A NOVEL EXTENSION OF INVERSE EXPONENTIAL DISTRIBUTIONS: A HEAVY-TAILED MODEL WITH UPSIDE DOWN BATHTUB SHAPED HAZARD RATE

Jabir Bengalath

•

Department of Statistics, Govt. Arts and Science College, Calicut, University of Calicut, Kerala, India.
jabirbengu@gmail.com

Bindu Punathumparambath

•

Department of Statistics, Govt. Arts and Science College, Calicut, University of Calicut, Kerala, India.
ppbindukannan@gmail.com

Abstract

Heavy-tailed distributions have garnered interest due to their advantageous statistical and reliability characteristics, rendering them valuable in applied fields such as economics, finance, and risk management. Such distributions offer robust properties, making them pertinent to studies in various areas like econometrics, statistics, and insurance. Thus, the primary objective of this paper is to propose a Two parameter right skewed- upside down bathtub type, heavy tailed distribution, which is a generalisation of Inverse Exponential distribution and is referred to as Modi Inverse Exponential distribution. We derive several mathematical and statistical features, including quantile function, mode, median, skewness, kurtosis, and mean deviation. Additionally, the reliability and hazard rate functions are also derived. Stochastic ordering and order statistics of the proposed distribution were derived. We also investigate the tail behaviour of the proposed model. Furthermore, estimation methods such as maximum likelihood estimation and its asymptotic confidence bound, percentile method, and Cramer-von-Mises method were examined. To demonstrate the appropriateness of the suggested model, we have considered two distinct real datasets along with three distinct models and concluded that the proposed model is more adaptable.

Keywords: Inverse Exponential distribution, Modi Inverse Exponential distribution, Moments, Tail Behaviour, Order Statistics, Parameter estimation.

1. Introduction

During the recent years, heavy-tailed distributions have gained attention as an attractive subject for various research and studies. References to some works on these distributions can be found in [1],[2],[3],[4],[5]. These distributions possess excellent statistical and reliability properties, making them practical for many applied sciences such as economics, finance, econometrics, statistics, risk management, and insurance. Several authors have developed inferential results under financial modeling, as seen in [6],[7],[8],[9],[10]. There exist various heavy-tailed distributions in many practical situations, such as financial sciences, reliability engineering and bio-medical science,data are usually positive, and their distribution is uni-modal hump shaped and extreme values yielding heavier tails than the classical models. For example, in health science research (1). The medical expenditure that exceed a given threshold and (2). Length of stay in hospital, presents highly skewed, heavy tailed data for which standard classical distributions and simple variable transform are insufficient to provide an adequate fit to such data. The Exponential distribution has been widely used for analyzing life-time data. However, its usefulness is restricted to scenarios with a constant hazard rate, which can be difficult to justify in practical situations. To address this problem, so many alternative models for life-time data have been developed. Among them, distributions like Weibull and gamma have been extensively employed when dealing with life-time data exhibiting a monotonically increasing or decreasing hazard rate. The UBT failure rate distributions commonly appear in medical and biological fields like in lung cancer patient data (see [11], in bladder cancer patient data (see [28]) and in breast carcinoma patient data (see [12]. The inverse transform method is a widely used approach to derive the inverse form of different lifetime distributions. The distribution family obtained by this method, known as the generalized inverted family, often exhibits the characteristic"upside-doun bathtub" hazard rate pattern. These distributions have the advantage of being the number of parameters required and are straightforward to apply. For instance, Notable examples of such a distribution include the inverted gamma distribution (IED) proposed by Lin et al[13] and the Inverse Lindley Distribution (ILD), introduced by Sharma et al[14]. In addition,The transmuted inverse exponential distribution was presented by Oguntunde and Adejumo [15]. In the same year, Khan et al [16] propose the transmuted inverse Weibull distribution. In addition, in 2014, Sharma et al [17]introduce the transmuted inverse Rayleigh distribution. Then, in 2016, Sharma et al [18] further introduced the generalized inverse Lindley distribution. The purpose of this study is to propose a new inverted probability model with UBT type of failure rate. For this purpose , we consider A one parameter Inverse exponential distribution. Let X be the random variable having the probability density function (pdf) is given by

$$g(x) = \frac{\theta}{x^2} e^{-\frac{\theta}{x}}, \quad x > 0, \quad \theta > 0. \tag{1}$$

and the cumulative distribution function (cdf) is

$$G(x) = e^{-\frac{\theta}{x}}, \quad x > 0, \quad \theta > 0. \tag{2}$$

for all $x > 0$, where $\theta > 0$. Furthermore, K. Modi et al [19] introduced a new family of distributions name Modi family of distribution in his paper titled a new family of distributions with applications on two real data sets on the survival problem. where he proposed a new probability distribution by taking base line distribution by exponential distribution with one parameter. This paper aims to substitute α^{β} with γ within the established family, resulting in a modified probability density function(pdf) and cumulative distribution function(cdf) is given by.

$$f_Y(x) = \frac{(1+\gamma)(\gamma g(x))}{(\gamma + G(x))^2}, \quad \gamma > 0. \tag{3}$$

and

$$F_Y(x) = \frac{(1+\gamma)G(x)}{\gamma + G(x)}, \quad \gamma > 0. \tag{4}$$

where $g(x)$ and $G(x)$ are the pdf and cdf of the baseline distribution. Moreover, We can easily verified that the given Modi family satisfies the identifiable properties and other properties which is required for a probability distribution. Hence, this familiy of distribution can be used to generate more flexible probability distributions. The hazard function of equation (1) is given by,

$$h_Y(x) = \left[\frac{1+\gamma}{\gamma + G(x)}\right] h_X(x), \quad \gamma > 0.$$

Where $h_X(x)$ is the hazard function of the baseline distribution.

Acknowledging the need for more flexible lifetime distributions, we introduce a new family of probability distribution known as the Modi Inverse Exponential distribution with two parameters. which can be extensively used to fit and analyze data in a variety of field. The paper is organized as follows: Section 2 specifies the Modi Inverse Exponential distribution, whereas Section 3 provides the Modi Inverse Exponential distribution Properties, which include the Hazard function, Survival function, Quantile function,Mode, Median, Skewness, Kurtosis, Mean Deviation, Stochastic Ordering, Order statistics. In Section 4, we look at the Modi inverted exponential distribution's tail behaviour. Section 5 We investigated different method of estimation which includes Maximum likelihood estimation and its asymptotic confidence bound, Percentile method and Cramer-von Mises method is discussed Section 5, we conduct a simulation study to validate the proposed model's estimations, and two real data sets are analysed to demonstrate the efficacy of the proposed model. The conclusion is provided in Section 6.

