

# THE POWER GENERALIZED KAVYA MANOHARAN TRANSFORMATION: APPLICATIONS IN RELIABILITY AND LIFETIME REGRESSION

JINET MARIYA JOJY<sup>1</sup>, SAJESH T. A.<sup>2</sup> AND NICY SEBASTIAN<sup>3\*</sup>

<sup>1,2,3</sup> Department of Statistics, St Thomas College, Thrissur  
affiliated to the University of Calicut, Kerala, India- 680 001

<sup>1</sup>jinetmariyajojoy@gmail.com, <sup>2</sup>sajesh.t.abraham@gmail.com,

<sup>3\*</sup>nicycms@gmail.com

\* Corresponding Author

## Abstract

*A new class of distributions is introduced through the exponentiated generalization of the Kavya Manoharan transformation, referred to as the Power Generalized Kavya Manoharan (PGKM) transformation. In parallel systems, applying power transformations to the distribution of individual components yields the system's overall distribution. This generalization enhances the model's flexibility and accuracy. In this context, we propose a transformation to generate a novel class of distributions, specifically by developing a distribution based on the Inverse Weibull distribution as the baseline. We investigate the behaviour of the hazard rates of these distributions, along with other analytical properties. Reliability measures for both single-component and multi-component stress-strength models are derived. The parameters of the proposed model are estimated using the maximum likelihood method, and simulation studies confirm the consistency of these estimates. Additionally, the new life distribution is applied to a real dataset, demonstrating superior fit when compared to various existing distributions in the literature. The distribution is further re-parametrized with location-scale parameters and employed in a lifetime regression analysis, with application provided to a practical dataset.*

**Keywords:** Lifetime, PGKM transformation, Inverse Weibull distribution, Lifetime Regression Analysis.

## 1. INTRODUCTION

Modelling and analysis of lifetime distributions have been widely used in many fields of science such as medicine, engineering and statistics. There are several methods to propose new classes of distributions using the existing distributions as the baseline distribution. For example, if  $G(x)$  is some baseline distribution function, then a new distribution function is obtained as  $F(x) = (G(x))^\alpha$ , where  $\alpha > 0$  is a shape parameter. This is the well-known Lehmann family found useful in modelling failure time data as revealed in several research works, see, Gupta et al. [12]. Another approach to generalizing a baseline distribution is to transform it using the quadratic rank transmutation map (QRTM), see Shaw and Buckley [23]. If  $H(x)$  be the cdf of transmuted distribution corresponding to the baseline distribution having cdf  $F(x)$ , then

$$H(x) = (1 + \lambda)F(x) - \lambda\{F(x)\}^2, |\lambda| \leq 1.$$

Recently, various generalizations have been introduced based on QRTM, see Aryal and Tsokos [2], Aryal [3], Khan and King [15] etc. Kumar et al. [17] proposed a method to get new lifetime

distributions called DUS(Dinesh-Umesh-Sanjay) transformation and is defined as follows: Let  $G_1(x)$  be the cumulative distribution function (cdf) and  $g_1(x)$  be the probability density function (pdf) of some baseline distribution. Then by using DUS transformation, a new distribution is obtained corresponding to the baseline distribution as

$$F_1(x) = \frac{1}{e-1} \left( e^{G_1(x)-1} \right), \quad x \in R \tag{1}$$

$$f_1(x) = \frac{1}{e-1} g_1(x) \left( e^{G_1(x)} \right), \quad x \in R. \tag{2}$$

There are many distributions introduced by using the DUS transformation see, Maurya et al. [20], Kavaya and Manoharan [13], Tripathi et al. [26].

Thomas and Chacko [25] a new class of distribution is introduced using an exponentiated generalization of the DUS transformation, called the power generalized DUS (PGDUS) transformation. Let  $X$  be a random variable with a baseline cdf  $G_2(x)$  and the corresponding pdf  $g_2(x)$ . Then, the cdf and pdf of the PGDUS distribution is defined as:

$$F_2(x) = \left( \frac{1}{e-1} (e^{G_2(x)-1}) \right)^\theta, \quad x \in R \quad \theta > 0 \tag{3}$$

$$f_2(x) = \frac{\theta}{(e-1)^\theta} g_2(x) \left( e^{G_2(x)} \right) \left( e^{G_2(x)-1} \right)^{\theta-1}, \quad x \in R \quad \theta > 0. \tag{4}$$

Kavaya and Manoharan [14] introduced a new transformation called Kavaya-Manoharan (KM) transformation. Let  $X$  be a random variable with cumulative distribution function (cdf)  $G_3(x)$  and probability density function (pdf)  $g_3(x)$  of some baseline distribution. Then the cdf and pdf of new distribution is defined as,

$$F_3(x) = \frac{e}{e-1} \left( 1 - e^{-G_3(x)} \right), \quad x \in R \tag{5}$$

$$f_3(x) = \frac{e}{e-1} g_3(x) \left( e^{-G_3(x)} \right), \quad x \in R. \tag{6}$$

### 1.1. Power Generalized Kavaya Manoharan (PGKM) transformation

We introduce a new transformation from now on called Power Generalized Kavaya Manoharan (PGKM) transformation in the present study. The new PGKM distribution can be obtained as follows: Let  $X$  be a random variable with baseline cdf  $G(x)$  and corresponding pdf  $g(x)$ . Then the cdf and pdf of the proposed PGKM distribution is

$$F(x) = \left( \frac{e}{e-1} (1 - e^{-G(x)}) \right)^\theta, \quad x \in R, \quad \theta > 0, \tag{7}$$

and the corresponding pdf is

$$f(x) = \theta \left( \frac{e}{e-1} \right)^\theta \left( 1 - e^{-G(x)} \right)^{\theta-1} \left( e^{-G(x)} \right) g(x), \quad x \in R, \quad \theta > 0. \tag{8}$$

The main objective of our study is to introduce a transformation that yields new lifetime models by using a given baseline distribution. When researchers deal with series systems with components distributed as KM-transformed lifetime distributions, the PGKM transformation is highly useful. So the investigation of the PGDKM transformation of various lifetime distributions is relevant in the sense of the selection of appropriate lifetime models for series systems. The procedure transforms the response variable to achieve a model with interesting ageing properties as revealed through the hazard rates. We choose inverse Weibull as the baseline distribution in the present work because this distribution has wide application in reliability theory and survival

analysis. If  $X$  follows inverse Weibull distribution with parameters  $\alpha$  and  $\beta$ , then the cdf and pdf are respectively given by

$$G(x) = e^{-\left(\frac{\alpha}{x}\right)^\beta}, \quad x > 0, \quad \alpha, \beta > 0 \tag{9}$$

and

$$g(x) = \left(\frac{\beta}{\alpha}\right) \left(\frac{\alpha}{x}\right)^{\beta+1} \left(e^{-\left(\frac{\alpha}{x}\right)^\beta}\right)^\beta, \quad x > 0, \quad \alpha, \beta > 0. \tag{10}$$

The inverse Weibull distribution which is also known as Fréchet distribution has been widely used in reliability engineering. It is a well-defined limiting distribution for the maximum of random variables with non-negative real support. It has application ranging from accelerated life testing through to earthquakes, floods, horse racing, rainfall, wind speeds etc. For more details see Kotz and Nadarajah [16].

