes. Simulation modelling of processes nowadays forms the basis of scientific research in several scientific fields. Even though modelling methods and simulation tools were developed based on different scientific disciplines, knowledge and methods of system simulation can be transferred from one discipline to another and applied in several scientific fields. The theoretical basis for the management of such a specific area is the knowledge from various scientific disciplines, which enter the research process and thus enable the research of the given area.

From the research possibilities of the given methods, it follows that computer simulation and modeling of systems has a primary position, which allows us to:

- describe and express processes whose analytical solution we do not know,
- to simplify the solution of complex mathematical problems,
- shorten time-consuming experiments,
- carry out experiments of new projects,
- quickly and efficiently analyze changes and assess their consequences,
- examine many failure states and assess their variants.

It is most appropriate to use the following procedure to examine the systems defined in this way:

1. We will determine the properties of system elements and the relationships between them and the environment through analysis. We select and consider only those that are essential from the point of view of the function of the system and the objectives of the investigation. We will neglect the unimportant features of the system.
2. We will acquire specific knowledge of the necessary statistical characteristics (e.g. the reliability of machines and equipment, which allow us to express the probability of the occurrence of random failures - failure-free, their groups, subgroups and requirements for restoring their operability - repairability). In this way, we can express the basic states out of several possible ones in which the system element is located.
3. We will describe the system theoretically. We create an idea about the system of the so-called theoretical model. We assign the modeled system to a different system based on similarity, according to certain criteria - called a model, or computer model.
4. We transform the system or its subsystem into a computer simulation model that imitates the current idea of the simulated system and its movement by implementing a simulation program on the computer.
5. We will perform simulation experiments with the simulation model, which allow us to monitor the values of important parameters of the model, implement changes in the studied model and determine their effects on the function of the system.
6. Based on the evaluation of the simulation experiments, we will propose or implement changes in the investigated system

The inference theory of reliability originates from a concept based on the comparison of the mutual connection of the selected reliability quantity. It is understood and described by deterministic principles and is based on determining and respecting the values of reliability indicators [1]. The approaches described by the inference theory of reliability are based on the assumption that a failure or a failure function occurs when the resistance limit of the object is exceeded, i.e. ability to withstand stress. An element of a technical system can fulfil its function if it must be sufficiently resistant to all loads that may act on it. From the internal and external environment that influence such an element.

This is an assessment approach that is based on the deliberate oversizing of the object with the expression of the safety factor against failure (SF - safety factor) or safety level (SM - safety margin). If the stress R exceeds the strength value S of the element given by the structural design, a failure will occur.

$$SF = \frac{\overline{S}}{\overline{R}} \; or \; SF = \frac{Smin}{Rmax} \tag{1}$$

$$SM = \frac{S\,min - Rmax}{Smin} \tag{2}$$

Where:
SF – safety factor
SM – safety margin
L – stress
R – strength

Numerical values of the required or acceptable SF and SM depend on the technical department, type of product, consequences of the failure on safety, risks, etc. They are determined by standards and regulations.

The concept of stochastic description assumes that stress as well as resistance to failure are generally of a random nature with a certain law of distribution of the probability of occurrence. They can occur in a wide range of values, with the possibility of non-overlapping or interference.

For a better characteristic of resistance to failure, we can express reliability as follows:

$$SM = \frac{\bar{R}-\bar{L}}{(\delta_R^2+\delta_L^2)^{\frac{1}{2}}} \quad (3)$$

and LR (Loading roughness)

$$LR = \frac{\delta_L}{(\delta_R^2+\delta_L^2)^{\frac{1}{2}}} \quad (4)$$

The SM factor characterizes the relative distance between the mean stress and resistance values. The LR factor characterizes the impact of the standard deviation of stress on safety. The input values of stress L and strength R are random in nature. This can be obtained from an experimental measurement, or they are determined by other methods / e.g. by expert estimate/. The result is either a non-parametric distribution of the obtained data, which we can statistically process in the form of a histogram or convert it to a usable parametric distribution. Both cases provide us with the possibility of generating input quantities and assessing the occurrence of decisive events for the statistical expression of failure rate or failure-freeness of elements using the interference method.

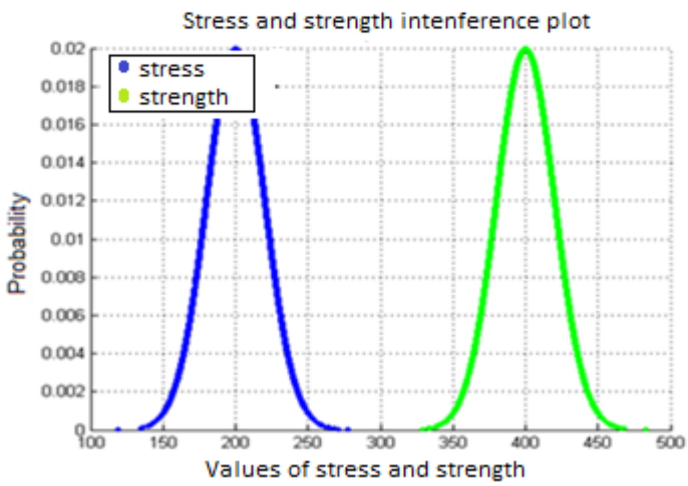


Figure 1: *Deterministic concept of load and resistance simulation*

2. Possibilities of reliability simulation

The simulation in figure 1 presents a situation that occurs very often, especially when applying high values of coefficients with small variances of R and L. Stress and resistance are very far from each other with medium values, they have a small variance, a low LR value and a high SM value. If it is possible to influence the variances and standard deviations of stress and resistance quantities and to move the mean values of both quantities as far as possible from each other, these quantities cannot interfere with each other, and then there is a high probability that the object will not fail during its entire life. Since the quantities R and L do not interfere with each other, the element is reliable. If there is a situation where the distributions of stress and resistance overlap, there is mutual interference and thus the possibility of a malfunction - the system thus becomes unstable, and a malfunction may occur. The SM and LR factors enable a deterministic analysis of the mutual interference of stress and strength and the possibility of failure. The interference theory of reliability is based on the analysis of regularities and properties of two random variables that characterize the elementary properties of reliability, as a distribution function. We will call the distribution function of the random variable X the real function of the real variable F(x) defined by the relation:

$$F(x) = P(X < x) = P(X \in (-\infty, x)) \text{ for any } x \in R. \tag{5}$$

The distribution function has the following properties:

1. for each $x \in R$: $0 \leq F(x) \leq 1$,
2. for each $x1, x2 \in R$: if $x1 < x2$, then $F(x1) \leq F(x2)$
3. $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$

Both the stress and the resistance are random variables, characterized by quantities or random processes. The mechanical structure Mk, which is subjected to operational load at the time, will be reliable if the operational stress L with a certain probability does not exceed the resistance R (bearing capacity, permissible stress, etc.). Quantities characterizing operational stress (load) and resistance (bearing capacity) of the structure can be expressed by probability densities and distribution functions. Let $f_L(L)$ denote the probability density for the stress random variable L and f_S (R) the probability density for the fault resistance random variable S. Let's denote the distribution function for the random stress variable L by $F_L(L)$ and $F_R(R)$ the distribution function for the random variable S against failure. The quantities L and R are random, and we assume that they have a specific probability distribution law (continuous or discrete). They can influence each other (interfere) [3,4].

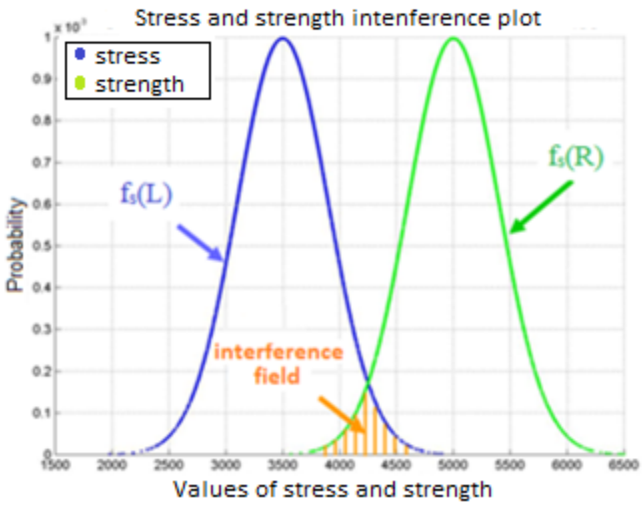


Figure 2: *Interference of operating load and permissible stress*

The simulation model of load probability density and resistance must follow the following steps:

1. Determining the appropriate number of simulations.
2. To analyze the input data of the distribution parameters of the probability density functions of the operating load and the resistance of the assessed system element.
3. Determine input parameters of histograms or probability distributions.
4. Generate the level of operational load and resistance.
5. Statistically process a set of values in the form of values of probability density functions and distribution functions.
6. Graphically display the courses of probability density functions and distribution functions.

Statistically process the generated data into the form of values of probability density functions and distribution functions, and by plotting them we get an idea of their interference.

Calculate

- probability of failure PF = n / N,
- probability of reliability PR = 1 - PF

7. Determine the minimum value of the resistance function and the maximum value of the load function, i.e. the interference interval of the generated values.
8. Graphically show the mutual dependence of L and R values.
9. Calculate the size of the area under the graph line.
10. Determine the resulting of reliability.

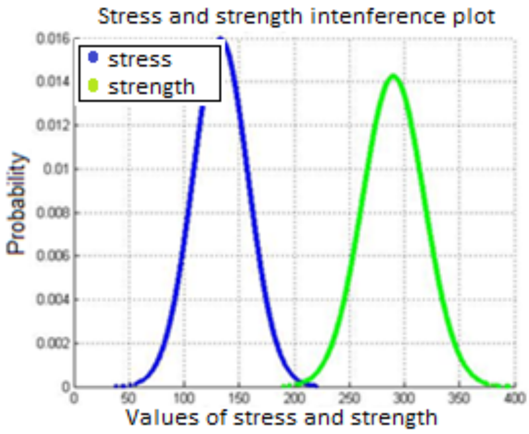


Figure 3: *Load and resistance probability density distribution functions*

The boundary between the failure and reliability areas is given by the condition R≥L and is expressed by a red line.

Basic conditions for SSI reliability modeling:

- The parameters for stress and force are statistically independent.
- Dynamic parameters for loading forces and stress are described by the Poisson distribution.
- Random strength degradation is described by a distribution.
- Random strength degradation and deterministic loading force.

Stress L and strength R are time-dependent random variables. The load values may have the character of an increasing gradient, while the resistance tends to decrease due to external factors dependent on the environment.

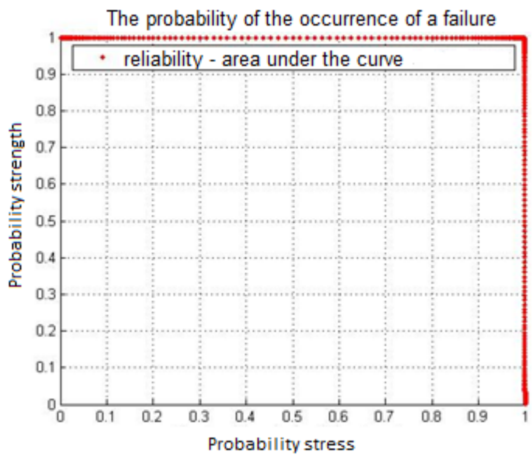


Figure 4: *Graphic representation the probability of failure*

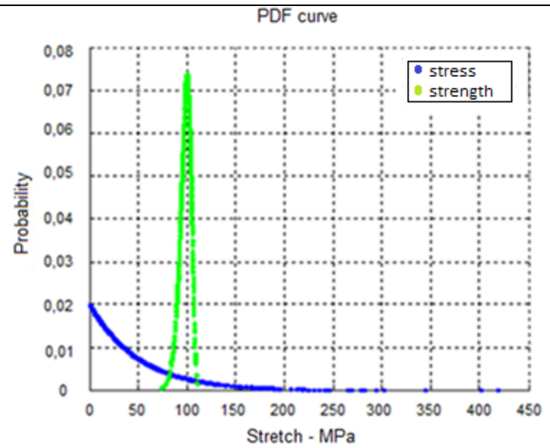


Figure 5: *PDF curve*

The MATLAB simulation language and its graphics make it possible to create interactive programs, the environment of which allows the user to dialogically change the parameters of the distributions and to judge what load and resistance values are acceptable for the structural design application [5]. The program in the basic window offers the option of choosing the number of simulations, the type of load distribution and the resistance of the element under investigation and the distribution parameters. If, from the input data used, the simulation results indicate that the required fault-free parameters do not meet, the simulations can be carried out by changing the parameters until an acceptable level of interference is reached.

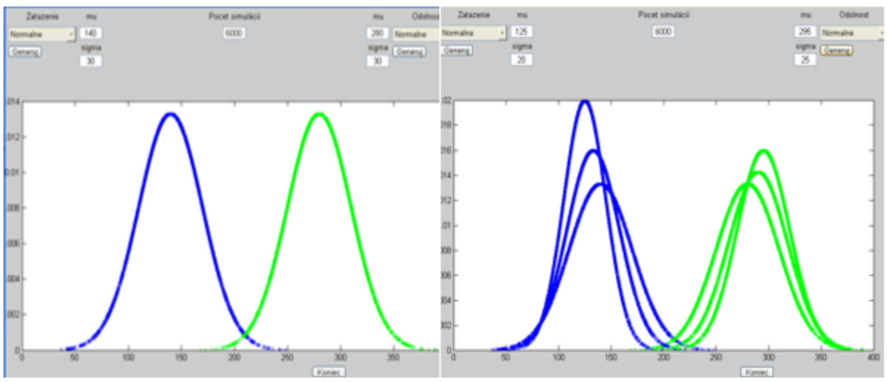


Figure 6: *Analysis of the impact of changes in stress and strength parameters*

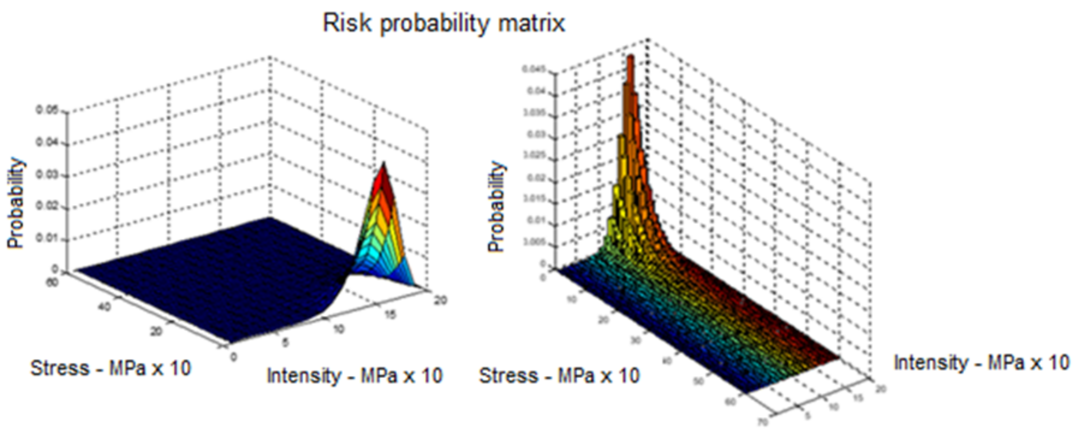


Figure 7: *Analysis 3D model*

3. Conclusion

The interference stress-strength reliability assessment model is a model for predicting the probability of failure for various engineering systems. For an objective modeling result, it is necessary to have known probability distributions of stress and strength available. Failure or failure is defined as the ratio of stress > strength. If both stress and strength distributions are acceptable, the simulation provides an analytical solution that provides an accurate result. The technical perfection of all means of computer simulation and modelling often does not allow the creation of a suitable simulation model and its correct application without deep knowledge of the nature of the investigated phenomena or systems [6,7]. The quality of the system determination depends on the solver of the model and the obtained input data necessary for its determination. A simulation model or any forecast is a very complex process. A complex part of the methodology, requiring good knowledge of the simulation process, is the processing of algorithms for individual events and program events. Nevertheless, let's know even the most complex algorithm is created from simple algorithms. Such a simulation model makes it possible to eliminate the imperfections of calculation methods.

It uses the results for few measurements of experimental methods in an appropriate way, can apply the results to determine the reliability of elements of diverse systems and determine the interference of different probability density distributions of randomly variable functions of operational loads and resistances.

References

[1] Kumar, N. and Malik, S. C. and Nandal, N. (2022). Stochastic analysis of a repairable system of non-identical units with priority and conditional failure of repairman, RT&A, No 1 (67), Volume 17.

[2] Saroj, A. and Sonker, P.K. and Kumar, M. (2022). Statistical properties and application of a transformed lifetime distribution: inverse muth distribution, RT&A, No 1 (67) Volume 17.

[3] Mathew, J. and George, S. (2020). Length biased exponential distribution as a reliability model: a bayesian approach, RT&A, No 3 (58) Volume 15.

[4] Chovanec A. and Kianicová M. and Bucha J. and Ambrózy A. (2005). Aplikácia simulačného modelovania pri riešení úloh interferenčnej teóre spoľahlivosti, In: Kvalita a spoľahlivosť technických systémov, ISBN 80-8069-518-0, Page 131-134.

[5] Leitner, B. (2020). The procedure of operational risks management in railway companies, In. Transport Means - Proceedings of the International Conference, Pages 57 - 622020 24th International Scientific Conference on Transport Means, Kaunas, ISSN 1822296X.

[6] Dvořák, Z. and Leitner, B. (2018). Software tool for railway traffic modelling under conditions of limited permeability, In. Transport Means - Proceedings of the International Conference, Page 448–454, ISSN 1822296X.

[7] Vališ, D. and Zak, L. and Vintr, Z. (2019). System Condition Assessment Based on Mathematical Analysis, Conference Proceedings, ISBN 978-153866786-6, IEEE International Conference on Industrial Engineering and Engineering Management, Pages 222 – 226.

SELECTION PROCEDURE FOR SKIP LOT SAMPLING PLAN OF TYPE SKSP-R BASED ON PERCENTILES OF EXPONENTIATED RAYLEIGH DISTRIBUTION

P. Umamaheswari
•
Assistant Professor,
Sona College of Arts & Science
uma_m2485@yahoo.com

K. Pradeepa Veerakumari
•
Assistant Professor, Department of Statistics,
Bharathiar University,
Coimbatore, Tamil Nadu, India
pradeepaveerakumari@buc.ac.in

S.Suganya
•
Assistant Professor, Department of Statistics,
PSG College of Arts & Science,
Coimbatore, Tamil Nadu, India
suganstat@gmail.com

Abstract

In this study, acceptance sampling techniques are effective for reducing the cost and time of the submitted lots. In this hectic environment, a high level of product reliability and quality assurance is expected. Use the abbreviated life tests in the acceptance sampling plan as a result. To make a choice on the product, sampling plans with time-truncated life tests are used. This study uses percentiles under the exponentiated Rayleigh distribution to build a skip lot sampling plan of the SkSP-R type for a life test. A truncated life test may be carried out to determine the minimum sample size to guarantee a specific percentage life time of products. In particular, this paper highlights the construction of the Skip lot Sampling Plan of the type SkSP-R by considering the Singe Sampling Plan (SSP) and Double Sampling plan(DSP) as reference plans for life tests based on percentiles of Exponentiated Rayleigh Distribution (ERD). Calculations are made for various quality levels to determine the minimum sample size, prescribed ratio, and operational characteristic values. The proposed sampling plan, which is appropriate for the manufacturing industries for the selection of samples, is also analyzed in terms of its parameters and metrics. Illustrations are provided to help you comprehend the plan. In addition, it addresses the feasibility of the new strategy.

Keywords: Exponentiated Rayleigh Distribution, Percentiles, Life tests, Single Sampling Plan, SkSP –R.

I. Introduction

The term used to define the collection of statistical instruments used by quality engineering professionals is Statistical Quality Control (SQC). Acceptance sampling is one of the main fields of statistical quality control. Acceptance sampling is a technique that deals with the procedures by which the acceptance or rejection decision is based on sample inspection. Between no inspection and 100% inspection, acceptance sampling is the "middle of the road" method. Two key classifications of acceptance plans are available: by attributes and variables. There is a possibility, therefore, of rejecting a good lot (the risk of the producer) and accepting a poor lot (the risk of the consumer). Skip-lot sampling proposals are intended to decrease the cost of inspection.

Samples can be drawn from just a fraction of the submitted lots under the skip-lot sampling inspection. The main aim of the skip-lot sampling plan is to reduce the pace of inspection of samples, thus reducing the overall cost of inspection. Dodge [6] initially suggested the Skip-lot sampling plan, requiring a single decision or review to assess the acceptability or non-acceptability of the lot. These plans are called Skip-lot sampling plans of type SkSP-1. The time-truncated acceptance sample technique has fascinated enormous writers such as Epstein [7], Sobel and Tischendr [16], Goode and Kao [8], Gupta and Groll [9], Kantam et al. [10,11], Baklizi [3], Tsai and Wu [18], Balakrishnan et al . [4], Aslam and Shahbaz [2], Rao et.al. [15], Aslam et al. [1], Pradeepa Veerakumari and Ponneeswari [13], Pradeepa Veerakumari et.al [14], Suganya and Pradeepa Veerakumari [17].

Balamurali et.al., [5] introduce new skip-lot sampling plan of resampling, it is named as SkSP-R. The SkSP-R plan is a sampling technique based on the concept of the continuous sampling procedure and the resampling scheme for continuous bulk product flow consistency inspection. As a percentile-based comparative plan with Exponential Rayleigh Distribution, the paper focuses on developing a SkSP-R life-test plan with a single sampling plan. Rayleigh Exponential Distribution is an important distribution and reliability study of life studies. It has some of the basic structural properties and exhibits great consistency in mathematics Most properties of the distribution of Exponentiated Rayleigh are similar to those of gamma, Weibull and exponential distribution. The functions of ERD for distribution and density are in comparable forms. This is easily applied to the truncated plans as a consequence. The ERD distribution's cumulative function is given by,

$$F(t;\tau,\theta) = \left[1 - e^{-1/2(t/\tau)^2}\right]^{\theta},\ t>0,\ 1/\tau>0,\ \theta>0 \quad (1)$$

Where, τ and θ are the scale and shape parameters respectively. Its probability density function is the first derivative of any cumulative distribution function. The likelihood density function of ERD can therefore be written as,

$$f(t;\tau,\theta) = \theta\left[1 - e^{-1/2(t/\tau)^2}\right]^{\theta-1}\left[\frac{t}{\tau^2} e^{-1/2(t/\tau)^2}\right] \quad (2)$$

SSP were suggested by Pradeepa Veerakumari and Ponneeswari [13] for life research based on the ERD percentiles. Subsequently, Skip-lot sampling plans for life testing based on the percentiles of ERD were developed by Pradeepa Veerakumari et.al [14].

II. Skip-lot Sampling Plan SkSP-R

When opposed to a single sampling plan, the skip-lot sampling plan (SkSP) offers a smaller sample size during the inspection. To minimise the inspection cost, skip-lot sampling plans have been commonly used in industries. In order to establish the sampling plan that is designated as

SkSP-R, the lot resubmission considered under the resampling method was introduced. In circumstances where resampling is allowed on lots not approved during the original inspection, the desired sampling plan may be used. Therefore, if the idea of the resampling concept is implemented, a skip-lot sampling strategy is supposed to be more successful. In order to decide the optimal parameters of the sampling plan, the optimum combination of plan parameters was derived to protect both the manufacturer and the user. With the aid of the reference attribute sampling method, the Skip-lot Sampling Plan-R demonstrates that it is similar to the SkSP-2 plan developed by Perry [12].

III. Operating procedure for SkSP-R

i. Use the reference plan to start a standard inspection. During the usual inspection, the lots are inspected in the order of application, one by one.

ii. On normal inspection, discontinue normal inspection when 'i' consecutive lots are admitted, and turn to skipping inspection.

iii. Inspect just a fraction of 'f' of the randomly picked lots during skipping inspection. Until a sampled lot is refused, the skipping inspection is continued.
iv. If a lot is rejected after 'k' has been approved consecutively by sampled lots, then go to the immediate next lot for the resampling procedure as in step (5) below.
v. Carry out the inspection using the reference plan during the resampling process. If the lot is approved, then proceed to skip the inspection. Resampling is performed on non-acceptance of the lot for 'm' times and the lot is rejected. If, it has not been approved on resubmission (m-1).
vi. If a lot is rejected in a resampling system, then return to regular inspection in phase (1) immediately.
vii. Replace or correct all non-conforming units discovered in the rejected lots with conforming units.

The proposed plan was developed with a reference plan and defined by four parameters, namely $'f'$ - the fraction of lots inspected in the skipping inspection mode, $'i'$ - the usual inspection clearance number, $'k'$ - the sampling inspection clearance number, and $'m'$ - the number of times the lots are submitted for resampling.

IV. Operating procedure of SkSP-R with SSP as a reference plan based on percentiles of ERD

Step 1: A random sample of size n is drawn and a random sample of size is drawn and positioned on a time test.

Step 2: The number of defects d is counted and the approval number c is compared.

i. If $d>c$, then the lot is refused.
ii. If $d\leq c$, then the lot is approved.

Step 3: If $d>c$, terminate the test and reject the lot if it is obtained before the stated time.

V. Operating characteristic function for SkSP-R using Single Sampling Plan

The OC function is the most commonly used tool for calculating the effectiveness of the sampling plan and from which the likelihood of acceptance is extracted. It provides the likelihood that it is possible to accept the lot. SSP OC feature for life tests based on the ERD percentiles is as follows,

$$L(p) = \sum_{i=0}^{c} \binom{n}{i} p^i (1-p)^{n-i} \tag{3}$$

Where $P = F(t, \delta_0)$ represents the failure probability at time t given a determined 100q[th] percentile of the lifetime t_q^0 and p depends only on $\delta_0 = t/t_q^0$. In Table 3, Pradeepa Veerakumari (2016), the OC values are tabulated.

For the lot value p, the OC function of SkSP-R is given by,

The operating feature function for the SkSP-R plan is given as per Balamurali.et.al (2014),

$$P_a(p) \frac{fP+(1-f)P^i+fP^k(P^i-P^k)(1-Q^m)}{f(1-P^i)(1-P^k(1-Q^m)+P^i(1+fQP^k)} \tag{4}$$

Then, the Average Sample number is

$$ASN\ (p) = ASN\ (R)F \tag{5}$$

Where, R- represents the Average Sample number of the reference plan, P represents the probability of acceptance of the reference plan.

5.1 Illustration

Suppose that the life time of the electric goods follows ERD. Skip lot sampling plan of type SkSP-3 with SSP as reference plan based on 50[th] percentile is applied for testing. The parameters for the life testing is as follows: θ=2, t=40hrs, $t_{0.50}$=20hrs, c=0, α=0.05 and β=0.05 then η = 1.56712 from the equation and the ratio is found to be $t/t_{0.50}$ =2.00 by applying the minimum sample size according to the requirements is $n=5$ and the corresponding OC values L(p) for the Single Sampling plan for the life tests based on percentiles of ERD (n,c, $t/t_{0.50}$ = 5,2,1.22) with $P^* = 0.95$. $L(p)$ is the P value for SkSP-R with SSP for life tests based on the percentiles of ERD as reference plan. For $i=1,k=3$ and $f=1/3$; the probability of acceptance $L(p)$ values of SkSP-R with SSP for life tests based on percentiles of ERD are found from eqn. 4 as,

$t\,/\,t^0_{0.50}$	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
$L(p)$	0.0237	0.4685	0.8460	0.9552	0.9864	0.9957	0.9985	0.9995	0.9998

From the illustrations, it is indicated that the actual 50[th] percentile is almost equal to the required 50[th] percentile ($t/\ t_{0.50}$=1.00) the producer's risk is approximately 0.9808 (1-0.0192). Also the producer's risk is nearly equal to 0.05 or less and the actual producer risk is large or nearly equal to 2.15 times of the required percentile. For the purpose of convenience OC values of the table are constructed and tabulated with parameters $i=1$, $k=3$, $f=1/3$ and $c=2$ in Table 1.1.

Table 1.1 : *Gives the OC values for sampling plan (n,c = 2, t/$t_{0.50}$) for a given P^* under ERD when θ=2*

P^*	$t / t^0_{0.50}$	$t / t^0_{0.50}$								
		0.7	**0.9**	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**
0.75	0.7	0.4655	0.8770	0.9711	0.9928	0.9981	0.9994	0.9998	0.9999	1.0000
0.75	0.9	0.4864	0.8696	0.9664	0.9909	0.9974	0.9992	0.9997	0.9999	1.0000
0.75	1	0.4710	0.8570	0.9612	0.9891	0.9968	0.9990	0.9997	0.9999	1.0000
0.9	0.7	0.2283	0.7741	0.9437	0.9851	0.9958	0.9988	0.9996	0.9999	0.9999
0.9	0.9	0.1618	0.6956	0.9151	0.9751	0.9924	0.9976	0.9992	0.9997	0.9999
0.9	1	0.3376	0.8004	0.9443	0.9837	0.9950	0.9983	0.9993	0.9997	1.0000
0.95	0.7	0.1276	0.6919	0.9206	0.9782	0.9938	0.9981	0.9994	0.9998	0.9999
0.95	0.9	0.1155	0.6446	0.8987	0.9698	0.9907	0.9970	0.9990	0.9996	0.9999
0.95	1	0.0921	0.5938	0.8770	0.9615	0.9876	0.9958	0.9985	0.9995	0.9998
0.99	0.7	0.0237	0.4685	0.8460	0.9552	0.9864	0.9957	0.9985	0.9995	0.9998
0.99	0.9	0.0259	0.4366	0.8198	0.9437	0.9817	0.9938	0.9978	0.9992	0.9997
0.99	1	0.0110	0.3168	0.7495	0.9173	0.9715	0.9899	0.9963	0.9986	0.9994

Table 1.2 : *Minimum sample size for the 50th percentile of ERD to exceed the given value* $t_{0.50}$

P^*	c	$t / t_{0.50}$								
		0.7	**0.9**	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**
0.75	**0**	4	3	3	2	2	1	1	1	1
0.75	**1**	9	6	4	2	2	2	2	2	2
0.75	**2**	14	7	5	3	3	3	3	3	3
0.75	**3**	19	9	6	5	4	4	4	4	4
0.9	**0**	10	5	3	1	1	1	1	1	1
0.9	**1**	17	8	6	3	3	2	2	2	2
0.9	**2**	23	11	9	4	4	4	3	3	3
0.9	**3**	30	14	11	5	5	4	4	4	4
0.9	**10**	71	35	28	14	12	12	11	11	11
0.95	**0**	13	6	4	2	2	1	1	1	1
0.95	**1**	21	9	8	3	3	2	2	2	2
0.95	**2**	28	14	11	6	5	4	3	3	3
0.95	**3**	36	17	13	6	4	4	4	4	4
0.99	**0**	20	10	6	3	3	2	1	1	1
0.99	**1**	29	13	10	4	3	3	2	2	2
0.99	**2**	39	17	14	6	4	4	3	3	3
0.99	**3**	45	21	17	7	5	5	4	4	4

Table 1.3 : *Gives the ratio* $d_{0.50}$ *for accepting the lot with the procedure's risk of 0.05 when θ=2*

P^*	c	$t / t_{0.50}$								
		0.7	**0.9**	**1**	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**
0.75	**0**	0.4411	0.4412	0.4421	0.4935	0.6021	0.6021	0.6021	0.6021	0.6021
0.75	**1**	0.5022	0.5691	0.6538	0.8637	0.8637	0.8637	0.8637	0.8637	0.8637
0.75	**2**	0.5426	0.6816	0.7764	0.8178	1.0061	1.0061	1.0061	1.0061	1.0061
0.75	**3**	0.5640	0.7283	0.8644	0.9499	1.1016	1.1016	1.1016	1.1016	1.1016
0.9	**0**	0.3185	0.3838	0.4411	0.6039	0.6021	0.6021	0.6021	0.6021	0.6021
0.9	**1**	0.4242	0.5415	0.5692	0.7241	0.7242	0.8637	0.8637	0.8637	0.8637
0.9	**2**	0.4656	0.5842	0.6492	0.8570	0.8573	1.0061	1.0061	1.0061	1.0061
0.9	**3**	0.4898	0.6229	0.6761	0.9499	0.9500	1.1016	1.1016	1.1016	1.1016
0.95	**0**	0.2977	0.3654	0.4080	0.6039	0.4935	0.6022	0.6021	0.6021	0.6021
0.95	**1**	0.3998	0.5023	0.5204	0.7230	0.7230	0.8637	0.8655	0.8637	0.8637
0.95	**2**	0.4402	0.5537	0.6029	0.7224	0.7762	1.0061	1.0061	1.0061	1.0061
0.95	**3**	0.4689	0.5845	0.6564	0.9519	1.1017	1.1016	1.1016	1.1016	1.1016
0.99	**0**	0.2675	0.3280	0.3653	0.4421	0.4408	0.4935	0.6022	0.6022	0.6022
0.99	**1**	0.3600	0.4508	0.4866	0.6539	0.7264	0.7242	0.8637	0.8637	0.8637
0.99	**2**	0.4038	0.5095	0.5536	0.7766	1.0089	0.8573	1.0060	1.0060	1.0060
0.99	**3**	0.4356	0.5467	0.5949	0.8644	0.9494	0.9499	1.1017	1.1017	1.1017

VI. Operating procedure of SkSP-R with DSP as a reference plan based on percentiles of ERD

The modus operandi of SkSP-R with DSP as a reference plan based on percentiles of ERD are as follows:

Step 1: A random sample of size n_1 is drawn and put on a life test.

Step 2: The number of defectives $d_1$is counted and a comparison is made with the acceptance number c.

i. If $d_1>c_1$, then reject the lot.

ii. If $d_1\leq c_1$, then accept the lot.

Step 3: If $d_1<c_2$, is obtained before the specified time t_0, terminate the test, and reject the lot.

Step 4: If $c_1< d_1 \leq r_1$, take a second sample of size n_2 from the remaining lot and put them on test for time $t_0$and count the number of non-conformities (d_2).

Step 5:

If $d_1+d_2 \leq r_1$, accept the lot.

If $d_1+d_2 > r_1$, reject the lot.

VII. Operating characteristic function for SkSP-R using Double Sampling Plan

OC function is the most applied technique to measure the efficiency of the sampling plan and from where the probability of acceptance is derived. It provides the probability that the lot can be accepted. The OC function of DSP for life tests based on the percentiles of ERD is as follows,

$$L(p) = \sum_{d_1}^{c_1} \binom{n_1}{d_1} p^{d_1}(1-p)^{n_1-d_1}. \sum_{d_1+c_{1+1}}^{c_2} \binom{n_1}{d_1} p^{d_1}(1-p)^{n_1-d_1}. \sum_{d_2=0}^{c_2-d_2} \binom{n_2}{d_2} p^{d_2}(1-p)^{n_2-d_2} \tag{6}$$

Where $P = F(t,\delta_0)$ represents the failure probability at time t given a determined 100q[th] percentile of the lifetime t_q^0 and p depends only on $\delta_0 = t/t_q^0$. The ASN Value of DSP is calculated from the equation,

$$ASN = n_1 p_1 + (n_1+n_2)(1-p_1) = n_1 + n_2(1-p_1) \tag{7}$$

The OC function of SkSP-R for the lot quality p is given by,

$$P_a(p) \frac{fP+(1-f)P^i+fP^k(P^i-P^k)(1-Q^m)}{f(1-P^i)(1-P^k(1-Q^m)+P^i(1+fQP^k)} \tag{8}$$

Then, the Average Sample number is

$$ASN\ (p) = ASN\ (R)F \tag{9}$$

Where ASN (R) represents the Average Sample number of the reference plan; P represents the probability of acceptance of the reference plan.

7.1 Illustration

Presume that the lifetime of the electric goods follows ERD. Skip lot sampling plan of type SkSP-R with DSP as a reference plan based on the 10[th] percentile is applied for testing. The parameters for the life testing is as follows: θ=2, t=40hrs, $t_{0.50}$=20hrs, c_1=0,c_2=1, α=0.01 and β=0.05 then η = 0.871929 from the equation and the ratio is found to be $t/t_{0.50}$ =2.00 by applying the minimum sample size according to the requirements is n_1=9,n_2=11 and the corresponding OC values $L(p)$ for the Double Sampling plan for the life tests based on percentiles of ERD $n_1, n_2, c_1, c_2, t/_{t_{0.1}}$ = (9,11,0,1,0.9379) with P^* = 0.99. $L(p)$ is the P value for SkSP-R with DSP for life tests as a reference plan defined on the percentiles of ERD. For i=1,k=3, and f=1/5; the probability that SkSP-R with DSP will consider $L(p)$ values for life tests based on percentiles of ERD is found from Equation 8 For,

$t/t_{0.1}^0$	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
$L(p)$	0.0398	0.3764	0.6235	0.9123	0.9668	0.9978	0.9993	0.9997	0.9999

From the illustrations, it is indicated that the actual 10th percentile is almost equal to the required 10th percentile $t/t_{0.1}^0$ the producer's risk is approximately 0.9602 (1-0.0398). Moreover, the producer's risk is closely equal to 0.05 or less and the actual producer risk is large or nearly equal to 2 times the required percentile. For the purpose of convenience OC values of the table are constructed and tabulated with parameters *i=1,k=3, f=1/5* and $c_1=0, c_2=1$ in Table 2.

Table 2: *Gives the OC values for Sampling Plan* $(n_1, n_2, c_1, c_2, t/t_{0.1} = (9,11,0,1,0.9379)$ *for a given P* under ERD when θ=2*

P^*	$t\,/\,t^0_{0.50}$	$t_{0.1}/t_{0.1}^0$								
		1	**1.25**	**1.5**	**1.75**	**2**	**2.25**	**2.5**	**2.75**	**3**
0.75	0.7	0.5839	0.8364	0.7766	0.9265	0.9736	0.9897	0.9957	0.9980	0.9990
0.75	0.9	0.4318	0.8912	0.7603	0.9209	0.9715	0.9888	0.9953	0.9978	0.9996
0.75	1	0.4797	0.8460	0.7440	0.9153	0.9694	0.9880	0.9949	0.9976	0.9995
0.9	0.7	0.2276	0.6008	0.7277	0.9097	0.9673	0.9872	0.9946	0.9975	0.9992
0.9	0.9	0.2755	0.6557	0.7114	0.9042	0.9652	0.9863	0.9942	0.9973	0.9991
0.9	1	0.2234	0.6105	0.6951	0.8986	0.9631	0.9855	0.9938	0.9971	0.9990
0.95	0.7	0.3713	0.0653	0.6788	0.8930	0.9610	0.9846	0.9935	0.9969	0.9999
0.95	0.9	0.1191	0.6201	0.6625	0.8874	0.9589	0.9838	0.9931	0.9968	0.9998
0.95	1	0.1670	0.5250	0.6462	0.8819	0.9569	0.9829	0.9928	0.9966	0.9997
0.99	0.7	0.0149	0.4702	0.6299	0.8763	0.9548	0.9821	0.9924	0.9964	0.9997
0.99	0.9	0.0628	0.3154	0.6136	0.8707	0.9527	0.9813	0.9920	0.9963	0.9996
0.99	1	0.0107	0.3605	0.5973	0.8651	0.9506	0.9804	0.9917	0.9961	0.9996

VIII. Conclusion

In this study, life testing plans based on percentiles of ERD for Skip-lot Sampling plan of type-R with Single Sampling Plan and Double Sampling Plan as reference plan are developed. Skip-lot Sampling plan of type-R with SSP and DSP as reference plan requires minimum sample size and also has better operating characteristics values. Thus, results in reduction of inspection cost and better efficient. The proposed plan can be further extended to other sampling plans for instance SkSP and other probability distribution.

Reference

[1] Aslam et.al., (2010): Optimal designing of a Skip lot Sampling Plan by two point method, *Pakistan Journal of Statistics*, 26 (4), 585-592.
[2] Aslam M, and M.Q. Shahbaz (2007). Economic Reliability plans using the Generalized Exponential Distribution, *J. Stat.*, 14, 53-60.
[3] Baklizi A (2003). Acceptance sampling based on truncated life tests in the pareto distribution of the second kind. *Advances and Applications in Statistics,* 3 (1), 33-48.

[4] Balakrishnan N. Leiva V. And Lopez J. (2007). Acceptance sampling plans from truncated tests based on the generalized Birnbaun-Sauders distribution. *Communications in statistics – Simulation and Computation*, 36, 643-656.

[5] Balamurali S and Chi-Hyuck Jun (2011): A new system of skip-lot sampling plans having a provision for reducing a normal inspection, *Applied stochastic Models Business Industries*, 27, 348-363.

[6] Dodge H.F. (1955). Skip-Lot sampling plan, *Industrial Quality Control*, 11(5), 3-5.

[7] Epstein B (1954). Truncated life tests in the exponential case, *Annals of Mathematical Statistics*, 25: 555-564.

[8] Goode. H P and Kao. J. H. K, (1961). Sampling plans based on the Weibull distribution, *Proceedings in the 7th National Symposium on Reliability and Quality Control*, 24-40, Philadelphia, USA.

[9] Gupta S. S., and Groll. P. A. (1961). Gamma distribution in acceptance sampling based on life tests, *Journal of the American Statistical Association*, 56, 942-970.

[10] Kantam R.R.L. and Rosaiah K. and Rao. G.S. (2001). Acceptance sampling based on life tests: Log-Logistic model. *Journal of Applied Statistics,* 28: 121-128.

[11] Kantam R.R.L. G. Srinivasa Rao and G. Sriram (2006). An economic reliability test plan: Log-Logistic distribution, *J. Appl. Statist*. 33(3): 291-296.

[12] PERRY, R.L. (1973): Skip-lot Sampling Plans, *Journal of Quality Technology*, 5(3), 123-130.

[13] Pradeepa Veerakumari K and Ponneeswari P (2016). Designing of Acceptance Sampling Plan for life tests based on Percentiles of Exponentiated Rayleigh Distribution, *International Journal of Current Engineering and Technology,* 6 (4):1148-1153.

[14] Pradeepa Veerakumari K , Umamaheswari P, Aruna H M , Suganya S (2018). Designing of Skip-Lot Sampling Plan of Type (Sksp-3) For Life Tests Based On Percentiles of Exponentiated Rayleigh Distribution. *International Journal of Research in Advent Technology*, 6(8)

[15] Rao G S, Ghitany M. E and Kantam R. R. L.(2009). Acceptance Sampling Plans for Marshall-Olkin extended Lomax distribution. *International Journal of Applied Mathematics*, 2(1), 139 – 148.

[16] Sobel, M and Tischendrof, J. A. (1959). Acceptance sampling with sew life test objective. *Proceedings of Fifth National Symposium on Reliability and Quality Control,* Philadelphia, 108-118.

[17] Suganya S and Pradeepa Veerakumari (2020). Skip-lot sampling plan of type SkSP-T for life test based on Percentiles of Exponentiated Rayleigh Distribution, *Life Cycle Reliability and Safety Engineering*, 9, 247-251.

[18] Tsai T.R. and Wu S.J. (2006). Acceptance Sampling Based on truncated life tests for generalized Rayleigh Distribution. *Journal of Applied Statistics.* 33,595 – 600.

APPLICATION OF FUZZY LINEAR PROGRAMMING APPROACH FOR SOLVING MIX-PRODUCT SELECTION PROBLEM

Mahesh M. Janolkar

•

Department of First Year Engineering
Prof Ram Meghe College of Engineering and Management Badnera-Amravati (MS India)
maheshjanolkar@gmail.com

Kirankumar L. Bondar

•

P. G. Department of Mathematics, Govt Vidarbh Institute of Science and Humanities, Amravati.
klbondar_75@rediffmail.com

Pandit U. Chopade

•

Research Supervisor, Department of Mathematics,
D.S.M's Arts Commerce and Science College ,Jintur.
chopadepu@rediffmail.com

Abstract

In this paper, the modified SMF system is used in the real MPS problem. This problem occurs in the production planning process where the decision maker plays an important role in making decisions in an uncertain environment. As researchers, we are trying to find the best solution for the final decision maker. SMF analyzed FLP production equipment using data actually collected from chocolate production companies. The problem of MPS incompatibility has been described. The aim of this article is to find the best UOP with high satisfaction and nonsense as the main thing. Since there are so many decisions to make, the best UOP table is defined in terms of insensitivity and satisfaction to find a solution with a high UOP level and a high level of satisfaction. OF indicates that a high UOP will not lead to a high level of satisfaction. The results of this work suggest that the best decision is based on the negative impact on the FS of the MPS. In addition, a high level of UOP is achieved when the blur is low.

Keywords: Linear Programming, Uncertainty, Fuzzy constraint, Mix-Product Selection.

Abbreviations:

MPS	:	mix product selection
FLP	:	fuzzy linear programming
SMF	:	s-curve membership function
FO	:	fuzzy outcome
FS	:	fuzzy system
UOP	:	units of product

1. Introduction

Non-SMF conversion function is used for problems related to FLP. The function S can be applied and tested for its effectiveness by applied pressure. In this example, the S function is applied to make a decision after binary, such as the number of technologies and equipment, of which MPS is complex. Solutions thus obtained can provide the decision maker and the coordinator for the final implementation. The wording described in this article is just one of the three FPS words that actually have an application. The above FPS term is considered to be the real-life situation when it comes to making chocolate. Data for this problem are provided in the database of Choco man Inc, USA. Choco man manufactures chocolate bars, candies and wafer using a variety of ingredients and formulas. The goal is to use the modified S function as a system to get the best UOP through the FLP system [1-3]
Compared with this FLP system. The recommended method is based on its relationship with the decision maker, developer and researcher to find satisfactory solutions for the FLP problem. In the decision-making process using the FLP model, modifications and source software can be complex, rather than providing exact numbers as in the net LP model. For example, machine hours, work, requirements, etc. and manufacturing, which is not always good, due to insufficient information and uncertainty among potential importers in the environment. Therefore, they should be considered as non-essential components and the FLP problem can be solved by using the FLP method. The problem of non-compliant MPS has been described. The aim of this article is to find the best UOP with high satisfaction and nonsense as the main thing. This problem is considered because all the parameters such as technology and hardware changes are uncertain. This is considered to be a major overall problem that includes 29 barriers and 8 barriers. Since there are so many decisions to make, the best UOP table is described for uncertainty and satisfaction to find a solution. with the highest UOP level and the highest satisfaction. It should be borne in mind that a high UOP does not mean it will lead to a high level of satisfaction. The best UOP was calculated at the satisfaction level using the FLP method. OF indicates that a high POU will not lead to a high level of satisfaction. The results of this work suggest that the best decision is based on the negative impact on the FS of the MPS. In addition, high levels of UOP are obtained when blur is low in the system [4-25].

2. Methodology of MF

A general model of classical LP is formulated as,

$$Max(w) = dy \quad \text{standard formulation;}$$

Subject to,

$$By \le c; y \ge 0 \tag{1}$$

Where d and y are the m-part vector, d is the m-part vector, and B is $n \times m$ matrix. Since we live in an uncertain environment, the number of objective functions (d), the number of matrix technologies (B) and the variability of assets (d) are complex. Therefore, an infinite number can be displayed, so that the problem can be solved by the FLP system. FLP problems are designed as follows:

$$Max(w) = d^* y \quad \text{The Fuzzy formulation;}$$

Subject to,

$$B^* y \le c^*; y \ge 0 \tag{2}$$

where x is the vector of the decision change; $B^*, c^* \& d^*$ are zero numbers; The function of addition and multiplication is explained by fact that in-depth numbers are derived from the extension principles of Li [26]; Njikọ Inequalities are provided by some relationship and work objectives, w must take into account the given LP problem. The approach of Mohammed [27] is being considered to solve the problem of FLP 2 depletion., which means that the solution will probably be to some satisfaction. First, design the team function for the zero parameter of $B^*, c^* \& d^*$. Here, non-existent team functions, such as logic, are used. vb_{kl} represents the work of members; vc_k and vd_l is the numerical function of matrix B for $k = 1, 2, ...n \& n = 1, 2, ...m$, c_k is the numerical variable for $k = 1, 2, ...n$ and d_l are the integers of purpose point w for $l = 1, 2, ...m$.

Then, with the appropriate change in the concept of agreement between the non- b_{kl}^* numbers; c_k^* and d_{kl}^* & l, words for b_{kl}^*, c_k^* and d_l^* will be obtained. When an agreement between b_{kl}^*; The solution c_k^* and d_l^* will be [28];

$$v = vd_l = vb_{kl} = vc_k \tag{3}$$

for all $k = 1, 2, ...n \& l = 1, 2, ...m$

Therefore, we can obtain;

$$D = pd(v), B = pb(v) \& c = pc(v) \tag{4}$$

Where $v \in [0,1]$ in $pd, pb \& pc$ are distinct functions [29] of $vd, vb \& vc$ respectively. Equation (2) would be,

$$Max(w) = [pd(v)]y \quad \text{Fuzzy formulation;}$$

Subject to,

$$Max(w) = [pd(v)]y; y \geq 0 \tag{5}$$

First, create a group function for the complex part of $B^* \& c^*$. Here, non-uniform functions are used as S-curve function [30]. vb_{kl} represents the work of members and vc_k, where b_{kl} is the coefficient of matrix B for $k = 1, 2, ...29$, $l = 1, 2, 3, ...8$ and c_k is the material variable for $k = 1, 2, ...29$. Group function is also obtained for b_k and beard time, $c_k b$ to $c_k c$ for c_k^*. Similarly, we can create team work for a number of non-core technologies and their production [31]. Due to the high cost of production and the need to meet certain production and demand conditions, the problem of inefficiency arises in the manufacturing process. This problem also arises in the production of chocolate when deciding on the combination of ingredients to create different types of products. This is called here the choice of product mix [32]).

3. The Fuzzy MPS

There are products that can be made by mixing different ingredients and using k type processing. It is expected that the infrastructure will be massive. There are also some restrictions by the retail department, such as the requirement for the product mix, the requirement of the main product line, as well as the minimum and maximum query for each product. Not everything that is needed in these circumstances is obvious. It is important to achieve maximum UOP and satisfaction using the FLP method. Since the number of technologies and equipment changes is running high, the results of the UOP would be foolish. FLP problem, customized in size. 2 can be written:

$$Max(w) = \sum_{l=1}^{8} y_l$$

Subject to,

$$\sum_{l=1}^{8} b_{kl}^{*} y_l \le c_k^{*}; \tag{6}$$

where $y_l \ge 0, i = 1,2,3,...8,\ b_{kl}^{*}\ \&\ c_k^{*}$ are fuzzy parameters.

3.1 Fuzzy Resource Variable

For an interval, $c_k b \prec c_k \prec c_k c$,

$$b_{c_k} = \frac{c}{1 + De^{\left(b\left(c_k - c_k^b\right)/c_k^c - c_k^b\right)}}$$

$$e^{\frac{\beta\left(c_k - c_k^b\right)}{\left(c_k^c - c_k^b\right)}} = \frac{1}{D}\left(\frac{c}{\theta_{c_k}} - 1\right) \tag{7}$$

$$\frac{\beta\left(c_k - c_k^b\right)}{c_k^c - c_k^b} = \ln\left[\frac{1}{D}\left(\frac{C}{\theta_{c_k}}\right) - 1\right]$$

$$c_k = c_k^a + \frac{1}{D}\left(\frac{c_k^c - c_k^b}{\beta}\right)\ln\left[\frac{1}{D}\left(\frac{C}{\theta_{c_k}}\right) - 1\right]$$

Since c_k is a non-trivial material change in size. 7, it is found vc_k. Therefore

$$cc_k^{*} = c_k^b + \left(\frac{c_k^c - c_k^b}{\beta}\right)\ln\left[\frac{1}{D}\left(\frac{C}{\theta_{c_k}}\right) - 1\right] \tag{8}$$

3.2 Fuzzy Constraints

The products, materials and equipment requirements are shown in Tables 1 as well as 2, respectively. Tables 3 as well as 4 provide the mix size and use the required material to make each product.

Table 1: *Products Demand*

Items	Fuzzy Interval $(\times 10000)$ units
Milk Chocolate, (200 gram)	[450-575) Gram
Milk Chocolate, (50 gram)	[750-950) Gram
Crunchy Chocolate, (200 gram)	[350-450) Gram
Crunchy Chocolate, (50 gram)	[550-700) Gram
Chocolate with Nuts (200 gram)	[250-325) Gram
Chocolate with Nuts (50 gram)	[450-575) Gram
Chocolate Candy	[150-200) Gram
Wafer	[350-450) Gram

Table 2: *Material and Ease of Access*

Raw Material	Fuzzy Interval (x1000 units)
Coco (Kilo Gram)	[75-125) Kilo Gram
Milk (Kilo Gram)	[90-150) Kilo Gram
Nuts (Kilo Gram)	[45-75) Kilo Gram
Sugar (Kilo Gram)	[150-450) Kilo Gram
Flour (Kilo Gram)	[15-25) Kilo Gram
Aluminum Foil (Kilo Gram)	[375-625) Kilo Gram
Paper (Per Feet Square)	[375-625) Per Feet Square
Plastic (Per Feet Square)	[375-625) Per Feet Square
Cooking (Ton per H)	[750-1250) Ton Per H
Mixing (Ton per H)	[150-250) Ton Per H
Forming (Ton per H)	[1125-1875) Ton Per H
Grinding (Ton per H)	[150-250) Ton Per H
Wafer Making (Ton per H)	[75-125) Ton Per H
Cutting (H)	[300-350) H
Packaging 1 (H)	[300-500) H
Packaging 2 (H)	[900-1500) H
Labor (H)	[750-1250) H

There are two unclear barriers such as access to the equipment and restrictions on the capacity of the equipment. These barriers are inevitable for any object and property depending on the consumption of the property, to trade and acquire property. These selections are based on the FLP resolution of Choco man Inc. Decision changes for the FPSP are defined as:

$y_1 =$ 250 grams of chocolate milk to be produced (in 1000)

$y_2 =$ 250 grams of chocolate milk to be produced (per 1000)

$y_3 =$ Chocolate Crispy of 250 grams to be produced (in 1000)

$y_4 =$ 100 grams of Chocolate Crispy to be produced (in 1000)

$y_5 =$ Chocolate with 250 grams of fruit to produce (in 1000)

$y_6 =$ Chocolate contains 100 grams per gram to produce (in 1000)

$y_7 =$ Chocolate candies will be produced (in 1000 packages)

$y_8 =$ Chocolate wafer production (in 1000 packages)

$$y_1 \leq 0.6 y_2 \tag{9}$$

$$y_3 \leq 0.6 y_4 \tag{10}$$

$$y_5 \leq 0.6 y_6 \tag{11}$$

The required product line is key. Total sales of confectionery products and wafers should not exceed 15% (uncertain value) of total confectionery product.

Table 3: *Mixing Proportions*

Material Required per 1000 units	Product types (Fuzzy interval)							
	AMC 150	AMC 50	ACC 150	ACC 50	ACN 150	ACN 50	Candy	Wafer
Coco (Kilo Gram)	[60-90)	[20-45)	[105-130)	[25-60)	[150-250)	[0-0)	[1200-1400)	[150-300)
Milk (Kilo Gram)	[0-0)	[0-0)	[60-90)	[0-0)	[78-101)	[35-80)	[230-500)	[0-0)
Nuts (Kilo Gram)	[325-456)	[78-105)	[230-280)	[34-87)	[0-0)	[0-0)	[110-230)	[73-130)
Sugar (Kilo Gram)	[172-201)	[0-0)	[78-99)	[0-0)	[321-436)	[103-120)	[0-0)	[54-90)
Flour (Kilo Gram)	[0-0)	[0-0)	[120-150)	[0-0)	[450-487)	[245-298)	[1001-1200)	[540-670)
Aluminum Foil (Kilo Gram)	[110-165)	[78-95)	[0-0)	[0-0)	[330-420)	[110-154)	[0-0)	[0-0)
Paper (Per Feet Square)	[156-185)	[0-0)	[190-245)	[0-0)	[100-150)	[56-89)	[0-0)	[0-0)
Plastic (Per Feet Square)	[0-0)	[0-0)	[170-240)	[40-82)	[510-725)	[120-179)	[0-0)	[0-0)

Table 4: *Facility Usage*

Facility Usage Required Per 1000 Units	Product types (fuzzy interval)							
	AMC 150	AMC 50	ACC 150	ACC 50	ACN 150	ACN 50	Candy	Wafer
Cooking (Ton per H)	[0.60-0.90)	[0.20-0.45)	[0.105-0.130)	[0.25-0.60)	[0.150-0.250)	[0-0)	[0.1200-0.1400)	[0.150-0.300)
Mixing (Ton per H)	[0-0)	[0-0)	[0.60-0.90)	[0-0)	[0.78-0.101)	[0.35-0.80)	[0.230-0.500)	[0-0)
Forming (Ton per H)	[0.325-0.456)	[0.78-0.105)	[0.230-0.280)	[0.34-0.87)	[0-0)	[0-0)	[0.110-0.230)	[0.73-0.130)
Grinding (Ton per H)	[0.172-0.201)	[0-0)	[0.78-0.99)	[0-0)	[0.321-0.436)	[0.103-0.120)	[0-0)	[0.54-0.90)
Wafer Making (Ton per H)	[0-0)	[0-0)	[0.120-0.150)	[0-0)	[0.450-0.487)	[0.245-0.298)	[0.1001-0.1200)	[0.540-0.670)
Cutting (H)	[0.110-0.165)	[0.78-0.95)	[0-0)	[0-0)	[0.330-0.420)	[0.110-0.154)	[0-0)	[0-0)
Packaging 1 (H)	[0.156-0.185)	[0-0)	[0.190-0.245)	[0-0)	[0.100-0.150)	[0.56-0.89)	[0-0)	[0-0)
Packaging 2 (H)	[0-0)	[0-0)	[0.170-0.240)	[0.40-0.82)	[0.510-725)	[0.120-0.179)	[0-0)	[0-0)
Labor (H)	[0.325-0.456)	[0.78-0.105)	[0.230-0.280)	[0.34-0.87)	[0-0)	[0-0)	[0.110-0.230)	[0.73-0.130)

Table 5: *Optimal UOP with a satisfaction degree*

Number	Satisfaction degree (**θ**)	Optimal UOP (w^*)
1	7.562	2438.54
2	14.076	2500.51
3	15.2145	2615.83
4	16.1148	2651.25
5	18.057	2701.67
6	24.8497	2845.48
7	28.9782	2848.79
8	30.3968	2889.39
9	31.7572	2923.44
10	42.6513	2955.9
11	50.0115	2965.11
12	52.1911	3001.89
13	52.8741	3057.48
14	59.6383	3152.55
15	63.3374	3160.55
16	63.538	3180.37
17	64.8241	3204.67
18	70.4424	3250.39
19	85.5813	3277.92
20	95.4286	3344.58

4. Results

The FPS problem is solved by using MATLAB and its LP application. It provides complexity and a degree of satisfaction. The LP application has two extras in addition to the non-existent. There is an output w^*, the best UOP.

Table 6: *The vagueness* β *as well as objective value* w^* *with* $\theta = 50\%$

Vagueness **β**	UOP w^*
1	2465.54
3	2533.72
5	2568.99
7	2631.09
9	2730.54
11	2740.35
13	2778.95
15	2784.04
17	2833.00
19	3011.15
21	3037.45
23	3080.78
25	3223.61
27	3239.79
29	3282.03
31	3352.45
33	3368.74
35	3438.1
37	3446.69

Table 7: *Optimal UOP* w^*

w^*	Vagueness β			
θ	1	3	5	7
7.562	2421.27	2478.47	2594.46	2488.84
14.076	2514.88	2502.54	2673.13	2509.44
15.2145	2638.86	2623.91	2765.32	2574.27
16.1148	2639.8	2632.57	2780.56	2604.7
18.057	2668.82	2675.98	2797.33	2618.06
24.8497	2686.3	2680.99	2919.95	2621.45
28.9782	2753.94	2747.67	2930.67	2652.31
30.3968	2827.54	2773.03	3028.05	2723.29
31.7572	2870.88	2807.2	3189.58	2753.75
42.6513	2957.06	2847.5	3230.2	2810.63
50.0115	2960.57	3010.7	3234.95	2838.32
52.1911	2981.24	3017.36	3248.8	2843.2
52.8741	3078.7	3080.9	3297.06	3039.16
59.6383	3079.57	3086.95	3298.37	3157.71
63.3374	3132.07	3162.39	3334.88	3206.49
63.538	3273.09	3202.78	3415.55	3315.88
64.8241	3443.79	3348.41	3426.19	3411.56
70.4424	3479.39	3434.25	3470.15	3476.37

Different standards of Chocolate production are transferred to the toolbox. The answer can be listed in the following tables. From Table 5, it can be seen that a high level of satisfaction provides a high UOP. But the best solution to the above problem is at a satisfaction rate of 50%, or 2833 minutes. From the tables below, we conclude that within the objective, w^* is an ever-increasing function. Increased [33].

Table 8: *Optimal UOP* w^*

w^*	Vagueness β			
θ	9	11	13	15
7.562	2517.93	2511.75	2700.82	2626.7
14.076	2555.17	2562	2817.03	2713.6
15.2145	2610.27	2712.45	2818.6	2730.28
16.1148	2694.71	2735.65	2917.06	2735.94
18.057	2704.95	2778.61	3015.94	2814.01
24.8497	2768.05	2785.92	3017.65	2843.42
28.9782	2803.52	2982.47	3019.4	2857.43
30.3968	2912.9	3162.64	3200.54	2919.49
31.7572	2959.22	3205.75	3210.48	2936.06
42.6513	3006.57	3238.42	3211.28	3082.57
50.0115	3106.2	3252.29	3236.27	3155.49
52.1911	3110.49	3312.54	3276.6	3166.6
52.8741	3155.25	3326.07	3285.56	3215.15
59.6383	3206.75	3341.22	3292.6	3306.44
63.3374	3367.82	3383.69	3312.35	3339.97
63.538	3432.71	3393.02	3319.99	3353.86
64.8241	3461.5	3394.43	3341.83	3462.87
70.4424	3478.85	3435.72	3421.66	3493.17

Table 9: *Optimal UOP* w^*

w^*	Vagueness β			
θ	17	19	21	23
7.562	2560.71	2591.74	2598.75	2569.53
14.076	2577.5	2681.47	2671.48	2712.04
15.2145	2827.45	2695.28	2725.3	2774.99
16.1148	2857.61	2745.12	2898.84	2857.97
18.057	2877.99	2760.14	2919.28	2910.07
24.8497	3081.74	2770.16	2962.64	2962.97
28.9782	3093.67	2858.84	2989.96	2977.2
30.3968	3157.45	3063.62	3018.63	2983.99
31.7572	3202.92	3087.9	3020.53	2988.83
42.6513	3279.76	3093.95	3025.39	3012.8
50.0115	3289.08	3100.34	3089.09	3119.28
52.1911	3329.94	3206.97	3105.94	3133.89
52.8741	3339.61	3249.02	3118.94	3212.27
59.6383	3343.42	3287.02	3159.21	3267.98
63.3374	3362.92	3361.71	3185.11	3331.74
63.538	3373.1	3417.77	3275.53	3457.72
64.8241	3440.06	3434.14	3397.49	3486.65
70.4424	3492.01	3471.26	3495.27	3498.94

Table 10: *Optimal UOP* w^*

w^*	Vagueness β			
θ	25	27	29	31
7.562	2557.26	2509.77	2624.58	2522.45
14.076	2639.95	2531.72	2637.73	2547.82
15.2145	2727.12	2561.53	2645.54	2584.66
16.1148	2785.23	2610.31	2745.36	2750.06
18.057	2845.05	2680.12	2766.93	2756.62
24.8497	2879.51	2758.1	2778.77	2762.94
28.9782	2937.4	2800.6	2817.91	2832.69
30.3968	2967.17	2840.55	2893.03	2886.01
31.7572	3057.98	2846.94	2961.62	2938.18
42.6513	3110.12	2866.61	3012.12	3001.32
50.0115	3128.99	2880.25	3060.57	3044.8
52.1911	3139.91	2957.15	3075.73	3135.83
52.8741	3240.09	3012.5	3126.45	3297.11
59.6383	3259.24	3066.82	3170.93	3305.56
63.3374	3263.83	3118.69	3292.42	3313.34
63.538	3378.55	3132.87	3296.45	3384.03
64.8241	3422.86	3324.07	3375.38	3404.9
70.4424	3483.18	3350.47	3470.84	3428.67

Table 11: *Optimal UOP* w^*

w^*	Vagueness β			
θ	33	35	37	39
7.562	2522.48	2523.96	2533.43	2519.95
14.076	2532.12	2608.62	2618.64	2611.46
15.2145	2571.52	2618.64	2717.62	2615.81
16.1148	2712.13	2739.13	2749.95	2652.37
18.057	2916.79	2771.39	2778.74	2857.52
24.8497	2943.77	2797.06	2979.54	2891.37
28.9782	3088.17	2828.98	3023.91	2963.05
30.3968	3126.97	2886.21	3082.34	3010.27
31.7572	3130.92	2887.8	3171.68	3020.85
42.6513	3144.28	2901.63	3220.44	3041.08
50.0115	3183.95	2934.68	3236.11	3068.4
52.1911	3202.9	3052.3	3264.69	3102
52.8741	3213.79	3204.34	3330.91	3109.29
59.6383	3342.85	3264.08	3393.05	3214.24
63.3374	3361.04	3270.6	3426.9	3242.07
63.538	3403.39	3377.37	3432.62	3352.56
64.8241	3406.28	3467.32	3455.09	3392.32
70.4424	3492.01	3471.26	3495.27	3498.94

4.1 UOP w* for different vagueness values

Reasonable solutions and some uncertainties in the zero parameter of the technical rate and the hardware change are $= 50\%$. Thus, the results for the 50% satisfaction level for $1 \le \beta \le 39$ and the principles corresponding to w^* are shown in Table 6. OFs of UOP reduce β imprecision and increase of the nonlinear parameter of the number of technologies. and asset exchange. This is clearly shown in Table 6. Table 6 is very important for the decision maker when choosing UOP so that the result is a perfect level.

4.2 Output for $\theta, \beta \,\&\, w^*$

The results in the table below show that when the inaccuracy of the increase results in a small UOP.

Table 12: w^* *w.r.to* $\beta \,\&\, \theta$

Satisfaction degree ($\boldsymbol{\theta}$)	Vagueness ($\boldsymbol{\beta}$)	Optimal UOP (w*)
7.562	1	2500.51
14.076	3	2615.83
15.2145	5	2651.25
16.1148	7	2701.67
18.057	9	2845.48
24.8497	11	2848.79
28.9782	13	2889.39
30.3968	15	2923.44
31.7572	17	2955.9
42.6513	19	2965.11
50.0115	21	3001.89
52.1911	23	3057.48
52.8741	25	3152.55

59.6383	27	3180.37
63.3374	29	3204.67
63.538	31	3250.39
64.8241	33	3277.92
70.4424	35	3338.54
83.3374	37	3344.58

It is also seen that SMF has a variety of standards that provide possible solutions with some satisfaction. Also, the link between w^*, θ is provided in Tables 7, 8, 9, 10 and 11. This is clearly shown in Table 6. Table 6 is very important for the decision maker when choosing UOP so that the result is a perfect level. From Tables 7, 8, 9, 10 and 11, we find that for each type of satisfaction θ, the optimal UP w^* decreases as the endpoint increases between 1 and 37. Similarly, with any positive value, the optimal UOP increases. as the degree of satisfaction increases. Table 12 is the result of the diagonal pattern of w^* in Table 6. The results of this result show that: when the inaccuracies are low $\beta = 1, 3 \,\&\, 5$, UOP w^* is best. reached the lowest satisfaction level, $\theta = 7.5\%$, $\theta = 14.1\%$ and $\theta = 15.2\%$. When the odds are high at $\beta = 33, 35 \,\&\, 37$, UOP w^* is best reached with high satisfaction level, i.e., $\theta = 64.8\%, \theta = 70.4\% \,\&\, \theta = 83.3\%$.

5. Selection of Parameter β and Decision Making

In order for the decision maker to get the best results for the UOP w^*, the researcher creates a production table. From the table above, the decision maker can select the negative value according to his preference. Hair volume is divided into w^* in three parts, namely short, medium and high. It can be slightly modified if the input data for the number of technologies and hardware changes. It can be called a bunch of empty vanities. The decision can be made by the decision maker by choosing the best UOP for w^* and providing solutions for its implementation.

5.1 Discussion

The results show that the POU minimum is 2,755.4 with a maximum of 3,034.9. It can be seen that when the understanding is between 0 and 1, the maximum value of w * 3 034.9 is obtained by the minimum value. Similarly, when over 39, the minimum gain of w * 2,755.4 and the maximum gain are obtained. Since the solution for MPS nonsense is the most satisfying solution with a high satisfaction degree, it is important to choose a blur between the minimum value and the maximum value of w *. The well-distributed value of w * belongs to a group of musicians.

6. Conclusion

The purpose of this research project was to find the most effective UOP for MPS problems that have not been identified. SMF was recently developed as a framework for the task of solving the above problems effectively. The decision-making process and its implementation will be easier if the decision maker and consultant can work with the analyst to get the best and most satisfactory results. There are two more cases to consider in future work of the running technology that is not negative and that the dynamic assets are running and not complicated. FS mathematical relationships can be developed for MPS problems to find satisfying solutions. The decision maker, researcher and practitioner can apply their knowledge and experience to get the best results.

References

[1] Azadeh A, Raoofi Z, Zarrin M. A multi-objective fuzzy linear programming model for optimization of natural gas supply chain through a greenhouse gas reduction approach. Journal of Natural Gas Science and Engineering. 2015;26:702-10.

[2] Chandrawat RK, Kumar R, Garg B, Dhiman G, Kumar S, editors. An analysis of modeling and optimization production cost through fuzzy linear programming problem with symmetric and right angle triangular fuzzy number. Proceedings of Sixth International Conference on Soft Computing for Problem Solving; 2017: Springer.

[3] Wan S-P, Wang F, Lin L-L, Dong J-Y. An intuitionistic fuzzy linear programming method for logistics outsourcing provider selection. Knowledge-Based Systems. 2015;82:80-94.

[4] Kumar D, Rahman Z, Chan FT. A fuzzy AHP and fuzzy multi-objective linear programming model for order allocation in a sustainable supply chain: A case study. International Journal of Computer Integrated Manufacturing. 2017;30(6):535-51.

[5] Rani D, Gulati T, Garg H. Multi-objective non-linear programming problem in intuitionistic fuzzy environment: Optimistic and pessimistic view point. Expert Systems with Applications. 2016;64:228-38.

[6] Govindan K, Sivakumar R. Green supplier selection and order allocation in a low-carbon paper industry: integrated multi-criteria heterogeneous decision-making and multi-objective linear programming approaches. Annals of operations research. 2016;238(1-2):243-76.

[7] Abdel-Basset M, Gunasekaran M, Mohamed M, Smarandache F. A novel method for solving the fully neutrosophic linear programming problems. Neural computing and applications. 2019;31(5):1595-605.

[8] Liao H, Jiang L, Xu Z, Xu J, Herrera F. A linear programming method for multiple criteria decision making with probabilistic linguistic information. Information Sciences. 2017;415:341-55.

[9] Yang X-P, Zhou X-G, Cao B-Y. Latticized linear programming subject to max-product fuzzy relation inequalities with application in wireless communication. Information Sciences. 2016;358:44-55.

[10] Edalatpanah S. A direct model for triangular neutrosophic linear programming. International journal of neutrosophic science. 2020;1(1):19-28.

[11] Rodger JA, George JA. Triple bottom line accounting for optimizing natural gas sustainability: A statistical linear programming fuzzy ILOWA optimized sustainment model approach to reducing supply chain global cybersecurity vulnerability through information and communications technology. Journal of cleaner production. 2017;142:1931-49.

[12] Talaei M, Moghaddam BF, Pishvaee MS, Bozorgi-Amiri A, Gholamnejad S. A robust fuzzy optimization model for carbon-efficient closed-loop supply chain network design problem: a numerical illustration in electronics industry. Journal of cleaner production. 2016;113:662-73.

[13] Alavidoost M, Babazadeh H, Sayyari S. An interactive fuzzy programming approach for bi-objective straight and U-shaped assembly line balancing problem. Applied Soft Computing. 2016;40:221-35.

[14] Ebrahimnejad A. An improved approach for solving fuzzy transportation problem with triangular fuzzy numbers. Journal of intelligent & fuzzy systems. 2015;29(2):963-74.

[15] Garg H. A linear programming method based on an improved score function for interval-valued Pythagorean fuzzy numbers and its application to decision-making. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 2018;26(01):67-80.

[16] Mirzaee H, Naderi B, Pasandideh SHR. A preemptive fuzzy goal programming model for generalized supplier selection and order allocation with incremental discount. Computers & Industrial Engineering. 2018;122:292-302.

[17] Singh SK, Yadav SP. Efficient approach for solving type-1 intuitionistic fuzzy transportation problem. International Journal of System Assurance Engineering and Management. 2015;6(3):259-67.

[18] Darbari JD, Kannan D, Agarwal V, Jha P. Fuzzy criteria programming approach for optimising the TBL performance of closed loop supply chain network design problem. Annals of operations research. 2019;273(1-2):693-738.

[19] Xu Y, Xu A, Wang H. Hesitant fuzzy linguistic linear programming technique for multidimensional analysis of preference for multi-attribute group decision making. International Journal of Machine Learning and Cybernetics. 2016;7(5):845-55.

[20] Li M, Fu Q, Singh VP, Ma M, Liu X. An intuitionistic fuzzy multi-objective non-linear programming model for sustainable irrigation water allocation under the combination of dry and wet conditions. Journal of Hydrology. 2017;555:80-94.

[21] Paydar MM, Saidi-Mehrabad M. Revised multi-choice goal programming for integrated supply chain design and dynamic virtual cell formation with fuzzy parameters. International Journal of Computer Integrated Manufacturing. 2015;28(3):251-65.

[22] Tirkolaee EB, Goli A, Weber G-W, editors. Multi-objective aggregate production planning model considering overtime and outsourcing options under fuzzy seasonal demand. International Scientific-Technical Conference Manufacturing; 2019: Springer.

[23] Zaidan A, Atiya B, Bakar MA, Zaidan B. A new hybrid algorithm of simulated annealing and simplex downhill for solving multiple-objective aggregate production planning on fuzzy environment. Neural computing and applications. 2019;31(6):1823-34.

[24] Nematian J. An Extended Two-stage Stochastic Programming Approach for Water Resources Management under Uncertainty. Journal of Environmental Informatics. 2016;27(2).

[25] Gholamian N, Mahdavi I, Tavakkoli-Moghaddam R. Multi-objective multi-product multi-site aggregate production planning in a supply chain under uncertainty: fuzzy multi-objective optimisation. International Journal of Computer Integrated Manufacturing. 2016;29(2):149-65.

[26] Li D-F. Linear programming models and methods of matrix games with payoffs of triangular fuzzy numbers: Springer; 2015.

[27] Mohammed A, Harris I, Soroka A, Nujoom R. A hybrid MCDM-fuzzy multi-objective programming approach for a G-resilient supply chain network design. Computers & Industrial Engineering. 2019;127:297-312.

[28] Mahmoudi A, Liu S, Javed SA, Abbasi M. A novel method for solving linear programming with grey parameters. Journal of intelligent & fuzzy systems. 2019;36(1):161-72.

[29] Oliveira C, Coelho D, Antunes CH. Coupling input–output analysis with multiobjective linear programming models for the study of economy–energy–environment–social (E3S) trade-offs: a review. Annals of operations research. 2016;247(2):471-502.

[30] Garg H. Non-linear programming method for multi-criteria decision making problems under interval neutrosophic set environment. Applied Intelligence. 2018;48(8):2199-213.

[31] Chen S-M, Huang Z-C. Multiattribute decision making based on interval-valued intuitionistic fuzzy values and linear programming methodology. Information Sciences. 2017;381:341-51.

[32] Afzali A, Rafsanjani MK, Saeid AB. A fuzzy multi-objective linear programming model based on interval-valued intuitionistic fuzzy sets for supplier selection. International Journal of Fuzzy Systems. 2016;18(5):864-74.

[33] Subulan K, Taşan AS, Baykasoğlu A. A fuzzy goal programming model to strategic planning problem of a lead/acid battery closed-loop supply chain. Journal of manufacturing system Knowledge-Based Systems. 2015; 37:243-64.

STOCHASTIC ANALYSIS OF DISCRETE PARAMETRIC MARKOV CHAIN SYSTEM MODEL

Manoj Kumar[1], Shiv Kumar[2]

•

[1]D.A.V. College, Muzaffarnagar, U.P., (India)
[2]J.V. College Baraut (Baghpat), U.P., (India)
manojdu26@gmail.com
s.shiv@gmail.com

Abstract

The present paper deals with the behavior of the parallel model system of two non-identical units, warm standby models have been developed by in view of all random variables are independent. Initially priority unit is working and the non-priority unit is warm standby. Two repairmen are always available with the system to carry out the system operation as soon as possible, skilled repairmen carry out phase-1 repair while ordinary repairmen carry out a phase-2 repair. The main unit is take two phases for his repair while the repair of the ordinary unit is completed in one phase. The statistical measures of the model are analyzed probabilistically by applying the regenerative point technique the distribution of failure and repair time of the system taken as a geometric distribution with different parameters.

Keywords: Geometric distribution, Steady state transition probability, MTSF, Availability, Busy period, and Cost-benefit analysis.

1. Introduction

The configuration of the stochastic model is very complex with the development of modern system models, minimizing the high maintenance cost and increasing the system efficiency by reducing the frequency of failures. The design and model of industrial systems such as communication systems, satellite systems, power plant systems mechanical engineering, aeronautical engineering software engineering, and gaming systems are more complex to design in the current scenario. Using the different probabilistic measures of a two-unit system model with various kinds of repair policy deals with the system model involving various general human failures. Kumar and Kadyan [1] analyzed a non-identical parallel unit system with a single repairman visit whenever the original unit requires a repair facility, to repair the original unit with immediate effect and the duplicate unit is replaced by a similar new one. The various reliability characteristic such as study state availability, MTSF, and busy period and profit analysis of the system model are estimated by applying the semi-Markov approach. Sureria at el. [2] analyzed a computer system model whenever a system failed, priority is given to software replacement against hardware repair purpose to determine a mean sojourn time, reliability, availability, and busy period of a computer system of two similar units, initially one is active and the other is kept into cold standby whenever operative unite is failed, the cold unit is operative. The failure rate of the computer system is independent having an exponential distribution with different parameters while the repair and replacement rates distribution are taken as common. Each unit has hardware and software components that may have independent complete failure from the normal mode.

There is always a possibility that any system model during its operative condition to failure condition by two or more kinds of failure with single repair, post-repair, or waits for repair facility has been analyzed under some common assumption. Various other reliability characteristics models have been discussed of two identical or non-identical system models applying various types of repair facilities. Using discrete distribution, Bhatti at el. [3] introducing the concept of inspection to detect the major or minor failure, the repairman perform dual role of inspection and repair of the system after detect the type of failure of dissimilar operative cold standby systems. Ahmed at el. [4] studies a two non-identical parallel cold standby redundant unit system models each unit has two possible mode normal *(N)* and total failure *(F)*. A repairman is always available to repair the system whenever it's required for preventive maintenance, priority to repair the failed unit is by given initially operative unit after the repair of a unit works as well as new. The one parametric geometric distribution with different random variables is taken for failure and repair rate of the each unit. Malik [5] studied a repairable system under different weather conditions. Singh at el. [6] applying a probabilistic assessment of parallel system with correlated lifetime under different inspection method. Kumar at el. [7] analyzed a redundant system with priority and weibull distribution for failure and repair rate. Kumar at el. [8] introduce a repairable system of non-identical units with priority and conditional failure of repairman.

