

The Transmuted Kumaraswamy Pareto Distribution

K. U. Urama¹, S. I. Onyeagu² and F. C. Eze^{2,*}

¹Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

²Department of Statistics, Nnamdi-Azikiwe University, Awka, Anambra State, Nigeria

Abstract

The generalization of probability distributions in view of improving their flexibility in capturing the shape and tail behavior of disparate data sets has become one of the most active aspects of statistical research. We developed a new Pareto distribution by using Kumaraswamy method which gave rise to a new distribution called Kumaraswamy Pareto distribution. This method is called transmutation. The mathematical properties of the new generalized distribution was presented using quartiles, moments, entropy, order statistics mean deviation and maximum likelihood method of parameter estimation. The new generalized distribution was applied to a real life data on the exceedances of flood peaks (in m^3/s) where it was observed to be superior to its sub models in terms of some goodness-of-fit test such as log likelihood criterion, Akaike Information Criterion (AIC) and Kolmogorov-Smirnov (K-S) measures.

1. Introduction

In model development, probability distributions have been used in real life applications such as in finance, medicine, engineering agriculture, environmental sciences to mention but a few. These distributions include Gamma, Pareto, Weibull, Exponential, Lindley, Kumaraswamy and their generalization.

In some cases, classical probability distributions do not provide adequate fit to some real life data in terms of goodness of fit measure, for instance the normal distribution cannot provide an adequate description of the patterns of asymmetric or skewed data (Urama et al. [25]).

Received: February 26, 2021; Accepted: April 4, 2021

2010 Mathematics Subject Classification: 60E05.

Keywords and phrases: transmutation, quartiles, moments, entropy, order statistics, maximum likelihood.

*Corresponding author

Copyright © 2021 the Authors

Kumaraswamy [19] studied the Kumaraswamy distribution for double bounded random process and applied in hydrology.

Azzalini [7] studied some class of new distributions and also some normal ones. This work gave rise to new methods for generalizing univariate continuous probability distributions.

Oguntunde et al. [22] remarked that attention has been shifted to comparing the performance of compound distributions to that of standard theoretical distributions.

1.1. The quadratic rank transmutation map (QRTM)

Let F_1 and F_2 be the cumulative distribution functions of two distributions that have a common sample space and F_1^{-1} and F_2^{-1} are their inverse functions also called the quantile functions respectively. The general rank transmutation are given by Shaw and Buckley [24] can be defined as $G_{R12}(u) = F_2(F_1^{-1}(u))$ and $G_{R21}(u) = F_1(F_2^{-1}(u))$. The First expression is called Sample transmutation map while the latter is known as the Quadratic transmutation map. The functions G_{12} and G_{21} both map the unit interval $[0,1]$ into it selves.

$G_{R12}(u) = u + \lambda u(1 - u)$, $|\lambda| \leq 1$, from which follows that the cumulative function (cdf) satisfies the relationship $F_2(x) = (1 + \lambda)F_1(x) - \lambda(F_1(x))^2$. Then, by differentiation, it yields $f_2(x) = f_1(x)(1 + \lambda - 2\lambda F_1(x))$, where f_1 and f_2 are the corresponding probability density functions associated with the cumulative distribution functions F_1 and F_2 respectively. Therefore, a random variable X is said to have a transmuted distribution if its cumulative distribution function (cdf) and probability density function (pdf) satisfies the following relationships

$$G(x) = (1 + \lambda)F(x) - \lambda(F(x))^2, \quad x > 0, \quad -1 < \lambda < 1 \quad (1.1)$$

and

$$g(x) = f(x)[(1 + \lambda - 2\lambda F(x))], \quad (1.2)$$

$f(x)$ and $f(x)$ are the corresponding probability density functions associated with the cumulative distribution functions $F(x)$ and $G(x)$ respectively.

Generally, quadratic rank transmutation map have been used as a convenient way of constructing new distributions, in particular, survival time models. According to Shaw and Buckley [24], transmutation maps comprise the functional composition of the

cumulative function of one distribution with the inverse cumulative distribution (quartile) function of another.

Differentiating equation (1.1) we will obtain the corresponding probability density function of the transmuted probability distribution $g(x) = f(x) (1 + \lambda - 2\lambda F(x))$, where, λ is the transmuted parameter and $F(x)$ is the baseline function. Equations (1.1) and (1.2) will reduce to the distribution function and probability density function of the baseline distribution when the transmuted parameter takes the value zero.