2. Modi Inverse Exponential Distribution

A random variable X is said to have Modi Inverse Exponential distribution (MIE) if its cumulative distribution function (cdf) is given by

$$F(x) = \frac{(\gamma+1)}{1+\gamma e^{\frac{\theta}{x}}} \quad , \quad x>0, \quad \gamma>0, \quad \theta>0. \tag{5}$$

and its pdf is,

$$f(x) = \frac{\theta\gamma(\gamma+1)e^{\frac{\theta}{x}}}{(x\gamma e^{\frac{\theta}{x}}+x)^2} \quad , \quad x>0, \quad \gamma>0, \quad \theta>0. \tag{6}$$

The shape of the distribution might provide important insights into its characteristics, such as whether it is symmetrical or skewed. In this context, the MIE(γ,θ) distribution is represented by its cumulative distribution function (cdf) in Figure 2 and its probability density function (pdf) in Figure 1 for different values parameter.

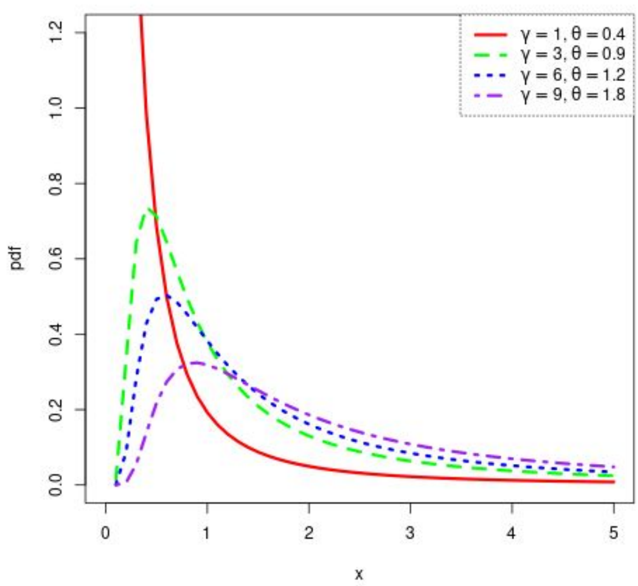

Figure 1: *Pdf plot of MIE distribution for different parameter values*

Theorem 1. Given that X follows the MIE (γ,θ) distribution with $f(x)$ and $F(x)$ as given in (6) and (5) respectively, then:

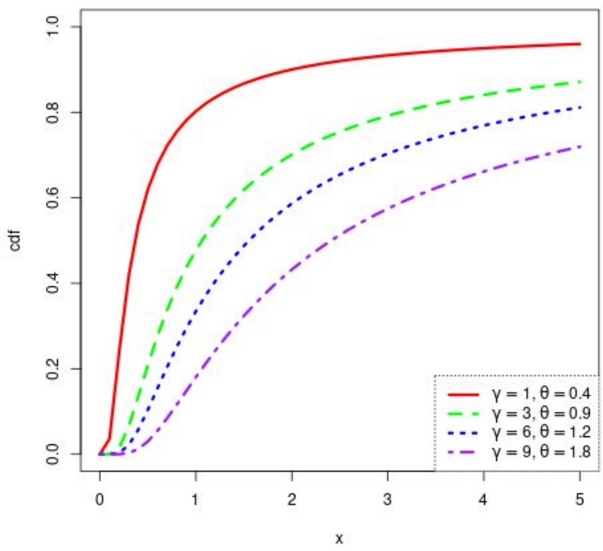

Figure 2: *cdf plot of MIE distribution for different parameter values*

1. $\lim_{x\to\infty} f(x) = 0$
2. $\lim_{x\to\infty} F(x) = 1$

proof. Trivial and hence omitted.

3. Some Statistical Properties of MIE Distribution

This section contains various statistical features associated with the new distribution that have been derived.

3.1. Hazard Rate Function

The hazard function characterizing a specific phenomenon elucidates the inherent nature of the failure rate that is associated with lifetime of the specific equipment. For the cdf and pdf provided in equations (5) and (6), respectively, the expression for $h(x)$ is as follows:

$$h(x) = \frac{(\gamma+1)\theta e^{\frac{\theta}{x}}}{x^2(e^{\frac{\theta}{x}}-1)(\gamma e^{\frac{\theta}{x}}+1)} \tag{7}$$

The hazard rate function plot in Figure 3 shows various curves indicating different values of the parameters γ and θ. We can gain useful insights about the nature of the model's failure rate by using this visual depiction, which reveals a distinct right-skewed pattern and UBT type failure model.The mathematical verification of this assertion may also be established through the utilization of the outcome presented by Glaser [20]. Glaser demonstrated that the Condition for UBT can be established if and only if the following conditions are met: $\varphi'(t) > 0$ for all $t \in (0, t_0)$, $\varphi'(t_0) = 0$, $\varphi'(t) < 0$ for all $t > t_0$, and satisfying $\lim_{t\to 0} f(t) = 0$ where φ is equal to $-\frac{f'(t)}{f(t)}$ and $f(t)$ is the first derivative of the density function with respect to t. For our proposed model, it is evident that.

$$\varphi(t) = -\frac{2t(1+e^{\frac{\theta}{t}}\gamma)+\theta-e^{\frac{\theta}{t}}\gamma\theta}{t^2(1+e^{\frac{\theta}{t}}\gamma)}$$

and

$$\varphi'(t) = \frac{2((t+e^{\frac{\theta}{t}}t\gamma)^2 - e^{\frac{\theta}{t}}\gamma\theta^2 + t(\theta - e^{\frac{2\theta}{t}}\gamma^2\theta)}{t^4(1+e^{\frac{\theta}{t}}\gamma)^2}$$