We organize the paper as follows: In Section 2, we introduce a new life distribution using the inverse Weibull distribution as the baseline distribution in the proposed transformation and study its various analytical characteristics such as the probability density function, cumulative distribution function, survival function and failure rate function. Shapes of the probability density function and failure rate function are also discussed in this Section. Statistical properties including moments, quantile function, order statistics, stress-strength reliability, Rényi entropy, residual life function, Lorenz curve, characterization results and estimation of the parameters using method of maximum likelihood are discussed in Section 3. Simulation study and real data application is discussed in Section 4. In Section 5, a lifetime regression analysis is investigated by using the log power generalized Kavya Manoharan inverse Weibull distribution, with an illustrative data example. Finally, some concluding remarks are given in Section 6.

## 2. POWER GENERALIZED KAVYA MANOHARAN INVERSE WEIBULL DISTRIBUTION

In this section, Power Generalized Kavya Manoharan Inverse Weibull(PGKMIW) Distribution has been proposed by using inverse Weibull distribution as baseline distribution. Putting cdf and pdf of Inverse Weibull distribution in (7) and (8), we get the cdf and pdf of newly proposed PGKMIW distribution. The cdf is

$$F(x) = \left(\frac{e}{e-1} \left(1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right)\right)^\theta, \quad x > 0, \quad \theta, \alpha, \beta > 0, \tag{11}$$

and the corresponding pdf is

$$f(x) = \alpha^\beta \beta \theta \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right)^{\theta-1} \left(e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right) \left(e^{-\left(\frac{\alpha}{x}\right)^\beta}\right) x^{-\beta-1}, \tag{12}$$

when  $x > 0, \theta, \alpha, \beta > 0$ . The corresponding survival function

$$S(x) = 1 - \left(\frac{e}{e-1} \left(1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right)\right)^\theta, \quad x > 0, \quad \theta, \alpha, \beta > 0 \tag{13}$$

and the hazard function is

$$h(x) = \frac{\alpha^\beta \beta \theta \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right)^{\theta-1} \left(e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right) \left(e^{-\left(\frac{\alpha}{x}\right)^\beta}\right) x^{-\beta-1}}{1 - \left(\frac{e}{e-1} \left(1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}}\right)\right)^\theta}, \tag{14}$$

$x > 0, \theta, \alpha, \beta > 0$ . The plots of pdf and hazard function of PGKMIW distribution are given in Figure 1. The graphical representation gives a better understanding of the shapes of the pdf and hazard function. As seen in the plot the hazard function have upside-down and decreasing curves for different combinations of  $\alpha, \beta$  and  $\theta$ .

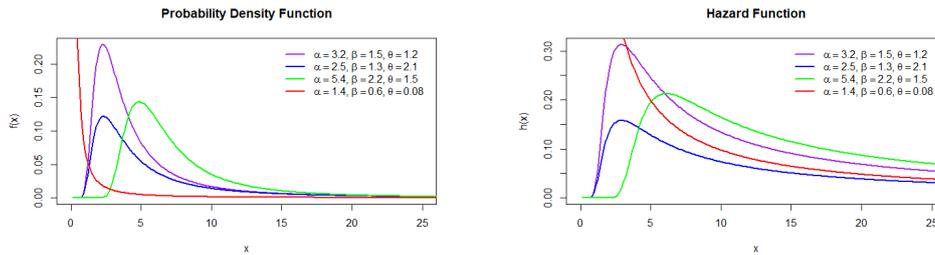


Figure 1: Plots of pdf(left) and hazard function (right) for different values of parameters.

### 3. STATISTICAL PROPERTIES

In this section, some statistical properties of PGKMIW distribution are discussed.

#### 3.1. Moments

The moments are a set of statistical parameters to measure a distribution. The  $r^{th}$  raw moment of the PGKMIW distribution is given by

$$\mu'_r = \sum_{k=0}^{\theta-1} \sum_{m=0}^{\infty} \alpha^r \theta \left( \frac{e}{e-1} \right)^\theta \binom{\theta-1}{k} (-1)^k \frac{(-k+1)^m}{m!} \frac{\Gamma\left(1 - \frac{r}{\beta}\right)}{(m+1)^{1-\frac{r}{\beta}}}. \tag{15}$$

By putting  $r = 1, 2, 3, \dots$  in (15), the raw moments can be viewed.

#### 3.2. Quantile Function

The quantile function of  $X$  is the real solution of the equation  $F(x_p) = p$ , where  $p \sim U(0,1)$ . Then by inverting (11), we get

$$x_p = \frac{-\alpha}{\left( \log \left( -\log \left( 1 - p^{\frac{1}{\theta}} \frac{e-1}{e} \right) \right) \right)^{\frac{1}{\beta}}}. \tag{16}$$

Put  $p = 0.5$  in (16), we get the median of the PGKMIW distribution.

#### 3.3. Order Statistics

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

$$f_r(x) = \frac{n!}{(r-1)!(n-r)!} F^{r-1}(x) (1-F(x))^{n-r} f(x),$$

and

$$F_r(x) = \sum_{j=r}^n \binom{n}{j} F^j(x) (1-F(x))^{n-j}.$$

And those of our newly proposed model are

$$f_r(x) = \frac{n!}{(r-1)!(n-r)!} \frac{\alpha^\beta \theta x^{-\beta-1}}{(e-1)^{\theta n}} \left( 1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}} \right)^{\theta r-1} \left( (e-1)^\theta - (e(1 - e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}})) \right)^{n-r} \left( e^{-e^{-\left(\frac{\alpha}{x}\right)^\beta}} - \left(\frac{\alpha}{x}\right) + \theta r \right), \tag{17}$$

$$F_r(x) = \sum_{j=r}^n \binom{n}{j} \frac{e^{\theta j}}{(e-1)^{\theta n}} \left(1 - e^{-e^{-\left(\frac{x}{\beta}\right)^\beta}}\right)^{\theta j} \left((e-1)^\theta - \left(e^{-e^{-\left(\frac{x}{\beta}\right)^\beta}}\right)^\theta\right)^{n-j}. \quad (18)$$

Then, the pdf and cdf of  $X_{(1)}$  and  $X_{(n)}$  are obtained by substituting  $r = 1$  and  $r = n$  respectively in  $f_r(x)$  and  $F_r(x)$ . It is nothing but the distribution of minimum and maximum in series and parallel reliability systems, respectively.