2. Methods System description and assumptions

The aim of the present paper deals with priority (unit-I) and non-priority (unit-II) parallel unit systems, each unit has two achievable modes normal *(N)* and total failure *(F)*, in the beginning one unit is operative and another unit is reserved in warm standby. Two repairmen are always available with the system to repair the failed unit. A master repairman carries out the phase-I repair while an assistant repairman is present to take out the phase-II repair. Initially, the failure unit-I goes to phase-I repair while completing phase-I repair it enters into phase-II for its final repair by the assistant repairman, and the repair of a non-priority unit is completed in one phase (phase-I) repaired by the master repairman. The operation priority is given to unit-I and repair priority is first come first serve *(FCFS)* bases. All the random variables are independent and uncorrelated under this study. The distributions of failure and repair times are taken as a discrete nature having a geometric distribution with different parameters. The system model is derived using the Markov-chain approach and using the regenerative point technique for various probabilistic analyses of the system effect such as mean sojourn time, reliability, availability, mean time to system failure *(MTSF)*, a busy period in the different repair facility and cost-benefit function have been derived. The system consists of the following assumptions:

- The system consists of priority and non-priority units, and they are connected in parallel. Initially, one unit is operative (unit-I) and the other is kept on warm standby (non-priority unit-II).
- Both units have two possible modes, normal *(N)* when the unit is operative and total failure *(F)* when the unit is in failure mode.
- Two repairmen are always with the system to carry out the repair facility, the repair of unit-I is completed in two phases while the repair of unit-II is completed in one phase. The master repairman perform phase-I repair while the assistant repairman perform phase-II repair.
- The priority unit failed than non-priority unit is loading warm standby unit into operation using switching device to be perfect, the repair of priority unit is completed in two phases (phase-I and phase-II) i.e., a failed unit first enters in phase-I for its repair and after the completion of phase-I repair it enters phase-II for finishing repair, and the repair of a non-priority unit is done in one phase (phase-I). After repair of a unit is work as well as a new one.

- The system transition rate from state S_i to S_j is independent having a one parametric discrete geometric distribution.
- The repair priority is first come first serve bases while the operation priority is given to unit-I.

3. Notations and states of the system

N_0^1/N_0^2	The unit-I/ unit-II is in normal-mode and operative.
F_r^1/F_r^2	The unit-I/unit-II is in failure-mode and under repair by master repairman.
F_w^1/F_w^2	The unit-I/unit-II is in failure-mode and waiting for repair.
F_R^1	The unit-I is in F-mode and under repair by assistant repairman.
N_s^2	The unit-II is in normal-mode and kept into standby.
pq^k/rs^k	Probability mass function of failure rate of unit-I/unit-II.
ab^k/cd^k	Probability mass function of repair rate of unit-I in phase-I/phase-II.
mn^k	Probability mass function of repair rate of unit-II.
q_{ij}, Q_{ij}	Probability mass function and cumulative density function of one step transition time from state $S_i$toS_j.
p_{ij}	Steady state transition probability from state $S_i$to S_j.
Ψ_i	Mean sojourn time in regenerative state S_i.
$Z_i(k)$	Probability that the system is operational, initially in state sojourns S_i up to time k.
h,*	Dummy variable used in geometric transformation and sign.
©	Symbol for ordinary convolution.

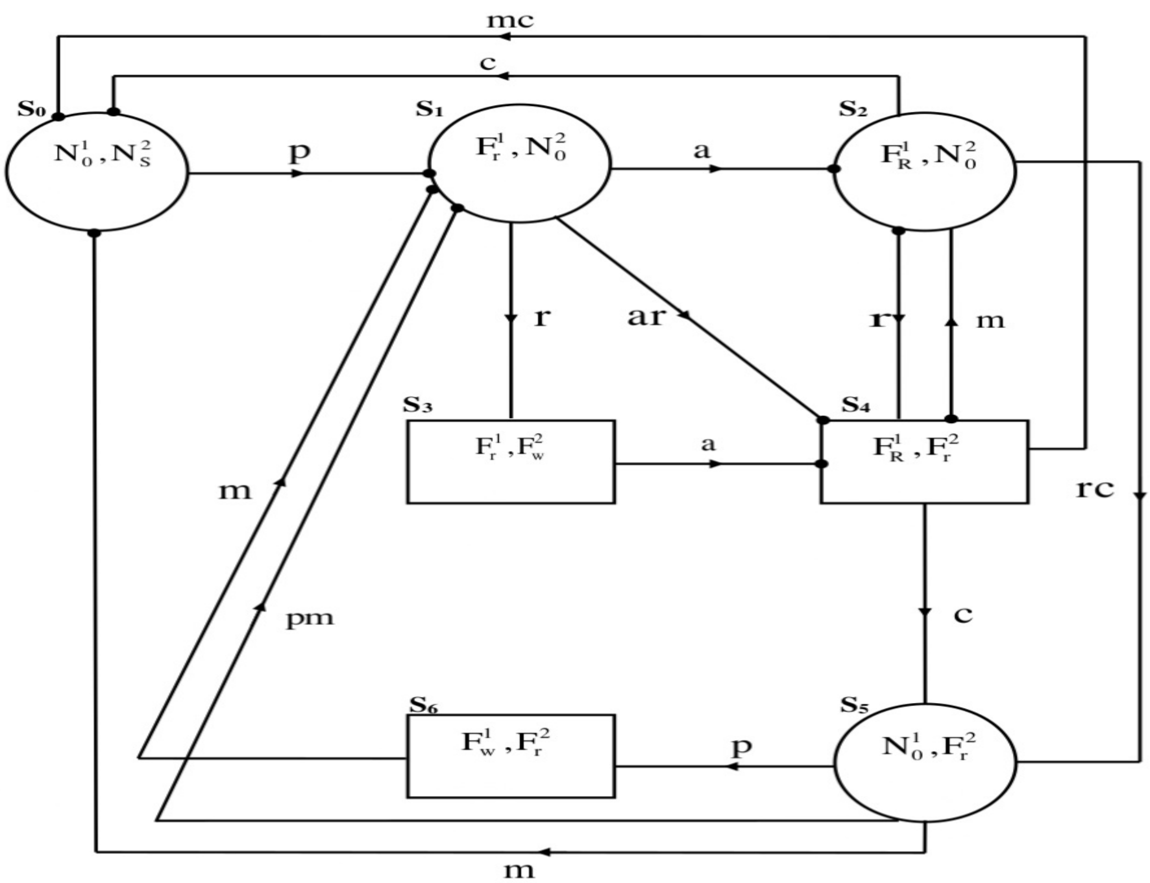


Figure 1: *Transitions Probability Diagram*

Operative States

$S_0 = (N_0^1, N_s^2);\ S_1 = (F_r^1, N_0^2);\ S_2 = (F_R^1, N_0^2)$ and $S_5 = (N_0^1, F_r^2)$

Failed States

$S_3 = (F_r^1, F_w^2);\ S_4 = (F_R^1, F_r^2)$ and $S_6 = (F_w^1, F_r^2)$

4. Transition probabilities and mean sojourn times

Using simple probabilistic arguments that the system transits from state S_i to S_j within time interval *(0, k)*, then $Q_{ij}(k)$ maybe obtain the following approach:

$$Q_{ij}(k) = [K_{n+1} = j, K_{n+1} - K_n < k | K_n = i]$$

$$Q_{01}(k) = 1 - q^{k+1} \qquad Q_{12}(k) = \frac{as\left[1-(bs)^{k+1}\right]}{1-bs}$$

$$Q_{13}(k) = \frac{rb\left[1-(bs)^{k+1}\right]}{1-bs} \qquad Q_{14}(k) = \frac{ar\left[1-(bs)^{k+1}\right]}{1-bs}$$

$$Q_{20}(k) = \frac{cs\left[1-(ds)^{k+1}\right]}{1-ds} \qquad Q_{24}(k) = \frac{rd\left[1-(sd)^{k+1}\right]}{1-sd}$$

$$Q_{25}(k) = \frac{rc\left[1-(sd)^{k+1}\right]}{1-sd} \qquad Q_{34}(k) = 1 - b^{k+1}$$

$$Q_{40}(k) = \frac{cm\left[1-(dn)^{k+1}\right]}{1-dn} \qquad Q_{42}(k) = \frac{md\left[1-(dn)^{k+1}\right]}{1-dn}$$

$$Q_{45}(k) = \frac{cn\left[1-(dn)^{k+1}\right]}{1-dn} \qquad Q_{50}(k) = \frac{mq\left[1-(nq)^{k+1}\right]}{1-nq}$$

$$Q_{51}(k) = \frac{mp\left[1-(nq)^{k+1}\right]}{1-nq} \qquad Q_{56}(k) = \frac{pn\left[1-(nq)^{k+1}\right]}{1-nq}$$

$$Q_{61}(k) = 1 - n^{k+1}$$

Similarly, using $p_{ij} = \lim_{k\to\infty} Q_{ij}(k)$, the steady state transition probability is:

$$p_{01} = p_{34} = p_{61} = 1,\ p_{12} = \frac{as}{1-bs},\ p_{13} = \frac{rb}{1-bs},\ p_{14} = \frac{ar}{1-bs},\ p_{20} = \frac{cs}{1-ds}, p_{24} = \frac{rd}{1-ds},\ p_{25} = \frac{rc}{1-ds}, p_{40} = \frac{cm}{1-dn}, p_{42} = \frac{md}{1-dn},\ p_{45} = \frac{cn}{1-dn},\ p_{50} = \frac{mq}{1-nq}, p_{51} = \frac{mp}{1-nq} \text{ and } p_{56} = \frac{pn}{1-nq}$$

We can easily verify that

$p_{12} + p_{13} + p_{14} = 1, p_{20} + p_{24} + p_{25} = 1, p_{40} + p_{42} + p_{45} = 1$ and $p_{50} + p_{51} + p_{56} = 1$

5. Mean sojourn time

The expected time a system spends in one state before moving onto another state is known as the mean sojourn time Ψ_i in state S_i; *i=0,1,2,3,4,5,6* is defined as:

$$\Psi_i = E[K_i] = \sum_{k=1}^{\infty} P[K_i \geq K]$$

So that

$$\Psi_0 = \frac{p}{q}, \Psi_1 = \frac{bs}{1-bs}, \Psi_2 = \frac{ds}{1-ds}, \Psi_3 = \frac{b}{a}, \Psi_4 = \frac{dn}{1-dn}, \Psi_5 = \frac{qn}{1-qn} \text{ and } \Psi_6 = \frac{n}{m}$$

6. Reliability of the system and mean time to system failure (MTSF)

The system originally starts operational from state $S_i \in E$. Then the system reliability, $R_i(k)$; *i = 0, 1, 2, 5;* have the following set of convolution equations is given by:

$$R_0(k) = q^k + \sum_{u=0}^{k-1} q_{01}(u)©R_1(k-1-u)$$

$R_0(k) = Z_0(k) + q_{01}(k-1)©R_1(k-1)$

Similarly,

$R_1(k) = Z_1(k) + q_{12}(k-1)©R_2(k-1)$

$R_2(k) = Z_2(k) + q_{20}(k-1)©R_0(k-1) + q_{25}(k-1)©R_5(k-1)$

$R_5(k) = Z_5(k) + q_{50}(k-1)©R_0(k-1)$

where,

$Z_1(k) = b^k s^k; \quad Z_2(k) = d^k s^k$ and $Z_5(k) = q^k n^k$

Using the geometric transformation of the above set of equations, get the algebraic solutions for $R_0^*(h)$. We get

$R_0^*(h) = \frac{N_1(h)}{D_1(h)}$

where,

$N_1(h) = Z_0^*(h) + hq_{01}^* Z_1^*(h) + h^2 q_{01}^* q_{12}^* Z_2^*(h) + h^3 q_{01}^* q_{12}^* q_{25}^* Z_5^*(h)$

$D_1(h) = 1 - h^3 q_{01}^* q_{12}^* q_{20}^* - h^4 q_{01}^* q_{12}^* q_{25}^* q_{50}^*$

The MTSF is given by:

$E(K_0) = \lim_{s\to 0} R_0^*(h) = \frac{N_1(0)}{D_1(0)}$

To determine $N_1(0)$ and $D_1(0)$, we apply the results

$Z_i^*(0) = \Psi_i$ and $q_{ij}(0) = p_{ij}$

We get,

$$\text{MTSF} = \frac{\Psi_0 + \Psi_1 + p_{12}\Psi_2 + p_{12}p_{25}\Psi_5}{1 - p_{12}p_{20} - p_{12}p_{25}p_{50}}$$

7. Availability analyses

Let $A_i(k)$; *i=0,1,2,3,4,5,6* be the probability that the system will be normal at epoch time *k*, when at the system start function from state $S_i \in E$. We observe the following recurrence relations can be easily developed for $A_i(k)$, using similar probabilistic arguments:

$$A_0(k) = q^k + \sum_{u=0}^{k-1} q_{01}(u)©A_1(k-1-u)$$

$A_0(t) = Z_0(k) + q_{01}(k-1)©A_1(k-1)$

Similarly,

$A_1(k) = Z_1(k) + q_{12}(k-1)©A_2(k-1) + q_{13}(k-1)©A_3(k-1) + q_{14}(k-1)©A_4(k-1)$

$A_2(k) = Z_2(k) + q_{20}(k-1)©A_0(k-1) + q_{24}(k-1)©A_4(k-1) + q_{25}(k-1)©A_5(k-1)$

$A_3(k) = q_{34}(k-1)©A_4(k-1)$

$A_4(k) = q_{40}(k-1)©A_0(k-1) + q_{42}(k-1)©A_2(k-1) + q_{45}(k-1)©A_5(k-1)$

$A_5(k) = Z_5(k) + q_{50}(k-1)©A_0(k-1) + q_{51}(k-1)©A_1(k-1) + q_{56}(k-1)©A_6(k-1)$

$A_6(k) = q_{61}(k-1)©A_1(k-1)$(17-23)

where,

$Z_1(k); Z_2(k);$ and $Z_5(k)$ same as in reliability.

After solving the set of algebraic equations that emerge from applying geometric transforms to the equations above, we have

$$A_0^*(s) = \frac{N_2(s)}{D_2(s)}$$

where,

$N_2(s) = \{1 - q_{24}^* q_{42}^* - (q_{51}^* + q_{56}^* q_{61}^*)[q_{45}^*(q_{13}^* q_{34}^* + q_{14}^*) + q_{25}^* q_{42}^*(q_{13}^* q_{34}^* + q_{14}^*) + q_{12}^* q_{25}^* + q_{12}^* q_{24}^* q_{45}^*]\}Z_0^* - (q_{01}^* q_{24}^* q_{42}^* - q_{01}^*)Z_1^* + [q_{01}^* q_{12} + q_{01}^* q_{42}^*(q_{13}^* q_{34}^* + q_{14}^*)]Z_2^* + [q_{01}^* q_{12}^*(q_{24}^* q_{45}^* + q_{25}^*) + q_{01}^* q_{45}^*(q_{13}^* q_{34}^* + q_{14}^*) + q_{01}^* q_{25}^* q_{42}^*(q_{13}^* q_{34}^* + q_{14}^*)]Z_5^*$

and

$D_2(s) = 1 - q_{01}^* q_{12}^*(q_{20}^* + q_{24}^* q_{40}^* + q_{24}^* q_{45}^* q_{50}^* + q_{25}^* q_{50}^*) - q_{01}^*(q_{13}^* q_{34}^* + q_{14}^*)(q_{20}^* q_{42}^* + q_{25}^* q_{42}^* q_{50}^* + q_{40}^* + q_{45}^* q_{50}^*) - q_{12}^*(q_{51}^* + q_{56}^* q_{61}^*)(q_{24}^* q_{45}^* + q_{25}^*) - (q_{13}^* q_{34}^* + q_{14}^*)(q_{51}^* + q_{56}^* q_{61}^*)(q_{25}^* q_{42}^* + q_{45}^*) -$

$q_{24}^{*}q_{42}^{*}$

Now, the steady state availability i.e. the probability that the system will be active in long run is known as:

$$A_0 = \lim_{k\to\infty} A_0(k) = \lim_{s\to 0} sA_0^{*}(s) = \lim_{s\to 0} s\frac{N_2(s)}{D_2(s)}$$

Since, $D_2(0) = 0$, therefore by applying L-Hospital rule;

$$A_0 = \lim_{s\to 0}\frac{N_2(s)}{D_2'(s)} = \frac{N_2(0)}{D_2'(0)}$$

where,

$$N_2(0) = [1 - p_{24}p_{42} - (p_{13} + p_{14})(p_{51} + p_{56})(p_{45} + p_{25}p_{42}) - p_{12}(p_{51} + p_{56})(p_{25} + p_{24}p_{45})]\Psi_0 + (1 - p_{24}p_{42})\Psi_1 + [p_{12} + p_{42}(p_{13} + p_{14})]\Psi_2 + [p_{12}(p_{24}p_{45} + p_{25}) + (p_{13} + p_{14})(p_{25}p_{42} + p_{45})]\Psi_5$$

and

$$D_2'(0) = [(p_{13} + p_{14})(p_{20}p_{42} + p_{25}p_{42}p_{50} + p_{40} + p_{45}p_{50}) + p_{12}(p_{20} + p_{25}p_{50}) + p_{12}p_{24}(p_{40} + p_{45}p_{50})]\Psi_0 + [p_{42}(p_{20} + p_{25}) + p_{40} + p_{45}]\Psi_1 + [p_{12} + p_{42}(p_{13} + p_{14})]\Psi_2 + [p_{13}p_{42}(p_{20} + p_{25}) + p_{13}(p_{40} + p_{45})]\Psi_3 + (p_{12}p_{24} + p_{13} + p_{14})\Psi_4 + [(p_{13} + p_{14})(p_{25}p_{42} + p_{45}) + p_{12}(p_{24}p_{45} + p_{25})] + [p_{56}(p_{13} + p_{14})(p_{25}p_{42} + p_{45}) + p_{12}p_{56}(p_{24}p_{45} + p_{25})]\Psi_6$$

8. Busy period for master repairman

Let $B_i^r(k)$; *i=0,1,2,3,4,5,6* be the probability that the master repairman is busy repairing the failed unit in phase-I at epoch time *k* when the system operational from the state $S_i \in E$. Now for $B_0^r(k)$, we have the sum of the probabilities of the following contingencies:

$$B_0^r(k) = \sum_{u=0}^{k-1} q_{01}(u)©B_1^k(k - 1 - u)$$

$$B_0^r(k) = q_{01}(k - 1)©B_1^r(k - 1)$$

Similarly,

$$B_1^r(k) = Z_1^r(k) + q_{12}(k - 1)©B_2^r(k - 1) + q_{13}(k - 1)©B_3^r(k - 1) + q_{14}(k - 1)©B_4^r(k - 1)$$
$$B_2^r(k) = q_{20}(k - 1)©B_0^r(k - 1) + q_{24}(k - 1)©B_4^r(k - 1) + q_{25}(k - 1)©B_5^r(k - 1)$$
$$B_3^r(k) = Z_3^r(k) + q_{34}(k - 1)©B_4^r(k - 1)$$
$$B_4^r(k) = Z_4^r(k) + q_{40}(k - 1)©B_0^r(k - 1) + q_{42}(k - 1)©B_2^r(k - 1) + q_{45}(k - 1)©B_5^r(k - 1)$$
$$B_5^r(k) = Z_5^r(k) + q_{50}(k - 1)©B_0^r(k - 1) + q_{51}(k - 1)©B_1^r(k - 1) + q_{56}(k - 1)©B_6^r(k - 1)$$
$$B_6^r(k) = Z_6^r(k) + q_{61}(k - 1)©B_1^r(k - 1)$$

where,

$Z_1^r(k) = b^k s^k$; $Z_3^r(k) = b^k$; $Z_4^r(k) = d^k n^k$; $Z_5^r(k) = q^k n^k$ and $Z_6^r(k) = n^k$

Using the inverse Laplace transform of $B_0^{r*}(s)$, we get:

$$B_0^{r*} = \lim_{s\to 0} s\frac{N_3(s)}{D_2(s)}$$

here,

$D_2(0) = 0$

Therefore, by L-hospital rule, we have

$$B_0^{r*} = \lim_{s\to 0}\frac{N_3(s)}{D_2'(s)} = \frac{N_3(0)}{D_2'(0)}$$

where,

$$N_3(0) = (1 - p_{24}p_{42})(\Psi_1 + \Psi_3) + [p_{12}p_{24} + (p_{13} + p_{14})]\Psi_4 + [p_{12}(p_{24}p_{45} + p_{25}) + (p_{13} + p_{14})(p_{25}p_{42} + p_{45})](\Psi_5 + p_{56}\Psi_6)$$

9. Busy period for assistant repairman

Let $B_0^R(k)$ *i=0,1,2,3,4,5,6* be the probability that the master repairman is busy repairing the failed unit in phase-I at epoch time *k* when the system operational from the state $S_i \in E$. Now for $B_0^R(k)$, we have the sum of the probabilities of the following contingencies:

$$B_0^R(k) = \sum_{u=0}^{k-1} q_{01}(u)©B_1^R(k - 1 - u)$$

$B_0^R(k) = q_{01}(k-1)©B_1^R(k-1)$

Similarly,

$B_1^R(k) = q_{12}(k-1)©B_2^R(k-1) + q_{13}(k-1)©B_3^R(k-1) + q_{14}(k-1)©B_4^R(k-1)$

$B_2^R(k) = Z_2^R(k) + q_{20}(k-1)©B_0^R(k-1) + q_{24}(k-1)©B_4^R(k-1) + q_{25}(k-1)©B_5^R(k-1)$

$B_3^R(k) = q_{34}(k-1)©B_4^R(k-1)$

$B_4^R(k) = Z_4^R(k) + q_{40}(k-1)©B_0^R(k-1) + q_{42}(k-1)©B_2^R(k-1) + q_{45}(k-1)©B_5^R(k-1)$

$B_5^R(k) = q_{50}(k-1)©B_0^R(k-1) + q_{51}(k-1)©B_1^R(k-1) + q_{56}(k-1)©B_6^R(k-1)$

$B_6^R(k) = q_{61}(k-1)©B_1^R(k-1)$

where,

$Z_2^R(k) = d^k s^k$ and $Z_4^R(k) = d^k n^k$

Using the inverse Laplace transform of $B_0^{R*}(s)$ we get:

$$B_0^{R*} = \lim_{s\to 0} s\frac{N_4(s)}{D_2(s)}$$

here,

$D_2(0) = 0$

Therefore, by L-hospital rule, we have

$$B_0^{R*} = \lim_{s\to 0}\frac{N_4(s)}{D_2^{'}(s)} = \frac{N_4(0)}{D_2^{'}(0)}$$

where,

$N_4(0) = [p_{12} + p_{42}(p_{13} + p_{14})]\Psi_2 + [p_{12}p_{24} + (p_{13} + p_{14})]\Psi_4$

10. Profit analysis

The system model net-expected profit during the time interval *(0, k)* is given below:

P(k) = Expected total revenue in *(0, k)* - Expected cost of repair in *(0, k)*

$P(k) = C_0\mu_{up}(k) - C_1\mu_b^r(k) - C_2\mu_b^R(k)$

WhereC_0 per-unit up time revenue by the system due to the operation of unit-I and unit-II, C_1 and C_2 are the repair cost per-unit of time when unit is repair by master repairman and assistant repairman respectively.

The expected total cost per-unit time in steady state is given by:

$$P = \lim_{k\to\infty}\frac{P(k)}{k}$$
$$= C_0A_0 - C_1B_0^r - C_2B_0^R$$

Where A_0 , B_0^r and B_0^R have been already defined.

11. Conclusion

This paper concludes with an analysis of stochastic modeling of various reliability measures such as MTSF, availability and busy period for a master repairman, assistant repairman, and profit analysis by different levels of performance. Let us suppose that the random variables follow a geometric distribution with dissimilar probability mass functions. The numerical analysis of MTSF, availability, and profit analysis have been studied at various levels of failure rate *(q)* of unit-I, and failure rate *(s)* of unit-II by fixing the values of certain parameters *a=0.8, b=0.2, c=0.6, d=0.4, m=0.4* and *n=0.6.* Table 1 and Figure 2 a show the variation in MTSF is decries by increasing the failure rate of unit-I and unit-II. The availability is linearly falling shown in Table 2 and Figure 3, for various values of the failure rate of unit-I and unit-II. Also putting the other parameters C_0*=10000,* C_1*=2000,* and C_2*=1000* the profit analysis concerning various values of failure rate *(q)* of unit-I, failure rate *(s)* of unit-II, and the fixing value of *a, b, c, d, m,* and *n* showed in a smooth curve in Figures 4 and Table 3.

References

[1] Kumar, J. and Kadyan, M.S., (2012). Profit analysis of a system of non-identical units with degradation and replacement, International journal of computer application, Vol. 40 (3): 19-25.

[2] Sureria, J.K., Malik, S.C. and Anand, J., (2012). Cost benefit analysis of a computer system with priority to software replacement over hardware repair, Applied Mathematical Sciences, Vol. 6 (75): 3723-3734.

[3] J. Bhatti, A. K. Chitkara, M. K. Kakkar, (2016). Stochastic analysis of dis-similar standby system with discrete failure, inspection and replacement policy, Demonstratio Mathematica, Vol. 49(2): 224-235.

[4] M.A. El-Damcese, N. H. El-Sodany, (2015). Discrete Time Semi-Markov Model of a Two Non-Identical Unit Cold Standby System with Preventive Maintenance with Three Modes, American Journal of Theoretical and Applied Statistics, Vol. 4 (4): 277-290.

[5] Malik, S.C., (2016). Stochastic Modeling of a Repairable System under Different Weather Conditions, Recent Advances in Mathematics Statistics and Computer Science, 155-163.

[6] Singh, V. V., Poonia, P.K., (2019). Probabilistic Assessment of Two-Unit Parallel System with Correlated Lifetime under Inspection Using Regenerative Point Technique, International Journal of Reliability, Risk and Safety: Theory and Application, Vol. 2 (1): 5-14.

[7] Kumar, A., Saini, M., Devi, K., (2016). Analysis of a redundant system with priority and weibull distribution for failure and repair, Cogent Mathematics, Vol. 3 (1).

[8] 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 & Application, No 1 (67), Vol. 17: 123-133.

Appendix

Table 1: *Effect of a, b, c, d, m and n on system performance with respect to various failure rate of unit-I and unit-II*

		a=0.8, b=0.2, c=0.6, d=0.4, m=0.4 and n=0.6		
Failure rate of unit-I *(p)*	Failure rate of unit-II *(r)*	MTSF	Availability	Profit Analysis
0.02	0.01	47.66470	0.99	8117.19246
0.04	0.02	23.79169	0.98	8044.72539
0.06	0.03	15.86021	0.97	7966.79472
0.08	0.04	11.91403	0.96	7885.26352
0.10	0.05	9.562128	0.95	7801.60347
0.12	0.06	8.007633	0.94	7716.98608
0.14	0.07	6.909151	0.93	7632.35110
0.16	0.08	6.096083	0.92	7548.45847
0.18	0.09	5.473738	0.91	7465.92800
0.20	0.10	4.985377	0.90	7385.27013

Table 2: *Effect of a, b, c, d, m and n on system performance with respect to various failure rate of unit-I and unit-II*

		a=0.8, b=0.2, c=0.6, d=0.4, m=0.4 and n=0.6		
Failure rate of unit-I *(p)*	Failure rate of unit-II *(r)*	MTSF	Availability	Profit Analysis
0.03	0.02	23.81516	0.98	8030.99201
0.05	0.04	11.88649	0.97	7855.23332
0.07	0.06	7.916658	0.95	7677.4412
0.09	0.08	5.938082	0.93	7499.77873
0.11	0.10	4.757167	0.91	7323.78275
0.13	0.12	3.976017	0.90	7150.55723
0.15	0.14	3.424099	0.88	6980.90371
0.17	0.16	3.016157	0.86	6815.41137
0.19	0.18	2.704848	0.85	6654.52024
0.21	0.20	2.461794	0.83	6498.56645

Table 3: *Effect of a, b, c, d, m and n on system performance with respect to various failure rate of unit-I and unit-II*

		a=0.8, b=0.2, c=0.6, d=0.4, m=0.4 and n=0.6		
Failure rate of unit-I *(p)*	Failure rate of unit-II *(r)*	MTSF	Availability	Profit Analysis
0.04	0.04	11.86965	0.97	7846.51242
0.06	0.08	5.897579	0.94	7493.47397
0.08	0.12	3.903218	0.91	7159.24093
0.10	0.16	2.905934	0.88	6841.64064
0.12	0.20	2.309357	0.85	6539.07996
0.14	0.24	1.914567	0.82	6250.37274
0.16	0.28	1.636242	0.79	5974.63725
0.18	0.32	1.431696	0.77	5711.23236
0.20	0.36	1.277211	0.74	5459.71658

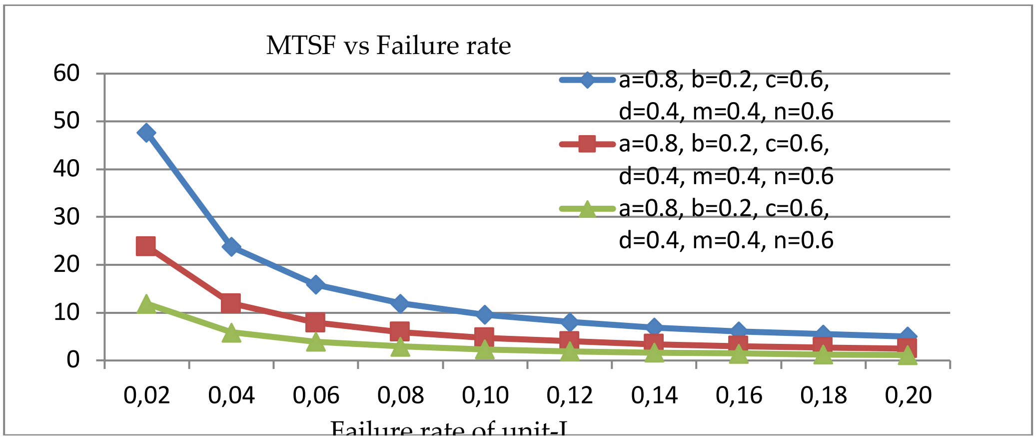


Figure 2: *MTSF vs failure rate of unit-I (p) and unit-II (r)*

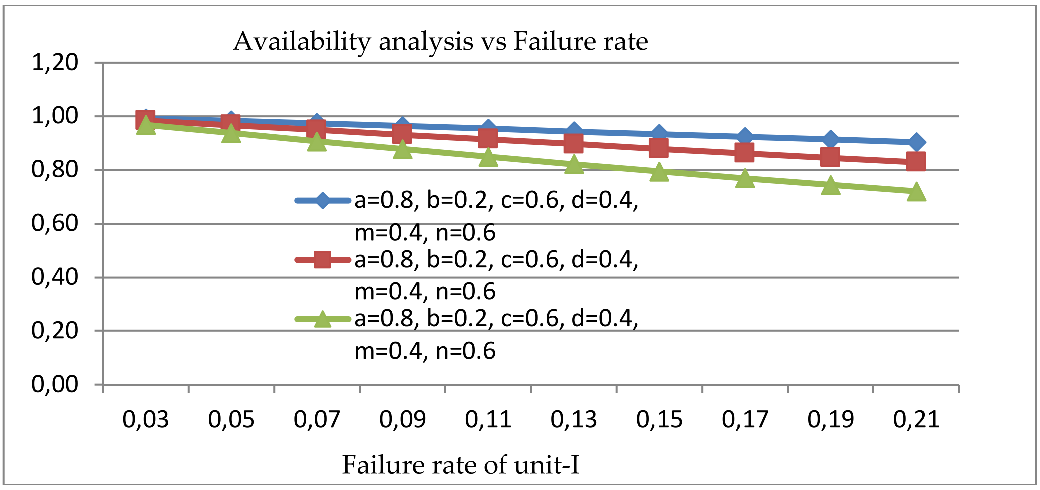


Figure 3: *Availability analysis vs failure rate of unit-I (p) and unit-II (r)*

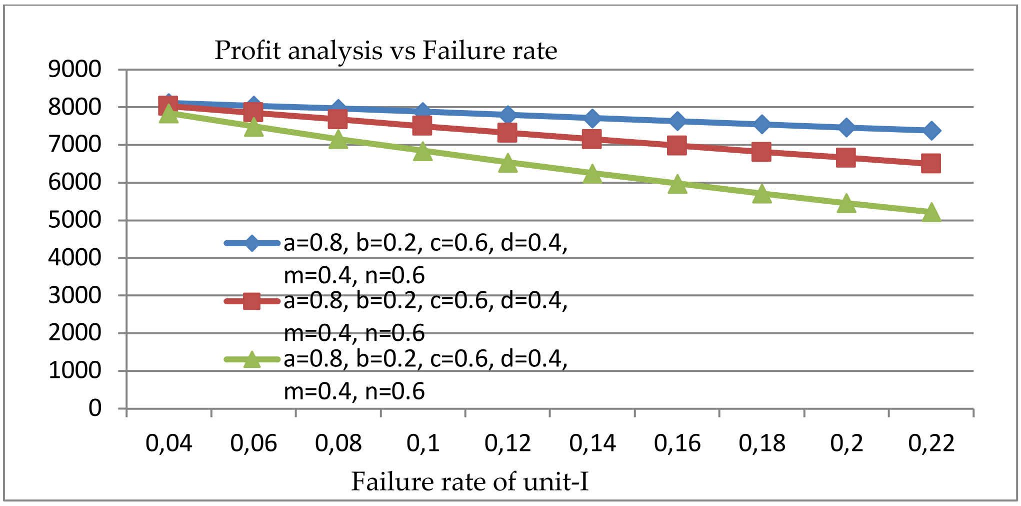


Figure 4: *Profit analysis vs failure rate of unit-I (p) and unit-II (r)*

RT&A, No 1 (72)
Volume 18, March 2023

PHASE TYPE QUEUEING MODEL OF SERVER VACATION, REPAIR AND DEGRADING SERVICE WITH BREAKDOWN, STARTING FAILURE AND CLOSE-DOWN

G. Ayyappan, S. Meena

•

Department of Mathematics,
Puducherry Technological University,
Puducherry, India.
ayyappan@ptuniv.edu.in, meenasundar2296@gmail.com

Abstract

We consider a single server phase type queueing model with server vacation, repair, breakdown, degrading service, starting failure and closedown. When the arrival rate of the customer follows the Markovian Arrival Process (MAP) and the service rate of the server follows the phase-type distribution. If no one is in the system when the server is back from the vacation, then the server will wait until the customer arrives. If the customer arrives at the moment with no starting failure, then he provides service, otherwise the server immediately goes to the repair process. Here, the service rate declining until degradation fixed. After completion of K services the degradation is addressed. During the period of service, the server may get a breakdown at any moment, and then the server immediately goes for a repair process. After completing the service, he switches to the close-down process, and then he goes on vacation. Using the Matrix-Analytic method, The stationary probability vector representing the total number of customers in the system is examined. The analysis of the busy period, the mean waiting time, and cost analysis are discussed. A few significant performance measures are attained. Finally, some numerical examples are given.

Keywords: Phase type Distribution, Markovian Arrival Process, Degrading Service, Server Vacation, Breakdown, Repair, Starting failure, Close-down, Matrix-analytic method.

AMS Subject Classification (2010): 60K25, 68M30, 90B22.

1. Introduction

The Markovian arrival process is one of the modelling techniques for studying point processes that is most flexible. In order to define arrival processes that are not fundamentally renewal processes, Neuts [13] proposed the concept of a versatile Markovian point process (VMPP). Neuts [14] first introduced and investigated the underlying Markov structure of the MAP, which fits perfectly into the framework of matrix-analytic methods and is one of its most notable properties. Qi-Ming He [16] investigated the foundations of matrix analytical methodologies in order to comprehend the idea of service and arrival process.

Chakravarthy [5] made a significant contribution to MAP. Markovian Arrival Process represents by (D_0, D_1) and the service times with representation (α, T) that follow phase type distribution and whose matrices of order m and n, respectively. He described several types of arrivals and services. The irreducible stochastic matrix $D = D_0 + D_1$ defines the generator D. If the irreducible generator D describes the Markov process, then π is the steady state probability vector, and it is defined as $\pi D = 0$ and $\pi e = 1$. Based on the Markovian arrival process, the constant $\lambda = \pi D_1 e$ represents the basic customer arrival rate per unit time.

MAP/PH/1-type queueing models with degradation and phase type vacation have been analysed by Alka et al. [6]. Degradation can be included in a service system in a number of ways. The service rate will decrease unless the degradation is addressed. In other words, the service rate will decrease as more services are provided. For vacation queueing models, we refer to Doshi's survey paper [7] and Tian and Zhang's book [20]. Li and Tian [12] investigated the M/M/1 model with working vacation and proposed an interruption in vacation, where the server returns without completing the ongoing vacation due to certain conditions. Krishna Kumar et al. [10] have analysed the several server model with server vacations under the Bernoulli schedule. Sreenivasan et al. [18] have examined the MAP/P H/1 queueing model with N-Policy, vacation interruption and working vacations.

One of the main queueing theory subfields has recently been queueing models with server breakdown. Wang et al. [21] have investigated the batch arrival queueing model with multiple vacations and the server struck with breakdown. Ayyappan and Nirmala [2] have explored the non-Markovian queueing model and the server provides service to the customers based on general bulk service rule with multiple vacations, breakdown and two-phase repair . Ayyappan and Deepa [1] have studied the batch arrival and bulk service queueing model with multiple vacations and optional repair. A single server queueing model with MAP arrival and phase type service, vacation, instantaneous feedback and breakdown has been looked into by Ayyappan and Thilagavathy [3]. In this model, they obtained stability condition and busy period analysis. Senthil Vadivu et al. [17] have performed a cost function of the bulk service queueing model of a single server with finite capacity and close-down times by using embedded Markov chain and supplementary variable techniques.

Yang et al. [22] have discussed the Markovian model of the retrial queue with multi-server and starting failure. They analyzed their model with the aid of the matrix geometric method. With respect to the stability condition, the cost analysis is built to calculate the ideal number of servers, the ideal average service rate, and the ideal average repair rate. Karpagam et al. [9] have been analysed the batch arrival and bulk service queueing system with starting failure and additional service. They obtained system performance measures and the stability condition. Ayyappan and Gowthami [4] has analysed a Phase type model with impatient customers, Setup time, vacation, feedback, Breakdown and Repair. In this article, they compute the average waiting time.

2. Description of the Model

Assume that there is a single server in a queueing model, and that customers arrive at the system according to the MAP with representation (D_0, D_1), where D_0 and D_1 are m-dimensional square matrices. Let $D = D_0 + D_1$ be the generator matrix, where D_0 governs for no arrival at the system and D_1 governs for an arrival at the system. The stationary vector of D is denoted by π, so we have $\pi D = 0$ and $\pi e = 1$. The arrival rate λ is given by $\lambda = \pi D_1 e$. The system is performed on an FCFS basis. With the notation (α, T), that is of order n, the length of the server's service is thought to be a PH-distribution, where $T^0 + Te = 0$ so that $T^0 = -Te$. The average service rate ξ is given by $\xi = [\alpha(-T)^{-1}e]^{-1}$. The service rate decreases after each service is completed. Let ξ be the first service rate and ξ_i be the i^{th} service rate such that $\xi = \xi_1 \geq \xi_2 \geq \xi_3 \geq \cdots \geq \xi_K$, where $\xi_i = \theta_i \xi$ and $0 < \theta_i \leq 1$ for all $i = 1, 2, 3, \ldots, K$. After K services are completed, the original rate of ξ is immediately applied to the degraded service rate. Because $\theta_1 = 1$, After the degradation has been corrected, the service rate for the first customer is always ξ. The server that customers use to access services could breakdown at any time and needs to be repaired.

The repair procedure is based on the PH-distribution with representation (β, S) of order n_2 and $S^0 + Se = 0$ so that $S^0 = -Se$. If no one is present in the system when the server's service is completed, the close-down process begins, and then the server goes on vacation. The vacation period is thought to be a PH-distribution with the notation (γ, V) of order n_1, where $V^0 + Ve = 0$ so that $V^0 = -Ve$. After completion of the vacation period if no customer present in the system, then the server is idle; otherwise the server starts the service. If a customer arrives while the server is idle, it may experience a starting failure with probability p or no starting failure with probability q, resulting in p+q=1. In the event of a server breakdown, the customer who is currently providing the service from the server will remain in a frozen state until the server gets rid of the repair process. After completion of the repair process, the server will serve a fresh service for the current frozen customer. The breakdown and close-down time follows an exponential distribution with the parameters σ and δ respectively. The average repair rate and vacation rate are given by ζ and η respectively.

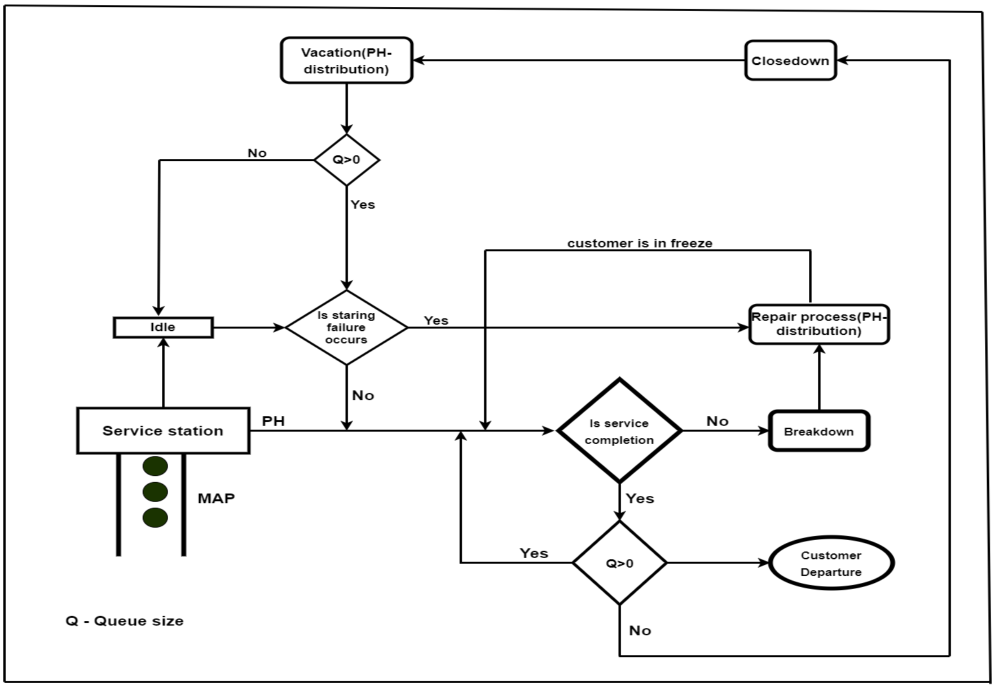


Figure 1: *Schematic Representation of the model*

3. The QBD Process of Matrix Generation

We have described our model's notation for the basis of generating the QBD process in this section as follows.

Matrix Generation Notations

- $\otimes$ - Kronecker product represents the product of any two different order matrices, can refer to the works in Steeb et al. [19].
- $\oplus$ - The Kronecker Sum represents the sum of any two of the different orders of matrices.
- I_k - An identity matrix of order k.
- $e_i'(m)$ - An m-dimensional row vector with 1 in the i^{th} position and 0 elsewhere.
- e-Each entry in a column vector of appropriate dimension is 1.
- The customer's arrival rate is denoted by λ and is defined by $\lambda = \pi D_1 e_m$
- The server's service rate is denoted by ξ and is defined by
  $\xi = [\alpha(-T)^{-1}e_n]^{-1}$

- The server's vacation rate is denoted by η and is defined by $\eta = [\gamma(-V)^{-1}e_{n_1}]^{-1}$
- The server's repair rate is denoted by ζ and is defined by $\zeta = [\beta(-S)^{-1}e_{n_2}]^{-1}$
- Define $\theta = (\theta_1, \theta_2, \ldots, \theta_K)^t$ and $\Delta(\theta) = \begin{pmatrix} \theta_1 & 0 & \ldots & \ldots \\ 0 & \theta_2 & \ldots & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \theta_K \end{pmatrix}$
- Let N(t) be the number of customers in the system at epoch t
- Let V(t) be the server's status at epoch t

$$V(t) = \begin{cases} 0, \text{ if the server is on vacation,} \\ 1, \text{ if the server is in idle,} \\ 2, \text{ if the server is in busy,} \\ 3, \text{ if the server is in repair process,} \\ 4, \text{ if the server is in closedown process.} \end{cases}$$

- I(t) is the type of service at time t
- $J_1(t)$ represents the vacation process as framed by phases.
- $J_2(t)$ represents the repair process as framed by phases.
- S(t) represents the service process as framed by phases.
- M(t) represents the arrival process as framed by phases.

Let { N(t),V(t),I(t),$J_1(t)$,$J_2(t)$,S(t),M(t):$t \geq 0$} denote the Continuous Time Markov Chain (CTMC) with state level independent Quasi-Birth and Death process, the state space of which is as follows:

$$\Omega = l(0) \cup l(q) \ ,$$

where

$$l(0) = \{(0,0,j_1,k) : 1 \leq j_1 \leq n_1, 1 \leq k \leq m\} \cup \{(0,1,k) : 1 \leq k \leq m\} \cup \{(0,4,k) : 1 \leq k \leq m\}$$

for $q \geq 1$,

$$l(q) = \{q,0,j_1,k) : 1 \leq j_1 \leq n_1, 1 \leq k \leq m\} \cup \{(q,2,l,j,k) : 1 \leq l \leq K, 1 \leq j \leq n, 1 \leq k \leq m\}$$
$$\cup \{(q,3,l,j_2,k) : 1 \leq l \leq K, 1 \leq j_2 \leq n_2, 1 \leq k \leq m\} \cup \{(q,4,k) : 1 \leq k \leq m\}.$$

The QBD process's infinitesimal matrix generation is given by

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & \ldots \\ B_{10} & A_1 & A_0 & 0 & 0 & 0 & \ldots \\ 0 & A_2 & A_1 & A_0 & 0 & 0 & \ldots \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

The entries in the block matrices of Q are defined as follows,

$$B_{00} = \begin{bmatrix} V \oplus D_0 & V^0 \otimes I_m & 0 \\ 0 & D_0 & 0 \\ \gamma \otimes \delta I_m & 0 & D_0 - \delta I_m \end{bmatrix},$$

$$B_{01} = \begin{bmatrix} I_{n_1} \otimes D_1 & 0 & 0 & 0 \\ 0 & e_1'(K) \otimes \alpha \otimes qD_1 & e_1'(K) \otimes \beta \otimes pD_1 & 0 \\ 0 & 0 & 0 & D_1 \end{bmatrix}, B_{10} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta \otimes T^0 \otimes I_m \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$A_0 = \begin{bmatrix} I_{n_1} \otimes D_1 & 0 & 0 & 0 \\ 0 & I_K \otimes I_n \otimes D_1 & 0 & 0 \\ 0 & 0 & I_K \otimes I_{n_2} \otimes D_1 & 0 \\ 0 & 0 & 0 & D_1 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} V \oplus D_0 & e_1'(K) \otimes qV^0\alpha \otimes I_m & e_1'(K) \otimes pV^0\beta \otimes I_m & 0 \\ 0 & (\Delta(\theta) \otimes T) \oplus D_0 - \sigma I_{Knm} & I_K \otimes (e_n \otimes \beta) \otimes \sigma I_m & 0 \\ 0 & I_K \otimes S^0\alpha \otimes I_m & (I_K \otimes S) \oplus D_0 & 0 \\ \gamma \otimes \delta I_m & 0 & 0 & D_0 - \delta I_m \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{22} = \begin{bmatrix} 0 & \theta_1 T^0\alpha \otimes I_m & 0 & \dots & 0 \\ 0 & 0 & \theta_2 T^0\alpha \otimes I_m & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \theta_{K-1} T^0\alpha \otimes I_m \\ \theta_K T^0\alpha \otimes I_m & 0 & \dots & \dots & 0 \end{bmatrix}.$$

4. Analysis of Stability Condition

We examined our model under the assumption that the system is stable.

4.1. Condition for Stableness

Let us specify the matrix A as $A = A_0 + A_1 + A_2$. It clearly demonstrates that the order of the square matrix A is $n_1m + Knm + Kn_2m + m$ and this matrix is an irreducible infinitesimal generator matrix. Let φ indicate the steady-state probability vector of A and it satisfying $\varphi A = 0$ and $\varphi e = 1$. The vector φ is partitioned by $\varphi = (\varphi_0, \varphi_1, \varphi_2, \varphi_3)$=$(\varphi_0, \varphi_{11}, \varphi_{12}, \varphi_{13}, \dots, \varphi_{1K-1}, \varphi_{1K}, \varphi_{21}, \varphi_{22}, \varphi_{23}, \dots, \varphi_{2K-1}, \varphi_{2K}, \varphi_3)$, where φ_0 is of dimension n_1m, φ_1 is of dimension Knm, φ_2 is of dimension Kn_2m, φ_3 is of dimension m. Our model's stability should satisfy the necessary and sufficient condition $\varphi A_0 e < \varphi A_2 e$ when the Markov Process is investigated using the Quasi-Birth-and-Death structure. The probability vector φ is calculated by solving the following equations

$$\begin{aligned}
&(V \oplus D)\varphi_0 + (\gamma \otimes \delta I_m)\varphi_3 = 0, \\
&(qV^0\alpha \otimes I_m)\varphi_0 + (\theta_1 T \oplus D - \sigma I_{nm})\varphi_{11} + (\theta_L T^0\alpha \otimes I_m)\varphi_{1K} + (S^0\alpha \otimes I_m)\varphi_{21} = 0, \\
&(\theta_{j-1} T^0\alpha \otimes I_m)\varphi_{1j-1} + (\theta_j T \oplus D - \sigma I_{nm})\varphi_{1j} + (S^0\alpha \otimes I_m)\varphi_{2j} = 0 \ \ for\ 2 \le j \le K, \\
&(pV^0\beta \otimes I_m)\varphi_0 + (e_n \otimes \beta \otimes \sigma I_m)\varphi_{11} + (S \oplus D)\varphi_{21} = 0, \\
&(e_n \otimes \beta \otimes \sigma I_m)\varphi_{1j} + (S \oplus D)\varphi_{2j} = 0 \ \ for\ 2 \le j \le K, \\
&(D - \delta I_n)\varphi_3 = 0.
\end{aligned}$$

subject to normalizing condition

$$\varphi_0 e_{n_1m} + \sum_{j=1}^{K} \varphi_{1j} e_{nm} + \sum_{j=1}^{K} \varphi_{2j} e_{n_2m} + \varphi_3 e_m = 1.$$

The stability condition $\varphi A_0 e < \varphi A_2 e$ is obtained after some algebraic manipulation, which turns out to be

$$\varphi_0(e_{n_1} \otimes D_1 e_m) + \sum_{j=1}^{K} \varphi_{1j}(e_n \otimes D_1 e_m) + \sum_{j=1}^{K} \varphi_{2j}(e_{n_2} \otimes D_1 e_m) + \varphi_3 D_1 e_m < \sum_{j=1}^{K} \varphi_{1j}(\theta_j T^0 \otimes e_m)$$

4.2. Analysis of Steady-State Probability Vector

Consider the steady-state probability vector x of Q and it is divided into $x = (x_0, x_1, x_2, \dots)$. x_0 has a dimension $2m + n_1 m$ while $x_1,\ x_2, \dots$ have a dimension $n_1 m + Knm + Kn_2 m + m$. Then x satisfied the condition $xQ = 0$ and $xe = 1$.

Furthermore, if the system is stable with the vector x, the following equation provides the remaining sub vectors except for the boundary states.

$$x_q = x_1 R^{q-1},\ q \geq 2$$

where the rate matrix R indicates the minimal non-negative solution of the matrix quadratic equation as $R^2 A_2 + RA_1 + A_0 = 0$, as referred by Neuts [15] and satisfies the relation $RA_2 e = A_0 e$.

The sub vectors of x_0 and x_1 were calculated by solving the subsequent equations.

$$x_0 B_{00} + x_1 B_{10} = 0$$

$$x_0 B_{01} + x_1(A_1 + RA_2) = 0$$

The normalizing condition is subject to

$$x_0 e_{2m+n_1 m} + x_1 (I - R)^{-1} e_{n_1 m + Knm + Kn_2 m + m} = 1$$

As a result, the rate matrix R could be mathematically calculated using crucial procedures in the Latouche algorithm for logarithmic reduction of R [11].

5. Busy period Analysis

- The time between customers entering into an empty system and the system becoming empty again after the first interval can be used to measure a busy period. This is the first passage in the transition from level 1 to 0. Thus, it is the first time returns to level 0, followed by at least one visit to a state at any other level is known as the busy cycle.
- We give an overview of the fundamental period before moving on to the busy period. The QBD process takes into account the first transition time, $q \geq 2$, from level q to level q-1.
- It is necessary to examine each of the cases q = 0, 1 that correspond to the boundary states individually. It should be noted that for each level j with $q \geq 2$, there are $(n_1 m + Lnm + Ln_2 m + m)$ states that correspond. Similarly, when the states are organised in lexicographic order, the state(q, j) at level j signifies that j^{th} state at the level q is mentioned.
- The variable $G_{jj'}(v, x)$ represents the conditional probability that the QBD process, which begins in the state (q, j) at time t=0 and visits the level q-1 but not before time x, can make changes v transition to the left and enter the state (q, j'). Let us first define the joint transform

  $$\tilde{G}_{jj'}(z, s) = \sum_{v=1}^{\infty} z^v \int_0^{\infty} e^{-sx} dG_{jj'}(v, x); |z| \leq 1, Re(s) \geq 0$$

  and the matrix is represented as $\tilde{G}(z, s) = \tilde{G}_{jj'}(z, s)$ [14]then the previously defined matrix $\tilde{G}(z, s)$ satisfied the equation

  $$\tilde{G}(z, s) = z(sI - A_1)^{-1} A_2 + (sI - A_1)^{-1} A_0 \tilde{G}^2(z, s).$$

- The matrix $G = G_{jj'} = \tilde{G}(1,0)$, excluding the boundary states, would be used for the first passage time. If we are already familiar with the matrix R, we can use the results to discover the matrix G

$$G = -(A_1 + RA_2)^{-1}A_2.$$

Or else, the idea of a logarithmic reduction algorithm method [11] could be used to determine the values of the G matrix.

Notations

- $G^{(1,0)}_{jj'}(v,x)$ shows that at time $t = 0$, the conditional probability has been discussed for the first time during the passage from level 1 to level 0.
- $G^{(0,0)}_{jj'}(v,x)$ shows that the conditional probability has been discussed for the return time to level 0.
- $\mathfrak{H}_{1q}$ shows the average first passage time between levels q and q-1, assuming the process is in the state (q, j) at time t=0.
- $\vec{\mathfrak{H}}_1$ identifies the column vector containing the entries $\mathfrak{H}_{1q}$.
- $\mathfrak{H}_{2q}$ shows the average number of customers expected to be served during the first passage time from level q to q-1, assuming that the state's first passage time has already begun (q, j).
- $\vec{\mathfrak{H}}_2$ identifies the column vector containing the entries $\mathfrak{H}_{2q}$.
- $\vec{\mathfrak{H}}^{(1,0)}_1$ shows the average first passage time between level 1 and level 0.
- $\vec{\mathfrak{H}}^{(1,0)}_2$ shows the expected number of services finished during the first passage time from level 1 to level 0.
- $\vec{\mathfrak{H}}^{(0,0)}_1$ shows the initial return time to level 0.
- $\vec{\mathfrak{H}}^{(0,0)}_2$ shows the expected number of services finished between the first return time and level 0.

The following equations, which are given by $\tilde{G}^{(1,0)}(z,s)$ and $\tilde{G}^{(0,0)}(z,s)$, are for the boundary levels 1 and 0 respectively.

$$\tilde{G}^{(1,0)}(z,s) = z(sI - A_1)^{-1}B_{10} + (sI - A_1)^{-1}A_0\tilde{G}(z,s)\tilde{G}^{(1,0)}(z,s),$$
$$\tilde{G}^{(0,0)}(z,s) = (sI - B_{00})^{-1}B_{01}\tilde{G}^{(1,0)}(z,s).$$

The matrices are used to calculate the following instances because G, $\tilde{G}^{(0,0)}(1,0)$ and $\tilde{G}^{(1,0)}(1,0)$ are all stochastic in nature. We can compute the instants as follows:

$$\begin{aligned}
\vec{\mathfrak{H}}_1 &= -\frac{\partial}{\partial s}\tilde{G}(z,s)\Big|_{z=1,s=0} e = -[A_1 + A_0(I+G)]^{-1}e,\\
\vec{\mathfrak{H}}_2 &= \frac{\partial}{\partial z}\tilde{G}(z,s)\Big|_{z=1,s=0} e = -[A_1 + A_0(I+G)]^{-1}A_2 e,\\
\vec{\mathfrak{H}}^{(1,0)}_1 &= -\frac{\partial}{\partial s}\tilde{G}^{(1,0)}(z,s)\Big|_{z=1,s=0} e = -[A_1 + A_0G]^{-1}(A_0\vec{\mathfrak{H}}_1 + e),\\
\vec{\mathfrak{H}}^{(1,0)}_2 &= \frac{\partial}{\partial z}\tilde{G}^{(1,0)}(z,s)\Big|_{z=1,s=0} e = -[A_1 + A_0G]^{-1}(A_0\vec{\mathfrak{H}}_2 + B_{10}e),\\
\vec{\mathfrak{H}}^{(0,0)}_1 &= -\frac{\partial}{\partial s}\tilde{G}^{(0,0)}(z,s)\Big|_{z=1,s=0} e = -B_{00}^{-1}[B_{01}\vec{\mathfrak{H}}^{(1,0)}_1 + e],\\
\vec{\mathfrak{H}}^{(0,0)}_2 &= \frac{\partial}{\partial z}\tilde{G}^{(0,0)}(z,s)\Big|_{z=1,s=0} e = -B_{00}^{-1}[B_{01}\vec{\mathfrak{H}}^{(1,0)}_2].
\end{aligned}$$

6. System Performance Measures

- The average system size

$$E_{system} = \sum_{q=1}^{\infty} q x_q e$$

- Probability of the server is busy

$$P_{busy} = \sum_{q=1}^{\infty} \sum_{l=1}^{K} \sum_{j=1}^{n} \sum_{k=1}^{m} x_{q2ljk}$$

- Probability of the server is in idle

$$P_{idle} = \sum_{k=1}^{m} X_{01k}$$

- Probability of the server is on vacation

$$P_{vac} = \sum_{q=0}^{\infty} \sum_{j_1=1}^{n_1} \sum_{k=1}^{m} x_{q0j_1k}$$

- Probability of the server is breakdown

$$P_{bd} = \sum_{q=1}^{\infty} \sum_{l=1}^{K} \sum_{j_2=1}^{n_2} \sum_{k=1}^{m} x_{q3lj_2k}$$

- Probability of the server is on closedown

$$P_{cd} = \sum_{q=0}^{\infty} \sum_{k=1}^{m} x_{q4k}$$

- The average system size during vacation

$$E_{vac} = \sum_{q=1}^{\infty} \sum_{j_1=1}^{n_1} \sum_{k=1}^{m} q x_{q0j_1k} e_{n_1m}$$

- The average system size of the server is busy

$$E_{busy} = \sum_{q=1}^{\infty} \sum_{l=1}^{K} \sum_{j=1}^{n} \sum_{k=1}^{m} q x_{q2ljk} e_{Knm}$$

- The average system size during breakdown

$$E_{bd} = \sum_{q=1}^{\infty} \sum_{l=1}^{K} \sum_{j_2=1}^{n_2} \sum_{k=1}^{m} q x_{q3lj_2k} e_{Kn_2m}$$

- The average system size when the server is close-down

$$E_{cd} = \sum_{q=1}^{\infty} \sum_{k=1}^{m} q x_{q4k} e_m$$

7. Waiting Time Distribution

The first passage time analysis is used in this section to analyse the distribution of a customer's waiting time when they enter the queueing line. Let W(t) be the waiting time distribution function, which takes into account new customers joining the queue. If the server is idle when a customer arrives, they will get service immediately; otherwise, if the server is busy or on vacation, they will have to wait in a queue to receive service from the server.

Let's look at the absorption time in the state space of a Markov chain, which is given by

$$\bar{\Omega} = (*) \cup \{\bar{0}, \bar{1}, \bar{2}, \dots\}$$

where

$$\bar{0} = \{(0,0,j_1) : 1 \le j_1 \le n_1\} \cup \{(0,4)\}$$

and for $q \ge 1$,

$$\bar{q} = \{(q,0,j_1) : 1 \le j_1 \le n_1\} \cup \{(q,2,l,j) : 1 \le l \le K, 1 \le j \le n\}$$
$$\cup \{(q,3,l,j_2) : 1 \le l \le K, 1 \le j_2 \le n_2\} \cup \{(q,4)\}$$

The state space $(*)$obtained by considering the states that have the server in the idle state at the instant of arrival is as below

$$(*) = \{(0,1)\}$$

Let this Markov process's transition matrix $\bar{Q}$ be

$$\bar{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots \\ J_0 & L_0 & 0 & 0 & 0 & 0 & \dots & \dots \\ J_1 & L_{10} & L & 0 & 0 & 0 & \dots & \dots \\ 0 & 0 & L_2 & L & 0 & 0 & \dots & \dots \\ 0 & 0 & 0 & L_2 & L & 0 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \end{bmatrix}$$

where

$$J_0 = \begin{bmatrix} V^0 \\ 0 \end{bmatrix}, \quad L_0 = \begin{bmatrix} V & 0 \\ \gamma \otimes \delta & -\delta \end{bmatrix}, \quad J_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad L_{10} = \begin{bmatrix} 0 & 0 \\ 0 & \theta \otimes T^0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$L = \begin{bmatrix} V & e_1'(K) \otimes qV^0\alpha & e_1'(K) \otimes pV^0\beta & 0 \\ 0 & \Delta(\theta) \otimes T - \sigma I_{Kn} & I_K \otimes (e_n \otimes \sigma\beta) & 0 \\ 0 & I_K \otimes S^0\alpha & I_K \otimes S & 0 \\ \gamma \otimes \delta & 0 & 0 & -\delta \end{bmatrix},$$

$$L_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & L_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad L_{22} = \begin{bmatrix} 0 & \theta_1 T^0\alpha & 0 & \dots & 0 \\ 0 & 0 & \theta_2 T^0\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 & \theta_{K-1} T^0\alpha \\ \theta_K T^0\alpha & 0 & \dots & 0 & 0 \end{bmatrix}.$$

With the aim of determining the arriving tagged customer's waiting time distribution W(t), where($t \ge 0$). To start, we search for the system size probability vector at the arrival epoch in a steady state and it is indicated by $Z(0) = (Z_0(0), Z_1(0), Z_2(0), \dots)$. The vector $Z_0(0)$ may be further partitioned as follows $Z_0(0) = (Z_{00}, Z_{04})$. The system size probability vector at the arrival epoch in the steady state is as follows because the arrival process abides by the Markovian property:

$$Z_{00} = x_{00}\left[I_{n_1} \otimes \frac{D_1 e_m}{\lambda}\right], \quad Z_{04} = x_{04}\left[\frac{D_1 e_m}{\lambda}\right],$$

$$Z_q(0) = x_q\left[I_{n_1+Kn+Kn_2+1} \otimes \frac{D_1 e_m}{\lambda}\right], \quad for\ q \ge 1$$

where λ denotes the fundamental arrival rate of Markovian Arrival Process.
Define $Z(t) = (Z_*(t), Z_0(t), Z_1(t), \dots)$,
where
$Z_q(t),\ q \ge 1$ - vector of order $1 \times (n_1 + Kn + Kn_2 + 1)$

$Z_0(t) = (Z_{00}, Z_{04})$ - vector of order $1 \times (n_1 + 1)$
and their components give the probability that at epoch t, the CTMC generator matrix is $\bar{Q}$, will be in the appropriate level q state. Since the tagged customer's probability of being in the absorbing state at epoch t is specified by $Z_*(t)$, we get $W(t) = Z_*(t)$, where $t \geq 0$.
The differential equation $Z'(t) = Z(t)\bar{Q}$ for $t \geq 0$ becomes

$$\begin{aligned} Z'_*(t) &= Z_0(t)J_0 + Z_1(t)J_1, \\ Z'_0(t) &= Z_0(t)L_0 + Z_1(t)L_{10}, \\ Z'_q(t) &= Z_q(t)L + Z_{q+1}(t)L_2,\ q \geq 1 \end{aligned}$$

where $'$ specifies the derivative concerning t. Let's use the method suggested by Neuts et al. [15] to compute the LST for W(t). The row vector $\omega(s)$ specifies the Laplace-Stieltjes Transform (LST) of the first passage time to level 1 by starting the process at state q and using $Z_q(0),\ q \geq 1$ as the initial probability vector. Neuts et al. [15] state that we get,

$$\omega(s) = \sum_{q=1}^{\infty} Z_q(0)[(sI - L)^{-1}L_2]^{q-1} \tag{1}$$

With the restriction that the process begins at level q = 0, 1, let the LST of the time to become absorbed into the state $(*)$ be specified by $\phi(q, s)$. Similar to Neuts et al. [15], we have

$$(0, s) = [sI - L_0]^{-1}J_0, \tag{2}$$

$$\phi(1, s) = [sI - L]^{-1}L_{10}\phi(0, s) + [sI - L]^{-1}J_1. \tag{3}$$

This allows us to quickly note that the LST for the distribution of sojourn time is as follows.

$$W(s) = Z_0(0)\phi(0, s) + \omega(s)\phi(1, s) \tag{4}$$

Expected Waiting Time

The mean waiting time is given as

$$E(W) = -W'(0) = -Z_0(0)\phi'(0, 0) - \omega'(0)e_{n_1+Kn+Kn_2+1} - \omega(0)\phi'(1, 0) \tag{5}$$

The initial term of the previous equation gives the expected time to reach the absorbing state $(*)$, assuming that the system is at level 0. The final two components of the previous equation provide the expected time for accessing the absorbing state $(*)$ if the system is resting at level $q \geq 1$. By differentiating (2) and (3) and making s=0,

$$\Phi'(0, 0) = -[-L_0]^{-2}J_0 \tag{6}$$

$$\Phi'(1, 0) = -[-L]^{-2}L_{10}\Phi(0, 0) + [-L]^{-1}L_{10}\Phi'(0, 0) - [-L]^{-2}J_1 \tag{7}$$

Using (6) and the vector $Z(0) = (Z_0(0), Z_1(0), \dots)$, it is simple to calculate the first term of (5).
From (1), we get

$$\omega(0) = \sum_{q=1}^{\infty} Z_q(0)M^{q-1} \tag{8}$$

where $M = [-L]^{-1}L_2$. As M is a stochastic matrix, we get

$$\omega(0)e_{n_1+Kn+Kn_2+1} = 1 - Z_0(0)e_{n_1+1} \tag{9}$$

Equations (7) and (8), as well as the vector $Z(0) = (Z_0(0), Z_1(0), \dots)$, allow us to quickly calculate the last term of (5).
We obtain by differentiating (1) and setting s=0,

$$\omega'(0) = (-1)\sum_{q=1}^{\infty} Z_{1+q}(0)\sum_{j=0}^{q-1} M^j[-L]^{-1}M^{q-j}. \tag{10}$$

Due to the stochastic nature of matrix M,

$$(-1)\omega'(0)e_{n_1+Kn+Kn_2+1} = \sum_{q=1}^{\infty} Z_{1+q}(0) \sum_{j=0}^{q-1} M^j [-L]^{-1} e_{n_1+Kn+Kn_2+1}. \tag{11}$$

Let's assess the value of $(-1)\omega'(0)e_{n_1+Kn+Kn_2+1}$ using the technique described in Neuts et al. [15] and Kao et al. [8]. We start by building a matrix M_2 that is generalised inverse of I-M and stochastic, with $I - M + M_2$ being non-singular and M_2 being stochastic. The matrix M_2 can be viewed as $M_2 = e_{n_1+Kn+Kn_2+1}m_0$, where m_0 is the invariant probability vector of M. Additionally, using the property $MM_2 = M_2M = M_2$, we have

$$\sum_{j=0}^{q-1} M^j (I - M + M_2) = I - M^q + qM_2 \quad for\ q \geq 1. \tag{12}$$

By using (12) in (11), we obtain

$$\begin{aligned}(-1)\omega'(0)e_{n_1+Kn+Kn_2+1} = &\left\{ x_1(I-R)^{-1}\left[I_{n_1+Kn+Kn_2+1} \otimes \frac{D_1 e_m}{\lambda}\right] - \omega(0) \right.\\ &\left. + x_1 R (I-R)^{-2}\left[I_{n_1+Kn+Kn_2+1} \otimes \frac{D_1 e_m}{\lambda}\right] M_2 \right\}\\ &\times [I - M + M_2]^{-1} [-L]^{-1} e_{n_1+Kn+Kn_2+1}.\end{aligned} \tag{13}$$

Since we have calculated all the terms in (5), we can easily calculate the average waiting time.

8. Cost Analysis

Our model's cost analysis has been created below by assuming the cost elements (per unit time) correspond to distinct measures of the system.

$$\begin{aligned}TC = &C_H E_{system} + C_{busy}P_{busy} + C_{idle}P_{idle} + C_{vac}P_{vac} + C_{bd}P_{bd} + C_{cd}P_{cd}\\ &+ \sum_{i=1}^{K} C_{1i}\theta_i \xi + \sigma C_2 + \zeta C_3 + \delta C_4\end{aligned}$$

where

- TC - Total cost per unit time
- C_H - Each customer's holding cost in the system
- C_{busy} - Cost acquired by the system during server being busy
- C_{idle} - Cost acquired due to server being idle
- C_{vac} - Cost acquired during server's vacation period
- C_{bd} - Cost acquired by the server during breakdown time
- C_{cd} - Cost acquired by the server during close-down process
- C_{1i} - Cost acquired by the server for offering i^{th} type service, $i = 1, 2, \ldots, K$
- C_2 - Cost acquired when the server caused by breakdowns
- C_3 - Cost acquired in carrying out the repair process
- C_4 - Cost acquired in carrying out the close-down process

9. Numerical

In this section, we are using numerical and graphical representations to analyze model behavior. The mean value of the subsequent five different MAP representations is 1, which is the same for all the various arrival processes. In published studies, these five sets of arrival values have been used as input data (see Chakravarthy [5]).

- **Arrival in Erlang(ERL-A):**

$$D_0 = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

- **Arrival in Exponential(EXP-A):**

$$D_0 = \begin{bmatrix} -1 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 1 \end{bmatrix}$$

- **Arrival in Hyper-exponential(HEX-A):**

$$D_0 = \begin{bmatrix} -1.90 & 0 \\ 0 & -0.19 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{bmatrix}$$

- **Arrival in MAP-Negative Correlation(MAPNC-A):**

$$D_0 = \begin{bmatrix} -1.25 & 1.25 & 0 \\ 0 & -1.25 & 0 \\ 0 & 0 & -2.5 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0.0125 & 0 & 1.2375 \\ 2.4750 & 0 & 0.0250 \end{bmatrix}$$

- **Arrival in MAP-Positive Correlation(MAPPC-A):**

$$D_0 = \begin{bmatrix} -1.25 & 1.25 & 0 \\ 0 & -1.25 & 0 \\ 0 & 0 & -2.5 \end{bmatrix}, \qquad D_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1.2375 & 0 & 0.0125 \\ 0.0250 & 0 & 2.4750 \end{bmatrix}.$$

Let's think about the service, repair, and vacation processes as three phase type distributions. In the literature, these sets of service, vacation, and repair values have been used as input data [5].

- **Service in Erlang(ERL-S):**

$$\alpha = (1, 0), \qquad T = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

- **Repair in Erlang(ERL-R):**

$$\beta = (1, 0), \qquad S = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

- **Vacation in Erlang(ERL-V):**

$$\gamma = (1, 0), \qquad V = \begin{bmatrix} -2 & 2 \\ 0 & -2 \end{bmatrix}$$

- **Service in Exponential(EXP-S):**

$$\alpha = [-1], \qquad T = [1]$$

- **Repair in Exponential(EXP-R):**

$$\beta = [-1], \qquad S = [1]$$

- **Vacation in Exponential(EXP-V):**

$$\gamma = [-1], \qquad V = [1]$$

- **Service in Hyper-exponential(HEX-S):**

$$\alpha = (0.8, 0.2), \qquad T = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

- **Repair in Hyper-exponential(HEX-R):**

$$\beta = (0.8, 0.2), \qquad S = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

- **Vacation in Hyper-exponential(HEX-V):**

$$\gamma = (0.8, 0.2), \qquad V = \begin{bmatrix} -2.8 & 0 \\ 0 & -0.28 \end{bmatrix}$$

9.1. Illustration 1

We investigated the consequence of the repair rate (ζ) on the average system size(E_{system}). We fix $\lambda = 2,\ \xi = 6,\ \eta = 10,\ \sigma = 1,\ \delta = 5,\ K = 10,\ \theta^t = [1, 0.97, 0.93, 0.9, 0.87, 0.83, 0.8, 0.75, 0.7, 0.6]$, $p = 0.6,\ q = 0.4$.

Table 1: *Repair rate (ζ) vs E_{system} - ERL-S*

ERL-S					
ζ	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
4	1.240013	1.428884	2.424636	1.337204	10.61745
5	1.100282	1.263127	2.098389	1.18343	8.801677
6	1.015791	1.16274	1.90369	1.090236	7.692035
7	0.959243	1.095599	1.775197	1.027872	6.946641
8	0.918743	1.047606	1.684428	0.98328	6.412992
9	0.888298	1.011625	1.617078	0.949845	6.012954
10	0.864568	0.983665	1.565214	0.923863	5.702437
11	0.845543	0.961321	1.524099	0.903103	5.454723

Table 2: *Repair rate (ζ) vs E_{system} - EXP-S*

EXP-S					
ζ	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
4	1.300513	1.477625	2.385188	1.39259	8.815389
5	1.14504	1.300952	2.072725	1.226593	7.266135
6	1.05166	1.194414	1.885579	1.126538	6.322526
7	0.989618	1.1235	1.761902	1.059955	5.692052
8	0.945487	1.073041	1.674513	1.012588	5.243369
9	0.912517	1.035369	1.609694	0.977232	4.909026
10	0.886957	1.006202	1.559813	0.949868	4.650976
11	0.866562	0.982972	1.520303	0.928081	4.446203

Table 3: *Repair rate (ζ) vs E_{system} - HEX-S*

HEX-S					
ζ	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
4	1.558351	1.660665	2.248675	1.59979	4.230083
5	1.325099	1.423599	1.946641	1.369357	3.457511
6	1.188053	1.283658	1.765584	1.233688	3.008707
7	1.099269	1.19259	1.646429	1.145591	2.721063
8	1.037696	1.12918	1.562789	1.084356	2.523394
9	0.992794	1.082778	1.501231	1.039609	2.380326
10	0.958759	1.047505	1.454251	1.005634	2.272559
11	0.932162	1.019875	1.417352	0.979044	2.188777

With the help of tables 1, 2 and 3, we can determine that increasing the repair rate reduces the average system size in various arrangement of services and arrivals of ERL-A, EXP-A, HEX-A, MAPNC-A and MAPPC-A. The positive correlation arrival decreases rapidly compared to all other arrivals.

9.2. Illustration 2

We investigated the consequence of the vacation rate (η) on the average waiting time E(W). We fix $\lambda = 2,\ \xi = 6,\ \sigma = 1,\ \zeta = 4,\ \delta = 5,\ K = 5,\ \theta^t = [1, 0.9, 0.8, 0.7, 0.6],\ p = 0.4,\ q = 0.6.$

Table 4: *Vacation rate (η) vs E(W) - ERL-S*

ERL-S					
η	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
10	0.513310803	0.652469261	1.316678245	0.597597919	6.305631788
11	0.502141873	0.641535425	1.304722472	0.586775589	6.294830579
12	0.493006347	0.632560036	1.294845632	0.577894144	6.285940644
13	0.485401455	0.625064256	1.286552181	0.570478428	6.27849855
14	0.478976132	0.618712658	1.279491771	0.564195772	6.272178893
15	0.4734782	0.613263522	1.273409875	0.55880656	6.266746629
16	0.46872206	0.608538394	1.268117234	0.554133955	6.262027841
17	0.46456821	0.604402741	1.263470256	0.550044688	6.257891127
18	0.460909846	0.600753272	1.259358071	0.546436465	6.254235397
19	0.457663866	0.597509402	1.255693735	0.543229496	6.250981623

Table 5: *Vacation rate (η) vs E(W) - EXP-S*

EXP-S					
η	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
10	0.557188985	0.69195187	1.314692759	0.63940605	5.40504825
11	0.545405475	0.680428504	1.302062559	0.627990745	5.393620067
12	0.535811233	0.671007926	1.291662266	0.618660721	5.384251971
13	0.527857453	0.663169128	1.282954639	0.610898654	5.376437568
14	0.521162512	0.656548669	1.27556107	0.604343883	5.369822843
15	0.515453367	0.650885673	1.269207338	0.598737649	5.364153128
16	0.510529823	0.645988247	1.263690113	0.593889693	5.359240685
17	0.506241951	0.641712219	1.258855539	0.589657117	5.354944174
18	0.502475357	0.637947256	1.254585115	0.585930588	5.351155205
19	0.49914132	0.634607531	1.250786132	0.58262508	5.347789301

Table 6: *Vacation rate (η) vs E(W) - HEX-S*

HEX-S					
η	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
10	0.765770257	0.873794399	1.367546979	0.829456953	2.90799282
11	0.749707796	0.858258947	1.350700199	0.814080592	2.891938923
12	0.736875156	0.84579312	1.337039635	0.801751095	2.879049871
13	0.726421715	0.835595592	1.325762738	0.791670904	2.868502654
14	0.717764748	0.827116374	1.316311418	0.783293156	2.859731284
15	0.710493246	0.819966703	1.30828648	0.77623166	2.852335034
16	0.704309955	0.813864754	1.301395404	0.770206698	2.846023015
17	0.698995159	0.808601698	1.29541931	0.765011173	2.840579451
18	0.694383327	0.804019799	1.290191394	0.760488778	2.835841216
19	0.690347624	0.799997883	1.285582472	0.756519514	2.831682906

With the help of tables 4, 5 and 6, we can determine that increasing the vacation rate reduces the average waiting time in various arrangement of services and arrivals of ERL-A, EXP-A, HEX-A, MAPNC-A and MAPPC-A.

9.3. Illustration 3

We examined the consequence of the vacation rate(η) on the Total cost(TC) of the system. We fix $\lambda = 2,\ \xi = 6,\ \zeta = 4,\ \sigma = 1,\ \delta = 5,\ K = 5,\ \theta^t = [1, 0.9, 0.8, 0.7, 0.6],\ p = 0.6,\ q = 0.4,\ C_H = 10,$ $C_{vac} = 2,\ C_{idle} = 1,\ C_{busy} = 4,\ C_{bd} = 2,\ C_{cd} = 2, C_{11} = 3,\ C_{12} = 2.9,\ C_{13} = 2.7,\ C_{14} = 2.5,$ $C_{15} = 2.2,\ C_2 = 1,\ C_3 = 2,\ C_4 = 2.$

Table 7: *Vacation rate (η) vs TC - ERL-S*

ERL-S					
η	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
10	99.68315738	102.0191734	114.5649299	100.8662306	220.3575308
11	99.58654131	101.9327141	114.4864353	100.7816754	220.2740852
12	99.50790684	101.8624352	114.422411	100.7129863	220.2062508
13	99.44275241	101.8042608	114.369262	100.6561557	220.150093
14	99.38794402	101.7553616	114.3244797	100.6084052	220.1028821
15	99.34123751	101.7137163	114.2862631	100.5677517	220.0626686
16	99.30098652	101.677845	114.253288	100.5327444	220.0280254
17	99.26595761	101.6466406	114.2245603	100.5022988	219.9978848
18	99.23520954	101.6192592	114.1993197	100.4755886	219.971433
19	99.20801222	101.595047	114.1769756	100.4519741	219.9480397

Table 8: *Vacation rate (η) vs TC - EXP-S*

EXP-S					
η	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
10	100.2339903	102.3592967	113.45068	101.3175747	196.2037767
11	100.1280443	102.2634003	113.3608307	101.2234308	196.1105677
12	100.0424323	102.1859477	113.28795	101.1474223	196.0352852
13	99.97195226	102.122204	113.2277544	101.0848836	195.973326
14	99.91300867	102.0689027	113.1772692	101.0325998	195.9215153
15	99.86304362	102.0237237	113.134369	100.9882889	195.8775984
16	99.8201921	101.9849774	113.0974976	100.9502904	195.8399333
17	99.78306479	101.9514061	113.0654919	100.9173687	195.8072978
18	99.75060709	101.9220556	113.0374655	100.8885871	195.7787647
19	99.72200511	101.8961899	113.0127327	100.8632232	195.7536187

Table 9: *Vacation rate (η) vs TC - HEX-S*

HEX-S					
η	ERL-A	EXP-A	HEX-A	MAPNC-A	MAPPC-A
10	102.3258123	103.4719568	110.0500597	102.7836354	135.054583
11	102.1616111	103.3188195	109.8998592	102.6325837	134.9021191
12	102.0324917	103.1981907	109.7806956	102.5136915	134.7822937
13	101.9288319	103.1011839	109.684308	102.4181401	134.6861809
14	101.8441297	103.0217915	109.6050482	102.3399749	134.6077393
15	101.7738585	102.9558273	109.5389409	102.2750516	134.5427562
16	101.7147844	102.9002977	109.4831181	102.2204095	134.4882179
17	101.6645447	102.8530131	109.4354652	102.1738855	134.4419202
18	101.621379	102.8123398	109.3943937	102.1338671	134.402219
19	101.5839523	102.7770377	109.3586901	102.0991312	134.3678669

With the help of tables 7, 8 and 9, we can determine that increasing the vacation rate reduces the total cost of the system in various arrangement of services and arrivals of ERL-A, EXP-A, HEX-A, MAPNC-A and MAPPC-A.