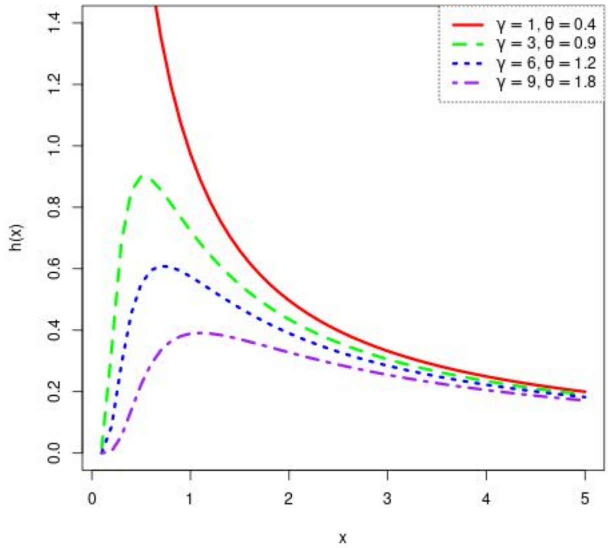

Figure 3: *Hazard plot of MIE distribution for different parameter values*

Since, the above equation is not provided in an explicit form to derive the solution, a simulation study was performed.It was observed that for $t_0 \approx 0.4488$, $\varphi'(t) > 0$ for all $t \in (0, t_0)$, $\varphi'(t_0) = 0$, $\varphi'(t) < 0$ for all $t > t_0$. Also, from Equation (6), we verified that $\lim_{x\to 0} f(t) = 0$. Therefore, it can be deduced that the MIE(γ,θ) distribution proposed exhibits a right-skewed distribution, which is characterized by an UBT shape of hazard rate. This distribution is particularly useful when analyzing medical and reliability data.

3.2. Survival Function

The survival function describes the probability that a unit, component, or individual will not fail at a given time. The expression for survival function $S(x)$ is stated as follows, and its corresponding survival plot is presented in Fiqure 4:

$$S(x) = \frac{\gamma(e^{\frac{\theta}{x}} - 1)}{\gamma e^{\frac{\theta}{x}} + 1}, \tag{8}$$

Theorem 2. The limit of the hazard rate function of MIE(γ, θ) distribution as $x \to \infty$ is zero.

$$i.e, \lim_{x\to\infty} h(x) = 0.$$

Proof. Trivial and hence omitted. ■

3.3. The Odd Function

The Odd Function is obtained using the relation $Q(x) = \frac{F(x)}{s(x)}$ and is given by

$$Q(x) = \frac{\gamma + 1}{\gamma(e^{\frac{\theta}{x}} - 1)}, \tag{9}$$

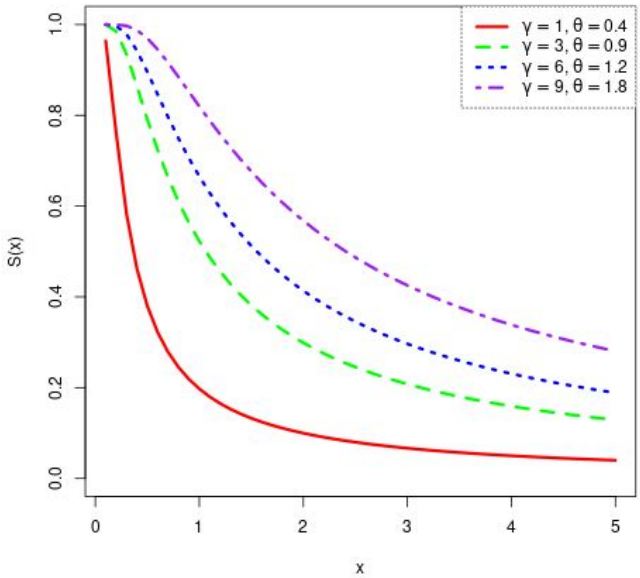

Figure 4: *Survival plot of MIE distribution for different parameter values*

3.4. Reverse Hazard Rate Function

The Revised Hazard Rate Function is obtained by using the relation,

$$\begin{aligned}\phi(x) &= \frac{f(x)}{F(x)}\\ &= \frac{\theta\gamma e^{\frac{\theta}{x}}}{x^2(\gamma e^{\frac{\theta}{x}}+1)},\end{aligned} \tag{10}$$

3.5. Cumulative Hazard Function

The Cumulative Hazard Function is obtained using the relation,

$$\begin{aligned}C(x) &= -\log(S(x))\\ &= -\log\left[\frac{\gamma(e^{\frac{\theta}{x}}-1)}{\gamma e^{\frac{\theta}{x}}+1}\right],\end{aligned} \tag{11}$$

3.6. Quantile Function, Skewness and Kurtosis

The MIE distribution can be simulated using the inverse cdf method,

$$X = \left[\frac{\theta}{\ln\left(\frac{\gamma+1-u}{\gamma u}\right)}\right], \tag{12}$$

where, u is a uniform random variable, $0 < u < 1$. The q^{th} quantile of the MIE distribution is obtained as:

$$x_q = \left[\frac{\theta}{\ln\left(\frac{\gamma+1-q}{\gamma q}\right)}\right], \tag{13}$$

By making use of equation (13), we are able to calculate the first and third quartiles by substituting $q = 0.25$ and $q = 0.75$, correspondingly. Once we have obtained these values, we can subsequently

calculate Galton's [[21]] skewness (S_k) and Moor's [[22]] Kurtosis (K_r) by means of the given formulae.
The measure of skewness S_k,

$$S_k = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)}, \tag{14}$$

and the measure of Kurtosis, K_r

$$K_r = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)}, \tag{15}$$

3.7. Median

Since the distribution proposed is a heavily tailed, right-skewed distribution, the most appropriate measure of central tendency is the median.The median of the proposed distribution can be obtained by utilizing $q = 0.5$ in the quantile function, as delineated in equation (13). So median M_d

$$M_d = \left[\frac{\theta}{\ln\left(\frac{\gamma+1-0.5}{\gamma(0.5)}\right)} \right], \tag{16}$$

3.8. Mode

if a random variable X has the PDF given by equation (6), then the corresponding mode is given by $f'(x) = 0$, thus we obtain

$$f'(x) = \frac{e^{\frac{\theta}{x}} \theta\gamma(1+\gamma)\left(e^{\frac{\theta}{x}}(\theta - 2x) - \gamma(\theta + 2x)\right)}{\left(e^{\frac{\theta}{x}} + \gamma\right)^3 x^4} = 0$$

$$\implies \left(e^{\frac{\theta}{x}}(\theta - 2x) - \gamma(\theta + 2x)\right) = 0$$

For various values of γ and θ, we can estimate the value of x by using an optimization technique in R. If $\gamma = 3$ and $\theta = 4$, we obtain the mode as 1.93755.