### 3.4. Stress-Strength Reliability

The Stress-Strength model plays a crucial role in reliability engineering. Stress-strength reliability is defined as the probability that a component's or system's random strength exceeds its random stress, Church and Harris [6]. Literature provides studies on point estimation of the stress-strength reliability parameter for parallel systems with independent and non-identical components, Pakdaman and Ahmadi [21]. This section focuses on reliability estimation for both single-component stress-strength models (SSS) and multicomponent stress-strength models (MSS).

#### 3.4.1 Single Component Stress-Strength Reliability

Suppose  $X$  and  $Y$  have the  $PGKMIW(\alpha, \beta, \theta_1)$  and  $PGKMIW(\alpha, \beta, \theta_2)$  distributions respectively, then the system reliability is

$$\begin{aligned} R &= P(Y < X) \\ &= \int_0^\infty f_X(x)F_Y(x)dx \\ &= \int_0^\infty \alpha^\beta \beta \theta_1 \left(\frac{e}{e-1}\right)^{\theta_1} \left(1 - e^{-e^{-\left(\frac{x}{\beta}\right)^\beta}}\right)^{\theta_1-1} \left(e^{-e^{-\left(\frac{x}{\beta}\right)^\beta}}\right) \left(e^{-\left(\frac{x}{\beta}\right)^\beta}\right) x^{-\beta-1} \\ &\quad \times \left(\frac{e}{e-1}(1 - e^{-e^{-\left(\frac{x}{\beta}\right)^\beta}})\right)^{\theta_2} dx \\ &= \sum_{m=0}^\infty \sum_{k=0}^{\theta_1+\theta_2-1} \binom{\theta_1+\theta_2-1}{k} (-1)^k \frac{\theta_1}{(m+1)!} \left(\frac{e}{e-1}\right)^{\theta_1+\theta_2} (-(k+1))^m. \end{aligned}$$

#### 3.4.2 Multicomponent Stress-Strength (MSS) Reliability

Suppose  $X$  and  $Y$  have the  $PGKMIW(\alpha, \beta, \theta_1)$  and  $PGKMIW(\alpha, \beta, \theta_2)$  distributions respectively, then the reliability in MSS is

$$\begin{aligned} R_{s,k} &= \sum_{i=s}^k \binom{k}{i} \int_{-\infty}^\infty (1 - F_X(x))^i F_X(x)^{k-i} dF_Y(x) \\ &= \sum_{r=0}^\infty \sum_{i=s}^k \sum_{j=0}^i \sum_{m=0}^{\theta_1(k-i)} \sum_{n=0}^{\theta_2-1} \sum_{p=0}^{\theta_1 j} \binom{k}{i} \binom{i}{j} \binom{\theta_1(k-i)}{m} \binom{\theta_2-1}{n} \binom{\theta_1 j}{p} \\ &\quad \times (-1)^{r+j+m+n+p} \frac{\theta_2}{(r+1)!} (e-1)^{\theta_1(i-j)} (e)^{\theta_1 j} \left(\frac{e}{e-1}\right)^{\theta_1(k-i)+\theta_2} (m+n+p+1)^r. \end{aligned}$$

### 3.5. Rényi Entropy

Entropy is interpreted as the degree of disorder or randomness in the system. An important measure of entropy is Rényi entropy, see Rényi [22]. The Rényi entropy is defined as

$$I_R(\delta) = \frac{1}{1-\delta} \left[ \log(I(\delta)) \right]$$

where  $I(\delta) = \int f^\delta(x)dx$ ,  $\delta > 0$  and  $\delta \neq 1$ . The function  $I(\delta)$  of PGKMIW distribution is given by

$$I(\delta) = \sum_{k=0}^{\delta(\theta-1)} \sum_{m=0}^{\infty} a^{1-\delta} \beta^{\delta-1} \theta^\delta \left(\frac{e}{e-1}\right)^{\delta\theta} \binom{\delta(\theta-1)}{k} (-1)^k \times \frac{(-k+1)^m}{m!} \frac{\Gamma\left(\frac{\delta\beta+\delta-1}{\beta}\right)}{(m+\delta)^{\frac{\delta\beta+\delta-1}{\beta}}}. \quad (19)$$

### 3.6. Residual Life Function

The Residual Life Function (RLF) is a powerful tool to quantify and manage the remaining time until failure or event occurrence for systems that have already survived some period of time. The  $n^{th}$  moment of the residual life is defined as

$$\begin{aligned} m_n(t) &= E((x-t)^n | X > t), \quad \forall n = 1, 2, 3, \dots \\ &= \frac{1}{1-F(t)} \int_t^\infty (x-t)^n dF(x) \\ &= \frac{1}{F(t)} \int_t^\infty (x-t)^n f(x) dx. \end{aligned}$$

The  $n^{th}$  moment of the residual life of PGKMIW distribution is given by

$$\begin{aligned} m_n(t) &= \frac{1}{1-F(t)} \sum_{k=0}^{\theta-1} \sum_{j=0}^n \sum_{m=0}^{\infty} \alpha^{n-j} \theta \left(\frac{e}{e-1}\right)^\theta \binom{\theta-1}{k} (-1)^{k+j} \frac{(-k+1)^m}{m!} \\ &\quad \times \binom{n}{j} t^j (m+1)^{\frac{n-j}{\beta}-1} \gamma\left(\frac{j-n}{\beta} + 1, (m+1)\left(\frac{\alpha}{t}\right)^\beta\right), \quad (20) \end{aligned}$$

where  $\gamma\left(\frac{j-n}{\beta} + 1, (m+1)\left(\frac{\alpha}{t}\right)^\beta\right)$  is the lower incomplete gamma function.