1.2. The transmuted Kumaraswamy Pareto distribution

In this study we proposed a four-parameter probability distribution called the transmuted Kumaraswamy Pareto distribution. Although Kumaraswamy pareto distribution is an extension of the Pareto distribution taking Kumaraswamy distribution function as the generator. Kumaraswamy Pareto is not common in the literature and has not been fully exploited. Although, its performance in modeling datasets exhibiting extreme value properties was remarked by Bourguignon et al. [9]. Although, they observed that Kumaraswamy Pareto provided a good fit in modeling exceedances of flood peaks data; much is still needed to be done to improve on its goodness of fit by adding extra parameter to the existing parameters of the model. According to Bourguignon et al. [8], a random variable, X , is distributed as Kumaraswamy Pareto if its probability density function is defined as follows,

$$f(x) = \frac{\alpha\beta\theta x_0}{x^{\beta+1}} \left[1 - \left(\frac{x_0}{x}\right)^\beta\right]^{\alpha-1} \left\{1 - \left[1 - \left(\frac{x_0}{x}\right)^\beta\right]^\alpha\right\}^\theta, \tag{1.3}$$

with the corresponding distribution function is given as

$$F(x) = 1 - \left\{1 - \left[1 - \left(\frac{x_0}{x}\right)^\beta\right]^\alpha\right\}^\theta. \tag{1.4}$$

The quadratic rank transmutation map (QRTM) proposed by Shaw and Buckley [24] was defined as

$$G(x) = (1 + \lambda)F(x) - \lambda[F(x)]^2. \tag{1.5}$$

The derivative of equation (1.5) yields the density function of a transmuted distribution given as,

$$g(x) = f(x)[(1 + \lambda) - 2\lambda F(x)], \tag{1.6}$$

where $G(x)$, is the cdf of the baseline distribution, $g(x)$ and $f(x)$ are the corresponding probability density functions (pdf's) associated with $G(x)$ and $F(x)$ respectively. It is important to note that if $\lambda = 0$, equation (1.5) becomes $F(x)$, the distribution function of the baseline random variable; otherwise we shall have the distribution function of the transmuted random variable. Similarly, the probability density function of the baseline distribution will be equal to the probability density function of the transmuted distribution associated with it when $\lambda = 0$. Substituting equation (1.3) and (1.4) into equation (1.6) we will obtain the probability density function of the Transmuted Kumaraswamy Pareto distribution given by

$$g(x) = \frac{\alpha\beta\theta x_0^\beta}{x^{\beta+1}} \left[1 - \left(\frac{x_0}{x}\right)^\beta \right]^{\alpha-1} \left\{ 1 - \left[1 - \left(\frac{x_0}{x}\right)^\beta \right]^\alpha \right\}^{\theta-1} \\ \times \left\{ (1 + \lambda) - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{x_0}{x}\right)^\beta \right]^\alpha \right\}^\theta \right\} \right\}, \quad (1.7)$$

for $x > x_0$ otherwise, $g(x) = 0$.

Note: If $\lambda = 0$, equation (1.7) reduces to the probability density function of the Kumaraswamy Pareto distribution.

Akinsete et al. [1] studied the beta-Pareto distribution and came up with the cumulative distribution function of the beta-Pareto random variable as

$$G(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^{F(x)} t^{\alpha-1} (1-t)^{\beta-1} dx \quad (1.8) \\ 0 < \alpha, \beta < \infty,$$

where α and β are the shape parameters.

Akinsete et al. [1] found that the beta-Pareto distribution is unimodal, exhibits either decreasing or a unimodal hazard rate. The authors derived the expressions for mean, mean deviation and variance, skewness, kurtosis and entropies. The method of maximum likelihood estimate was proposed to estimate the parameters of the model while the application of the model to flood dataset was carried out to show that it can model heavy-tailed distributions.