3.9. Mean Deviation

The mean deviation from the median is a statistical measure, serves as an indicator of population dispersion. Let "M" stand in for the median of the MIE Distributions specified in equation (16). The mean deviation from the median may be computed as follows:

$$\rho(x) = E|x - M| = \int_0^\infty |x - M| f(x)\,\mathrm{dx},$$

it can be obtained the following equation $\rho = \mu - 2W(M)$ where W(M) is

$$W(M) = \theta\gamma(\gamma + 1)\int_0^M \frac{e^{\frac{\theta}{x}}}{x^2(\gamma e^{\frac{\theta}{x}} + 1)^2}\mathrm{dx}, \tag{17}$$

This integral may be readily computed numerically using tools such as R, MATLAB, Mathcad, and others. Thus, obtaining the mean deviation from the median is desired.

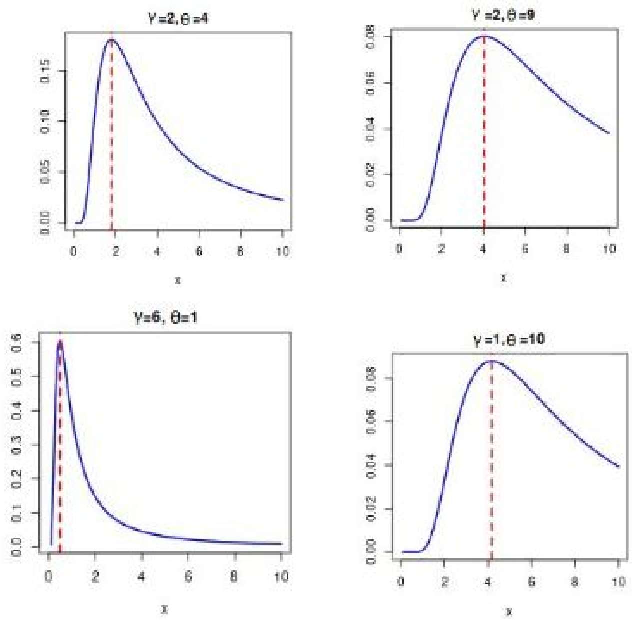

Figure 5: *Mode plot of MIE distribution for different parameter value*

3.10. Stochastic Ordering

Let X_1 and X_2 be random variables with cumulative distribution functions (cdf's) $F_1(x)$ and $F_2(x)$, respectively. X_1 is said to be stochastically greater than or equal to X_2 if $F_1(x) \leq F_2(x) \quad$ for all x. (see Gupta et al [23] for more detail).

Theorem 3. Let $X_1 \sim \text{MIE}(\gamma_1, \theta_1)$ and $X_2 \sim \text{MIE}(\gamma_2, \theta_2)$. X_1 is said to be stochastically greater than X_2 if $\gamma_1 = \gamma_2 = \gamma$ and $\theta_1 > \theta_2$.
Proof. Let's consider $\theta_1 > \theta_2$ and $\gamma_1 = \gamma_2 = \gamma$, the ratio simplifies to:

$$\frac{F_1(x)}{F_2(x)} = \frac{\frac{\gamma+1}{1+\gamma e^{\frac{\theta_1}{x}}}}{\frac{\gamma+1}{1+\gamma e^{\frac{\theta_2}{x}}}}$$

Since $\theta_1 > \theta_2$, we have:

$$\frac{e^{\frac{\theta_1}{x}}}{e^{\frac{\theta_2}{x}}} > 1$$

Therefore, when $\theta_1 > \theta_2$ and $\gamma_1 = \gamma_2 = \gamma$, $F_1(x)$ is stochastically smaller than $F_2(x)$ for all $x > 0$. ■

3.11. Order Statistics

In this section, we derive a compact expression for the pdf of the i^{th} order statistic of the Modi inverse exponential distribution. Let $X_1, X_2, X_3, ..., X_n$ be a simple random sample from the Modi inverse exponential distribution with cdf and pdf given by equations (5) and (6), respectively. Let $X_{1:n} \leq X_{2:n} \leq X_{3:n} \leq ... \leq X_{n:n}$ denote the order statistics obtained from this sample. We now give the pdf of $X_{r:n}$, denoted as $f_{r:n}(x)$, and the r^{th} moments of $X_{r:n}$, for $i = 1, 2, ..., n$, which are given by:

$$f_{r:n}(x) = C_{r:n} \left[F(x;\gamma,\theta)\right]^{r-1} \left[1 - F(x;\gamma,\theta)\right]^{n-r} f(x;\gamma,\theta), \tag{18}$$

for all $x > 0$, where F and f are given by equations (5) and (6), respectively, and $C_{r:n} = \frac{n!}{(n-1)!(n-r)!}$. Thus, using the binomial series expansion:

$$(1-x)^{\alpha-1} = \sum_{j=0}^{\infty}(-1)^j \binom{\alpha-1}{j} x^j.$$

We obtain:

$$\begin{aligned} f_{r:n}(x) &= C_{r:n}\sum_{s=0}^{\infty}(-1)^s \binom{n-r}{s} \left[F(x;\gamma,\theta)\right]^{r+s-1} f(x;\gamma,\theta), \\ &= C_{r:n}\sum_{s=0}^{\infty}(-1)^s \binom{n-r}{s} \left[\frac{\gamma+1}{1+\gamma e^{\frac{\theta}{x}}}\right]^{r+s-1} \frac{\theta\gamma(\gamma+1)e^{\frac{\theta}{x}}}{(x\gamma e^{\frac{\theta}{x}}+x)^2}, \\ &= C_{r:n}\sum_{s=0}^{\infty}(-1)^s \binom{n-r}{s} \theta\gamma(\gamma+1)^{r+s}(1+\gamma e^{\frac{\theta}{x}})^{r+s-1} e^{\frac{\theta}{x}}(x\gamma e^{\frac{\theta}{x}}+x)^{-2}. \end{aligned} \quad (19)$$