### 3.7. Lorenz Curve

The Lorenz Curve is a graphical representation used to describe the distribution of a probability density function (PDF), particularly in contexts like economics and inequality analysis. It plots the cumulative proportion of a variable (like income or wealth) against the cumulative proportion of the population, showing how concentrated or evenly distributed that variable is. The Lorenz curve for a positive random variable  $X$  is defined as the graph of the ratio

$$L(F(x)) = \frac{E(X|X \leq x)F(x)}{E(x)} = \frac{\int_0^x xf(x)dx}{\int_0^\infty xf(x)dx} = \frac{1}{\mu} \int_0^x xf(x)dx,$$

against  $F(x)$  with the property  $L(p) \leq p$ ,  $L(0) = 0$  and  $L(1) = 1$ . If  $X$  represents annual income,  $L(p)$  is the proportion of total income that accrues to individuals having the 100 $p$  percent lowest incomes. If all individuals earn the same income then  $L(p) = p$  for all  $p$ . The area between the line  $L(p) = p$  and the Lorenz curve may be regarded as a measure of inequality of income, or more generally, of the variability of  $X$ . The Lorenz curve for the PGKMIW distribution is given by

$$\begin{aligned} L(p) &= \frac{1}{\mu} \sum_{k=0}^{\theta-1} \sum_{m=0}^{\infty} \alpha \theta \left(\frac{e}{e-1}\right)^\theta \binom{\theta-1}{k} (-1)^k \frac{(-k+1)^m}{m!} (m+1)^{\frac{1}{\beta}-1} \\ &\quad \times \Gamma\left(1 - \frac{1}{\beta}, (m+1)\left(\frac{\alpha}{x}\right)^\beta\right), \quad (21) \end{aligned}$$

where  $\Gamma\left(1 - \frac{1}{\beta}, (m+1)\left(\frac{\alpha}{x}\right)^\beta\right)$  is the upper incomplete gamma function.

Figure 2 displays the Lorenz Curve, when applied to PGKMIW distribution when parameters  $\alpha$  and  $\beta$  are fixed, helps in analyzing and visualizing the degree of concentration or inequality in the distribution of values of a random variable.

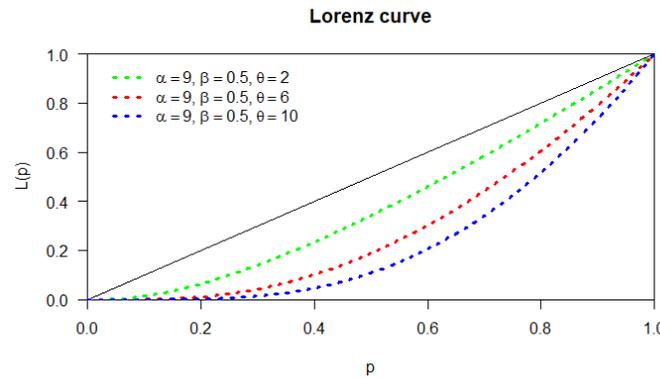


Figure 2: Lorenz curve of PGKMIW distribution

### 3.8. Characterization

Characterization theorems for a probability density function (PDF) provide conditions under which a random variable or a probability distribution is uniquely identified or described by specific properties. These theorems are crucial in probability theory and statistics, as they help establish the uniqueness of distributions based on certain features or properties. This section deals with the characterization of the PGKMIW distribution based on the ratio of two truncated moments. To present the characterization of the distribution, consider the theorem presented in Glänzel [9]. Note that the result also holds when the interval  $H$  is not closed. It could also be applied when the cdf  $F$  does not have a closed form. As shown in Glänzel [10], this characterization is stable in the sense of weak convergence.

**Theorem 1.** Let  $(\Omega, \mathcal{F}, P)$  be a given probability space and let  $H = [a, b]$  be an interval for some  $d < b$  ( $a = -\infty, b = \infty$  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 defined on  $H$  such that

$$E(q_2(X)|X \geq x) = E(q_1(X)|X \geq x)\eta(x), \quad x \in H, \tag{22}$$

is defined with some real function  $\eta$ . Assume that  $q_1, q_2 \in C^1(H)$ ,  $\eta \in C^2(H)$  and  $F$  is twice continuously differentiable and strictly monotone function on the set  $H$ . Finally, assume that the equation  $\eta q_1 = q_2$  has no real solution in the interior of  $H$ . Then  $F$  is uniquely determined by the functions  $q_1, q_2$  and  $\eta$ , particularly

$$F(x) = \int_0^x C \left| \frac{\eta'(u)}{\eta(u)q_1(u) - q_2(u)} \right| e^{-s(u)} du, \tag{23}$$

where the function  $s$  is a solution of the differential equation  $s' = \frac{\eta' q_1}{\eta q_1 - q_2}$  and  $C$  is the normalization constant, such that  $\int_H dF = 1$ .

#### 3.8.1 Characterization based on truncated moments

**Proposition 1.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let

$$q_1(x) = \left( 1 - e^{-e^{-\left(\frac{x}{\lambda}\right)^\beta}} \right)^{1-\theta} \left( e^{-\left(\frac{x}{\lambda}\right)^\beta} \right)$$

and

$$q_2(x) = q_1(x) \left( e^{-\left(\frac{x}{\lambda}\right)^\beta} \right)$$

for  $x > 0$ . The random variable  $X$  belongs to the PGKMIW family if and only if the function  $\eta$  defined in Theorem 1 has the form

$$\eta(x) = \frac{\left( e^{e^{-\left(\frac{x}{\alpha}\right)} } \right)}{\left( e^{-\left(\frac{x}{\alpha}\right)^\beta} \right)}, \quad x > 0. \tag{24}$$

**Proof.** Let  $X$  be a random variable with pdf given in (12) , then

$$(1 - F(x))E(q_1(X)|X \geq x) = \theta \left( \frac{e}{e-1} \right)^\theta \left( e^{-\left(\frac{x}{\alpha}\right)^\beta} \right),$$

and

$$(1 - F(x))E(q_2(X)|X \geq x) = \theta \left( \frac{e}{e-1} \right)^\theta \left( e^{e^{-\left(\frac{x}{\alpha}\right)^\beta}} \right).$$

Further,

$$\eta(x)q_1(x) - q_2(x) = q_1(x) \left( e^{e^{-\left(\frac{x}{\alpha}\right)^\beta}} \right) \left( e^{-\left(\frac{x}{\alpha}\right)^\beta} \right).$$