9.4. Illustration 4

We investigated the consequence of the breakdown rate (σ) on the average system size(E_{system}). We fix $\lambda = 2,\ \xi = 6,\ \eta = 10,\ \zeta = 4,\ \delta = 5,\ K = 10,\ \theta^t = [1, 0.97, 0.93, 0.9, 0.87, 0.83, 0.8, 0.75, 0.7, 0.6],\ p = 0.6,\ q = 0.4$.

With the help of figures 2, 3, 4, 5 and 6, we analyze the breakdown rate versus the average system size with the combination of arrival and service time groupings. The breakdown rate increases then the corresponding average system size is also increases rapidly in Erlang services and, increases gradually in Exponential services and slowly in Hyper-exponential services but in case of MAP positive correlation arrival increases rapidly than compared to all other arrivals.

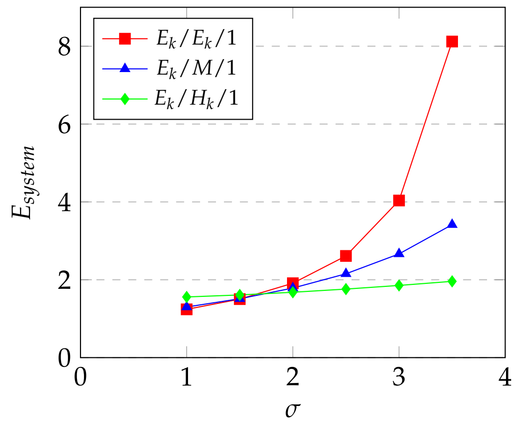


Figure 2: *Breakdown rate(σ) vs E_{system} - ERL-A*

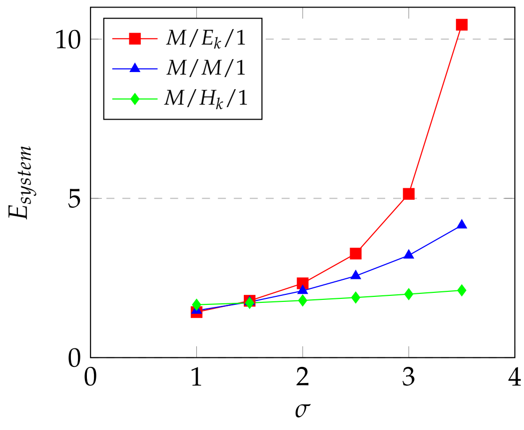


Figure 3: *Breakdown rate(σ) vs E_{system} - EXP-A*

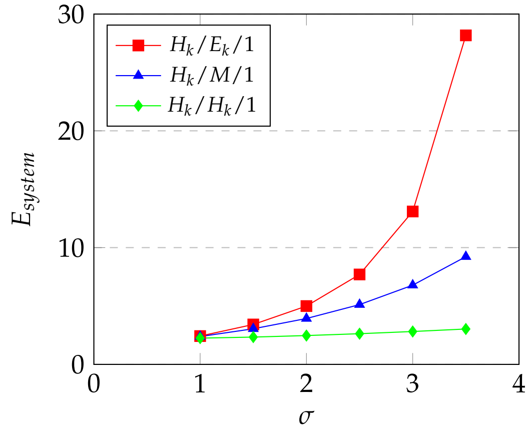


Figure 4: *Breakdown rate(σ) vs E_{system} - HEX-A*

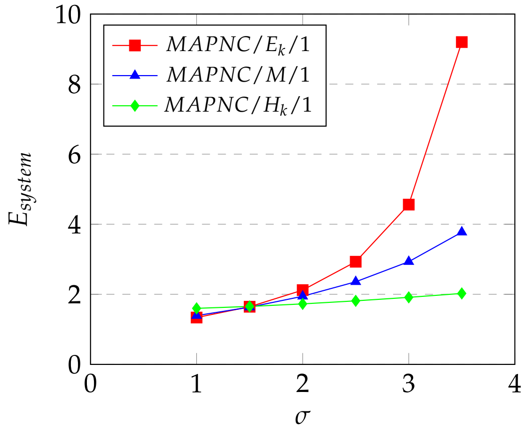


Figure 5: *Breakdown rate(σ) vs E_{system} - MAPNC-A*

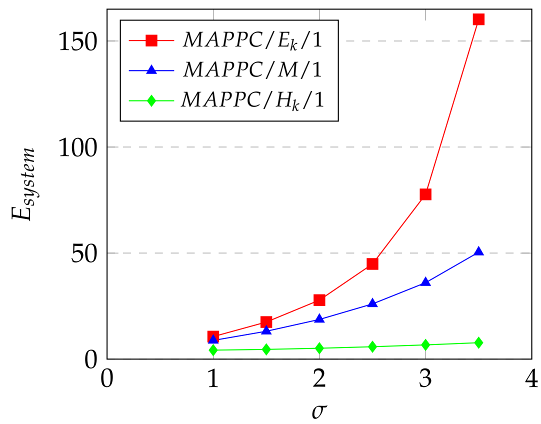


Figure 6: *Breakdown rate(σ) vs E_{system} - MAPPC-A*

9.5. Illustration 5

We have examined both the vacation rate(η) and repair rate(ζ) against the average system size(E_{system}). We fix $\lambda = 2,\ \xi = 6,\ \sigma = 1,\ \delta = 5,\ K = 10,\ \theta^t = [1, 0.97, 0.93, 0.9, 0.87, 0.83, 0.8, 0.75, 0.7, 0.6],\ p = 0.6,\ q = 0.4.$

With the help of figures 7 to 11, we analyze the both vacation rate and repair rate versus the average system size with the combination of arrival and service time groupings. Both the vacation rate and repair rate increases then the corresponding average system size is decreases rapidly in MAP positive correlation compared to all other arrivals.

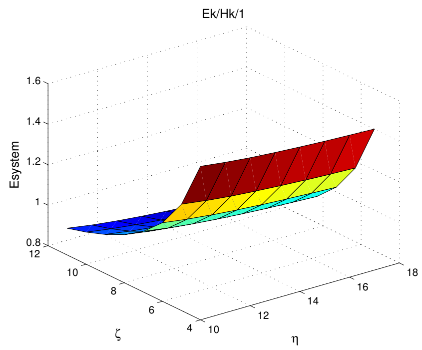


Figure 7: *$E_k/H_k/1$ - Vacation rate(η) and Repair rate(ζ) vs E_{system}*

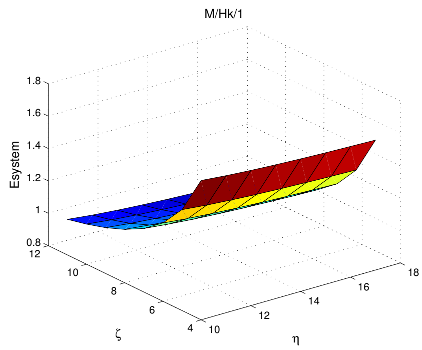


Figure 8: *$M/H_k/1$ - Vacation rate(η) and Repair rate(ζ) vs E_{system}*

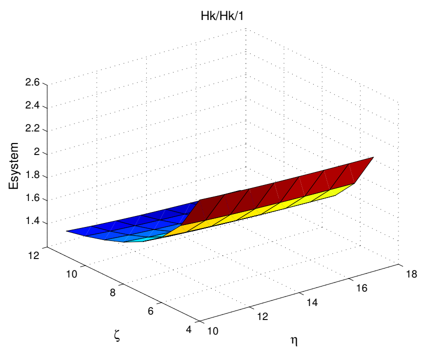


Figure 9: *$H_k/H_k/1$ - Vacation rate(η) and Repair rate(ζ) vs E_{system}*

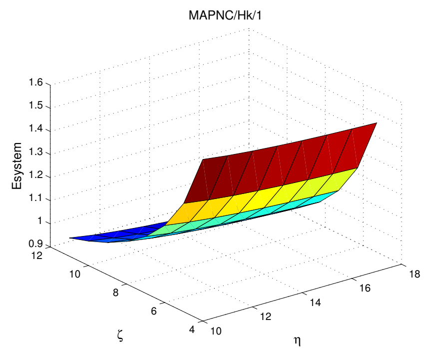


Figure 10: *$MAPNC/H_k/1$ - Vacation rate(η) and Repair rate(ζ) vs E_{system}*

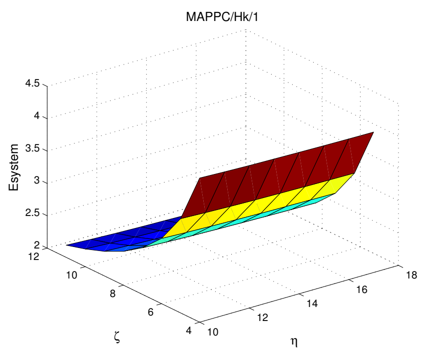


Figure 11: *$MAPPC/H_k/1$ - Vacation rate(η) and Repair rate(ζ) vs E_{system}*

10. Conclusion

In our paper, customers arrive in a Markovian Arrival Process and the service process follows a phase-type distribution with degrading service, server breakdown, vacation process in phase type distribution, repair process in phase type distribution, starting failure and close-down. We also perform the busy period analysis, waiting time distribution and cost analysis in our work. Using numerical values of arrival and service times, we tabulated the repair rate versus expected system size and the vacation rate versus the expected waiting time numerically. We compared the breakdown rate to the expected system size, as well as the vacation and repair rates to the expected system size, as shown by the graphical demonstrations.

References

[1] Ayyappan, G. and Deepa, T. (2018). Analysis of batch arrival bulk service queue with multiple vacation closedown essential and optional repair, *Applications and Applied Mathematics*, 13(2):578–598.

[2] Ayyappan, G. and Nirmala, M. (2018). An M[X]/G(a, b)/1 queue with breakdown and delay time to two phase repair under multiple vacation, *Applications and Applied Mathematics*, 13(2):639–663.

[3] Ayyappan, G. and Thilagavathy, K. (2020). Analysis of MAP/PH/1 Queueing Model with Breakdown, Instantaneous Feedback and Server Vacation, *Applications and Applied Mathematics*, 15(2):673–707.

[4] Ayyappan, G. and Gowthami, R. (2021). A MAP/PH/1 Queue with Setup time, Bernoulli vacation, Reneging, Balking, Bernoulli feedback, Breakdown and repair, *Reliability: Theory and Applications*, 16(2):191–221.

[5] Chakravarthy, S. R. (2010). Markovian arrival process, Wiley Encyclopaedia of Operations Research and Management Science.

[6] Choudhary, A., Chakravarthy, S. R. and Sharma, D. C. (2021). Analysis of MAP/PH/1 Queueing System with Degrading Service Rate and Phase Type Vacation, *Mathematics*, 9, 2387.

[7] Doshi, B. T. (1986). Queueing System with vacations-A Survey, *Queueing Systems*, 1:29–66.

[8] Kao, E. P. C. and Narayanan, K. S. (1991). Analysis of an M/M/N queue with server's vacations, *European Journal of Operational Research*, 54(2):256–266.

[9] Karpagam, S., Ayyappan, G. and Somasundaram, B. (2020). A Bulk Queueing System with Rework in Manufacturing Industry with Starting Failure and Single Vacation, *International Journal of Applied and Computational Mathematics*, 6(6):1-22.

[10] Krishna Kumar, B., Rukumani, R. and Thangaraj, V. (2008). Analysis of MAP/P H(1), P H(2)/2 queue with Bernoulli vacations, *Journal of Applied Mathematics and Stochastic Analysis*, Article ID: 396871, 1-20.

[11] Latouche, G. and Ramaswami, V. Introduction of Matrix Analytic Methods in Stochastic Modeling, Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[12] Li, J. and Tian, N. (2007). The M/M/1 queue with working vacations and vacation interruptions, *The M/M/1 queue with working vacations and vacation interruptions*, 16:121–127.

[13] Neuts M. F. (1979). A Versatile Markovian point process *Journal of Applied Probability*, 16:764–779.

[14] Neuts M. F. Matri geometric Solutions in Stochastic Models: an algorithmic approach. John Hopkins Series in Mathematical Sciences, John Hopkins University Press, Baltimore, Md, USA, 1981.

[15] Neuts, M. F. and Lucantoni, D. M. (1979). A Markovian Queue with N Servers Subject toBreakdowns and Repairs, *Management Science*, 25(9):849–861.

[16] Qi-Ming He. Fundamentals of Matrix - Analytic Methods, Springer, New York, 2004.

[17] Senthil Vadivu, A. and Arumuganathan, R. (2015). Cost Analysis of MAP/G(a,b)/1/N Queue with Multiple Vacations and Closedown Times, *Quality Technology and Quantitative Management*, 12(4):605–626.

[18] Sreenivasan, C., Chakravarthy, S.R. and Krishnamoorthy, A. (2013). A MAP/PH/1 queue with working vacations, vacation interruptions and N-Policy, *Applied Mathematical Modelling*,37:3879–3893.

[19] Steeb, W. H. and Hardy, Y. Matrix Calculus and Kronecker Product: A Practical Approach to Linear and Multilinear Algebra; World Scientific Publishing: Singapore, 2011.

[20] Tian, N. and Zhang, Z. G. Vacation Queueing Models: Theory and Applications; Springer Publishers: New York, NY, USA, 2006.

[21] Wang, K. H. Chan, M. C. and Ke, J. C. (2007). Maximum entropy analysis of the M[x]/M/1 queueing system with multiple vacations and server breakdowns, *Computers and Industrial Engineering*, 52(2):192–202.

[22] Yang, D. Y. Ke, J. C. and Wu, C. H. (2014). The multi-server retrial system with Bernoulli feedback and starting failures, *International Journal of Computer Mathematics*, 92(5):954–969.

COMPARISON OF MAXIMUM LIKELIHOOD ESTIMATION AND BAYESIAN ESTIMATION ON EXPONENTIATED POWER LOMAX DISTRIBUTION

S. Amirtha Rani Jagulin[1] and A. Venmani[2]

•

[1,2] Department of Mathematics and Statistics, College of Science and Humanities,
SRM Institute of Science and Technology,
Kattankulathur, Tamil Nadu, India
as4371@srmist.edu.in[1], venmania@srmist.edu.in[2]

Abstract

The aim of the paper is to apply the Bayesian estimation under squared error loss function to the Exponentiated Power Lomax (EPOLO) distribution to estimate the parameters and then compare with maximum likelihood estimation. The reliability of the distribution is analyzed by computing survival and hazard function for the exponentiated power lomax distribution. The mean square error will help to compare the different estimates like Bayesian and maximum likelihood estimation to decide the best one. Bayesian estimation and the maximum likelihood estimation are discussed for the distribution. A simulation study is done using the R programming software to generate random values and estimate the parameters for different n = 15, 30, 50, 100 and parameter values taken as 0.5 and 1.5. At n=100 and c= 0.5 the mean square error of bayes and mle are same and survival and hazard function mean square error are decreases when the sample size increases, this indicates the distribution is fitted good in all areas.

Keywords: Exponentiated Power Lomax (EPOLO) distribution, Bayesian estimation, Reliability, Mean Square Error (MSE), Monte Carlo simulation.

1. Introduction

According to El-Monsef [12], the EPOLO distribution performs superior fits than several well-known distributions to the data such as the number of ball bearing revolutions, the lung cancer patient's tumor size, and the total COVID-19 deaths in Egypt. The EPOLO (Exponentiated Power Lomax) distribution is an extension of the POLO (Power Lomax) distribution. By exponentiating the POLO cumulative distribution function to positive real power c we get this distribution. The CDF of the EPOLO distribution is given by

$$F(x) = (1 - \gamma^{\alpha}\left(\gamma + x^{\beta}\right) - \alpha)^{c} \quad x > 0; \alpha, \beta, \gamma, c > 0 \tag{1}$$

The PDF of EPOLO distribution is given by

$$f(x) = \alpha\beta c\gamma^{\alpha}x^{\beta-1}\left(\gamma + x^{\beta}\right)^{-\alpha-1}(1 - \gamma^{\alpha}(\gamma + x^{\beta})^{-\propto})^{c-1} \quad x > 0;\ \alpha, \beta, \gamma, c > 0 \tag{2}$$

When $c=\beta=1$, it will result to Lomax distribution and if $c=1$, then (2) will leads to POLO distribution. Lomax [10] has proposed by Lomax distribution is a mixture of gamma and exponential distributions. It is also denoted as Pareto II distribution. Applying the Lomax distribution to simulate lifetime data was common. It is helpful for income modeling, business failure data, and biological issues. The random variable X follows a L(α,γ) distribution, the cumulative distribution function (CDF) is given as

$$F(x) = 1 - (\frac{x}{\gamma} + 1)^{-\alpha} \qquad x > 0;\ \alpha, \gamma > 0 \tag{3}$$

The probability density function (PDF) is given as

$$f(x) = \frac{\alpha}{\gamma}\left(\frac{x}{\gamma} + 1\right)^{-\alpha-1} x > 0;\ \alpha, \gamma > 0 \tag{4}$$

Lingappaiah [9] created different ways of estimation procedures for Lomax distribution. Myhre and Saunders [13] were described right censored data were subjected to Lomax distribution. According to Hassan [5], Lomax distribution can be utilized for dependability modeling and life testing. Balakrishnan and Ahsanullah [3] investigated the record moments and distribution properties of the Lomax distribution. Kilany [7] discussed the weighted Lomax distribution. Lemonte and Cordeiro [8] used the Lomax distribution to study its expansion. This paper is divided as follows: section 2 provides the information about the Bayesian analysis with prior, likelihood, posterior function and also the parameter estimation of the parameter c. Section 3, derivation of posterior of the maximum likelihood function using the likelihood function, section 4 provides the knowledge about the reliability measures, survival and hazard function. Section 5, describes the quantile function to generate random numbers. In section 6 and 7 a simulation study is applied and analyzed for EPOLO distribution, section 8 gives the conclusion about the result.

2. Bayesian estimation

The Reverend Thomas Bayes (1701–1761), who used a subjective method to quantify probability, is the topic of the philosophy known as Bayesian. The most famous accomplishment of Bayes was never published. Richard Price updated his notes and published them after his death (1763). The statistician using Bayesian, he is free to interpret probability as a frequency and a level of belief or a function that calculated with the mathematical principles of probability, depending on which interpretation best fits the task. The Bayesian technique gives the option of adding earlier knowledge of the pertinent parameters. Mahmoud et al. [11] and Tierney [17] introduces the Bayesian concept and describes how it is computationally implemented using MCMC algorithms. Amal s.Hassan et al [2] and Shrestha and Kumar [15] analyzed Bayesian estimation takes parameters as probabilistic variables with random variables. The Bayesian is helpful in analysis since it can take the prior knowledge into account.
In this EPOLO distribution, we have four parameters, but estimating the four parameters analytically is not possible, so we have estimated the shape parameter c alone here for the better understanding of Bayesian analysis by Hesham and Soha [6]. We assume the prior for the parameter c as a gamma distribution with the pdf.

$$\pi(c) \propto c^{a-1} e^{-bc}\, c > 0, \qquad a, b > 0 \tag{5}$$

The likelihood functions of EPOLO for the random samples X_1, X_2… X_n of independent and identically distributed random variables is given by

$$L\,(\alpha,\beta,\lambda,c) = (\alpha\beta c\gamma^{\alpha})^{n}\prod_{i=1}^{n} x^{\beta-1}\,\left(\gamma+x^{\beta}\right)^{-\alpha-1}\,\left(1-\gamma^{\alpha}\left(\gamma+x^{\beta}\right)^{-\alpha}\right)^{c-1}\; x > 0;\quad \alpha,\beta,\gamma,c > 0 \tag{6}$$

The likelihood only for the shape parameter c is taken as

$$L(\frac{x}{c}) = c^{n}\prod_{i=1}^{n}(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c} \tag{7}$$

Using the bayes theorem the posterior is given by

$$\pi(\frac{c}{x}) = \frac{L(\frac{x}{c})\,\pi(c)}{\int_0^{\infty} L(\frac{x}{c})\,\pi(c)dc} \tag{8}$$

The posterior density of the shape parameter c can be obtained as

$$\pi(\frac{c}{x}) = \frac{[b-w]^{n+a}}{\Gamma(n+a)}\, c^{n+a-1}exp\{-c(b\, -\, w)\} \tag{9}$$

Where $w = \sum_{i=1}^{n} log(1-\gamma^{\alpha}\,(\gamma+x_i^{\beta})^{-\alpha}$

The Bayes estimate of $\hat{c}$, Sowbhagya [16] and Nada S. Karam [14] provides a knowledge on the squared error loss function, is derived as

$$\hat{c} = \int_{-\infty}^{\infty} c\, f(\frac{c}{x})dc \tag{10}$$

$$\hat{c} = \frac{n+a}{b+w} \tag{11}$$

3. Maximum likelihood estimation

The likelihood function is obtained by applying the product to the density function of the given distribution over i = 1 to n. That is,

$$L(\alpha,\beta,\gamma,c) = (\alpha\beta c\gamma^{\alpha})^{n}\prod_{i=1}^{n} x^{\beta-1}\,(\gamma+x^{\beta})^{-\alpha-1}\,(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c-1} \tag{12}$$

The log-likelihood function for the parameters α, β, λ, and c is

$$ln\, L(\alpha,\beta,\gamma,c) = n\, log\,(\alpha\beta c\,\gamma^{\alpha}) + (c - 1)\sum_{i=1}^{n}\log\,(1-\gamma^{\alpha}\left(\gamma+x_i^{\beta}\right)^{-\alpha}) - (\alpha + 1)\sum_{i=1}^{n}\log\,(x_i^{\beta} + \gamma) + (\beta - 1)\sum_{i=1}^{n} log\,(x_i) \tag{13}$$

To find the maximum likelihood estimator of the parameter $\hat{c}$ we equate the log-likelihood to zero, then we get,

$$n\,c + \sum_{i=1}^{n} log(1-\gamma^{\alpha}\,(\gamma+x_i^{\beta})^{-\alpha} = 0 \tag{14}$$

Taking $w = \sum_{i=1}^{n} log(1-\gamma^{\alpha}\,(\gamma+x_i^{\beta})^{-\alpha}$

$$\hat{c} = \frac{n}{-w} \tag{15}$$

4. Reliability Measures

Survival and hazard functions referred to the cumulative distribution function (1) of EPOLO distribution are obtained. The examination of organism or technological unit breakdowns that take place after a certain time is aided by the survival function. The hazard rate is used to track the specific unit over the course of its lifespan distribution. The likelihood to fail or die, depending on the age attained, is measured by the hazard rate (HR), which is a critical factor in categorizing lifespan distributions. According to Arun Kumar Rao [1] and Hare and Sharma [4] the hazard rates are often monotonic or non-monotonic. The possibility of being alive for a specific period of time or the likelihood that an important event will take place before a certain amount of time (t) is the definition of the survival function (x). The survival mechanism is expressed as

$$S\,(x) = 1-\left(1-\gamma^{\alpha}\left(\gamma + x^{\beta}\right)^{-\propto}\right)^{c} \quad x > 0;\ \alpha,\beta,\gamma,c > 0 \tag{16}$$

Where *S(0) = 1 and lim x→∞ S(x) = 0.*

To find out MSE for the survival rate, we need to calculate

$$\widehat{S\,(x)} = \int_0^{\infty} S\,(x)\,\pi\,(\frac{c}{x})\,dc \tag{17}$$

$$\widehat{S\,(x)} = \int_0^{\infty} 1-(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c}\,\frac{[b-w]^{n+a}}{\Gamma(n+a)}\; c^{n+a-1}exp\{-c(b - w)\}\; dc \tag{18}$$

The failure rate, also known as the hazard function h(x), is the ratio of the likelihood that an event will occur in a given amount of time, t, to the likelihood that it will pass off successfully. When expressing the hazard function it is given as,

$$h(x) = \frac{f(x)}{S(x)} = \frac{\alpha\beta c\gamma^{\alpha} x^{\beta-1} \left(\gamma + x^{\beta}\right)^{-\alpha-1} (1 - \gamma^{\alpha} (\gamma + x^{\beta})^{-\alpha})^{C-1}}{1 - (1-\gamma(\gamma+x^{\beta})^{-\alpha})^{C}} \quad x > 0; \tag{19}$$

$$\alpha, \beta, \gamma, c > 0$$

To calculate the MSE of the hazard function

$$\widehat{h(x)} = \int_0^{\infty} h(x)\, \pi\, (\frac{c}{x})\, dc \tag{20}$$

$$\widehat{h(x)} = \int_0^{\infty} \frac{\alpha\beta c\gamma^{\alpha} x^{\beta-1} (\gamma + x^{\beta})^{-\alpha-1} (1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{C-1}}{1-(1-\gamma(\gamma+x^{\beta})^{-\alpha})^{C}} \frac{[b-w]^{n+a}}{\Gamma(n+a)} c^{n+a-1} exp\{-c(b-w)\}\ dc \tag{21}$$

After finding the $\widehat{S(x)}$ *and* $\widehat{h(x)}$ using the formula then MSE is calculated to compare the errors.

5. Quantile function

Using the c.d.f, the quantile function can be found by solving the equation F(x) = u, 0< u < 1. Assume that X is a random variable with an EPOLO distribution. The equation F(Q(u)) = u defines the quantile function Q(u) and represents it.

$$F(x) = \left(1 - \gamma^{\alpha}\left(\gamma + x^{\beta}\right)^{-\propto}\right)^{c} x > 0;\ \alpha, \beta, \gamma, c > 0 \tag{24}$$

$$u^{1/c} = 1 - \gamma^{\alpha}(\gamma + x^{\beta})^{-\alpha} \tag{25}$$

$$[(\frac{1 - u^{\frac{1}{c}}}{\gamma^{\alpha}})^{-1/\alpha} - \gamma]^{1/\beta} = x \tag{26}$$

6. Monte Carlo simulation study

The effectiveness of the parameters is tested using a simulation study, which is conducted for the shape parameter c of the EPOLO distribution. Mean Square Error (MSE) using the two familiar estimation maximum likelihood estimation and Bayesian estimation under squared error loss function was calculated. The following algorithm is proposed by Shrestha and Kumar [12] for the simulation studies.

To simulate the random samples of x which follows the EPOLO distribution, the equation F(x) = u is used, where u (0, 1) follows uniform distribution and F(x) is the CDF of the EPOLO distribution. Fix the values of the parameters α, β, γ, c. Here we fix other parameters to estimate the shape parameter c.

- Determine the different sample sizes
- Calculate $\hat{c}$ for each of the n samples
- Calculate the MSE for the parameter *c* using the below formula

The average squared difference between the estimator and parameter, measured by the MSE, is a reliable indicator of an estimator's performance. Generally speaking, any growing function of the absolute distance $|\hat{c} - c|$ would be used to assess an estimator's quality. But compared to other distance metrics, MSE offers at least two advantages: First, it can be analyzed, and second, it has the interpretation.

$$MSE(c) = \sum_{i=1}^{n} \frac{(\hat{c} - c)^2}{n}$$

A simulation were conducted for samples n=15, 30, 50, 100 and repeated for 1,000 times with the parameter c values 0.5, 1.5.

7. Analyses

According to Venables et al. [18], all the analysis is done using the R software. R software is free source software helps a lot in statistical analysis.

Table 1: *MSE of Maximum likelihood estimate of the shape parameter c and MSE of Bayesian estimate of the shape parameter under SELF loss function is given below*

n	c	MSE (MLE)	MSE (BAYES)
15	0.5	0.0229	0.0165
	1.5	0.2322	0.1087
30	0.5	0.0093	0.0081
	1.5	0.086	0.0635
50	0.5	0.0059	0.0053
	1.5	0.0489	0.0403
100	0.5	0.0021	0.0021
	1.5	0.034	0.0263

Table 2: *MSE of survival and hazard function of c along with the Bayesian posterior is given below*

n	c	survival	hazard
15	0.5	0.031	1.03e-04
	1.5	0.00267	5.77e-07
30	0.5	0.0092	2.11e-05
	1.5	0.00076	5.02e-08
50	0.5	0.0045	7.28e-06
	1.5	0.00027	8.43e-09
100	0.5	0.0012	9.58e-07
	1.5	5.49e-05	3.72e-10

8. Conclusion

In this article, we looked at the Bayesian estimates for the exponentiated power Lomax distribution's of shape parameter c. Based on the findings in Tables 1 and 2, we note the following:

- For the shape parameter c, the MSE of the Bayesian and maximum likelihood estimate have been determined. In that MSE of Bayesian is less than MLE method that indicates the Bayesian estimation is a good estimator for estimating the parameters of any distribution. As sample size increases both will be same.
- The survival and hazard function using the Bayesian posterior is calculated and it indicates the distribution is of good fit. The error decreases as the sample size increases for survival and hazard functions.

References

[1] Arun Kumar Rao. (2016). Estimation of Reliability Function of Lomax Distribution via Bayesian Approach. International Journal of Mathematics Trends and Technology (IJMTT) – Volume 40 Number

[2] Amal S Hassan., Said g., and Nassr. (2018). Power lomax poisson distribution: properties and estimation. journal of data science, 105-128.

[3] Balakrishnan, N., and Ahsanullah, M. (1994). Relations for single and product moments of record values from lomax distribution. Sankhya: indian j statist. Ser b, 56:140-146.

[4] Hare Krishna., and Ranjeet Sharma. (2007). Estimation of reliability characteristics of general system configuration. Chaudhary Charan Singh University, Meerut, India, IJQRM.

[5] Hassan, A.S., and Al-ghamdi, A.S. (2009). Optimum step stress accelerated life testing for lomax distribution. J appl sci res,5:2153-2164

[6] Hesham mohamed reyad., and Soha othman ahmed. (2016). Bayesian and e-bayesian estimation for the kumaraswamy distribution based on type-ii censoring. international journal of advanced mathematical sciences, 4 (1).

[7] Kilany, N.M. (2016). Weighted lomax distribution. Springerplus, 5(1):1862.

[8] Lemonte, A.J., and Cordeiro, G.M. (2013). An extended lomax distribution. Statistics, 47(4):800-816.

[9] Lingappaiah, G. S. (1986). On the pareto distribution of the second kind (lomax distribution). Revista de mathemtica e estatistica, 4:63–68.

[10] Lomax, K.S. (1954). Business failures: another example of the analysis of failure data. J am statist assoc, 45:21–29.

[11] Mahmoud, M. A., Soliman, A. A., Abd ellah, A. H., and El-sagheer, R. M. (2013). Mcmc technique to study the bayesian estimation using record values from the lomax distribution. International journal of computer applications, 73(5).

[12] Mohamed mohamed ezzat abd el-monsef. (2020). The exponentiated power lomax distribution and its applications, Qual Reliab Engng Int, 1–24.

[13] Myhre, J., and saunders, S. (1982). Screen testing and conditional probability of survival. Lect notes-monogr series, 166–178.

[14] Nada S. Karam. (2014). Bayesian analysis of five exponentiated distributions under different priors and loss functions, iraqi journal of science, vol 55, no.3b, pp: 1353-1369.

[15] Shrestha, S. K., and kumar, V. (2014). Bayesian analysis of extended lomax distribution. International journal of mathematics trends and technology, 7(1), 33-41.

[16] Sowbhagya S prabhu. (2020). E-bayesian estimation of the shape parameter of lomax model under symmetric and asymmetric loss functions, International Journal of Statistics and Applied Mathematics, 5(6): 142-146

[17] Tierney, l. (1994). Markov chains for exploring posterior distributions. The annals of statistics, 1701-1728.

[18] Venables, W. N., Smith D. M., and the R core team. (2022). An introduction to r notes on r: a programming environment for data analysis and graphics version 4.2.1.

HUNTSBERGER TYPE SHRINKAGE ENTROPY ESTIMATOR FOR VARIANCE OF NORMAL DISTRIBUTION UNDER LINEX LOSS FUNCTION

Priyanka Sahni

•

Department of Mathematics, Maharshi Dayanand University, Rohtak.
sahnipriyanka2@gmail.com

Rajeev Kumar

•

Department of Mathematics, Maharshi Dayanand University, Rohtak.
profrajeevmdu@gmail.com

Abstract

The aim of the paper is to develop a better estimator for the entropy function of variance of the normal distribution. The present paper proposes a Huntsberger type shrinkage estimator of the entropy function for the variance of normal distribution. This Huntsberger type shrinkage entropy estimator is based on test statistic, which eliminates arbitrariness of choice of shrinkage factor. For the proposed estimator risk expressions under LINEX loss function have been calculated. Numerical computations and graphical analysis is carried out for risk and relative risks for the proposed estimators. It is also compared with the existing best estimator for distinct degrees of asymmetry and different levels of significance. Based on the criteria of relative risk, it is found that the proposed Huntsberger type shrinkage estimator is better than the existing estimator for the entropy function of variance of normal distribution for smaller values of level of significance and degrees of freedom..

Keywords: Normal distribution, entropy function, shrinkage estimation, LINEX loss function, level of significance, relative risk.

1. Introduction

Normal distribution plays a vital role in theory of statistics. Its testimation and estimating its parameters have been acknowledged and refined by researchers. Pandey et al. [9] proposed some shrinkage testimators of variance under the mean square error criterion. Parsian and Farsipour [10], Mishra and Meulen [7], Ahmadi et al. [1], Singh et al. [14], Prakash et al [12], Prakash and Pandey [11] and others have studied the estimation methods under LINEX loss function in distinct contexts. The concept of entropy was introduced by Shannon [13] and is given as

$$H(f) = E[-\ln(f(X))], \tag{1}$$

where X is a random variable having probability density function f and distribution function F. For sharply peaked distribution entropy is very low and is much higher when the probability is spread out. Many authors worked on the estimation entropy for different life distributions. Misra et al. [8] proposed an entropy estimator for a multivariate normal distribution while Jeevanand and Abdul- Sathar [4] also obtained estimators for the residual entropy function of exponential distribution from censored samples. Lazo and Rathee [6] and Kayal and Kumar [5] also worked in this direction.

Suppose the random variable X has the probability distribution $f(x,\theta)$ where interest is to estimate entropy function as a function of θ. Thomson [17] proposed a shrinkage type estimator $k\hat{\theta}+(1-k)\theta_0$, where k is constant and is designed to shrink the usual estimator $\hat{\theta}$ of the parameter θ towards a natural origin θ_0 and Huntsberger [3] introduced weighted shrinkage estimator of the form

$$\tilde{\theta}_f = \phi(\hat{\theta})\hat{\theta} + (1-\phi(\hat{\theta}))\theta_0 ,$$

where $\phi(.)$, $0 \le \phi(.) \le 1$, represents a weighted function specifying the degree of belief in θ_0. In this paper, we shall concentrate on obtaining Huntsberger type shrinkage estimation of entropy function with respect to asymmetric loss function for a random sample $x_1, x_2, ..., x_m$ of size m from a normal distribution.

The form of normal density we consider is

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} ; -\infty < x < \infty,\ -\infty < x < \infty,\ \sigma > 0 \tag{2}$$

For the normal distribution, the entropy function can be obtained as

$$H(f) = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(\sigma^2) \tag{3}$$

Since $H(f)$ is linear function of $\frac{1}{2}\ln(\sigma^2)$, estimating $H(f)$ is correspondent to estimating $\frac{1}{2}\ln(\sigma^2)$. We shall write $I(\sigma^2) = \frac{1}{2}\ln(\sigma^2)$ so that $H(f) = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + I(\sigma^2)$. Now we shall discuss estimation of $I(\sigma^2)$. Since $I(\sigma^2)$ is continous function of σ^2, the MLE of $I(\sigma^2)$ is obtained by replacing σ^2 by its MLE $\hat{\sigma}^2$ in $I(\sigma^2)$. Then, the MLE of entropy function for the exponential distribution is

$$\begin{aligned} H(f) &= \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(\hat{\sigma}^2) \\ &= \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(s^2) \end{aligned} \tag{4}$$

where $s^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - \bar{x}\right)^2$ is MLE of σ^2, when μ is unknown.

It can be shown that s^2 has distribution as

$$f(s^2) = \frac{1}{2^{\frac{m-1}{2}}\Gamma\left(\frac{m-1}{2}\right)} e^{-\frac{ms^2}{2\sigma^2}} \left(s^2\right)^{\frac{m-1}{2}-1} \left(\frac{m}{\sigma^2}\right)^{\frac{m-1}{2}} \tag{5}$$

Although the SELF (squared error loss function) is commonly used for estimating various statistics parameters, it may not be convenient in actual situations, particularly in insurance claims, estimating any health statistics parameter, over-estimation and under-estimaton have distinct impacts. An analysis of various superior properties of asymmetric loss function over squared error loss function has been presented by several authors. Basu and Ebrahimi [2] derived Bayes estimators for mean lifetime and reliability function for the exponential model using asymmetric loss function. Srivastava and Tanna [16] as well as Srivastava and Shah [15] also derived estimators and studied their properties under asymmetric loss function.

A fruitful asymmetric loss function, LINEX loss function was recommended by Varian [18] as:

$$L(\Delta)=b(e^{a\Delta}-a\Delta-1),\ b>0,\ a\neq 0. \tag{6}$$

The magnitude and sign of 'a' shows the degree of asymmetry and direction respectively. When overestimation is more severe than underestimation then positive values of 'a' are taken, whereas in reverse situations its negative values are usually preferred and 'b' is constant of proportionality.

2. The Shrinkage Estimator

From a normal population with mean μ and variance σ^2, a random sample $x_1, x_2,...,x_m$ of size m is taken. Let the initial guess value for σ^2 is presumed to be $\sigma_0{}^2$, available from the past knowledge or some other reliable origins. It is noted that MLE of σ^2 is s^2 having variance of $s^2=\frac{2(n-1)\sigma^4}{n^2}$.

Here, we test the null hypothesis H_0: $\sigma^2=\sigma_0{}^2$ against the alternative H_1: $\sigma^2\neq\sigma_0{}^2$ using the test statistic $\frac{ms^2}{\sigma_0^2}$, where $s^2=\frac{1}{m}\sum_{i=1}^{m}(x_i-\bar{x})^2$ which follows χ^2- distribution with degree of freedom $m-1$. If $\chi_1{}^2\leq\frac{ms^2}{\sigma_0^2}\leq\chi_2{}^2$ then H_0 may be approved at $\alpha\%$ level of significance, where lower and upper α^{th} percentile values of χ^2 distribution are $\chi_1{}^2$ and $\chi_2{}^2$ respectively with degree of freedom $m-1$. Then by taking shrinkage factor $k=\frac{ms^2}{\sigma_0^2\chi^2}$, which is negatively associated with $\chi^2(=(\chi_2{}^2-\chi_1{}^2))$, a shrinkage entropy estimator may be considered. If data does not hold, H_0 it may be dropped and in this case it is recommended to use $\frac{1}{2}\ln(s^2)$, the MLE of $\frac{1}{2}\ln(\sigma^2)$.

Thus, the proposed shrinkage entropy estimator $\tilde{I}_1(\sigma^2)$ of $\frac{1}{2}\ln(\sigma^2)$ is as under:

$$\tilde{I}_1(\sigma^2)=\begin{cases} k(\frac{1}{2}\ln(s^2))+(1-k)(\frac{1}{2}\ln(\sigma_0^2)), & \text{if} \quad \chi_1{}^2\leq\frac{ms^2}{\sigma_0^2}\leq\chi_2{}^2 \\ \frac{1}{2}\ln(s^2) \quad , & \text{Otherwise}\end{cases} \tag{7}$$

3. Risk of Estimator

Risk of estimator $\tilde{I}_1(\sigma^2)$ under LLF is obtained as under:

$$R_{LLF}(\tilde{I}_1(\sigma^2))=E(\tilde{I}_1(\sigma^2)/LLF)$$

$$=\int_{\frac{\sigma_0^2\chi_1{}^2}{m}}^{\frac{\sigma_0^2\chi_2{}^2}{m}}\begin{pmatrix}\exp(a((\frac{ms^2}{2\sigma_0^2\chi^2})\ln(s^2)+(1-(\frac{ms^2}{\sigma_0^2\chi^2}))\frac{\ln(\sigma_0^2)}{2}-\frac{\ln(\sigma^2)}{2})) \\ -a((\frac{ms^2}{2\sigma_0^2\chi^2})\ln(s^2)+(1-(\frac{ms^2}{\sigma_0^2\chi^2}))\frac{\ln(\sigma_0^2)}{2}-\frac{\ln(\sigma^2)}{2})-1\end{pmatrix}f(s^2)ds^2$$

$$+\int_0^{\infty}\left(\exp(\frac{a}{2}(\ln(s^2)-\ln(\sigma^2)))-\frac{a}{2}(\ln(s^2)-\ln(\sigma^2))-1\right)f(s^2)ds^2$$

$$-\int_{\frac{\sigma_0^2\chi_1{}^2}{m}}^{\frac{\sigma_0^2\chi_2{}^2}{m}}\left(\exp(\frac{a}{2}(\ln(s^2)-\ln(\sigma^2)))-\frac{a}{2}(\ln(s^2)-\ln(\sigma^2))-1\right)f(s^2)ds^2$$

Straight forward integration gives

$$R_{LLF}(\hat{I}_1(\sigma^2)) = I_1 + I_2 + I_3 + I_4 - \frac{au\log(\frac{2}{m\phi})}{\phi\chi^2}\left[I(r_2', u+1) - I(r_1', u+1)\right] + v\ln(\frac{2}{m\phi})\left[I(r_2', u) - I(r_1', u)\right]$$

$$-\frac{2^v\Gamma(u+v)}{m^v\Gamma(u)}\left[I(r_2', u+v) - I(r_1', u+v)\right] - v\ln(\frac{2}{m}) - 1 + \frac{2^v\Gamma(u+v)}{m^v\Gamma(u)}, \tag{8}$$

where $r_1' = \frac{\phi\chi_1^2}{2}$, $r_2' = \frac{\phi\chi_2^2}{2}$, $\phi = \frac{\sigma_0^2}{\sigma^2}$, $u = \frac{m-1}{2}$, $v = \frac{a}{2}$ and I(x, n) is the cumulative distribution function of gamma distribution given as

$$I(x, n) = \int_0^x \frac{e^{-t}t^{n-1}}{\Gamma(n)}dt$$

and

$$I_1 = \int_{r_1'}^{r_2'} \phi^v \left(\frac{2t}{m\phi}\right)^{\frac{at}{\phi\chi^2}} \frac{e^{-t}t^{u-1}}{\Gamma(u)}dt, \quad I_2 = \frac{-au}{\phi\chi^2}\int_{r_1'}^{r_2'}(\log t)\frac{e^{-t}t^u}{\Gamma(u+1)}dt, \quad I_3 = v\int_{r_1'}^{r_2'}(\log t)\frac{e^{-t}t^{u-1}}{\Gamma(u)}dt$$

and

$$I_4 = -v\int_0^\infty(\log t)\frac{e^{-t}t^{u-1}}{\Gamma(u)}dt$$

4. Relative Risk

A common way of analyzing risk of considered estimator, is to examine its work relative to the best possible estimator $\hat{I}(\sigma^2)$ in this case. With this motto, we calculate risk of $\hat{I}(\sigma^2)$ as:

$$R_{LLF}\left(\hat{I}(\sigma^2)\right) = E\left(\hat{I}(\sigma^2)\setminus LLF\right)$$

$$= \int_0^\infty (\exp(a(\frac{1}{2}\ln(s^2) - \frac{1}{2}\ln(\sigma^2))) - a(\frac{1}{2}\ln(s^2) - \frac{1}{2}\ln(\sigma^2)) - 1)f(s^2)ds^2$$

$$= \int_0^\infty (\exp(\frac{a}{2}\ln(\frac{s^2}{\sigma^2})) - \frac{a}{2}\ln(\frac{s^2}{\sigma^2}) - 1)f(s^2)ds^2$$

Now, taking the transformation $t = \frac{ms^2}{2\sigma^2}$ and then solving the integral, we get

$$R_{LLF}\left(\hat{I}(\sigma^2)\right) = \frac{2^v\Gamma(u+v)}{m^v\Gamma(u)} - v\ln(\frac{2}{m}) - 1 - v\psi(u), \tag{9}$$

where

$$\psi(n) = \frac{\frac{d}{dn}\Gamma(n)}{\Gamma(n)}$$

Now, we determine relative risk of $\tilde{I}_1(\sigma^2)$ under LLF as

$$RR_{LLF}(\tilde{I}_1(\sigma^2)) = \frac{R_{LLF}(\hat{I}(\sigma^2))}{R_{LLF}(\tilde{I}_1(\sigma^2))} \tag{10}$$

Using (8) and (9) the expression given in (10) can be obtained. It is observed that relative risk given above is a function of m, a, ϕ and α.

5. Numerical Computations And Graphical Analysis

To examine the performance of $\tilde{I}_1(s^2)$, a few values of these parameters have been taken as α = 0.01, 0.05, 0.1, m=5, 8, 11, ϕ = 0.2(0.2)2 and a=-2, -1, 1, 1.5, 1.75 i.e. both positive and negative values, as a is the best essential component that determines the seriousness of over/under estimation in the actual cases. Several tables and graphs for relative risk calculation are represented in Table 1, Figure 1 to Figure 6. However our recommendations depending on all these analyses are as follows:

i. For $m = 5$, $\alpha = 1\%$ and for considered values of 'a', $\tilde{I}_1(\sigma^2)$ gives better results than the existing estimator for all values of 'a' and for the whole scale of ϕ i.e. $0.2 \le \phi \le 2$.

ii. Further if we switch α to 5%, the same type of behavior noticed for relative risk (RR). However, magnitude of RR is smaller as computed to $\alpha = 1\%$ values.

iii. Taking $\alpha = 10\%$ in order to observe the pattern for higher values of α and it is found that $\tilde{I}_1(\sigma^2)$ still gives the better results as compared to the existing estimator whereas magnitude of relative risks values become lower but even though it remains above unity.

iv. After comparing these relative risks, a lower value of α is preferred. Similarly, as 'm' raises there is a fall in RR values for distinct values of α and a. However the best result of $\tilde{I}_1(\sigma^2)$ is observed at $\alpha = 1\%$ for $a = 1$ and $\alpha = 1\%$ for $a = 1.75$.

It is therefore suggested to take up a smaller values of $\alpha\,(=1\%)$ as well as m (=5 or 8) for better results for a, in particular $a(=1\;\&\;1.75)$.

Table 1: *Relative risk of estimator $\tilde{I}_1(\sigma^2)$ under LLF*

α=0.01		ϕ									
m	a	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2
	-2	1.4182	2.3334	3.5356	4.3068	4.2010	3.6270	3.0222	2.5295	2.1539	1.8697
	-1	1.2002	1.9735	3.5429	5.7055	6.5919	5.5134	4.0969	3.0576	2.3684	1.9069
5	1	1.0382	1.4929	2.6592	5.1746	8.1241	7.0627	4.4323	2.7966	1.9057	1.3929
	1.5	1.0226	1.4256	2.4849	4.8294	7.8372	6.9390	4.2635	2.623	1.7512	1.2593
	1.75	1.0166	1.3967	2.4065	4.6584	7.6447	6.8323	4.1657	2.5355	1.6775	1.1972
	-2	0.9844	1.3487	2.3739	4.2491	5.384	4.3037	2.9433	2.0695	1.5473	1.2225
	-1	0.9594	1.2222	2.0956	4.0243	5.9406	4.8704	3.0967	2.0288	1.4387	1.0928
8	1	0.9493	1.0861	1.7061	3.2695	5.6045	4.9869	2.9165	1.7353	1.1407	0.8166
	1.5	0.9502	1.0664	1.6375	3.0932	5.3591	4.8641	2.8201	1.6510	1.0697	0.7563
	1.75	0.9509	1.0580	1.6064	3.0095	5.2275	4.7898	2.7692	1.6097	1.0358	0.7279
	-2	0.9615	1.0777	1.7002	3.2243	4.8735	3.9475	2.4526	1.5938	1.1345	0.872
	-1	0.9634	1.0316	1.5541	2.9483	4.8576	4.1132	2.4478	1.5165	1.0407	0.7778
11	1	0.9715	0.983	1.3555	2.4472	4.3008	3.9708	2.2651	1.3114	0.8494	0.6054
	1.5	0.9736	0.9764	1.3197	2.3413	4.117	3.8711	2.2019	1.2593	0.8058	0.5681
	1.75	0.9745	0.9738	1.3034	2.2913	4.0237	3.8159	2.1695	1.2339	0.7849	0.5504
α= 0.05											

	-2	1.4071	1.8922	2.2377	2.3004	2.1503	1.9251	1.7058	1.5193	1.3679	1.2466
	-1	1.2576	1.7852	2.4289	2.8421	2.7907	2.4451	2.0534	1.7221	1.4656	1.2712
5	1	1.1105	1.4803	2.1447	2.9731	3.3998	3.0605	2.3984	1.8212	1.4072	1.1213
	1.5	1.0933	1.4267	2.05	2.8825	3.3938	3.1	2.4085	1.7949	1.3598	1.0644
	1.75	1.0864	1.4029	2.0043	2.8294	3.3726	3.1043	2.4045	1.777	1.3339	1.0353
	-2	1.0733	1.3764	1.9092	2.4083	2.4372	2.0694	1.6564	1.3379	1.1146	0.96
	-1	1.0421	1.2813	1.7875	2.4168	2.6348	2.2633	1.7518	1.3543	1.0839	0.9028
8	1	1.0163	1.1634	1.5537	2.1943	2.6789	2.4327	1.8167	1.3106	0.9792	0.768
	1.5	1.0134	1.1445	1.5057	2.1206	2.6324	2.4297	1.8097	1.2887	0.9488	0.734
	1.75	1.0123	1.1362	1.4832	2.0833	2.6063	2.4223	1.8035	1.2768	0.9334	0.7174
	-2	1.0267	1.1768	1.5449	2.0735	2.2885	1.9517	1.4932	1.1549	0.9373	0.7999
	-1	1.0187	1.1351	1.4549	1.9908	2.3257	2.0333	1.5248	1.1407	0.8974	0.7464
11	1	1.0111	1.0846	1.3150	1.78	2.2286	2.0732	1.5306	1.0868	0.8091	0.6410
	1.5	1.01	1.0764	1.2877	1.7272	2.1801	2.0596	1.5219	1.07	0.7869	0.6161
	1.75	1.0095	1.0728	1.2749	1.7013	2.1535	2.0498	1.5162	1.0612	0.7759	0.6040

5.1. Graphs of Relative Risk for $\tilde{I}_1(\sigma^2)$

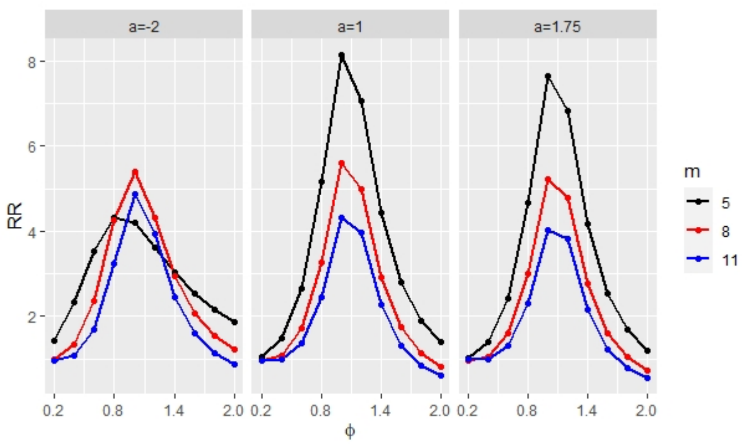


Figure 1: *For α=0.01*

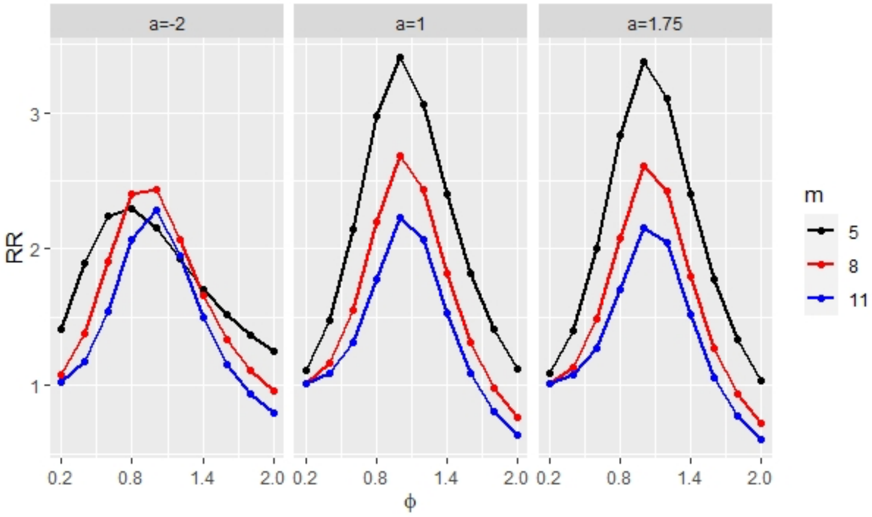


Figure 2: *For α=0.05*

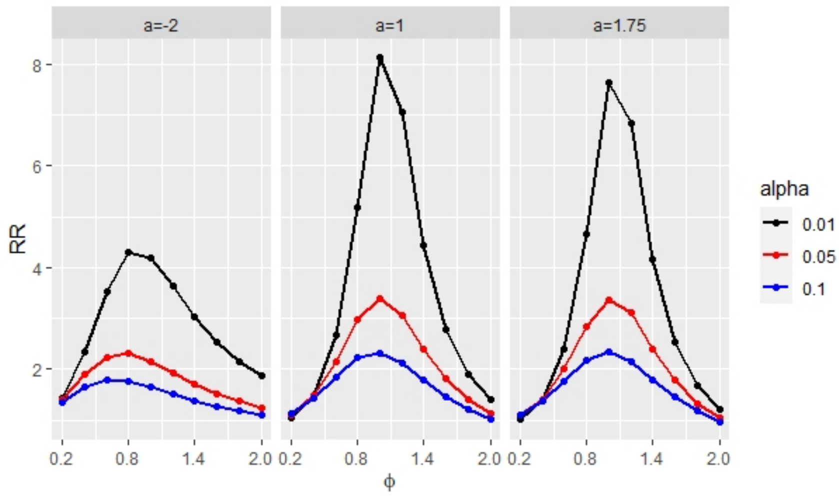


Figure 3: For m=5

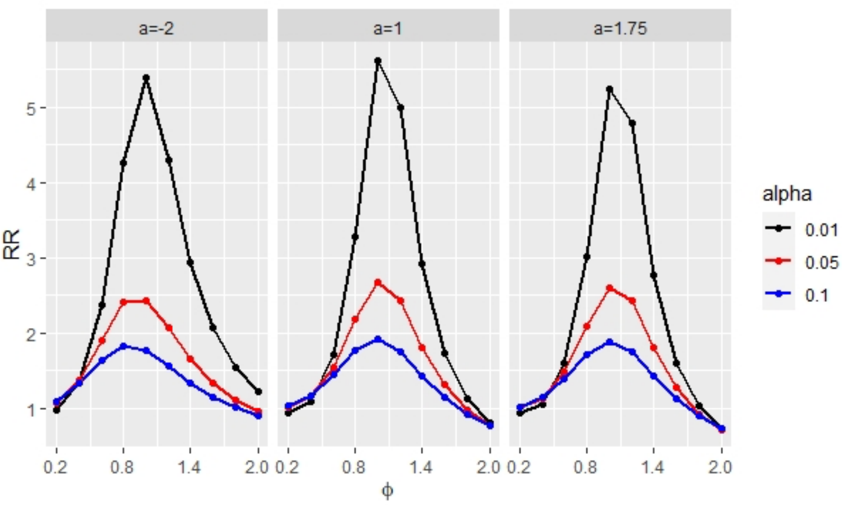


Figure 4: *For m=8*

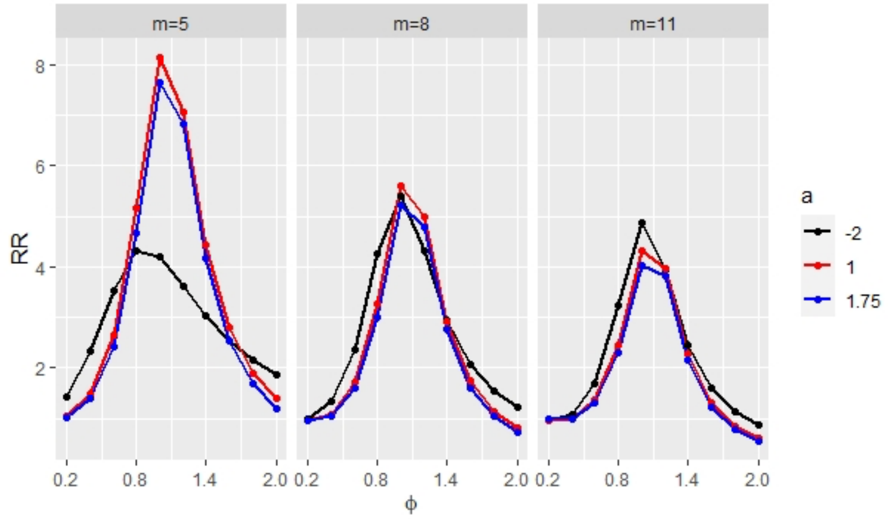


Figure 5: *For α=0.01*

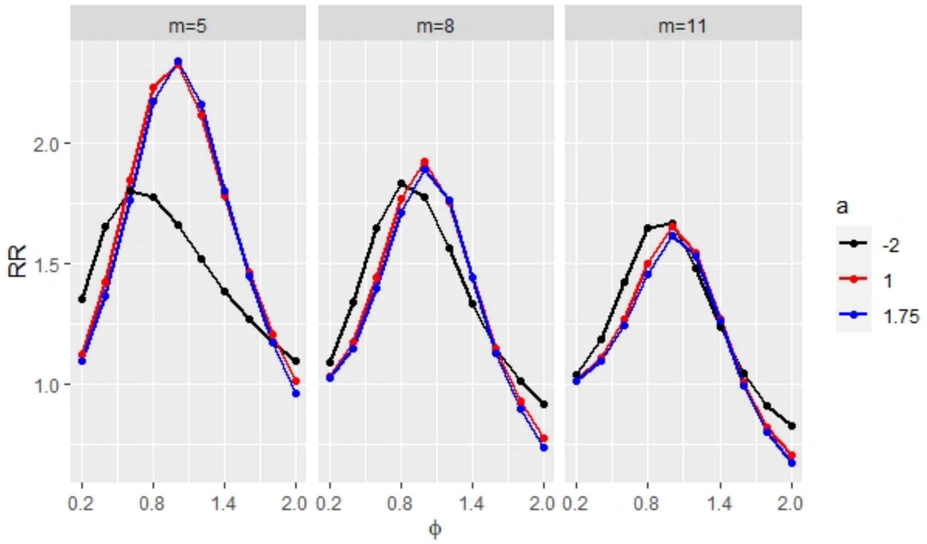


Figure 6: *For α=0.1*

6. Conclusion

In this paper, a Huntsberger Type shrinkage entropy estimator for normal distribution have been proposed and its properties have been investigated under LINEX Loss function. On the basis of relative risk, it is concluded that the proposed estimator gives better results for smaller values of level of significance and degrees of freedom.

References

[1] Ahmadi, J., Doostparast, M. and Parsian, A. (2005). Estimation and prediction in a two parameter exponential distribution based on k-record values under LINEX loss function. *Communications in Statistics: Theory and Methods*, 34:795-805.

[2] Basu, A.P. and Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of Statistical Planning and Inference*, 29:21-31.

[3] Huntsberger, D.V. (1955). A generalization of a preliminary testing procedure for pooling data. *The Annals of Mathematical Statistics*, 26:734-743.

[4] Jeevanand, E.S. and Abdul-Sathar, E.I. (2009). Estimation of residual entropy function for exponential distribution from censored samples. *ProbStat Forum.*, 2:68-77.

[5] Kayal, S. and Kumar, S. (2011). Estimating the entropy of an exponential population under the LINEX loss function. *Journal of Indian Statistical Association*, 49:91-112.

[6] Lazo, A.C.G.V. and Rathie, P.N. (1978). On the entropy of continous distribution. *IEEE Transactions on Information Theory*, 24:120-122.

[7] Misra, N. and Meulen, E.C.V.D. (2003). On estimating the mean of the selected normal population under the LINEX loss function. *Metrika: International Journal for Theoretical and Applied Statistics,* 58:173-184.

[8] Misra, N., Singh, H. and Demchuk, E. (2005). Estimation of the entropy of multivariate normal distribution. *Journal of Multivariate Analysis*, 92:324-342.

[9] Pandey, B.N., Malik, H.J. and Srivastava, R. (1988). Shrinkage testimator for the variance of a normal distribution at single and double stages. *Microelectron Reliability*, 28:929-944.

[10] Parsian, A. and Farsipour, N.S. (1999). Estimation of the mean of the selected population under asymmetric loss function. *Metrika*, 50:89-107.

[11] Prakash, G. and Pandey, B.N. (2010). Shrinkage estimation of the variance for the exponential type-II censored data. *International Journal of Statistics and Systems*, 5:251-263.

[12] Prakash, G., Singh, D.C. and Sinha, S.K. (2008). On shrinkage estimation for the scale parameter of weibull distribution. *Data Science Journal*, 7:125-136.

[13] Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379-423.

[14] Singh, D.C., Prakash, G. and Singh, P. (2007). Shrinkage testimators for the shape parameter of pareto distribution using the LINEX loss function. *Communications in Statistics-Theory and Methods*, 36:741-753.

[15] Srivastava, R. and Shah, T. (2010). Shrinkage testimators of scale parameter for exponential model under asymmetric loss function. *Journal of Reliability and Statistical Studies*, **3:**11-25.

[16] Srivastava, R. and Tanna, V. (2001). An estimation procedure for error variance incorporating PTS for random effects model under LINEX loss function. *Journal-Communication in Statistics- Theory and Methods,* 30:2583-2599.

[17] Thompson, J.R. (1968). Some shrinkage techniques for estimating the mean. *Journal of the American Statistical Association*, 63:113-122.

[18] Varian, H.R. (1975). A bayesian approach to real estate assessment studies in bayesian econometrics and statistics in honour of L.J. Savage Eds. S. E. Fienberg and A. Zeliner. Amsterdam. North Holland, 195-208.

THE INVERSE BURR LOG-LOGISTIC DISTRIBUTION: PROPERTIES, APPLICATIONS AND DIFFERENT METHODS OF ESTIMATION

Festus C. Opone[1]
Kadir Karakaya[2]
Francis E.U. Osagiede[3]
[1] Department of Statistics, Delta State University of Science and Technology, Ozoro, Nigeria.
[2] Department of Statistics, Selcuk University, Konya, Turkey.
[3] Department of Mathematics, University of Benin, Benin City, Nigeria.

festus.opone@physci.uniben.edu[1]
kkarakaya@selcuk.edu.tr[2]
francis.osagiede@uniben.edu[3]

Abstract

Lifetime distributions have played a significant role in lifetime data analysis. Despite the numerous distributions in literature, there have been several motivations for developing new ones. In this paper, a new lifetime distribution is proposed. Some important functions of the new distribution, such as probability density, cumulative distribution, survival, hazard, and quantile are derived in closed form. Some distributional properties such as moments, moment generating function, linear representation, probability weighted moments, etc. are obtained. Some estimators such as the least square estimator (LSE), the weighted least square estimator (WLSE), the Anderson-Darling estimator (ADE) and the Cramer-von Mises estimator (CvME) are investigated for three unknown parameters. The efficiency of the estimators is checked via Monte Carlo simulation based on the bias and mean square error criteria. The usability of the new distribution is investigated with two real data sets and empirical results obtained reveal that the new distribution offers a promising fit for the data sets under study.

Keywords: Bur distribution, log-logistic distribution, parameter estimation, quantile

1. INTRODUCTION

Statistical distributions have played a significant role in lifetime data analysis. Despite the numerous distributions in literature, there have been several motivations for developing new ones. In all, the central goal has remained to develop a more flexible and tractable distribution in fitting real-world problems. In the last decades, researchers have introduced different methodologies for generating new statistical distributions which are hoped to provide a better fit than the existing distributions in lifetime data analysis. Some of these methods are the Beta-G family by [7], Marshall-Olkin extended family by [11], Transmuted-G family by [14], Kumaraswamy-G family by [5], Transformer (T-X) family by [2], Weibull-G family by [4], Odd Burr-G family by [1], Type II Topp-Leone generated family by [6], etc.

Recently, [13] introduced the Inverse Burr-G family of distributions using the idea of [15]. By considering the inverse Burr as the generator, they defined the cumulative distribution function of the inverse Burr-G family of distribution as

$$F(x,\xi) = \alpha\beta\int_0^{-\log[1-G(x,\xi)]} x^{-(\alpha+1)}\left(1+x^{-\alpha}\right)^{-(\beta+1)}dx, \\ = \left[1+\{-\log[1-G(x,\xi)]\}^{-\alpha}\right]^{-\beta}, \qquad x>0,\ \alpha,\ \beta>0, \tag{1}$$

where α and β are the shape parameters and $G(x,\xi)$ is the baseline distribution which depends on a parameter vector ξ.

The corresponding density function associated with (1) is given by

$$f(x,\xi) = \alpha\beta\, g(x,\xi)[1-G(x,\xi)]^{-1}\{-\log[1-G(x,\xi)]\}^{-(\alpha+1)}\left[1+\{-\log[1-G(x,\xi)]\}^{-\alpha}\right]^{-(\beta+1)}. \tag{2}$$

In this paper, we employed the technique defined in (1) and consider in particular, the case where the baseline distribution $G(x,\xi)$ follows the log-logistic distribution.

The cumulative distribution function (cdf) and probability density function (pdf) of the log-logistic distribution with shape parameter $\lambda > 0$ are respectively defined as

$$G(x) = 1-\left(1+x^{\lambda}\right)^{-1}, \tag{3}$$

and

$$g(x) = \lambda x^{\lambda-1}\left(1+x^{\lambda}\right)^{-2}, \qquad x>0,\ \lambda>0. \tag{4}$$

Inserting (3) and (4) into (1) and (2), we define the cdf and pdf of a new statistical distribution as

$$F(x) = \left[1+\left\{\log\left(1+x^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}, \qquad x>0,\ \alpha,\lambda,\beta>0, \tag{5}$$

and

$$f(x) = \alpha\beta\lambda x^{\lambda-1}\left(1+x^{\lambda}\right)^{-1}\left\{\log\left(1+x^{\lambda}\right)\right\}^{-(\alpha+1)}\left[1+\left\{\log\left(1+x^{\lambda}\right)\right\}^{-\alpha}\right]^{-(\beta+1)}. \tag{6}$$

Suppose a random variable *X* has the density function in (6), then we say that *X* follows the Inverse Burr Log-Logistic ("IBLL" for short) distribution with shape parameters α, β and λ.

The motivation of this paper is to develop a tractable distribution that spans all the various forms of the hazard rate properties and provides a consistently better fit than most available statistical distribution in the literature.

The rest sections of the paper are structured as follows. In Section 2, we discuss in detail, some basic mathematical properties of the proposed distribution. Section 3 presents some methods of estimation of the unknown parameters of the proposed distribution. The asymptotic behavior of unknown parameters through a Monte Carlo simulation study are investigated in Section 4. In Section 5, we illustrate the applicability of the proposed distribution in lifetime data analysis two data sets and compared its fit alongside with fit attain by some existing non-nested distributions. Finally, in Section 6, we gave a concluding remark.

2. MATHEMATICAL PROPERTIES OF THE IBLL DISTRIBUTION

In this Section, some of the mathematical properties of the IBLL distribution are discussed. These include survival, hazard, quantile functions, the linear representation of the distribution, moments, moment generating function, probability weighted moment, Rényi entropy and distribution of order statistics.

2.1 Survival, Hazard and Quantile Functions

The survival, hazard and quantile functions of the IBLL distribution are respectively derived from (5) and (6) as follows.

$$S(x) = 1 - \left[1 + \left\{\log\left(1 + x^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}, \tag{7}$$

$$h(x) = \frac{\alpha\,\beta\, g(x,\xi)[1 - G(x,\xi)]^{-1}\left\{-\log[1 - G(x,\xi)]\right\}^{-(\alpha+1)}\left[1 + \left\{-\log[1 - G(x,\xi)]\right\}^{-\alpha}\right]^{-(\beta+1)}}{1 - \left[1 + \left\{\log\left(1 + x^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}}, \tag{8}$$

and

$$Q_x(p) = \left[\exp\left(p^{-1/\beta} - 1\right)^{-1/\alpha} - 1\right]^{1/\lambda}, \qquad 0 < p < 1. \tag{9}$$

The quantile function in (9) is derived by simply inverting the distribution function in (5). This is one of the most important properties of any distribution, as it allows for generating random numbers from the distribution for the simulation study. Substituting $p = 0.5$ in (9), we obtain the median of the IBLL distribution as

$$Q_x(0.5) = \left[\exp\left((0.5)^{-1/\beta} - 1\right)^{-1/\alpha} - 1\right]^{1/\lambda}. \tag{10}$$

Some graphical presentations of the pdf and hazard function of the IBLL distribution are displayed in Figures 1 and 2 respectively.

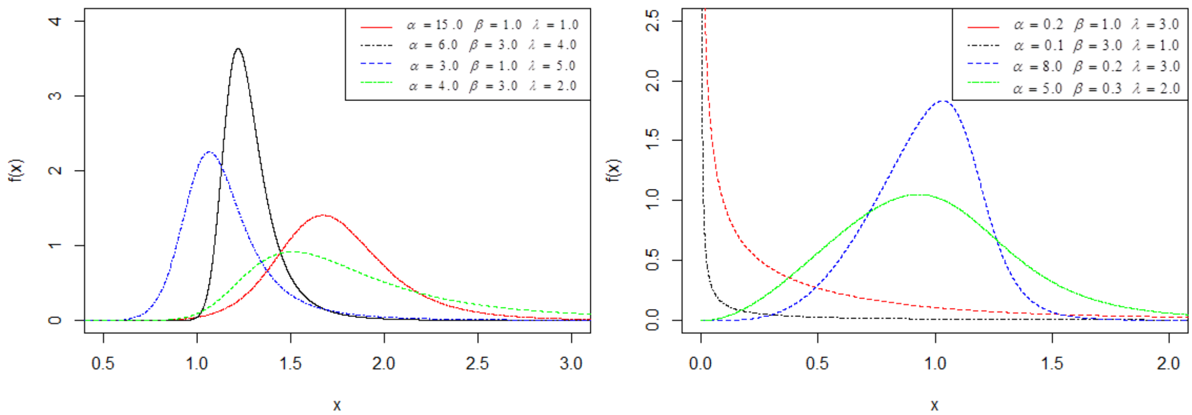


Figure 1: Density Plots of the IBLL Distribution

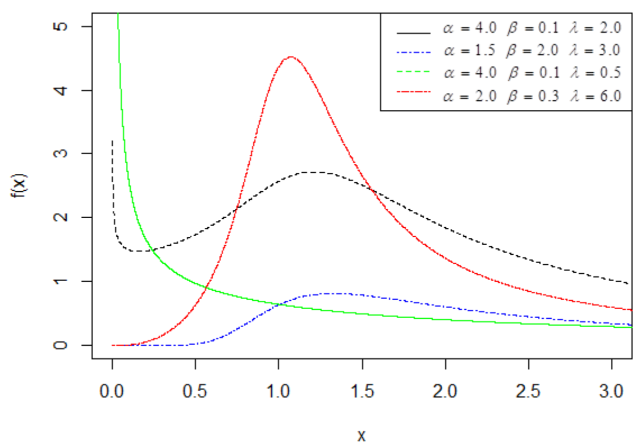


Figure 2: Hazard Plots of the IBLL Distribution

Figure 1 shows that the density plot of the IBLL distribution accommodates decreasing, left-skewed, right-skewed and symmetric shapes, while the plots displayed in Figure 2 indicates that the hazard function of the IBLL distribution exhibits a decreasing, increasing, upside down bathtub and decreasing-increasing-decreasing hazard properties.

2.2 Linear Representation

The linear representation of the density and distribution functions allow for easy derivation of some properties such as the moments, probability weighted moment, moment generating function, distribution of order statistics, etc. The following lemmas will guide us in the derivation of the linear representation of the density and distribution functions of the IBLL distribution.

Lemma 1:
For any positive real non-integer $s > 0$, consider the generalized binomial series expansion (see [12]).

$$(1+x)^{-s} = \sum_{k=0}^{\infty} \binom{s+k-1}{k} (-1)^k x^k .$$

Lemma 2:
For any real parameter $\alpha > 0$, the convergent series holds.

$$(-\log[1-y])^{\alpha-1} = y^{\alpha-1}\left[\sum_{m=0}^{\infty} \binom{\alpha-1}{m} y^m \left(\sum_{s=0}^{\infty} \frac{y^s}{s+2}\right)^m\right], \quad 0<y<1.$$

Applying the result on power series raised to a positive integer, with $a_s = (s+2)^{-1}$ that is,

$$\left(\sum_{s=0}^{\infty} a_s y^s\right)^m = \sum_{s=0}^{\infty} b_{s,m} y^s ,$$

so that,

$$(-\log[1-y])^{\alpha-1} = \sum_{m=0}^{\infty}\sum_{s=0}^{\infty} \binom{\alpha-1}{m} b_{s,m}\, y^{\alpha+m+s-1} ,$$

where $b_{s,m} = (sa_0)^{-1} \sum_{j=0}^{s} \{m(j+1)-s\} a_j b_{s-q,m}$ *and* $b_{0,m} = a_0^m$ (see [8]).

Now, applying the above two lemmas in (5),

$$\left[1+\left(\log[1+x^{\lambda}]\right)^{-\alpha}\right]^{-\beta} = \sum_{k=0}^{\infty} \binom{\beta+k-1}{k} (-1)^k \left(\log[1+x^{\lambda}]\right)^{-\alpha k},$$

$$\left(\log[1+x^{\lambda}]\right)^{-\alpha k} = \sum_{m=0}^{\infty}\sum_{s=0}^{\infty} \binom{-\alpha k}{m} b_{s,m} \left[1-\left(1+x^{\lambda}\right)^{-1}\right]^{m+s-\alpha k} ,$$

so that (5) now becomes,

$$\begin{aligned} F(x) &= \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{s=0}^{\infty} \binom{\beta+k-1}{k}\binom{-\alpha k}{m} (-1)^k b_{s,m} \left[1-\left(1+x^{\lambda}\right)^{-1}\right]^{m+s-\alpha k} \\ &= \sum_{k,m,s=0}^{\infty} \psi_{m+s-\alpha k} H_{m+s-\alpha k}(x) \end{aligned} \qquad (11)$$

where,

$$\psi_{m+s-\alpha k} = \sum_{k,m,s=0}^{\infty} \binom{\beta+k-1}{k}\binom{-\alpha k}{m}(-1)^k b_{s,m}$$

and $H_{m+s-\alpha k}(x) = [G(x)]^{m+s-\alpha k}$ is the distribution function of the log-logistic distribution with $m + s - \alpha k$ as the power parameter.
Differentiating (11), we obtain its associated density function as

$$f(x) = \sum_{k,m,s=0}^{\infty} \psi_{m+s-\alpha k} h_{m+s-\alpha k+1}(x). \tag{12}$$

where $h_{m+s-\alpha k}(x) = \lambda(m+s-\alpha k+1)x^{\lambda-1}(1+x^{\lambda})^{-2}\left[1-(1+x^{\lambda})^{-1}\right]^{m+s-\alpha k}$ is the density function of the log-logistic distribution with $m + s - \alpha k + 1$ as the power parameter.
Other useful properties such as the moments and moment generating function can be directly obtained from (12).

2.3 The Moments and Moment Generating Function

Let *X* be a continuous random variable following a known probability distribution with density function $f(x)$, then the r^{th} ordinary moment of *X* is defined as

$$E[X^r] = \mu_r^{'} = \int_{-\infty}^{\infty} x^r f(x) dx. \tag{13}$$

Substituting (6) into (12), the r^{th} ordinary moment of the IBLL distribution is obtained as

$$\begin{aligned} E[X^r] &= \sum_{k,m,s=0}^{\infty} \psi_{m+s-\alpha k+1} \int_{-\infty}^{\infty} x^r h_{m+s-\alpha k+1}(x) dx, \\ &= \lambda \sum_{k,m,s=0}^{\infty} \psi_{m+s-\alpha k+1}(m+s-\alpha k+1) \int_{-\infty}^{\infty} x^{r+\lambda-1}(1+x^{\lambda})^{-2}\left[1-(1+x^{\lambda})^{-1}\right]^{m+s-\alpha k} dx, \end{aligned} \tag{14}$$

using lemma 1,

$$\left[1-(1+x^{\lambda})^{-1}\right]^{m+s-\alpha k} = \sum_{q=0}^{\infty} \binom{m+s-\alpha k}{q}(-1)^q (1+x^{\lambda})^{-q},$$

Substituting this expression into (14), we have

$$E[X^r] = \lambda \sum_{k,m,s=0}^{\infty} \sum_{q=0}^{\infty} \binom{m+s-\alpha k}{q}(-1)^q (m+s-\alpha k+1)\psi_{m+s-\alpha k+1} \int_{-\infty}^{\infty} x^{r+\lambda-1}(1+x^{\lambda})^{-(q+2)} dx, \tag{15}$$

Further simplification of (15) and invoking the beta function, yields

$$E[X^r] = \sum_{k,m,s,q=0}^{\infty} \binom{m+s-\alpha k}{q}(-1)^q (m+s-\alpha k+1)\psi_{m+s-\alpha k+1} B\left[1+\frac{r}{\lambda}, q+1-\frac{r}{\lambda}\right]. \tag{16}$$

When $r = 1$ in (16) we obtain the mean of *X*. The variance, skewness and kurtosis of *X* can be computed from (16), using the following mathematical relationships.

$$\text{variance}\left(\sigma^2\right) = \mu_2^{'} - \left(\mu_1^{'}\right)^2,$$

$$\text{skewness}\left(S_k\right) = \frac{\mu_3^{'} - 3\mu_2^{'}\mu_1^{'} + 2\left(\mu_1^{'}\right)^3}{\left(\mu_2^{'} - \left(\mu_1^{'}\right)^2\right)^{\frac{3}{2}}},$$

$$\text{kurtosis}\left(K_s\right) = \frac{\mu_4^{'} - 4\mu_3^{'}\mu_1^{'} + 6\mu_3^{'}\left(\mu_1^{'}\right)^2 - 3\left(\mu_1^{'}\right)^4}{\left(\mu_2^{'} - \left(\mu_1^{'}\right)^2\right)^2},$$

where $\mu_1^{'}$, $\mu_2^{'}$, $\mu_3^{'}$ and $\mu_4^{'}$ are the first four ordinary moments of the IBLL distribution.