4. Tail Area Property

According to Klugman et al [24]and Nair et al [25] a distribution is classified as a heavy-tailed distribution when it displays the heavy tail property. These types of distributions are characterized by the lack of one or more orders of moments. Specifically, the absence of the first moment, which represents the distribution's arithmetic mean, indicates the presence of the distribution's heavy tail property. The proposed distribution's arithmetic mean can be derived by solving:

$$\theta\gamma(\gamma+1)\int_0^{\infty} \frac{e^{\frac{\theta}{x}}}{x(\gamma e^{\frac{\theta}{x}}+1)^2}\mathrm{dx},$$

which is a divergent integral,then the arithmetic mean of the corresponding distribution cannot be determined. Consequently, based on this criterion, the proposed distribution can be classified as a heavy-tailed distribution. Another method for evaluating the heavy tail attribute of the distribution is to examine whether the ratio of the hazard rate to x approaches zero as x approaches infinity; if it does, then the distribution displays the characteristic of a heavy-tailed distribution. For the proposed distribution:

$$\frac{(\gamma+1)\theta e^{\frac{\theta}{x}}}{x^3(e^{\frac{\theta}{x}}-1)(\gamma e^{\frac{\theta}{x}}+1)} \to 0$$

This fact can be proven by implementing L'Hpital's rule. As a result, the distribution put forward exhibits a heavy-tailed distribution.

In our analysis, we additionally consider the examination of the heavy tail characteristic of the distribution via an alternative methodology. This methodology involves the observation of the ratio of two survival functions. If the ratio of survivals approaches infinity as x approaches infinity, then one survival function is considered to be heavier than the other. Moreover, the limiting case of the ratio of two survival functions provides the limiting case of two probability density functions. Therefore, this ratio also indicates the ratio of two density functions.

$$\lim_{x\to\infty}\frac{S_1(x)}{S_2(x)} = \lim_{x\to\infty}\frac{S_1'(x)}{S_2'(x)} = \lim_{x\to\infty}\frac{f_1(x)}{f_2(x)}$$

here, we conduct a comparison between the proposed distribution and the Pareto Type II distribution. The survival function of the Pareto Type II distribution is expressed as $S(x) = P(X > x) = \left(1+\frac{x}{\lambda}\right)^{-\alpha}$. For $\alpha > 1$, the ratio between the two distributions tends to infinity as x approaches infinity. This suggests that the tail of the proposed distribution. Figure 6 provides the plot of the tail density for the proposed distribution in comparison with two other distributions, namely the normal distribution and the Parato Type II distribution.

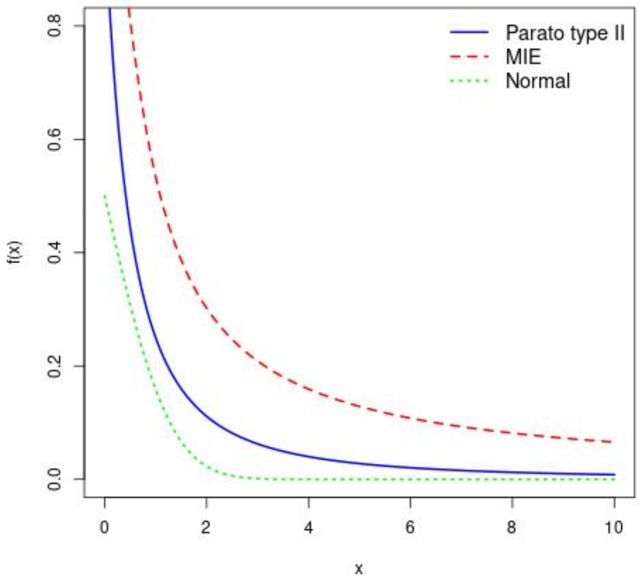

Figure 6: *Tail behaviour of Normal, Parato type II and MIE densities*

Remark 1: A distribution is said to be heavy-tailed distribution if and only if

$$\int_0^\infty e^{\lambda x} f(x)\, dx = \infty \quad \text{for all } \lambda > 0.$$

Hence MIE(γ,θ) is heavy tailed becuase it satisfies condition.

Remark 2: if a distribution is said to be heavy-tailed, if the right probabilities are heavier than the exponential distribution if

$$\lim_{x\to\infty} \frac{1 - F(x)}{e^{-\lambda x}} = \infty \quad \text{for all} \quad \lambda > 0.$$

Hence MIE(γ,θ) distribution also satisfies this condition.

Definition: An ultimately positive function f is called regularly varying at infinity with index $\gamma \in \mathbb{R}$ if for any fixed $c > 0$:

$$\lim_{x\to\infty} \frac{f(cx)}{x} = c^\gamma.$$

The aforementioned theorem demonstrates that the density function derived from the MIE(γ,θ) in equation (6) exhibits regularly varying tails.

Theorem 4. The density function of the MIE(γ,θ) distribution is a function with regularly varying tails.

Proof. Using the density function (6), we have:

$$\lim_{x\to\infty} \frac{f(cx)}{f(x)} = \lim_{x\to\infty} \frac{e^{\frac{\theta}{cx}} x^2 (\gamma e^{\frac{\theta}{x}} + 1)^2}{e^{\frac{\theta}{x}} (x + cx)^2 (\gamma e^{\frac{\theta}{cx}} + 1)^2} = 1,$$

applying limits, the above simplifies to:

$$\lim_{x\to\infty} \frac{f(cx)}{f(x)} = 1,$$

Hence the proof. ■

Definition: An ultimately positive function f is long-tailed and is said to belong to class L if and only if:

$$\lim_{x\to\infty}\frac{f(x+y)}{f(x)}=1,\quad \text{for all } y>0.$$

Theorem 5. The MIE(γ,θ)Distribution belongs to the class L.