Conversely, if  $\eta$  is given as above, then

$$s'(x) = \frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = -\alpha^\beta \beta x^{-\beta-1}, \quad x > 0,$$

and hence

$$s(x) = \alpha^\beta x^{-\beta}.$$

Now, according to Theorem 1,  $X$  has density which is given in (12). ■

**Corollary 1.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let  $q_1$  be as in Proposition 1. Then  $X$  has pdf in (12) if and only if there exist functions  $q_2$  and  $\eta$  defined in Theorem 1 satisfying the differential equation

$$\frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = -\alpha^\beta \beta x^{-\beta-1}, \quad x > 0. \tag{25}$$

The general solution of the differential equation in Corollary 1 is

$$\eta(x) = \frac{1}{\left( e^{-\left(\frac{x}{\alpha}\right)^\beta} \right)} \left[ \int \alpha^\beta \beta x^{-\beta-1} \left( e^{-\left(\frac{x}{\alpha}\right)^\beta} \right) (q_1(x))^{-1} q_2(x) + D \right], \tag{26}$$

where  $D$  is a constant. Note that a set of functions satisfying the above differential equation is given in Proposition 1 with  $D = 0$ . Clearly, there exist other triplet of functions  $(q_1, q_2, \eta)$  satisfying the conditions of Theorem 1.

### 3.9. Maximum Likelihood Estimation

There are several methods in the literature for estimating unknown parameters. In this section, maximum likelihood method of estimation is used for estimating the parameters of PGKMIW distribution. Let us consider  $x_1, x_2, \dots, x_n$  be a random sample taken from PGKMIW distribution. Then the likelihood function is given by,

$$L(x) = \prod_{i=1}^n \alpha^\beta \beta \theta \left( \frac{e}{e-1} \right)^\theta \left( 1 - e^{-e^{-\left(\frac{x_i}{\alpha}\right)^\beta}} \right)^{\theta-1} \left( e^{-e^{-\left(\frac{x_i}{\alpha}\right)^\beta}} \right) \left( e^{-\left(\frac{x_i}{\alpha}\right)^\beta} \right) x_i^{-\beta-1}.$$

Then the log-likelihood function,  $l$  becomes,

$$\begin{aligned}
 l = & n\beta \ln \alpha + n \ln \beta + n \ln \theta + n\theta \ln e - n\theta \ln(e - 1) \\
 & + (\theta - 1) \sum_{i=1}^n \ln \left( 1 - e^{-e^{-\left(\frac{\alpha}{x_i}\right)^\beta}} \right) - \sum_{i=1}^n \left( e^{-\left(\frac{\alpha}{x_i}\right)^\beta} \right) \\
 & - \sum_{i=1}^n \left( \frac{\alpha}{x_i} \right)^\beta - (\beta + 1) \sum_{i=1}^n \ln x_i. \quad (27)
 \end{aligned}$$

The maximum likelihood estimators (MLEs) are obtained by maximizing the log-likelihood with respect to the unknown parameters  $\alpha, \beta$  and  $\theta$ . To obtain the maximum likelihood estimates of the unknown parameters, we equate,

$$\frac{\partial l}{\partial \alpha} = 0; \frac{\partial l}{\partial \beta} = 0; \frac{\partial l}{\partial \theta} = 0.$$

The ML estimators are found through the solution of the non-linear system of equations. Hence, using R or MATLAB, a numerical approximation of the software’s solution to this system of equations is possible.

#### 4. NUMERICAL ILLUSTRATIONS

##### 4.1. Simulation Study

In order to illustrate the accuracy of the maximum likelihood estimation procedure for PGKMIW distribution, Monte Carlo simulation is carried out using inversion method. For the parameters  $\alpha, \beta$  and  $\theta$ , sample of sizes  $n = 250, 500, 1000, 1200$  are generated from the PGKMIW model. We calculated the bias and the MSEs of the parameter estimates. The simulation is conducted for the parameter values  $\alpha = 0.8, \beta = 0.5$  and  $\theta = 1.5$ .

**Table 1:** Simulation study at  $\alpha = 0.8, \beta = 0.5$  and  $\theta = 1.5$

n	Estimates	Bias	MSE
250	$\hat{\alpha} = 0.50070$	-0.29930	0.08958
	$\hat{\beta} = 0.53776$	0.03776	0.00143
	$\hat{\theta} = 1.88482$	0.38482	0.14809
500	$\hat{\alpha} = 0.553287$	-0.26713	0.07136
	$\hat{\beta} = 0.53497$	0.03497	0.00122
	$\hat{\theta} = 1.78101$	0.28101	0.07897
1000	$\hat{\alpha} = 1.01974$	0.21974	0.04829
	$\hat{\beta} = 0.51735$	0.01735	0.00030
	$\hat{\theta} = 1.24612$	-0.25388	0.06445
1200	$\hat{\alpha} = 0.78181$	-0.01819	0.00033
	$\hat{\beta} = 0.50654$	0.00654	0.00004
	$\hat{\theta} = 1.46821$	-0.03179	0.00101

As the sample size increases, the mean square error decreases for the selected parameter values (see Table 1). The bias caused by the estimates is nearer to zero. Also, when the sample size increases, absolute bias decreases. Thus, the estimates tend to the true parameter values with the increase in sample size.

##### 4.2. Data Analysis

In this section, we examine vinyl chloride data collected from clean up-gradient monitoring wells. Vinyl chloride, a volatile organic compound, is particularly significant in environmental

investigations due to its anthropogenic origin and carcinogenic properties. However, trace amounts of this compound are often detected in many background monitoring wells. These low-level detections in clean up-gradient wells can be attributed to cross-contamination from air, gas, or the analytical process itself. The data is obtained from Bhaumik et al. [4] and it represents 34 observations of the vinyl chloride data obtained from clean up gradient ground - water monitoring wells in mg/L is given in Table 2.