The r^{th} incomplete moment of *X* is obtained from (16) as

$$\varphi_r(t) = \sum_{k,m,s,q=0}^{\infty} \binom{m+s-\alpha k}{q} (-1)^q (m+s-\alpha k+1)\psi_{m+s-\alpha k+1} B_z\left[1+\frac{r}{\lambda}, q+1-\frac{r}{\lambda}\right], \tag{17}$$

where $B(\alpha,\beta) = \int_0^{\infty} x^{\alpha-1}(1+x)^{-(\beta+\alpha)}dx$ and $B_z(\alpha,\beta) = \int_0^{z} x^{\alpha-1}(1+x)^{-(\beta+\alpha)}dx$ are respectively the beta function of the second kind and the incomplete beta function of the second kind.

The moment generating function of *X* is define using the Maclaurin series expansion of the exponential function as

$$M_X(t) = E\left[e^{tx}\right] = \sum_{n=0}^{\infty} \frac{t^n}{n!} E\left[X^n\right]. \tag{18}$$

Inserting (16) into (18), we define the moment generating function of IBLL distribution as

$$M_X(t) = \sum_{k,m,s,q,n=0}^{\infty} \frac{t^n}{n!} \binom{m+s-\alpha k}{q} (-1)^q (m+s-\alpha k+1)\psi_{m+s-\alpha k+1} B\left[1+\frac{n}{\lambda}, q+1-\frac{n}{\lambda}\right]. \tag{19}$$

Table 1 presents the numerical computation of the mean, variance, skewness and kurtosis of the IBLL distribution at varying values of the parameters.

Table 1: *Moments of the IBLL Distribution at varying values of the Parameters*

λ	α	β	μ	σ^2	S_k	K_s
0.5	6	6	3.2316	11.9924	0.4073	1.5944
		8	2.7794	12.6009	0.6770	1.7675
	8	6	4.3219	10.0142	-0.1784	1.7717
		8	4.1435	11.8485	-0.0735	1.5029
1.0	6	6	3.3973	2.5320	1.1591	5.3969
		8	3.6548	2.8901	0.8689	4.6926
	8	6	2.9264	1.1707	2.0539	10.003
		8	3.1257	1.3177	1.8620	8.8985

From Table 1, we observed that the IBLL distribution is positively skewed ($S_k > 0$), negatively skewed ($S_k < 0$) and approximately symmetric ($S_k \approx 0$). This result is consistent with the plots of the density function displayed in Figure 1. Also, at some selected values of the parameters, the IBLL distribution is both leptokurtic ($K_s > 3$) and platykurtic ($K_s < 3$).
Figure 3 displays the plot of the skewness and kurtosis of IBLL distribution for $\lambda = 1$.

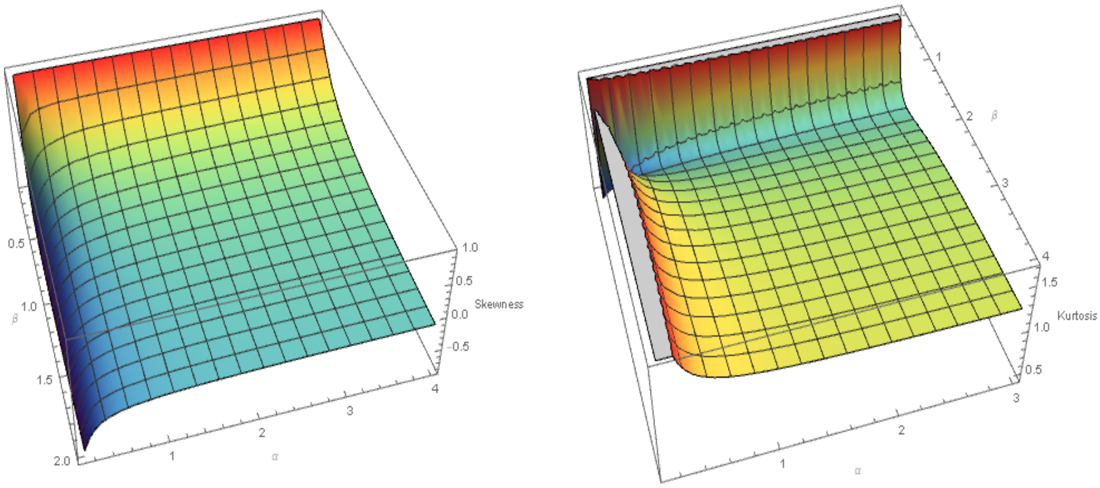


Figure 3: The Skewness and Kurtosis for IBLL Distribution $(\alpha,\ \beta,\ 1)$

2.4 The Probability Weighted Moments

The probability weighted moments (PWMs) are generally used to construct the estimator of the parameters as well as the quantiles of a known statistical distribution whose cdf is invertible. For a random variable X, [9] defined the $(q,r)^{th}$ PWMs as

$$\rho_{q,r} = E\left[X^r F(x)^q\right] = \int_{-\infty}^{\infty} x^r f(x) F(x)^q \, dx, \tag{20}$$

combining (5) and (6), we have

$$f(x)F(x)^q = \alpha \beta \lambda x^{\lambda-1} \left(1+x^{\lambda}\right)^{-1} \left\{\log\left(1+x^{\lambda}\right)\right\}^{-(\alpha+1)} \left[1+\left\{\log\left(1+x^{\lambda}\right)\right\}^{-\alpha}\right]^{-(\beta[q+1]+1)}, \tag{21}$$

applying the lemmas in (21), we have

$$\left[1+\left(\log\left[1+x^{\lambda}\right]\right)^{-\alpha}\right]^{-(\beta[q+1]+1)} = \sum_{k=0}^{\infty} \binom{\beta[q+1]+k}{k} (-1)^k \left(\log\left[1+x^{\lambda}\right]\right)^{-\alpha k},$$

$$\left(\log\left[1+x^{\lambda}\right]\right)^{-(\alpha[k+1]+1)} = \sum_{m=0}^{\infty}\sum_{s=0}^{\infty} \binom{-\alpha[k+1]+1}{m} b_{s,m} \left[1-\left(1+x^{\lambda}\right)^{-1}\right]^{m+s-\alpha[k+1]-1},$$

$$\left[1-\left(1+x^{\lambda}\right)^{-1}\right]^{m+s-\alpha[k+1]-1} = \sum_{p=0}^{\infty} \binom{m+s-\alpha[k+1]-1}{p} (-1)^p \left(1+x^{\lambda}\right)^{-p}.$$

Substituting these expressions into (21), we have

$$f(x)F(x)^q = \alpha\beta\lambda\sum_{m=0}^{\infty}\sum_{s=0}^{\infty}\sum_{k=0}^{\infty}\sum_{p=0}^{\infty}\binom{\beta[q+1]+k}{k}\binom{-\alpha[k+1]+1}{m}\binom{m+s-\alpha[k+1]-1}{p}(-1)^{k+p}b_{s,m}\,x^{\lambda-1}\left(1+x^{\lambda}\right)^{-(p+1)} \quad (22)$$

By inserting (22) into (20) and further simplification, we obtain the PWMs of the IBLL distribution as

$$\rho_{q,r} = \alpha\beta\sum_{m=0}^{\infty}\sum_{s=0}^{\infty}\sum_{k=0}^{\infty}\sum_{p=0}^{\infty}\binom{\beta[q+1]+k}{k}\binom{-\alpha[k+1]+1}{m}\binom{m+s-\alpha[k+1]-1}{p}(-1)^{k+p}b_{s,m}\,B\left[1+\frac{r}{\lambda},p-\frac{r}{\lambda}\right].$$
(23)

From (23), we remark that the PWMs of the inverse Burr log-logistic distribution can be expressed as a linear combination of the log-logistic densities.

2.5 Distribution of Order Statistics

Let $X_1, X_2, \ldots, X_n$ be random samples of size *n* from a known probability distribution. Suppose $X_{r:n}$ denotes the r^{th} order statistics, then the density function of $X_{r:n}$ is defined by

$$f_{r:n}(x) = \frac{1}{B(r,n-r+1)}\sum_{j=0}^{n-r}\binom{n-r}{j}(-1)^j f(x)F(x)^{r+j-1}. \quad (24)$$

Inserting (5) and (6) into (24), we define the r^{th} order statistics of the density of IBLL distribution as follows.

$$f(x)F(x)^{r+j-1} = \alpha\beta\lambda x^{\lambda-1}\left(1+x^{\lambda}\right)^{-1}\left\{\log\left(1+x^{\lambda}\right)\right\}^{-(\alpha+1)}\left[1+\left\{\log\left(1+x^{\lambda}\right)\right\}^{-\alpha}\right]^{-(\beta[r+j]+1)}. \quad (25)$$

We further simplify (25) using a similar approach in (21) as

$$f(x)F(x)^{r+j-1} = \alpha\beta\lambda\sum_{m=0}^{\infty}\varpi_m\,x^{\lambda-1}\left(1+x^{\lambda}\right)^{-(p+1)}, \quad (26)$$

Substituting (26) into (24), we have

$$f_{r:n}(x) = \frac{\alpha\beta\lambda}{B[r,n-r+1]}\sum_{m=0}^{\infty}\sum_{j=0}^{n-r}\binom{n-r}{j}(-1)^j\varpi_m\,x^{\lambda-1}\left(1+x^{\lambda}\right)^{-(p+1)} \quad (27)$$

where

$$\varpi_m = \sum_{s=0}^{\infty}\sum_{k=0}^{\infty}\sum_{p=0}^{\infty}\binom{\beta[q+1]+k}{k}\binom{-\alpha[k+1]+1}{m}\binom{m+s-\alpha[k+1]-1}{p}(-1)^{k+p}b_{s,m}.$$

(27) is readily the r^{th} order statistics of the density function of IBLL distribution.

An expression for the q^{th} moment of the r^{th} order statistics of the density of IBLL distribution is obtained using (27) as

$$E\left[X_{r:n}^{q}\right]= \frac{\alpha\beta}{B[r,n-r+1]}\sum_{m=0}^{\infty}\sum_{j=0}^{n-r}\binom{n-r}{j}(-1)^{j}\varpi_{m}\,B\left[1+\frac{r}{\lambda},p-\frac{r}{\lambda}\right]. \tag{28}$$

Again, we show that the q^{th} moment of the r^{th} order statistics of the density of IBLL distribution can be expressed as a linear combination of the log-logistic densities.

2.6 Rényi Entropy

The entropy of a random variable *X* represents the measure of randomness associated with the random variable *X*. The Rényi entropy of *X* is defined by

$$\tau_R(\gamma)= \frac{1}{1-\gamma}\log\int_{-\infty}^{\infty}f^{\gamma}(x)dx, \qquad \gamma>0,\ \gamma\neq 1. \tag{29}$$

The Rényi entropy of a random variable X following the IBLL distribution is derived by inserting (6) into (29) as

$$\tau_R(\gamma)= \frac{1}{1-\gamma}\log\left[(\alpha\beta\lambda)^{\gamma}\int_{-\infty}^{\infty}x^{\gamma(\lambda-1)}\left(1+x^{\lambda}\right)^{-\gamma}\left\{\log\left(1+x^{\lambda}\right)\right\}^{-\gamma(\alpha+1)}\left[1+\left\{\log\left(1+x^{\lambda}\right)\right\}^{-\alpha}\right]^{-\gamma(\beta+1)}\right]. \tag{30}$$

Applying the lemmas in (30), we have

$$\left[1+\left(\log\left[1+x^{\lambda}\right]\right)^{-\alpha}\right]^{-\gamma(\beta+1)} = \sum_{k=0}^{\infty}\binom{\gamma(\beta+1)+k-1}{k}(-1)^{k}\left(\log\left[1+x^{\lambda}\right]\right)^{-\alpha k},$$

$$\left(\log\left[1+x^{\lambda}\right]\right)^{-(\alpha[k+\gamma]+\gamma)} = \sum_{m=0}^{\infty}\sum_{s=0}^{\infty}\binom{-\alpha[k+\gamma]-\gamma}{m}b_{s,m}\left[1-\left(1+x^{\lambda}\right)^{-1}\right]^{m+s-\alpha[k+\gamma]-\gamma},$$

$$\left[1-\left(1+x^{\lambda}\right)^{-1}\right]^{m+s-\alpha[k+\gamma]-\gamma} = \sum_{p=0}^{\infty}\binom{m+s-\alpha[k+\gamma]-\gamma}{p}(-1)^{p}\left(1+x^{\lambda}\right)^{-p}.$$

Substituting these expressions into (30), we have

$$\tau_R(\gamma) = \frac{1}{1-\gamma}\log\left[(\alpha\beta\lambda)^{\gamma}\sum_{k=0}^{\infty}\omega_k\int_{0}^{\infty}x^{\gamma(\lambda-1)}\left(1+x^{\lambda}\right)^{-(p+\gamma)}dx\right]. \tag{31}$$

Evaluating the integral function in (31) yields,

$$\tau_R(\gamma) = \frac{1}{1-\gamma}\log\left[(\alpha\beta)^{\gamma}\lambda^{\gamma-1}\sum_{k=0}^{\infty}\omega_k\,B\left[\frac{\gamma(\lambda-1)+1}{\lambda},\frac{\gamma+\lambda p-1}{\lambda}\right]\right]. \tag{32}$$

where,

$$\omega_k = \sum_{m=0}^{\infty}\sum_{s=0}^{\infty}\sum_{p=0}^{\infty}\binom{\gamma(\beta+1)+k-1}{k}\binom{-\alpha[k+\gamma]-\gamma}{m}\binom{m+s-\alpha[k+\gamma]-\gamma}{p}(-1)^{k+p}b_{s,m}.$$

3. METHODS OF PARAMETER ESTIMATION

In this section, five estimators, i.e., maximum likelihood, least squares, weighted least squares, Anderson-Darling, and Cramer-von Mises, in order to estimate the unknown parameters of the IBLL distribution are investigated. Let $X_1, X_2, \ldots, X_n$ be a random sample from the $\mathrm{IBLL}(\Xi)$ distribution, $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ represent the associated order statistics and $x_{(i)}$ indicates the observed values of $X_{(i)}$ for $i=1,2,\ldots,n,$ where $\Xi=(\alpha,\beta,\lambda)$. The likelihood and log-likelihood functions are obtained, respectively, by,

$$L(\Xi)=\alpha^n \beta^n \lambda^n \prod_{i=1}^{n} x_i^{\lambda-1}\left(1+x_i^{\lambda}\right)^{-1}\left\{\log\left(1+x_i^{\lambda}\right)\right\}^{-(\alpha+1)}\left[1+\left\{\log\left(1+x_i^{\lambda}\right)\right\}^{-\alpha}\right]^{-(\beta+1)} \tag{33}$$

and

$$\begin{aligned}\ell(\Xi)&=n\log(\alpha)+n\log(\beta)+n\log(\lambda)+(\lambda-1)\sum_{i=1}^{n}\log(x_i)\\&-\sum_{i=1}^{n}\log\left(1+x_i^{\lambda}\right)-(\alpha+1)\sum_{i=1}^{n}\log\left(\log\left(1+x_i^{\lambda}\right)\right)-(\beta+1)\sum_{i=1}^{n}\log\left(1+\left\{\log\left(1+x_i^{\lambda}\right)\right\}^{-\alpha}\right)\end{aligned} \tag{34}$$

Then, the maximum likelihood estimator (MLE) of Ξ is obtained by

$$\hat{\Xi}_1 = \underset{\Xi}{argmax}\,\ell(\Xi). \tag{35}$$

Let us give the following functions that give us the four different estimators:

$$Q_{LS}(\Xi)=\sum_{i=1}^{n}\left(\left[1+\left\{\log\left(1+x_{(i)}^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}-\frac{i}{n+1}\right)^2, \tag{36}$$

$$Q_{WLS}(\Xi)=\sum_{i=1}^{n}\frac{(n+2)(n+1)^2}{i(n-i+1)}\left(\left[1+\left\{\log\left(1+x_{(i)}^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}-\frac{i}{n+1}\right)^2, \tag{37}$$

$$\begin{aligned}Q_{AD}(\Xi)=-n-\sum_{i=1}^{n}\frac{2i-1}{n}\Bigg[&\log\left\{\left[1+\left\{\log\left(1+x_{(i)}^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}\right\}\\&+\log\left\{1-\left[1+\left\{\log\left(1+x_{(i)}^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}\right\}\Bigg]\end{aligned} \tag{38}$$

and

$$QCvM(\Xi)=\frac{1}{12n}+\sum_{i=1}^{n}\left[\left[1+\left\{\log\left(1+x_{(i)}^{\lambda}\right)\right\}^{-\alpha}\right]^{-\beta}-\frac{2i-1}{2n}\right]^2. \tag{39}$$

Then, the least square estimator (LSE), the weighted least square estimator (WLSE), the Anderson-Darling estimator (ADE) and the Cramer-von Mises estimator (CvME) of the Ξ are achieved, respectively, by

$$\hat{\Xi}_2 = \underset{\Xi}{argmin}\,Q_{LS}(\Xi), \tag{40}$$

$$\hat{\Xi}_3 = \underset{\Xi}{argmin}\,Q_{WLS}(\Xi), \tag{41}$$

$$\hat{\Xi}_{44} = \underset{\Xi}{argmin}\,Q_{AD}(\Xi) \tag{42}$$

and

$$\hat{\Xi}_5 = \underset{\Xi}{argmin}\,Q_{CvM}(\Xi). \tag{43}$$

All of the maximization and minimization problems in Equations (35), (40), (41), (42), and (43) can be obtained by **optim** function in the R software.

4. SIMULATION EXPERIMENTS

In this section, the bias and mean square errors (MSEs) of the estimators are calculated with 5000 reputations based on the Monte Carlo simulation. The quantile function given in Equation (9) is used to generate data from the $\text{IBLL}(\Xi)$ distribution by taking $U(0,1)$ instead of p, where $U(0,1)$ is the standard uniform distribution. Eight parameters setting are chosen based on Table 1 as $\Xi=(6,6,0.5)(S1)$, $\Xi=(6,8,0.5)(S2)$, $\Xi=(8,6,0.5)(S3)$, $\Xi=(8,8,0.5)(S4)$, $\Xi=(6,6,1)(S5)$, $\Xi=(6,8,1)(S6)$, $\Xi=(8,6,1)(S7)$ and $\Xi=(8,8,1)(S8)$.

The sample size $n = 50, 100, 150, 200, 250, 500, 1000$ is selected in the simulation experiment. The simulation results are given in Tables 2-3. It can be inferred from Tables 2 and 3 that as sample the size increases, bias and MSEs for all estimators decrease and converge to zero. When the sample size increases, the bias and MSE values of the estimators converge. Although the bias and MSE of the estimators converge with each other when the sample size increases, generally the LSE in bias gives better results than the others.

Table 2: *Average bias for all estimators*

		MLE			LSE			WLSE			ADE			CvME		
Ξ	n	A	B	λ	A	β	λ	α	β	λ	α	β	λ	α	β	λ
S1	50	0.9375	4.1957	0.0704	0.5942	0.2506	0.0135	1.0350	7.8711	0.1362	1.2633	1.1027	0.0256	0.8859	0.2643	0.0002
	100	0.7669	1.2045	0.0116	0.5610	0.0525	-0.0018	0.7021	3.3303	0.0527	0.8007	1.1953	0.0228	0.6974	0.1023	-0.0068
	150	0.7121	0.2743	-0.0095	0.4901	0.0965	0.0011	0.5846	1.3907	0.0236	0.6546	0.7200	0.0108	0.5870	0.1106	-0.0029
	200	0.6581	-0.0165	-0.0167	0.4909	-0.0139	-0.0032	0.5050	0.6594	0.0090	0.5543	0.4239	0.0035	0.5754	-0.0416	-0.0077
	250	0.6475	-0.2638	-0.0238	0.5789	-0.2129	-0.0129	0.4872	0.2729	-0.0004	0.5208	0.1843	-0.0032	0.6399	-0.2247	-0.0160
	500	0.5943	-0.5326	-0.0327	0.3437	0.1976	0.0017	0.4036	-0.1782	-0.0143	0.4037	-0.1478	-0.0131	0.3731	0.1995	0.0003
	1000	0.5817	-0.6845	-0.0380	0.2675	0.0253	-0.0040	0.3801	-0.3884	-0.0218	0.3547	-0.3211	-0.0187	0.2826	0.0241	-0.0047
S2	50	1.1512	4.9371	0.0428	0.3964	0.0485	0.0064	1.1587	10.6314	0.1139	1.4767	0.3899	-0.0102	0.6971	0.1149	-0.0078
	100	0.9526	1.6496	-0.0035	0.5254	-0.1075	-0.0100	0.8302	5.5053	0.0471	0.9739	1.0317	0.0011	0.6672	-0.0897	-0.0162
	150	0.8626	0.3856	-0.0203	0.4366	-0.1390	-0.0078	0.6930	2.9162	0.0208	0.7759	0.9129	0.0003	0.5498	-0.1608	-0.0130
	200	0.8322	-0.2357	-0.0301	0.4681	-0.3456	-0.0151	0.6290	1.0993	0.0021	0.6780	0.5895	-0.0046	0.5559	-0.3655	-0.0193
	250	0.8124	-0.5348	-0.0359	0.5519	-0.5409	-0.0222	0.6234	0.4238	-0.0090	0.6521	0.2444	-0.0114	0.6099	-0.5278	-0.0246
	500	0.7788	-1.1148	-0.0479	0.3783	0.0574	-0.0067	0.5499	-0.5278	-0.0268	0.5343	-0.4227	-0.0236	0.4093	0.0538	-0.0083
	1000	0.7426	-1.2944	-0.0518	0.3472	-0.1384	-0.0121	0.5063	-0.8184	-0.0333	0.4556	-0.6417	-0.0276	0.3623	-0.1365	-0.0128
S3	50	0.6404	6.6985	0.1097	0.6121	0.0227	0.0017	1.0037	11.7552	0.1591	1.4091	0.9300	0.0156	0.9293	0.1141	-0.0063
	100	0.3057	3.7878	0.0679	0.4339	0.4604	0.0109	0.4177	7.2752	0.1038	0.6287	1.8584	0.0394	0.6244	0.5242	0.0068
	150	0.2074	2.3201	0.0453	0.4792	0.2701	0.0040	0.2730	4.4954	0.0717	0.3839	1.9209	0.0406	0.5920	0.2946	0.0011
	200	0.1603	1.6798	0.0345	0.4903	0.1043	-0.0018	0.1926	3.0110	0.0533	0.2641	1.8022	0.0377	0.5702	0.1333	-0.0036
	250	0.1602	1.1944	0.0254	0.5091	-0.0343	-0.0059	0.1992	2.0675	0.0389	0.2532	1.4024	0.0295	0.5717	-0.0091	-0.0073
	500	0.1270	0.4670	0.0100	0.1819	0.5339	0.0128	0.1055	0.8581	0.0196	0.1325	0.7466	0.0169	0.2199	0.5181	0.0112
	1000	0.1233	0.1356	0.0015	0.1057	0.5177	0.0119	0.0784	0.3467	0.0079	0.0893	0.3217	0.0071	0.1247	0.5136	0.0113
S4	50	0.7510	7.9644	0.0905	0.3044	-0.0100	0.0053	1.0016	17.0148	0.1599	1.4711	0.6111	-0.0013	0.6719	0.0439	-0.0061
	100	0.3105	5.8365	0.0692	0.3377	0.1576	0.0021	0.4839	12.2917	0.1128	0.7972	1.4888	0.0181	0.5115	0.2013	-0.0031
	150	0.3283	3.4028	0.0406	0.4856	-0.0614	-0.0064	0.3546	8.6056	0.0796	0.5630	1.8400	0.0227	0.6109	-0.0267	-0.0097
	200	0.2099	2.9491	0.0373	0.4213	-0.0560	-0.0062	0.2381	6.0182	0.0632	0.3530	2.4052	0.0327	0.5157	-0.0249	-0.0085
	250	0.2275	1.8875	0.0244	0.4184	-0.1450	-0.0077	0.2428	4.3372	0.0472	0.3183	2.2085	0.0290	0.4790	-0.0962	-0.0089
	500	0.1828	0.7228	0.0081	0.1918	0.2020	0.0015	0.1329	1.6237	0.0218	0.1614	1.3081	0.0179	0.2385	0.1780	-0.0003
	1000	0.1476	0.2229	0.0007	0.1720	0.4073	0.0044	0.0791	0.6492	0.0096	0.0887	0.6102	0.0088	0.1914	0.4067	0.0038
S5	50	0.6470	5.5044	0.2278	0.5451	0.4398	0.0442	0.9605	8.6682	0.3141	1.0958	1.4447	0.0834	0.8335	0.4091	0.0121
	100	0.3333	2.5247	0.1206	0.4569	0.2635	0.0169	0.4849	4.1211	0.1599	0.5393	1.7448	0.0941	0.6018	0.3162	0.0062
	150	0.2436	1.4603	0.0759	0.4206	0.2336	0.0163	0.3433	2.2477	0.1053	0.3757	1.3905	0.0772	0.5227	0.2319	0.0067
	200	0.1539	1.0055	0.0589	0.4435	0.0369	-0.0027	0.2277	1.4496	0.0796	0.2551	1.1010	0.0633	0.5127	0.0303	-0.0101
	250	0.1581	0.6795	0.0412	0.4363	-0.0472	-0.0088	0.2019	1.0266	0.0603	0.2300	0.8008	0.0485	0.4926	-0.0507	-0.0145
	500	0.0856	0.3212	0.0204	0.1461	0.5693	0.0383	0.0978	0.4501	0.0300	0.1111	0.4035	0.0262	0.1750	0.5688	0.0352
	1000	0.0604	0.1184	0.0067	0.0730	0.3805	0.0257	0.0541	0.1913	0.0133	0.0625	0.1720	0.0115	0.0878	0.3796	0.0242
S6	50	0.6140	7.6142	0.2150	0.3106	0.2057	0.0289	0.9251	12.6867	0.3072	1.1159	1.0707	0.0333	0.5886	0.2560	0.0003
	100	0.3562	4.3665	0.1293	0.3879	0.1270	0.0016	0.5263	7.6389	0.1810	0.6175	2.0318	0.0689	0.5437	0.1983	-0.0100
	150	0.2326	2.6693	0.0876	0.3632	-0.1900	-0.0137	0.3453	4.6530	0.1209	0.3940	2.0958	0.0704	0.4666	-0.1442	-0.0223
	200	0.1950	1.6498	0.0573	0.3831	-0.2584	-0.0208	0.2838	2.8226	0.0844	0.3148	1.6056	0.0552	0.4609	-0.2513	-0.0278
	250	0.1257	1.4433	0.0550	0.3322	-0.1799	-0.0165	0.1810	2.2602	0.0784	0.2129	1.5243	0.0583	0.3892	-0.1762	-0.0221
	500	0.1055	0.5230	0.0188	0.1714	0.4174	0.0160	0.1191	0.8336	0.0312	0.1332	0.7311	0.0268	0.2057	0.3961	0.0118
	1000	0.0716	0.2093	0.0065	0.1057	0.5346	0.0210	0.0632	0.3884	0.0160	0.0725	0.3517	0.0137	0.1214	0.5303	0.0191
S7	50	0.5456	7.2943	0.2396	0.5129	0.1875	0.0157	0.8056	12.7427	0.3507	1.2360	1.1529	0.0444	0.8404	0.3053	-0.0004
	100	0.2945	4.1175	0.1439	0.6107	0.3300	0.0101	0.4796	7.3733	0.2060	0.6888	1.8438	0.0748	0.7686	0.4205	0.0035
	150	0.1463	2.6347	0.1066	0.5751	0.2452	0.0028	0.2828	5.0317	0.1533	0.3928	1.9872	0.0843	0.6942	0.2156	-0.0055
	200	0.1257	1.8119	0.0778	0.5207	0.0253	-0.0082	0.2434	2.8158	0.1010	0.2986	1.7793	0.0742	0.6036	0.0326	-0.0128
	250	0.0829	1.4882	0.0665	0.3985	0.1005	-0.0025	0.1842	2.2312	0.0830	0.2281	1.5328	0.0649	0.4688	0.1045	-0.0066
	500	0.0838	0.6181	0.0287	0.2951	0.2163	0.0051	0.1192	0.8505	0.0375	0.1453	0.7482	0.0323	0.3361	0.1959	0.0018
	1000	0.0161	0.3224	0.0170	0.0801	0.5628	0.0267	0.0286	0.4301	0.0217	0.0454	0.3875	0.0190	0.0986	0.5609	0.0255
S8	50	0.6720	8.8634	0.2082	0.3249	-0.0013	0.0092	1.0153	17.5598	0.3332	1.3750	0.9431	0.0138	0.6687	0.0618	-0.0120
	100	0.2777	6.5039	0.1560	0.4394	0.2362	0.0066	0.5192	12.4466	0.2254	0.8599	1.3023	0.0289	0.6414	0.1311	-0.0095
	150	0.1709	4.2887	0.1125	0.4596	0.0644	-0.0083	0.2647	9.4685	0.1790	0.4694	2.0652	0.0577	0.5743	0.1148	-0.0138
	200	0.1214	3.1787	0.0867	0.4665	-0.1327	-0.0159	0.2429	6.0748	0.1277	0.3481	2.4949	0.0680	0.5472	-0.0944	-0.0200
	250	0.1023	2.5330	0.0718	0.3430	0.1276	-0.0047	0.1813	4.7362	0.1024	0.2613	2.3733	0.0639	0.4220	0.1297	-0.0094
	500	0.0812	1.0523	0.0322	0.2729	-0.0724	-0.0093	0.1200	1.6726	0.0455	0.1514	1.3147	0.0370	0.3129	-0.0768	-0.0120
	1000	0.0486	0.5072	0.0161	0.1995	0.4641	0.0092	0.0792	0.7237	0.0212	0.0927	0.6599	0.0187	0.2190	0.4620	0.0078

Table 3: *Average MSEs for all estimators*

Ξ	n	MLE			LSE			WLSE			ADE			CvME		
		A	B	λ	α	β	λ	α	β	λ	α	β	λ	α	β	λ
S1	50	8.2297	164.7673	0.0863	5.0257	11.1636	0.0284	13.5481	510.6877	0.1856	11.8089	28.9090	0.0530	5.6933	10.612	0.0253
	100	4.1556	45.3542	0.0337	3.6519	7.9128	0.0172	6.0824	245.3384	0.0766	5.8939	32.5961	0.0414	3.9543	8.3217	0.0166
	150	2.8888	16.3444	0.0184	3.2908	6.6964	0.0158	3.9830	64.0081	0.0397	3.9870	21.7640	0.0294	3.5216	6.7264	0.0152
	200	2.1614	8.9423	0.0134	2.9582	5.8082	0.0143	2.8287	29.2083	0.0246	2.8786	14.9217	0.0222	3.1087	5.6329	0.0137
	250	1.7531	5.1202	0.0097	2.8691	5.1904	0.0127	2.2584	12.0601	0.0174	2.3131	9.9242	0.0166	2.9603	5.0758	0.0123
	500	1.0220	2.4029	0.0054	1.6475	6.4392	0.0127	1.1197	3.8997	0.0078	1.1672	4.0569	0.0082	1.6769	6.4718	0.0125
	1000	0.6702	1.4077	0.0035	0.8923	3.8779	0.0075	0.6056	1.7534	0.0039	0.6217	1.8678	0.0041	0.9029	3.8582	0.0074
S2	50	8.8167	217.3302	0.0646	3.0613	10.8334	0.0187	13.7269	786.3757	0.1495	12.2312	31.7471	0.0364	3.5605	12.515	0.0171
	100	4.4242	90.7887	0.0323	2.7898	13.0720	0.0137	6.2139	478.1817	0.0770	6.0499	44.3473	0.0340	2.9681	11.4890	0.0128
	150	3.1118	41.5761	0.0211	2.3110	10.2296	0.0116	4.1584	249.6613	0.0476	4.1946	43.6199	0.0292	2.5483	9.0402	0.0112
	200	2.5062	19.6554	0.0143	2.0040	6.8737	0.0091	3.0617	69.6146	0.0271	3.1435	31.6162	0.0230	2.1611	6.7379	0.0090
	250	2.0678	14.2230	0.0117	2.0892	7.0033	0.0093	2.5515	36.3777	0.0200	2.6280	24.2292	0.0188	2.1822	6.9496	0.0091
	500	1.3012	5.4485	0.0066	1.4503	9.4806	0.0100	1.2919	8.5585	0.0084	1.3441	9.0206	0.0090	1.4850	9.4552	0.0098
	1000	0.9071	3.5782	0.0048	0.9311	7.4636	0.0074	0.7359	4.1769	0.0048	0.7322	4.5383	0.0050	0.9446	7.5026	0.0074
S3	50	12.8929	251.5244	0.0902	6.4170	9.1597	0.0163	21.4541	828.8226	0.1873	17.4964	23.7414	0.0328	7.1741	8.4168	0.0140
	100	6.2261	120.4071	0.0488	5.0548	10.9579	0.0147	9.3815	515.7800	0.1039	8.4520	36.4127	0.0343	5.6795	11.9140	0.0145
	150	4.1157	58.1565	0.0294	4.7730	9.5840	0.0120	6.2919	260.9974	0.0624	5.7657	39.1557	0.0315	4.9602	9.3776	0.0117
	200	3.1126	33.9692	0.0208	4.0023	7.0750	0.0096	4.5631	137.2141	0.0415	4.3428	36.5044	0.0271	4.0819	7.1352	0.0094
	250	2.5849	21.0839	0.0155	3.5837	5.8694	0.0088	3.8436	88.9140	0.0295	3.7146	28.4078	0.0220	3.6651	5.9658	0.0086
	500	1.3443	5.8824	0.0064	2.3211	7.6789	0.0096	1.9815	14.4027	0.0119	1.9674	11.5138	0.0109	2.3334	7.5252	0.0093
	1000	0.6629	2.3354	0.0029	1.4838	6.5431	0.0074	0.9678	4.3373	0.0049	0.9659	4.1395	0.0048	1.4899	6.4957	0.0073
S4	50	12.7339	323.5838	0.0714	4.3827	7.9723	0.0099	21.7192	1370.9550	0.1690	17.0598	29.9717	0.0248	5.0078	7.4320	0.0083
	100	6.1949	233.2069	0.0480	3.4387	11.5590	0.0093	10.0976	1050.0127	0.1098	8.5251	41.0548	0.0258	3.5890	11.5626	0.0088
	150	4.3135	114.9577	0.0294	3.6243	11.2110	0.0084	6.6965	786.8659	0.0729	5.9103	46.8700	0.0239	3.8311	11.7176	0.0083
	200	3.2894	91.4859	0.0245	2.9855	10.5398	0.0077	5.0877	423.5659	0.0518	4.6558	63.3257	0.0259	3.1638	10.8964	0.0077
	250	2.7217	49.9875	0.0172	2.7306	8.8887	0.0071	4.2953	289.5450	0.0391	4.0458	62.1376	0.0235	2.8216	9.0300	0.0070
	500	1.3877	15.3314	0.0074	1.4374	7.0680	0.0054	2.0690	49.0908	0.0146	2.0385	29.1442	0.0125	1.4884	6.9479	0.0052
	1000	0.7350	5.9916	0.0036	1.4175	9.4144	0.0061	1.0540	11.8543	0.0062	1.0497	11.2937	0.0061	1.4281	9.4462	0.0061
S5	50	8.1641	195.7027	0.4446	5.0974	12.6878	0.1316	14.0401	549.4940	0.8162	11.6743	29.8464	0.2232	5.7930	11.6472	0.1094
	100	3.5454	74.3962	0.2060	3.6045	9.2761	0.0764	5.5268	251.6307	0.3654	5.0756	40.0784	0.1980	3.9402	10.0963	0.0746
	150	2.3815	34.2574	0.1209	3.3376	7.5331	0.0704	3.6924	108.0385	0.2098	3.5101	27.8111	0.1456	3.4972	7.3340	0.0670
	200	1.6541	15.8051	0.0797	2.8503	5.8383	0.0564	2.5552	36.9380	0.1343	2.4107	19.5292	0.1090	2.9013	5.6422	0.0531
	250	1.3997	9.0560	0.0595	2.5443	5.2502	0.0516	2.1205	22.5280	0.1001	2.0613	13.3906	0.0848	2.5984	5.1582	0.0496
	500	0.6750	3.4245	0.0268	1.4751	7.3287	0.0553	0.9738	5.7074	0.0412	0.9556	5.2053	0.0389	1.4930	7.3074	0.0543
	1000	0.3345	1.4597	0.0126	0.7874	4.6437	0.0347	0.4722	2.2476	0.0187	0.4699	2.1680	0.0183	0.7923	4.6257	0.0343
S6	50	7.8179	310.9795	0.3789	3.1524	13.3794	0.0889	13.7931	949.0247	0.7276	11.1650	34.4827	0.1613	3.5131	13.2201	0.0762
	100	3.6646	177.4051	0.2185	2.7496	14.5714	0.0580	5.8716	604.6772	0.4157	5.3194	56.4762	0.1663	3.0771	16.3857	0.0592
	150	2.2539	83.5288	0.1373	2.1236	8.1668	0.0425	3.5175	351.2598	0.2552	3.2827	57.1584	0.1426	2.2514	9.5506	0.0429
	200	1.6717	41.3850	0.0887	1.9406	7.3814	0.0383	2.6789	137.1735	0.1655	2.5211	40.1409	0.1089	2.0465	7.5400	0.0381
	250	1.3400	29.7434	0.0714	1.5852	6.9718	0.0360	2.1197	89.1849	0.1257	2.0253	33.1473	0.0929	1.6268	6.9089	0.0351
	500	0.6809	8.6834	0.0308	1.1749	9.5075	0.0411	0.9909	18.2522	0.0507	0.9762	14.8899	0.0473	1.1975	9.4081	0.0403
	1000	0.3562	3.7743	0.0154	0.8043	9.7517	0.0362	0.5063	6.3992	0.0241	0.5032	6.1113	0.0234	0.8103	9.7011	0.0358
S7	50	12.9186	286.8685	0.3886	6.3044	9.3243	0.0684	20.5358	896.3676	0.7858	16.6925	24.0555	0.1260	7.0636	9.8315	0.0612
	100	6.2518	140.0123	0.2125	6.1765	11.6613	0.0624	9.7176	495.3834	0.4156	8.6520	37.3454	0.1412	6.4745	13.3001	0.0617
	150	4.1932	66.0174	0.1328	5.1175	10.5405	0.0526	6.3932	326.9165	0.2768	5.8300	40.5734	0.1309	5.2672	9.7465	0.0491
	200	3.2774	37.2524	0.0905	3.9563	6.8701	0.0395	4.9175	118.5219	0.1600	4.6726	37.1928	0.1107	4.0352	6.7307	0.0383
	250	2.6473	27.0774	0.0716	3.0137	5.7995	0.0334	3.8568	94.0344	0.1251	3.7213	30.0171	0.0914	3.0889	5.7940	0.0325
	500	1.3850	7.6326	0.0304	2.2580	5.8288	0.0312	1.9551	13.9261	0.0473	1.9263	11.9486	0.0439	2.2817	5.6618	0.0302
	1000	0.6859	2.7299	0.0132	1.4708	6.8897	0.0303	0.9413	4.6254	0.0202	0.9364	4.3075	0.0194	1.4775	6.8678	0.0300
S8	50	13.5950	362.1085	0.3166	4.4474	9.4767	0.0446	22.7560	1438.6746	0.7077	17.3753	32.9939	0.1059	4.8009	8.7280	0.0384
	100	6.4227	265.6164	0.2145	4.3368	16.3254	0.0513	10.5993	1111.1386	0.4408	8.9670	38.2959	0.0998	4.6587	14.2683	0.0449
	150	4.3203	140.1165	0.1392	3.5661	14.5328	0.0375	6.6212	864.3592	0.3178	5.7875	48.2825	0.0983	3.7190	14.6579	0.0373
	200	3.3002	97.5383	0.1041	3.2768	10.2293	0.0329	5.2374	430.0210	0.2116	4.8011	66.7278	0.1074	3.3280	10.6299	0.0326
	250	2.7079	68.5264	0.0814	2.4830	10.6151	0.0310	4.0239	343.1026	0.1598	3.7871	63.5251	0.0945	2.6327	10.5669	0.0303
	500	1.4285	19.2085	0.0353	1.6788	6.0184	0.0197	2.1088	52.8106	0.0608	2.0804	29.9620	0.0515	1.7198	6.0311	0.0195
	1000	0.7436	7.1666	0.0165	1.5655	10.8404	0.0273	1.0989	13.4083	0.0267	1.0918	12.3481	0.0255	1.5768	10.8464	0.0271

5. Real Data Analysis

In this section, two applications to the real data sets are examined to demonstrate the applicability of the IBLL distribution. We compare the IBLL distribution with Log-Logistic (LL), Burr III (BIII), Burr XII (BXII), Weibull (W) and Lindley (L) for two data sets. The pdfs of these distributions are given in Table 4. The MLEs of parameters and standard (SE) of MLEs are obtained and reported in Tables 5-6 for two datasets. To select the best distribution some criteria and goodness of-fit statistics such as the estimated log-likelihood values $\hat{\ell}(\Xi)$, Akaike information criteria (AIC), Bayesian information criteria (BIC), consistent Akaike information criteria (CAIC), Hannan–Quinn information criterion (HQIC), Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Cramer von Mises (CvM) statistics and related p values (KS-pval, AD-pval, CvM-pval) are calculated for all distributions. The fitted cdfs for two data sets are plotted in Figures 4 and 5. It is easily seen from Tables 5–6 and Figures 4–5 that the IBLL distribution gives the best modeling for both datasets, according to all criteria. The IBLL distribution can be used and be a good alternative in the literature because of its superior modeling capability.

The first data set represents the remission times (in months) of a random sample of 128 bladder cancer patients and it can be found in [10]. The first data is given by 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39,10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

The second data set represents the survival times (in days) of guinea pigs injected with different amount of tubercle bacilli and it can be consulted detail information in [3]. The second data is given by 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.

Table 4: *The pdfs list of the all distributions*

Distribution	Pdf	Range of the parameters
IBLL	$f(x)=p_1\,p_2\,p_3\,x^{p_3-1}\left(1+x^{p_3}\right)^{-1}\left\{\log\left(1+x^{p_3}\right)\right\}^{-(p_1+1)}\left[1+\left\{\log\left(1+x^{p_3}\right)\right\}^{-p_1}\right]^{-(p_2+1)}$	$p_1,p_2,p_3>0$
LL	$f(x)=p_1\,x^{p_1-1}\left(1+x^{p_1}\right)^{-2}$	$p_1>0$
BIII	$f(x)=p_1p_2x^{-(p_1+1)}\left(1+x^{-p_1}\right)^{-(p_2+1)}$	$p_1,p_2>0$
BXII	$f(x)=p_1p_2x^{p_1-1}\left(1+x^{p_1}\right)^{-(p_2+1)}$	$p_1,p_2>0$
W	$f(x)=(p_1/p_2)(x/p_2)^{p_1-1}\exp\left(-(x/p_2)^{p_1}\right)$	$p_1,p_2>0$
L	$f(x)=p_1^2(x+1)/(1+p_1)\exp(-p_1x)$	$p_1>0$

Table 5: *The modelling results for first data set*

	IBLL	LL	BIII	BXII	W	L
$\hat{\ell}(\Xi)$	-409.8744	-504.8603	-426.6864	-453.5166	-414.0869	-419.5299
AIC	825.7488	1011.7206	857.3729	911.0332	832.1738	841.0598
BIC	834.3048	1014.5726	863.0769	916.7372	837.8778	843.9118
CAIC	825.9423	1011.7523	857.4689	911.1292	832.2698	841.0916
HQIC	829.2251	1012.8794	859.6905	913.3508	834.4913	842.2186
KS	0.0345	0.5260	0.1017	0.2507	0.0700	0.1164
AD	0.1184	63.3436	2.9190	13.3638	0.9577	2.7853
CvM	0.0179	13.4609	0.4508	2.7195	0.1537	0.5191
KS-pval	0.9980	0.0000	0.1413	0.0000	0.5570	0.0623
AD-pval	0.9998	0.0000	0.0302	0.0000	0.3801	0.0353
CVM-pval	0.9986	0.0000	0.0531	0.0000	0.3789	0.0355
$\hat{p}_1$	15.8105	0.7897	1.0333	2.3349	1.0478	0.1960
$\hat{p}_2$	0.4855		4.2070	0.2337	9.5607	
$\hat{p}_3$	0.2312					
SE of $\hat{p}_1$	3.2570	0.0556	0.0604	0.3541	0.0676	0.0123
SE of $\hat{p}_2$	0.1236		0.4054	0.0400	0.8529	
SE of $\hat{p}_3$	0.0182					

Table 6: *The modelling results for second data set*

	IBLL	LL	BIII	BXII	W	L
$\hat{\ell}(\Xi)$	-389.6891	-526.9707	-395.5659	-490.5493	-397.1477	-394.5197
AIC	785.3781	1055.9415	795.1318	985.0986	798.2953	791.0394
BIC	792.2081	1058.2182	799.6852	989.6519	802.8487	793.3160
CAIC	785.7311	1055.9986	795.3057	985.2725	798.4693	791.0965
HQIC	788.0972	1056.8478	796.9445	986.9113	800.1080	791.9457
KS	0.0871	0.7210	0.1512	0.4813	0.1465	0.1326
AD	0.5375	57.3053	1.4907	23.5566	2.3730	1.8706
CvM	0.0898	12.0148	0.2500	5.0353	0.4312	0.3452
KS-pval	0.6450	0.0000	0.0745	0.0000	0.0911	0.1592
AD-pval	0.7083	0.0000	0.1787	0.0000	0.0580	0.1085
CVM-pval	0.6384	0.0000	0.1884	0.0000	0.0596	0.1011
$\hat{p}_1$	29.2029	0.3533	1.4165	48.7608	1.3932	0.0198
$\hat{p}_2$	1.2525		286.9916	0.0047	110.5552	
$\hat{p}_3$	0.1294					
SE of $\hat{p}_1$	5.8536	0.0322	0.1163	18.8737	0.1184	0.0016
SE of $\hat{p}_2$	0.5107		125.6638	0.0017	9.9344	
SE of $\hat{p}_3$	0.0081					

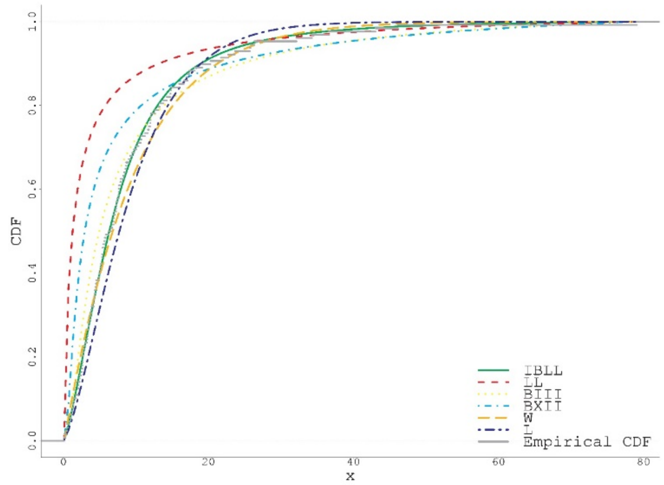


Figure 4: Empirical and fitted cdf plots for the first data

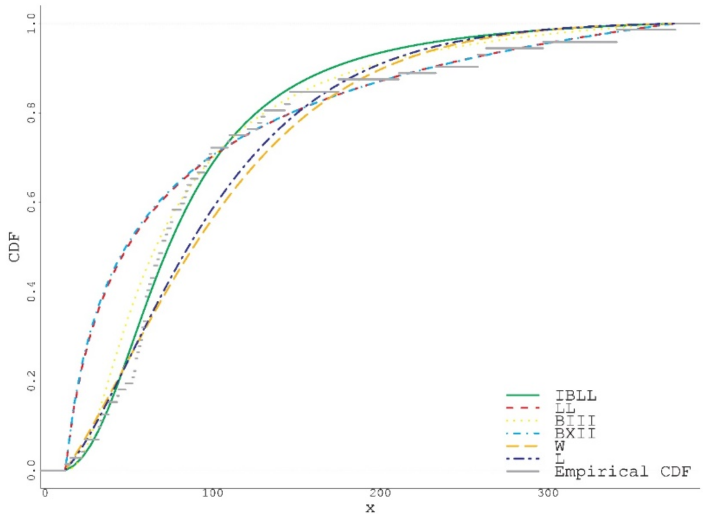


Figure 5: Empirical and fitted cdf plots for the second data

6. CONCLUSION

In this article, a new flexible distribution is introduced. The density and hazard functions of the new model are illustrated by various plots that are very flexible. Many properties related to the distribution have been obtained. Rényi entropy, which is an important measure of randomness, is examined. Some estimation techniques are used to investigate the parameter estimation problem. The performances of the estimators are examined by Monte Carlo simulation. Finally, bladder cancer and survival times of guinea pig data are modeled. As a result of the modeling, it was seen that the best distribution according to all criteria is the IBLL distribution for both data.

References

[1] Alizadeh, M., Cordeiro, G. M., C. Nascimento, A. D., Lima, M. D. C. S., Ortega, E. M. (2017). Odd-Burr generalized family of distributions with some applications. *Journal of Statistical Computation and Simulation*, 1-23.

[2] Alzaartreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1): 63-79.

[3] Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am. J. Hyg.*, 72:130–148.

[4] Bourguignon, M., Silva, R. B., and Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science*, 12: 53-68.

[5] Cordeiro, G. M., and de Castro, M. (2011). A new family of generalized distributions, *Journal of Statistical Computation and Simulation*, 81(7): 883-898.

[6] Elgarhy, M., Nasir, M. A., Farrukh Jamal, F. and Ozel, G., The type II Topp-Leone generated family of distributions: Properties and applications, *Journal of Statistics and Management Systems*, **21** (2018), 1529-1551. https://doi.org/10.1080/09720510.2018.1516725

[7] Eugene, N., Lee, C., and Famoye, F. (2002). The beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, 31(4): 497-512.

[8] Gradshteyn, I. and Ryzhik, I. (2007). Table of Integrals, Series and Products, *Elsevier/Academic Press*.

[9] Greenwood, J. A., Landwehr, J. M., and Matalas, N. C. (1979). Probability weighted moments: Definitions and relations of parameters of several distributions expressible in inverse form. *Water Resources Research*, 15: 1049-1054.

[10] Lee, E. and Wang, J. (2003). Statistical Methods for Survival Data Analysis. *New York: Wiley & Sons.*

[11] Marshall, A. W., and Olkin, I. (1997). A New Method for Adding a Parameter to a Family of Distributions with Application to the Exponential and Weibull Families. *Biometrika,* 84: 641-652.

[12] Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). Integrala and Series, 4. *Gordon and Breach Science Publishers*, Amsterdam.

[13] Uyi, S., Osagie, S. and Osemwenkhae, J. E. (2022). The inverse Burr Generated family of distributions: Properties and applications. *Earthline Journal of Mathematical Sciences*. Forthcoming.

[14] Shaw, W., Buckley, I. (2007). The alchemy of probability distributions: beyond Gram Charlier expansions and a skew-kurtotic normal distribution from a rank transmutation map. arXivprereprint, arXiv, 0901.0434.

[15] Zografos, K. and Balakrishnan, N. (2009). On families of beta and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6, 344–362

RT&A, No 1 (72)
Volume 18, March 2023

ON STRESS STRENGTH RELIABILITY ESTIMATION OF EXPONENTIAL INTERVENED POISSON DISTRIBUTION

K. Jayakumar[1], C. J. Rehana[2]

•

Department of Statistics, University of Calicut, Kerala- 673 635, India.
jkumar19@rediffmail.com[1] and rehanacjoshy1992@gmail.com[2]

Abstract

Aim. Inferences on stress strength reliability has many applications in reliability theory. In this paper, we made a comparative study of Simple random sampling, Ranked set sampling and Percentile ranked set sampling by considering the estimation of stress strength reliability when the stress and strength are independently following Exponential Intervened Poisson distribution. Methods. We used the method of Maximum likelihood estimation for finding the estimate of stress strength reliability. The efficiency of the proposed estimators of stress strength reliability using three sampling schemes are compared via a Monte Carlo simulation study. Also at the end of the study a real life data set is analyzed to understand the usefulness of the study. Results. The findings in this study are the stress strength reliability estimates under Percentile ranked set sampling performs better than the corresponding ones under Simple random sampling and Ranked set sampling. Conclusion. So we can conclude that making refinements in Ranked set sampling increases the efficiency of estimators by minimizing the chance of incorrect ranking.

Keywords: maximum likelihood estimation, percentile ranked set sampling, ranked set sampling, stress strength reliability.

1. Introduction

The estimation of stress strength reliability has applications in a variety of fields like engineering, healthcare, transportation etc. The stress strength reliability is defined as $R = P(X < Y)$, where X is the strength and Y is the applied stress against strength. Obviously the system will fail if the applied stress exceeds the strength of the component. Many researchers are interested to work in this area. A review of the works related to stress strength reliability until 2001 are given in Kotz et al. [10] . Krishnamoorthy et al. [11], Kundu and Gupta [12] and Raqab et al. [20] studied the estimation of R for the Exponential, two-parameter and three-parameter generalized Exponential distributions respectively. Al-Mutairi et al. [3], Ghitany et al. [7] and Rezaei et al. [21] considered the same problem in case of Lindley, power Lindley and generalized Lindley type 5, respectively.

McIntyre [14] introduced the concept of Ranked Set Sampling (RSS). The sampling units in RSS are more representative of population than Simple Random Sampling (SRS) with same sample size. Sengupta and Mukhuti [23] and Muttlak et al. [17] considered the estimation of R when the distribution of stress and strength are Exponential under RSS. Hassan et al. [8] considered the estimation of R under RSS in case of Burr type XII distribution. Akgul and Senoglu [1], Akgul et al. [2] and Al-Omari et al. [5] addressed the same problem in case of Weibull, Lindly and Exponentiated Pareto distribution respectively.

The main characteristic which determines the performance of RSS is the chance of committing error in ranking. The error in ranking increases due to the incorrect measurement of sampling observations. To control this trouble several modifications of RSS have been suggested. see, Samawi et al. [22] suggested Extreme Ranked Set Sampling (ERSS), Muttlak [15] developed Median Ranked Set Sampling (MRSS), Al-Saleh and Al-Kadiri [6] introduced Double Ranked Set Sampling (DRSS). Also Muttlak [16] and Al-Nasser [4] suggested Percentile Ranked Set Sampling (PRSS), L Ranked Set Sampling (LRSS) respectively. Recently Zamanzade and Al-Omari [25] suggested Neoteric Ranked Set Sampling (NRSS).

Intervened distributions has wide range of applications in many areas like life testing experiments, quality control and epidemiological studies etc. Shanmugam [24] developed Intervened Poisson distribution(IPD) to study the effect of some preventive actions or interventions in a system. Recently a family of distributions is generated using IPD, which contain Marshall and Olkin [13] extended families of distribution, families of distributions generated through truncated negative binomial studied by Nadarajah et al. [18] and families of distributions generated through truncated binomial distribution as sub families, see Jayakumar and Sankaran [9]. Also they introduced Exponential Intervened Poisson (EIP) distribution, which is obtained by taking Exponential distribution as the baseline distribution in the above family. Here we consider a comparative study of SRS, RSS and PRSS based on the stress strength reliability estimation of EIP distribution. That is the stress and strength are independently following EIP distribution.

A continuous random variable X on $(0,\infty)$ is said to have an EIP distribution with parameters λ,ρ and θ and write $X \sim$ EIP (λ,ρ,θ) if its probability density function is

$$f(x;\lambda,\rho,\theta) = \frac{\lambda\theta e^{-\theta x}}{e^{\lambda\rho}(e^{\lambda}-1)}\left[(1+\rho)e^{\lambda(1+\rho)e^{-\theta x}} - \rho e^{\lambda\rho e^{-\theta x}}\right] \tag{1}$$

where $\lambda > 0, \rho \geq 0$ and $\theta > 0$.

The cumulative distribution function of X is

$$F(x) = \left[1 - \left(\frac{e^{\lambda(1+\rho)e^{-\theta x}} - e^{\lambda\rho e^{-\theta x}}}{e^{\lambda\rho}(e^{\lambda}-1)}\right)\right]. \tag{2}$$

The corresponding survival (or reliability) and the hazard (or failure rate) functions, at any time $x > 0$, are respectively given by

$$\bar{F}(x) = \left(\frac{e^{\lambda(1+\rho)e^{-\theta x}} - e^{\lambda\rho e^{-\theta x}}}{e^{\lambda\rho}(e^{\lambda}-1)}\right) \tag{3}$$

and

$$h_F(x) = \lambda\theta e^{-\theta x}\left[\frac{1}{(1-e^{-\lambda e^{-\theta x}})} + \rho\right]$$

For a detailed view of properties of EIP distribution, we refer the interested readers to [9]. From [9], we can see that the distribution is under dispersed and leptokurtic. According to the value of the parameters, the distribution behave as positively skewed or negatively skewed.

The rest of this paper is organized as follows: Stress strength reliability for EIP distribution is computed in Section 2. The ML estimation of R based on SRS is considered in section 3. When RSS and PRSS are considered the ML estimation of R are considered in section 4 and section 5 respectively. An extensive Monte-Carlo simulation study is conducted in section 6. In section 7, we present a real data application. Finally conclusions are given in section 8.

2. Stress Strength Reliability

Let X and Y be the stress and strength random variables independently following $EIP(\lambda_1,\rho_1,\theta_1)$ and $EIP(\lambda_2,\rho_2,\theta_2)$, respectively. Then the system reliability is calculated as given below

$$\begin{aligned} R &= P(X < Y) \\ &= \int_0^\infty F_X(x) f_Y(x)dx \end{aligned}$$

$$= \int_0^\infty \left(1 - \left[\frac{e^{\lambda_1(1+\rho_1)e^{-\theta_1 x}} - e^{\lambda_1\rho_1 e^{-\theta_1 x}}}{e^{\lambda_1\rho_1(e^{\lambda_1}-1)}}\right]\right) \frac{\lambda_2\theta_2 e^{-\theta_2 x}}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)} \left[(1+\rho_2)e^{\lambda_2(1+\rho_2)e^{-\theta_2 x}} - \rho_2 e^{\lambda_2\rho_2 e^{-\theta_2 x}}\right] dx \tag{4}$$

We can not solve the above integral directly. Therefore, we use some numerical techniques to solve the equation.

3. Maximum Likelihood Estimation of R based on SRS

To obtain the Maximum likelihood estimates (MLE) of R first we need to find MLE's of the parameters. Let $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_m$ be two independent SRS samples from $EIP(\lambda_1, \rho_1, \theta_1)$ and $EIP(\lambda_2, \rho_2, \theta_2)$, respectively. Then the likelihood function based on SRS is given by,

$$\begin{aligned} L &= \prod_{i=1}^{n} f(x_i) \prod_{j=1}^{m} f(y_j) \\ &= \left(\frac{\lambda_1\theta_1}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)}\right)^n e^{-\sum_{i=1}^n \theta_1 x_i} \prod_{i=1}^{n}\left[(1+\rho_1)e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}} - \rho_1 e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}\right] \\ &\quad \left(\frac{\lambda_2\theta_2}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)}\right)^m e^{-\sum_{j=1}^m \theta_2 y_j} \prod_{j=1}^{m}\left[(1+\rho_2)e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}} - \rho_2 e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}\right] \end{aligned}$$

The log likelihood function is given by,

$$\begin{aligned} &\log L \\ &= n\left[\log\lambda_1 + \log\theta_1 - \lambda_1\rho_1 - \log\left(e^{\lambda_1}-1\right)\right] - \theta_1\sum_{i=1}^{n} x_i + \\ &\sum_{i=1}^{n}\log\left[(1+\rho_1)e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}} - \rho_1 e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}\right] + m\left[\log\lambda_2 + \log\theta_2 - \lambda_2\rho_2 - \log\left(e^{\lambda_2}-1\right)\right] - \\ &\theta_2\sum_{j=1}^{m} y_j + \sum_{j=1}^{m}\log\left[(1+\rho_2)e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}} - \rho_2 e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}\right] \end{aligned}$$

The partial derivatives of the log likelihood function with respect to the parameters are,

$$\frac{\partial \log L}{\partial \lambda_1} = \frac{n}{\lambda_1} - n\rho_1 - \frac{ne^{\lambda_1}}{e^{\lambda_1}-1} + \sum_{i=}^{n} \frac{(1+\rho_1)^2 e^{-\theta_1 x_i} e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}} - \rho_1^2 e^{-\theta_1 x_i} e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}}{(1+\rho_1)e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}} - \rho_1 e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}}$$

$$\frac{\partial \log L}{\partial \rho_1} = -n\lambda_1 + \sum_{i=1}^{n} \frac{e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}}\left(1+\lambda_1(1+\rho_1)e^{-\theta_1 x_i}\right) - e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}\left(1+\lambda_1\rho_1 e^{-\theta_1 x_i}\right)}{(1+\rho_1)e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}} - \rho_1 e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}}$$

$$\frac{\partial \log L}{\partial \theta_1} = \frac{n}{\theta_1} - \sum_{i=1}^{n} \frac{\left[\lambda_1(1+\rho_1)^2 x_i e^{\left(\lambda_1(1+\rho_1)e^{-\theta_1 x_i} - \theta_1 x_i\right)} - \lambda_1\rho_1^2 x_i e^{\left(\lambda_1\rho_1 e^{-\theta_1 x_i} - \theta_1 x_i\right)}\right]}{(1+\rho_1)e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_i}} - \rho_1 e^{\lambda_1\rho_1 e^{-\theta_1 x_i}}}$$

$$\frac{\partial \log L}{\partial \lambda_2} = \frac{m}{\lambda_2} - m\rho_2 - \frac{me^{\lambda_2}}{e^{\lambda_2}-1} + \sum_{j=}^{m} \frac{(1+\rho_2)^2 e^{-\theta_2 y_j} e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}} - \rho_2^2 e^{-\theta_2 y_j} e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}}{(1+\rho_2)e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}} - \rho_2 e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}}$$

$$\frac{\partial \log L}{\partial \rho_2} = -m\lambda_2 + \sum_{j=1}^{m} \frac{e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}}\left(1+\lambda_2(1+\rho_2)e^{-\theta_2 y_j}\right) - e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}\left(1+\lambda_2\rho_2 e^{-\theta_2 y_j}\right)}{(1+\rho_2)e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}} - \rho_2 e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}}$$

$$\frac{\partial \log L}{\partial \theta_2} = \frac{m}{\theta_2} - \sum_{j=1}^{m} \frac{\left[\lambda_2(1+\rho_2)^2 y_j e^{\left(\lambda_2(1+\rho_2)e^{-\theta_2 y_j} - \theta_2 y_j\right)} - \lambda_2\rho_2^2 y_j e^{\left(\lambda_2\rho_2 e^{-\theta_2 y_j} - \theta_2 y_j\right)}\right]}{(1+\rho_2)e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_j}} - \rho_2 e^{\lambda_2\rho_2 e^{-\theta_2 y_j}}}$$

So the ML estimates of the parameters are obtained by maximizing the log-likelihood function with respect to the parameters. Which is equivalent to the simultaneous solution of $\frac{\partial \log L}{\partial \lambda_1} = 0, \frac{\partial \log L}{\partial \rho_1} = 0, \frac{\partial \log L}{\partial \theta_1} = 0, \frac{\partial \log L}{\partial \lambda_2} = 0, \frac{\partial \log L}{\partial \rho_2} = 0$ and $\frac{\partial \log L}{\partial \theta_2} = 0$. The solutions of these equations cannot be obtained in closed form, so we used $optim()$ function in $R\ software$ to solve them numerically. Hence using the invariance property of MLE, the ML estimate of system reliability based on SRS, namely $\hat{R}_{SRS}$, is obtained by substituting the ML estimates of $(\lambda_1, \rho_1, \theta_1, \lambda_2, \rho_2, \theta_2)$ in equation 4.

4. Maximum Likelihood Estimation of R based on RSS

Let $X_{(i)ik}, (i = 1, 2, ..., m_x); (k = 1, 2, ..., r_x)$ be a ranked set sample observed from $EIP(\lambda_1, \rho_1, \theta_1)$ with sample size $n = m_x r_x$, where m_x is the set size and r_x is the number of cycles respectively. Similarly, let $Y_{(j)jl}, (j = 1, 2, ..., m_y); (l = 1, 2, ..., r_y)$ be a ranked set sample observed from $EIP(\lambda_2, \rho_2, \theta_2)$ with sample size $m = m_y r_y$, where m_y is the set size and r_y is the number of cycles respectively. Then the likelihood function based on RSS is given by,

$$
\begin{aligned}
L = & \prod_{k=1}^{r_x}\prod_{i=1}^{m_x} f(x_{ik}) \prod_{l=1}^{r_y}\prod_{j=1}^{m_y} f(y_{jl}) \\
= & C\left[\frac{\lambda_1\theta_1}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)}\right]^n \prod_{k=1}^{r_x}\prod_{i=1}^{m_x}\left[1-\frac{A_{ik}}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)}\right]^{i-1}\left[\frac{A_{ik}}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)}\right]^{m_x-i} e^{-\theta_1 x_{ik}}\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-\rho_1 A_{ik}\right) \\
& \left[\frac{\lambda_2\theta_2}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)}\right]^m \prod_{l=1}^{r_y}\prod_{j=1}^{m_y}\left[1-\frac{B_{jl}}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)}\right]^{j-1}\left[\frac{B_{jl}}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)}\right]^{m_y-j} e^{-\theta_2 y_{jl}}\left(e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}}-\rho_2 B_{jl}\right)
\end{aligned}
$$

where $C = \prod_{k=1}^{r_x}\prod_{i=1}^{m_x}\frac{m_x!}{(i-1)!(m_x-i)!}\prod_{l=1}^{r_y}\prod_{j=1}^{m_y}\frac{m_y!}{(j-1)!(m_y-j)!}$, $A_{ik} = \left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}} - e^{\lambda_1\rho_1 e^{-\theta_1 x_{ik}}}\right)$ and $B_{jl} = \left(e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}} - e^{\lambda_2\rho_2 e^{-\theta_2 y_{jl}}}\right)$
Also $f(x_{ik})$ and $f(y_{jl})$ are defined as,

$$
\begin{aligned}
f(x_{ik}) &= \frac{m_x!}{(i-1)!(m_x-i)!}\left[F_X(x_{ik})\right]^{i-1}\left[1-F_X(x_{ik})\right]^{m_x-i} f_X(x_{ik}) \\
f(y_{jl}) &= \frac{m_y!}{(j-1)!(m_y-j)!}\left[F_Y(x_{jl})\right]^{j-1}\left[1-F_Y(x_{jl})\right]^{m_y-j} f_Y(y_{jl})
\end{aligned}
$$

The log likelihood function is given by,

$$
\begin{aligned}
\log L = & \\
& \log C + n\log\lambda_1 + n\log\theta_1 - n\lambda_1\rho_1 - n\log(e^{\lambda_1}-1) + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(i-1)\log\left[1-\frac{A_{ik}}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)}\right] \\
& + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(m_x-i)\log\left[\frac{A_{ik}}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)}\right] - \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}\theta_1 x_{ik} + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}\log\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-\rho_1 A_{ik}\right) \\
& + m\log\lambda_2 + m\log\theta_2 - m\lambda_2\rho_2 - m\log(e^{\lambda_2}-1) + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}(j-1)\log\left[1-\frac{B_{jl}}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)}\right] \\
& + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}(m_y-j)\log\left[\frac{B_{jl}}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)}\right] - \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}\theta_2 y_{jl} + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}\log\left(e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}}-\rho_2 B_{jl}\right)
\end{aligned}
$$

Then the partial derivatives of the log likelihood function with respect to the parameters are,

$$\begin{aligned}\frac{\partial \log L}{\partial \lambda_1} =& \frac{n}{\lambda_1} - n\rho_1 - \frac{ne^{\lambda_1}}{e^{\lambda_1}-1} - \\ & \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(i-1)\frac{e^{-\theta_1 x_{ik}}(e^{\lambda_1}-1)\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}+\rho_1 A_{ik}\right) - A_{ik}\left(1+\rho_1(e^{\lambda_1}-1)\right)}{\left(e^{\lambda_1\rho_1}(e^{\lambda_1}-1)-A_{ik}\right)(e^{\lambda_1}-1)} \\ & + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(m_x-i)\frac{e^{-\theta_1 x_{ik}}(e^{\lambda_1}-1)\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}+\rho_1 A_{ik}\right) - A_{ik}\left(1+\rho_1(e^{\lambda_1}-1)\right)}{A_{ik}(e^{\lambda_1}-1)} \\ & + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}\frac{e^{-\theta_1 x_{ik}}\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-\rho_1^2 A_{ik}\right)}{e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-\rho_1 A_{ik}}\end{aligned}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \rho_1} =& -n\lambda_1 + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(i-1)\frac{\lambda_1 A_{ik}(1-e^{-\theta_1 x_{ik}})}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)-A_{ik}} + \lambda_1\sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(m_x-i)(1-e^{-\theta_1 x_{ik}}) \\ & + \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}\frac{\lambda_1 e^{-\theta_1 x_{ik}-\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}} - A_{ik}(1+\lambda_1\rho_1 e^{-\theta_1 x_{ik}})}{e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-\rho_1 A_{ik}}\end{aligned}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \theta_1} =& \frac{n}{\theta_1} + \lambda_1\sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(i-1)\frac{x_{ik}e^{-\theta_1 x_{ik}}\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}+\rho_1 A_{ik}\right)}{e^{\lambda_1\rho_1}(e^{\lambda_1}-1)-A_{ik}} \\ & - \lambda_1\sum_{k=1}^{r_x}\sum_{i=1}^{m_x}(m_x-i)\frac{x_{ik}e^{-\theta_1 x_{ik}}\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}+\rho_1 A_{ik}\right)}{A_{ik}} - \sum_{k=1}^{r_x}\sum_{i=1}^{m_x}x_{ik} \\ & + \lambda_1\rho_1^2\sum_{k=1}^{r_x}\sum_{i=1}^{m_x}\frac{x_{ik}e^{-\theta_1 x_{ik}}\left(e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-A_{ik}\right)}{e^{\lambda_1(1+\rho_1)e^{-\theta_1 x_{ik}}}-\rho_1 A_{ik}}\end{aligned}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \lambda_2} =& \frac{m}{\lambda_2} - m\rho_2 - \frac{me^{\lambda_2}}{e^{\lambda_2}-1} - \\ & \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}(j-1)\frac{e^{-\theta_2 y_{jl}}(e^{\lambda_2}-1)\left(e^{\lambda_2(1+\rho_2)e^{-\theta_2 x_{jl}}}+\rho_2 B_{jl}\right) - B_{jl}\left(1+\rho_2(e^{\lambda_2}-1)\right)}{\left(e^{\lambda_2\rho_2}(e^{\lambda_2}-1)-B_{jl}\right)(e^{\lambda_2}-1)} \\ & + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}(m_y-j)\frac{e^{-\theta_2 y_{jl}}(e^{\lambda_2}-1)\left(e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}}+\rho_2 B_{jl}\right) - B_{jl}\left(1+\rho_2(e^{\lambda_2}-1)\right)}{B_{jl}(e^{\lambda_2}-1)} \\ & + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}\frac{e^{-\theta_2 y_{jl}}\left(e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}}-\rho_2^2 B_{jl}\right)}{e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}}-\rho_2 B_{jl}}\end{aligned}$$

$$\begin{aligned}\frac{\partial \log L}{\partial \rho_2} =& -m\lambda_2 + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}(j-1)\frac{\lambda_2 B_{jl}(1-e^{-\theta_2 y_{jl}})}{e^{\lambda_2\rho_2}(e^{\lambda_2}-1)-B_{jl}} + \lambda_2\sum_{l=1}^{r_y}\sum_{j=1}^{m_y}(m_y-j)(1-e^{-\theta_2 y_{jl}}) \\ & + \sum_{l=1}^{r_y}\sum_{j=1}^{m_y}\frac{\lambda_2 e^{-\theta_2 y_{jl}-\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}} - B_{jl}(1+\lambda_2\rho_2 e^{-\theta_2 y_{jl}})}{e^{\lambda_2(1+\rho_2)e^{-\theta_2 y_{jl}}}-\rho_2 B_{jl}}\end{aligned}$$

$$\frac{\partial \log L}{\partial \theta_2} = \frac{m}{\theta_2} + \lambda_2 \sum_{l=1}^{r_y} \sum_{j=1}^{m_y} (j-1) \frac{y_{jl} e^{-\theta_2 y_{jl}} \left(e^{\lambda_2 (1+\rho_2) e^{-\theta_2 y_{jl}}} + \rho_2 B_{jl} \right)}{e^{\lambda_2 \rho_2} (e^{\lambda_2} - 1) - B_{jl}}$$
$$- \lambda_2 \sum_{l=1}^{r_y} \sum_{j=1}^{m_y} (m_y - j) \frac{y_{jl} e^{-\theta_2 y_{jl}} \left(e^{\lambda_2 (1+\rho_2) e^{-\theta_2 y_{jl}}} + \rho_2 B_{jl} \right)}{B_{jl}} - \sum_{l=1}^{r_y} \sum_{j=1}^{m_y} y_{jl}$$
$$+ \lambda_2 \rho_2^2 \sum_{l=1}^{r_y} \sum_{j=1}^{m_y} \frac{y_{jl} e^{-\theta_2 y_{jl}} \left(e^{\lambda_2 (1+\rho_2) e^{-\theta_2 y_{jl}}} - B_{jl} \right)}{e^{\lambda_2 (1+\rho_2) e^{-\theta_2 y_{jl}}} - \rho_2 B_{jl}}$$

The the ML estimates of the unknown parameters under RSS are calculated by equating above equations to zero and solving simultaneously. But it is difficult so solve these equations analytically, so similar to estimation of parameters in SRS, we used $optim()$ function in $R\ software$. Then using the invariance property of MLE, the ML estimate of R based on RSS, namely $\hat{R}_{RSS}$, is obtained by substituting the ML estimates of the parameters in equation (4).

5. Maximum Likelihood Estimation of R based on PRSS

This section deals with the ML estimation of stress strength reliability measure R based on PRSS. Here we consider inference procedure for odd and even set sizes separately.

Case 1: Odd set size Let a_x, b_x, a_y and b_y are the nearest integer values of $p[m_x+1], q[m_x+1], p[m_y+1]$ and $q[m_y+1]$, where $0 < p < 1$ and $q = 1-p$. Also ϑ and ω are defined as $\frac{m_x+1}{2}$ and $\frac{m_y+1}{2}$.

Let $\{X_{(a_x)ik},\ i = 1,2,...,\vartheta-1;\ k = 1,2,...,r_x\} \bigcup \{X_{(\vartheta)ik},\ i = \vartheta;\ k = 1,2,...,r_x\} \bigcup \{X_{(b_x)ik},\ i = \vartheta+1,...,m_x;\ k = 1,2,...,r_x\}$ be the percentile ranked set samples selected from $EIP(\lambda_1, \rho_1, \theta_1)$ with sample size $n = m_x r_x$, where m_x and r_x be the set size and number of cycles respectively. Similarly let $\{Y_{(a_y)jl},\ j = 1,2,...,\omega-1;\ l = 1,2,...,r_y\} \bigcup \{Y_{(\omega)jl},\ j = \omega;\ l = 1,2,...,r_y\} \bigcup \{Y_{(b_y)jl},\ j = \omega+1,...,m_y;\ l = 1,2,...,r_y\}$ be the percentile ranked set samples selected from $EIP(\lambda_2, \rho_2, \theta_2)$ with sample size $m = m_y r_y$, where m_y and r_y be the set size and number of cycles respectively.

Then, the likelihood function is obtained as follows:

$$L = \prod_{k=1}^{r_x} \prod_{i=1}^{\vartheta-1} f(x_{(a_x)ik}) \prod_{k=1}^{r_x} f(x_{(\vartheta)\vartheta k}) \prod_{k=1}^{r_x} \prod_{i=\vartheta+1}^{m_x} f(x_{(b_x)ik})$$
$$\prod_{l=1}^{r_y} \prod_{j=1}^{\omega-1} f(y_{(a_y)jl}) \prod_{l=1}^{r_y} f(y_{(\omega)\omega l}) \prod_{l=1}^{r_y} \prod_{j=\omega+1}^{m_y} f(y_{(b_y)jl})$$

where

$$f(x_{(a_x)}) = \frac{m_x!}{(a_x-1)!(m_x-a_x)!} \left[F_X(x_{a_x})\right]^{a_x-1} \left[1 - F_X(x_{a_x})\right]^{m_x-a_x} f_X(x_{a_x})$$
$$f(x_{(b_x)}) = \frac{m_x!}{(b_x-1)!(m_x-b_x)!} \left[F_X(x_{b_x})\right]^{b_x-1} \left[1 - F_X(x_{b_x})\right]^{m_x-b_x} f_X(x_{b_x})$$
$$f(x_{(\vartheta)}) = \frac{m_x!}{(\vartheta-1)!(m_x-\vartheta)!} \left[F_X(x_{\vartheta})\right]^{\vartheta-1} \left[1 - F_X(x_{\vartheta})\right]^{m_x-\vartheta} f_X(x_{\vartheta})$$

Similarly we can define $f(y_{(a_y)}), f(y_{(b_y)})$ and $f(y_{(\omega)})$.

Case 2: Even set sizes: Here, the reliability estimator is investigated when both X and Y are drawn based on PRSS from EIP with even set size.

Let $\{X_{(a_x)ik},\ i = 1,2,...,\frac{m_x}{2};\ k = 1,2,...,r_x\} \bigcup \{X_{(b_x)ik},\ i = \frac{m_x}{2}+1,...,m_x;\ k = 1,2,...,r_x\}$ and $\{Y_{(a_y)jl},\ j = 1,2,...,\frac{m_y}{2};\ l = 1,2,...,r_y\} \bigcup \{Y_{(b_y)jl},\ j = \frac{m_y}{2}+1,...,m_y;\ l = 1,2,...,r_y\}$ be percentile

ranked set samples from EIP with even set sizes.
Therefore the likelihood function is,

$$L = \prod_{k=1}^{r_x} \prod_{i=1}^{\frac{m_x}{2}} f(x_{(a_x)ik}) \prod_{k=1}^{r_x} \prod_{\frac{m_x}{2}+1}^{m_x} f(x_{(b_x)ik})$$
$$\prod_{l=1}^{r_y} \prod_{j=1}^{\frac{m_y}{2}} f(y_{(a_y)jl}) \prod_{l=1}^{r_y} \prod_{j=\frac{m_y}{2}+1}^{m_y} f(y_{(b_y)jl})$$

For finding the ML estimate of the parameters based on PRSS for both odd and even set sizes, we equate the partial derivatives of the log-likelihood equation to zero and solve them simultaneously. For this we used $optim()$ function in $R\ software$. Hence using the invariance property of MLE, the ML estimate of system reliability based on PRSS, namely $\hat{R}_{PRSS}$, is obtained by substituting the ML estimates in equation (4).

6. Simulation Study

In this section, we conducted a simulation study to assess the potentiality of system reliability estimates based on SRS, RSS and PRSS . We generate 1000 replications of the stress and strength random variables from EIP distribution with parameters $(\lambda_1, \rho_1, \theta_1, \lambda_2, \rho_2, \theta_2) = (.1, .5, 1, 1, 2, 1), (1, .5, 1, 1, 1, 1)$ and $(1, .8, 2, .5, .2, 1)$ using SRS, RSS and PRSS. Using these true values of the parameters we obtain the stress strength reliability R as 0.2634, 0.4518 and 0.7501 respectively. For selecting samples using SRS we set the sample sizes as $(n, m) = (40, 40), (40, 60), (60, 60), (60, 80)$ and $(80, 80)$. Similarly for RSS and PRSS, $(m_x, m_y) = (4, 4), (4, 6), (6, 6), (6, 8), (8, 8)$ and $r_x = r_y = 10$. Also we fix $p = .4$ for PRSS. From these generated samples we compute the estimates of stress strength reliability. Mean square error (MSE) and Relative efficiency (RE) are used to compare the estimated stress strength reliability measures. The results are reported in Table 1. In this table, RE_1, RE_2 and RE_3 is the relative efficiency of RSS over SRS, PRSS over SRS and PRSS over RSS respectively. For all sampling methods, the MSE decreases when the sample size increases, which indicates the consistency property of MLE. According to the values of relative efficiencies we can say that RSS and PRSS performs better than SRS in all cases. Moreover PRSS performs better than RSS in almost everywhere.