Proof.

$$\lim_{x\to\infty}\frac{f(x+y)}{f(x)}=\frac{e^{\frac{\theta}{x+y}}x^2(\gamma e^{\frac{\theta}{x}}+1)^2}{e^{\frac{\theta}{x}}(x+y)^2(\gamma e^{\frac{\theta}{x+y}}+1)^2}=1,$$

Hence, f belongs to the class L. ■

Definition: An ultimately positive function f belongs to the class D of dominated variation distributions if there exists $c>0$ such that:

$$\lim_{x\to\infty}\frac{f(x)}{f(2x)}=c,\quad \text{for all } x>0.$$

Theorem 6. The MIE(γ,θ) Distribution belongs to the class D of dominated variation distributions.

Proof.

$$\lim_{x\to\infty}\frac{f(x)}{f(2x)}=\frac{2^2e^{\frac{\theta}{x}}\left(\gamma e^{\frac{\theta}{2x}}+1\right)^2}{e^{\frac{\theta}{2x}}\left(\gamma e^{\frac{\theta}{x}}+1\right)^2}=4,$$

hence f belongs to the class of dominated variation distributions. ■

4.1. Different method of Estimation

In this section, we are looking at three estimation methods for estimating the unknown model parameters of the proposed model. The procedures are detailed below

4.2. Maximum Likelihood Estimation

Let X be a random variable with the pdf of the Modi inverse exponential distribution defined in equation (6), then its log-likelihood function is defined by:

$$\begin{aligned}\log L(x;\gamma,\theta) &= n\log\theta+n\log\gamma+n\log(\gamma+1)\\ &\quad+\theta\sum_{i=0}^{\infty}\frac{1}{x_i}-2\sum_{i=0}^{\infty}\log(x_i\gamma e^{\frac{\theta}{x_i}}+x_i)\end{aligned} \tag{20}$$

Thus, the non-linear normal equations are given as follows:

$$\frac{\partial\log L(x;\gamma,\theta)}{\partial\theta}=\frac{n}{\theta}+\sum_{i=0}^{n}\frac{1}{x_i}-2\sum_{i=0}^{n}\frac{e^{\frac{\theta}{x_i}}\gamma}{(x_i+e^{\frac{\theta}{x_i}}\gamma x_i)} \tag{21}$$

$$\frac{\partial\log L(x;\gamma,\theta)}{\partial\gamma}=\frac{n}{\gamma}+\frac{n}{1+\gamma}-2\sum_{i=0}^{n}\frac{e^{\frac{\theta}{x_i}}x_i}{(x_i+e^{\frac{\theta}{x_i}}\gamma x_i)} \tag{22}$$

The equations from (21) to (22) above are not in closed form. We refer to using some iterative procedure such as Newton Raphson, Bisection methods, or some other method to obtain the approximate maximum likelihood estimates (MLE) of these parameters for the solution of these explicit equations.

4.3. The Asymptotic Confidence Bounds:

The maximum likelihood estimators (MLE) of the unknown parameters γ, θ in the MIE (γ, θ) do not have closed-form solutions. As a result, the exact distribution of the MLE cannot be derived. However, asymptotic confidence bounds can be obtained based on the asymptotic distribution of the MLE. The information matrix is calculated by taking the second partial derivatives of equations (21) to (22) and is given as:

$$\frac{\partial^2 \log L(x;\gamma,\theta)}{\partial\theta^2} = -\frac{n}{\theta^2} - 2\sum_{i=0}^{n}\left(\frac{e^{\frac{2\theta}{x_i}}\gamma^2}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)^2} + \frac{e^{\frac{\theta}{x_i}}\gamma}{x_i(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)}\right) \tag{23}$$

$$\frac{\partial^2 \log L(x;\gamma,\theta)}{\partial\gamma^2} = -\frac{n}{\gamma^2} - \frac{n}{(1+\gamma)^2} - 2\sum_{i=0}^{n}\left(-\frac{e^{\frac{2\theta}{x_i}}x_i^2}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)^2}\right) \tag{24}$$

$$\frac{\partial^2 \log L(x;\gamma,\theta)}{\partial\gamma,\theta} = -2\sum_{i=0}^{n}\left(-\frac{e^{\frac{2\theta}{x_i}}\gamma x_i}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)^2} + \frac{e^{\frac{\theta}{x_i}}}{(x_i + e^{\frac{\theta}{x_i}}\gamma x_i)}\right) \tag{25}$$

So that the observed information matrix is given by:

$$I = -\begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{bmatrix}$$

Hence, the variance-covariance matrix is approximated as:

$$V \approx I^{-1} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

To obtain the estimate of V, we replace the parameters by their corresponding maximum likelihood estimators (MLEs) to get:

$$\hat{V} \approx \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix}$$

Using the above variance-covariance matrix, one can derive the $(1-\beta)100\%$ confidence intervals for the parameters θ and γ as follows:

$$\hat{\gamma} \pm Z_{\frac{\gamma}{2}}\sqrt{\text{var}(\hat{\gamma})}, \quad \hat{\theta} \pm Z_{\frac{\gamma}{2}}\sqrt{\text{var}(\hat{\theta})}.$$

4.4. The Percentile Method

Let $X_{(j)}$ be the jth order statistic, i.e., $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$. if p_j denote some estimate of $F(x_j;\theta,\gamma)$, then the estimate of θ and γ can be obtained by minimizing

$$\sum_{j=1}^{n}\left(x_j - \left[\frac{\theta}{\ln\left(\frac{\gamma+1-p_j}{\gamma p_j}\right)}\right]\right)^2;$$

with respect to θ and γ. Several types of estimators for p_j can be used [26], and this paper considers $p_j = \frac{j}{n+1}$.

4.5. Method of Cramer-von Mises

Cramer-von-Mises type minimum distance estimators aim to minimize the disparity between the theoretical and empirical cumulative distribution functions. It has been demonstrated empirically that these estimators have a lesser bias than other minimum distance estimators. The Cramer-von-Mises estimators $\hat{\gamma}_{\text{CME}}$, and $\hat{\theta}_{\text{CME}}$, are the values of γ, and θ that minimizing

$$W(\gamma, \theta) = \frac{1}{12n} + \sum_{i=1}^{n} \left[F\left(t_i\,;\, \gamma, \theta\right) - \frac{2i-1}{2n} \right]^2 \tag{26}$$

where t_i is the i-th ordered observation, and $F(t_i; \gamma, \theta)$ is the cumulative distribution function of the MIED with parameters γ and θ. To estimate the parameters, we differentiate the above equation partially with respect to the parameters γ and θ, respectively, and equate them to zero to get the normal equations. Since the normal equations are nonlinear, we can use iterative methods to obtain the solutions.