**Table 2:** Vinyl chloride data

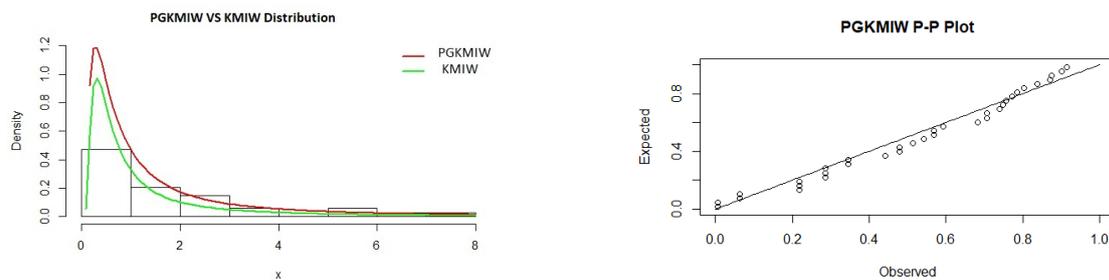
5.1	1.2	1.3	0.6	0.5	2.4	0.5	1.1	8.0	0.8
0.4	0.6	0.9	0.4	2.0	0.5	5.3	3.2	2.7	2.9
2.5	2.3	1.0	0.2	0.1	0.1	1.8	0.9	2.0	4.0
6.8	1.2	0.4	0.2						

We consider two distributions for comparison purpose and they are Kavya Manoharan Inverse Weibull (KMIW) distribution (see Gauthami et al. [8]) and Inverse Weibull (IW) distribution. We have used log-likelihood function, AIC, K-S test value and  $p$ -value for the comparison purpose.

**Table 3:** Parameter estimates and goodness of fit statistics for models fitted to the data

Model	MLE	-Log L	AIC	KS	p-value
PGKMIW	$\hat{\alpha} = 6.9946$ $\hat{\beta} = 0.8458$ $\hat{\theta} = 0.1411$	-57.2529	120.5058	0.1005	0.8821
KMIW	$\hat{\alpha} = 0.9107$ $\hat{\beta} = 0.7685$	-59.0510	122.1020	0.1090	0.8144
IW	$\hat{\alpha} = 0.6173$ $\hat{\beta} = 0.8804$	-58.6266	121.2532	0.1134	0.7745

Table 3 shows that the proposed distribution gives the lowest AIC, K-S test values,  $p$ -values and the largest log-likelihood value. So we can conclude that our proposed distribution provides a better fit for the data set than the other distributions given in Table 3. Figure 3 represents the



**Figure 3:** (a) Plot of the estimated pdfs over the histogram and (b) pp- plot of the PGKMIW model for dataset.

estimated pdfs over the histogram and pp-plot of the PGKMIW model for the dataset.

### 5. LOG POWER GENERALIZED KAVYA MANOHARAN INVERSE WEIBULL DISTRIBUTION (LPGKMIW)

Let the random variable  $T$  follows the PGKMIW distribution with parameters  $\alpha$  and  $\beta$ . Let us consider a re-parametrization by  $\sigma = \frac{1}{\beta}$  and  $\mu = \log \alpha$ . Then, the random variable  $Y = \log T$  will

have LPGKMIW distribution with the following cdf and pdf respectively as,

$$F(y) = \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right)^\theta, \quad -\infty < y < \infty, \theta > 0, \tag{28}$$

$$f(y) = \frac{\theta}{\sigma} \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right)^{\theta-1} \left(e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right) \left(e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}\right) \left(e^{-\left(\frac{y-\mu}{\sigma}\right)}\right), \tag{29}$$

$-\infty < y < \infty, \theta > 0$ , where  $\sigma > 0$  and  $-\infty < \mu < \infty$  are the scale and location parameters respectively. The plot of pdf of LPGKMIW distribution is given in Figure 4. The survival and

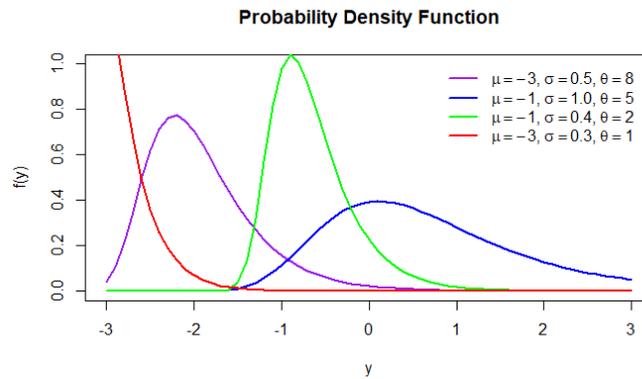


Figure 4: pdf of LPGKMIW distribution

hazard functions of LPGKMIW distribution respectively are as follows:

$$S(y) = 1 - \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right)^\theta, \tag{30}$$

and

$$h(y) = \frac{\frac{\theta}{\sigma} \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right)^{\theta-1} \left(e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right) \left(e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}\right) \left(e^{-\left(\frac{y-\mu}{\sigma}\right)}\right)}{1 - \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}}}\right)^\theta}, \tag{31}$$

where  $-\infty < y < \infty, \sigma > 0, -\infty < \mu < \infty, \theta > 0$ . Then the standardized random variable  $Z = (y - \mu)/\sigma$  has density function given by

$$f(z) = \theta \left(\frac{e}{e-1}\right)^\theta \left(1 - e^{-e^{-e^{-z}}}\right)^{\theta-1} \left(e^{-e^{-e^{-z}}}\right) \left(e^{-e^{-z}}\right) \left(e^{-z}\right), \tag{32}$$

$-\infty < z < \infty, \theta > 0$ .

### 5.1. LPGKMIW Regression Model

In real-world situations, survival time is affected by various factors that contribute to its variability. To evaluate the impact of these factors on survival time, it is crucial to use appropriate regression models that account for censored and time-to-failure data. In essence, constructing such models requires a probabilistic framework for survival time. Several types of regression models are available for this purpose, among which the location-scale regression model, as emphasized by Lawless [18], is particularly noteworthy and commonly applied in clinical trials. Log-lifetime models are based on a linear combination assumption, maintaining the structure of standard

regression models while also offering the benefit of managing censored data. Numerous studies have utilized log-location-scale regression models in their research. Among these, de Gusmão et al. [11] proposed a location-scale regression model based on the log-generalized inverse Weibull distribution for modeling lifetime data. Al-Dawsari and Sultan [1] studied the classical and Bayesian regression models for use in conjunction with the inverted Weibull (IW) distribution. Recently, a regression model is introduced by using the unit inverse Weibull distribution, see Lucas [19].