Aryal and Tsokos [2] studied the Transmuted Generalized Extreme value by using the transmutation map technique on the cumulative generalized extreme value

distribution. The Transmuted Generalized Extreme value distribution function given by the cdf of the form:

$$G(x) = (1 + \lambda) \exp \left\{ - \left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\gamma}} \right\} - \lambda \exp \left\{ - 2 \left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\gamma}} \right\}, \quad (1.9)$$

$$1 + \frac{\gamma(x - \mu)}{\sigma} > 0, -\infty < \mu < \infty, \sigma > 0, |\lambda| \leq 1.$$

Aryal and Tsokos [2] observed that a random variable, X is said to have generalized extreme value (GEV) distribution if its cdf is given by:

$$F(X) = \exp \left\{ - \left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\gamma}} \right\}, \quad (1.10)$$

for $1 + \frac{\gamma(x - \mu)}{\sigma} > 0$ where $-\infty < x < \infty$ is the location parameter, σ is the scale parameter and $-\infty < \gamma < \infty$ is the shape parameter. The corresponding density function is given by

$$f(x) = \frac{1}{\sigma} \left[1 + \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\gamma-1}} \exp \left\{ - \left[1 + \gamma \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\gamma}} \right\} \quad (1.11)$$

where γ determines the tail behaviour of the distribution.

A particular case of the Generalized Extreme Value (GEV) distribution for $\gamma = 0$ and $-\infty < x < \infty$ is the Gumbel distribution. The cases for $\gamma > 0$ and $\gamma < 0$ the GEV distribution tends to be Frechet and the negative Weibull distributins respectively.

The authors studied the transmuted Gumbel distribution with cdf

$$G(x) = (1 + \lambda) \exp \left[- \exp \left\{ - \frac{x - \mu}{\sigma} \right\} \right] - \lambda \exp \left[- 2 \exp \left\{ - \frac{x - \mu}{\sigma} \right\} \right], \quad (1.12)$$

Equation (1.10) yields, equation (1.11) on differentiation, the probability density function (pdf) of a transmuted Gumbel random variable

$$g(x) = \frac{1}{\sigma} \exp \left[- \frac{x - \mu}{\sigma} - \exp \left\{ \frac{x - \mu}{\sigma} \right\} \right] \times \left[1 + \lambda - 2\lambda \exp \left(- \exp \left\{ - \frac{x - \mu}{\sigma} \right\} \right) \right]. \quad (1.13)$$

Aryal and Tsokos [2] further used the transmuted Gumbel distribution to model snow fall data and it was observed to be a very good model for the data.

Aryal and Tsokos [2] developed the analytical framework of the transmuted extreme value probability distribution and derived the expression for basic statistical measures and also provided the maximum likelihood equations of the parameter inherent in the subject distribution. They also illustrated the usefulness and effectiveness of the said transmuted Gumbel probability distribution and applied it in the snow fall data. The goodness of it tests they carried out reviews that the data is well described by this distribution.

Aryal and Tsokos [3] defined and studied the transmuted Weibull distribution with the cumulative distribution function of the form

$$G(x) = \left[1 - \exp\left(-\left(\frac{x}{\sigma}\right)^\eta\right) \right] \left[1 + \lambda \exp\left(-\left(\frac{x}{\sigma}\right)^\eta\right) \right], \quad (1.14)$$

$x > 0, \eta, \sigma > 0, |\lambda| \leq 1$, x , is the transmuted Weibull random variable, η , σ and λ are parameters of the model. They provided the comprehensive description of the mathematical properties of the proposed distribution with its reliability behavior. The author's also used two real world datasets to illustrate the usefulness of this distribution for modeling reliability data and compared the result with Exponentiated Weibull through maximum likelihood estimate and log likelihood procedures. It was also observed that the transmuted Weibull can be a good competitor of other generalized Weibull distributions.

Merovci [20] proposed a generalization of the Lindley distribution using the quadratic rank transmutation map by Shaw and Buckley [24] to construct the transmuted Lindley distribution given by the cumulative distribution function (cdf) of the form:

$$G(x) = \left(1 - \frac{1 + \theta + \theta x}{1 + \theta} e^{-\theta x} \right) \left(1 + \lambda \frac{1 + \theta + \theta x}{1 + \theta} e^{-\theta x} \right), \quad (1.15)$$

$x > 0, \theta > 0, |\lambda| \leq 1$. Where x is the transmuted Lindley-distributed random variable associated with the Transmuted Lindley distribution. The author studied several Mathematical properties of the new distribution alongside with the reliability behaviour. The researcher also observed that the transmuted Lindley distribution is an extended probability model that can analyze more complex data and generalize some of the widely used distributions. An uncensored set involving the remission times (in months) of a random sample of 128 bladder cancer patients was modeled using the transmuted Lindley distribution, Exponential distribution and Lindley distribution. According to the result of the performance measure reported by the author, the transmuted Lindley is a better model than the Lindley distribution.