Table 1: *Bias, MSE and RE of $\hat{R}$ based on SRS, RSS and PRSS.*

			SRS		RSS		PRSS				
R	(m_x, m_y)	(n, m)	Bias	MSE	Bias	MSE	Bias	MSE	RE_1	RE_2	RE_3
0.2634	(4,4)	(40,40)	-0.0014	0.0020	-0.0083	0.0014	-0.0066	0.0012	1.47	1.69	1.15
	(4,6)	(40,60)	0.0035	0.0017	-0.0032	0.0011	-0.0078	0.0009	1.51	1.87	1.24
	(6,6)	(60,60)	0.0162	0.0011	-0.0006	0.0007	-0.0067	0.0006	1.64	1.91	1.16
	(6,8)	(60,80)	-0.0003	0.0010	-0.0035	0.0006	-0.0069	0.0005	1.86	1.91	1.03
	(8,8)	(80,80)	0.0004	0.0009	-0.0038	0.0005	-0.0053	0.0004	1.97	2.12	1.08
0.4518	(4,4)	(40,40)	-0.0030	0.0025	-0.0013	0.0019	-0.0023	0.0019	1.31	1.37	1.04
	(4,6)	(40,60)	-0.0200	0.0023	-0.0010	0.0015	-0.0017	0.0014	1.49	1.62	1.09
	(6,6)	(60,60)	-0.0325	0.0018	-0.0007	0.0011	0.0004	0.0010	1.69	1.84	1.09
	(6,8)	(60,80)	-0.0284	0.0017	-0.0004	0.0010	-0.0016	0.0009	1.74	1.95	1.12
	(8,8)	(80,80)	-0.0351	0.0014	0.0004	0.0007	0.0003	0.0007	1.97	2.03	1.03
0.7502	(4,4)	(40,40)	-0.0023	0.0018	0.0049	0.0013	0.0073	0.0013	1.37	1.44	1.05
	(4,6)	(40,60)	-0.0101	0.0014	0.0009	0.0010	0.0061	0.0009	1.41	1.62	1.15
	(6,6)	(60,60)	0.0012	0.0013	0.0037	0.0008	0.0064	0.0007	1.66	1.88	1.13
	(6,8)	(60,80)	-0.0227	0.0013	0.0017	0.0007	0.0049	0.0005	1.82	2.32	1.28
	(8,8)	(80,80)	-0.0146	0.0011	0.0015	0.0005	0.0045	0.0005	2.19	2.38	1.09

7. Data Analysis

Here we analyzed a real life data set to illustrate the use of our proposed methodology. We consider two real life data sets which contain times to breakdown of an insulating fluid between electrodes recorded at different voltages see, [19]. These are the failure times (in minutes) for an insulating fluid between two electrodes subject to a voltage of 34 kV (X) and 36 kV (Y) are given in Table 2 and Table 3.

Table 2: *Data X: (34 kV)*

0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67	4.85	6.5
7.35	8.01	8.27	12.06	31.75	32.52	33.91	36.71	72.89	

Table 3: *Data Y: (36 kV)*

0.35	0.59	0.96	0.99	1.69	1.97	2.07	2.58	2.71	2.9
3.67	3.99	5.35	13.77	25.50					

Now to identify the behaviour of the hazard rate function of the data, we examined total time on test transform plot of the data sets. For this we use $TTT()$ function in $R\ Software$. The total time on test transform plots for both data sets are given in Figure1 and Figure 2. From these figures we can say that the hazard rate function of both data sets show decreasing nature.

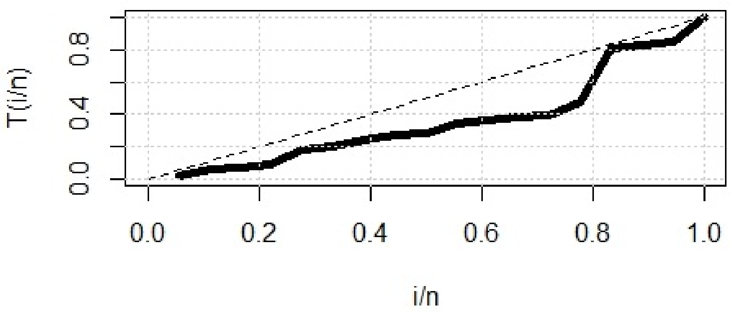


Figure 1: *The scaled TTT plot of Data X.*

Moreover the hazard rate function of EIP distribution also shows decreasing behaviour, see [9]. So we fit EIP distribution for both data sets separately. For fitting, we first find MLE's of the parameters. Also we need to check the goodness of fit of the NGP distribution for the data. For this purpose we use $-\log L$ and Kolmogorov Smirnov (KS) statistic along with p-value. The values of the estimated parameters, $-\log L$, KS, p value for both the data sets are reported in Table 4.

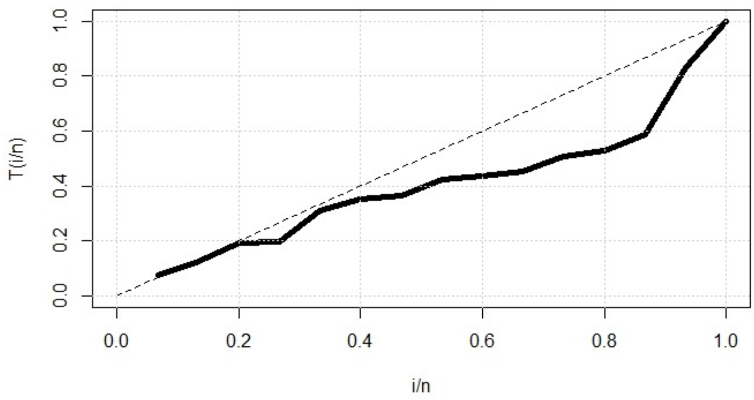


Figure 2: *The scaled TTT plot of Data Y.*

Table 4: *Estimates of the parameters, -log L, KS and P values for data sets.*

Data Set	Sample Size	λ	ρ	θ	$-\log L$	K-S	p value
X	19	0.96842829	0.54290810	0.04633776	68.54817	0.16834	0.5963
Y	15	0.9729146	0.9806698	0.1147315	36.97626	0.16411	0.7559

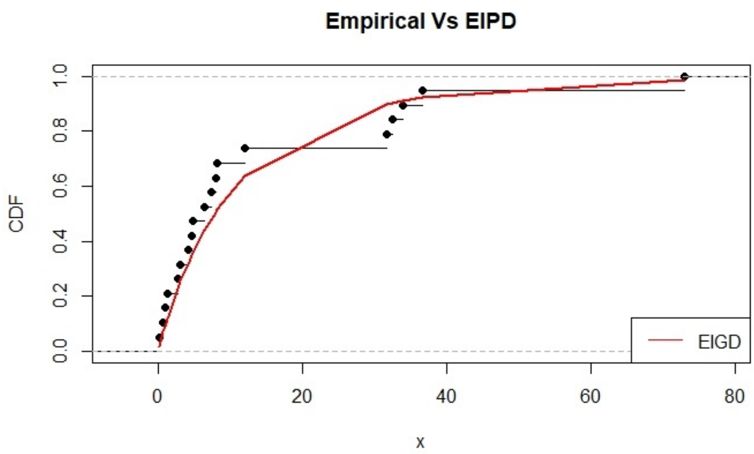


Figure 3: *The empirical distribution function and fitted distribution functions for Data X.*

From Table 4 and Figures 2 and 3, we can say that EIP distribution fits well for both data sets. So we are choosing these data sets to select samples from EIP distribution based on SRS, RSS and PRSS. For selecting the samples via SRS we take the sample sizes for X and Y as $n = 12, m = 8$. In case of RSS and PRSS, we take $m_x = 4$ and $r_x = 3$ for data X and $m_y = 2$ and $r_y = 4$ for data Y. Also R based on $n = 19$ and $m = 15$ observations is calculated as 0.27257. The mean, bias and MSEs of the estimates of R based on 10,000 replications of each sampling method is given in Table 5.

From Table 5 we can say that the estimated values of R based on $n = 12$ and $m = 8$ sampling

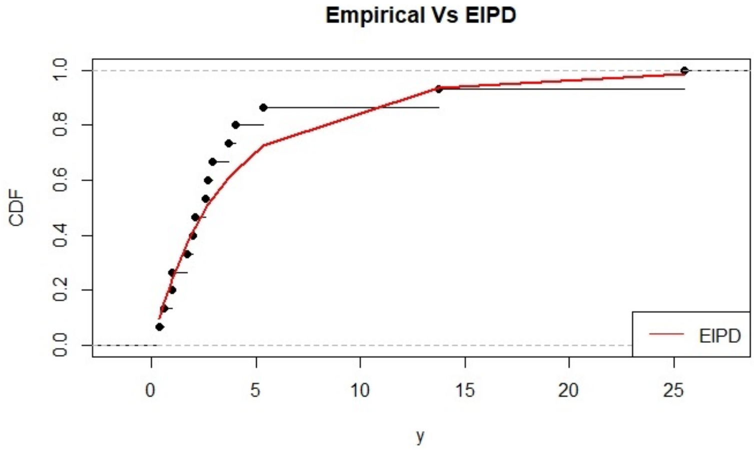


Figure 4: *The empirical distribution function and fitted distribution functions for Data Y.*

units using SRS, RSS and PRSS are close to the estimated value of R calculated from the entire data set. However, in view of MSEs, we can seen that $\hat{R}_{PRSS}$ and $\hat{R}_{RSS}$ perform better than $\hat{R}_{SRS}$.

8. Conclusions

In this paper, the ML estimates of the stress strength reliability R based on SRS, RSS and PRSS are obtained, when the stress and strength are independently following EIP distribution. The performance of the proposed estimators are compared using a Monte Carlo simulation study. From the simulation study it is clear that PRSS performs better than RSS and SRS. Also we can see that, the efficiency of all estimates increases as the set size increases. The results from the simulation study is supported by a real life data set. So if our aim is to choose a sampling procedure which minimizes the error in ranking, then we can consider PRSS than RSS and SRS.

9. Funding

This work was supported by the University Grants Commission, New Delhi, India [Junior Research Fellowship].

References

[1] Akgul, F. G. and Senoglu, B. (2017). Estimation of $P(X < Y)$ using ranked set sampling for the Weibull distribution. *Quality Technology and Quantitative Management* 14: 296-309.

[2] Akgul, F. G. , Acitas, S. and Senoglu, B. (2018). Inference on stress-strength reliability based on ranked set sampling data in case of Lindley distribution. *Journal of Statistical Computation and Simulation*,8 : 3018 - 32. DOI:10.1080/00949655.2018.1498095

[3] Al-Mutairi, D. K., Ghitany, M. E. and Kundu, D. (2013). Inferences on stress-strength reliability from Lindley distribution. *Communications in Statistics - Theory and Methods*, 42 : 1443 - 63. DOI:10.1080/03610926.2011.563011.

[4] Al-Nasser, A. D. (2007). L ranked set sampling: A generalized procedure for robust visual sampling. *Communications in Statistics - Simulation and Computation*, 36 : 33 - 43. DOI:10.1080/03610910601096510.

[5] Al-Omari, A. M., Almanjahie, I. M., Hassan, A. S. and Nagy, H. F. (2020). Estimation of the Stress-Strength Reliability for Exponentiated Pareto Distribution Using Median and Ranked Set Sampling Methods. *Computers, Materials and Continua.*, 64 : 835 - 857.

[6] Al-Saleh, M. F. and Al-Kadiri, M. (2000). Double ranked set sampling. *Statistics and Probability Letters*, 48 : 205 - 212.

[7] Ghitany, M. E., Al-Mutairi, D. K. and Aboukhamseen, S. M. (2015). Estimation of the reliability of a stress-strength system from power Lindley distributions. *Communications in Statistics -Simulation and Computation*, 44 : 118 - 36. DOI:10.1080/03610918.2013.767910.

[8] Hassan, A. S., Assar, S. M. and Yahya, M. (2015). Estimation of for Burr Type XII Distribution under Several Modifications for Ranked Set Sampling. *Australian Journal of Basic and Applied Sciences*, 9 : 124 - 138.

[9] Jayakumar, K. and Sankaran, K. K. (2019). Exponential intervened Poisson distribution. *Communications in Statistics - Theory and Methods*, 50 : 3063 - 3093. DOI:10.1080/03610926.2019.1682161.

[10] Kotz, S., Lumelskii, Y. and Pensky, M. The Stress Strength Model and Its Generalizations Theory and Applications, World scientific, 2003.

[11] Krishnamoorthy, K., Mukherjee, S. and Guo, H. (2007). Inference on reliability in two-parameter exponential stress-strength model. *Metrika*, 65 : 261 - 273.

[12] Kundu, D. and Gupta, R.D. (2005). Estimation of $P[Y < X]$ for generalized exponential distribution. *Metrika*, 61 : 291 - 308.

[13] Marshall, A. W., and I. Olkin. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*. 84 : 641 - 52.

[14] McIntyre, G. A. (1952). A method of unbiased selective sampling using ranked sets. *Australian Journal of Agricultural Research*, 3 : 385 - 390.

[15] Muttlak, H. A. (1997). Median ranked set sampling. *Journal of Applied Statistical Science*, 6 : 245 - 255.

[16] Muttlak, H. A. (2003). Modified ranked set sampling methods. *Pakistan Journal of Statistics*, 19 : 315 - 23.

[17] Muttlak, H. A., Abu-Dayyeh, W. A., Saleh, M. F. and Al-Sawi, E. (2010). Estimating $P(Y < X)$ using ranked set sampling in case of the exponential distribution. *Communications in Statistics Theory Methods*, 39: 1855 - 68.

[18] Nadarajah, S., K. Jayakumar and M. M. Ristic. (2013). A new family of lifetime models. *Journal of Statistical Computation and Simulation*. 83 : 1389 - 404.

[19] Nelson, W. Applied Life Data Analysis, Wiley, NewYork, 1982.

[20] Raqab, M.Z., Madi, M.D. and Kundu, D. (2008). Estimation of $P(Y < X)$ for the 3-parameter generalized exponential distribution. *Communications in Statistics-Theory and Methods*, 37: 2854 - 2864.

[21] Rezaei, A., Sharafi, M., Behboodian, J. and Zamani, A. (2018). Inferences on stress-strength parameter based on GLD5 distribution. *Communications in Statistics - Simulation and Computation*, 47 : 1251 - 63. DOI:10.1080/03610918.2017.1309666.

[22] Samwi, H., Ahmad, M. and Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. *Biometrical Journal*, 38 : 577 - 586.

[23] Sengupta, S., and Mukhuti, S. (2008). Unbiased estimation of $P(X > Y)$ using ranked set sampling data. *Statistics*, 42 : 223 - 30. DOI:10.1080/02331880701823271.

[24] Shanmugam, R. (1985). An intervened Poisson distribution and its medical application. *Biometrics*, 41: 1025 - 29.

[25] Zamanzade, E. and Al-omari, A. I. (2016). New ranked set sampling for estimating the population mean and variance. *Hacettepe journal of Matematics and Statistics*, 45 : 1891 - 1905.

A DISCRETE PARAMETRIC MARKOV-CHAIN SYSTEM MODEL OF A TWO-UNIT STANDBY SYSTEM WITH TWO TYPES OF REPAIR

Laxmi Raghuvanshi, Rakesh Gupta and Pradeep Chaudhary
•
Department of Statistics
Ch. Charan Singh University , Meerut-250004 (India)
laxya.raghav@gmail.com; smprgccsu@gmail.com; pc25jan@ gmail.com

Abstract

The aim of the present paper is to deal with the cost-benefit analysis of a two identical unit cold standby system model. There are two modes of a unit say Normal(N) and Total failure(F). When a unit operates then any of the two types of failure minor or major may occur some fixed known probabilities. Two repairmen are always available with the system to repair a unit failed with minor or major fault respectively. Upon failure of an operative unit the cold standby unit starts operations instantaneously with a perfect switching device. After minor and major repair of a failed unit it becomes as good as new. The distributions of failure times of minor and major faults and each type of repair time are assumed to follow geometric distributions with different parameters. Using regenerative point technique with the basic tools of probabilistic argument and Laplace Transform various important measures of system effectiveness useful to system designers and operations managers have been obtained.

Keywords: Cold standby, transition probabilities, mean sojourn time, regenerative pint, MTSF, geometric distribution.

1. Introduction

The stochastic models pertaining to two-unit standby redundant systems have been frequently analysed in the field of reliability theory due to their wide applicability in modern business and industrial units. The consideration of repair is one of the important criteria to enhance the system effectiveness such as reliability, expected life and availability of the system etc. Various authors during past many decades have analysed the two identical and non-identical unit standby system models by using different repair policies. Some of the authors have analysed models by assuming two types of failure say minor and major in an operating unit and so accordingly they have considered two types of repair with different repair time distributions.

Levitin et al. [7] studied a series-parallel reparable system model with two types of failure states-failure safe and failure dangerous. Ram and Singh [8] performed the stochastic analysis of a complex system model assuming that a unit can fail in n-mutually exclusive ways of total failure or common cause failure. Choudhary and Kumar [1] analysed a system model consisting of two units-one is main unit and other is supporting unit. The main unit passes through two types of failure-partial failure and total failure. A single repairman is always available with the system for the repair of each type of failure. Gupta and Vinodiya [6] have analysed a two non-identical unit cold standby system model assuming two types of failure in one of the unit. They have considered a single repairman for the repair of each type of failures. Most recently Chaudhary and Tyagi [2] analysed a two non-identical unit parallel system model assuming that one of the unit can fail either due to hardware or due to human error. All the above system models are based on the continuous parametric Markov-chain.

The purpose of the present study is to analyse a stochastic model based on discrete parametric Markov-chain system composed of two identical units in cold standby configuration. Two types of failure (minor and major) have been considered in an operating unit. Two different repairmen are always available with the system. One is considered to attend a failed unit due to minor fault and other is to attend a failed unit due to major fault. Some authors including [3-5] infact analysed the system models by taking geometric distributions of failure and repair times but not much work is done in this direction. By using regenerative point technique the following economic related measures of system effective are obtained -

- Transition probabilities and mean sojourn times.
- Reliability and mean time to system failure.
- Point-wise and steady-state availabilities of the system and expected up time of the system by the epoch (t-1).
- Expected busy period of the repairman by the epoch (t-1).
- Net expected profit incurred by the system by the epoch (t-1) and in steady state.

2. Model Description and Assumptions

The system under study is based on the following assumptions

i. The system is composed of two identical units. Initially one unit is operative and other is kept as cold standby. The cold standby never lose its operational ability in its standby state.

ii. Each unit has two modes: Normal (N) and Total failure (F).

iii. Two types of failure minor and major may occur in an operating unit with respective probabilities a and b. (a+b=1).

iv. Two different repairmen are always available with the system. One is ordinary repairman for minor fault and other is skilled repairman for major fault in an operating unit.

v. Upon either type of failure in an operating unit, the standby unit starts working immediately with a perfect and instantaneous switching device.

vi. After each type of repair, a unit becomes as good as new.

vii. The time to failure and each type of repair time follow geometric distribution with different parameters.

3. Notations and States of the System

3.1 Notations used in the paper :

pq^x : p.m.f. of failure time of a unit $(p+q=1)$.

$r_i s_i^x$: p.m.f. of repair time of minor and major types respectively for i=1 and 2.

a, b : probability that the failed unit requires minor and major repairs. $(a+b=1)$.

$q_{ij}(\cdot), Q_{ij}(\cdot)$: p.m.f. and c.d.f. of one step or direct transition time from state S_i to S_j.

p_{ij} : Steady state transition probability from state S_i to S_j.

$$p_{ij} = Q_{ij}(\infty)$$

$Z_i(t)$: Probability that the system sojourn in state S_i up to the cycles 0,1,2,……..t-1.

ψ_i : Mean sojourn time in state S_i.

$*, h$: Symbol and dummy variable used in geometric transform e. g.

$$GT\left[q_{ij}(t)\right] = q_{ij}^*(h) = \sum_{t=0}^{\infty} h^t q_{ij}(t)$$

© : Symbol of ordinary convolution i.e.,

$$A(t)©B(t) = \int_0^t A(u)B(t-u)du$$

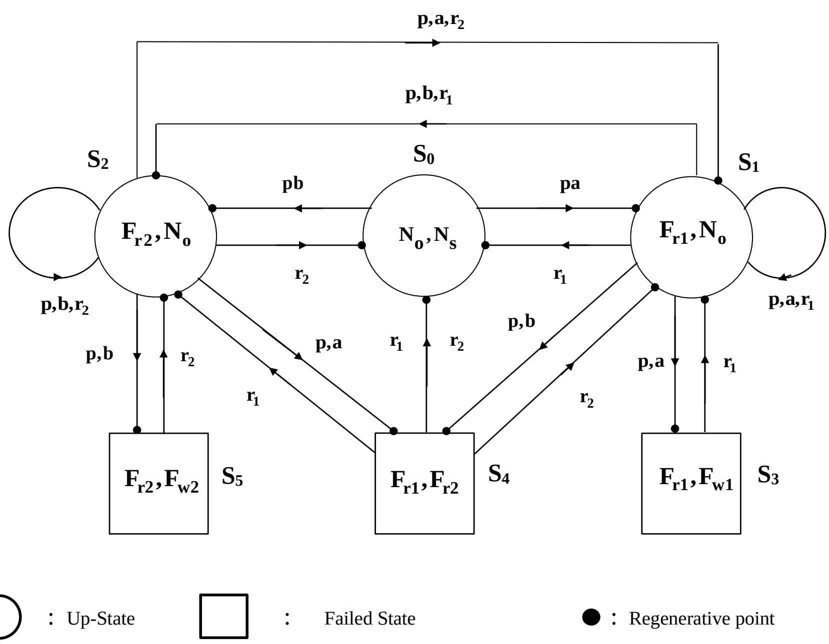


Figure.1 *Transition Diagram*

3.2 Symbols used for the states of the system:

N_o , N_s : Unit in normal mode and operative/cold standby.

F_{r1}, F_{w1} : Unit in total failure mode and under minor repair/waiting for minor repair.

F_{r2}, F_{w2} : Unit in total failure mode and under major repair/waiting for major repair.

With the help of above symbols and keeping into consideration the stated assumptions, the possible states of the system are as follows-

Up States: $S_0 \equiv (N_0, N_s)$, $S_1 \equiv (F_{r1}, N_o)$, $S_2 \equiv (F_{r2}, N_o)$

Failed States: $S_3 \equiv (F_{r1}, F_{w1})$, $S_4 \equiv (F_{r1}, F_{r2})$, $S_5 \equiv (F_{r2}, F_{w2})$

The transition diagram of the system model alongwith the transition rates is shown in fig.1. We observe that all the entrance epochs into the states are regenerative epochs.

4. Transition Probabilities

Let $Q_{ij}(t)$ be the probability that the system transits from state S_i to S_j during time interval (0, t) i.e., if T_{ij} is the transition time from state S_i to S_j then

$$Q_{ij}(t) = P\left[T_{ij} \le t\right]$$

By considering the elementary probabilistic arguments we have,

$$Q_{01}(t) = \sum_{u=0}^{t} apq^u = a\left(1-q^{t+1}\right) \tag{1}$$

Similarly,

$$Q_{02}(t) = \sum_{u=0}^{t} bpq^u = b\left(1-q^{t+1}\right) \tag{2}$$

$$Q_{10}(t) = \frac{r_1 q}{(1-qs_1)}\left[1-(qs_1)^{t+1}\right] \tag{3}$$

$$Q_{11}(t) = \frac{apr_1}{(1-qs_1)}\left[1-(qs_1)^{t+1}\right] \tag{4}$$

$$Q_{12}(t) = \frac{bpr_1}{(1-qs_1)}\left[1-(qs_1)^{t+1}\right] \tag{5}$$

$$Q_{13}(t) = \frac{aps_1}{(1-qs_1)}\left[1-(qs_1)^{t+1}\right] \tag{6}$$

$$Q_{15}(t) = \frac{bps_1}{(1-qs_1)}\left[1-(qs_1)^{t+1}\right] \tag{7}$$

$$Q_{20}(t) = \frac{r_2 q}{(1-qs_2)}\left[1-(qs_2)^{t+1}\right] \tag{8}$$

$$Q_{21}(t) = \frac{apr_2}{(1-qs_2)}\left[1-(qs_2)^{t+1}\right] \tag{9}$$

$$Q_{22}(t) = \frac{bpr_2}{(1-qs_2)}\left[1-(qs_2)^{t+1}\right] \tag{10}$$

$$Q_{24}(t) = \frac{bps_2}{(1-qs_2)}\left[1-(qs_2)^{t+1}\right] \tag{11}$$

$$Q_{25}(t) = \frac{aps_2}{(1-qs_2)}\left[1-(qs_2)^{t+1}\right] \tag{12}$$

$$Q_{31}(t) = \left[1-(s_1)^{t+1}\right] \tag{13}$$

$$Q_{42}(t)=\left[1-(s_2)^{t+1}\right] \quad (14)$$

$$Q_{50}(t)=\frac{r_1r_2}{(1-s_1s_2)}\left[1-(s_1s_2)^{t+1}\right] \quad (15)$$

$$Q_{51}(t)=\frac{r_2s_1}{(1-s_1s_2)}\left[1-(s_1s_2)^{t+1}\right] \quad (16)$$

$$Q_{52}(t)=\frac{r_1s_2}{(1-s_1s_2)}\left[1-(s_1s_2)^{t+1}\right] \quad (17)$$

From the transient-state transition probabilities (1-17), the steady state t.p.m can be obtained as follows

$$((p_{ij}))=\begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & p_{05} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} & p_{25} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} & p_{35} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} & p_{45} \\ p_{50} & p_{51} & p_{53} & p_{54} & p_{55} & p_{56} \end{pmatrix}$$

$$p_{01}=\lim_{t\to\infty}Q_{01}(t)=a$$

Similary,

$p_{02}=b$, $p_{10}=\frac{r_1q}{(1-qs_1)}$, $p_{11}=\frac{apr_1}{(1-qs_1)}$, $p_{12}=\frac{bpr_1}{(1-qs_1)}$

$p_{13}=\frac{aps_1}{(1-qs_1)}$, $p_{15}=\frac{bps_1}{(1-qs_1)}$, $p_{20}=\frac{r_2q}{(1-qs_2)}$, $p_{21}=\frac{apr_2}{(1-qs_2)}$

$p_{22}=\frac{bpr_2}{(1-qs_2)}$, $p_{24}=\frac{bps_2}{(1-qs_2)}$, $p_{25}=\frac{aps_2}{(1-qs_2)}$, $p_{31}=1$

$p_{42}=1$, $p_{50}=\frac{r_1r_2}{(1-s_1s_2)}$, $p_{51}=\frac{r_2s_1}{(1-s_1s_2)}$, $p_{52}=\frac{r_1s_2}{(1-s_1s_2)}$

We observe that the following relations hold-

$$p_{01}+p_{02}=1 \quad (18)$$

$$p_{10}+p_{11}+p_{12}+p_{13}+p_{15}=1 \quad (19)$$

$$p_{20}+p_{21}+p_{22}+p_{24}+p_{25}=1 \quad (20)$$

$$p_{31}=p_{42}=1 \quad (21)$$

$$p_{50}+p_{51}+p_{52}=1 \quad (22)$$

5. Mean Sojourn Times

Let T_i be the sojourn time in state S_i (i=0,1,2,3,4,5,6) then mean sojourn time ψ_i in state S_i is given by

$$\psi_i=\sum_{t=1}^{\infty}P[T_i>t-1]=\sum_{t=1}^{\infty}P[T_i\geq t]$$

In particular,

$$\psi_0=\frac{q}{p} \quad (23)$$

$$\psi_1=\frac{qs_1}{(1-qs_1)} \quad (24)$$

$$\psi_2 = \frac{qs_2}{(1-qs_2)} \tag{25}$$

$$\psi_3 = \frac{s_1}{r_1} \tag{26}$$

$$\psi_4 = \frac{s_2}{r_2} \tag{27}$$

$$\psi_5 = \frac{s_1 s_2}{(1-s_1 s_2)} \tag{28}$$

6. Methodology For Developing Equations

In order to obtain various interesting measures of system effectiveness we developed the recurrence relations for reliability, availability and busy period of repairman as follows-

6.1 Reliability of the system-

Let us define $R_i(t)$ as the probability that the system does not fail up to the epochs (t-1) when it is initially starts from state S_i. To determine it, we assume the failed states S_3, S_4 and S_5 as absorbing state. Using the simple probabilistic reasoning in regenerative point technique we have the following set of convolution equations $R_i(t); i = 0,1,2$.

$$\begin{aligned} R_0(t) &= q^t + \sum_{u=0}^{t-1} q_{01}(u) R_1(t-1-u) + \sum_{u=0}^{t-1} q_{02}(u) R_2(t-1-u) \\ &= Z_0(t) + q_{01}(t-1)©R_1(t-1) + q_{02}(t-1)©R_2(t-1) \end{aligned} \tag{29}$$

Similarly,

$$R_1(t) = Z_1(t) + q_{10}(t-1)©R_0(t-1) + q_{11}(t-1)©R_1(t-1) + q_{12}(t-1)©R_2(t-1) \tag{30}$$

$$R_2(t) = Z_2(t) + q_{20}(t-1)©R_0(t-1) + q_{21}(t-1)©R_1(t-1) + q_{22}(t-1)©R_2(t-1) \tag{31}$$

Where,

$Z_0(t) = q^t$, $\quad Z_1(t) = q^t s_1^t \quad$ and $\quad Z_2(t) = q^t s_2^t$

6.2 Availability of the system-

Let $A_i(t)$ be the probability that the system is up (operative) during the t^{th} cycle (t-1, t), when it initially started from state S_i. Using elementary probabilistic arguments as in case of reliability, we have the following recurrence relations-

$$A_0(t) = Z_0(t) + q_{01}(t-1)©A_1(t-1) + q_{02}(t-1)©A_2(t-1) \tag{32}$$

$$\begin{aligned} A_1(t) = Z_1(t) &+ q_{10}(t-1)©A_0(t-1) + q_{11}(t-1)©A_1(t-1) + q_{12}(t-1)©A_2(t-1) \\ &+ q_{13}(t-1)©A_3(t-1) + q_{15}(t-1)©A_5(t-1) \end{aligned} \tag{33}$$

$$\begin{aligned} A_2(t) = Z_2(t) &+ q_{20}(t-1)©A_0(t-1) + q_{21}(t-1)©A_1(t-1) + q_{22}(t-1)©A_2(t-1) \\ &+ q_{24}(t-1)©A_4(t-1) + q_{25}(t-1)©A_5(t-1) \end{aligned} \tag{34}$$

$$A_3(t) = q_{31}(t-1)©A_1(t-1) \tag{35}$$

$$A_4(t) = q_{42}(t-1)©A_2(t-1) \tag{36}$$

$$A_5(t) = q_{50}(t-1)©A_0(t-1) + q_{51}(t-1)©A_1(t-1) + q_{52}(t-1)©A_2(t-1) \quad (37)$$

Where, the values of $Z_i(t)$; i=0,1,2 are same as given in section 6.1.

6.3 Busy period of Repairman-

Let $B_i^1(t)$ and $B_i^2(t)$ be the respective probabilities that the repairman is busy in the minor and major repair of a failed unit during the t[th] cycle (t-1, t), when it initially started from state S_i. Then, by using simple probabilistic arguments as in case of reliability, the following recurrence relations can be easily developed for $B_i^j(t)$; i= 0 to 5. The dichotomous variable δ takes value 1 and 0 respectively for j=1 and 2.

$$B_0^j(t) = q_{01}(t-1)©B_1^j(t-1) + q_{02}(t-1)©B_2^j(t-1) \quad (38)$$

$$B_1^j(t) = \delta Z_1(t) + q_{10}(t-1)©B_0^j(t-1) + q_{11}(t-1)©B_1^j(t-1) + q_{12}(t-1)©B_2^j(t-1)$$
$$+q_{13}(t-1)©B_3^j(t-1) + q_{15}(t-1)©B_5^j(t-1) \quad (39)$$

$$B_2^j(t-1) = (1-\delta)Z_2(t) + q_{20}(t-1)©B_0^j(t-1) + q_{21}(t-1)©B_1^j(t-1) + q_{22}(t-1)©B_2^j(t-1)$$
$$+q_{24}(t-1)©B_4^j(t-1) + q_{25}(t-1)©B_5^j(t-1) \quad (40)$$

$$B_3^j(t) = \delta Z_3(t) + q_{31}(t-1)©B_1^j(t-1) \quad (41)$$

$$B_4^j(t) = Z_4(t) + q_{42}(t-1)©B_2^j(t-1) \quad (42)$$

$$B_5^j(t-1) = (1-\delta)Z_5(t) + q_{50}(t-1)©B_0^j(t-1) + q_{51}(t-1)©B_1^j(t-1) + q_{52}(t-1)©B_2^j(t-1) \quad (43)$$

Where, $Z_3(t) = s_1^t$, $Z_4(t) = s_2^t$ and $Z_5(t) = s_1^t s_2^t$.

7. Analysis of Reliability and MTSF

Taking geometric transform of (29-31) and simplifying the resulting set of algebraic equations for $R_0^*(h)$ we get

$$R_0^*(h) = \frac{N_1(h)}{D_1(h)} \quad (44)$$

Where,

$$N_1(h) = \left[\left(1-hq_{11}^*\right)\left(1-hq_{22}^*\right) - h^2 q_{12}^* q_{21}^*\right]Z_0^* + \left[hq_{01}^*\left(1-hq_{22}^*\right) + h^2 q_{02}^* q_{21}^*\right]Z_1^*$$
$$+\left[h^2 q_{01}^* q_{12}^* + hq_{02}^*\left(1-hq_{11}^*\right)\right]Z_2^*$$

and

$$D_1(h) = \left(1-hq_{11}^*\right)\left(1-hq_{22}^*\right) - h^2 q_{12}^* q_{21}^* - hq_{10}^*\left[hq_{01}^*\left(1-hq_{22}^*\right) + h^2 q_{02}^* q_{21}^*\right]$$
$$-hq_{20}^*\left[h^2 q_{01}^* q_{12}^* + hq_{02}^*\left(1-hq_{11}^*\right)\right]$$

Taking the inverse geometric transform of (44). one can get the expression of reliability of the system.

The MTSF is given by

$$E(T) = \sum_{t=1}^{\infty} R(t) = \lim_{h\to 1} \sum_{t=1}^{\infty} h^t R(t)$$

$$= \lim_{h\to 1} R^*(h) - R(0) = \frac{N_1(1)}{D_1(1)} - 1 \quad (45)$$

Where observing $q_{ij}^*(1)=p_{ij}$ and using results (23-28), we have

$$N_1(1)=\left[(1-p_{11})(1-p_{22})-p_{12}p_{21}\right](1+\psi_0)+\left[p_{01}(1-p_{22})+p_{02}p_{21}\right](1+\psi_1)$$
$$+\left[p_{01}p_{12}+p_{02}(1-p_{11})\right](1+\psi_2)$$

and

$$D_1(1)=(1-p_{11})(1-p_{22})-p_{12}p_{21}-p_{10}\left[p_{01}(1-p_{22})+p_{02}p_{21}\right]-p_{20}\left[p_{01}p_{12}+p_{02}(1-p_{11})\right]$$

8. Availability Analysis

On taking geometric transform of (32-37) and simplifying the resulting equations for $A_0^*(h)$, we get

$$A_0^*(h)=\frac{N_2(h)}{D_2(h)} \tag{46}$$

Where,

$$N_2(h)=\Big[\left(1-hq^*_{11}-h^2q^*_{13}q^*_{31}\right)\left\{1-hq^*_{22}-h^2q^*_{24}q^*_{42}-h^2q^*_{25}q^*_{52}\right\}$$
$$-hq^*_{21}\left\{hq^*_{12}+h^2q^*_{15}q^*_{51}\right\}-hq^*_{15}\left\{h^2q^*_{12}q^*_{25}+hq^*_{51}\left(1-hq^*_{22}-h^2q^*_{24}q^*_{42}\right)\right\}\Big]Z^*_0$$
$$+\Big[hq^*_{01}\left\{1-hq^*_{22}-h^2q^*_{24}q^*_{42}-h^2q^*_{25}q^*_{52}\right\}$$
$$+hq^*_{02}\left\{hq^*_{21}+h^2q^*_{25}q^*_{51}\right\}\Big]Z^*_1+\Big[hq^*_{01}\left\{hq^*_{12}+h^2q^*_{15}q^*_{52}\right\}$$
$$+hq^*_{02}\left\{1-hq^*_{11}-h^2q^*_{13}q^*_{31}-h^2q^*_{15}q^*_{51}\right\}\Big]Z^*_2$$

and

$$D_2(h)=\Big[\left(1-hq^*_{11}-h^2q^*_{13}q^*_{31}\right)\left\{1-hq^*_{22}-h^2q^*_{24}q^*_{42}-h^2q^*_{25}q^*_{52}\right\}$$
$$-hq^*_{21}\left\{hq^*_{12}+h^2q^*_{15}q^*_{51}\right\}-hq^*_{15}\left\{h^2q^*_{12}q^*_{25}+hq^*_{51}\left(1-hq^*_{22}-h^2q^*_{24}q^*_{42}\right)\right\}\Big]$$
$$-hq^*_{10}\Big[hq^*_{01}\left\{1-hq^*_{22}-h^2q^*_{24}q^*_{42}-h^2q^*_{25}q^*_{52}\right\}+hq^*_{02}\left\{hq^*_{21}+h^2q^*_{25}hq^*_{51}\right\}\Big]$$
$$-hq^*_{20}\Big[hq^*_{01}\left\{hq^*_{12}+h^2q^*_{15}q^*_{52}\right\}+hq^*_{02}\left\{1-hq^*_{11}-h^2q^*_{13}q^*_{31}-h^2q^*_{15}q^*_{51}\right\}\Big]$$
$$-hq^*_{50}\Big[hq^*_{01}\left\{h^2q^*_{12}q^*_{25}+hq^*_{15}\left(1-hq^*_{22}-h^2q^*_{24}q^*_{42}\right)\right\}$$
$$+hq^*_{02}\left\{hq^*_{25}\left(1-hq^*_{11}-h^2q^*_{13}q^*_{31}\right)+h^2q^*_{21}q^*_{15}\right\}\Big]$$

The steady-state availability of the system is given by

$$A_0=\lim_{t\to\infty}A_0(t)=\lim_{h\to 1}(1-h)\frac{N_2(h)}{D_2(h)}$$

Observing $q^*_{ij}(1)=p_{ij}$ and using relations (18-22), we get $D_2(h)$ at h=1 is zero, therefore by applying L. Hospital rule, we get

$$A_0=-\frac{N_2(1)}{D'_2(1)}$$

Where,

$$N_2(1)=u_0(1+\psi_0)+u_1(1+\psi_1)+u_2(1+\psi_2)$$
$$u_0=(1-p_{11}-p_{13})(1-p_{22}-p_{24}-p_{25}p_{52})-p_{21}(p_{12}+p_{15}p_{51})-p_{15}\left\{p_{12}p_{25}+p_{51}(1-p_{22}-p_{24})\right\}$$
$$u_1=p_{01}(1-p_{22}-p_{24}-p_{25}p_{52})+p_{02}(p_{21}+p_{25}p_{51})$$

$$u_2 = p_{01}\left(p_{12} + p_{15}p_{52}\right) + p_{02}\left(1 - p_{11} - p_{13} - p_{15}p_{51}\right)$$

and

$$D_2'(1) = -\left[u_0\left(1+\psi_0\right) + u_1\left\{\left(1+\psi_1\right) + p_{13}\left(1+\psi_3\right)\right\} + u_2\left\{\left(1+\psi_2\right) + p_{24}\left(1+\psi_4\right)\right\} + u_5\left(1+\psi_5\right)\right]$$

Where

$$u_5 = p_{01}\left\{p_{12}p_{25} + p_{51}\left(1 - p_{22} - p_{24}\right) + p_{02}p_{25}\left(1 - p_{11} - p_{13}\right) + p_{02}p_{21}p_{15}\right\}$$

Now the expected-up time of the system by the epoch (t-1) is given by

$$\mu_{up}(t) = \sum_{x=0}^{t-1} A_0(x)$$

So that,

$$\mu_{up}^*(h) = \frac{A_0^*(h)}{(1-h)}$$

9. Busy Period Analysis of Repairman

On taking geometric transforms of relation (38-43) and simplifying the resulting equations for minor and major repair i.e., for $\delta = 0,1$ we get,

$$B_0^{1*}(h) = \frac{N_3(h)}{D_2(h)} \quad \text{and} \quad B_0^{2*}(h) = \frac{N_4(h)}{D_2(h)}$$

Where,

$$N_3(h) = \left[hq_{01}^*\{1 - hq_{22}^* - h^2q_{24}^*q_{42}^* - h^2q_{25}^*q_{52}^*\} + hq_{02}^*\left\{hq_{21}^* + h^2q_{25}^*q_{51}^*\right\}\right]\left(Z_1^* + hq_{13}^*Z_3^*\right)$$
$$+\left[hq_{01}^*\{h^2q_{12}^*q_{25}^* + hq_{15}^*\left(1 - hq_{22}^* - h^2q_{24}^*q_{42}^*\right)\}\right.$$
$$\left.+h^2q_{02}^*q_{25}^*\left(1 - hq_{11}^* - h^2q_{13}^*q_{31}^*\right) + h^3q_{02}^*q_{21}^*q_{15}^*\right]Z_5^*$$

$$N_4(h) = \left[hq_{02}^*\{hq_{12}^* + h^2q_{15}^*q_{52}^*\} + hq_{02}^*\left\{1 - hq_{11}^* - h^2q_{13}^*q_{31}^* - h^2q_{15}^*q_{51}^*\right\}\right]\left(Z_2^* + hq_{24}^*Z_4^*\right)$$
$$+\left[hq_{01}^*\{h^2q_{12}^*q_{25}^* + hq_{15}^*\left(1 - hq_{22}^* - h^2q_{24}^*q_{42}^*\right)\}\right.$$
$$\left.+h^2q_{02}^*q_{25}^*\left(1 - hq_{11}^* - h^2q_{13}^*q_{31}^*\right) + h^3q_{02}^*q_{21}^*q_{15}^*\right]Z_5^*$$

and $D_2(h)$ is same as in availability analysis.

In the long run the respective probabilities that the repairman is busy in the minor and major repair of a failed unit are respectively given by-

$$B_0^1 = \lim_{t\to\infty} B_0^1(t) = \lim_{h\to 1}(1-h)\frac{N_3(h)}{D_2(h)}$$

$$B_0^2 = \lim_{t\to\infty} B_0^2(t) = \lim_{h\to 1}(1-h)\frac{N_4(h)}{D_2(h)}$$

But $D_2(h) = 0$ at h=1, therefore by applying L. Hospital rule, we get

$$B_0^1 = -\frac{N_3(1)}{D_2'(1)}, \qquad B_0^2 = -\frac{N_4(1)}{D_2'(1)} \tag{47}$$

Where,

$$N_3(1) = u_1\left\{\left(1+\psi_1\right) + p_{13}\left(1+\psi_3\right)\right\} + u_5\left(1+\psi_5\right)$$

$$N_4(1) = u_2\left(\psi_2 + p_{24}\psi_4\right) + u_5\psi_5$$

and $D_2'(1)$ is same as in availability analysis.

Now the expected busy period duration of the repairman in minor repair and major repair of a failed unit by the epoch (t-1) are respectively given by -

$$\mu_b^1(t)=\sum_{x=0}^{t-1}B_0^1(x), \qquad \mu_b^2(t)=\sum_{x=0}^{t-1}B_0^2(x)$$

So that,

$$\mu_b^{1*}(h)=\frac{B_0^{1*}(h)}{(1-h)}, \qquad \mu_b^{2*}(h)=\frac{B_0^{2*}(h)}{(1-h)} \tag{48}$$

10. Profit Function Analysis

We are now in the position to obtain the net expected profit incurred by the epoch (t-1) by considering the characteristics obtained in earlier section.
Let us consider,

K_0 = revenue per-unit time by the system due to its operation.
K_1 = cost per-unit time when repairman is busy in the minor repair
K_2 = cost per-unit time when repairman is busy in the minor repair

Then, the net expected profit incurred by the epoch (t-1) is given by

$$P(t)=K_0\mu_{up}(t)-K_1\mu_b^1(t)-K_2\mu_b^2(t) \tag{49}$$

The expected profit per-unit time in steady state is given by

$$P=\lim_{t\to\infty}\frac{P(t)}{t}=\lim_{h\to1}(1-h)^2P^*(h)$$

$$=K_0\lim_{h\to1}(1-h)^2\frac{A_0^*(h)}{(1-h)}-K_1\lim_{h\to1}(1-h)^2\frac{B_0^{1*}(h)}{(1-h)}-K_2\lim_{h\to1}(1-h)^2\frac{B_0^{2*}(h)}{(1-h)}$$

$$=K_0A_0-K_1B_0^1-K_2B_0^2 \tag{50}$$

11. Graphical Representation and conclusions

The curves for MTSF and profit function have been drawn for different values of failure parameters. Fig. 2 depicts the variation in MTSF with respect to failure rate (p) for different values of repair rate r_1 and r_2 when values of other parameters are kept fixed as $a=0.8$. From the curves we conclude that expected life of the system decrease with increase in p. Further, it increases with the increase of the values of r_1 and r_2. Also to achieve at least MTSF at 475 units, we conclude from smooth curves that the value of p must be less than 0.044, 0.050, 0.059 for $r_1=0.6, 0.75, 0.95$ where $r_2=0.25$. Whereas from dotted curves we conclude that the value of p must be less than 0.041, 0.046, 0.053 for $r_1=0.6, 0.75, 0.95$ when $r_2=0.20$.

Similarly, Fig. 3 reveals the variations in profit (P) with respect to p for varying values of r_1 and r_2, when other parameters are kept fixed as $a=0.01$, $K_0=175$, $K_1=195$ and $K_2=180$. From the figure, it is clearly observed from the smooth curves, that the system is profitable if the value of parameter p is less than 0.038, 0.056 and 0.077 respectively for $r_1=0.6, 0.75, 0.95$ when $r_2=0.24$. From dotted curves, we conclude that system is profitable only if value of parameter p is less than 0.029, 0.044 and 0.062 respectively for $r_1=0.6, 0.75, 0.95$ when $r_2=0.22$.

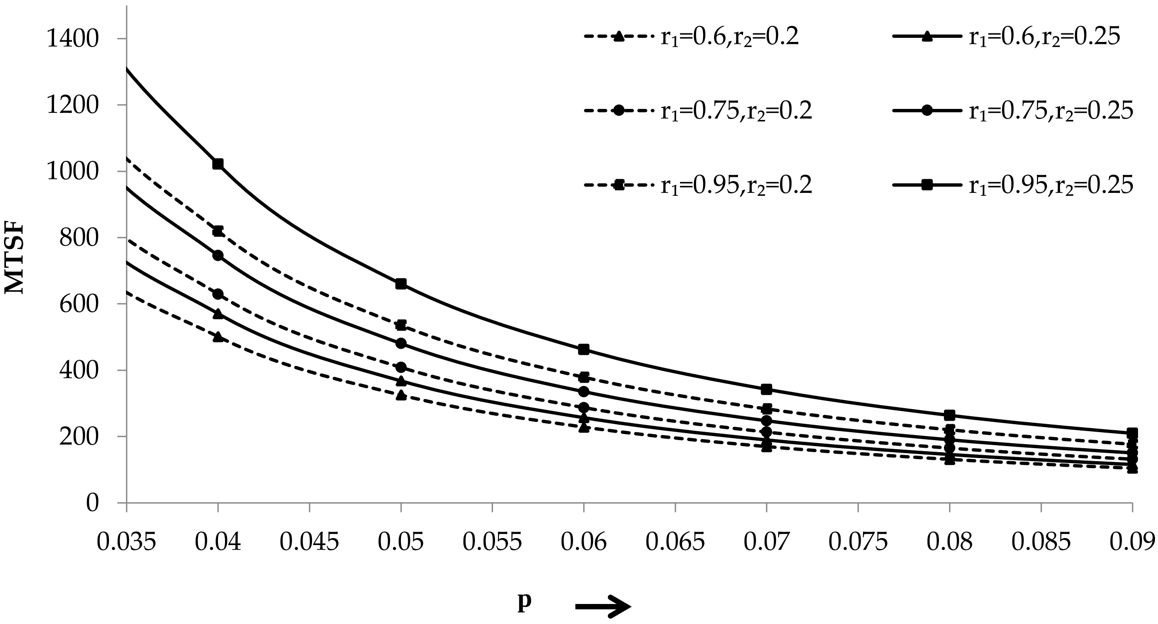


Figure.2 *Behavior of MTSF with respect to* p, r_1 *and* r_2

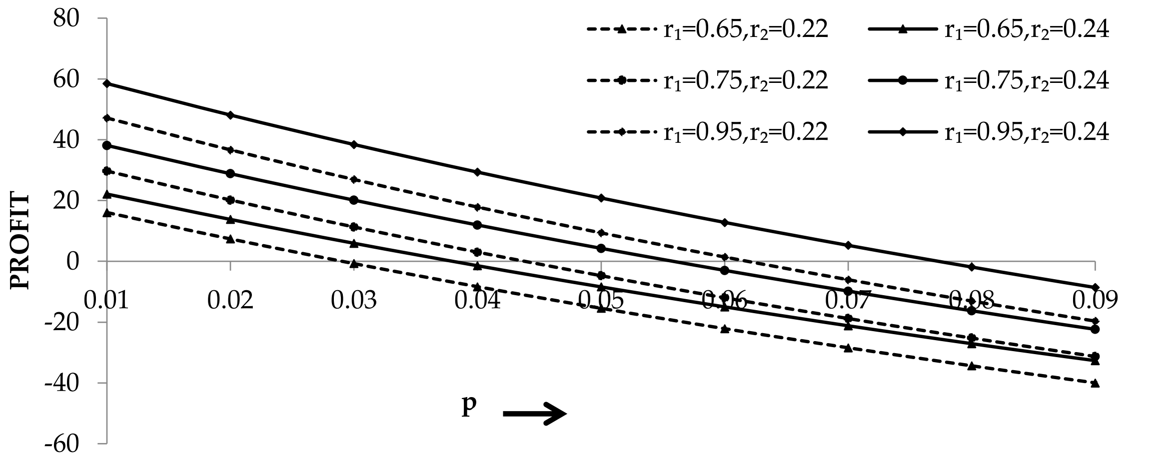


Figure.3 *Behavior of Profit (P) with respect to* p, r_1 *and* r_2

References

[1] Chowdhary, S. and Kumar, H. (2017). Profit analysis of a complex system having two type of failure and rest period of repairman with gamma repair time distribution. *International Journal of Advanced Science and Research*, 2(5):115-118.

[2] Chaudhary P. and Tyagi, L. (2021). A Two Non-Identical Unit Parallel System Subject to Two Types of Failure and Correlated Life Times. *Reliability: Theory & Applications*, 16(2):247-258.

[3] Gupta, R. and Varshney, G. (2006). A Two Non-Identical Unit Parallel System with geometric failure and repair time distributons. *IAPQR Trans.*, 31(2):127-139.

[4] Gupta, R. and Bhardwaj , P. (2014). Analysis of a discrete parametric Markov-chain model of a two-unit cold standby system with repair machine failure. *Int. J. of Scientific and Engineering research*, 5(2):1763-1770.

[5] Gupta, R. and Tyagi, A. (2014). A two identical unit cold standby system with switching device and geometric failure and repair time distributions. *Aligarh J. of Statistics*, 34:55-56.

[6] Gupta, P. and Vinodiya, P. (2018). Analysis of Reliability of a two-non identical units cold standby repairable system has two types of failure. *International J. of Computer Sciences and Engineering*, 6(11):907-913.

[7] Levitin G., Zhang T. and Xie, M. (2006). State probability of a series-parallel repairable system with two types of failure states. *International J. of System Science*, 37:1011-1020.

[8] Ram M. and Singh, S.B. (2006). Analysis of a complex system with common cause failure and two types of repair facilities with different distributions in failure. *International J. of Reliability and Safety*, 4(4):381-392.

Transmuted Exponentiated Kumaraswamy Distribution

Jeena Joseph[1] and Meera Ravindran[2]

•

Department of Statistics
St. Thomas' College (Autonomous), Thrissur, India

sony.jeena@gmail.com, meeraravindran798@gmail.com

Abstract

In this paper, a generalization of the Exponentiated Kumaraswamy distribution referred to as the Transmuted Exponentiated Kumaraswamy distribution is proposed. The new transmuted distribution is developed using the quadratic rank transmutation map. The mathematical properties of the new distribution is provided. Explicit expressions are derived for the moments, incomplete moments, moment generating function, quantile function, entropy, mean deviation and order statistics. Survival analysis is also performed. The distribution parameters are estimated using the method of maximum likelihood. Simulation of random variables is performed in order to investigate the performance of the estimates. An analysis using real life data is conducted to demonstrate the usefulness of the proposed distribution.

Keywords: Bonferroni and Lorenz curves; Hazard function; Maximum likelihood estimation; Moments; Transmuted Exponentiated Kumaraswamy Distribution; Transmuted family.

1. Introduction

In probability theory and Statistics, a probability distribution is a mathematical function that provides the probabilities of the occurrence of various possible outcomes in an experiment.In modelling our world, probability distributions helps us, thus allowing to obtain estimates of the probability of a certain event to occur, or estimate it's variability of happening. Many distributions have been discovered suitable for many different purposes.The recognition of the proper distribution will allow a correct application of a model that would easily forecast the probability of an event.

The Kumaraswamy probability distribution was developed by Kumaraswamy [11] which is closely related to the beta distribution. It is often termed as a Beta-like distribution. But, in some situations the Kumaraswamy distribution is simpler to use and more amenable. Since it's cumulative distribution function (cdf) has a closed form, it is often preferred over the Beta distribution. Moreover, unlike the beta cdf, the cdf of Kumaraswamy distribution does not contain the incomplete Beta function, which makes it much simple to work with and the new properties of Kumaraswamy distribution such as the Kumaraswamy variables show closeness under exponentiation and under linear transformation was studied by Mitnik [16]. In numerous areas such as hydrology, electrical, civil, mechanical and financial engineering, Kumaraswamy distribution has secured appreciable interest, see Mohammed [17]. Some generalized beta distributions of the second kind having desirable application features in hydrology and meteorology was studied by Mielke and Johnson [15] and Fletcher and Ponnambalam [7]. Several authors studied more general properties of Kumaraswamy Distribution, see Silva et. al. [18], ZeinEldin et.al. [23], Dey et. al. [6], Hassan and Elgarhy [8] and Simbolan et. al. [19]. Usman et. al. [22] derived a new Weibull-Kumaraswamy distribution and studied its properties and applications. Another distribution named Kumaraswamy- Pareto distribution was derived by Bourguignon et. al. [5]. A bivariate Kumaraswamy (BVK) distribution with marginals being Kumaraswamy distributions was introduced by Barreto-Souza and Lemonte [4]. Another new three-parameter probability model named Exponentiated Kumaraswamy distribution and its basic statistical properties and its applications using real-life datasets were studied by Lemonte et. al. [12].Also, a three-parameter weighted kumaraswamy distribution was proposed by Abd El-Monsef et. al. [1] for modeling some biological data, which could accommodate increasing and decreasing hazard rate function with bathtub shape. AL-Fattah et. al. [2] introduced the Inverted Kumaraswamy Distribution, it's properties and estimation.

The transmuted family of distributions has been receiving a high attention over the past few years. A new technique for adding a new parameter to an already existing distribution that would provide more flexibility to this distribution by Shaw and Buckley [20]. The method is named as quadratic rank transmutation map (QRTM). It includes the parent distribution as a special case and makes it more flexible to model different types of data. The generated family is also called the transmuted extended distribution. This method have been considered by several authors for different disributions, see Al-Kadim et. al. [3], Khan et. al. [10], F. Merovci [13,14] and Sherwaia et. al. [21]. Transmuted Kumaraswamy distribution and its basic statistical properties were discussed by Khan et. al. [9]. The new model was found to outperform some existing baseline distributions when applied to real-life data sets.

A random variable X is said to have an Exponentiated Kumaraswamy distribution with parameters $\alpha, \beta, \gamma > 0$ if its probability density function (pdf) is given by

$$f(x;\alpha,\beta,\gamma) = \alpha\beta\gamma x^{\alpha-1}(1-x^{\alpha})^{\beta-1}[1-(1-x^{\alpha})^{\beta}]^{\gamma-1}, \qquad 0 < x < 1, \quad \alpha,\beta,\gamma > 0 \tag{1}$$

and the respective cdf is

$$F(x;\alpha,\beta,\gamma) = [1-(1-x^{\alpha})^{\beta}]^{\gamma-1}, \;\; 0 < x < 1, \;\; \alpha,\beta,\gamma > 0 \tag{2}$$

In this paper, a generalization of the Exponentiated Kumaraswamy distribution referred to as the Transmuted Exponentiated Kumaraswamy distribution is proposed. The new transmuted distribution is obtained using the quadratic rank transmutation map introduced by Shaw and Buckley [20]. According to this method, transmutation maps consists of the functional composition of the cumulative distribution function of one distribution with the inverse cumulative distribution (quantile) function of another. A comprehensive account of the mathematical properties of the new distribution is provided.

The organization of this paper is as follows: Section 2 explains the quadratic rank transmutation method. In section 3, the pdf and cdf of our new model, Transmuted Exponentiated Kumaraswamy distribution is given and provide the graphical presentation of its pdf, cdf, survival function and hazard rate function for selected values of the parameters. Section 4 provides its statistical properties such as moments, moment generating function, characteristic function, quantile function, incomplete moments, entropy, mean deviation and order statistics. Estimation of parameters of the distribution is done using maximum likelihood estimation is also included in this section. In section 5, a simulation study is included which is done to validate the estimates and a real data analysis illustrates the practicability of the proposed distribution. Finally, the summary and conclusions are stated in section 6.

2. Transmuted Distribution

A random variable X is said to have transmuted distribution if its cumulative distribution function(cdf) satisfy the relation,

$$F(x) = G(x)[(1+\lambda) - \lambda G(x)], \quad |\lambda| < 1 \tag{3}$$

which on differentiation yields the corresponding pdf

$$f(x) = g(x)[1 + \lambda - 2\lambda G(x)] \tag{4}$$

where G(x) and g(x) are the cdf and pdf of the base distribution. Observe that at λ = 0, we have the distribution of the base random variable.

3. Transmuted Exponentiated Kumaraswamy Distribution

Using (3) and (4) we have the cdf of Transmuted Exponentiated Kumaraswamy (TEKw) distribution

$$F(x;\alpha,\beta,\gamma,\lambda) = (1+\lambda)[1-(1-x^{\alpha})^{\beta}]^{\gamma} - \lambda[1-(1-x^{\alpha})^{\beta}]^{2\gamma} \tag{5}$$

with shape parameters $\alpha, \beta, \gamma > 0$ and the transmuting parameter $|\lambda| < 1$.
Hence, the pdf of TEKw distribution is given as,

$$f(x;\alpha,\beta,\gamma,\lambda) = \alpha\beta\gamma x^{\alpha-1}(1-x^{\alpha})^{\beta-1}[1-(1-x^{\alpha})^{\beta}]^{\gamma-1}[(1+\lambda)-2\lambda[1-(1-x^{\alpha})^{\beta}]^{\gamma}] \tag{6}$$

where $\alpha, \beta, \gamma > 0$ and $|\lambda| < 1$.

Note that the transmuted Exponentiated Kumaraswamy distribution is an extended model to analyze more complex data and it generalizes some of the widely used distributions.The Exponentiated Kumaraswamy distribution is clearly a special case for λ = 0.

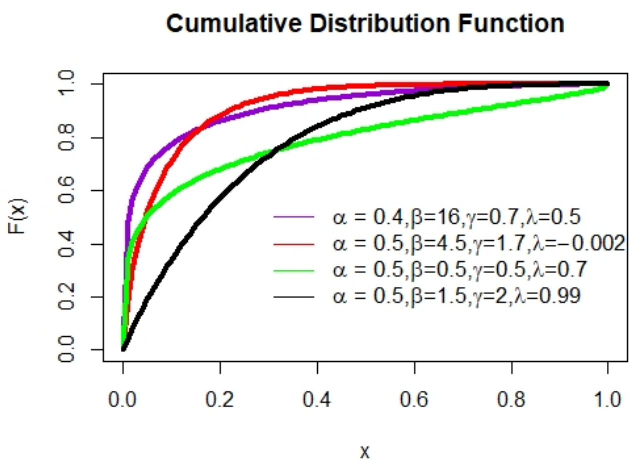


Figure 1: *Plot of the cumulative distribution function for different values of parameters.*

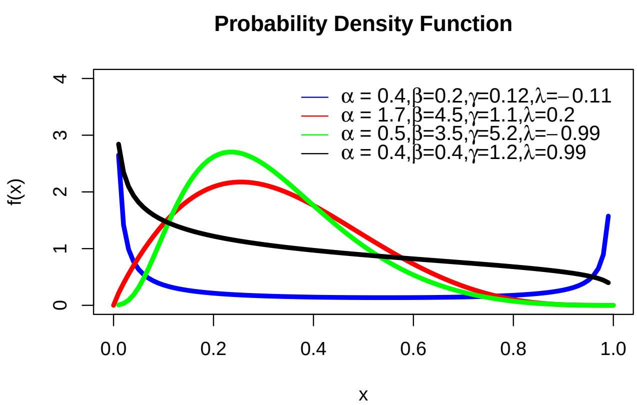


Figure 2: *Plot of the density function for different values of parameters.*

4. Statistical Properties

4.1. Moments

Let X be a random variable with its pdf given by (6), then its r^{th} raw moment is given by,

$$\begin{aligned} \mu_r' &= E(X^r) \\ &= \int_0^1 x^r f(x) dx \\ &= \int_0^1 x^r \alpha\beta\gamma x^{\alpha-1}(1-x^\alpha)^{\beta-1}[1-(1-x^\alpha)^\beta]^{\gamma-1}[(1+\lambda)-2\lambda[1-(1-x^\alpha)^\beta]^\gamma]dx \end{aligned} \tag{7}$$

splitting into two parts,

$$= \alpha\beta\gamma[\int_0^1 x^{r+\alpha-1}(1-x^\alpha)^{\beta-1}(1+\lambda)[1-(1-x^\alpha)^\beta]^{\gamma-1}dx - \int_0^1 2\lambda x^{r+\alpha-1}(1-x^\alpha)^{\beta-1}[1-(1-x^\alpha)^\beta]^{2\gamma-1}dx]$$

Substituting u=x^α, du=$\alpha x^{\alpha-1}$ dx

and evaluating both parts using the series expansions

$$\begin{aligned} [1-(1-x^\alpha)^\beta]^{\gamma-1} &= \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma\gamma}{\Gamma(\gamma-i)i!}(1-x^\alpha)^{\beta i} \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma\gamma}{\Gamma(\gamma-i)i!} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\beta j+1)}{\Gamma(\beta j+1-j)j!} x^{\alpha j} \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{i+j}\Gamma\gamma\Gamma(\beta j+1)}{i!j!\Gamma(\gamma-i)\Gamma(\beta i-j+1)} x^{\alpha j} \\ &= M x^{\alpha j} \end{aligned} \tag{8}$$

and,

$$\begin{aligned} [1-(1-x^\alpha)^\beta]^{2\gamma-1} &= \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma 2\gamma}{\Gamma(2\gamma-i)i!}(1-x^\alpha)^{\beta i} \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma 2\gamma}{\Gamma(2\gamma-i)i!} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\beta i+1)}{\Gamma(\beta i+1-j)j!} x^{\alpha j} \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{i+j}\Gamma 2\gamma\Gamma(\beta i+1)}{i!j!\Gamma(2\gamma-i)\Gamma(\beta i-j+1)} x^{\alpha j} \\ &= N x^{\alpha j} \end{aligned} \tag{9}$$

The r^{th} moment is given by,

$$\begin{aligned} E(X^r) &= \alpha\beta\gamma[\frac{M(1+\lambda)}{\alpha}B(\frac{r}{\alpha}+j+1,\beta) - \frac{2\lambda N}{\alpha}B(\frac{r}{\alpha}+j+1,\beta)] \\ &= \frac{\alpha\beta\gamma}{\alpha}[M(1+\lambda)B(\frac{r}{\alpha}+j+1,\beta) - 2\lambda N B(\frac{r}{\alpha}+j+1,\beta)] \\ &= \beta\gamma B(\frac{r}{\alpha}+j+1,\beta)[M+\lambda(M-2N)] \end{aligned} \tag{10}$$

where,
$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ is the Beta function and,

$$\begin{aligned} M &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{i+j}\Gamma\gamma\Gamma(\beta j+1)}{i!j!\Gamma(\gamma-i)\Gamma(\beta i-j+1)} \\ N &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{i+j}\Gamma 2\gamma\Gamma(\beta i+1)}{i!j!\Gamma(2\gamma-i)\Gamma(\beta i-j+1)} \end{aligned} \tag{11}$$

Therefore, the expected value E(X) and variance Var(X) of a transmuted exponentiated Kumaraswamy random variable X are respectively, given by

$$E(X) = \beta\gamma B(\frac{1}{\alpha}+j+1,\beta)[M+\lambda(M-2N)] \tag{12}$$

and

$$V(X) = \beta\gamma B(\frac{2}{\alpha}+j+1,\beta)[M+\lambda(M-2N)] - (\beta\gamma B(\frac{1}{\alpha}+j+1,\beta)[M+\lambda(M-2N)])^2 \tag{13}$$

4.2. Moment Generating Function

The moment generating function of TEKw$(\alpha,\beta,\gamma,\lambda)$ is given by,

$$M_X(t) = E(e^{tx}) = \int_0^1 e^{tx} f(x)dx$$

Using the Taylor series expansion,

$$e^{tx} = \sum_{n=0}^{\infty} \frac{(tx)^n}{n!}$$

$$\begin{aligned} M_X(t) &= \int_0^1 \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} f(x)dx \\ &= \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} \int_0^1 x^n f(x)dx \\ &= \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} \int_0^1 x^n \alpha\beta\gamma x^{\alpha-1}(1-x^\alpha)^{\beta-1}[1-(1-x^\alpha)^\beta]^{\gamma-1}[(1+\lambda)-2\lambda[1-(1-x^\alpha)^\beta]^\gamma]dx \end{aligned} \tag{14}$$

Splitting into two parts, and evaluating, the moment generating function of TEKw$(\alpha,\beta,\gamma,\lambda)$ is given by,

$$M_X(t) = \beta\gamma \sum_{n=0}^{\infty} \frac{t^n}{n!} B(\frac{n}{\alpha}+j+1)[M+\lambda(M-2N)] \tag{15}$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ is the Beta function and M and N are given by (11).

4.3. Characteristic Function

The characteristic function of TEKw$(\alpha,\beta,\gamma,\lambda)$ is given by,

$$\phi_X(t) = \beta\gamma \sum_{n=0}^{\infty} \frac{(it)^n}{n!} B(\frac{n}{\alpha}+j+1)[M+\lambda(M-2N)] \tag{16}$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ is the Beta function and M and N are given by (11).

4.4. Quantile Function

$$Q(p) = [1 - [1 - [\frac{(1+\lambda) - \sqrt{(1+\lambda)^2 - 4\lambda p}}{2\lambda}]^{1/\gamma}]^{1/\beta}]^{1/\alpha} \tag{17}$$

Then,
Median, 2^{nd} quartile of TEKw(α,β, γ,λ) is obtained by substituting p =1/2 in (17). If U is a standard uniform variate, we can generate random variables using the following expression.

$$X = [1 - [1 - [\frac{(1+\lambda) - \sqrt{(1+\lambda)^2 - 4\lambda u}}{2\lambda}]^{1/\gamma}]^{1/\beta}]^{1/\alpha} \tag{18}$$

Then the random variable X follows TEKw(α,β,γ,λ).

4.5. Incomplete Moments

The s^{th} incomplete moment, say $\phi_s(t)$ of TEKw is,

$$\begin{aligned}\phi_s(t) &= \int_0^t x^s f(x)dx \\ &= \int_0^t x^s \alpha\beta\gamma x^{\alpha-1}(1-x^\alpha)^{\beta-1}[1-(1-x^\alpha)^\beta]^{\gamma-1}[(1+\lambda) - 2\lambda[1-(1-x^\alpha)^\beta]^\gamma]dx \\ &= \int_0^t x^{s+\alpha-1} \alpha\beta\gamma(1-x^\alpha)^{\beta-1}[1-(1-x^\alpha)^\beta]^{\gamma-1}[(1+\lambda) - 2\lambda[1-(1-x^\alpha)^\beta]^\gamma]dx\end{aligned}$$

After some algebra,

$$\begin{aligned}\phi_s(t) &= \alpha\beta\gamma\Big[\frac{M(1+\lambda)}{\alpha}B(t^\alpha; \frac{s}{\alpha}+j+1, \beta) - \frac{2\lambda N}{\alpha}B(t^\alpha; \frac{s}{\alpha}+j+1, \beta)\Big] \\ &= \beta\gamma B(t^\alpha; \frac{s}{\alpha}+j+1, \beta)\big[M + \lambda(M-2N)\big]\end{aligned} \tag{19}$$

where $B(w; a, b) = \int_0^w t^{a-1}(1-t)^{b-1}dt$ is the incomplete beta function.The first incomplete moment can be obtained by substituting s=1 in (19).

4.6. Mean Deviations

The mean deviation is a measure of amount of scatter in a random variable. Let X follow TEKw(α,β,γ,λ) with mean μ and median M.

- Mean Deviation from the mean is given by,

$$\delta_1(x) = \int_{-\infty}^{+\infty} |x-\mu| f(x)dx = 2\mu F(\mu) - 2\phi(\mu) \tag{20}$$

- Similarly, the Mean Deviation from the median is,

$$\delta_2(x) = \int_{-\infty}^{+\infty} |x-M| f(x)dx = \mu - 2\phi(M) \tag{21}$$

where F(μ) can be determined from (5) and $\phi(q) = \int_{-\infty}^{q} xf(x)dx$ is the first incomplete moment.
The mean deviations about mean and median are obtained by substituting median obtained from (17), first incomplete moment (19) with $s = 1$ and cdf (5) in (20) and (21).

Application of these equations can be made to obtain the Bonferroni curve, $B(x) = \frac{\phi_1(X)}{E(X)}$ and the Lorenz curve, $L(X) = \frac{\phi_1(X)}{F(X)E(X)}$ where $\phi_1(X)$ is the first incomplete moment from (19), F(x) is the cdf of TEKw distribution and E(X) is the mean.
These curves are very useful in economics, reliability, medicine, insurance and demography.

4.7. Entropy

The Renyi Entropy (Alfred Renyi) of a random variable X represents a measure of variation of the uncertainity which is defined by,

$$R_p(x) = \frac{1}{1-p} log \int_{-\infty}^{+\infty} f(x)^p dx$$

where $p > 0$ and $p \neq 1$

We have,

$$\begin{aligned} [1-(1-x^\alpha)^\beta]^{(\gamma-1)p} &= \sum_{i=0}^{\infty} \frac{\Gamma(\gamma-1)p+1}{\Gamma((\gamma-1)p+1-i)}[(1-x^\alpha)^\beta]^i \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-1)^{i+j}\Gamma((\gamma-1)p+1)\Gamma(\beta i+1)}{\Gamma((\gamma-1)p+1-i)\Gamma(\beta i+1-j)} x^{\alpha j} \\ &= \eta x^{\alpha j} \end{aligned} \tag{22}$$

and

$$\begin{aligned} [1-\lambda[2(1-(1-x^\alpha)^\beta)^\gamma-1]]^p &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{\Gamma(p+1)}{\Gamma(p+1-i)} [\lambda[2(1-(1-x^\alpha)^\beta)^\gamma-1]^i \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{\Gamma(p+1)}{\Gamma(p+1-i)} \lambda^i[2(1-(1-x^\alpha)^\beta)^\gamma-1]^i \end{aligned}$$

and,

$$\begin{aligned} [1-\lambda[2(1-(1-x^\alpha)^\beta)^\gamma-1]]^p &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{\Gamma(p+1)}{\Gamma(p+1-i)} [\lambda[2(1-(1-x^\alpha)^\beta)^\gamma-1]^i \\ &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{\Gamma(p+1)}{\Gamma(p+1-i)} \lambda^i[2(1-(1-x^\alpha)^\beta)^\gamma-1]^i \end{aligned}$$

$$\begin{aligned} &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{(-1)^{2i+j+k+l}}{i!j!k!l!} \frac{\Gamma(p+1)\Gamma(i+1)\Gamma(\gamma j+1)\Gamma(\beta k+1)}{\Gamma(p+1-i)\Gamma(i+1-j)\Gamma(\gamma j+1-k)\Gamma(\beta k+1-l)} \lambda^i 2^{\gamma j} x^{\alpha l} \\ &= \theta x^{\alpha l} \end{aligned} \tag{23}$$

Evaluating the above integral, Rennyi entropy is,

$$R_p(x) = \frac{1}{1-p} log[\alpha^{p-1}(\beta\gamma)^p \eta\theta B(\frac{1}{\alpha}[(\alpha+1)p+1]+j+l,(\beta-1)p+1)] \tag{24}$$

where, η and θ are given by (22) and (23). The δ entropy,$\delta > 0$,$\delta \neq 1$,say $H_\delta(x)$ is defined as,

$$H_p(x) = \frac{1}{p-1} log[1 - \int_{-\infty}^{+\infty} f(x)^p dx]$$

where $p > 0$ and $p \neq 1$

Using (24),

$$H_p(x) = \frac{1}{p-1} log[1-\alpha^{p-1}(\beta\gamma)^p \eta\theta B(\frac{1}{\alpha}[(\alpha+1)p+1]+j+l,(\beta-1)p+1)] \tag{25}$$

where η and θ are given by (22) and (23).

The Rennyi entropy converge to the Shannon entropy when $\delta \to 1$.

4.8. Survival Function

Survival function is the probability that a system will survive beyond a given time. Mathematically, the survival function of TEKw(α,β,γ,λ)) is defined by:

$$S(x;\alpha,\beta,\gamma,\lambda) = 1 - [1-(1-x^{\alpha})^{\beta}]^{\gamma}[(1+\lambda) - \lambda[1-(1-x^{\alpha})^{\beta}]^{\gamma}] \tag{26}$$

where $\alpha, \beta, \gamma > 0$ and $|\lambda| < 1$. By choosing some arbitrary values for parameters, we provide some possible shapes for the survival function of the TEKw as shown in Figure 3:

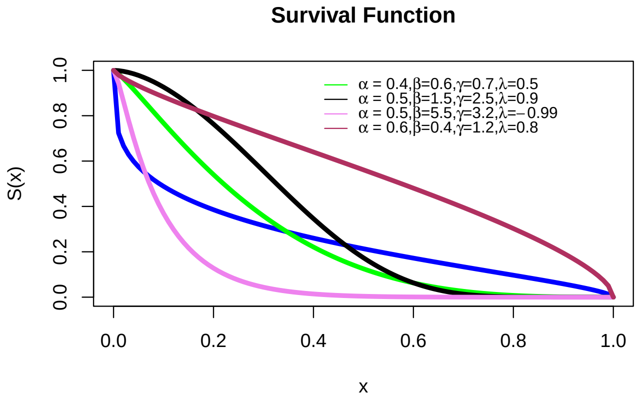


Figure 3: *Plot of the survival function for different values of parameters.*

4.9. Hazard Rate Function

The hazard rate function h(x) of TEKw($\alpha, \beta, \gamma, \lambda$) is given as,

$$h(x;\alpha,\beta,\gamma,\lambda) = \frac{\alpha\beta\gamma x^{\alpha-1}(1-x^{\alpha})^{\beta-1}[(1+\lambda) - 2\lambda[1-(1-x^{\alpha})^{\beta}]^{\gamma}]}{1-[(1+\lambda)[1-(1-x^{\alpha})^{\beta}]^{\gamma} - \gamma[1-(1-x^{\alpha})^{\beta}]^{2\gamma}]} \tag{27}$$

where $\alpha, \beta, \gamma > 0$ and $|\lambda| < 1$.

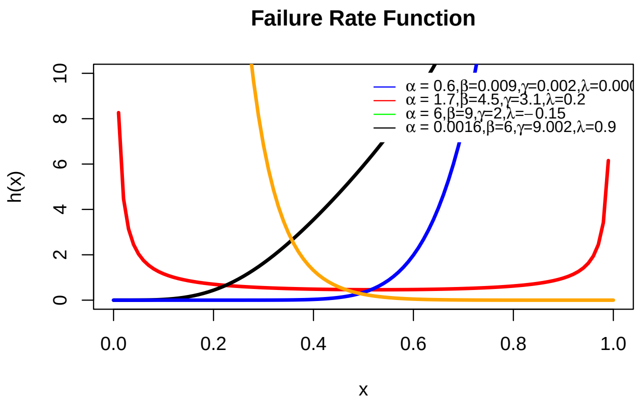


Figure 4: *Plot of the hazard rate function for different values of parameters.*

4.10. **Order Statistics**

Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ denote the order statistics of a random sample $X_1, X_2, ..., X_n$ from a population with cdf $F_X(x)$ and pdf $f_X(x)$given by (5) and (6). The pdf of k^{th} order statistic

$$f_{X(k)}(x) = \frac{n!}{(k-1)!(n-k)!}\alpha\beta\gamma x^{\alpha-1}(1-x^{\alpha})^{\beta-1}[1-(1-x^{\alpha})^{\beta}]^{\gamma-1}[(1+\lambda)-2\lambda[1-(1-x^{\alpha})^{\beta}]^{\gamma}]$$
$$\times\left\{[(1+\lambda)[1-(1-x^{\alpha})^{\beta}]^{\gamma}-\lambda[1-(1-x^{\alpha})^{\beta}]^{2}\gamma]^{k-1}\right\}$$
$$\times\left\{[1-(1+\lambda)[1-(1-x^{\alpha})^{\beta}]^{\gamma}+[\lambda[1-(1-x^{\alpha})^{\beta}]^{2\gamma}]]^{n-k}\right\}$$

4.11. Maximum Likelihood Estimation

The estimation of parameters α, β, γ and λ is done using the maximum likelihood estimation method. Let $X_1, X_2, ..., X_n$ be an observed random sample from TEKw$(\alpha, \beta, \gamma, \lambda)$ distribution with unknown parameters α,β,γ and λ. The likelihood function is,

$$L(x) = \prod_{i=1}^{n} f(x_i; \alpha, \beta, \gamma, \lambda)$$

i.e.,

$$L(x) = \prod_{i=1}^{n} \alpha\beta\gamma x_i^{\alpha-1}(1-x_i^{\alpha})^{\beta-1}[1-(1-x_i^{\alpha})^{\beta}]^{\gamma-1}[(1+\lambda)-2\lambda[1-(1-x_i^{\alpha})^{\beta}]^{\gamma}]$$

Then the log-likelihood function is given by

$$\begin{aligned} lnL = nln\alpha + nln\beta + nln\gamma + (\alpha-1)\sum_{i=1}^{n} ln(x_i) + (\beta-1)\sum_{i=1}^{n} ln(1-x_i^{\alpha}) \\ + (\gamma-1)\sum_{i=1}^{n} ln(1-(1-x_i^{\alpha})^{\beta}) + \sum_{i=1}^{n} ln[(1+\lambda)-2\lambda(1-(1-x_i^{\alpha})^{\beta})^{\gamma}] \end{aligned} \quad (28)$$

Therefore, the MLEs of $\alpha, \beta, \gamma, \lambda$ which maximize (28) must satisfy the following normal equations;

$$\begin{aligned} \frac{n}{\alpha} + \sum_{i=1}^{n} lnx_i - (\beta-1)\sum_{i=1}^{n} \frac{x_i^{\alpha} lnx_i}{1-x_i^{\alpha}} + \beta(\gamma-1)\sum_{i=1}^{n} \frac{(1-x_i^{\alpha})^{\beta-1} x_i^{\alpha} lnx_i}{[1-(1-x_i^{\alpha})^{\beta}]} \\ -2\lambda\beta\gamma\sum_{i=1}^{n} \frac{[1-(1-x_i^{\alpha})^{\beta}]^{\gamma-1}(1-x_i^{\alpha})^{\beta-1} x_i^{\alpha} lnx_i}{[(1+\lambda)-2\lambda[1-(1-x_i^{\alpha})^{\beta}]^{\gamma}]} = 0 \end{aligned} \quad (29)$$

$$\begin{aligned} \frac{n}{\beta} + \sum_{i=1}^{n} ln(1-x_i^{\alpha}) - (\gamma-1)\sum_{i=1}^{n} \frac{(1-x_i^{\alpha})^{\beta} ln(1-x_i^{\alpha})}{[1-(1-x_i^{\alpha})^{\beta}]} \\ +2\lambda\gamma\sum_{i=1}^{n} \frac{[1-(1-x_i^{\alpha})^{\beta}]^{\gamma-1}(1-x_i^{\alpha})^{\beta} ln(1-x_i^{\alpha})}{[(1+\lambda)-2\lambda[1-(1-x_i^{\alpha})^{\beta}]^{\gamma}]} = 0 \end{aligned} \quad (30)$$

$$\frac{n}{\gamma} + \sum_{i=1}^{n} ln[1-(1-x_i^{\alpha})^{\beta}] - 2\lambda\sum_{i=1}^{n} \frac{[1-(1-x_i^{\alpha})^{\beta}]^{\gamma} ln[1-(1-x_i^{\alpha})^{\beta}]}{[(1+\lambda)-2\lambda[1-(1-x_i^{\alpha})^{\beta}]^{\gamma}]} = 0 \quad (31)$$

$$\sum_{i=1}^{n} \frac{1-2[1-(1-x_i^{\alpha})^{\beta}]^{\gamma}}{[(1+\lambda)-2\lambda[1-(1-x_i^{\alpha})^{\beta}]^{\gamma}]} = 0 \quad (32)$$

Hence, the MLEs of the parameters are obtained by solving these nonlinear system of equations. Solving these system of nonlinear equations are complicated, we can therefore use statistical software to solve the equations numerically.