We have now introduced a linear regression model using the LPGKMIW distribution to analyze the relationship between the response variable  $y_i$  and the explanatory variable vector  $\mathbf{x}_i = (x_1, x_2, \dots, x_p)^T$  in the following way:

$$y_i = \boldsymbol{\beta}^T \mathbf{x}_i + \sigma z_i, \quad i = 1, 2, 3, \dots, n, \tag{33}$$

where  $z_i$  is the random error with density function in (32),  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$  is a  $p \times 1$  vector of unknown parameters,  $\sigma > 0$  is an unknown scale parameters and  $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^T$  is the explanatory variable vector. The parameter  $\mu_i = \boldsymbol{\beta}^T \mathbf{x}_i$  is the location of  $Y_i$ . The location parameter vector  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$  can be represented as a linear model  $\boldsymbol{\mu} = \boldsymbol{\beta}^T \mathbf{X}$ , where  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T$  is a known model matrix.

### 5.2. Maximum Likelihood Estimation

Let  $(y_1, \delta_1, \mathbf{x}_1), \dots, (y_n, \delta_n, \mathbf{x}_n)$  be a right censored random sample of  $n$  observations, where

$$y_i = \begin{cases} \log(t_i), & \text{if } \delta = 1 \\ \log(c_i), & \text{if } \delta = 0, \end{cases}$$

$t_i$  and  $c_i$  are lifetimes and censoring times respectively and  $\mathbf{x}_i$  is an explanatory variable. Assuming that the lifetimes and censoring times are random and independent, the log likelihood function is given by

$$\begin{aligned} l(\boldsymbol{\eta}) = & r \log \theta - r \log \sigma + \theta r \log e - \theta r \log(e - 1) + (\theta - 1) \sum_{i=1}^n \delta_i \log(1 - e^{-e^{-z_i}}) \\ & - \sum_{i=1}^n \delta_i (e^{-e^{-z_i}}) - \sum_{i=1}^n \delta_i (e^{-z_i}) - \sum_{i=1}^n \delta_i z_i \\ & + \sum_{i=1}^n (1 - \delta_i) \log \left( 1 - \left( 1 - e^{-e^{-z_i}} \right)^\theta \left( \frac{e}{e-1} \right)^\theta \right), \tag{34} \end{aligned}$$

where  $r$  denotes the no.of uncensored observations,  $\boldsymbol{\eta} = (\sigma, \boldsymbol{\beta})^T$  and  $z_i = (y_i - \boldsymbol{\beta}^T \mathbf{x}_i) / \sigma$ . By maximizing the log likelihood in (34), the maximum likelihood estimate (MLE) for the parameter vector  $\boldsymbol{\eta}$  can be obtained.

### 5.3. Real-Life Data Analysis

Stone [24] describes an experiment involving solid epoxy electrical insulation specimens subjected to an accelerated voltage life test. A total of 20 specimens were tested at each of three voltage levels: 52.5, 55.0, and 57.5 kV. The failure times, measured in minutes, are presented in Table 4, with censored data marked by asterisks. The sample size is  $n = 60$ , the percentage of censored observations is 10%.

**Table 4:** Failure Times for Epoxy Insulation Specimens at Three Voltage Levels

Voltage (kV)	Failure Times (min)
52.5	4690, 740, 1010, 1190, 2450, 1390, 350, 6095, 3000, 1458, 6200*, 550, 1690, 745, 1225, 1480, 245, 600, 246, 1805
55.0	258, 114, 312, 772, 498, 162, 444, 1464, 132, 1740*, 1266, 300, 2440*, 520, 1240, 2600*, 222, 144, 745, 396
57.5	510, 1000*, 252, 408, 528, 690, 900*, 714, 348, 546, 174, 696, 294, 234, 288, 444, 390, 168, 558, 288

Let  $y_i$  denote the failure time for epoxy insulation specimens (min) and  $x_{i1}$  be the Voltage (kV). The voltage dataset is analyzed using the following regression model:

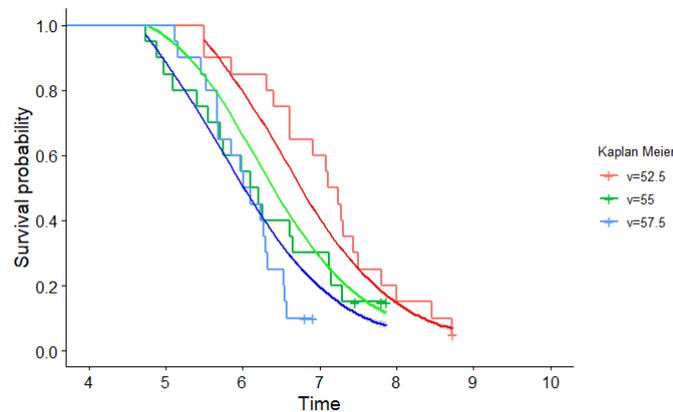
$$y_i = \beta_0 + \beta_1 x_{i1} + \sigma z_i, \tag{35}$$

where the random variable  $y_i$  follows the LPGKMIW distribution given in (29). For comparison, we also consider that  $y_i$  follows the Log Weibull (LW) (see, Lawless [18]), the Log Gamma Extended Weibull (LGEW) distribution (see, Cordeiro et al. [7]), the Log Topp Leone Generated Burr XII (LTLGBXII) distribution (see, Yousof et al. [27]), and the Marshal Olkin Weibull (LGMOW) distribution (see, Chesneau et al. [5]).

**Table 5:** MLEs of the parameters, SEs in second line, p values in [·] and the AIC statistics

Model	Parameters						AIC
LPGKMIW	$\beta_0$	$\beta_1$	$\sigma$	$\theta$			165.5
	10.9660	-1.1488	0.8601	78.0224			
	2.3896	0.0433	0.0885	5.7106			
	[< 0.001]	[0.0006]					
LGMOW	$\beta_0$	$\beta_1$	$\sigma$	$\lambda$	$\alpha$		166.4
	18.2532	-0.1798	0.4480	0.8969	0.0110		
	3.0054	0.0549	0.0631	0.0665	0.0101		
	[1.25 × 10 <sup>-9</sup> ]	[0.0011]					
LTLGBXII	$\beta_0$	$\beta_1$	$\sigma$	$\gamma$	$\theta$	$\beta$	168.4
	14.4513	-0.1790	0.8024	0.6860	7.7247	0.7089	
	4.876	0.074	0.827	0.685	7.027	0.984	
	[0.003]	[0.0155]					
LGEW	$\beta_0$	$\beta_1$	$\sigma$	$\beta$	$\alpha$		168.6
	31.859	-0.21	9.320	298.90	81.891		
	3.730	0.058	0.778	33.045	0.081		
	[< 0.0001]	[< 0.0001]					
LW	$\beta_0$	$\beta_1$	$\sigma$				173.4
	22	-0.274	0.845				
	3.046	0.055	0.09				
	[< 0.0001]	[< 0.0001]					

The MLEs of the model parameters, the asymptotic SEs of these estimates and the values of the AIC measures are used to compare the LPGKMIW with LGMOW, LTLGBXII, LGEW and LW regression models and their values are given in Table 5. It can be seen that the LPGKMIW model has the lowest AIC value, which suggests that the LPGKMIW regression model provides a better fit for the voltage data compared to the LGMOW, LTLGBXII, LGEW, and LW models. Figure 5 presents the empirical survival function by Kaplan Meier estimates and the estimated survival



**Figure 5:** Empirical and fitted survival function.

function based on fitting the LPGKMF regression model. These plots suggest that the regression model fits the data well.