5. Simulation Study and Data Analysis

5.1. Simulation Study

In this section, we conduct a comprehensive Monte Carlo simulation study that is repeated 1000 times in order to compare the performance of the previously discussed estimators. We evaluate these estimators using Mean Squared Error (MSE) and make comparisons across sample sizes $n = 50, 100, 150, 200$ for two distinct parameter settings: $\gamma = 1$ and $\theta = 0.05$, and $\gamma = 0.9$ and $\theta = 2.5$. The simulation-based outcomes provide estimates denoted as $\gamma\hat{P}M, \theta\hat{P}M, \gamma\hat{M}L, \theta\hat{M}L, \gamma\hat{C}M$, and $\theta\hat{C}M$ for the Percentile Method (PM), Maximum Likelihood Estimation (MLE), and Cramer-von Mises (CVM) method. The corresponding MSE values are displayed in parentheses. Notably, as sample size n increases, the Mean Squared Error (MSE) tends to decrease, indicating improved estimation accuracy. Table 1 and Table 2 present the simulation results.

n	PM		MLE		CVM	
	$\gamma\hat{P}M$	$\theta\hat{P}M$	$\gamma\hat{M}L$	$\theta\hat{M}L$	$\gamma\hat{C}M$	$\theta\hat{C}M$
50	0.0836 (0.3486)	0.0203 (6.918×10^{-04})	0.0013 (8.709×10^{-04})	0.0206 (5.277×10^{-04})	0.0900 (0.4250)	0.0102 (0.0050)
100	0.0406 (0.1623)	0.0097 (1.303×10^{-04})	0.0005 (6.057×10^{-04})	0.0102 (1.144×10^{-04})	0.0425 (0.1819)	0.0049 (0.0026)
150	0.0269 (0.1069)	0.0065 (6.499×10^{-05})	0.0004 (1.357×10^{-04})	0.0067 (5.123×10^{-05})	0.0282 (0.1195)	0.0033 (0.0017)
200	0.0203 (0.0804)	0.0049 (3.696×10^{-05})	0.0004 (4.004×10^{-05})	0.0050 (3.308×10^{-05})	0.0213 (0.0896)	0.0024 (0.0013)

Table 1: *Simulation outcomes obtained for the parameter value of $\gamma = 1$ and $\theta = 0.05$ are presented herein. The values enclosed within the parentheses denote the Mean Squared Error (MSE) values.*

5.2. Data Analysis

In this section, we demonstrate the usefulness of the proposed Modi Inverse Exponential distribution with parmeter γ and θ . We fit this distribution to a real-life data set and compare the results with some recent efficient models, namely the Inverse Generalized Weibull distribution, Generalized Inverse Generalized Weibull distribution. The corresponding PDFs are presented below:

- Inverse Generalized Weibull Distributions:

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

- Generalized Inverse Generalized Weibull Distribution:

n	PM		MLE		CVM	
	$\hat{\gamma_{PM}}$	$\hat{\theta_{PM}}$	$\hat{\gamma_{ML}}$	$\hat{\theta_{ML}}$	$\hat{\gamma_{CM}}$	$\hat{\theta_{CM}}$
50	0.0564 (3.926×10^{-03})	0.0725 (3.08×10^{-02})	1.9823 (0.2345)	0.05313 (3.09×10^{-02})	0.0065 (4.036×10^{-03})	0.0362 (3.103×10^{-02})
100	0.02762 (2.591×10^{-05})	0.0354 (2.044×10^{-04})	1.2356 (0.0083)	0.0256 (2.055×10^{-04})	0.0029 (2.688×10^{-05})	0.0177 (2.064×10^{-04})
150	0.01835 (2.536×10^{-07})	0.0232 (2.024×10^{-06})	0.8017 (0.0072)	0.0168 (2.048×10^{-06})	0.0019 (2.682×10^{-07})	0.0116 (2.053×10^{-06})
200	0.0138 (4.744×10^{-09})	0.0174 (3.928×10^{-08})	0.2211 (0.0037)	0.0127 (3.991×10^{-08})	0.0014 (5.322×10^{-09})	0.0087 (4.046×10^{-08})

Table 2: *Simulation outcomes obtained for the parameter value of $\gamma = 0.9$ and $\theta = 2.5$ are presented herein. The values enclosed within the parentheses denote the Mean Squared Error (MSE) values.*

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

Data Set 1: This data set has been taken from [27]. The data on survival of 40 patients suffering from leukemia, from the Ministry of Health Hospitals in Saudi Arabia, was taken from Abouammoh et al. (1994):

115	181	255	418	441	461	516	739	743	789	807	865	924	983
1024	1062	1063	1165	1191	1222	1222	1251	1277	1290	1357	1369	1408	1455
1478	1549	1578	1578	1599	1603	1605	1696	1735	1799	1815	1852		

Table 3: *Estimates and Goodness-of-fit measures based on AIC, BIC, AICC, and CAIC for Data Set 1*

Distribution	Estimates	Log-Likelihood	AIC	BIC	AICC	CAIC
MIE	$\gamma = 0.1350$ $\theta = 1.7856$	-326.978	656.822	660.199	657.146	660.199
IGWD	$\alpha = 0.0426$ $\beta = 1.3546$ $\theta = 2.7048$	-346.170	698.340	703.406	699.006	703.406
GIGWD	$\alpha = 0.0323$ $\beta = 0.9067$ $\theta = 1.2614$ $c = 4.2107$	-367.455	742.909	749.664	744.052	749.664

From Table 3, it shows that the proposed Modi Inverse Exponential distribution model has the lowest AIC, BIC, AICC, and CAIC values among the other distributions, suggesting that it provides the best fit to the dataset.

Data Set 2: This data set represents survival times in Days, from a Two-Arm Clinical Trial considered by [28] and [29]. The survival time in days for the 31 patients from Arm B are:

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	

From Table 4, we can see that our proposed model MIE has minimum AIC, BIC, AICC, and CAIC values compared to IGWD and GIGWD distributions. Thus, we can infer that the newly proposed model is a better fit for the given data compared to the other models.