5. Simulation Study and Data Analysis

5.1. Simulation Study

Considering (18), the simulation is done for two instances using different parameter values.The chosen parameter values here are,

- $\alpha = 0.1, \beta = 0.5, \gamma = 0.8, \lambda = 0.01$
- $\alpha = 0.8, \beta = 1.5, \gamma = 3, \lambda = -0.9$

As the n increases, mean square error decreases for the selected parameter values given in table 1 and 2. Moreover, the bias is close to zero as the sample size increases. Thus, as the sample size increases. the estimates become closer to the true parameter values.

Table 1: *Simulation study at α = 0.1,β = 0.5,γ = 0.8,λ= 0.01*

n	Parameter	Estimate	Bias	MSE
25	α	0.0188	-0.0812	0.0066
	β	0.7373	0.2373	0.0563
	γ	2.9549	2.155	4.6439
	λ	-0.2509	-0.2609	0.0681
50	α	0.0212	-0.0788	0.0062
	β	0.6358	0.1359	0.0185
	γ	2.1147	1.3147	1.7284
	λ	-0.1819	-0.2929	0.03681
100	α	0.1132	0.0132	0.0002
	β	0.4486	-0.0514	0.0026
	γ	0.6076	-0.1924	0.0370
	λ	0.1659	0.1559	0.02430
500	α	0.0949	-0.0051	2.55e-05
	β	0.5348	0.0348	0.0012
	γ	0.8485	0.0485	0.0024
	λ	0.1121	0.1021	0.01041
1000	α	0.1027	0.003	7.56e-06
	β	0.5112	0.0116	0.0001
	γ	0.7835	-0.0164	0.0003
	λ	0.0129	0.003	8.81e-06

Table 2: *Simulation study at α = 0.8,β = 1.5,γ = 3,λ= -0.9*

n	Parameter	Estimate	Bias	MSE
25	α	12.3475	11.5475	133.3457
	β	5.6831	4.1831	17.4990
	γ	0.1973	-2.8026	7.8549
	λ	-0.2146	0.6853	0.56964
50	α	5.3842	4.5842	21.0153
	β	2,2110	0.7110	0.50556
	γ	0.3943	-2.6056	6.7896
	λ	-0.3780	0.5219	0.27241
100	α	0.4486	-0.3513	0.1234
	β	1.3816	-0.1183	0.0140
	γ	8.1035	2.1035	4.42471
	λ	-0.4335	0.4665	0.2176
500	α	0.6669	-0.1333	0.0117
	β	1.5409	0.0409	0.0017
	γ	5.6903	2.0577	4.2344
	λ	-0.5098	0.39018	0.1522
1000	α	0.7799	0.0027	7.51e-06
	β	1.4994	0.0012	0.0001
	γ	3.0002	-0.0164	0.0021
	λ	-0.8024	0.0055	0.0004

5.2. Data Analysis

In this section, we demonstrate the usefulness of the proposed Transmuted Exponentiated Kumaraswamy TEKw$(\alpha, \beta, \gamma, \lambda)$ distribution. We fit this distribution to a real life data set and compare the results with some recent efficient models: those corresponding to the Kumaraswamy (Kw) distribution (Kumaraswamy [11]), Transmuted Kumaraswamy (TKw) distribution (Khan et.al. [9]) and Exponentiated Kumaraswamy (EKw) distribution (Lemonte et.al. [12]). The corresponding pdfs are presented below.

- The pdf of the Kw distribution is given by

$$f(x;\alpha,\beta) = \alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}, 0 < x < 1$$

where $\alpha, \beta > 0$

- The pdf of the TKw distribution is given by,

$$f(x;\alpha,\beta,\lambda) = \alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}[(1-\lambda)+2\lambda(1-x^{\alpha})^{\beta}], \quad 0 < x < 1$$

where $\alpha, \beta > 0$ and $|\lambda| < 1$.

- The pdf of EKw distribution is given by,

$$f(x;\alpha,\beta,\gamma) = \alpha\beta\gamma x^{\alpha-1}(1-x^{\alpha})^{\beta-1}[1-(1-x^{\alpha})^{\beta}]^{\gamma-1}, \qquad 0 < x < 1$$

where $\alpha, \beta, \gamma > 0$

Shasta Reservoir capacity data is used for the purpose. The reservoir is located in California, United States. The reservoir has a height of 602 ft (183 m), a length of 3460 ft (1050 m), and a total capacity of 4.552 million acre-ft (5.615 million dam^3). The capacity of the Reservoir (after transformation) for each February from 1991 to 2010 is given by table 3, see Simbolan et. al. [19]. The analysis is carried out using R software.The parameters are estimated by maximum

Table 3: *Shasta Reservoir Capacity Data Each February from 1991 to 2010*

Year	Transformed Capacity	Year	Transformed Capacity
1991	0.338936	2001	0.768007
1992	0.431915	2002	0.843485
1993	0.759932	2003	0.787408
1994	0.724626	2004	0.849868
1995	0.757583	2005	0.69597
1996	0.811556	2006	0.842316
1997	0.785339	2007	0.828689
1998	0.78366	2008	0.580194
1999	0.815627	2009	0.430681
2000	0.847413	2010	0.742563

likelihood method. Akaike information criterion (AIC), the correct Akaike information criterion (CAIC), Bayesian information criterion (BIC), Hannan information criterion (HQIC), -2 ln L, the Kolmogorov-Smirnov (K-S), Cramer-von Mises and Anderson-Darling goodness-of-fit statistic and the p-values are considered to compare the four models which are defined as follows.

$$AIC = -2l + 2k$$

$$BIC = -2l + k\ log(n)$$

$$HQIC = -2l + k\,log(log(n))$$

$$CAIC = -2l + 2kn/(n-k-1)$$

where, l denotes the log-likelihood function, k is the number of parameters and n is the sample size.
The Kolmogorov-Smirnov test is used to decide if a sample comes from a population with a specific distribution. The test statistic is given by,

$$K-S \; statistic \;\; D_n = sup|F(x) - F_n(x)|$$

where $F_n(x)$ is the empirical distribution function.

The Cramer-von Mises criterion for testing that a sample $x_1, x_2, ..., x_n$ has been drawn from a specified continuous distribution F(x)is

$$\omega^2 = \int_{-\infty}^{+\infty} [F_n(x) - F(x)]^2 dF(x) \tag{33}$$

The Anderson-Darling is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and is given by,

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[lnF(X_i) + ln(1 - F(X_{n-i+1}))] \tag{34}$$

where, n is the sample size, F(x) is the cdf for the specified distribution, and i is the i^{th} sample, calculated when the data is sorted in ascending order.
The parameter estimates based on the reservoir capacity data for the four models considered are given by table 4.

Table 4: *The MLEs and log-likelihood (l) estimate of the model parameters for reservoir capacity data.*

Distribution	Parameter Estimates				log-likelihood
	α	β	γ	λ	
Kw	6.891239	5.215555	-	-	15.90481
TKw	6.2451572	5.4432386	-0.5010172	-	16.67048
EKw	24.0913041	64.7302856	0.2273644	-	17.97171
TEKw	24.8114998	83.6984162	0.1786838	-0.5454838	20.32666

The tables 5 gives the estimates of the model parameters, AIC , BIC, CAIC and the HQIC values.

Table 5: *AIC, BIC, CAIC and HQIC statistics of the fitted model in data set*

Distribution	-2l	AIC	BIC	CAIC	HQIC
Kw	-31.80962	-27.80962	-25.62753	-27.17804	-25.53863
TKw	-33.34096	-27.34096	-24.06783	-26.00763	-26.56991
EKw	-35.94341	-29.94341	-26.67028	-28.61008	-29.17236
TEKw	**-40.65332**	**-32.65332**	**-28.28915**	**-30.30038**	**-31.62525**

From table 4, it shows that the proposed Transmuted Exponentiated Kumaraswamy model has a maximum value of log likelihood. Table 5 shows that the proposed model has a minimum values of statistics AIC, BIC, CAIC and HQIC compared to other models. In order to compare the distributions, we had considered the Kolmogorov-Smirnov (K-S) test, Cramer-von Mises and

Table 6: *Test statistic values and corresponding p values*

Distribution	K-S Statistic (p-value)	Anderson-Darling Statistic (p-value)	Cramer-Von Statistic (p-value)
Kw	0.22384 (0.1892)	0.96123 (0.01243)	0.13848 (0.03077)
TKw	0.19077 (0.3543)	0.80139 (0.03188)	0.10962 (0.0768)
EKw	0.18032 (0.4221)	0.65474 (0.07574)	0.085587 (0.1656)
TEKw	**0.16457 (0.5365)**	**0.50999 (0.1762)**	**0.060932 (0.3527)**

Anderson-Darling goodness-of-fit statistics for the Shastha Reservoir Capacity data. From table 6, it is seen that Transmuted Exponentiated Kumaraswamy model has largest p-value based on K-S Statistic, Cramer-von Mises and Anderson-Darling statistic. As the results indicate, the proposed model performed better than other models.

6. Conclusion

In this paper, we have introduced a new generalization of the exponentiated Kumaraswamy distribution called the transmuted exponentiated Kumaraswamy distribution. The graphical representations of its density function,cumulative distribution function, hazard rate function and survival function are obtained. We derived the moments, moment generating function, characteristic function, entropy, mean deviations, quantile function, etc. of the proposed distribution. Estimation of parameters of the distribution is performed using maximum likelihood method. A simulation study is performed to validate the estimates of the model parameters. Finally, TEKw(α,β,γ,λ) distribution is applied to a real data set and compared with other distributions. It is empirically verified that the new TEKw model is a better model than the other competing models.

References

[1] Abd El-Monsef, M. M. E. and Ghoneim, S. A. E. (2015). The weighted kumaraswamy distribution. *Information : An International Interdisciplinary Journal*, 18(8): 3289-3300.

[2] Al-Fattah, A.M. and EL-Helbawy, Abeer and Al-Dayian, Gannat. (2017). Inverted Kumaraswamy distribution: Properties and estimation. *Pakistan Journal of Statistics*, 33(1): 37-61.

[3] Al-Kadim, K. A. and Mahdi, A. A. (2018). Exponentiated transmuted exponential distribution. *Journal of University of Babylon for Pure and Applied Sciences*, 26(2): 78-90.

[4] Barreto-Souza, W. and Lemonte, A. J. (2013). Bivariate Kumaraswamy distribution: properties and a new method to generate bivariate classes. *Statistics*, 47(6): 1321-1342.

[5] Bourguignon, M., Silva, R. B., Zea, L. M. and Cordeiro, G. M. (2013). The Kumaraswamy Pareto distribution.*Journal of Statistical Theory and Applications*, 12(2): 129-144.

[6] Dey, S., Mazucheli, J. and Nadarajah, S. (2018). Kumaraswamy distribution: different methods of estimation. *Computational and Applied Mathematics*, 37(2): 2094-2111.

[7] Fletcher, S. C. and Ponnambalam, K. (1996). Estimation of reservoir yield and storage distribution using moments analysis. *Journal of Hydrology*, 182(1): 259-275.

[8] Hassan, A. S. and Elgarhy, M. (2016). Kumaraswamy Weibull-generated family of distributions with applications. *Advances and Applications in Statistics*, 48(3): 205-239.

[9] Khan, M. S., King, R. and Hudson, I. L. (2016). Transmuted Kumaraswamy Distribution. *Statistics in Transition new series*, 17(2): 183-210.

[10] Khan, M. S., King, R. and Hudson, I. L. (2017). Transmuted Weibull distribution: Properties and estimation. *Communications in Statistics-Theory and Methods*, 46(11): 5394-5418.

[11] Kumaraswamy, P. (1980). A generalized probability density functions for double-bounded random processes. *Journal of Hydrology*, 46(1-2): 79-88.

[12] Lemonte, A. J., Barreto-Souza, W. and Cordeiro, G. M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27(1): 31-53.

[13] Merovci, F. (2013). Transmuted Exponentiated Exponential Distribution. *Mathematical Sciences and Applications E-Notes*, 1(2): 112-122.

[14] Merovci, F. (2016). Transmuted Rayleigh Distribution. *Austrian Journal of statistics*, 42 (1): 21-32.

[15] Mielke, P. W., and Johnson, E. S. (1974). Some generalized beta distributions of the second kind having desirable application features in hydrology and meteorology.*Water Resources Research*, 10(2): 223-226.

[16] Mitnik, P. A. (2013). New properties of the Kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, 42(5): 741-755.

[17] Mohammed, A. S. (2019). Theoretical Analysis of the Exponentiated Transmuted Kumaraswamy Distribution with Application. *Annals of Statistical Theory and Applications (ASTA)*,1: 51-60.

[18] Silva, R., Gomes-Silva, F., Ramos, M., Cordeiro, G., Marinho, P. and De Andrade, T. A. (2019). The exponentiated Kumaraswamy-G class: general properties and application. *Revista Colombiana de Estadstica*, 42(1): 1-33.

[19] Simbolon, H. G., Fithriani, I. and Nurrohmah, S. (2017). Estimation of shape parameter in Kumaraswamy distribution using Maximum Likelihood and Bayes method. *AIP Conference Proceedings*, 1862(1): 030160.

[20] Shaw, W. T. and Buckley, I. R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skewkurtotic-normal distribution from a rank transmutation map. *arXiv preprint* arXiv: 0901.0434.

[21] Sherwania, R. A. K., Waqas, M., Saeed, N., Farooq, M., Ali Raza, M. and Jamal, F. (2021). Transmuted Inverted Kumaraswamy Distribution: Theory and Applications. *Punjab University Journal of Mathematics*, 53(3): 29-45.

[22] Usman, A., Ishaq, A., Tasiu, M., Aliyu Y. and Idris, F.A. (2017). A new Weibull-Kumaraswamy distribution: theory and applications. *Nigerian Journal of Scientific Research*, 16(2): 158-166.

[23] ZeinEldin, R. A., Hashmi, S., Elsehety, M. and Elgarhy, M. (2020). Type II half logistic Kumaraswamy distribution with applications. *Journal of Function Spaces*, 2020: 1-15.

A NONPARAMETRIC CONTROL CHART FOR JOINT MONITORING OF LOCATION AND SCALE

Vijaykumar Ghadage[1] and Vikas Ghute[2*]

•

[1]Department of Statistics
Vidnyan Mahavidyalaya Sangola (MS)-413307, India
vkghadage.stat@ gmail.com

[2]Department of Statistics
Punyashlok Ahilyadevi Holkar
Solapur University Solapur (MS)-413255, India
vbghute_stats@rediffmail.com

[*] Corresponding Author

Abstract

Traditional control charts are based on the assumption that the process observations are normally distributed. However, in many applications, there is insufficient information to justify this assumption. Thus, nonparametric control charts have been designed in literature to monitor location parameter and scale parameter of a process. In this paper, a single nonparametric control chart based on modified Lepage test is proposed for simultaneously monitoring of location and scale parameters of any continuous process distribution. The charting statistic combines two nonparametric test statistics namely Baumgartner test for location and Ansari-Bradely test for scale. The performance of the proposed chart is examined through simulation studies in terms of the mean, the standard deviation, the median and some percentiles of the run length distribution. The average run length (ARL) performance of the proposed chart is compared with that of the existing nonparametric Shewhart-Cucconi (SC) and Shewhart-Lepage (SL) charts for joint monitoring of location and scale.

Keywords: Control chart; average run length; joint monitoring; nonparametric tests; location parameter; scale parameter.

1. Introduction

Control charts are the most important statistical process control tool used to monitor manufacturing processes with the objective of detecting any change in process parameters that may affect the quality of the output. Shewhart $\bar{X}$ and Ror S control charts are most popular control charts for monitoring process mean and process variability. These control charts are easy to implement but are based on the fundamental assumption that the distribution of quality characteristic is normal. In real applications, there are many situations in which process data come from non-normal distribution. In such situations, it is desirable to use nonparametric control charts. The main advantage of nonparametric control chart is that it does not assume any probability distribution for the characteristic of interest. A formal definition of nonparametric or distribution-free control chart is given in terms of its run-length distribution. The number of samples that need to be collected before the first out-of-signal is given by a chart is a random variable called the run-length; the probability distribution of the run-length is referred to as run-length distribution. If the in-control run-length distribution is same for every continuous distribution then the chart is called as distribution-free or nonparametric (Chakraborti and Eryilmaz [1]). The location and scale of a process are the two main parameters most often monitored in nonparametric control charts. The problem of monitoring the location of a process is important in many applications. The location parameter could be the mean or the median or some percentiles of the distribution. Many authors have developed nonparametric control charts to monitor location parameter of the process some of these includes Bakir [2-3], Chakraborti and Eryilmaz [1], Khilare and Shirke [4], Human et al. [5]. These charts are based on sign and/or rank statistics. Chakraborti et al. [6] and Chakraborti and Graham [7] presented an extensive overview of literature on nonparametric control charts and discussed their advantages.

The problem of monitoring the scale parameter of a process is also important in many applications. For monitoring scale parametric of a process very few nonparametric are available in literature. Amin et al. [8] proposed a sign chart for process variation based on quartiles. Das [9] proposed a nonparametric control chart for controlling variability based on squared rank test. Das [10] developed a nonparametric control chart based on rank test. Das and Bhattacharya [11] proposed a control chart for controlling variability based on some nonparametric tests. Murakami and Matsuki [12] developed a nonparametric control chart based on Mood statistic for dispersion. Khilare and Shirke [13] developed a nonparametric synthetic control chart for process variability based on sign statistic. Zombade and Ghute [14] provided nonparametric control charts for process variation based on Sukhatme's test and Mood's test. Shirke and Barale [15] proposed a nonparametric cumulative sum control chart for process dispersion using in-control deciles.

The existing nonparametric control charts are designed for monitoring location and scale by using separate control charts. Using two separate charts for monitoring location and scale can sometimes be difficult in practice for the interpretation of signals because the effect of changes in one of the parameters can affect the changes in other one. The joint monitoring scheme with single chart has received more attention in the recent literature due to simplicity and clarity. A single control chart uses a statistic that is a combination of two separate statistics one each for mean and variance. Joint monitoring of a process involves two parameters, the mean (location) and variance (scale) and typically uses an efficient statistic for monitoring each parameter. The control charts currently available for jointly monitoring the mean and variance are focused on parametric control chart. Cheng and Thaga [16] provided a review of literature on joint monitoring of control charts up to 2005. McCracken and Chakraborti [17] presented an overview of literature on joint monitoring control charts. They also discussed some of the joint monitoring schemes for multivariate processes, autocorrelated data, and individual observations. Most of the parametric control charts for joint monitoring the mean and variability of a process are based on the assumption that process distribution is normal. However, in many applications there is not always enough knowledge or information to support the assumption that process distribution is of specific shape or form such as normal. In such cases nonparametric control charts can be useful. The literature in the area of nonparametric joint monitoring schemes is currently very limited. A few nonparametric joint monitoring schemes are available in the literature. Zou and Tsung [18] developed EWMA control chart based on goodness-of-fit test. It has been shown that the proposed chart is effective for detecting changes in location, scale and shape. Mukherjee and Chakraborti [19] developed a single distribution-free control chart for joint monitoring of location and scale. The chart is based on nonparametric test for location-scale by Lepage [20] which combines the Wilcoxon rank sum (WRS) location statistic and with Ansari-Bradely scale statistic. Chowdhury et al. [21] proposed distribution-free chart based on Cucconi statistic, for joint monitoring of location and scale parameters of continuous distribution. Nonparametric joint monitoring scheme is an important area for research and literature in this area is currently very limited and thus presents a great opportunity for further research. The purpose of this paper is to contribute the research on nonparametric joint monitoring scheme.

In this paper, a single nonparametric Shewhart-type control chart is developed for joint monitoring of location and scale parameters of a continuous process distribution. The proposed chart is based on nonparametric two sample modified Lepage-type test proposed by Neuhäuser [22]. The test combines the Baumgartner statistic and Ansari-Bradely statistic for jointly detecting location and scale changes. The in-control and out-of-control performance of the proposed control chart is evaluated through average run length for the normal and double exponential distributions. The rest of the paper is organized as follows. The nonparametric Baumgartner and Ansari-Bradely tests for location and scale respectively are modified Lepage-type test proposed by Neuhäuser [22] for joint location and scale is are discussed in Section 2. A single nonparametric control chart for simultaneously monitoring the location parameter and the scale parameter of a process based on modified Lepage-type test statistic is presented in Section 3. In-control and out-of-control performance of the proposed control chart is studied in detail in Section 4. Performance of the proposed control chart is compared with the existing nonparametric charts in Section 5. Some conclusions are given in Section 6.

2. Nonparametric Tests for Location and Scale

In this section, we briefly discuss the nonparametric tests for location parameter, scale parameter and jointly location scale parameters.

2.1 Baumgartner two sample test for location

Baumgartner test is a two-sample test can be applied for location and scale parameters. Let $(X_{1,}X_2,\dots,X_n)$ and $(Y_{1,}Y_2,\dots,Y_m)$ denote two random samples. The observations within each sample are independent and identically distributed, and we assume independence between two samples. Let F and G be continuous distribution functions corresponding two populations1 and2 respectively. In location shift, model considered first the distribution functions are same except perhaps for change in their location; that is$F(x) = F(x-\theta)$. The null hypothesis is $H_0: \theta = 0$, whereas alternative is $H_1: \theta \neq 0$. Baumgartner et al. [23] proposed a distribution-free two-sample rank test for general alternative. For combined samples, let $R_1 < R_2 < \dots < R_n$ and $H_1 < H_2 < \dots < H_m$ denote the ranks of the $X-$values and $Y-$values in increasing order of magnitude, respectively. Baumgartner et al. [23] defined a nonparametric two-sample rank statistic B as follows:

$$B = \frac{B_X+B_Y}{2} \tag{1}$$

where $B_X = \dfrac{1}{n}\sum_{i=1}^{n}\dfrac{\left(R_i - \frac{N}{n}i\right)^2}{\frac{i}{n+1}\left(1-\frac{i}{n+1}\right)\left(\frac{m\,N}{n}\right)}$ and $B_Y = \dfrac{1}{m}\sum_{j=1}^{m}\dfrac{\left(H_j - \frac{N}{m}j\right)^2}{\frac{j}{m+1}\left(1-\frac{j}{m+1}\right)\left(\frac{n\,N}{m}\right)}$

The larger value of statistic B gives evidence to reject the null hypothesis. Baumgartner et al. [23] also provided asymptotic distribution of test statistic B.

2.2 Ansari-Bradely test for scale

The Ansari-Bradely test is a two-sample rank test applied for scale parameter. The test statistic is defined as follows: In the combined samples, the observations less than or equal to the median are replaced by their ranks in the increasing order and those larger than the median are replaced by their ranks in descending order. The statistic is the sum of these ranks for the Y sample. The corresponding test statistic is defined as (Gibbons and Chakraborti [24]),

$$AB = \sum_{k=1}^{n}\left(k - \frac{N+1}{2}\right)Z_k \tag{2}$$

The mean and variance of statistic AB is given by,

$$E(AB) = \begin{cases} \frac{m\,(N+1)}{4}, & \text{when N is even} \\ \frac{m\,(N+1)^2}{4N}, & \text{when N is odd} \end{cases} \quad \text{and} \quad V(AB) = \begin{cases} \frac{m\,n\,(N^2-4)}{48\,(N-1)}, & \text{when N is even} \\ \frac{m\,n\,(N+1)\,(N^2+3)}{48\,N^2}, & \text{when N is odd} \end{cases}$$

2.3 Modified Lepage-type test for location and scale

After Lepage statistic was proposed, various Lepage-type statistics have been proposed and discussed by many authors in the literature. One of the most famous and powerful modified Lepage-type statistic proposed by Neuhäuser [22] is a combination of the Baumgartner and Ansari-Bradely statistic given as:

$$L_M = \left(\frac{B-E_0(B)}{\sqrt{Var_0(B)}}\right)^2 + \left(\frac{AB-E_0(AB)}{\sqrt{Var_0(AB)}}\right)^2 \tag{3}$$

where B is Baumgartner statistic for location shift and AB is Ansari-Bradely statistic for scale shift. In this paper, we useL_M test statistic as a charting statistic for detecting simultaneous location and scale shifts in a continuous process distribution.

3. Control chart based on modified Lepage-type statistic

In this Section, we develop a nonparametric control chart based on modified Lepage-type test statistic proposed by Neuhäuser [22] for simultaneously monitoring the location and the scale parameters of a continuous process. The single plotting statistic for the joint monitoring of location and scale is given by L_M in Eq. (3) and chart is called LM chart. To adopt the idea of two sample test for control chart implementation, m independent observations $(X_1, X_2, \ldots, X_m)$ from an in-control process are used as reference sample and compared to future sample subgroups of n independent observations $(Y_1, Y_2, \ldots, Y_n)$ an arbitrary test sample.

The proposed LM control chart for joint monitoring of location and scale is constructed as follows:

Step1: Collect Phase-I reference sample $X = (X_1, X_2, \ldots, X_m)$ of size m from an in-control process.
Step2: Let $Y = (Y_1, Y_2, \ldots, Y_n)$ be j^{th} Phase-II (test) sample of size n, $j = 1, 2, 3, \ldots$
Step 3: Calculate B_j and $(AB)_j$ using (1) and (2) for j^{th} test sample.
Step 4: Compute means and standard deviations of B and AB statistics respectively
Step 5: Calculate the standardized B and AB statistics as

$$T_{1j} = \left(\frac{B - E_0(B)}{\sqrt{Var_0(B)}}\right) \text{ and } T_{2j} = \left(\frac{AB - E_0(AB)}{\sqrt{Var_0(AB)}}\right) \text{ respectively.}$$

Step 6: Calculate the control chart statistic LM chart as $T_j = T_{1j}^2 + T_{2j}^2$, $j = 1, 2, 3, \ldots$
Step 7: Plot T_j against an upper control limit $(UCL), H > 0$.
Step 8: If T_j exceed H, the process is out-of-control at the j^{th} test sample. If not, the process is thought to be in-control and testing continues to the next sample.

4. Performance evaluation and analysis of LM chart

Implementation of the proposed LM chart requires the upper control limit H. Typically, in practice, it is determined for some specified in-control average run length (ARL_0), say, 370 or 500. A Monte-Carlo simulation approach based on sufficiently large number of possible samples is used to determine H. For a given pair of (m, n) values, a search is conducted with different values of H, and that value of H is obtained for which ARL_0 is equal to nominal (target) value. We choose $m = 30, 50, 100$ for the reference sample size and $n = 5, 11, 25$ as the test sample size and target values $ARL_0 = 200, 370, 500$. The results are presented in Table 1.

Table 1: *Charting constant H for the LM chart for some standard (target) values of* ARL_0

Reference sample size (m)	Test sample size (n)	Upper control limit (H)		
		$ARL_0 = 200$	$ARL_0 = 370$	$ARL_0 = 500$
30	5	29.540	35.242	37.960
30	11	25.050	33.128	37.312
30	25	16.985	22.089	24.820
50	5	14.510	19.510	22.390
50	11	15.389	18.712	20.752
50	25	15.798	19.123	20.910
100	5	20.020	29.050	32.800
100	11	20.740	27.490	31.305
100	25	18.540	24.450	28.023

The performance of a control chart is generally studied through its runlength distribution. If the runlength distribution is skewed to the right, it is useful to come across at various measures such as average run length (ARL), the standard deviation of run length (SDRL) and several percentiles including the first and third quartiles to characterize the distribution. We study the performance of the proposed LM chart both under in-control and out-of-control setup. For the in-control setup, we simulate both the reference and the test sample from standard normal distribution. We choose $m = 30, 50, 100$ and $n = 5, 11, 25$. For a given pair of (m, n) values, we obtain upper control limits H for nominal (target) $ARL_0 = 500$ and simulate different characteristics of the in-control run-length distribution. The results of simulation are shown in Table 2.
It indicates that the target $ARL_0 = 500$ is much larger than the median (Q_2) for all (m, n) combinations. Hence, in-control run-length distribution of the LM chart is highly skewed to the right.

In order to investigate the out-of-control performance of the proposed LM chart, we consider the underlying process distributions as normal and double exponential. The double exponential distribution is considered as process distribution to study the effect of heavy tailed distribution on the performance of the LM chart. The distribution of observations from the process is considered to have mean zero and variance one for both the process distributions under study.

Table 2: *In-control performance characteristics of the LM chart for* $ARL_0 = 500$.

m	n	H	ARL_0	$SDRL_0$	P_5	Q_1	Q_2	Q_3	P_{95}
30	5	37.960	501.0	500.5	26	146	350	694	1484
30	11	37.312	499.7	499.2	27	145	346	692	1493
30	25	24.820	500.4	499.9	26	144	347	695	1481
50	5	22.390	499.5	499.0	27	143	344	691	1508
50	11	20.752	501.4	500.9	26	144	351	698	1499
50	25	20.910	501.3	500.8	26	143	348	692	1502
100	5	32.800	500.6	500.1	26	144	348	696	1506
100	11	31.305	500.1	499.6	26	145	345	694	1499
100	25	28.023	502.9	502.4	26	147	349	695	1508

4.1 Performance analysis of LM chart under normal distribution

In order to investigate the out-of-control performance of the proposed LM chart, we consider the underlying process distribution as normal; samples are taken from $N(\theta, \lambda)$ distribution, with in-control samples coming from $N(0, 1)$distribution. To examine the effects of shifts in process parameters, 30 combinations of (θ, λ)values are considered with $\theta = 0, 0.25, 0.5, 1.0, 1.5, 2.0$ and $\lambda = 1.0, 1.25, 1.5, 1.75, 2.0$.

Tables 3 and 4 present the performance characteristics of the LM chart when underlying process distribution is normal with combinations of the reference and test sample sizes $m = 50, 100$ and $n = 5$.

The results in Table 3 and Table 4 indicate that the out-of-control run-length distributions are also skewed to right. It is observed that, for a fixed m, n and a given ARL_0, the out-of-control ARL values as well as the percentiles all decrease sharply with increasing shift in the location and also with the increasing shift in the scale. It indicates that the proposed LM chart is effective in detecting shifts in location and/or in the scale. The proposed LM chart detect shift in the scale more quickly than that in the location. For example, from Table 3, we observe that for 25% increase in location when scale is in-control, the ARL decreases by 68%, whereas for a 25% increase in a scale when the location is in-control, ARL decreases by 78%. Finally, when location and scale increases by 25% the ARL decreased by 88%. The pattern is same for SDRL; it decreases for an increase in the shift in both parameters, but decreases more for a shift in scale. For example, from Table 3, for 25% increase in location, the SDRL decreases by 68% but for 25% increase in scale, the SDRL decreases by 78%.

Table 3: *Performance characteristics of the LM chart for the normal distribution.*
$(ARL_0 = 500, m = 50 \text{ and } n = 5)$.

θ	λ	ARL	SDRL	P_5	Q_1	Q_2	Q_3	P_{95}
0.0	1.0	499.5	499.0	27	143	344	691	1508
0.25	1.0	160.3	159.8	9	46	111	223	478
0.5	1.0	42.4	41.9	3	13	30	59	126
1.0	1.0	5.9	5.4	1	2	4	8	16
1.5	1.0	1.9	1.3	1	1	1	2	5
2.0	1.0	1.2	0.5	1	1	1	1	2
0.0	1.25	108.1	107.6	6	31	76	150	322
0.25	1.25	58.4	57.9	3	17	41	81	175
0.5	1.25	24.1	23.6	2	7	17	33	72
1.0	1.25	5.7	5.1	1	2	4	8	16
1.5	1.25	2.2	1.7	1	1	2	3	6
2.0	1.25	1.3	0.7	1	1	1	2	3
0.0	1.5	43.0	42.5	3	13	30	60	127

0.25	1.5	30.6	30.1	2	9	21	42	90
0.5	1.5	16.6	16.1	1	5	12	23	48
1.0	1.5	5.5	4.9	1	2	4	7	15
1.5	1.5	2.5	1.9	1	1	2	3	6
2.0	1.5	1.5	0.9	1	1	1	2	3
0.0	1.75	22.8	22.3	2	7	16	31	68
0.25	1.75	18.8	18.2	1	6	13	26	55
0.5	1.75	12.7	12.2	1	4	9	17	37
1.0	1.75	5.3	4.8	1	2	4	7	15
1.5	1.75	2.7	2.2	1	1	2	4	7
2.0	1.75	1.7	1.1	1	1	1	2	4
0.0	2.0	14.5	14.0	1	5	10	20	42
0.25	2.0	13.1	12.6	1	4	9	18	39
0.5	2.0	9.9	9.4	1	3	7	13	29
1.0	2.0	5.1	4.6	1	2	4	7	14
1.5	2.0	2.9	2.3	1	1	2	4	8
2.0	2.0	1.9	1.3	1	1	1	2	4

Table 4: *Performance characteristics of the LM chart for normal distribution.*
($ARL_0 = 500,\ m = 100\ and\ n = 5$).

θ	λ	ARL	SDRL	P_5	Q_1	Q_2	Q_3	P_{95}
0.0	1.0	500.6	500.1	26	144	348	696	1506
0.25	1.0	273.6	273.1	15	79	191	380	814
0.5	1.0	66.4	65.9	4	20	46	91	199
1.0	1.0	7.5	7.0	1	3	5	10	21
1.5	1.0	2.1	1.5	1	1	2	3	5
2.0	1.0	1.2	0.5	1	1	1	1	2
0.0	1.25	119.9	119.4	7	35	83	166	355
0.25	1.25	83.8	83.3	5	25	58	116	251
0.5	1.25	33.3	32.8	2	10	23	46	98
1.0	1.25	6.8	6.3	1	2	5	9	19
1.5	1.25	2.4	1.9	1	1	2	3	6
2.0	1.25	1.4	0.7	1	1	1	2	3
0.0	1.5	48.6	48.1	3	14	34	67	145
0.25	1.5	39.6	39.1	2	12	27	55	118
0.5	1.5	21.5	21.0	2	7	15	30	63
1.0	1.5	6.4	5.9	1	2	5	9	18
1.5	1.5	2.7	2.2	1	1	2	4	7
2.0	1.5	1.6	1.0	1	1	1	2	4
0.0	1.75	26.3	25.8	2	8	18	36	78
0.25	1.75	23.3	22.8	2	7	16	32	69
0.5	1.75	15.5	15.0	1	5	11	21	45
1.0	1.75	6.1	5.6	1	2	4	8	17
1.5	1.75	2.9	2.4	1	1	2	4	8
2.0	1.75	1.8	1.2	1	1	1	2	4
0.0	2.0	16.5	16.0	1	5	12	23	49
0.25	2.0	15.4	14.9	1	5	11	21	45
0.5	2.0	11.8	11.3	1	4	8	16	35
1.0	2.0	5.8	5.2	1	2	4	8	16

1.5	2.0	3.1	2.5	1	1	2	4	8
2.0	2.0	2.0	1.4	1	1	1	2	5

4.2 Performance analysis of LM chart under double exponential distribution

To study the effect of heavy tailed distribution on the performance of the proposed LM chart, double exponential distribution is included in the study as heavy tailed process distribution. We conduct simulation study with data from double exponential distribution. The performance characteristics of the run-length are evaluated when the in-control sample is from double exponential with mean 0 and variance 1, and test samples are generated from the double exponential distribution with mean θ and standard deviation λ.

To examine the effects of shifts in location and scale, as in normal case, we studied 30 combinations of (θ, λ)values. Table 5 and Table 6 presents the performance characteristics of the proposed LM chart when underlying process distribution is double exponential with combinations of reference and test samples of size $m = 50, 100$and $n = 5$.

Table 5: *Performance characteristics of LM chart for double exponential distribution.* $(ARL_0 = 500, m = 50 \ and \ n = 5)$.

θ	λ	ARL	SDRL	P_5	Q_1	Q_2	Q_3	P_{95}
0.0	1.0	499.6	499.1	25	143	345	693	1505
0.25	1.0	91.5	91.0	5	27	64	127	272
0.5	1.0	18.2	17.7	1	6	13	25	53
1.0	1.0	3.0	2.4	1	1	2	4	8
1.5	1.0	1.5	0.8	1	1	1	2	3
2.0	1.0	1.1	0.4	1	1	1	1	2
0.0	1.25	222.2	221.7	12	64	155	307	663
0.25	1.25	59.7	59.2	4	18	42	82	178
0.50	1.25	17.0	16.5	1	5	12	23	50
1.0	1.25	3.6	3.0	1	1	3	5	10
1.5	1.25	1.7	1.1	1	1	1	2	4
2.0	1.25	1.3	0.6	1	1	1	1	2
0.0	1.5	128.6	128.1	7	38	90	177	381
0.25	1.5	45.5	45.0	3	14	32	63	136
0.5	1.5	16.1	15.6	1	5	11	22	47
1.0	1.5	4.2	3.6	1	2	3	6	11
1.5	1.5	2.0	1.5	1	1	2	3	5
2.0	1.5	1.4	0.8	1	1	1	2	3
0.0	1.75	87.1	86.6	5	26	61	121	261
0.25	1.75	37.0	36.5	2	11	26	51	110
0.50	1.75	15.8	15.3	1	5	11	22	46
1.0	1.75	4.7	4.2	1	2	3	6	13
1.5	1.75	2.3	1.8	1	1	2	3	6
2.0	1.75	1.6	1.0	1	1	1	2	4
0.0	2.0	65.0	64.5	4	19	45	90	193
0.25	2.0	31.8	31.3	2	10	22	44	94
0.50	2.0	15.4	14.9	1	5	11	21	45
1.0	2.0	5.2	4.7	1	2	4	7	14
1.5	2.0	2.6	2.1	1	1	2	3	7
2.0	2.0	1.8	1.2	1	1	1	2	4

Table 6: *Performance characteristics of LM chart for double exponential distribution* $(ARL_0 = 500, \quad m = 100 \ and \ n = 5)$.

θ	λ	ARL	SDRL	P_5	Q_1	Q_2	Q_3	P_{95}
0.0	1.0	497.9	497.4	25	142	344	691	1502
0.25	1.0	789.9	789.4	42	225	546	1098	2376
0.5	1.0	194.6	194.1	10	56	134	269	583
1.0	1.0	10.5	10.0	1	3	7	14	30
1.5	1.0	2.2	1.6	1	1	2	3	5
2.0	1.0	1.2	0.5	1	1	1	1	2
0.0	1.25	179.6	179.1	10	52	125	249	536
0.25	1.25	246.2	245.7	13	71	170	341	739
0.5	1.25	91.5	91.0	5	27	64	126	272
1.0	1.25	9.6	9.1	1	3	7	13	27
1.5	1.25	2.5	2.0	1	1	2	3	6
2.0	1.25	1.4	0.7	1	1	1	2	3
0.0	1.5	91.1	90.6	5	26	64	127	272
0.25	1.5	111.4	110.9	6	32	77	154	332
0.5	1.5	55.5	55.0	3	16	38	77	164
1.0	1.5	9.3	8.7	1	3	7	13	27
1.5	1.5	2.8	2.3	1	1	2	4	7
2.0	1.5	1.5	0.9	1	1	1	2	3
0.0	1.75	54.4	53.9	3	16	38	75	162
0.25	1.75	63.3	62.8	4	19	44	88	188
0.5	1.75	37.8	37.3	2	11	26	52	112
1.0	1.75	8.7	8.2	1	3	6	12	25
1.5	1.75	3.1	2.5	1	1	2	4	8
2.0	1.75	1.7	1.1	1	1	1	2	4
0.0	2.0	36.7	36.2	2	11	26	51	109
0.25	2.0	41.2	40.7	3	12	29	57	123
0.5	2.0	27.9	27.4	2	8	20	39	83
1.0	2.0	8.3	7.8	1	3	6	11	24
1.5	2.0	3.3	2.7	1	1	2	4	9
2.0	2.0	1.8	1.2	1	1	1	2	4

From Tables 5 and 6, it is observed that when underlying process distribution is doubling exponential, the general pattern remains the same as in the case of normal distribution. However, the out-of-control ARL values for detecting a shift in the mean and/or variance under double exponential distribution are larger than that of the ARL values under normal process distribution. For example, from Table 6, mean shift is 50% ($\theta = 0.50$) and dispersion shift is 50% ($\lambda = 1.5$), The ARL is 194.6 which is larger than 66.4 in the normal case of Table 4. It indicates that the proposed LM chart detects shifts in process location and scale slower under heavy tailed distribution. Moreover, the percentiles as well as SDRL all increase under double exponential distribution as compared with normal distribution.

5. Performance comparison with existing control charts

In this section, the performance of the proposed LM chart is compared with that of the SL chart by Mukherjee and Chakraborti [19] and SC chart by Chowdhury et al. [21] when underlying process distributions are normal and double exponential. Table 7 presents the ARL performance of SC chart, SL chart and LM chart for normal distribution with reference sample size $m = 50, 100$ and test sample of size $n = 5$.

Table7: *Performance comparisons between SC, SL and LM charts for the normal distribution with $ARL_0 = 500$.*

θ	λ	$m = 50, n = 5$			$m = 100, n = 5$		
		SC chart	SL chart	LM chart	SC chart	SL chart	LM chart
0.0	1.0	497.3	499.6	499.5	509.4	513.0	500.6
0.5	1.0	92.2	94.7	42.4	68.6	66.5	66.4
1.0	1.0	8.5	9.3	5.9	7.7	7.7	7.5
1.5	1.0	2.2	2.3	1.9	2.1	2.1	2.1
2.0	1.0	1.2	1.3	1.2	1.2	1.2	1.2
0.0	1.25	71.1	106.2	108.1	74.5	102.9	119.9
0.5	1.25	27.6	35.4	24.1	26.2	30.9	33.3
1.0	1.25	6.6	7.4	5.7	6.2	6.7	6.8
1.5	1.25	2.4	2.6	2.2	2.4	2.5	2.4
2.0	1.25	1.4	1.4	1.3	1.3	1.4	1.4
0.0	1.5	22.8	36.82	43.0	24.3	37.5	48.6
0.5	1.5	13.3	19.0	16.6	13.4	17.8	21.5
1.0	1.5	5.2	6.5	5.5	5.3	6.1	6.4
1.5	1.5	2.4	2.8	2.5	2.4	2.7	2.7
2.0	1.5	1.5	1.6	1.5	1.5	1.6	1.6
0.0	1.75	10.9	18.5	22.8	11.7	19.1	26.3
0.50	1.75	8.1	12.1	12.7	8.4	12.1	15.5
1.0	1.75	4.4	5.7	5.3	4.4	5.5	6.1
1.5	1.75	2.5	2.9	2.7	2.4	2.8	2.9
2.0	1.75	1.6	1.8	1.7	1.6	1.8	1.8
0.0	2.0	6.6	11.3	14.5	7.1	11.5	16.5
0.5	2.0	5.5	8.5	9.9	5.8	8.6	11.8
1.0	2.0	3.7	4.9	5.1	3.8	4.8	5.8
1.5	2.0	2.4	2.9	2.9	2.4	2.9	3.1
2.0	2.0	1.7	1.9	1.9	1.7	1.9	2.0

Examination of Table 7 that for normal distribution leads the following findings:

- For location shifts only when the scale parameter is in-control, the proposed LM chart performs better than the SL and SC charts.
- For scale shifts only when the location parameter is in-control, the proposed LM chart is not as much better as the SL and SC charts.
- For reference sample of size $m = 50$, for any given shift in location parameterθ with a fixed shift in scale parameter as $\lambda = 1.25$, the proposed LM chart performs better than the SL and SC charts. As shift in scale parameter λincreases to 1.5 with any given shift in location parameter θ, the proposed LM chart is efficient than the SL chart only. For scale shift of size $\lambda = 1.25$ and location shift $\theta = 1.5$ and $2.0$the proposed LM chart is efficient than the SL chart only. For scale shift $\lambda = 2.0$ and location shift $\theta = 1.5$ and $2.0$the proposed LM chart is equally efficient to the SL chart only.

- For reference sample of size $m = 100$, for detecting shift in location parameter as $\theta = 1.5$ and $2.0$and shift in scale parameter $\lambda = 1.5\ and\ 1.5$ the proposed LM chart is equally efficient to the SL chart only.

Table 8: *Performance comparison between the SC, SL and LM charts for the double exponential distribution with* $ARL_0 = 500$.

θ	λ	$m = 50, n = 5$			$m = 100, n = 5$		
		SC chart	SL chart	LM chart	SC chart	SL chart	LM chart
0.0	1.0	492.7	493.2	499.6	509.6	508.3	497.9
0.5	1.0	240.0	235.2	18.2	191.0	159.2	194.6
1.0	1.0	41.4	36.1	3.0	26.5	19.9	10.5
1.5	1.0	7.2	5.93	1.5	4.8	4.1	2.2
2.0	1.0	2.1	2.0	1.1	1.8	1.7	1.2
0.0	1.25	118.0	156.8	222.2	124.5	153.2	179.6
0.5	1.25	69.7	79.8	17.0	61.7	66.19	91.5
1.0	1.25	20.1	19.9	3.6	14.6	14.0	9.6
1.5	1.25	5.1	5.2	1.7	4.4	4.2	2.5
2.0	1.25	2.1	2.2	1.3	2.0	2.0	1.4
0.0	1.5	43.3	65.9	128.6	47.8	66.8	91.1
0.5	1.5	29.3	42.1	16.1	29.6	36.8	55.5
1.0	1.5	12.0	14.2	4.2	10.7	11.1	9.3
1.5	1.5	4.5	4.7	2.0	4.0	4.1	2.8
2.0	1.5	2.2	2.3	1.4	2.1	2.2	1.5
0.0	1.75	22.8	35.6	87.1	24.4	36.4	54.4
0.5	1.75	16.7	24.5	15.8	16.9	23.2	37.8
1.0	1.75	8.5	10.4	4.7	7.9	9.2	8.7
1.5	1.75	4.0	4.5	2.3	3.7	4.0	3.1
2.0	1.75	2.2	2.4	1.6	2.1	2.3	1.7
0.0	2.0	13.8	22.1	65.0	14.5	22.9	36.7
0.5	2.0	11.1	17.0	15.4	11.3	16.6	27.9
1.0	2.0	6.5	8.6	5.2	6.3	7.9	8.3
1.5	2.0	3.5	4.3	2.6	3.5	3.9	3.3
2.0	2.0	2.2	2.5	1.8	2.1	2.3	1.8

Examination of Table 8 that for double exponential distribution leads the following findings:

- For reference sample of size $m = 50$, for location shifts only when the scale parameter is in-control, the proposed LM chart performs better than the SL and SC charts. For scale shifts only when the location parameter is in-control, the proposed LM chart is not as better as the SL and SC charts. For any given shift in location parameter θ with any shift in scale parameter λ, the proposed LM chart performs better than the SL and SC charts.
- For reference sample of size $m = 100$, for detecting shift in location parameter as $\theta = 1.0$ and 2.0 and shift in scale parameter $\lambda = 1.25, 1.5$ and 1.75, the proposed LM chart is efficient than the to the SL and SC charts. For detecting shift in location parameter as $\theta = 1.0$ and 2.0 and shift in scale parameter $\lambda = 2.0$, the proposed LM chart is efficient than the to the SL and SC charts.

6. Conclusions

In this paper, a single nonparametric control chart based on modified Lepage-type test statistic is developed for joint monitoring of location and scale parameters of a continuous process distribution. Both in-control and out-of-control performance of the chart are studied under normal and heavy tailed double exponential distributions. The various performance characteristics such as mean, median and some percentiles of the run-length distribution are examined. It is observed that the proposed LM chart maintain its designed in-control ARL under the considered process distributions. The chart is found to be more efficient under normal distribution as compared to double exponential distribution. The performance of the proposed chart is compared with SL chart by Mukherjee and Chakraborti [19] and SC chart by Chowdhury [21]. It is observed that the proposed LM chart for joint monitoring of location and scale performs better than the SL and SC charts in some situations.

Acknowledgment

The authors would like to thank the editor for his comments and suggestions which contributed to the improvement of this article.

References

[1] Chakraborti, S. and Eryilmaz, S. (2007). A nonparametric Shewhart-type signed-rank control chart based on runs. *Communications in Statistics-Simulation and Computation*, 36: 335-356.

[2] Bakir, S. T. (2004). A distribution-free Shewhart quality control chart based on signed-ranks.*Quality Engineering*, 16: 613-623.

[3] Bakir, S. T. (2006). Distribution-free quality control charts based on signed rank like statistics. *Communications in Statistics-Theory and Methods*, 35: 743-757.

[4] Khilare, S. K. and Shirke, D. T. (2010). A nonparametric synthetic control chart using sign statistic. *Communications in Statistics-Theory and Methods*, 39: 3282-3293.

[5] Human, S. W., Chakraborti, S. and Smit, C. F. (2010). Nonparametric Shewhart-type sign control chart based on runs. *Communications in Statistics-Theory and Methods*, 39: 2046-2062.

[6] Chakraborti, S., van der Laan, P. and Bakir, S. T. (2001). Nonparametric control charts: An overview and some results. *Journal of Quality Technology*, 33(3): 304-315.

[7] Chakraborti, S. and Graham, M. A. (2007). Nonparametric control charts. *Encyclopedia of Statistics in Quality and Reliability*, John Wiley: New York, 1:415-429.

[8] Amin, R. W., Reynolds, M. R. Jr. and Bakir, S. T. (1995). Nonparametric quality control charts based on sign statistic. *Communications in Statistics-Theory and Methods*, 24: 1579-1623.

[9] Das, N. (2008a). A nonparametric control chart for controlling variability based on squared rank test. *Journal of Industrial and System Engineering*, 2(2): 114-125.

[10] Das, N. (2008b). Nonparametric control chart for controlling variability based on rank test. *Economic Quality Control*, 23(2): 227-242.

[11] Das, N. and Bhattacharya, A. (2008). A new nonparametric control chart for controlling variability.*Quality Technology and Quantitative Management*, 5(4): 351-361.

[12] Murakami, M. and Matsuki, T. (2010). A nonparametric control chart based on Mood statistic for dispersion. *International Journal of Advanced Manufacturing Technology*, 49: 757-763.

[13] Khilare, S. K. and Shirke, D. T. (2012). Nonparametric synthetic control chart for process variation. *Quality and Reliability Engineering International*, 28: 193-202.

[14] Zombade, D. M. and Ghute, V. B. (2014). Nonparametric CUSUM charts for process variability. *Journal of Academia and Industrial Research*, 3(1): 53-57.

[15] Shirke, D. T. and Barale, M. S. (2018). A nonparametric CUSUM chart for process dispersion.*Quality and Reliability Engineering International*, 34(5): 858-866.

[16] Cheng, S. W. and Thaga, K. (2006). Single variable control charts: An overview. *Quality and Reliability Engineering International*, 22(7): 811-820.

[17] McCracken, A. K. and Chakraborti, s. (2013). Control charts for joint monitoring of mean and variance: An overview. *Quality Technology and Quantitative Management*, 10(1): 17-36.

[18] Zou, C. and Tsung, F. (2010). Likelihood ratio based distribution-free EWMA control charts. *Journal of Quality Technology*, 42(2): 174-196.

[19] Mukherjee, A. and S Chakraborti, S. (2012). A Distribution-free control chart for the joint monitoring of location and scale. *Quality and Reliability Engineering International*, 28: 335-352.

[20] Lepage, Y. A. (1971). Combinations of Wilcoxon's and Ansari-Bradley's statistics, *Biometrika,* 58: 213-217.

[21] Chowdhury, S., Mukherjee, A. and Chakraborti, S. (2013). A new distribution-free control chart for joint monitoring of unknown location and scale parameters of continuous distributions.*Quality and Reliability Engineering International*, 30(2): 191-204.

[22] Neuhäuser, M. (2000). An exact two-sample test based on the Baumgartner–Weiss–Schindler statistic and a modification of Lepage's test.*Communications Statistics- Theory and Methods*, 29: 67–78.

[23] Baumgartner, W., Weiß, P. and Schindler, H. A. (1998). A nonparametric test for the general two-sample problem.*Biometrics*, 54: 1129-1135.

[24] Gibbons and Chakraborti, S. (2003). Nonparametric Statistical Inference, 4th Edition, New York: Dekker

Maintenance policy costs considering imperfect repairs

Allan Jonathan da Silva

•

Production engineering, CEFET/RJ
Federal Center for Technological Education, Itaguaí, 23812-101, RJ, Brazil.
allan.jonathan@cefet-rj.br

Abstract

Objective: This paper extends the analysis of imperfect preventive maintenance interspaced with minimal repairs. The aim is to find the intervals of future scheduled maintenance actions considering different recovery factors and costs. Methods: The optimal preventive maintenance scheduling are obtained by minimizing the overall maintenance costs. Minimal repairs interspersed with scheduled imperfect preventive maintenance actions are considered. The model parameters of the power law process are estimated using the maximum likelihood estimation method and a Differential Evolution algorithm is used to solve the maximization problem. Results: The optimal preventive maintenance periods for different levels of maintenance restoration with respect to corrective and preventive maintenance costs are found. Graphs are drawn to highlight the effect of future maintenance costs and the hazard function paths. It is shown that the preventive maintenance becomes more frequent as the equipment ages and the hazard function increases. Also, it is perceived that the scheduled maintenance intervals become shorter as the corrective maintenance becomes more expensive. Conclusion: A hazard rate model which considers minimal repairs interspersed with scheduled imperfect preventive maintenance provides a useful tool for defining the optimal maintenance policy. The results obtained in this paper show that maintenance cost varies widely according to the recovery factor of the maintenance action and that the optimal interval of two consecutive preventive maintenance actions strongly depends on the costs.

Keywords: Reliability, imperfect maintenance, proportional age reduction model, maintenance costs, power law model.

1. Introduction

Recently statisticians and engineers have paid a lot of attention to reliability centered maintenance and its cost assessment. As stated by Löfsten (2000), the overall costs of maintenance, estimated to be between 15% and 40% of production costs, and the trend toward industry automation has forced engineers and managers to pay more attention to maintenance policies.

There are several papers in the recent literature that have attempted to estimate the failure probability distributions implied by different maintenance policies. Researchers have developed a wide variety of models to deal with maintenance policy optimization. Performance and condition-based maintenance models can be found in Dui et al. (2023), Azizi and Salari (2023) and Chen et al. (2022). Preventive maintenance policies with degradation models can be found in Wei et al. (2023) and Li et al. (2023). Predictive maintenance models can be found in Huynh et al. (2022) and Guo and Liang (2022) and new probability distributions have been studied, such as in S. and Sebastian (2022). Also, maintenance models are under constantly development, as can be found in Tijjani A. Waziri (2022) for a replacement policy, in Naveen Kumar (2022), Shanti Parkash (2022) and Neetu Dabas (2022) for priority repair policies and in Nse Udoh (2022) studying maintenance policies for non-repairable products. Even modern techniques, such as Artificial-intelligence-based model can be found in Nguyen et al. (2022).

In the present study, it is analyzed the consequence of minimal corrective repairs interspersed with imperfect schedulued preventive maintenance. Once a scheduled maintenance is performed on an equipment, it will be restored to a state that is between as good as new and as bad as old. Under this assumption, the Proportional Age Reduction model was proposed in Malik (1979). The paper's objective is to refine the study made by Shin et al. (1996) including the cost analysis and the possibility to change the level of the equipment's regeneration after maintenance actions.

The remaining part of this paper is organized as follows. In Section 2 it is made a review of the Reliability theory, it is shown the maximum likelihood estimation of the power law process under the PAR model and it is presented the expected cost of the maintenance policy. In Section 3 it is described the failure and maintenance actions data. Also, the maximum likelihood estimators are calculated. In Section 4 the results is extended. It is analyzed the optimal preventive maintenance interval under different recovery parameters. The overall cost is also analyzed. Section 5 concludes the paper.

2. Reliability Theory

The purpose of the use of reliability theory is to assist management in decision making by using known quantitative facts effectively and by reducing the reliance on subjective judgement (Löfsten (2000)).

A formal definition of reliability is given by Elsayed (2021): *"Reliability is the probability that a product will operate or a service will be provided properly for a specified period of time (design life) under the design operating conditions (such as temperature, load, volt. . .) without failure."*

As the probability theory is the foundation of the reliability engineering and of the reliability centered maintenance, we review the following definitions that can be found in William Q. Meeker (2021) and in in Elsayed (2021).

Let $f(t)$ be a real function such that

$$f(t) \geq 0 \quad \forall t \geq 0$$

and

$$\int_0^\infty f(s)ds = 1.$$

Then, $f(t)$ is a failure probability density function. The probability of failure up to time t is given by

$$F(t) = \int_0^t f(s)ds, \tag{1}$$

so that the reliability function is given by

$$C(t) = 1 - F(t) = \int_t^\infty f(s)ds. \tag{2}$$

The following function

$$h(t) = \lim_{\Delta t \to 0} \frac{C(t) - C(t + \Delta t)}{\Delta t C(t)} = \frac{1}{C(t)} \left[-\frac{d}{dt} C(t) \right] = \frac{f(t)}{C(t)} \tag{3}$$

is called the hazard function. Considering that only minimal repairs are performed when the equipment fails, the expected number of failures in $[0, t]$ is given by

$$H(t) = \int_0^t h(s)ds. \tag{4}$$

The reliability function can be also calculated by

$$C(t) = e^{\int_0^t h(s)ds}, \tag{5}$$

and the expectation of T is defined as the mean time to failure (MTTF):

$$MTTF = \int_0^\infty C(s)ds = \int_0^\infty sf(s)ds. \quad (6)$$

As an example, the Weibull distribution is given by the following probability density function

$$f(t;\alpha,\beta) = \alpha\beta t^{\beta-1}\exp\{-\alpha t^\beta\}1_{\{t>0\}}, \quad (7)$$

for fixed parameters α and β, scale and shape, respectively. A plot of the probability density function $f(t)$, the reliability function 5, the hazard function 3 and cumulative distribution function 1 for the Weibull distributions with different parameters are shown in Figure 1.

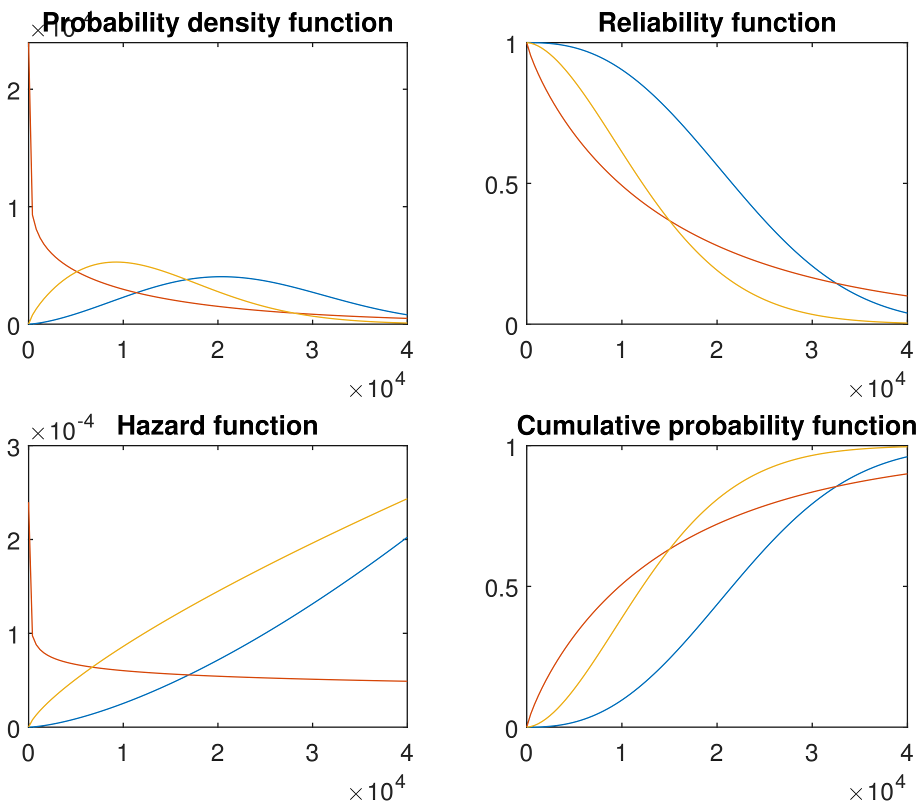


Figure 1: *Probability functions for the Weibull distribution*

2.1. Maintenance policy cost

Following Pham (2003), maintenance can be defined as actions to control the deterioration process leading to failure of a system - called preventive maintenance, and to restore the system to its operational state through corrective actions after a failure - called corrective maintenance. The behavior of the equipment after a repair depends on the type of repair carried out.

As stated by Huynh et al. (2022), maintenance is an effective solution to improve not only the system availability, but also the system safety, the product quality, as well as the customer satisfaction. An appropriate preventive maintenance policy is an effective way to save cost by reducing the probability of failure (Li et al. (2023)).

Unscheduled or corrective maintenance refers to maintenance actions carried out after the occurrence of component's failure. Suppose that the equiment are minimally repaired at failure and the effect of imperfect preventive maintenance is modeled according to the Proportional Age Reduction (PAR) criterion. In the case of perfect maintenance, the action restores the equipment to be as new. The PAR approach introduced by Malik (1979), assumes that each preventive maintenance action reduces the age of the equipment by a quantity proportional to the operating time elapsed form the most recent scheduled maintenance.

The failure pattern in each maintenance cycle is described by a Non-homogeneous Poisson process in which the age of the equipment in the k-th maintenance cycle is reduced by a fraction ρ of the most recent scheduld maintenance action τ_{k-1}. The hazard function (3) at a time t is

$$h(t) = h(t - \rho\tau_{k-1}), \quad \tau_{k-1} < t < \tau_k. \tag{8}$$

The following are the hypotheses assumed by Shin et al. (1996) to model minimal repairs interspersed with scheduled imperfect preventive maintenance actions:

1. Suppose that l units are observed until T_i, $i = 1, ..., l$.
2. Suppose that each equipment i is subjected to m_i scheduled maintenance actions at $\tau_{i,1} < \tau_{i,m_i} \leq T_i$.
3. The i-th equipment experiences $r_{i,k}$ failures during the k-th preventive maintenance cycle $(k = 1, ..., m_i + 1)$
4. Let $t_{i,k,j}$ be the time of the j-th failure of the i-th equipment that occurs in the k-th maintenance cycle.

Pham and Wang (1996) presents another interesting approach to imperfect maintenance via quasi-renewal process.

The maintenance model here adopted considers that the preventive action is executed periodically at a prespecified times and different policies for treat the failures may be employed.

Let

- c_p be the preventive maintenance cost;
- c_m be the corrective maintenance cost.

Adapting the maintenance policy for repairable equipments of Pham (2003), the maintenance expected cost per unit time for the period $[t_{m+1}, t_{m+2}]$ is given by

$$V_1(t_{m+1}, t_{m+2}) = \frac{c_m H(t_{m+1}, t_{m+2}) + c_p}{t_{m+2} - t_{m+1}}, \tag{9}$$

where $H(t_{m+1}, t_{m+2}) = \int_{t_{m+1}}^{t_{m+2}} h(s)ds$.

2.2. Maximum Likelihood Estimation

The maximum likelihood estimation (MLE) is a method of estimating the parameters of an assumed probability density function given some observed data by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable. The MLE of the model (8) is given by

$$\begin{aligned} L = \quad & \prod_{i=1}^{l} \left\{ \prod_{k=1}^{m_i+1} \left[\prod_{j=1}^{r_{i,k}} h(t_{i,k,j} - \rho\tau_{i,k-1}) \right] \times \right. \\ & \left. \exp\left[-\sum_{k=1}^{m_i+1} \int_{\tau_{i,k-1}}^{\tau_{i,k}} h(x - \rho\tau_{i,k-1}) dx \right] \right\}, \end{aligned} \tag{10}$$

where $\tau_{i,0} = 0$ and $\tau_{i,m_i+1} = T_i$ (see more details in Shin et al. (1996)).

The most frequently used non-homogeneous Poisson process (NHPP) is the power law process (Pham (2003)), whose intensity, applied in (8), is given by

$$h(t) = \frac{\beta}{\alpha} \left(\frac{t - \rho\tau_{k-1}}{\alpha} \right)^{\beta - 1}. \tag{11}$$

Using (11) in (10) we have

$$L = \left(\frac{\beta}{\alpha}\right)^n \left[\prod_{i=1}^{l}\prod_{k=1}^{m_i+1}\prod_{j=1}^{r_{i,k}} \left(\frac{t_{i,k,j}-\rho\tau_{i,k-1}}{\alpha}\right)^{\beta-1}\right] \times \exp\left\{-\sum_{i=1}^{l}\sum_{k=1}^{m_i+1}\left[\left(\frac{t_{i,k}-\rho\tau_{i,k-1}}{\alpha}\right)^{\beta} - \left(\frac{t_{i,k-1}-\rho\tau_{i,k-1}}{\alpha}\right)^{\beta}\right]\right\}, \tag{12}$$

where n is the total number of failures for l units during the whole observation period. The maximum likelihood estimation for α is given by the following analytical solution:

$$\alpha = \left\{\sum_{i=1}^{l}\sum_{k=1}^{m_i+1} \frac{\left[(t_{i,k}-\hat{\rho}\tau_{i,k-1})^{\hat{\beta}} - (t_{i,k-1}-\hat{\rho}\tau_{i,k-1})^{\hat{\beta}}\right]}{n}\right\}^{\frac{1}{\hat{\beta}}}. \tag{13}$$

The estimators for β e ρ are found by maximizing the modified two-parameter loglikelihood function:

$$l(\beta,\rho) = n\ln\beta - n\ln\left\{\sum_{i=1}^{l}\sum_{k=1}^{m_i+1}\left[(t_{i,k}-\rho\tau_{i,k-1})^{\beta} - (t_{i,k-1}-\rho\tau_{i,k-1})^{\beta}\right]\right\} + n\ln n + (\beta-1) \times \left[\sum_{i=1}^{l}\sum_{k=1}^{m_i+1}\sum_{j=1}^{r_{i,k}} \ln(t_{i,k,j}-\rho\tau_{i,k-1})\right] - n \tag{14}$$

In order to solve the maximization problem (14), which does not have analytical solution, it is applied a Differential Evolution method. As it is described in Kienitz and Wetterau (2012), the Differential Evolution method is a population-based search algorithm which belongs to the class of genetic algorithms. It mimics the process of Darwinian evolution using techniques such as inheritance, mutation, recombination, selection and crossover. The algorithm is designed to converge to the global optimal solution. The algorithm is described in Chapter 9 of Kienitz and Wetterau (2012).

3. Failure Data

In this paper it is considered the failure data of a central cooler system of a nuclear power plant analyzed first by Shin et al. (1996) and discussed in Pham (2003). The data consist of $n = 15$ failure times and $m = 3$ preventive maintenance epochs - highlighted by $(*)$, observed over 612 days. The data are given in Table 1.

Table 1: *Failure data*

116	151	154*	213	263*	386	387	395	407
463	492	494	501	512*	537	564	590	609

Minimal repairs are performed at failures. No information are given with respect to the level of recovery of the preventive maintenance actions.

The Differential Evolution method described in Kienitz and Wetterau (2012) was used to solve the modified two-parameter loglikelihood function (14). The parameters shown in the Table 7 of Shin et al. (1996) were exactly recovered, namely

$$\alpha = 141, \quad \beta = 2.91, \quad \rho = 0.77.$$

The parameters suggest that the equipment is restored to a level corresponding to 33% of the actual age. Also, as the shape parameter β is greater that 2, it is known that the hazard rate increases in a convex form towards infinite.

The objective is to find the scheduled maintenance interval that minimizes the overall cost given by (9). This work also aims to analyze the impact of the recovery parameter ρ for future maintenance actions.

4. Results

Let t_{m+1} be the actual time, when the last preventive maintenance action was performed with $\rho = 0.77$. In order to calculate the optimal preventive maintenance interval which minimizes the cost (9), we need to find

$$\begin{aligned}\frac{\partial V_1(t_{m+1}, t_{m+2})}{\partial t_{m+2}} &= \frac{[c_m h(t_{m+1}, t_{m+2})]\,(t_{m+2} - t_{m+1}) - [c_m H(t_{m+1}, t_{m+2}) + c_p]}{(t_{m+2} - t_{m+1})^2} = 0 \\ &= c_m(t_{m+2} - t_{m+1}) \left(\frac{\beta}{\alpha}\right) \left[\left(\frac{t_{m+2} - \rho t_{m+1}}{\alpha}\right)^{(\beta-1)}\right] - \\ &\quad \left\{c_m \left[\left(\frac{t_{m+2} - \rho t_{m+1}}{\alpha}\right)^{\beta} - \left(\frac{(1-\rho) t_{m+1}}{\alpha}\right)^{\beta}\right] + c_p\right\}. \end{aligned} \tag{15}$$

Applying the Power Law process (11), we cannot find the solution for (15) analytically. We find the root of the nonlinear function (15) using the fzero function of Matlab, which is a combination of bisection, secant, and inverse quadratic interpolation methods (Forsythe G. E. (1976)).

Suppose that $c_m = 1.25c_p$, the next maintenance epoch which minimizes the overall maintenance cost under the hypotheses that a major overhaul was performed at $t_{m+1} = 612$ days is in $t_{m+2} = 678$ days. The next optimal scheduled maintenance actions is shown in Table 2. We observe that the first optimal scheduled maintenance interval is 66 days. We can note that the preventive maintenance becomes more frequent as the equipment ages and the hazard function increases. The next intervals are the following: 64, 63, 61, 59 and 58 days. Figure 2 compares the time between scheduled maintenance actions for $c_c = 0.75c_p$, $c_c = c_p$ and $c_c = 1.25c_p$. As expected, the scheduled maintenance interval is shorter as the corrective maintenance becomes more expensive.

It can be analyzed in Figure 3 that higher level of recovery results in larger scheduled maintenance intervals. It is only true for $\beta > 2$, which is the actual case. For $\beta = 2$ the optimal preventive maintenance interval would be the same for any value of ρ. The opposite behavior is found for $1 < \beta < 2$. It can also be seen in Figure 3 that as the preventive maintenance cost becomes greater than the corrective maintenance costs, that is, $\frac{CM}{CP} < 1$, the scheduled maintenance interval increases fastly.

Table 2: *Tempos de parada programada*

t_{m+3}	742
t_{m+4}	805
t_{m+5}	866
t_{m+6}	925
t_{m+7}	983

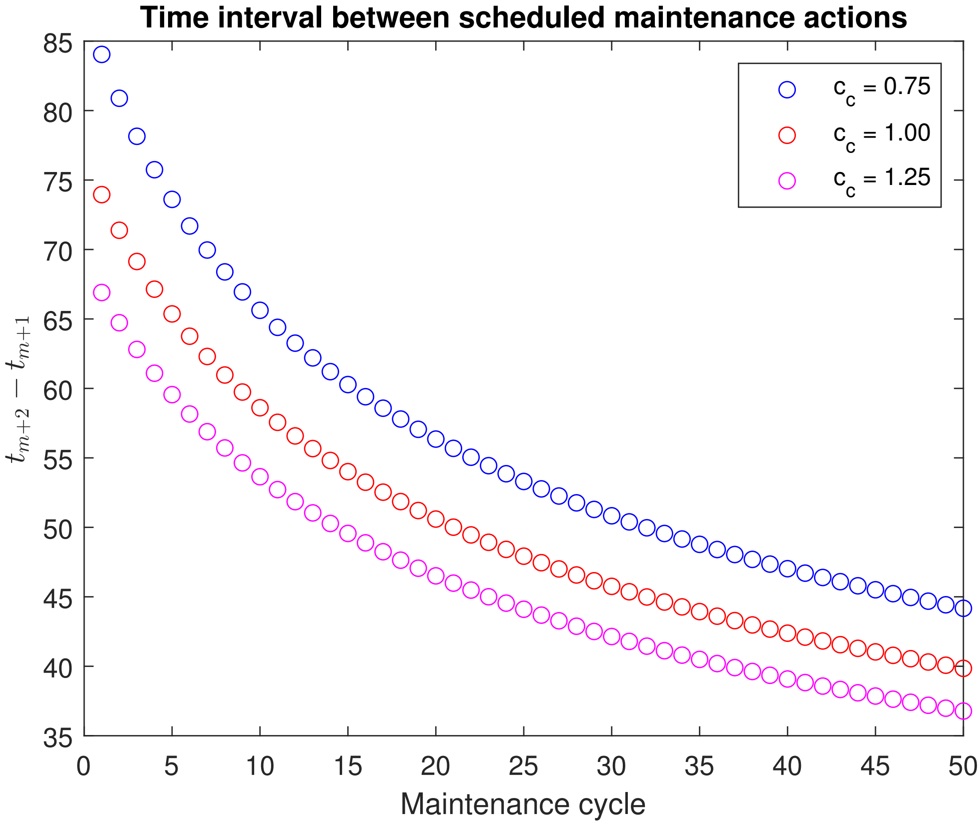


Figure 2: *Time period of the next maintenance actions*

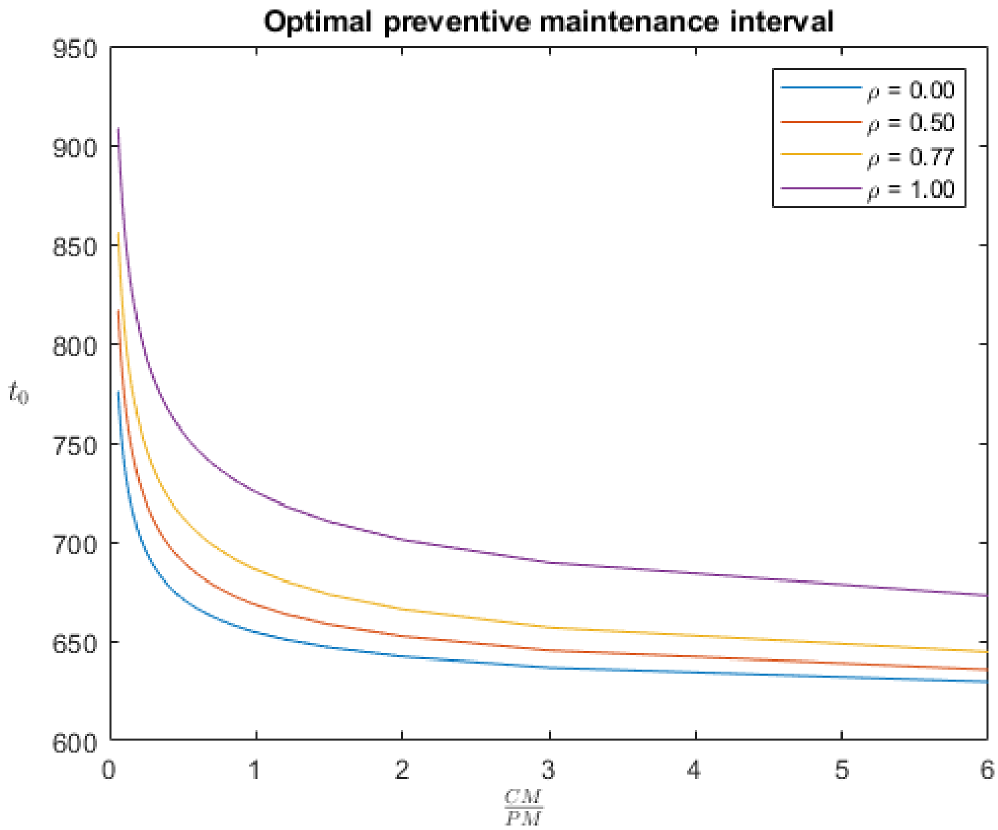


Figure 3: *Optimal preventive maintenance next interval as a function of the corrective repair cost*

In Figure 4 it is exhibited the hazard rate paths. In the left panel, it is shown the in-sample path. We can note an unorganized maintenance epochs in the original data which results in a high value for the hazard function before the third action. In the right panel it is shown the out of sample predicted hazard function. It is considered that the scheduled maintenance actions are performed at the optimal interval, that is, at each 58 days. It is noteworthy that, respecting the historical recovery factor of $\rho = 0.77$, in the 11-th action the hazard function finds in a lower level than it was observed in the third maintenance interval.

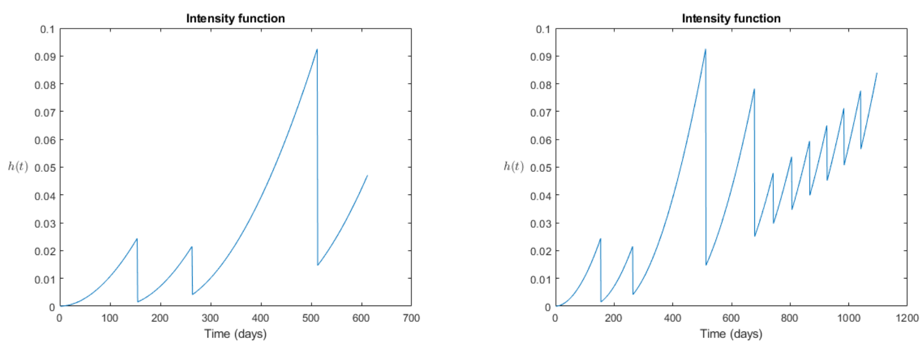


Figure 4: *In-sample and out of sample intensity function*

In Figure 5 we fix $c_m = 1$ and analyze the cost of the next maintenance cycle. We note that cost varies widely according to the recovery parameter ρ.