## 6. CONCLUSION

A new class of distribution by generalizing the KM transformation, called the PGKM transformation, is introduced. A new lifetime distribution called the PGKMIW distribution with Inverse Weibull distribution as the baseline distribution is also proposed. Characterization based on truncated moments are presented. Different statistical properties are derived. The parameter estimation has been done through the method of maximum likelihood. Monte Carlo simulation is carried out. Real data analysis is performed to show that the proposed generalization can be used to provide better fits. A new regression model based on the logarithm of the PGKMIW distribution is proposed. Real data analysis is also performed in new regression model to show that this regression model is more appropriate.

## REFERENCES

- [1] Al-Dawsari, S. R. and Sultan, K. S. (2021). Inverted Weibull Regression Models and Their Applications. *Stats*, 4(2):269–290.
- [2] Aryal, G. R. and Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods and Applications*, 71(12):1401–1407.
- [3] Aryal, G. R. (2013). Transmuted log-logistic distribution. *Journal of Statistics Applications and Probability*, 2(1):11–20.
- [4] Bhaumik, D. K., Kapur, K., and Gibbons, R. D. (2009). Testing Parameters of a Gamma Distribution for Small Samples. *Technometrics*, 51(3):326–334.
- [5] Chesneau, C., Karakaya, K., Bakouch, H. S. and Ku?. C. (2022). An Alternative to the Marshall-Olkin Family of Distributions: Bootstrap, Regression and Applications. *Communications on Applied Mathematics and Computation*, 4(4):1229–1257.
- [6] Church, J. D. and Harris, B. (1970). The estimation of reliability from stress-strength relationships. *Technometrics*, 12(1):49–54.
- [7] Cordeiro, G. M., Nadarajah, S., Ortega, E. M. M. and Ramires, T. G. (2018). An alternative two-parameter gamma generated family of distributions: properties and applications. *Hacetatepe Journal of Mathematics and Statistics*, 47(1):145–173.
- [8] Gauthami P., Mariyamma, K. D. and Kavya, P. (2023). A modified inverse Weibull distribution using KM transformation. *Reliability: Theory and Applications*, 18 (72(1))(3):35–42.

- [9] Glänzel, W. (1987). A Characterization Theorem Based on Truncated Moments and its Application to Some Distribution Families. In: Bauer, P., Konecny, F., Wertz, W. (eds) *Mathematical Statistics and Probability Theory*. Springer, Dordrecht:75–84.
- [10] Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments. *Statistics*, 21(4):613–618.
- [11] Gusmão, F.R.S., Ortega, E.M.M. and Cordeiro, G.M. (2011). The generalized inverse Weibull distribution. *Stat Papers*, 52:591–619.
- [12] Gupta, R. C., Gupta, P. L., Gupta, R. D. (1998). Modeling failure time data by lehman alternatives. *Communications in Statistics - Theory and Methods*, 27(4):887–904.
- [13] Kavya, P., Manoharan, M. (2020). On a Generalized Lifetime Model Using DUS Transformation. In: Joshua, V., Varadhan, S., Vishnevsky, V. (eds) *Applied Probability and Stochastic Processes*. Infosys Science Foundation Series. Springer, Singapore:281–291.
- [14] Kavya, P., and Manoharan, M. (2021). Some parsimonious models for lifetimes and applications. *Journal of Statistical Computation and Simulation*, 91(18):3693–3708.
- [15] Khan, M. S. and King, R. (2013). Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. *Europe. J. of Pure Appl. Math.*, 6 (1):66–88.
- [16] Kotz, S., and Nadarajah, S. *Extreme value distributions Theory and Applications*. Imperial College Press, London, 2000.
- [17] Kumar, D., Singh, U., & Singh S. K. (2015). A method of proposing new distribution and its application to bladder cancer patients data. *J. Stat. Appl. Pro. Lett*, 2(3):235–245.
- [18] Lawless, J. F. *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons, New Jersey, 2011.
- [19] Lucas, D. R. R. (2023). The Unit Inverse Weibull Distribution and its Associated Regression Model. *Journal of Econometrics and Statistics*, 3(1):55–68.
- [20] Maurya, S. K., Kaushik, A., Singh, S. K., and Singh, U. (2017). A new class of distribution having decreasing, increasing, and bathtub-shaped failure rate. *Communications in Statistics-Theory and Methods*, 46 (20):10359–10372.
- [21] Pakdaman, Z. and Ahmadi, J. (2017). Point estimation of the stress–strength reliability parameter for parallel system with independent and non-identical components. *Communications in Statistics-Simulation and Computation*, 47(4):1193–1203.
- [22] Rényi, A. (1961). On measures of entropy and information. In: Neyman, J., Ed., 4th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, 1:547–561.
- [23] Shaw, W. T., Buckley, I. R., (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv preprint arXiv:0901.0434.
- [24] Stone, G. C. Statistical analysis of accelerated aging tests on solid electrical insulation. Unpublished M.A.Sc. Thesis, University of Waterloo, Waterloo, Ont., Canada, 1978.
- [25] Thomas, B. and Chacko, V. M. (2021). Power Generalized DUS Transformation of Exponential Distribution. *International Journal of Statistics and Reliability Engineering*, 9(1):150–157.
- [26] Tripathi, A., Singh, U., and Singh, S. K. (2021). Inferences for the DUS-exponential distribution based on upper record values. *Annals of Data Science*, 8:387–403.
- [27] Yousof, H. M., Altun, E., Rasekhi, M., Alizadeh, M., Hamedani, G. G., and Ali, M. M. (2017). A new lifetime model with regression models, characterizations and applications. *Communications in Statistics - Simulation and Computation*, 48(1):264–286.