Elbatal and Elgarhy [15] introduced the Transmuted Quasi-Lindley distribution which is a generalized form of quasi-Lindley distribution. They used the quasi-Lindley distribution as the baseline distribution and define the cumulative distribution function of the Transmuted Quasi-Lindley probability model of the form:

$$G(x) = \left[1 - e^{-\theta x} \left(1 + \frac{\theta x}{1 + \alpha} \right) \right] \times \left[1 + \lambda e^{-\theta x} \left(1 + \frac{\theta x}{1 + \alpha} \right) \right], \tag{1.16}$$

$$x > 0, \alpha > -1, \theta > 0, |\lambda| \leq 1.$$

They derived the moments and moment generating function of the distribution and also the least squares, weighted least square and maximum likelihood estimate of the parameters of the distribution.

Ashour and Eltehiwy [5] proposed and studied a generalization of the Exponentiated Lomax distribution known as the transmuted Exponentiated Lomax distribution using the transmutation techniques and defined the Cumulative distribution function (cdf) of the form:

$$G(x) = \left[1 - (1 + \gamma x)^{-\theta} \right]^\alpha \left\{ (1 + \lambda) - \lambda \left[1 - (1 + \gamma x)^{-\theta} \right]^\alpha \right\}, \tag{1.17}$$

$$x > 0, \alpha > 0, \theta > 0, \gamma > 0, |\lambda| \leq 1.$$

They derived various structural properties of the distribution such as the moments, quantiles, mean deviations, and also obtained the model parameters by maximum likelihood method. They suggested that this distribution can be used to model reliability data.

Elbatal [11] defined the transmuted modified inverse Weibull distribution by the cdf of the form:

$$G(x) = e^{-\frac{\alpha}{x-\eta} - \gamma \left(\frac{1}{x-\eta}\right)^\beta} \left(1 + \lambda - 2\lambda e^{-\frac{\alpha}{x-\eta} - \gamma \left(\frac{1}{x-\eta}\right)^\beta} \right), \tag{1.18}$$

$$\eta < x < \infty, \alpha > 0, \beta > 0, \gamma > 0, |\lambda| \leq 1.$$

Structural properties of the distribution were examined by the author in view of real-life applications of the distribution.

In a similar fashion, Elbatal [12] also defined the Transmuted Generalized Inverted

Exponential distribution by the cumulative distribution function (cdf) given by:

$$G(x) = \left(1 - \left(1 - e^{-\frac{\theta}{x}}\right)^\alpha\right) \left(1 + \lambda \left(1 - e^{-\frac{\theta}{x}}\right)^\alpha\right), \quad (1.19)$$

$$x > 0, \alpha > 0, \theta > 0, |\lambda| \leq 1.$$

They discussed some properties of the distribution of derived moment and other statistics. They obtained the parameters of the model through the maximum likelihood estimate Method and also derived the information matrix.

Ashour and Eltehiwy [6] studied the Transmuted Exponentiated Modified Weibull distribution taking the Exponentiated Modified Weibull as the baseline distribution and obtained the cumulative distribution function (cdf) as:

$$G(x) = \left(1 + \lambda - \lambda \left(1 - e^{-(\theta x + \gamma x^\beta)}\right)\right), \quad (1.20)$$

$$x > 0, \alpha > 0, \theta > 0, \beta > 0, |\lambda| \leq 1.$$

Where x the Transmuted Exponential Modified Weibull random variable while α , β and λ are the model parameters. They derived various structural properties including explicit expression for the moments, quantiles and moment generating function of the distribution and obtained the estimate of the model parameter by the least square method. They concluded that the distribution can be used to model reliability data. The method of least squares was proposed for estimating the parameters of the distribution by the authors.