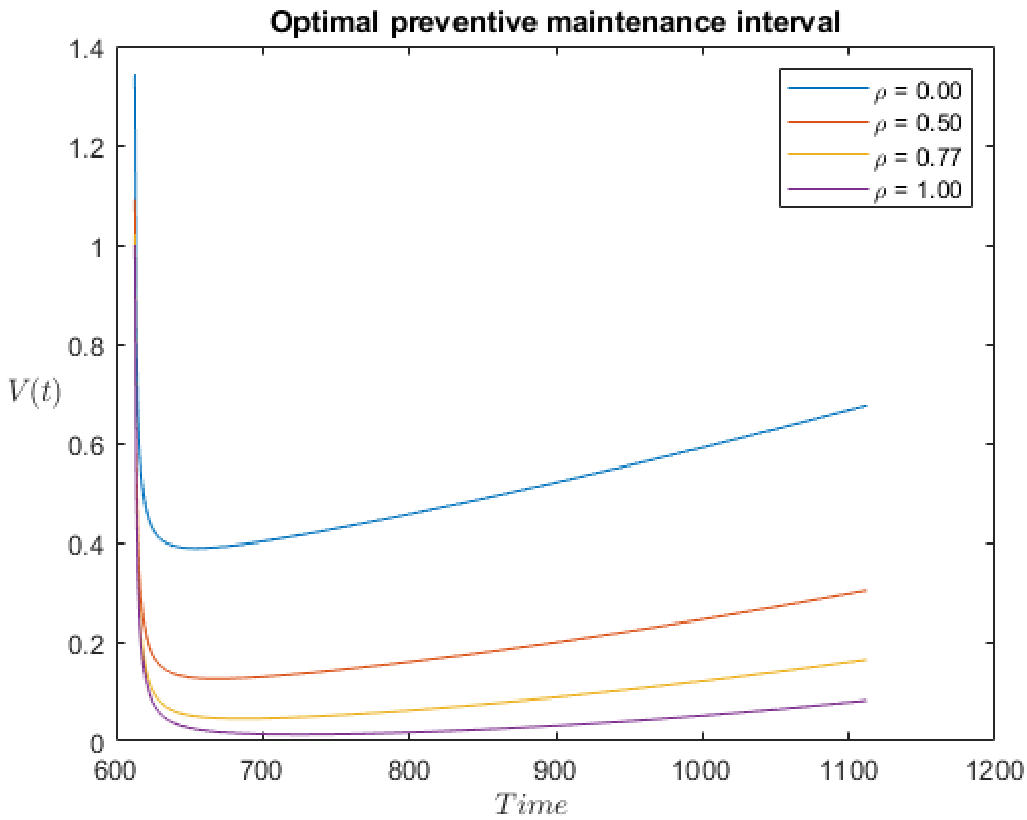


Figure 5: *Maintenance cost*

In Figure 6 it is exhibited the coupled in-sample and out of sample hazard function for different recovery parameters ρ after the observed period of 612 days. With $t_{m+2} = 678$, and t_{m+n} according to Table 2, we can compare the hazard function for $\rho = 0$ (minimal repairs), $\rho = 0.5$ (intermediary repairs), $\rho = 0.77$ (actual policy) and $\rho = 1$ (perfect repairs).

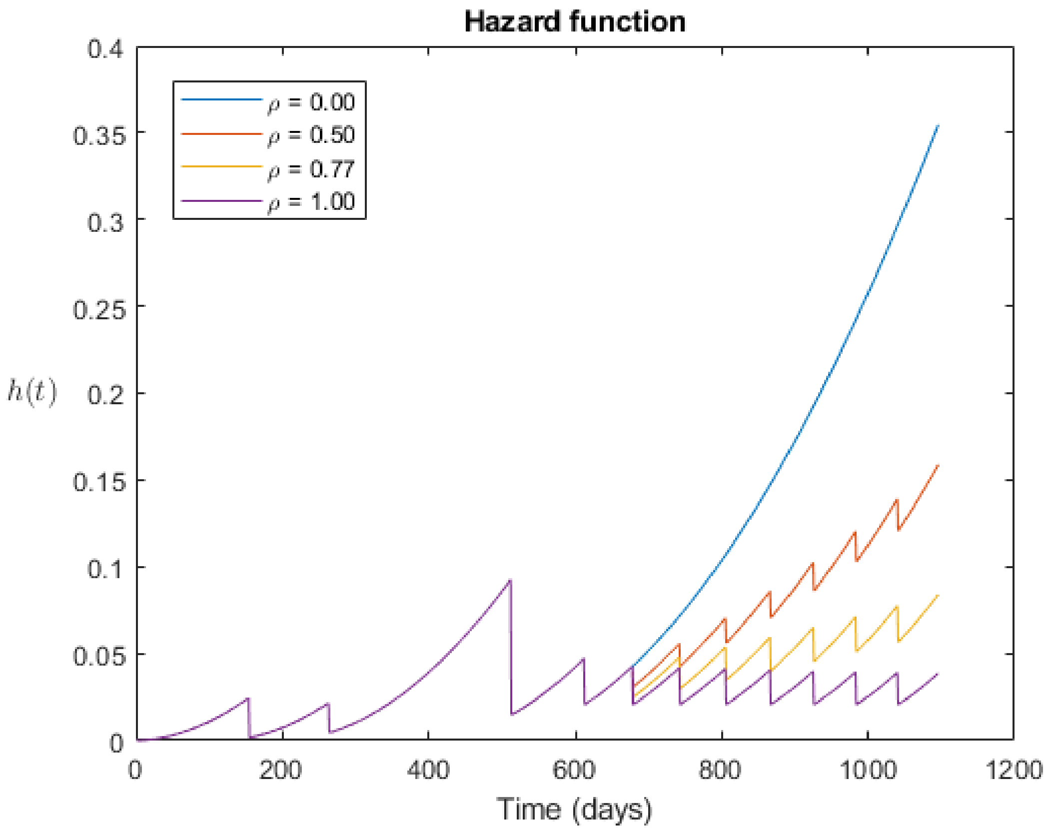


Figure 6: *Hazard function*

Finally,

5. Conclusion

A maintenance model under the assumption of imperfect preventive maintenance interspaced with minimal repairs was considered. This study analyzed a real failure database available in the literature consisting of 15 failure times and 3 unequally-spaced preventive maintenance actions. The model parameters of the power law process were estimated using the MLE and a Differential Evolution algorithm. The results obtained in this paper showed that maintenance cost varies widely according to the recovery parameter. It is also clear from the results here obtained that the optimal interval of two consecutive preventive maintenance actions strongly depends on the costs. Hence, a proper estimation of repairs expenditures are needed. The proposed cost model can be extended to consider random unscheduled repair costs.

References

Azizi, F. and Salari, N. (2023). A novel condition-based maintenance framework for parallel manufacturing systems based on bivariate birth-birth-death processes. *Reliability Engineering & System Safety*, 229:108798.

Chen, Y., Qiu, Q., and Zhao, X. (2022). Condition-based opportunistic maintenance policies with two-phase inspections for continuous-state systems. *Reliability Engineering & System Safety*, 228:108767.

Dui, H., Wei, X., Xing, L., and Chen, L. (2023). Performance-based maintenance analysis and resource allocation in irrigation networks. *Reliability Engineering & System Safety*, 230:108910.

Elsayed, E. A. (2021). *Reliability engineering*. WILEY SERIES IN SYSTEMS ENGINEERING AND MANAGEMENT. Wiley, 3rd edition.

Forsythe G. E., M. A. Malcolm, C. B. M. (1976). *Computer Methods for Mathematical Computations*. Prentice-Hall.

Guo, C. and Liang, Z. (2022). A predictive markov decision process for optimizing inspection and maintenance strategies of partially observable multi-state systems. *Reliability Engineering & System Safety*, 226:108683.

Huynh, K., Vu, H., Nguyen, T., and Ho, A. (2022). A predictive maintenance model for k-out-of-n:f continuously deteriorating systems subject to stochastic and economic dependencies. *Reliability Engineering & System Safety*, 226:108671.

Kienitz, J. and Wetterau, D. (2012). *Financial Modelling: Theory, Implementation and Practice with MATLAB Source*. The Wiley Finance Series. Wiley.

Li, Y., Xia, T., Chen, Z., and Pan, E. (2023). Multiple degradation-driven preventive maintenance policy for serial-parallel multi-station manufacturing systems. *Reliability Engineering & System Safety*, 230:108905.

Löfsten, H. (2000). Measuring maintenance performance - in search for a maintenance productivity index. *International Journal of Production Economics*, 63(1):47–58.

Malik, M. A. K. (1979). Reliable preventive maintenance scheduling. *A I I E Transactions*, 11(3):221–228.

Naveen Kumar, S.C. Malik, N. N. (2022). Stochastic analysis of a repairable system of non-identical units with priority and conditional failure of repairman. *Reliability Theory & Applications*, 17(1):123–133.

Neetu Dabas, R. R. (2022). Parallel system analysis with priority and inspection using semi-markov approach. *Reliability Theory & Applications*, 17(2):56–66.

Nguyen, V.-T., Do, P., Vosin, A., and Iung, B. (2022). Artificial-intelligence-based maintenance decision-making and optimization for multi-state component systems. *Reliability Engineering & System Safety*, 228:108757.

Nse Udoh, Iniobong Uko, A. U. (2022). Optimal economic age replacement models for non-repairable systems with sudden but non-constant failure rate. *Reliability Theory & Applications*, 17(3):108–120.

Pham, H. (2003). *Handbook of Reliability Engineering*. Springer London, 1st edition.

Pham, H. and Wang, H. (1996). Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438.

S., D. G. and Sebastian, N. (2022). A new life time distribution: Burr iii modified weibull distribution and its application in burn in process. *Reliability Theory & Applications*, 17(1):76–86.

Shanti Parkash, P. (2022). Performance modeling and dss for assembly line system of leaf spring manufacturing plant. *Reliability Theory & Applications*, 17(2):403–412.

Shin, I., Lim, T., and Lie, C. (1996). Estimating parameters of intensity function and maintenance effect for repairable unit. *Reliability Engineering & System Safety*, 54(1):1–10.

Tijjani A. Waziri, Bashir M. Yakasai, R. S. A. (2022). Analysis of some proposed replacement policies. *Reliability Theory & Applications*, 17(1):87–103.

Wei, S., Nourelfath, M., and Nahas, N. (2023). Analysis of a production line subject to degradation and preventive maintenance. *Reliability Engineering & System Safety*, 230:108906.

William Q. Meeker, Luis A. Escobar, F. G. P. (2021). *Statistical Methods for Reliability Data*. Wiley Series in Probability and Statistics. Wiley, 2nd edition.

RT&A, No 1 (72)
Volume 18, March 2023

A Generalization of Lindley Distribution: Characterizations, Methods of Estimation and Applications

A. Mohammed Shabeer[1], Bindu Krishnan[2], K. Jayakumar[3]

•

[1,3] Department of Statistics, University of Calicut, India
[2] Department of Data Science, CS & IT, Jain(Deemed-to-be-University), India
md.shabeer.edv@gmail.com, bindusruthy@gmail.com, jkumar19@rediffmail.com

Abstract

Lindley distribution is a lifetime model with application in survival analysis and reliability theory problems often centred around its increasing hazard rate function and flexibility over exponential distribution. In this paper, we introduce a new generalization of the Lindley distribution referred to as Lindley Truncated Negative binomial (LTNB) distribution. The LTNB model has increasing, decreasing and upside-down bathtub(UBT) shapes for the hazard rate function. Various properties of the LTNB distribution are studied including moments, quantiles, and stochastic ordering. Characterizations of the new distribution are obtained. Maximum likelihood, Cramer-von-Mises, ordinary and weighted least squares methods of estimation are utilized to obtain the estimators of the model parameters. A simulation study is carried out to assess and compare the performance of different estimates. An autoregressive time series model with the LTNB as marginal is developed. The model is fitted to bladder cancer data set to show how the proposed model works in practice.

Keywords: Autoregressive Models, Characterizations, Lindley distribution, Maximum Likelihood, Stochastic Ordering.

1. Introduction

As a counter-example to fiducial statistics, the Lindley distribution, first put forth by [16], is one of the crucial lifetime distributions in the framework of Bayesian statistics. The Lindley distribution has an increasing hazard rate function; however, in real-life situations, the models exhibit non-monotone hazard rate shapes. For example, in the case of a serious illness condition such as cancer, the hazard rate increases and then decreases. Similarly, in the engineering field, the quality of production by untrained workers follows a similar pattern. As a result, many researchers have developed models with non-monotone hazard rates over time. Models with bathtub shapes can be found in [22]. Most often, upside-down bathtub (UBT) shape hazard rate models are used to model medical data, such as patient data for bladder cancer and lung cancer (see, [5] and [15]). In this paper, we introduce a generalized Lindley distribution with increasing, decreasing and UBT shapes for the hazard rate function. A positive random variable X is said to have Lindley distribution with the parameter α, denoted as $L(\alpha)$, if it has the probability density function (pdf)

$$f(x) = \frac{\alpha^2}{\alpha+1}(1+x)e^{-\alpha x}; x > 0, \alpha > 0, \tag{1}$$

The cumulative distribution function (cdf) is given by

$$F(x) = 1 - \frac{\alpha+1+\alpha x}{\alpha+1}e^{-\alpha x}; x \geq 0. \tag{2}$$

Ghitany[8] studied the statistical characteristics of the Lindley distribution and demonstrated its flexibility in modelling lifetime data over exponential distribution. Generalized Lindley distribution was later developed by [20], who also studied its mathematical characteristics. Weighted Lindley distribution was introduced by [9]. An extended Lindley distribution was proposed and investigated by [3]. In 2013, [6] introduced a generalized Lindley distribution. Later Beta Lindley distribution was examined by [18]. The Lindley-Exponential distribution was first proposed by [4]. Inverse Lindley distribution was extended by [23]. The Odd Lindley Burr XII distribution was proposed by [14]. The generalized two-parameter Lindley distribution was introduced by [7]. Unit-Lindley distribution was studied in [2].

The rest of the paper is as follows: The Lindley Truncated Negative binomial (LTNB) distribution is introduced as a compounded distribution in Section 2. Section 3 discuss the statistical properties of the LTNB distribution such as moments, quantile function, skewness, and kurtosis. Characterizations of the LTNB distribution are obtained in section 4.In Section 5, we investigate the stochastic ordering property of LTNB. In section 6, we use the maximum likelihood method, least squares method, weighted least squares, and Cramer-von-Mises-estimator for estimating the parameters of the new distribution. In section 7 we carry out a simulation study to evaluate the performance of Maximum likelihood estimation and other estimation methods. The LTNB distribution is used to model remission times of bladder cancer patients in Section 8. It is demonstrated that the LTNB distribution fits this data set better than the other well-known competitors. In section 9, we develop a first-order autoregressive minification process with the LTNB distribution as the marginal distribution.

2. Lindley Truncated Negative binomial distribution

We propose a new generalization for Lindley distribution with relatively simple expressions for the survival function, hazard rate function, and quantile's. The hazard rate function is more flexible than the previous generalizations and the proposed model has a closed-form for distribution function. In addition, the new distribution fits some real-world data sets better than existing Lindley distribution generalizations.

In 1997, [17] developed a method of adding a tilt parameter to a distribution to extend existing distributions. Many distributions were extended and published in the literature as a result of this technique. As a generalization of the above technique, [21] proposed a novel family of distributions through truncated negative binomial distribution. It's important to remember that these distributions arise as the distribution of random minimum or random maximum. The cdf is given by

$$G(z) = \frac{\gamma^n}{1-\gamma^n}[(\bar{F}(z)+\gamma F(z))^{-n}-1]; z \in \mathbf{R}, \gamma, \theta > 0, \tag{3}$$

where $F(.)$ is the cdf of baseline distribution . If $F(x)$ follows Lindly with cdf (2) we have Lindley Truncated Negative binomial(LTNB) distribution, LTNB(α, γ, θ), having pdf

$$g(z;\alpha,\gamma,\theta) = \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta\gamma^\theta(1+z)e^{-\alpha z}}{(1-\gamma^\theta)\left[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}\right]^{\theta+1}}; z>0 \quad \alpha,\gamma,\theta>0. \tag{4}$$

The cumulative distribution function of LTNB(α, γ, θ) is given by

$$G(z;\alpha,\gamma,\theta) = \frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta}\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}}\right]^\theta \tag{5}$$

and hence the survival function is

$$\bar{G}(z;\alpha,\gamma,\theta) = \frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}}\right]^\theta - 1\right\} \tag{6}$$

Remark: The LTNB(α, γ, θ) distribution reduces to the Lindley distribution when $\theta = 1$ and $\gamma \longrightarrow 1$.

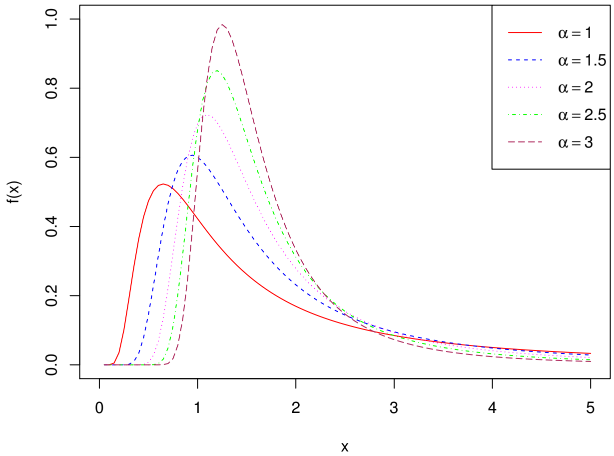


Figure 1: *LTNB*(α, γ, θ) *pdf when* $\gamma = 0.5, \theta = 0.5$.

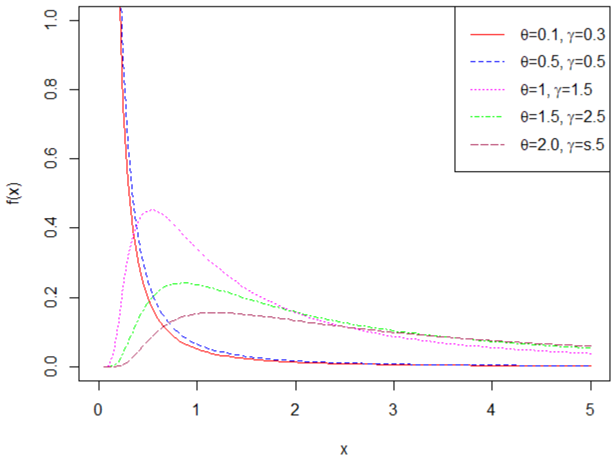


Figure 2: *LTNB*(α, γ, θ) *pdf when* $\alpha = 1$.

The hazard rate function of LTNB distribution is

$$h(z;\alpha,\gamma,\theta) = \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta(1+z)e^{-\alpha z}}{[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]\left[(\alpha+1)^\theta - [(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]^\theta\right]} \quad (7)$$

and the reverse hazard rate function is

$$r(z;\alpha,\gamma,\theta) = \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta(1+z)e^{-\alpha z}}{[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]\left[[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha z)e^{-\alpha z}]^\theta - \gamma^\theta(\alpha+1)^\theta\right]}. \quad (8)$$

Figure 3 depicts hazard rate function graphs for various parameter values. The graphs showing decreasing, non-decreasing, and UBT shapes for hazard rate function.

3. Statistical Properties

We discuss about moments, simulation, quantiles, skewness, and kurtosis in this section.

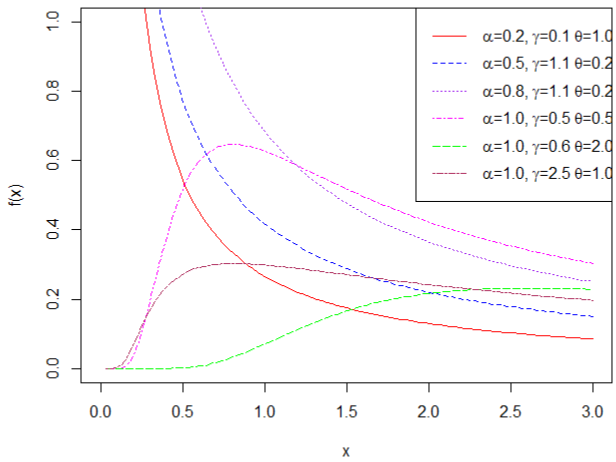


Figure 3: *LTNB(α, γ, θ) hazard rate function for various parameter values.*

3.1. Moments

Let $X \sim LTNB(\alpha, \gamma, \theta)$, then its r^{th} moment with respect to the origin is given by,

$$E[X^r] = \int_0^\infty x^r \frac{\alpha^2(\alpha+1)^\theta(1-\gamma)\theta\gamma^\theta(1+x)e^{-\alpha x}}{(1-\gamma^\theta)\left[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}\right]^{\theta+1}}dx$$
$$= \sum_{k=0}^\infty \frac{\alpha^2(1-\gamma)^{k+1}\theta\gamma^\theta}{(1-\gamma^\theta)(\alpha+1)^{k+1}}\binom{k+\theta}{\theta}\int_0^\infty x^r(1+x)e^{-\alpha x(1+k)}(\alpha+1+\alpha x)^k dx.$$

3.2. Simulation, Quantiles and Median

In order to generate random numbers from a LTNB(α, γ, θ) distribution, we use

$$X = \frac{-1}{\alpha} - 1 - \frac{1}{\alpha}W_{-1}\left(-e^{-(\alpha+1)}\frac{(\alpha+1)}{1-\gamma}(1-\gamma((1-\gamma^\theta)Y+\gamma^\theta)^{-1/\theta})\right) \tag{9}$$

where $W_{-1}(.)$ is the negative Limbert W function and $Y \sim$ U(0,1).
The qth quantile of the LTNB(α, γ, θ) is

$$X = \frac{-1}{\alpha} - 1 - \frac{1}{\alpha}W_{-1}\left(-e^{-(\alpha+1)}\frac{(\alpha+1)}{1-\gamma}(1-\gamma((1-\gamma^\theta)q+\gamma^\theta)^{-1/\theta})\right), \tag{10}$$

and thus the median of LTNB(α, γ, θ) is,

$$X = \frac{-1}{\alpha} - 1 - \frac{1}{\alpha}W_{-1}\left(-e^{-(\alpha+1)}\frac{(\alpha+1)}{1-\gamma}(1-\gamma((1-\gamma^\theta)\frac{1}{2}+\gamma^\theta)^{-1/\theta})\right).$$

3.3. Skewness and Kurtosis

According to [13], the distribution skewness can be calculated using the below equation

$$S = \frac{Q\left(\frac{3}{4}\right) - 2Q(\frac{1}{2}) + Q(\frac{1}{4})}{Q(\frac{3}{4}) - Q(\frac{1}{4})}$$

and according to [19], the kurtosis of the LTNB distribution is as follows

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

where $Q(.)$ is the quantile function of X as defined by (10). These metrics are less susceptible to outliers.

4. Characterizations of LTNB distribution

Many characterization results, established in various ways, can be found in the literature. In this section, the ratio of truncated moments and the hazard rate function are used to characterize the LTNB distribution.

4.1. Characterizations based on two truncated moments

This section discusses how to characterise the LTNB distribution using the ratio of two truncated moments. In our first characterization, we apply the Theorem 1 below established in [10].

Theorem 1. Let $(\Omega, \mathcal{F}, P)$ be a given probability space and let $H = [d, e]$ be an interval for some $d < e \quad (d = -\infty, e = \infty$might as well be allowed). Let $X : \Omega \rightarrow H$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real functions on H such that

$$E[q_2(X)|X \geq x] = E[q_1(X)|X \geq x]\zeta(x), x \in H,$$

is defined with some real function ζ. Assume that $q_1, q_2 \in C^1(H), \zeta \in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $\zeta q_1 = q_2$ has no real solution in the interior of H. Then F is uniquely determined by the functions q_1, q_2 and ζ, particularly

$$F(x) = \int_a^x C|\frac{\zeta'(u)}{\zeta(u)q_1(u) - q_2(u)}|exp(-s(u))du,$$

where the function s is a solution of the differential equation $s' = \frac{\zeta' q_1}{\zeta q_1 - q_2}$ and C is the normalization constant, such that $\int_H dF = 1$.

Theorem 2. Let X be a positive real valued continuous random variable with pdf $f(x)$. Denote $q_1(x) = (1-(1-\gamma)\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x})^{\theta+1}$ and $q_2(x) = q_1(x)\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}$ for $x > 0, \alpha > 0, \theta > 0, \gamma > 0$. Then, the random variable X has pdf (4) if and only if the function $\zeta(.)$ defined in Theorem 1 is of the form

$$\zeta(x) = \frac{1}{2}\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}$$

Proof. Suppose the random variable X has the pdf (4). Then

$$(1-F(x))E\left[q_1(X)|X \geq x\right] = C\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}, \quad x > 0$$

and

$$(1-F(x))E\left[q_2(X)|X \geq x\right] = \frac{C}{2}\left(\frac{1+\alpha+\alpha x}{\alpha+1}\right)^2 e^{-2\alpha x}, \quad x > 0$$

where $C = \left(\frac{\gamma^\theta \theta(1-\gamma)}{1-\gamma^\theta}\right)$.Further,

$$\zeta(x)q_1(x) - q_2(x) = -\frac{1}{2}q_1(x)\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x} < 0, \quad x > 0.$$

Conversely, if ζ is of the above form, then

$$s'(x) = \frac{\zeta'(x)q_1(x)}{\zeta(x)q_1(x) - q_2(x)} = \frac{\alpha^2(1+x)}{1+\alpha+\alpha x}, \quad x > 0.$$

Thus

$$s(x) = -log(\frac{1+\alpha+\alpha x}{\alpha+1}e^{-\alpha x}), \quad x > 0.$$

Now from Theorem 1, result follows.

■

4.2. Characterization based on the hazard rate function

For $\theta = 1$, we characterise the LTNB distribution using the hazard rate function.

Theorem 3. Let X be a positive real valued continuous random variable with hazard rate function $h(x)$. The random variable X has LTNB distribution, for $\theta = 1$, if and only if $h(x)$ satisfies the differential equation

$$h^{'}(x) - \frac{h(x)}{(1+x)(1+\alpha+\alpha x)} = \frac{\alpha^4(1-\gamma)(1+x)^2 e^{-\alpha x}}{(1+\alpha+\alpha x)(1-\frac{(1-\gamma)(1+\alpha+\alpha x)e^{-\alpha x}}{1+\alpha})^2}, \quad x > 0.$$

Proof. If X has pdf (4) , the above differential equation clearly holds.
Now suppose the differential equation holds then,

$$\frac{d}{dx}\left\{\frac{h(x)(1+\alpha+\alpha x)}{1+x}\right\} = \alpha^2 \frac{d}{dx}\left\{\frac{1}{1-(\frac{(1-\gamma)(1+\alpha+\alpha x)e^{-\alpha x}}{1+\alpha})}\right\}$$

which implies,

$$h(x) = \frac{\alpha^2(1+x)}{(1+\alpha+\alpha x)} \frac{1}{1-(\frac{(1-\gamma)(1+\alpha+\alpha x)e^{-\alpha x}}{1+\alpha}}$$

which is the hazard rate function of LTNB distribution.

■

5. Stochastic Ordering

Let X and Y be two random variables with distribution functions of F_1 and F_2, respectively, with corresponding pdfs of f_1 and f_2. Then it is stated that X is less than Y in ,
i) stochastic order (denoted as $X \leq_{st} Y$) if $F_1(x) \geq F_2(x)$ for all x;
ii) hazard rate order (denoted as $X \leq_{hr} Y$) if $(1 - F_1(x)/(1 - F_2(x))$ is decreasing in $x \geq 0$;
ii) likelihood ratio order (denoted as $X \leq_{lr} Y$) if $f_1(x)/f_2(x)$ is decreasing in $x \geq 0$;
iv) reverse hazard rate order (denoted as $X \leq_{rhr} Y$) if $F_1(x)/F_2(x)$ is decreasing in $x \geq 0$.
The four orderings have a relationship with one another; for further information, see [24],

$$X \leq_{lr} Y \implies X \leq_{hr} Y \implies X \leq_{rhr} Y \implies X \leq_{st} Y. \tag{11}$$

For $\gamma_2 > \gamma_1$,
let $X \sim LTNB(\alpha, \gamma_1, \theta)$ and $Y \sim LTNB(\alpha, \gamma_2, \theta)$. Then

$$\frac{f_X(y)}{f_Y(y)} = \frac{\gamma_1^\theta(1-\gamma_2^\theta)(1-\gamma_1)\left[(\alpha+1)-(1-\gamma_2)(\alpha+1+\alpha z)e^{-\alpha y}\right]^{\theta+1}}{\gamma_2^\theta(1-\gamma_1^\theta)(1-\gamma_2)\left[(\alpha+1)-(1-\gamma_1)(\alpha+1+\alpha y)e^{-\alpha y}\right]^{\theta+1}}$$

Since $\gamma_2 > \gamma_1$,

$$\frac{d}{dx}\left[\frac{f_X(y)}{f_Y(y)}\right] = \frac{\gamma_1^\theta(1-\gamma_2^\theta)(1-\gamma_1)\alpha^2 e^{-\alpha y}(1+y)(\alpha+1)(\gamma_1-\gamma_2)}{(1-\gamma_2)\gamma_2^\theta(1-\gamma_1^\theta)}$$
$$\times \frac{\left[(\alpha+1)-(1-\gamma_2)(\alpha+1+\alpha y)e^{-\alpha y}\right]^\theta}{\left[(\alpha+1)-(1-\gamma_1)(\alpha+1+\alpha y)e^{-\alpha y}\right]^{\theta+2}},$$
$$< 0.$$

$\implies f_X(y)/f_Y(y)$ is decreasing in y.
That is $X \leq_{lr} Y$. From (11), the remaining ordering follows.

6. Different methods of estimation

In this section we use maximum likelihood estimation, least squares, weighted least squares and cramer-von-Mises estimation methods to estimate the parameters of the LTNB distribution.

6.1. Maximum likelihood Estimation

Let $x_1, x_2, ..., x_n$ be an observed random sample from LTNB distribution with unknown parameter vector $\nu = (\alpha, \gamma, \theta)^T$. Then the log-likelihood function is

$$\begin{aligned} log\ \ell &= 2nlog\alpha + nlog\theta + nlog(1-\gamma) + n\theta log\gamma + n\theta log(\alpha+1) - nlog(1-\gamma^\theta) - \alpha\sum_{i=1}^{n} x_i \\ &+ \sum_{i=1}^{n} log(1+x_i) - (\theta+1)\sum_{i=1}^{n} log\left[(\alpha+1) - (1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}\right] \end{aligned} \tag{12}$$

The partial derivatives are given by

$$\begin{aligned} \frac{\partial log\ \ell}{\partial \alpha} &= \frac{2n}{\alpha} + \frac{n\theta}{\alpha+1} - \sum_{i=1}^{n} x_i \\ &-(\theta+1)\sum_{i=1}^{n}\left[\frac{1-(1-\gamma)e^{-\alpha x_i}(1-\alpha x_i - \alpha x_i^2)}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}}\right], \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{\partial log\ \ell}{\partial \gamma} &= \frac{-n}{1-\gamma} + \frac{n\theta}{\gamma} + \frac{n\theta\gamma^{\theta-1}}{1-\gamma^\theta} \\ &-(\theta+1)\sum_{i=1}^{n}\left[\frac{(\alpha+1+\alpha x_i)e^{-\alpha x_i}}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}}\right], \end{aligned} \tag{14}$$

$$\begin{aligned} \frac{\partial log\ \ell}{\partial \theta} &= nlog\gamma + \frac{n}{\theta} + nlog(\alpha+1) + \frac{n\gamma^\theta log\gamma}{1-\gamma^\theta} \\ &- \sum_{i=1}^{n} log\left[(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x_i)e^{-\alpha x_i}\right] \end{aligned} \tag{15}$$

Set the score vector to zero,

$$U(\nu) = \left(\frac{\partial log\ell}{\partial\alpha}, \frac{\partial log\ell}{\partial\gamma}, \frac{\partial log\ell}{\partial\theta}\right)^T.$$

$U(\nu) = 0$, and solve them simultaneously to obtain the ML estimators $\hat{\alpha}$, $\hat{\gamma}$, and $\hat{\theta}$. These equations can be numerically solved using statistical software using iterative techniques like the Newton-Raphson algorithm as they cannot be solved analytically. All of the second order derivatives exist for the three-parameter LTNB distribution as well.

6.2. Least squares and weighted least squares estimators

Let $t_1 < t_2 < t_3 < ... < t_n$ be the n ordered random sample from any distribution with cdf $F(t)$. Then we have,

$$E[F(t_i)] = \frac{i}{n+1}.$$

The least squares method minimizes

$$P_{LSE}(\alpha,\gamma,\theta) = \sum_{i=1}^{n}\left(F(t_i) - \frac{i}{n+1}\right)^2 \tag{16}$$

with respect to the unknown parameters. Here the least squares estimates are obtained by minimizing the following equation with respect to α, γ, θ.

$$P_{LSE}(\alpha,\gamma,\theta) = \sum_{i=1}^{n}\left(\frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta}\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - \frac{i}{n+1}\right)^2. \tag{17}$$

Weighted least squares estimates of α, γ, θ are obtained by minimizing the following equation with respect to α, γ, θ.

$$P_{WLSE}(\alpha,\gamma,\theta) = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left(\frac{1}{1-\gamma^\theta} - \frac{\gamma^\theta}{1-\gamma^\theta} \left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}} \right]^\theta - \frac{i}{n+1} \right)^2. \tag{18}$$

6.3. Cramer-von-Mises-estimator(CME)

CME is obtained by minimizing the following equation with respect to α, γ, θ. Here $F(.)$ is the distribution function of LTNB distribution given by (5) .

$$P_{CME}(a,b,\gamma,\theta) = \frac{1}{12n} + \sum_{i=1}^{n} \left(F(t_i) - \frac{2i-1}{2n} \right)^2. \tag{19}$$

7. Simulation

Monte Carlo simulation is used to evaluate the performance of the maximum likelihood estimation procedure for estimating the LTNB parameters. From the LTNB model we generates samples of sizes $n = 100, 200, 300, 400$, and 500 for various combinations of α, γ, and θ. We ran the simulation 1000 times to determine the MLE's and MSE's of the parameter estimates. Table 1 displays the results. We can observe that the ML estimates are consistent. To investigate the efficiency of least squares estimators, we took samples from the LTNB distribution with $n = 100, 200, 300, 400$, and 500. We repeated the simulation 1000 times and computed the estimates and corresponding MSE's for the same set of parameter values using three methods. Table 2 displays the results. Table 2 shows that least squares and CME estimators perform similarly and both methods are giving smaller MSE's compared to weighted least squares estimates.

8. Data Analysis

Table 3 contains data on the remission times (in months) of a random sample of 128 bladder cancer patients provided by [15]. Table 4 displays the descriptive statistics for the data.

We will now look at the Total Time on Test (TTT) plot, a graphical method for determining the shape of the data's hazard rate function. The empirical TTT plot is defined as,

$$G(r/n) = \left(\sum_{i=1}^{r} x_{(1)} + (n-r)x_{(r)} \right) / \sum_{i=1}^{n} x_{(i)}, \quad r = 1, 2, .., n$$

where $x_{(i)}$ denote the i^{th} order statistic of the sample. The shape of the hazard rate function would be increasing, decreasing, bathtub-shaped, and UBT if the TTT transform is concave, convex, convex then concave, and concave then concave (see [1]). Figure 3 depicts the data's TTT plot. The hazard rate function is clearly UBT.

The Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion(HQIC) are s the goodness-of-fit metrics that we take into consideration,

$$\begin{aligned} AIC &= 2k - 2log\ \hat{\ell}, \\ AICC &= \frac{2kn}{n-k-1} - 2log\ \hat{\ell}, \\ BIC &= klog\ n - 2log\ \hat{\ell}, \\ HQIC &= 2klog(log\ n) - 2log\ \hat{\ell}, \end{aligned} \tag{20}$$

Table 1: *MLE's and their corresponding MSE's of LTNB distribution parameters.*

α	γ	θ	n	$\hat{\alpha}(MSE(\hat{\alpha}))$	$\hat{\gamma}(MSE(\hat{\gamma}))$	$\hat{\theta}(MSE(\hat{\theta}))$
.8	.2	1.5	100	.9196(.1598)	.2055(.0266)	1.5251(.7527)
			200	.8551(.0888)	.1989(.0158)	1.5342(.7232)
			300	.8813(.1184)	.2001(.0193)	1.5428(.6812)
			400	.8347(.0506)	.2015(.0091)	1.5034(.6691)
			500	.8171(.0382)	.1986(.0064)	1.5098(.5061)
.5	.3	2	100	.5698(.0539)	.2824(.0449)	1.9985(1.2211)
			200	.5298(.0313)	.2726(.0309)	1.9952(.9892)
			300	.5202(.0230)	.2836(.0245)	2.0922(.7610)
			400	.5122(.0192)	.2856(.0222)	2.0962(.5948)
			500	.5087(.0142)	.2942(.0164)	2.0962(.5484)
.3	.2	1.5	100	.3403(.0229)	.2075(.0309)	1.4718(.6438)
			200	.3179(.0111)	.2088(.0201)	1.6118(.4881)
			300	.3141(.0087)	.2047(.0149)	1.5339(.2959)
			400	.3109(.0065)	.2026(.0112)	1.4966(.2363)
			500	.3078(.0049)	.2073(.0092)	1.5387(.1401)

where $\hat{\ell}$ is the likelihood function evaluated at the maximum likelihood estimates, the sample size is n, and k is the number of parameters.
We fitted the LTNB distribution to the data and compared it to the Lindley distribution with pdf (1) and the New Generalized Lindley distribution (NGLD) with pdf

$$f(x) = \frac{e^{-\theta x}}{1+\theta}\left(\frac{\theta^{\alpha+1}x^{\alpha-1}}{\Gamma(\alpha)} + \frac{\theta^{\beta}x^{\beta-1}}{\Gamma(\beta)}\right),$$

Beta Lindely(BL) distribution having pdf,

$$f(x) = \frac{\theta^2(\theta+1+\theta x)^{\beta-1}(1+x)e^{-\theta\beta x}}{B(\alpha,\beta)(\theta+1)^{\beta}}\left(1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right)^{\alpha-1},$$

new generalized two-parameter Lindley distribution (NG2PLD) distribution having pdf,

$$f(x) = \frac{\theta^2}{\theta+1}\left(1+\frac{\theta^{\alpha-2}x^{\alpha-1}}{\Gamma(\alpha)}\right)e^{-\theta x}$$

transmuted Lindley distribution(TLD) having pdf,

$$f(x) = \frac{\theta^2(1+x)e^{-\theta x}}{(\theta+1)}\left(1-\lambda+2\lambda\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right),$$

and exponential distribution(ED) having pdf,

$$f(x) = \alpha e^{-\alpha x}$$

Table 2: *LSE, WLSE, CVME and corresponding MSEs of parameters of LTNB distribution.*

α γ θ	n	$\hat{\alpha}(MSE(\hat{\alpha}))$	$\hat{\gamma}(MSE(\hat{\gamma}))$	$\hat{\theta}(MSE(\hat{\theta}))$
Least squares	100	.6262(.0015)	.2178(.0007)	1.5386(.0043)
.6 .2 1.5	200	.6206(.0009)	.2151(.0005)	1.5408(.0045)
	300	.6173(.0005)	.2141(.0004)	1.5408(.0038)
	400	.6162(.0005)	.2136(.0003)	1.5437(.0039)
	500	.6150(.0004)	.2122(.0002)	1.5441(.0039)
Weighted Least squares	100	.6194(.1305)	.1308(.0288)	1.3082(2.3691))
.6 .2 1.5	200	.6062(.0893)	.1474(.0186)	1.3082(1.7783)
	300	.5902(.0641)	.1595(.0134)	1.3271(.7774)
	400	.5870(.0514)	.1660(.0102)	1.3618(.6277)
	500	.5938(.0478)	.1729(.0089)	1.3942(.5434)
CVM	100	.6261(.0015)	.2178(.0008)	1.5386(.0042)
.6 .2 1.5	200	.6206(.0009)	.2151(.0006)	1.5440(.0046)
	300	.6172(.0005)	.2141(.0004)	1.5408(.0038)
	400	.6162(.0005)	.2135(.0004)	1.5434(.0041)
	500	.6150(.0004)	.2122(.0002)	1.5441(.0039)

The table 5 lists the parameter estimates and goodness of fit statistics for bladder cancer patient data.

The LTNB distribution is more suitable for these data since the values of $-log\hat{l}$, AIC, AICC, BIC, and HQIC for the LTNB distribution are lower than those of the other competing models. Figure 4 displays the fitted densities.

9. Time series Models with LTNB marginals

Several researchers have developed and studied time series models with non-Gaussian marginals (see, for example, [11], [12], and [25]). The following definition is required in order to create the time series model with LTNB marginal distribution

Definition 9.1. A positive real valued random variable X is said to have Marshall-Olkin Lindley Truncated Negative binomial distribution and write $X \stackrel{d}{=}$MOLTNB$(\vartheta, \alpha, \gamma, \theta)$ if it has the survival function

$$\bar{F}_X(x) = \frac{1}{1+\frac{1}{\vartheta}\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^{\theta}-\gamma^{\theta}(\alpha+1)^{\theta}}{\gamma^{\theta}[(\alpha+1)^{\theta}-((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^{\theta}]}\right]}. \tag{21}$$

Theorem 4. The first order auto regressive (AR(1)) process given by

$$X_n = \begin{cases} \varepsilon_n & \text{w.p} \quad \delta \\ min(X_{n-1}, \varepsilon_n) & \text{w.p} \quad 1-\delta \end{cases} \tag{22}$$

Table 3: *Remission times of Bladder Cancer Patient Data.*

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23	3.52	4.98
6.97	9.02	13.29	0.40	2.26	3.57	5.06	7.09	9.22	13.80	25.74	0.50
2.46	3.64	5.09	7.26	9.47	14.24	25.82	0.51	2.54	3.70	5.17	7.28
9.74	14.76	26.31	0.81	2.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64
3.88	5.32	7.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
15.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01	1.19	2.75
4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33	5.49	7.66	11.25	17.14
79.05	1.35	2.87	5.62	7.87	11.64	17.36	1.40	3.02	4.34	5.71	7.93
11.79	18.10	1.46	4.40	5.85	8.26	11.98	19.13	1.76	3.25	4.50	6.25
8.37	12.02	2.02	3.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76
12.07	21.73	2.07	3.36	6.93	8.65	12.63	22.69				

Table 4: *Descriptive Statistics of Cancer data.*

Min.	Q_1	Median	Mean	Q_3	Max.	Var.
0.08	3.348	6.395	9.366	11.838	79.05	110.425

where $0 < \delta < 1; n \geq 1$, defines a stationary AR(1) minification process with LTNB(α, γ, θ) as marginal distribution if and only if ε_n's are i.i.d MOLTNB$(\delta^{-1}, \alpha, \gamma, \theta)$ with $X_0 \stackrel{d}{=} CL(\alpha, \gamma, \theta)$.

Proof. If $\{X_n\}$ is stationary with LTNB(α, γ, θ) marginals, then

$$\begin{aligned}
\bar{F}_{\varepsilon_n}(x) &= \frac{\bar{F}_X(x)}{\delta + (1-\delta)\bar{F}_X(x)} \\
&= \frac{\frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - 1\right\}}{\delta + (1-\delta)\frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - 1\right\}} \\
&= \frac{1}{1+\delta\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}. \qquad (23)
\end{aligned}$$

That is, ε_n's are i.i.d MOLTNB$(\delta^{-1}, \alpha, \gamma, \theta)$.

Conversely, if ε_n's are i.i.d MOLTNB$(\delta^{-1}, \alpha, \gamma, \theta)$ with $X_0 \stackrel{d}{=} CL(\alpha, \gamma, \theta)$, then,

$$\begin{aligned}
\bar{F}_{X_1}(x) &= \delta\bar{F}_{\varepsilon_1}(x) + (1-\delta)\bar{F}_{\varepsilon_1}(x)\bar{F}_{X_0}(x) \\
&= \delta\left\{\frac{1}{1+\delta\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}\right\} + \\
&\quad (1-\delta)\left\{\frac{1}{1+\delta\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}\right\} \\
&\quad \left\{\frac{1}{1+\left[\frac{((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta - \gamma^\theta(\alpha+1)^\theta}{\gamma^\theta[(\alpha+1)^\theta - ((\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x})^\theta]}\right]}\right\} \\
&= \frac{\gamma^\theta}{1-\gamma^\theta}\left\{\left[\frac{\alpha+1}{(\alpha+1)-(1-\gamma)(\alpha+1+\alpha x)e^{-\alpha x}}\right]^\theta - 1\right\}.
\end{aligned}$$

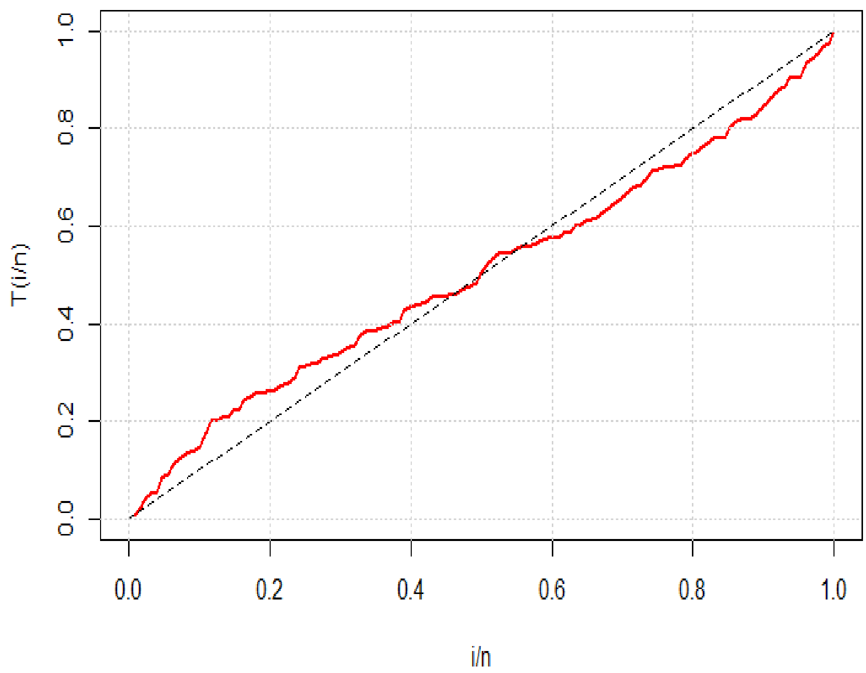


Figure 4: *The empirical TTT plot of the data.*

Table 5: *Parameter estimates and goodness of fit statistics for various models fitted to the data.*

Model	Estimates	$-log\hat{l}$	AIC	AICC	BIC	HQIC
LTNB	$\hat{\alpha} = 0.0617, \hat{\gamma} = 0.1218,$ $\hat{\theta} = 1.5407$	409.23	824.47	824.67	833.03	827.95
NGLD	$\hat{\alpha} = 4.679, \hat{\beta} = 1.324,$ $\hat{\theta} = 0.180$	412.75	831.50	831.69	840.06	834.98
BL	$\hat{\alpha} = 1.340, \hat{\beta} = 0.065,$ $\hat{\theta} = 1.861$	412.80	831.60	831.80	840.16	835.08
NG2PLD	$\hat{\alpha} = 0.8303, \hat{\theta} = 0.2942$	413.37	830.73	830.83	836.44	833.05
TLD	$\hat{\theta} = 0.156, \hat{\lambda} = -0.617$	415.15	834.31	834.41	840.01	836.62
LD	$\hat{\alpha} = 0.1960$	419.53	841.06	841.09	843.89	842.20
ED	$\hat{\alpha} = 0.097$	426.76	855.53	855.56	858.37	856.68

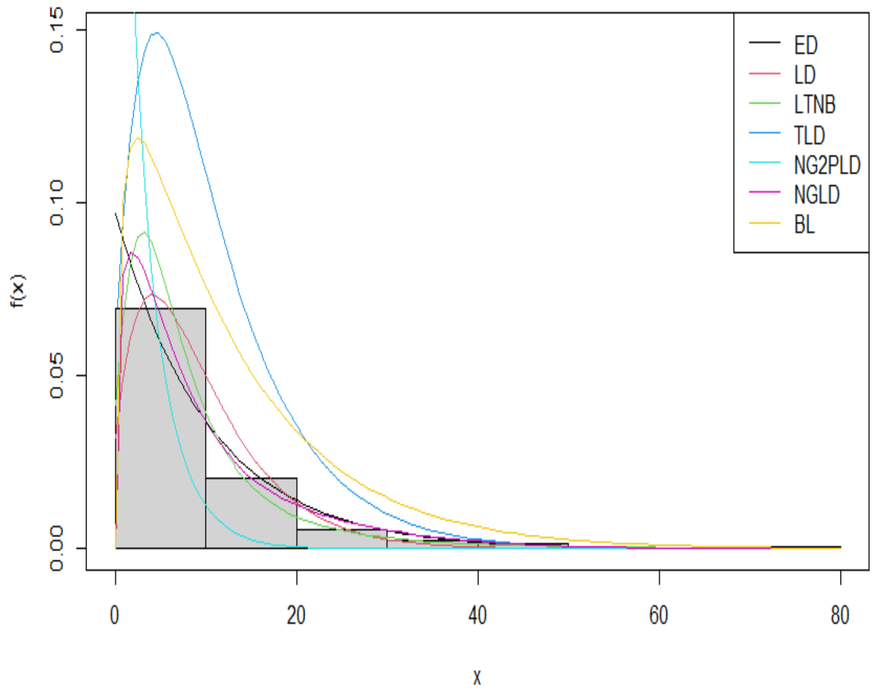


Figure 5: *Estimated pdf*

That is, $X_1 \stackrel{d}{=} LTNB(\alpha, \gamma, \theta)$.

If we assume that $X_{n-1} \stackrel{d}{=} LTNB(\alpha, \gamma, \theta)$, then by induction, we can establish that $X_n \stackrel{d}{=} LTNB(\alpha, \gamma, \theta)$. Hence the process $\{X_n\}$ is stationary with LTNB marginals.

Funding

The first author wishes to thank the University Grants Commission, New Delhi for the financial assistance in the form of Junior Research Fellowship.

References

[1] Aarset, M. V. (1987). How to identify bathtub hazard rate. *IEEE Transactions on Reliability* , 36: 106-108.

[2] Bapat S. R. and Bhardwaj R.(2021). On an inflated Unit-Lindley distribution. *Journal of Statistical Research* , 55: 299-311.

[3] Bakouch H., Al-Zahrani B., Al-Shomrani A., Marchi V. and Louzad F.(2012). An extended Lindley distribution. *Journal of the Korean Statistical Society* ,41: 75–85.

[4] Bhati D., Malik M. A. and Vaman H. J.(2015). Lindley-Exponential distribution: Properties and applications. *Metron* ,73: 335–357.

[5] Bennett S. (1983). Log-logistic regression models for survival data. *Applied Statistics*, 32: 165.

[6] Elbatal I., Merovci F. and Elgarhy M.(2013). A new generalized Lindley distribution. *Mathematical Theory and Modeling* , 3: 30-47.

[7] Ekhosuehi N., Opone F. and Odobaire F.(2018). New generalised two parameter Lindley distribution. *Journal of Data Science* ,16: 549-566.

[8] Ghitany M.E., Atieh B., and Nadarajah S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation* , 78: 493-506.

[9] Ghitany M.E., Alqallaf F., Al-Mutairi D.K. and Husain H.A.(2011). A two-parameter weighted Lindley distribution and its applications to survival data.*Mathematics and Computers in Simulation* , 81: 1190-1201.

[10] Glanzel (1987). A characterization theorem based on truncated moments and its application to some distribution families. *Mathematical statistics and probability theory, (Bad Tatzmannsdorf, 1986)*, B: 75-84.

[11] Jayakumar K. and Pillai R.N (1993). The first-order autoregressive Mittag–Leffler process. *Journal of Applied Probability* , 30: 462-466.

[12] Jayakumar K., Bindu Krishnan, and Hamedani G.G (2020). On a new generalization of Pareto distribution and its applications. *Communications in Statistics: Computation and Simulation* , 49: 1264-1284.

[13] Kenny, J. F. and Keeping, E., Mathematics of Statistics D. Van Nostrand Company, Princeton. 1962.

[14] Korkmaz C. M., Yousof H.M., Rasekhi M. and Hamedani G.G.(2018). The Odd Lindley Burr XII Model. Bayesian Analysis, Classical Inference and Characterizations. *Journal of Data Science* , 16: 327-354.

[15] Lee E. T. and Wang J. Statistical Methods for Survival Data Analysis, New York, Wiley, 2003.

[16] Lindley D.V. (1958). Fiducial distributions and Bayes' theorem . *Journal of the Royal Statistical Society B* , 20: 102-107.

[17] Marshall A.W and Olkin I.A (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika* , 84: 641-652.

[18] Merovci, F. and Sharma, V.K.(2014). The beta Lindley distribution: Properties and Applications . *Journal of Applied Mathematics* , 2014: 1-10.

[19] Moors J. J. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society, Series D*, 37: 25–32.

[20] Nadarajah S., Bakouch H.S., and Tahmasbi R. (2011). A Generalized Lindley distribution . *Sankhya B* ,73: 331-359.

[21] Nadarajah S., Jayakumar K. and Ristic M. M (2013). A new family of lifetime models. *Journal of Statistical Computation and Simulation* , 83: 1389-1404.

[22] Rajarshi S. and Rajarshi M. B. (1988). Bathtub distributions: A review . *Communications in Statistics - Theory and Methods* , 17: 2597-2621.

[23] Sharma V. K. and Khandelwal P.(2017). On the extension of the Inverse Lindley distribution *Journal of Data Science* , 15: 205-220.

[24] Shaked M., and Shanthikumar J.G. Stochastic orders, Springer, New York, 2007.

[25] Yeh H.C., Arnold B.C. and Robertson C.A. (1988). Pareto processes *Journal of Applied probability* , 25: 291-301.

RECOVERY PERIOD OF AIR TRANSPORTATION: A FORECAST WITH VECTOR ERROR CORRECTION MODEL

Tüzün Tolga İnan
•
Bahcesehir University
tuzuntolga.inan@sad.bau.edu.tr

Abstract

Air transport is the primary module of civil aviation and because of its nature, air transport has been simultaneously affected by Pandemics and crises. The influence of COVID-19 was more devastating than the other Pandemics and crises due to its global effect. This effect has continued a long period that still this effect exists now with a slight trend. The aim of this study is to analyse the selected variables that shows the past and future trend of air transportation related to operational and financial status. These variables are the primary ones that can define the countries' general status in air transport. The forecasting results are examined by 9-months forecasting with Vector Error Correction Model. It is forecasted that slightly decreasing trend will proceed in the following 9-months for passenger transportation due to fall and winter seasons. It is forecasted that slightly upward trend will proceed in the following 3-months and slightly decreased in the other 6-months for cargo transportation due to potential economic crisis in 2023. The originality of this paper is the first research related to analyse passenger and freight transportation together with the operational and financial parameters that defined in the sample of data and methodology sections.

Keywords: passenger load factor, cargo load factor, vector error correction model, air transportation, recovery period.

1. Introduction

Air transportation is a staminal facilitator for the countries' development. Air transportation includes air passenger and air freight modules. Especially air passenger module is related to the economic status and welfare of people. Trade and tourism primarily affect the country's development rates, so the development of air passenger transportation increases revenues with the development interest of demanded services in high amounts [1]. This study forecasts the following three quarters (9-month period) to analyse the negative impact of COVID-19 on air transportation. When time passes and the COVID-19 effect has decreased, air passenger demand will probably rise, and the growing trend of civil aviation will rise too. Nearly all countries prohibited travel to other countries because of the increment in COVID-19 cases. Several strategies regarding restrictions and procedures have been applied in countries, so air passenger traffic numbers decreased [2]. Before COVID-19 emerged, the annual development rate of the civil aviation industry in 21st century has 4.2%. More than 5000 airlines and 40 million flights clearly show that the global airline market value is measured by billions of dollars (USD). Furthermore, civil aviation affects the Gross Domestic Product (GDP) by more than 1% globally [3]. So, the importance of civil aviation is better than the other transportation modules. Nowadays, there have more than 1200 big-scale international airports globally and these airports provide the transportation of more than 4 billion passengers annually [4].

Besides air passenger, air freight is described as the transportation process of products by using an air carrier. Air freight is evaluated as an important transportation module when these goods are carried globally [5]. Air transportation has a significant development trend globally, so the revenues and populations have increased with the alteration of industrial structure. This status has affected the progress of free trade which enhanced this development trend globally [6]. With the development of air freight transportation demand, air traffic models have also altered and turned into a more complicated structure [7]. The profit level of this transportation module reached 40% in 2009, while this profit was only 5% in 2000 [8, 9]. In 2014, Boeing Company [10] forecasted that the air freight industry has proceeded to develop by a 4.7% per year. The forecasts show that this number will triple its income in 2033. The estimations specified that from 2013 to 2033, the billion ton-kilometers (RTKs) have risen from 207.8 to 521.8. There have several strategies for the sustainability of the stunning development of expanding international trade [11, 12]. So, airlines take place in the operation process of air freight. They enhance their strategic plans to reflect changes in the global competitive landscape [9; 13]. Thus, increasing theoretical researches have tried to solve the difficulties in the operational process in air freight since the 1990s [14]. Besides, there was a general supposition about the global aviation industry with an improvement trend in 2020. So, the status was the same as the growing years that has in a consecutive growth in air passenger, freight, and incomes in the latest years. Following this developing pattern, several airlines have made investments in purchasing and leasing more aircraft from well-known manufacturers like Airbus and Boeing, offering significant savings on new orders with long-term contracts [15].

In addition to these definitions related to air passenger and air freight transport, this study aims to analyze the 11 variables that mentioned below according to all active countries. These variables are;

- Gross Domestic Product (GDP),
- Total National Air Passenger Numbers,
- Total International Air Passenger Numbers,
- Total National Air Freight Numbers.
- Total International Air Freight Numbers
- Available Seat Kilometer for Air Passenger Transportation
- Revenue per Passenger Kilometer for Air Passenger Transportation
- Passenger Load Factor for Air Passenger Transportation
- Total Available Cargo Tonne Kilometer for Air Freight Transportation
- Revenue per Cargo Tonne Kilometer for Air Freight Transportation
- Cargo Load Factor for Air Freight Transportation

In this study, these 11 variables are gathered to figure out significant factors that define the operational and financial level of air transport (passenger and freight) for countries. This study also analyses the COVID-19 effect on air passengers and air freight transport (not affected as passenger transportation) as an unprecedented worldwide crisis globally.

2. Literature Review

When it is examined the related studies, first study is related to predict the relationship between the degree of economic shocks and the temporal recovery of the global air transport business was published by Gudmundson et al. [16]. According to this study's results, the global flight demand, particularly for passengers, will take at least 2.4 years (recovery by the end of 2022) to reach pre-

COVID-19 levels. The most optimistic forecast for the recovery has continued approximately two years (recovery by the third quarter of 2022), while the most pessimistic forecast for the recovery has continued approximately for six years (recovery in 2026). According to the methodology related to passenger and freight transportation, the recovery period forecast calculated by using a univariate approach called ARIMAX. This study was published at the beginning of 2021, so the recovery is not reflecting the up-to-date data at. Second study is related to air travel was released by Truong [17]. Truong analyzed the quantity of both domestic and foreign flights. This study specified that although the number of flights during the pandemic may not have reached 2019 levels, in 2022, the number of flights will not far away about the demand for travel in 2019. In that research, neural network models for estimating domestic and international air transportation in the medium and long terms were developed and tested. To estimate passenger demand, economic factors following COVID-19 that placed restrictions on transportation were examined. In the third study, Wang and Gao [18] analyzed 87 research about air transportation. They took these studies between 2010 to 2020. They used preliminary analytical techniques to analyze the input data. In the input data, three analyses were created and collected. This analysis revealed the relationship between the reviewed studies by forecasting airlines' socioeconomic and operational characteristics in time series modelling, the study reviewed air transportation demand on international scale.

In the fourth study, Dube et al. [19] specified that because of air travel is in full swing, general problems will not resolve right away. He proposed to aviation specialists that safeguarding passengers' health and safety, planning ticket prices, boosting efficiency, ensuring high-quality in-flight amenities, and maintaining such safeguards is likewise essential for the development of air traffic. In the fifth study, Li et al. [20] specified that there was a sharp drop in passenger air travel because of COVID-19 for two reasons. These are the breakdown of demand and supply of restrictions. The study segmented passengers according to their characteristics by applying several simulations and predictions of demand for each segment. In the sixth study, Zhang et al. [21] determined econometric and subjective models related to patterns of Hong Kong's tourism revival. These models described how the COVID-19 Pandemic affected Hong Kong's tourism industry economically. They also analyzed the airline revenues due to COVID-19 effect. In the seventh and last study, Xuan et al. [22] forecasted the air transport recovery period. He used the vector autoregression approach to calculate the period of recovery. The results showed that the decisive factors of the recovery are gross domestic product (GDP) and air freight traffic.

If such a Pandemic had not happened, there was a general supposition in the worldwide aviation industry to anticipate an increasing development trend in 2020. This status was the same as the consecutive growing years for air passenger and freight transportation. The forecast reports have shown this trend for airlines about the investments in buying new aircraft from widely known aircraft manufacturers such as Airbus and Boeing by directing new order deliveries. In addition to the manufacturing companies in civil aviation, the connection between the tourism and transportation sectors has broadly been debated academically. This debate is commonly focused on air transportation. Despite the existence of other modules in transportation such as railway, maritime, road vehicles, etc., tourism issue is the most decisive one for air transportation. Correspondingly, tourism is the preferential factor for the development of air transportation and the increment of GDP related to the economic level of countries [23, 24].

Civil aviation has turned into the most important one among the other industries in its contribution to worldwide economic development. However, the development and continuation of COVID-19 will cause economic issues related to excessive tourism costs on a global scale. Correspondingly, the implementation of measures is also about the operation process planning in a global fleet of narrow and wide-body aircraft [25]. International tourism, which stands on air transportation with a ratio of (58%) for arriving passengers to countries, has reached a standstill with important ratios of negative outcomes for tourism-linked activities and numbers of employment [26]. The necessity for the airlines related to stopping their flight operations has negatively affected the demand of passengers, so the airlines do not have enough options in the usage of airports for the sustainability of the flight planning strategy rather than to use aircraft for cargo flights. Besides, these selections are required to assist the airports and airlines by ensuring a particular field for using new aircraft parking positions. The crisis in the civil aviation industry is related to interminable, connected, and unresolved problems designated for the worldwide air transportation framework [27].

When benchmarking air passenger and freight transportation, air freight is more complicated than passenger transportation because this transportation process includes more strategies and more detailed processes. When compare with passenger transportation, a compound of weight and volume, various types of services, combinations, and network planning design are more detailed. The obvious differences between freight and passenger operations demonstrate that the multidisciplinary nature of freight transportation is more complicated than passenger transportation [28, 29, 11, 30]. In general, air freight transportation has higher ambiguity than passenger transportation because of its volume capacity. In passenger transportation passengers may cancel bookings, so the passengers that do not come to the aircraft have no place on the passenger list. Because of this, International Civil Aviation Organization (ICAO) permits airlines for selling tickets at more than %10 of their capacity. The booking capacity of air freight transportation is related to the freight forwarders' planning, and it can be assigned to the volume of cargo capacity. The cargo capacity volume plans the determined flights for six or twelve months [31]. The planning of the number of goods shipped with on-time performance is more important rather than booked reservations. These reservations compose elevated fluctuations due to the management of capacity. This situation shows freight forwarders generally do not require to give the price for unserviceable freights. Without punishment charges for unserviceable freights, the forwarder can fulfil the need to reduce risks to compete with the other companies. This status actualizes with several reservations in air freight that have been planned, reserved, and cancelled after the plan because cancelled flights are not an expense for the airlines. Consequently, the reservation process can show substantial volatility in air freight more than air passenger transportation [32].

It is more challenging to predict air freight capacity than air passenger capacity. Passenger aircraft capacity is related to its total seats, but freight capacity relies on the types and volume of containers and pallets named unit load devices (ULDs). The main problems with ULDs include variety of issues such as, pivot weight, volume, type of the product, and center of gravity [29]. For example, capacity has a connection with volume, and solely weight is not a determining factor. The basic specifications of air freight include complicated decision-making models for the management of air freight capacity. Transfer routes between the origin and destination (OD) pair are significant for air freight transportation. They serve the airline for passenger transportation. Generally, big-scale airlines known as full-service carriers control the process of alleged hub-and-spoke networks. Freight and passengers are carried from diversified origins to several hubs where freight and passengers are unified and afterward carry to other hubs for using wide-body aircraft. Passenger transportation can have a problem with inadmissible passengers. These passengers cannot embark to the aircraft due to the prohibitions. In air freight transportation, the freights can transfer via numerous midpoint airports such as origin point to destination point with a quick delivery time [31].

Airlines report the origin, stopover (transit), and destination airports for both passenger and freight traffic to build the transfer route plans. These processes cover network capacity usage and have a connection between conceptual and empirical modelling. These models have an increasing trend aiming to analyse quantitative decision methods related to the operational process especially for air freight [14]. Furthermore, the literature review covers the information about air transportation with defining the decreasing trend about all active countries because of COVID-19. Also, the literature review shows the decreasing and increasing trend of air transportation numbers by making a forecast (for passenger and freight) before and after the COVID-19 Pandemic.

3. Sample of Data

For the 11 selected variables in this study, time series modeling was applied. The primary thing that affects air transportation statistics is gross domestic product (GDP). All finished products and services produced within the country for a particular period (often one year) are included in the GDP [33]. Additionally, the Federal Reserve Bank of St. Louis provided the GDP data [34] The study's introduction and literature review sections both include descriptions of air passenger and freight transportation that indicate both the domestic and international numbers involved.

The six airline variables that are still in effect are: available passenger kilometers (ASK), revenue passenger kilometers (RPK), passenger load factor for air passenger transportation (PLF), available tonne kilometers (ATK), revenue tonne kilometers (RTK), and cargo load factor (CLF) for air freight transportation. All selected variables are used for the time series modeling analysis between January 2016 and August 2022 due to the availability of data.

First, RPK and the airline's total kilometer passenger capacity are connected. The total number of seats flown, and the distance are added to determine RPK. RPK reimburses all miles flown by paying passengers. The total distance traveled and the number of passengers who make revenue are multiplied to determine RPK. Because it assesses the current demand for air travel, sometimes known as airline "traffic." RPK determines the amount of demand for air travel regarding the labor force or workforce by calculating the number of passengers and distance traveled [35]. Secondly, ASK includes the available seat capacity of the aircraft, and it is a decisive data for the calculation of the airline’s transporting passenger capacity. When a seat is available for carrying, it can calculate for the ASK [36]. Thirdly, PLF is the last widely accepted variable. By dividing RPK by ASK, it is determined with the airline's capability for carrying passengers. The capacity of the passenger seats has an impact on the rate at which RPK and ASK are rising. It means PLF directly affects an increase in terms of ASK. So, PLF covers supply, demand, the total passenger numbers, and seat capacity [37].

Fourthly, ATK is related to the airline’s total kilometer freight capacity. It is obtained by multiplying the distance by the total volume of capacity flown. The number of available freights is significant for the evaluation of ATK and the calculation of the airline’s transporting freight capacity. When a compartment is available for carrying, it can calculate for ATK [38]. Fifthly, RTK covers the number of kilometers flown by paid freights. It is defined as the total distance traveled with the number of freights that generate revenue. It evaluates the actual demand for air travel as an airline's "capacity." RTK calculates the amount of labor or work power used to measure the level of demand for air travel. RTK is multiplied by the distance traveled and the freight expense [39]. LF is the final extensively used variable. By dividing RTK by ATK, it is computed with the airline's capacity to move freight. RTK and ATK both rise in response to an expansion in this freight transport capacity. This indicates that CLF has a direct impact on an increase in ATK. So, CLF includes total freight capacity, total carriage volume, total supply, and total demand [40].

These factors reflect how well airlines are performing financially and operationally. IATA has shared these six variables since January 2016, so the period of the analysis starts from January 2016 [41, 42]. In this study, air passenger and freight transportation variables named RPK, ASK, PLF, ATK, RTK, and CLF analyzed in the time series modelling [43], and they obtained from the IATA Air Passenger Monthly Analysis Reports [41] and Air Freight Monthly Analysis Reports [42]. Additionally, the air passenger and freight numbers took from the Bureau of Transportation Statistics (BTS) for Air Passenger Data [44] and Air Freight Data [45].

4. Methodology

As previously determined, this research uses IATA-released monthly data from January 2016 through August 2022. These factors are listed as follows: Gross Domestic Product (GDP), Domestic Passenger Numbers, International Passenger Numbers, Domestic Freight Numbers, International Freight Numbers, Available Seat Kilometer (ASK), Revenue Passenger Kilometer (RPK), Passenger Load Load Factor (PLF), Available Cargo Tonne Kilometers (ACTK), Revenue Tonne Kilometers (RTK), and Cargo Load Factor (CLF) [35, 36]. This study includes time series modeling to forecast PLF and CLF while analyzing the contributing variables. Therefore, endogenous factors ASK, RPK, GDP, domestic, and international passenger numbers are considered in the Vector Error Correction Model (VECM), while endogenous variables CTK, ACTK, GDP, domestic, and international cargo numbers are thought to affect the CLF. After the VECM model has been applied, the relationship between PLF, CLF, and the other endogenous variables is made known to impulse response functions. In time series modeling, the series that are connected to the selected variables are transformed with a logarithmic transformation and normalized to remove variability by subtracting the mean and separating the standard deviation. To avoid NaN results caused by logarithmic transformations with negative values, plus 1 is added to all series. The time series analysis is carried out using the following packages: "TSA, vars, urca, forecast, and tsDyn" in R 4.0.2. [46]. Understanding the stationarity and connection of the series can be accomplished by looking at the standardized series (Figure 1 and Figure 2).

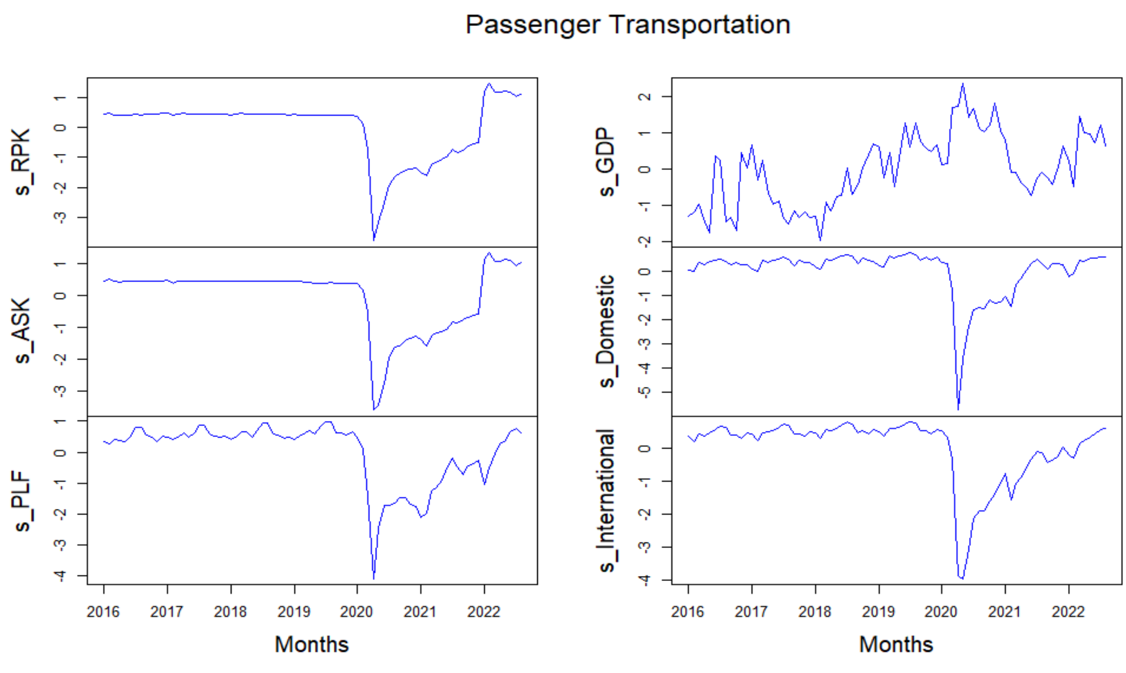


Figure 1: *The standardized time series plot of comparing s_RPK, s_GDP, s_Domestic, s_International, s_PLF and s_ASK*

In the period of January 2016 to August 2022, s_PLF, s_ASK, s_RPK, s_GDP, domestic, and international passenger number variables are shown in Figure 1. It is obvious that a major drop is reported in April 2020 due to COVID-19 impact on all series. Although the increase in every variable has slightly started and continue after May 2020. s_RPK and s_ASK variables pass their previous level at the beginning of 2022. Besides, other series approximately reach their previous level (slightly lower) at the beginning of 2022 and pass their previous level at the third quarter (July and August) of 2022.

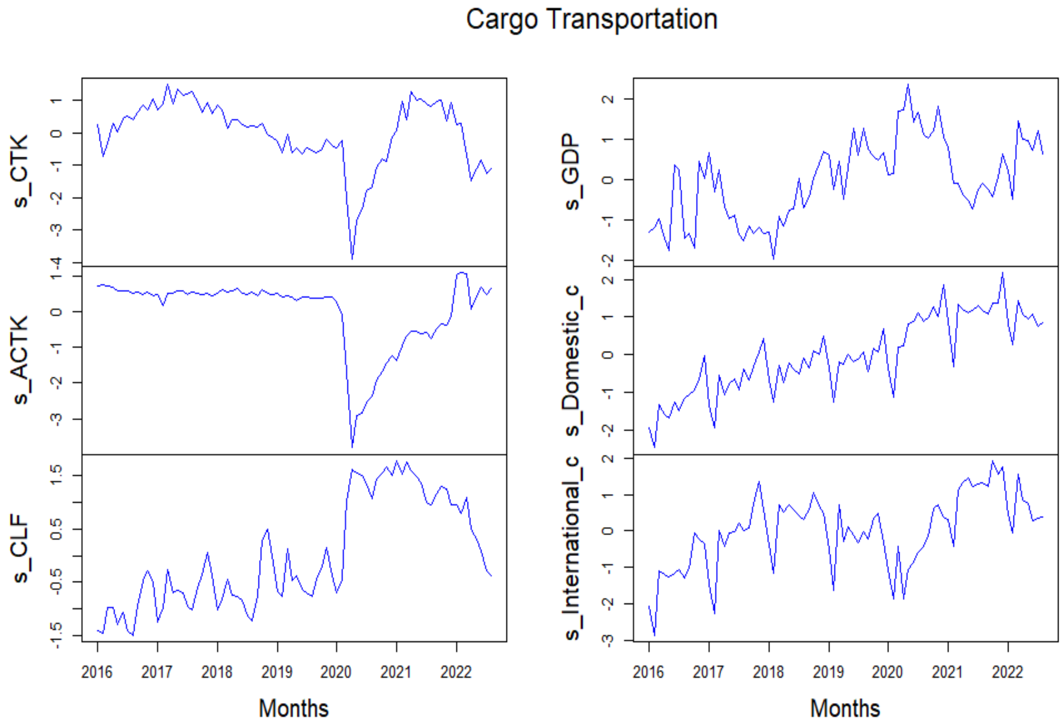


Figure 2: *The standardized time series plot of comparing s_CTK, s_ACTK, s_GDP, s_Domestic, s_International, and s_CLF*

In the period of January 2016 to August 2022, s_CTK, s_ACTK, s_GDP, s_Domestic, s_International, s_CLF parameters are shown in Figure 2. The COVID-19 Pandemic's impact is expected to result in a major drop in all data in April 2020, except for GDP. On May 1, 2020, the GDP will have increased and reached its peak value. In terms of s CLF, the series barely surpassed its prior level, but s_Domestic and s_International reached their top level in the third quarter (July to September) of 2022.

4.1. Unit root test and the lag-length determination

The idea of stationarity is very important in time series analysis. The covariance between two series values is expressed as the number of lags in the series, and a time series has a fixed mean, variance, and stationarity. Once trend and seasonality modifications have been made, time series models must be used (not changing over time). A subjective way of assessing stationarity is to use Augmented Dickey-Fuller (ADF) test statistics (47; 48]. The test is applied to a new model called the ADF test because the lags of the dependent series are expected to be added to the right of the equation to solve the autocorrelation problem. A test was suggested by Dickey et al. [49] related to issues in autoregressive time series. Here is how this test is shown.

$$\Delta y(t) = \alpha + \rho y(t-1) + \beta T + \sum_{s=1}^{p} \mathrm{d}_s \Delta y(t-s) + u_t \qquad (1)$$

In the formula, $\Delta y(t)$ is K-dimensional vector of observed variables, α is Kx1-dimensional constant vector, T is a time trend, u_t is the error term which has 0 mean and constant variance called white noise. Both the level of the series and their initial differences are tested in this procedure. In contrast to the alternative, the null hypothesis is that the series under examination has a unit root. In each example, the final prediction error (FPE) from Akaike is minimized to determine the lag-length [50]. Additionally, the Likelihood Ratio Test (LR), Akaike Information Criteria (AIC), Schwarz Information Criteria (SC), and Hannan-Quinn Information Criteria (HQ) can define the optimum lag-length. In every test except the LR test, the smallest value indicates the ideal lag-length. By putting the likelihood ratio statistics to the test at the selected level of significance, the LR test is discovered. The proper lag-length should be long enough to prevent autocorrelation between error terms but short enough to prevent any loss of information regarding the interaction of the series [51].

4.2. The cointegration test

The cointegration test determines whether there is a long-term relationship between the series after the stationary study between the series. Three alternative approaches are utilized in the literature to incorporate the cointegration test into the model. These methods were improved by Engle and Granger [52], Johansen and Juselius [53], and Pesaran, Shin, and Smith [54]. Johansen and Juselius' [53] technique was selected for this study because it allows for the examination of more than two variables. The absence of a cointegration vector in the series is the null hypothesis (H 0:r=0). A different possibility for the cointegration vector (H 0:r0) is that the time series have one (cointegrated series I(r)). R is a symbol for the quantity of cointegration vectors. If there is at least one cointegrated vector in the study, it means there is a long-term relationship between the series in the model. In other words, the order of these series remains constant. The VECM ought to be incorporated into the model because it is determined whether a long-term relationship exists at this point in the investigation.

4.3. Vector error correction model (VECM)

Long-term and short-term relationships between series are separated by VECM. Engle and Granger [46] created it to distinguish between short-term connections. It is attempted to assess whether the series experiences any shock over the long-term using VECM. According to Engle and Granger [52], the operating model for the VECM is as follows:

$$\Delta y(t) = \alpha\beta' \, y(t-1) + \Gamma_1 \, \Delta y(t-1) + \cdots + \Gamma_{p-1} \, \Delta y(t-p+1) + u_t \quad (2)$$

In the formula, $\Delta y(t)$ is K-dimensional vector of observed variables, α is Kxr-dimensional coefficient matrix, β is Kxr dimensional cointegration matrix, Γ_i is kxk-dimensional short_term coefficient matrix, and u_t is the error term which has 0 mean and constant variance called white noise. The error correction variable (β) acts to maintain the model dynamics in equilibrium and compels the variables to converge on the Error Correction Term (ECT), which is the long-term equilibrium value. When the error correction term's coefficient is statistically significant, bias is present. The coefficient size is a measure of how quickly the value of the long-term equilibrium is approaching. In actual, it is anticipated that the error correction variable will be statistically significant and negative. The variables in this instance are said to move in the direction of the long-term equilibrium value. Short-term departures from equilibrium will be rectified based on the size of the error correction variable's coefficient [55]. The lag order is p. The maximum lag and minimal Akaike Information Criteria are used to choose the lag-length p in the VECM.

4.4. Impulse response functions and the decomposition of forecast error variance

The estimated coefficients in VECM are quite challenging to comprehend. Impulse-response function (IRF) graphs, which are graphical representations of the reactions to varied shocks, are thus utilized to analyze the model's results. The vertical axis is used to generate the graphs of the impulse-response functions. The amplitude and direction of other series' responses indicate an increase in the standard deviation's response to the pertinent series. In a 12-month period, the shock is given a horizontal axis. Red-dashed lines reflect confidence intervals with 2 standard errors for how the variables will respond, and they are crucial in establishing the data' statistical significance. Indicating that the reaction is statistically significant at a 95% confidence level, the bottom and higher bands both had the same sign. The dotted lines on the graphs show the confidence intervals, and the straight lines on the graphs reflect the point estimates of the effect-reaction coefficients. To verify the relationship between the series, the Forecast Error Variance Decomposition (FEVD) approach is used. A specific variable's response to its own shock and the shock from other variables are evaluated in the VECM model, or FEVD. FEVD breaks down a variable's fluctuation into its individual shocks. It simply divides the variance of each variable's forecast mistakes between its own shocks and those of the other variables in the VECM [56].