Table 4: *Estimates and Goodness-of-fit measures based on AIC, BIC, AICC, and CAIC for Data Set 2*

Distribution	Estimates	Log-Likelihood	AIC	BIC	AICC	CAIC
MIE	$\gamma = 0.0618$	-206.657	417.315	420.183	417.744	420.183
	$\theta = 1.1340$					
IGWD	$\alpha = 0.0596$	-217.366	440.733	445.035	441.622	445.035
	$\beta = 1.8165$					
	$\theta = 4.5081$					
GIGWD	$\alpha = 0.0665$	-217.015	442.032	447.767	443.570	447.767
	$\beta = 0.8546$					
	$\theta = 4.9719$					
	$c = 1.0867$					

6. Conclusion

In this article, We establishes the Modi Inverse Exponential distribution which a right skewed heavy tailed UBT shaped probability model. The related structural properties are derived and represented in the respective sections. Furthermore, we explore the tail behavior of the suggested model and conclude that it is heavy-tailed. To estimate the distribution's parameters, different estimation methods such as method of maximum likelihood, method percentile and Method of Cramer-von Mises are used. For the simulated data set, the results are shown in Table 1 and 2. We can see that the estimated values obtained are near to the predefined parameters, and as n increases, MSE decreases, confirming the law of large numbers. However, the application to two real-life data sets shows that the MIE distribution has a better fit than other competing models, such as the Inverse Generalized Weibull distribution (IGWD) and Generalized Inverse Generalized Weibull distribution, based on goodness-of-fit measures AIC, BIC, AICC, and CAIC.

References

[1] Venturini, S., Dominici, F. and Parmigiani, G. (2012). Gamma shape mixtures for heavy-tailed distributions. *Annals of Applied Statistics*, 2(2):756–776.

[2] Ickowicz, A. and Sparks, R. (2017). Modelling hospital length of stay using convolutive mixtures distributions. *Statistics in Medicine*, 36(1):122–135.

[3] Harini, S., Subbiah, M. and Srinivasan, M. R. (2017). Fitting length of stay in hospitals using transformed distributions. *Communications in Statistics - Case Studies, Data Analysis and Applications*, 4(1):1–8.

[4] Karagrigoriou, A. and Vonta, I. (2014). Statistical Inference for Heavy-Tailed Distributions in Technical Systems. In *2014 Ninth International Conference on Availability, Reliability and Security* :412–419.

[5] Jayakumar, K. and Fasna, K. (2022). On a New Generalization of Cauchy Distribution. *Asian Journal of Statistical Science*, 2(1): 61–81.

[6] Bhati, D. and Ravi, S. (2018). Evolution by gene duplication. On generalized log-Moyal distribution: A new heavy-tailed size distribution. *Insurance: Mathematics and Economics*, 79:247–259.

[7] Beirlant, J., Matthys, G. and Dierckx, G. (2001). Heavy-tailed distributions and rating. *Astin Bulletin*, 31(1):37–58.

[8] Resnick, S. I. (1997). Discussion of the Danish data on large fire insurance losses. *Astin Bulletin*, 27(1):139–151.

[9] Dutta, K. and Perry, J. (2006). Evolution by gene duplication. A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital.

[10] Adcock, C., Eling, M. and Loperfido, N. (2015). Skewed distributions in finance and actuarial science: a review. *European Journal of Finance*, 21:1253–1281.

[11] Bennett, S. (1983). Log-logistic regression models for survival data.*Applied Statistics*, 32(2):165.

[12] Lamberson, L. R. (1974). An evaluation and comparison of some tests for the validity of the assumption that the underlying distribution of life is exponential. *AIIE Transactions*, 6(4):327–337.

[13] LIN, B. S. DURAN, T. O. LEWIS (1989). Inverted gamma as a life distribution. *Microelectronics Reliability*, 29(4):619–626.

[14] V. K. SHARMA, S. K. SINGH, U. SINGH, V. AGIWAL (2015). The inverse Lindley distribution: A stress-strength reliability model with application to head and neck cancer data. *Journal of Industrial and Production Engineering*, 32(3): 162–173.

[15] Oguntunde, P. and Adejumo, O. (2014). The transmuted inverse exponential distribution. *International Journal of Advanced Statistics and Probability*, 3(1):1–7.

[16] Khan, M. S., King, R., and Hudson, I. (2014). Characterizations of the transmuted inverse Weibull distribution. *ANZIAM Journal*, 55:197–217.

[17] Sharma, V. K., Singh, S. K., and Singh, U. (2014). A new upside-down bathtub shaped hazard rate model for survival data analysis. *Applied Mathematics and Computation*, 239:242–253.

[18] Sharma, V. K., Singh, S. K., Singh, U., and Merovci, F. (2016). The generalized inverse Lindley distribution: A new inverse statistical model for the study of upside-down bathtub data. *Communications in Statistics - Theory and Methods*, 45(19):5709–5729.

[19] Modi, K., Kumar, D., and Singh, Y. (2020). A new Family of Distribution with Application on Two Real Datasets on Survival Problem. *Science and Technology Asia*, 25: 1–10.

[20] Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 75(371): 667-672.

[21] Galton, F. (1883). Inquiries into Human Faculty and its Development,*Macmillan and Company*, London. (25).

[22] Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society, Series D* , 37(1):25–32.

[23] Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by Lehmann alternatives.*Communications in Statistics - Theory and Methods*, 27(4):887–904.

[24] Klugman, S.A, Panjer, H.H, Willmot, G.E.(2012) Loss Models: From Data to Decisions, vol. 715,*John Wiley and Sons*, New York.

[25] Nair,J., Wierman, A.and Zwart,B.(2013). The fundamentals of heavy-tails: Properties, emergence, and identification. In *ACM SIGMETRICS Performance Evaluation Review*, 41:387–388.

[26] Mann, N., Schafer, R., and Singpurwalla, N. (1974). Methods for Statistical Analysis of Reliability and Life Data. *Wiley*.

[27] Abouammoh, A., Abdulghani, S., Qamber, I. (1994). On partial orderings and testing of new better than renewal used classes.*Reliability Engineering and System Safety*, 43:37–41.

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

[29] 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.