4.5. Forecasting

Recursively, forecasts are made for the series' levels. By converting VECM to a VAR in R (using the VARrep function), forecasts for VECM can be obtained. Since a VECM with a lag of p corresponds to a VAR with a lag of p + 1, the new data for a VECM with a lag of p should have p + 1 rows.

5. Results

ADF test results for evaluating stationarity, cointegration test results, estimated VECM model results (for s PLF), s PLF impulse response, forecast error variance decomposition, forecasting of s PLF, and forecast values for passenger load factor are shown in the results section with tables and figures.

5.1. Unit root test and determining delay

The unit root of the series is the null hypothesis. The time series being stationary is an alternate theory (or trend-stationary). As shown in Table 2, the number of s_ASK, s_GDP, s_Domestic, and s_ International variables all reject the null hypothesis at the 0.05 significant level. The initial difference between s_PLF and s_RPK is discovered stationary. As shown in Table 1, the null hypothesis is rejected at the 0.05 significance level.

Table 1: *ADF test results to show the evoluation of stationary*

Variables	**s_PLF**	**s_RPK (1st diff.)**	**s_ASK**	**s_GDP**	**s_Domestic**	**s_International**
ADF test value	-2.368	-6.344	-2.399	-2.497	-3.118	-2.68
p	0.039	<0.001	0.009	<0.001	0.009	0.007
Variables	**s_CLF**	**s_CTK (1st diff.)**	**s_ACTK**	**s_GDP**	**s_Domestic_c**	**s_International_c**
ADF test value	-2.331	-6.491	-2.204	-2.497	-2.798	-3.724
p	0.034	<0.001	0.018	<0.001	<0.001	<0.001

The AIC, HQ, SC, and FPE tests are considered to estimate lag-length. The structure of VECM model is used in these tests to establish the ideal lag-length for this data set, which is 10. Despite the small sample size and the series' structure, the lag-length is set at 2. So, to forecast s_ PLF and s_CLF, the cointegration test is required. It can be utilized with the series in the estimate of regressions.

5.2. The test of cointegration

r indicates the cointegration equations number. The test statistic is low for r=1 for s_PLF and r=3 for s_CLF models at 5% significance level, rejecting the hypothesis. Table 2 shows the cointegration's presence that VECM was used to find the result.

Table 2: *Cointegration test results*

s_PLF	test	10pct	5pct	1pct	s_CLF	test	10pct	5pct	1pct
r <= 5 |	2.78	6.50	8.18	11.65	r <= 5 |	1.54	6.50	8.18	11.65
r <= 4 |	6.63	12.91	14.90	19.19	r <= 4 |	5.13	15.66	17.95	23.52
r <= 3 |	12.14	18.90	21.07	25.75	r <= 3 |	**20.61**	28.71	**31.52**	37.22
r <= 2 |	19.54	24.78	27.14	32.14	r <= 2 |	50.72	45.23	48.28	55.43
r <= 1 |	**29.01**	30.84	**33.32**	38.78	r <= 1 |	83.57	66.49	70.60	78.87
r = 0 |	49.63	36.25	39.43	44.59	r = 0 |	143.89	85.18	90.39	104.20

5.3. Vector error correction model (VECM)

The acquired results led to the formation of a VECM with r=1 and r=3 cointegration vectors. As previously indicated, the model's lag length is assumed to be 2. In Table 3, the estimated VECM results are shown.

Table 3: *VECM model results with an estimation (for s_PLF)*

Response	**s_PLF**	**Response**	**s_CLF**
Variables (lags)	Estimate (Standard Error)	Variables (lags)	Estimate (Standard Error)
s_RPK(-1)	1.886 (1.510)	s_CTK(-1)	-0.021 (0.100)
s_ASK(-1)	-1.571 (1.586)	s_ACTK(-1)	-0.079 (0.130)
s_PLF(-1)	0.888 (0.280)**	s_CLF(-1)	0.231 (0.135)·
s_GDP(-1)	0.003 (0.069)	s_GDP(-1)	-0.138 (0.065)*
s_Domestic (-1)	-0.726 (0.306)*	s_Domestic_c(-1)	0.071 (0.130)
s_International(-1)	-0.168 (0.470)	s_International_c(-1)	-0.202 (0.106)·
p<0.001		p<0.001	
ECT1= 0.181 (1.424)		ECT1=0.122 (0.042)** ECT2=-0.111 (0.068) **ECT3=-0.297 (0.105)****	

·p<0.10,*p<0.05, **p<0.001, ***p<0.0001

The Error Correction Term (ECT) controls how quickly long-run equilibrium returns. ECT1>0, ECT2>0, and ECT3>0, or at least one of them cannot be equal to 0, are necessary for a long-term connection to be stable. Table 3 shows that it does meet the prerequisite for a long-term stable connection. The error correction model will re-establish s_PLF's and s_CLF's long-term equilibrium because it is statistically significant and negative. Approximately 18% (s_PLF) and 30% (s_CLF) of the variances are corrected when there is a departure from equilibrium. The impulse-response functions are discussed in the next part to further explore the long-term impacts.

5.4. Impulse response function

For the results to be regarded as valid, both IRF confidence intervals must remain in the area above (or below) the zero band. As a result, conclusions from the research may only be drawn if the confidence intervals fall within the same range.

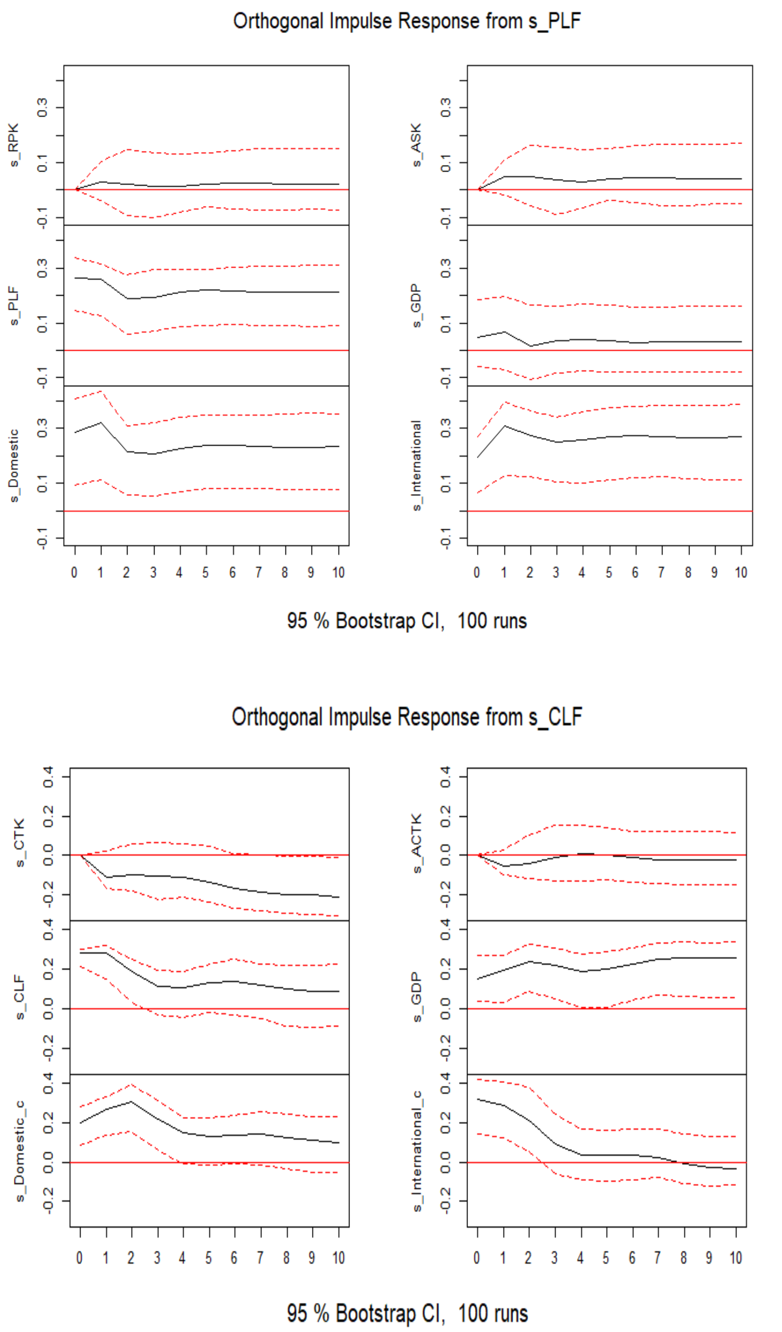
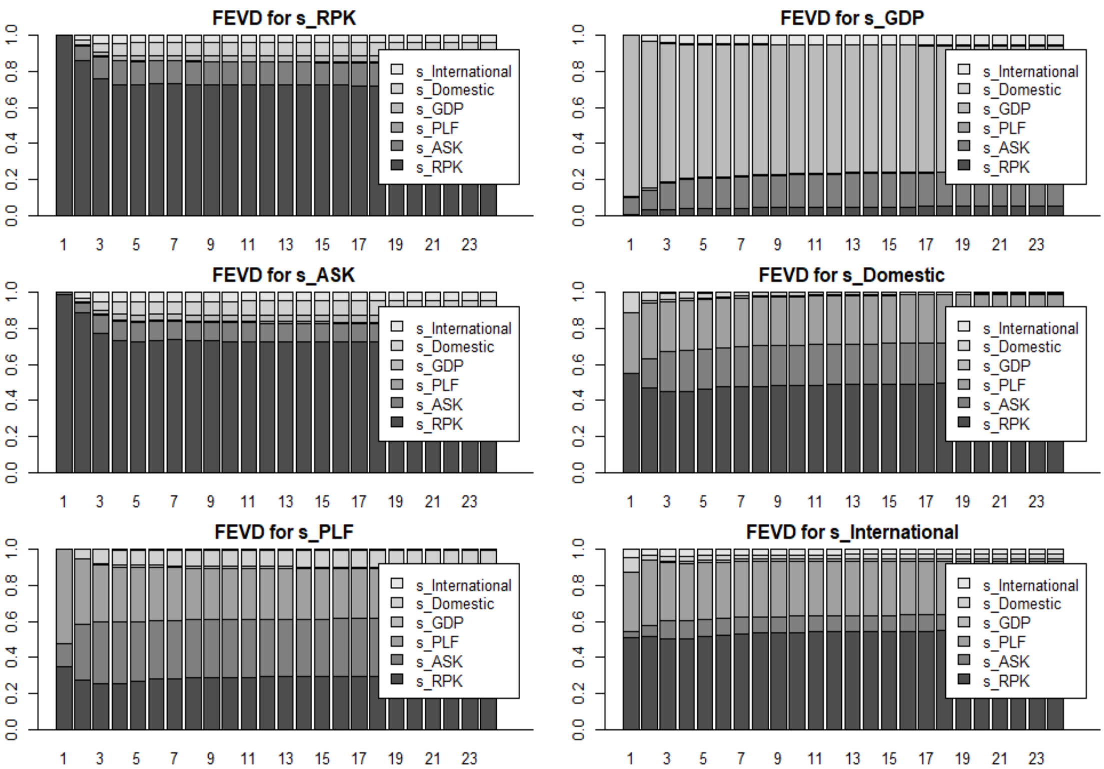


Figure 3: *s_PLF and s_CLF impulse response*

The top graph, in the middle of the left side, demonstrates that s_PLF started to fall after being temporarily touched by the effect, and then the effect disappears. According to the bottom-left graph, the short-term influence of s_Domestic_c on s_PLF increased, while the long-term effect of this impact gradually declined. The bottom right-hand graph demonstrates that after being touched by s_ International, s_PLF started to rise, and this impact gradually subsided in long-term. The bottom graph in the middle left-hand column demonstrates that s_CLF started to rise after being negatively affected by itself for a while, and eventually the effect disappears. The bottom right-hand graph demonstrates that the short-term influence of s_International_c on s_CLF was followed by an increase before the impact disappeared. The graph on the bottom left shows that s_CLF started to increase after being temporarily touched by s_Domestic_c, and subsequently this impact disappears. After being affected by s_ GDP in the near term, s_CLF started to rise, and this impact gradually lessened in long-term (Figure 3).

5.5. Decomposition of forecasting error variance

By dividing the variance of forecast error, it is possible to observe the impact of independent variables on s_PLF and s_CLF. The variance decomposition of forecast error examines the relative contributions of various variables and shocks to changes in a series. To find the impact of other series on a shock that happens in any of the series, variance decomposition is used. It defines the percentage of a shock unit that happened in one series that was brought on by changes in another.



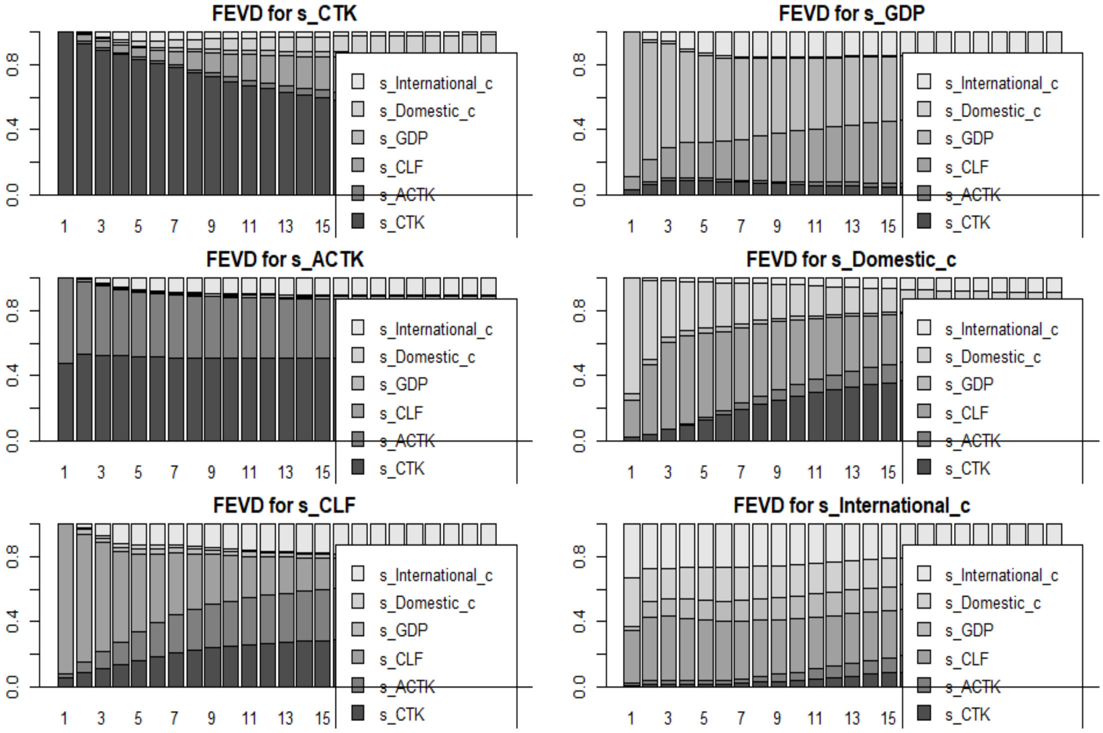
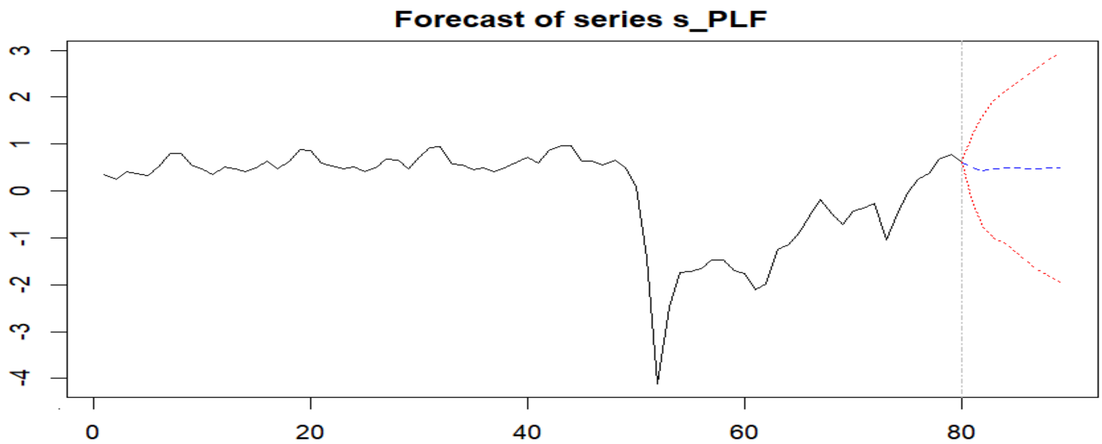


Figure 4: *Error variance decomposition of forecasting*

Figure 4 shows the FEVD results, which indicate that depending on how the series is coupled to the shock, the key variable influencing the s_PLF can alternate between the s_GDP, s_Domestic, and s_International. The primary factor influencing the s_CLF can either be s_International_c, s_GDP, or s_ CLF by itself depending on how the specific series is tied to the shock.

5.6. Forecasting

The VECM produces projected outcomes for the next nine months. Long-term results will be more accurate when the model is expanded to include the s_PLF and s_CLF values for each month. The findings of the forecast are as follows.



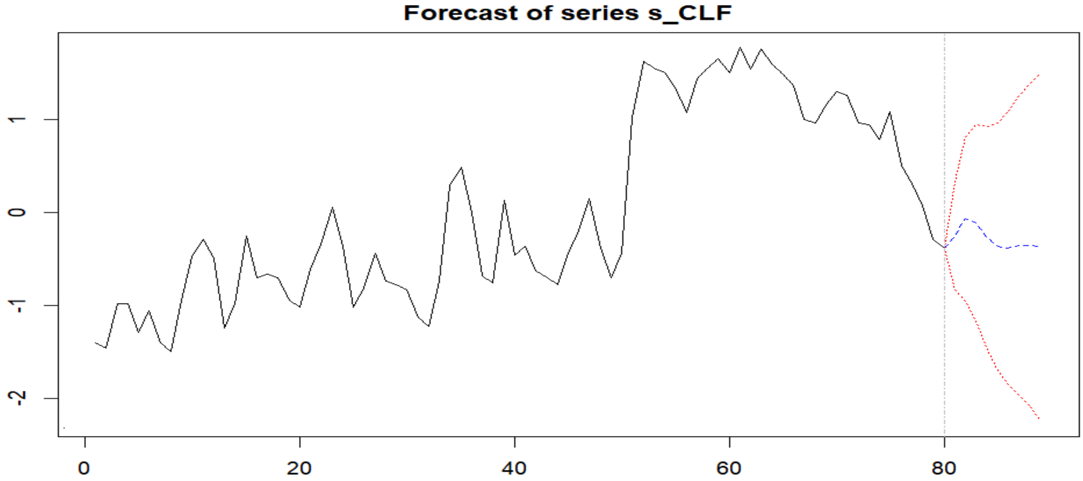


Figure 5: *The forecast of s_PLF and s_CLF*

In Figure 5, the red lines show the upper and lower boundaries, while the blue line show the conditional forecast. The s_PLF series are somewhat decreased after August whereas the s_CLF series are significantly raised after August that seen in the forecasting results in Figure 5.

Table 4: *The Forecasting Values of Passenger and Cargo Load Factor*

Date	**Actual PLF**	**95% CI Lower Limit PLF**	**95% CI Upper Limit PLF**	**Actual CLF**	**95% CI Lower Limit CLF**	**95% CI Upper Limit CLF**
August 2022	0.818	-	-	0.467	-	-
Date	**Forecast PLF**	**95% CI Lower Limit PLF**	**95% CI Upper Limit PLF**	**Forecast CLF**	**95% CI Lower Limit CLF**	**95% CI Upper Limit CLF**
September 2022	0.804	0.731	0.876	0,475	0,443	0,508
October 2022	0.797	0.676	0.919	**0,485**	0,436	0,534
November 2022	0.801	0.654	0.948	0,483	0,424	0,542
December 2022	0.804	0.640	0.969	0,474	0,407	0,540
January 2023	0.804	0.623	0.986	0,469	0,395	0,542
February 2023	0.803	0.603	1.003	0,468	0,387	0,549
March 2023	0.802	0.586	1.019	0,469	0,380	0,559
April 2023	0.803	0.571	1.035	0,470	0,373	0,566
May 2023	**0.804**	**0.558**	**1.050**	0,468	**0,365**	**0,572**

The predicted outcomes can be analyzed because model assumptions are revealed. With the VECM, the nine months forecasting can be applied to the analysis. Table 4 also includes the real data from December for comparison's sake. Due to the fall and winter seasons, it is anticipated that the slightly negative trend for s_PLF will persist for the next nine months. Therefore, the impact of the first and fourth quarters (October to December and January to March) stands a very low level. It is forecasted that the slightly upward trend will proceed in the following three months and slightly decreased in the other 6 months for s_CLF according to the potential of economic crisis in the first quarter (January to March) of 2023.

6. Conclusion

COVID-19 effect has negatively differentiated from all Pandemics and additionally crises due to its dissemination globally. Due to its nature, air transportation is the industry that mostly effected from crises and Pandemics globally. Air transportation includes air passenger and freight transportation, and they have different characteristics from each other. In general, air freight is more sophisticated than air passenger transportation because it includes more complicated processes. The calculation weight and volume, diversified services, combination, and reinforcement of network planning include more diversified specifications than passenger transportation. The significant distinctions between passenger and freight operations reveals that freight transportation have a more multidisciplinary system and higher uncertainty than passenger transportation due to the capacity problem related to volume. However, in passenger transportation passengers may cancel bookings, so the passengers that do not come to the aircraft have no place on the passenger list. Forecasting air freight capacity is also more complicated than forecasting air passenger capacity. Passenger aircraft capacity is related to total seat number in the aircraft, but freight capacity relies on the types, volume of containers and pallets named ULDs. Besides the differences between two modules of transportation, gross domestic product (GDP), total national air passenger numbers, total international air passenger numbers, total national air freight numbers, total international air freight numbers, available seat kilometer, revenue passenger kilometer, passenger load factor, available cargo tonne kilometer, revenue cargo tonne kilometer, and cargo load factor are the selected parameters for the application of time series modelling.

In the methodology part related to the analysis of air passenger transportation between January 2016 and August 2022, s_PLF, s_ASK, and s_RPK variables dramaticly decrease in April 2020 because of the COVID-19 Pandemic except s_GDP variable. Even if the increase has gradually begun and will continue to rise beyond May 2020, the series has not yet returned to its prior level (slightly lower). In the methodology part for air freight transportation between January 2016 and August 2022, s_CTK, s_ACTK, s_GDP, s_Domestic, s_International, s_CLF variables dramaticly decrease in April 2020 because of the COVID-19 Pandemic except s_GDP variable. The series returns to its former level for s_CLF, s_Domestic, and s_International with the rise trend in GDP on May 2020, although s_ GDP value has peaked. In the forecasting of a 9-month period using VECM, it is predicted that the slightly downward trend for s_PLF in air passenger transportation will continue in the coming 9 months due to weather conditions in the fourth and first quarters, whereas the slightly upward trend for s_CLF in air freight transportation will continue in the coming 3 months and slightly decrease in the remaining 6 months due to the potential economic problems globally in the first quarter. In future studies, a different analysis named Multidimensional Scaling can be applied with the same variables for all active countries in civil aviation on yearly basis.

Acknowledgments
Author has no personal interests from other parties.

Author Contributions
The author confirms contribution to the paper as follows: Data curation, Conceptualization, Investigation, Writing, Original draft preparation, Reviewing and Editing, Supervision, Resources, Methodology, Validation, Software, Formal analysis, Visualization.

References

[1] Schäfer, A. W., and Waitz, I. A. (2014). Air transportation and the environment. *Transport policy*, 34:1-4. doi: 10.1016/j.tranpol.2014.02.012

[2] Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Schubert, J., Bania, J., and Khosrawipour, T. (2020). The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *Journal of travel medicine*, 27(3). doi: 10.1093/jtm/taaa037

[3] International Air Transport Association (IATA). (2020). Airline expectations for 2020 improve ahead of virus outbreak, https://www.iata.org/en/iata-repository/publications/economic-reports/airline-expectations-for-2020-improve-ahead-of-virus-outbreak/ [retrieved 30 May 2022]

[4] International Air Transport Association (IATA). (2020). Economic Performance of the Airline Industry. iata.org/en/iata-repository/publications/economic-reports/airline-industry-economic-performance---december-2019---report/ [retrieved 30 May 2022]

[5] Air Freight. (2022). Retrieved From https://www.saloodo.com/logistics-dictionary/air-freight/ [retrieved 28 August 2022]

[6] Alekseev, K. P. G., and Seixas, J. M. (2009). A multivariate neural forecasting modeling for air transport-preprocessed by decomposition". *Journal of Air Transport Management*, 15(5):212-216.

[7] Xie, G., Wang, S., and Lai, K. K. (2014). Short-term forecasting of air passenger by using hybrid seasonal decomposition and least squares support vector regression approaches. *Journal of Air Transport Management*, 37:20-26. doi: 10.1016/j.jairtraman.2014.01.009

[8] Han, D. L., Tang, L. C., and Huang, H. C., (2010). A Markov model for single-leg air cargo revenue management under a bid-price policy. *European Journal of Operational Research*, 200(3): 800-811. doi: 10.1016/j.ejor.2009.02.001

[9] Nobert, Y., and Roy, J. (1998). Freight handling personnel scheduling at air cargo terminals. *Transportation Science*, 32(3):295-301.

[10] Boeing Company. (2016). World air cargo forecast 2014–2015.
http://www.boeing.com/assets/pdf/ commercial/cargo/wacf.pdf [retrieved 31 August 2021]

[11] Li, Y., Tao, Y., and Wang, F. A. (2009). Compromised large-scale neighborhood search heuristic for capacitated air cargo loading planning. *European Journal of Operational Research*, 199(2):553-560. doi:10.1016/j.ejor.2008.11.033

[12] Ou, J., Hsu, V. N., and Li, C. L. (2010). Scheduling truck arrivals at an air cargo terminal. *Production and Operations Management*, 19(1):83-97. doi: 10.3401/poms.1080.01068

[13] Ferguson, J., Kara, A. Q., Hoffman, K., and Sherry, L. (2013). Estimating domestic US airline cost of delay based on European model. *Transportation Research Part C: Emerging Technologies*, 33:311-323. doi: 10.1016/j.trc.2011.10.003

[14] Feng, B., Li, Y., and Shen, Z. J. M. (2015). Air cargo operations: Literature review and comparison with practices. *Transportation Research Part C: Emerging Technologies*, 56:263-280.
doi: 10.1016/j.trc.2015.03.028

[15] International Air Transport Association (IATA). (2020). COVID-19. Third Impact Assessment, 24 March 2020. https://www.iata.org/en/iata-repository/publications/economic-reports/third-impact-assessment/ [retrieved 25 August 2022]

[16] Gudmundsson, S. V., Cattaneo, M., and Redondi, R. (2021). Forecasting temporal world recovery in air transport markets in the presence of large economic shocks: The case of COVID-19. *Journal of Air Transport Management*, 91:102007. doi: 10.1016/j.jairtraman.2020.102007

[17] Truong, D. (2021). Estimating the impact of COVID-19 on air travel in the medium and long term using neural network and Monte Carlo simulation. *Journal of Air Transport Management*, 96:102126. doi: 10.1016/j. jairtraman.2021.102126

[18] Wang, S., and Gao, Y. A. (2021). Literature review and citation analyses of air travel demand studies published between 2010 and 2020. *Journal of Air Transport Management*, 97:102135.

doi: 10.1016/j. jairtraman.2021.102135

[19] Dube, K., Nhamo, G., and Chikodzi, D. (2021). COVID-19 pandemic and prospects for recovery of the global aviation industry. *Journal of Air Transport Management,* 92:102022. doi: 10.1016/j.jairtraman.2021.102022

[20] Li, X., Groot, M. de, and Bäck, T. (2021). Using forecasting to evaluate the impact of COVID-19 on passenger air transport demand, *Decision Sciences.* doi: 10.1111/deci.12549

[21] Zhang, H., Song, H., Wen, L., and Liu, C. (2021). Forecasting tourism recovery amid COVID-19. *Annals of Tourism Research*, 87:103149. https://doi.org/10.1016/j.annals.2021.103149

[22] Xuan, X., Khan, K., Su, C. W., and Khurshid, A. (2021). Will COVID-19 Threaten the Survival of the Airline Industry?. *Sustainability,* 13(21):11666. doi: 10.3390/su132111666

[23] Khan, S. A. R., Qianli, D., SongBo, W., Zaman, K., and Zhang, Y. (2017). Travel and tourism competitiveness index: The impact of air transportation, railways transportation, travel and transport services on international inbound and outbound tourism. *Journal of Air Transport Management,* 58:125-134. doi: 10.1016/j.jairtraman.2016.10.006

[24] World Travel and Tourism Council (WTTC). (2018). Travel & tourism: Economic Impact 2017. [retrieved 30 July 2021]

[25] International Air Transport Association (IATA). (2020). May Passenger Demand Shows Slight Improvement. https://www.iata.org/en/pressroom/pr/2020-07-01-02/ [retrieved 10 August 2021]

[26] Transport&Environment (T&E). (2021) Greenpeace, Carbon Market Watch Airline Bailout Tracker. https://www.transportenvironment.org/sites/te/files/Airline-bailout-tracker [retrieved 10 August 2022]

[27] Browne, A., St-Onge Ahmad, S., Beck, C. R., and Nguyen-Van-Tam, J. S. (2016). The roles of transportation and transportation hubs in the propagation of influenza and coronaviruses: a systematic review. *Journal of travel medicine,* 23(1), tav002.

[28] Bartodziej, P., Derigs, U., Malcherek, D., and Vogel, U. (2009). Models and algorithms for solving combined vehicle and crew scheduling problems with rest constraints: an application to road feeder service planning in air cargo transportation. *OR spectrum,* 31(2):405-429. doi: 10.1007/s00291-007-0110-7

[29] Leung, L. C., Van Hui, Y., Wang, Y., and Chen, G. A. (2009). 0–1 LP model for the integration and consolidation of air cargo shipments. *Operations Research,* 57(2), 402-412.

[30] Wang, Y. J., and Kao, C. S. (2008). An application of a fuzzy knowledge system for air cargo overbooking under uncertain capacity. *Computers & Mathematics with Applications,* 56(10):2666- 2675. doi: 10.1016/j.camwa.2008.02.049

[31] Amaruchkul, K., Cooper, W. L., and Gupta, D. (2011). A note on air-cargo capacity contracts. *Production and Operations Management,* 20(1):152, 162. doi: 10.1111/J.1937-5956.2010.01158.x

[32] Petersen, J. (2007). Air freight industry-white paper. Georgia Institute of Technology.

[33] Convention on the Organisation for Economic Co-operation and Development (OECD) Data. (2022). Gross Domestic Product (GDP). https://data.oecd.org/gdp/gross-domestic-product-gdp.htm [retrieved 31 October 2022]

[34] Federal Reserve Bank of St. Louis. (2022). https://fred.stlouisfed.org/series/GEPUCURRENT [retrieved 31 October 2022]

[35] Revenue Passenger Kilometers. (2022). https://airlinegeeks.com/2016/01/17/airline-metrics-revenue-passenger-kilometers/ [retrieved 31 October 2022]

[36] Available Seat Kilometers. (2022). https://airlinegeeks.com/2015/12/28/airline-metrics-available-seat-kilometers/ [retrieved 31 October 2022]

[37] Passenger Load Factor. (2022). https://airlinegeeks.com/2016/01/29/airline-metrics-passenger-load-factor/ [retrieved 31 October 2022]

[38] Available Tonne Kilometers. (2022). https://airlinegeeks.com/2015/12/28/airline-metrics-

available-tonne-kilometers/ [retrieved 31 October 2022]

[39] Revenue Tonne Kilometers. (2022). https://airlinegeeks.com/2016/01/17/airline-metrics-revenue-tonne-kilometers/ [retrieved 31 October 2022]

[40] Cargo Load Factor. (2022). https://airlinegeeks.com/2016/01/29/airline-metrics-passenger-load-factor/ [retrieved 31 October 2022]

[41] International Air Transport Association (IATA). (2022). Air Passenger Monthly Analysis Reports.https://www.iata.org/en/publications/economics/?Search=&EconomicsL2=146&Ordering=DateDesc [retrieved 01 November 2022]

[42] International Air Transport Association (IATA). (2022). Air Freight Monthly Analysis Reports.https://www.iata.org/en/publications/economics/?Search=&EconomicsL2=147&Ordering=DateDesc [retrieved 01 November 2022]

[43] Tsai, W. H., and Kuo, L. (2004). Operating costs and capacity in the airline industry. *Journal of air transport management,* 10(4):269-275.
doi: 10.1016/j.jairtraman.2004.03.004

[44] Bureau of Transportation Statistics (BTS). (2022). Air Passenger Data.
https://www.transtats.bts.gov/Data_Elements.aspx?Data=1 [retrieved 01 November 2022]

[45] Bureau of Transportation Statistics (BTS). (2022). Air Freight Data.
https://www.transtats.bts.gov/freight.asp [retrieved 01 November 2022]

[46] Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of statistical software,* 27:1-32.

[47] Dickey, D. A., and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association,* 74(366a):427-431.
doi: 10.1080/01621459.1979.10482531

[48] Dickey, D. A., and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: journal of the Econometric Society,* 1057-1072.

[49] Dickey, D. A., Hasza, D. P., and Fuller, W. A. (1984). Testing for unit roots in seasonal time series. *Journal of the American statistical Association,* 79(386):355-367.

[50] Akaike, H. (1969). Fitting autoregressive models for prediction, *Annals of the institute of Statistical Mathematics,* 21(1):243, 247.

[51] Kasapoğlu, Ö., and Müdürlüğü, P. G. (2007). *Parasal aktarim mekanizmalari: Türkiye için uygulama.* TCMB Uzmanlık Yeterlilik Tezi, Ankara: Şubat. 2007.

[52] Engle, R. F., and Granger, C. W. (1987). *Co-integration and error correction: representation, estimation, and testing.* Econometrica: journal of the Econometric Society, 251, 276.

[53] Johansen, S., and Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration—with appucations to the demand for money. *Oxford Bulletin of Economics and statistics,* 52(2):169, 210.

[54] Pesaran, M. H., Shin, Y., and Smith, R. J., (2001). Bounds testing approaches to the analysis of level relationships. Journal of applied econometrics, 16(3):289-326. doi: 10.1002/jae.616

[55] Enders, W. (1995). *Applied Econometric Time Series, John Wiley&Sons. Inc.,* New York, 365, 366.

[56] Omisakin, D., and Olusegun, A. (2008). Oil price shocks and the Nigerian economy: a forecast error variance decomposition analysis. *Journal of Economic Theory,* 2(4):124, 130.

RELIABILITY MODELING OF TWO-UNIT GAS TURBINE SYSTEM CONSIDERING THE EFFECT OF HUMIDITY

Pinki
•
Department of Mathematics, Maharshi Dayanand University Rohtak, India
pinkichahar95@gmail.com

Dalip Singh
•
Department of Mathematics, Maharshi Dayanand University Rohtak, India
dsmdur@gmail.com

Abstract

Aim. The purpose of this paper is to find the reliability measures and profit of a two-unit gas turbine power generating system incorporated with one gas turbine and one steam turbine. Effects of different humidity condition (humidity ≤/> 50%) are taken into consideration by fixing the range of temperature (5°C-25°C) for developing the model. At initial stage, both units (gas turbine and steam turbine) are in operative mode. If steam turbine fails, gas turbine remains in operative mode but if gas turbine fails, system goes to down state and when both unit fails, system fails. In this system we assume that failure time distribution is exponentially distributed while repair time distribution is arbitrary. Methods. In this paper we use the Laplace transform for mathematical analysis, and semi-markov process and regenerative point technique to investigate reliability measures and profit of the system. Findings. The system is analysed in steady state and different reliability measures such as mean time to system failure, availability for different cycles and for different humidity conditions, busy period, down time of the system etc. are calculated and the graphs have been drawn to see the effect of different transition rates such as failure rate and repair rate of the units for different humidity conditions on reliability measures and the profit for particular case is evaluated using the information/data collected from gas turbine power generating system located at Bawana, Delhi, India. Conclusion. Our finding shows that mean time to system failure and availability when both turbines are working decreases with increase in any one of failure rate while availability when only gas turbine is working increases with increase in steam turbine failure rate and profit for plant decreases with increase in failure rates. From this we concluded that availability for the fixed range of temperature (5°C-25°C) is higher when humidity is >50% as compared to when humidity is ≤50%. Thus, a comprehensive study of gas turbine system may be helpful to those who are involved in power generating industry.

Keywords: Gas Turbine, Steam Turbine, Failure Rate, Reliability Measures

1. Introduction

In this era of high competition, reliability has been widely progressed over the years to attain the needs and requirements of global competition. Reliability analysis of system plays pivotal role in deciding the productivity and profitability of the system. Gas turbine power systems are regarded as key element in industrial production and any deficiency in power supplying may lead to significant financial detriment as large capital investments are required for industrial production. Several studies have been conducted so far for gas turbine power system by assuming different failure and repair polices. A number of researcher have studied the system of two or more dissimilar/similar units including [1-3] in which one unit is operation mode and another one is standby state by using different repairs. El-Berry [4] discussed reliability based on failure of data of a gas turbine power plant. Singh and Taneja [5-6] introduced a situation of gas turbine power plant where units are dissimilar but nature of output is same by considering effect of random/scheduled inspection. Farouk and Sheng [7] proposed a situation for power plant to study effect of ambient temperature on gas turbine. Abigail et al. [8] studied the combined cycle power plant considering effect of ambient temperature with post-combustion CO_2 capture. Saleh et al. [9] discussed the performance of gas turbine in Saudi weather condition. Rajesh et al. [10-11] focused on reliability analysis of gas turbine power plant with effect of ambient temperature with different repair polices. Fernandez et al. [12] studied the effect of temperature in tropical climate on performability of gas turbine. Bird and Grabe [13] discussed humidity effects on gas turbine performance. Hanachi et al. [14] focused on the effects of intake air humidity on monitoring of gas turbine system but they did not find reliability measures like Mean time to system failure, Availability, Down time, etc.

Taking all above into consideration, in present paper we discuss stochastic modeling of gas turbine system (One Gas Turbine and One Steam Turbine). Here we fix the range of temperature (5°C-25°C) and develop model for two different humidity conditions (i.e. humidity less than or equal to 50% and humidity greater than 50%) are taken into consideration while developing model. Initially, system works with full capacity that is both the gas and the steam turbine are in operative state. If steam turbine fails, gas turbine continues to work then system is said to be working in single cycle but if gas turbine fails then steam turbine is put to down mode and system is said to be in down state. When both the gas and the steam turbine fails, the system is claimed to be in failed state. Different system effectiveness measures have been obtained by using semi-markov process and regenerative point technique. We obtained system effectiveness measures like mean time to system failure, availability analysis for full capacity as well as for single cycle for different humidity conditions. Interesting conclusions have been drawn by using assessed values on the basis of information/data gathered from a gas turbine power plant situated at Bawana, Delhi, India.

2. System Description and Assumptions

2.1. Assumptions

A two-unit gas turbine system is developed under some reasonable assumptions which are as follows:

- The failure time distribution is taken exponentially while repair time distribution is arbitrary.
- After each repair unit is claimed to be good as new unit.
- Repair pattern of the system rely on first come first serve.
- Complete failure of system is claimed on failure of both the units.

2.2. System Description of the Model

State transition diagram for gas turbine system consisting of a gas and a steam turbine is shown in Figure 1. States 0,1,2,3,5, and 6 are the regenerative points and hence 0,1,2,3,5 and 6 are regenerative states. States 4 and 7 corresponds to complete failure of the system. States 0 and 1 are the states representing production in combined cycle when humidity is less than equal to 50% and greater than 50% respectively. States 2 and 5 are the states at which point only gas turbine is working, representing production in single cycle when humidity is less than equal to 50% and greater than

50% respectively. At state 3 and 6, it is required to put steam turbine in down mode as on failure of gas turbine it can't work and hence we say state 3 and 6 are down state.

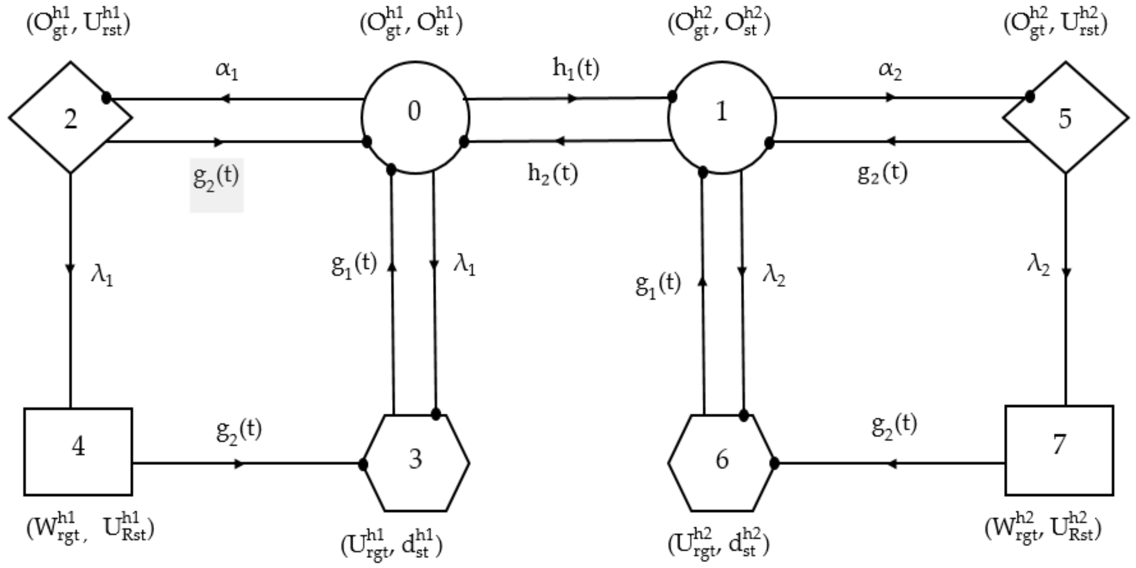


Figure 1: *State Transition Diagram of the System*

2.3. Notations

$O_{gt}^{h_1}/O_{st}^{h_1}$: Gas Turbine/Steam Turbine is operative when humidity ≤ 50%

$O_{gt}^{h_2}/O_{st}^{h_2}$: Gas Turbine/Steam Turbine is operative when humidity > 50%

$U_{rgt}^{h_1}/U_{rst}^{h_1}$: Gas Turbine/Steam Turbine is under repair when humidity ≤ 50%

$U_{rgt}^{h_2}/U_{rst}^{h_2}$: Gas Turbine/Steam Turbine is under repair when humidity > 50%

$U_{Rst}^{h_1}/U_{Rst}^{h_2}$: Continuing repair from previous state of Steam Turbine when humidity is ≤/> 50%

$d_{st}^{h_1}/d_{st}^{h_2}$: Steam Turbine is put to down mode when humidity is ≤/> 50%

$W_{rgt}^{h_1}/W_{rgt}^{h_2}$: Gas Turbine waiting for repair when humidity is ≤/> 50%

λ_1/λ_2 : Failure rate of Gas Turbine when humidity is ≤/> 50%

α_1/α_2 : Failure rate of Steam Turbine when humidity is ≤/> 50%

$g_1(t)/g_2(t)$: Probability density function of repair time of Gas Turbine/Steam Turbine

$G_1(t)/G_2(t)$: Cumulative distribution function of repair time of Gas Turbine/Steam Turbine

$h_1(t)/h_2(t)$: Probability density function for changing the humidity from ≤ 50% to > 50% / from > 50% to ≤ 50%

$q_{ij}(t), Q_{ij}(t)$: Probability density function/Cumulative distribution function of first passage time from regenerative state I to a regenerative state j or to a failed state j without visiting any other regenerative state in (0, t]

$q_{ij}^{(k)}(t)/Q_{ij}^{(k)}$: Probability density function/Cumulative distribution function of first passage time from regenerative state i to a regenerative state j or visiting state k one time in (0, t]

©/ Ⓢ : Laplace convolution/ Laplace Stieltjes convolution

3. State Transition Probabilities and Mean Sojourn Time

Based on state transition diagram, expression $dQ_{ij} = q_{ij}(t)dt$ for all essential combinations of i and j are derived and the transition probabilities p_{ij} are obtained by using Laplace transform and using $p_{ij} = \lim_{s\to 0} q_{ij}^{*}(s)$. Table 1 represents the state transition probabilities of the system.

Table 1: *State Transition Probabilities*

$dQ_{01} = e^{-(\alpha_1+\lambda_1)t}h_1(t)dt$	$dQ_{02} = \alpha_1 e^{-(\alpha_1+\lambda_1)t}\bar{H}_1(t)dt$	$dQ_{03} = \lambda_1 e^{-(\alpha_1+\lambda_1)t}\bar{H}_1(t)dt$
$dQ_{10} = e^{-(\alpha_2+\lambda_2)t}h_2(t)dt$	$dQ_{15} = \alpha_2 e^{-(\alpha_2+\lambda_2)t}\bar{H}_2(t)dt$	$dQ_{16} = \lambda_2 e^{-(\alpha_2+\lambda_2)t}\bar{H}_2(t)dt$
$dQ_{20} = e^{-\lambda_1 t}g_2(t)dt$	$dQ_{24} = \lambda_1 e^{-\lambda_1 t}\bar{G}_2(t)dt$	$dQ_{24}^{(4)} = (1-e^{-\lambda_1 t})g_2(t)dt$
$dQ_{30} = g_1(t)dt$	$dQ_{51} = e^{-\lambda_2 t}g_2(t)dt$	$dQ_{57} = \lambda_2 e^{-\lambda_2 t}\bar{G}_2(t)dt$
$dQ_{56}^{(7)} = (1-e^{-\lambda_1 t})g_2(t)dt$	$dQ_{61} = g_1(t)dt$	$p_{01} = h_1^{*}(\alpha_1+\lambda_1)$
$p_{02} = \frac{\alpha_1}{\alpha_1+\lambda_1}[1- h_1^{*}(\alpha_1+\lambda_1)]$	$p_{03} = \frac{\lambda_1}{\alpha_1+\lambda_1}[1- h_1^{*}(\alpha_1+\lambda_1)]$	$p_{10} = h_2^{*}(\alpha_2+\lambda_2)$
$p_{15} = \frac{\alpha_2}{\alpha_2+\lambda_2}[1- h_2^{*}(\alpha_2+\lambda_2)]$	$p_{16} = \frac{\lambda_2}{\alpha_2+\lambda_2}[1- h_2^{*}(\alpha_2+\lambda_2)]$	$p_{20} = g_2^{*}(\lambda_1)$
$p_{24} = 1-g_2^{*}(\lambda_1)$	$p_{23}^{(4)} = 1-g_2^{*}(\lambda_1)$	$p_{30} = 1 = p_{61}$
$p_{51} = g_2^{*}(\lambda_2)$	$p_{57} = 1-g_2^{*}(\lambda_2)$	$p_{56}^{(7)} = 1-g_2^{*}(\lambda_2)$

Mean Sojourn Time (μ_i) is the amount of time expected to spend in state i by the system. The expressions for μ_i are obtained by using $\mu_i = \int_0^{\infty} P[T_i>t]dt$ where T_i represents stay time of the system in state i. Table 2 represents means sojourn time of the system.

Table 2: *Mean Sojourn Time*

$\mu_0 = \frac{1}{\alpha_1+\lambda_1}[1- h_1^{*}(\alpha_1+\lambda_1)]$	$\mu_1 = \frac{1}{\alpha_2+\lambda_2}[1- h_2^{*}(\alpha_2+\lambda_2)]$	$\mu_2 = \frac{1}{\lambda_1}[1- g_2^{*}(\lambda_1)]$
$\mu_3 = \int_0^{\infty}\overline{G_1}(t)dt$	$\mu_5 = \frac{1}{\lambda_2}[1- g_2^{*}(\lambda_2)]$	$\mu_6 = \int_0^{\infty}\overline{G_1}(t)dt$
$\mu_2' = \int_0^{\infty}\overline{G_2}(t)dt$		

4. Mean Time to System Failure

$\phi_i(t)$ denotes the cumulative distribution function of first passage time to a failed state from regenerative state i. Recursive relations to obtain mean time to system failure of the system are:

$$\phi_0(t)= Q_{01}(t) Ⓢ \phi_1(t)+Q_{02}(t) Ⓢ \phi_2(t)+Q_{03}(t) Ⓢ \phi_3(t) \quad (1)$$

$$\phi_1(t)= Q_{10}(t) Ⓢ \phi_0(t)+Q_{15}(t) Ⓢ \phi_5(t)+Q_{16}(t) Ⓢ \phi_6(t) \quad (2)$$

$$\phi_2(t)= Q_{20}(t) Ⓢ \phi_0(t)+Q_{24}(t) \quad (3)$$

$$\phi_3(t)= Q_{30}(t) Ⓢ \phi_0(t) \quad (4)$$

$$\phi_5(t)= Q_{51}(t) Ⓢ \phi_1(t)+Q_{57}(t) \quad (5)$$

$$\phi_6(t)= Q_{61}(t) Ⓢ \phi_1(t) \quad (6)$$

Applying Laplace Stieltjes Transform on both sides of above relations and solve them using Cramer's Rule, we get

$$MTSF = \lim_{s\to 0}\frac{1-\phi_0^{**}(s)}{s} = \frac{N}{D} \quad (7)$$

Where, $N = \mu_1 p_{01}+\mu_3 p_{01}p_{16}+\mu_5 p_{01}p_{15}+\mu_0 p_{15}p_{57}+\mu_2 p_{02}p_{15}p_{57}+\mu_3 p_{03}p_{15}p_{57}+\mu_0 p_{10}+\mu_2 p_{02}p_{10}+\mu_3 p_{03}p_{10}$

$D = p_{02}p_{24}p_{15}p_{57}+p_{02}p_{10}p_{24}+p_{01}p_{15}p_{57}$

5. Availability in System when Humidity is ≤ 50%

$AH_i^1(t)$/ $AH_i^{1s}(t)$ denotes the probability that system available in full capacity/single cycle when humidity is ≤ 50% at any instant of time t provided it has entered regenerative state i at time t=0. By analyzing probabilistic arguments, we derive the expressions for availability in combined cycle as well as for single. The expression for combined cycle are:

$$AH_0^1(t) = MH_0(t)+Q_{01}(t)©AH_1^1(t) + Q_{02}(t)©AH_2^1(t)+Q_{03}(t)©AH_3^1(t) \quad (8)$$

$$AH_1^1(t) = Q_{10}(t)©AH_0^1(t) + Q_{15}(t)©AH_5^1(t)+Q_{16}(t)©AH_6^1(t) \quad (9)$$

$$AH_2^1(t) = Q_{20}(t)©AH_0^1(t) + Q_{23}^{(4)}(t)©AH_3^1(t) \quad (10)$$

$$AH_3^1(t) = Q_{30}(t)©AH_0^1(t) \quad (11)$$

$$AH_5^1(t) = Q_{51}(t)©AH_1^1(t) + Q_{56}^{(7)}(t)©AH_6^1(t) \quad (12)$$

$$AH_6^1(t) = Q_{61}(t)©AH_1^1(t) \quad (13)$$

Solving them we get,

$$AH_0^1(t) = \lim_{s\to 0} sAH_0^{1*}(s)= \frac{U_1}{V} \quad (14)$$

Where, $MH_0(t)= e^{-(\alpha_1+\lambda_1)t}\bar{H}_1(t)$ and $U_1= \mu_0 p_{10}$

$V= \mu_0 p_{10}+\mu_1 p_{01}+\mu_2^{'}(p_{02}p_{10}+p_{01}p_{15})+\mu_3(p_{03}p_{10}+p_{02}p_{10}p_{23}^{(4)}+p_{01}p_{16}+p_{01}p_{15}p_{56}^{(7)})$

Similarly, we derive expressions for single cycle availability as:

$$AH_0^{1s}(t) = Q_{01}(t)©AH_1^{1s}(t) + Q_{02}(t)©AH_2^{1s}(t)+Q_{03}(t)©AH_3^{1s}(t) \quad (15)$$

$$AH_1^{1s}(t) = Q_{10}(t)©AH_0^{1s}(t) + Q_{15}(t)©AH_5^{1s}(t)+Q_{16}(t)©AH_6^{1s}(t) \quad (16)$$

$$AH_2^{1s}(t) = MH_2(t)+Q_{20}(t)©AH_0^{1s}(t) + Q_{23}^{(4)}(t)©AH_3^{1s}(t) \quad (17)$$

$$AH_3^{1s}(t) = Q_{30}(t)©AH_0^{1s}(t) \quad (18)$$

$$AH_5^{1s}(t) = Q_{51}(t)©AH_1^{1s}(t) + Q_{56}^{(7)}(t)©AH_6^{1s}(t) \quad (19)$$

$$AH_6^{1s}(t) = Q_{61}(t)©AH_1^{1s}(t) \quad (20)$$

Solving them we get,

$$AH_0^{1s}(t) = \lim_{s\to 0} sAH_0^{1s^*}(s)= \frac{U_2}{V} \quad (21)$$

Where, $MH_2(t)= e^{-\lambda_1 t}\bar{G}_2(t)$, $U_2= \mu_2 p_{10}p_{02}$ and V = as defined above

6. Availability of the System when Humidity is > 50%

$AH_i^2(t)$/ $AH_i^{2s}(t)$ denotes the probability that is available in full capacity/single cycle when humidity is > 50% at any instant of time t provided it has entered regenerative state i at time t=0. By analyzing probabilistic arguments, we subsequently derive the expression for availability in combined as well as for single cycle. The following are expression for combined cycle:

$$AH_0^2(t) = Q_{01}(t)©AH_1^2(t) + Q_{02}(t)©AH_2^2(t)+Q_{03}(t)©AH_3^2(t) \quad (22)$$

$$AH_1^2(t) = MH_1(t)+Q_{10}(t)©AH_0^2(t) + Q_{15}(t)©AH_5^2(t)+Q_{16}(t)©AH_6^2(t) \quad (23)$$

$$AH_2^2(t) = Q_{20}(t)©AH_0^2(t) + Q_{23}^{(4)}(t)©AH_3^2(t) \quad (24)$$

$$AH_3^2(t) = Q_{30}(t)©AH_0^2(t) \quad (25)$$

$$AH_5^2(t) = Q_{51}(t)©AH_1^2(t) + Q_{56}^{(7)}(t)©AH_6^2(t) \quad (26)$$

$$AH_6^2(t) = Q_{61}(t)©AH_1^2(t) \quad (27)$$

Solving them we get,

$AH_0^2(t) = \lim_{s\to 0} sAH_0^{2*}(s) = \frac{U_3}{V}$ (28)

Where, $MH_1(t) = e^{-(\alpha_2+\lambda_2)t}\bar{H}_2(t)$, $U_3 = \mu_1 p_{01}$ and V = already defined

Similarly, we derive the recursive relations for availability in single cycle as:

$AH_0^{2s}(t) = Q_{01}(t)©AH_1^{2s}(t) + Q_{02}(t)©AH_2^{2s}(t) + Q_{03}(t)©AH_3^{2s}(t)$ (29)

$AH_1^{2s}(t) = Q_{10}(t)©AH_0^{2s}(t) + Q_{15}(t)©AH_5^{2s}(t) + Q_{16}(t)©AH_6^{2s}(t)$ (30)

$AH_2^{2s}(t) = Q_{20}(t)©AH_0^{2s}(t) + Q_{23}^{(4)}(t)©AH_3^{2s}(t)$ (31)

$AH_3^{2s}(t) = Q_{30}(t)©AH_0^{2s}(t)$ (32)

$AH_5^{2s}(t) = MH_5(t) + Q_{51}(t)©AH_1^{2s}(t) + Q_{56}^{(7)}(t)©AH_6^{2s}(t)$ (33)

$AH_6^{2s}(t) = Q_{61}(t)©AH_1^{2s}(t)$ (34)

Solving them we get,

$AH_0^{2s}(t) = \lim_{s\to 0} sAH_0^{2s*}(s) = \frac{U_4}{V}$ (35)

Where, $MH_5(t) = e^{-\lambda_2 t}\bar{G}_2(t)$, $U_4 = \mu_5 p_{15} p_{01}$ and V = already defined

7. System's Expected Down Time

$DH_i^1(t)/DH_i^2(t)$ denotes that the system is in down state at specific instant of time t when humidity is ≤ /> 50%. The recursive expressions for $DH_i^1(t)/DH_i^2(t)$ are given below:

$DH_0^1(t) = Q_{01}(t)©DH_1^1(t) + Q_{02}(t)©DH_2^1(t) + Q_{03}(t)©DH_3^1(t)$ (36)

$DH_1^1(t) = Q_{10}(t)©DH_0^1(t) + Q_{15}(t)©DH_5^1(t) + Q_{16}(t)©DH_6^1(t)$ (37)

$DH_2^1(t) = Q_{20}(t)©DH_0^1(t) + Q_{23}^{(4)}(t)©DH_3^1(t)$ (38)

$DH_3^1(t) = NH_3(t) + Q_{30}(t)©DH_0^1(t)$ (39)

$DH_5^1(t) = Q_{51}(t)©DH_1^1(t) + Q_{56}^{(7)}(t)©DH_6^1(t)$ (40)

$DH_6^1(t) = Q_{61}(t)©DH_1^1(t)$ (41)

$DH_0^1(t) = \lim_{s\to 0} sDH_0^{1*}(s) = \frac{U_5}{V}$ (42)

Where, $NH_3(t) = \bar{G}_1(t)$, $U_5 = \mu_3 p_{10}(p_{03} + p_{02}p_{23}^{(4)})$ and V = already defined

$DH_0^2(t) = Q_{01}(t)©DH_1^2(t) + Q_{02}(t)©DH_2^2(t) + Q_{03}(t)©DH_3^2(t)$ (43)

$DH_1^2(t) = Q_{10}(t)©DH_0^2(t) + Q_{15}(t)©DH_5^2(t) + Q_{16}(t)©DH_6^2(t)$ (44)

$DH_2^2(t) = Q_{20}(t)©DH_0^2(t) + Q_{23}^{(4)}(t)©DH_3^2(t)$ (45)

$DH_3^2(t) = Q_{30}(t)©DH_0^2(t)$ (46)

$DH_5^2(t) = Q_{51}(t)©DH_1^2(t) + Q_{56}^{(7)}(t)©DH_6^2(t)$ (47)

$DH_6^2(t) = NH_6(t) + Q_{61}(t)©DH_1^2(t)$ (48)

$DH_0^2(t) = \lim_{s\to 0} sDH_0^{2*}(s) = \frac{U_6}{V}$

Where, $NH_6(t) = \bar{G}_1(t)$, $U_5 = \mu_3 p_{01}(p_{16} + p_{15}p_{56}^{(7)})$ and V = already defined

8. Time Period for which Repairman is Busy

$BH_i^1(t)/BH_i^2(t)$ denotes the probability that repairman is busy at an instant t when humidity is ≤/> 50%. By analyzing probabilistic arguments, the recursive expressions for $BH_i^1(t)/BH_i^2(t)$ are given below:

$BH_0^1(t) = Q_{01}(t)©BH_1^1(t) + Q_{02}(t)©BH_2^1(t) + Q_{03}(t)©BH_3^1(t)$ (49)

$BH_1^1(t) = Q_{10}(t)©BH_0^1(t) + Q_{15}(t)©BH_5^1(t) + Q_{16}(t)©BH_6^1(t)$ (50)

$BH_2^1(t) = RH_2(t) + Q_{20}(t)©BH_0^1(t) + Q_{23}^{(4)}(t)©BH_3^1(t)$ (51)

$BH_3^1(t) = RH_3(t) + Q_{30}(t)©BH_0^1(t)$ (52)

$BH_5^1(t) = Q_{51}(t)©BH_1^1(t) + Q_{56}^{(7)}(t)©BH_6^1(t)$ (53)

$BH_6^1(t) = Q_{61}(t)©BH_1^1(t)$ (54)

$BH_0^1(t) = \lim_{s\to 0} sBH_0^{1*}(s) = \frac{U_6}{V}$ (55)

Where, $RH_2(t) = e^{-\lambda_1 t}\overline{G}_2(t)$, $RH_3(t) = \overline{G}_1(t)$,

$U_6 = p_{10}[p_{02}\mu_2 + (p_{03} + p_{02}p_{23}^{(4)})\mu_3]$ and V = already defined

$BH_0^2(t) = Q_{01}(t)©BH_1^2(t) + Q_{02}(t)©BH_2^2(t) + Q_{03}(t)©BH_3^2(t)$ (56)

$BH_1^2(t) = Q_{10}(t)©BH_0^2(t) + Q_{15}(t)©BH_5^2(t) + Q_{16}(t)©BH_6^2(t)$ (57)

$BH_2^2(t) = Q_{20}(t)©BH_0^2(t) + Q_{23}^{(4)}(t)©BH_3^2(t)$ (58)

$BH_3^2(t) = Q_{30}(t)©BH_0^2(t)$ (59)

$BH_5^2(t) = RH_5(t) + Q_{51}(t)©BH_1^2(t) + Q_{56}^{(7)}(t)©BH_6^2(t)$ (60)

$BH_6^2(t) = RH_6(t) + Q_{61}(t)©BH_1^2(t)$ (61)

$BH_0^2(t) = \lim_{s\to 0} sBH_0^{2*}(s) = \frac{U_7}{V}$ (62)

Where, $RH_5(t) = e^{-\lambda_2 t}\overline{G}_2(t)$, $RH_6(t) = \overline{G}_1(t)$

$U_7 = p_{01}p_{15}\mu_5 + p_{01}\mu_6(p_{16} + p_{15}p_{56}^{(7)})$ and V = already defined

9. Number of Visits to be Expected by Repairman

$VH_i^1(t)/VH_i^2(t)$ denotes the expected number of visits by the repairman when humidity is ≤/> 50%. By analyzing probabilistic arguments, the recursive expressions for $VH_i^1(t)/VH_i^2(t)$ are given below:

$VH_0^1(t) = Q_{01}(t) Ⓢ VH_1^1(t) + Q_{02}(t) Ⓢ [1+VH_2^1(t)] + Q_{03}(t) Ⓢ [1+VH_3^1(t)]$ (63)

$VH_1^1(t) = Q_{10}(t) Ⓢ VH_0^1(t) + Q_{15}(t) Ⓢ VH_5^1(t) + Q_{16}(t) Ⓢ VH_6^1(t)$ (64)

$VH_2^1(t) = Q_{20}(t) Ⓢ VH_0^1(t) + Q_{23}^{(4)}(t) Ⓢ VH_3^1(t)$ (65)

$VH_3^1(t) = Q_{30}(t) Ⓢ VH_0^1(t)$ (66)

$VH_5^1(t) = Q_{51}(t) Ⓢ VH_1^1(t) + Q_{56}^{(7)}(t) Ⓢ VH_6^1(t)$ (67)

$VH_6^1(t) = Q_{61}(t) Ⓢ VH_1^1(t)$ (68)

$VH_0^1(t) = \lim_{s\to 0} sVH_0^{1*}(s) = \frac{U_8}{V}$ (69)

Where, $U_8 = p_{10}(p_{02} + p_{03})$ and V = already defined

$VH_0^2(t) = Q_{01}(t) Ⓢ VH_1^2(t) + Q_{02}(t) Ⓢ VH_2^2(t) + Q_{03}(t) Ⓢ VH_3^2(t)$ (70)

$VH_1^2(t) = Q_{10}(t) Ⓢ VH_0^2(t) + Q_{15}(t) Ⓢ [1+VH_5^2(t)] + Q_{16}(t) Ⓢ [1+VH_6^2(t)]$ (71)

$VH_2^2(t) = Q_{20}(t) Ⓢ VH_0^2(t) + Q_{23}^{(4)}(t) Ⓢ VH_3^2(t)$ (72)

$VH_3^2(t) = Q_{30}(t) Ⓢ VH_0^2(t)$ (73)

$VH_5^2(t) = Q_{51}(t) Ⓢ VH_1^2(t) + Q_{56}^{(7)}(t) Ⓢ VH_6^2(t)$ (74)

$VH_6^2(t) = Q_{61}(t) Ⓢ VH_1^2(t)$ (75)

$VH_0^2(t) = \lim_{s\to 0} sVH_0^{2*}(s) = \frac{U_9}{V}$ (76)

Where, $U_9 = p_{01}(p_{15} + p_{16})$ and V = already defined

10. Profit Analysis of the System

Expected profit expressions induced per unit time in steady state for the system are:

$P = CA_1 * AH_0^1 + CA_2 * AH_0^2 + CA_{1s} * AH_0^{1s} + CA_{2s} * AH_0^{2s} - CB_1 * BH_0^1 - CB_2 * BH_0^2 - CV_1 * VH_0^1 - CV_2 * VH_0^2 - C_0$

CA_1/CA_2 : Revenue generated per unit uptime when humidity is ≤ 50%, system works in Combined Cycle/Single Cycle

CA_{1s}/CA_{2s} : Revenue generated per unit uptime when humidity is > 50%, system works in Combined Cycle/Single Cycle

CB_1/CB_2 : Cost per unit time when humidity is ≤/> 50% while repairman is busy in doing repair

CV_1/CV_2 : Cost per visit by repairman when humidity is ≤/> 50%

C_0 : Other expenses in plant operation

11. Results and Graphical Representation of Reliability Measures

Now, by considering the particular cases $g_1(t)=\beta_1 e^{-\beta_1 t}$, $g_2(t)=\beta_2 e^{-\beta_2 t}$, $h_1(t)=\gamma_1 e^{-\gamma_1 t}$, $h_2(t)=\gamma_2 e^{-\gamma_2 t}$ and using information collected from gas turbine power plant we get $\alpha_1=\lambda_1=0.001$, $\alpha_2=\lambda_2=0.0014$, $\beta_1=\beta_2=0.028$, $\gamma_1=0.333$, $\gamma_2=0.317$, $CA_1 = 1372000$, $CA_2 = 1250000$, $CA_{1s} = 882000$, $CA_{2s} = 800000$, $CB_1 = 12000$, $CB_2 = 20500$, $CV_1 = 8000$, $CV_2 = 12500$, $C_0 = 650000$ and plotted different graphs according to these esteemed values and studied various measures of reliability through graphs as explained below:

11.1. Mean Time to System Failure Vs Different Failure Rates

Figure 2 illustrates the behavior of MTSF with different failure rates α_1, α_2,λ_1,λ_2.

- MTSF decreases with increase in any one of the failure rates.
- MTSF increases with increase in repair rates and when both repair rates (β_1, β_2) are increased by same amount MTSF is higher in case of β_2 as compared to β_1.

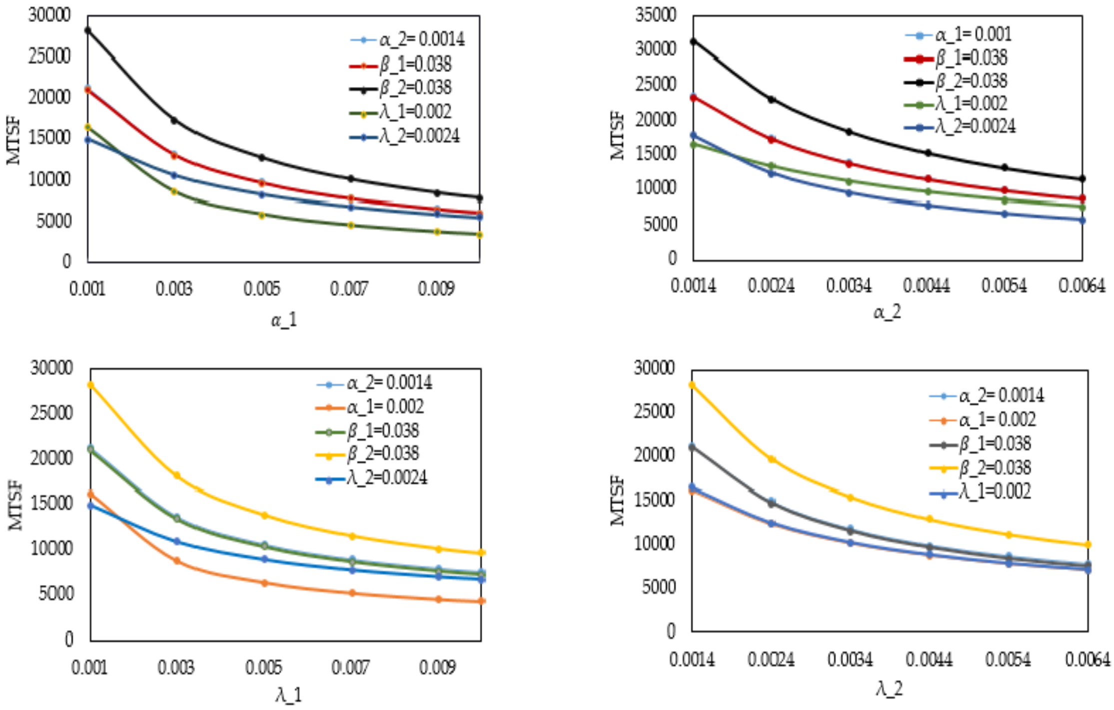


Figure 2: *MTSF Vs Failure rates* α_1, α_2,λ_1,λ_2

11.2. Availability in Combined Cycle Vs Failure Rate of Steam Turbine when Humidity is ≤ 50%

Figure 3 demonstrates the availability in combined cycle when humidity is ≤ 50% and when humidity is > 50%

- Both availabilities (when humidity is ≤/> 50%) of combined cycle decreases as we increase in any one of failure rate.
- Both availabilities (when humidity is ≤/> 50%) in combined cycle increases with increase in repair rates.
- Availability in Combined Cycle when humidity is > 50% is higher than Availability in Combined Cycle when Humidity is ≤ 50%.

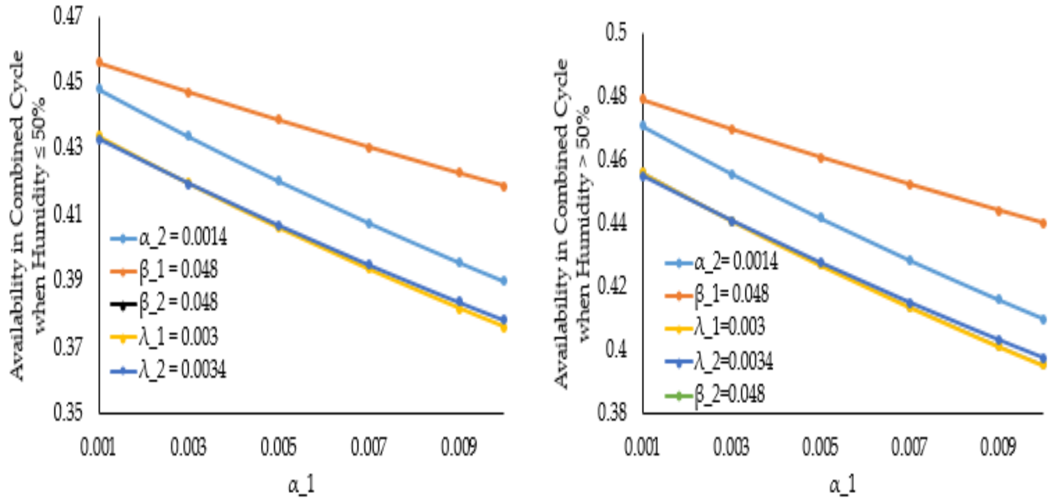


Figure 3: *Availability in combined cycle Vs Failure rate* $\propto_1$

11.3. Availability in Single Cycle Vs Failure Rate of Steam Turbine when Humidity is ≤ 50%

Figure 4 demonstrates the availability in single cycle when humidity is ≤ 50% and when humidity is > 50%

- Availability in single cycle when humidity is ≤ 50% increases with increase in α_1 but decreases with increase in any other failure rate while availability in single cycle when humidity is > 50% decreases with increase in α_1, λ_1, λ_2 but increases with increase in α_2.
- Availability in single cycle when humidity is ≤ 50% decrease with increase in β_1 while availability in single cycle when humidity is > 50% increases with increase in β_1.

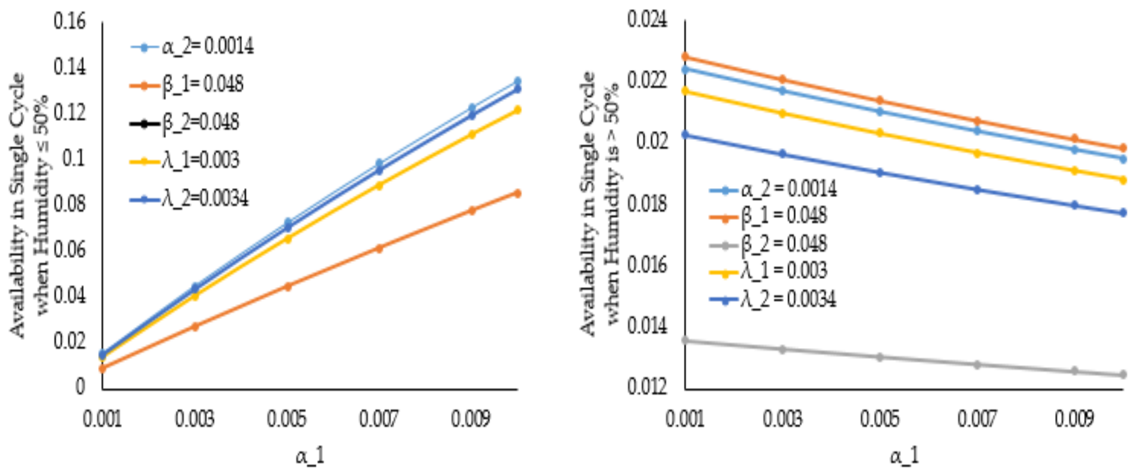


Figure 4: *Availability in single cycle Vs Failure rate* $\propto_1$

11.4. Profit Vs Failure Rate of Steam Turbine when Humidity is ≤ 50%

Figure.5 illustrates the behavior of Profit of plant with respect to failure rate α_1.

- Profit decreases with increase in any one failure rates $\alpha_1, \alpha_2, \lambda_1, \lambda_2$.

- Profit increases as we increase in repair rates and when both repair rates (β_1, β_2) are increased by same amount, Profit in case of β_2 is higher than Profit in case of β_1 (when α_1 = 0.001 to 0.006) and Profit in case of β_1 is higher than Profit in case of β_2 (when α_1 = 0.006 to 0.01).

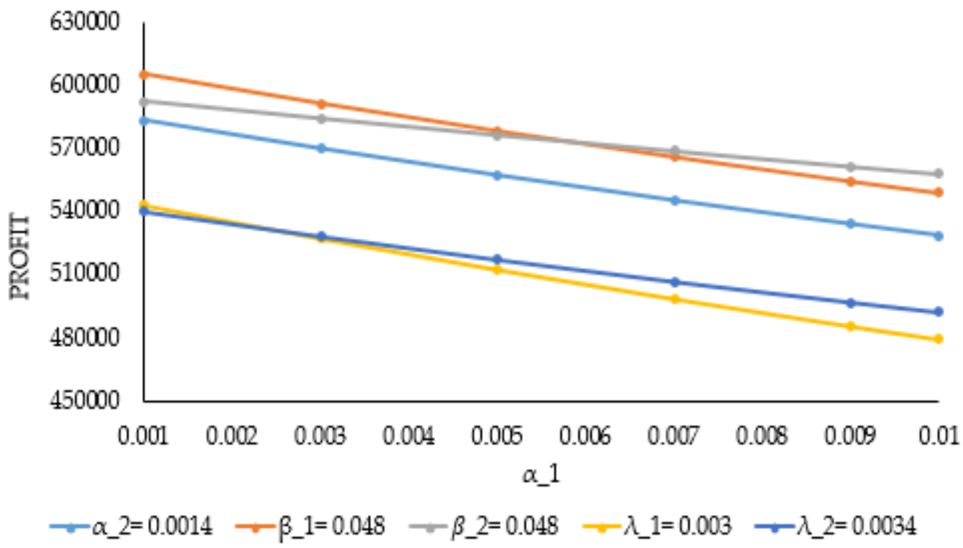


Figure 5: *Profit Vs Failure rate* $\propto_1$

12. Conclusion

A stochastic model for a two-unit gas turbine system has been developed by fixing the range of temperature and using the idea of different humidity conditions. Various reliability measures like mean time to system failure, availability for combined as well as for single cycle when humidity is ≤/> 50% have been obtained for particular cases using information gathered from gas turbine power plant. Simultaneous effects of failure rates of gas and steam turbines when humidity is ≤/> 50% on mean time to system failure have been graphically analyzed and from them we concluded that mean time to system failure decreases as any failure rate increases. Trends in the availability for both cycles and different humidity conditions i.e. when humidity is ≤/> 50% has been illustrated with respect to failure rate of steam turbine and various interesting results have been obtained regarding availability. At last, profit for plant is also depicted which decreases with increase in failure rates. Here we see that for this fixed range of temperature availability is higher when humidity is >50% as compared to when humidity is ≤50% which further impacts on profit of the plant. Furthermore, a comprehensive examination of gas turbine system may be helpful to those who are involved in power generating industry.

References

[1] Baohe, Su (1997). On a two-dissimilar-unit system with three modes and random check. Microelectron Reliab, 37:1233-1238.
[2] Chow, D.K. (1973). Reliability of some redundant systems with repair. IEEE Transactions on Reliability, 22:223-228.
[3] Tuteja, R.K., Arora R.T. and Taneja, G. (1991). Analysis of a two-unit system with partial failures and three types of repairs. Reliability Engineering and System Safety, 33:199-214.

[4] El-Berry, A. (2020). Reliability analysis of gas turbine power plant based on failure data. International Journal of Mechanical and Mechatronics Engineering, 2:13-25.

[5] Singh, D. and Taneja, G. (2012). Reliability analysis of a power generating system through gas and steam turbines with scheduled inspection. Aryabhatta Journal of Mathematics and Informatics, 5:373-380.

[6] Singh, D. and Taneja, G. (2014). Reliability and economic analysis of a power generating system comprising one gas and one steam turbine with random inspection. International Journal of Mathematics and Statistics, 4:30-37.

[7] Farouk, N., Sheng, L. and Hayat, Q. (2013). Effects of ambient temperature on performance of gas turbines power plant. International Journal of Computer Science, 10:30-37.

[8] Abigail, G.D., Agustin, M.A.C., Maria, O.G.D., Angel, M.A., Mathieu, L. and Jose, M.G.S., (2017). Effect of the ambient conditions on gas turbine combined cycle power plants with post-combustion CO_2 capture. Elsevier Ltd., 134:221-233.

[9] Saleh, S., Baakeem, J.O. and Hany, A. (2015). Performance of a typical simple gas turbine unit under Saudi weather conditions. International Journal of Energy and Power Engineering, 1:59-71.

[10] Rajesh, Taneja G. and Prasad, J. (2018). Reliability of a gas turbine system with change in weather and optimization of electricity price when working in single cycle. International Journal of Agricultural and Statistical Sciences. 14:119-128.

[11] Rajesh, Taneja G. and Prasad, J. (2018). Reliability and availability analysis for a three-unit gas turbine power generating system with seasonal effect and FCFS repair pattern. International Journal of Applied Engineering Research. 13:10948-10964.

[12] Fernandez, D.A.P., Foliaco, B., Padilla, R.V., Bula, A. and Quiroga, A.G. (2021). High ambient temperature effects on the performance of a gas turbine-based cogeneration system with supplementary fire in a tropical climate. Case Studies in Thermal Engineering, 26:101206-101219.

[13] Bird, J. and Grabe, W. (1991). Humidity effects on gas turbine performance. The American Society of Mechanical Engineers, 329.

[14] Hanachi, H., Liu, J., Banerjee, A. and Chen, Y., (2015). Effects of the intake air humidity on gas turbine performance monitoring. Proceedings of ASME Turbo 6.