Elbatal et al. [14] investigated a four- parameter generalized distribution called the Transmuted Generalized Linear Exponential distribution with cumulative distribution function (cdf):

$$G(x) = \left(1 - e^{-(\alpha x + \frac{\beta}{2}x^2)}\right)^\theta \left(1 + \lambda e^{-(\alpha x + \frac{\beta}{2}x^2)}\right)^\theta, \quad (1.21)$$

$$x > 0, \alpha > 0, \theta > 0, \beta > 0, |\lambda| \leq 1.$$

They derived the following mathematical properties of the distribution: a closed form of expression for the density, relative distribution, quantiles, median, moments and moment generating function and also obtained the estimate of the unknown parameters through the maximum likelihood estimation method. They tested the goodness-of-fit of some selected distributions with the Transmuted Generalized Linear Exponential with the Kolgomorov-Smirnov (K-S) distance test statistics and their corresponding P-value on

two real datasets where they found that the Transmuted Generalized Linear Exponential compete favourably with some well-known distribution in modeling lifetime data.

Elbatal and Aryal [13] proposed and studied the Transmuted Additive Weibull distribution that extends the Additive Weibull distribution and other distributions. They defined the cumulative distribution function (cdf) as:

$$G(x) = \left(1 - e^{-\alpha x^\theta - \gamma x^\beta}\right) \left(1 + \lambda e^{-\alpha x^\theta - \gamma x^\beta}\right), \tag{1.22}$$

$$x > 0, \alpha > 0, \theta > 0, \beta > 0, \gamma > 0, |\lambda| \leq 1.$$

They derived the explicit expression for moments, random number generation and order statistics of the subject distribution. They applied the maximum likelihood estimation procedure to obtain the unknown parameters of the model. They also used the analytical results to model real-life data. They compared the model with the Weibull and Additive Weibull distributions to analyze the data and the results indicated that the Transmuted Additive Weibull distribution has the lowest Alkaike Information Criteria (AIC). So, it fits the real-life data better than the other models.

Khan and King [17] defined the Transmuted Modified Weibull distribution with cdf:

$$G(x) = \left(1 - e^{-\alpha x - \gamma x^\beta}\right) \left(1 + \lambda e^{-\alpha x - \gamma x^\beta}\right), \tag{1.23}$$

$$x > 0, \alpha > 0, \beta > 0, \gamma > 0, |\lambda| \leq 1.$$

Several properties of the distribution were studied by the authors and both the method of least squares and maximum likelihood were proposed for the estimation of the parameters of the distribution. The Transmuted Modified Weibull distribution was applied to a real dataset and compared with the Transmuted Exponential, Transmuted Weibull, Modified Weibull and Weibull distributions and the Transmuted Modified Weibull distribution did better than the rest.

2. Methodology

Urama et al. [25] have developed a new family of probability distribution called the transmuted Kumaraswamy Pareto distribution (TKPD) whose cdf is of the form:

$$G(x) = (1 + \lambda) \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} - \lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\}^2, \tag{2.1}$$

which can be factorized to get,

$$G(x) = (1 + \lambda) - \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\}^2 \quad (2.2)$$

$$x \geq \beta; a, b, \beta, k > 0, |\lambda| \leq 1,$$

and the corresponding pdf is

$$g(x) = \frac{abk\beta^k}{x^{k+1}} \left\{ (1 + \lambda) - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} \right\} \left[1 - \left(\frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1} \quad (2.3)$$

$$x \geq \beta; a, b, \beta, k > 0, |\lambda| \leq 1,$$

where a, b, λ and k controls the shape of the density and β the location parameter.

The survival $S(x)$ and hazard $h(x)$ functions of the TKPD are expressed as follows. The survival function is expressed as

$$S(x) = 1 - G(x). \quad (2.4)$$

Thus

$$S(x) = 1 - (1 + \lambda) \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} - \lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\}^2, \quad x \geq \beta, a, b, \beta, k > 0, |\lambda| \leq 1.$$

$$h(x) = \frac{g(x)}{S(x)}, \quad (2.5)$$

$$= \frac{\frac{abk\beta^k}{x^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} \right\} \left[1 - \left(\frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}}{1 - (1 + \lambda) \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} - \lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\}^2}, \quad (2.6)$$

$$x \geq \beta, a, b, \beta, k > 0, |\lambda| \leq 1.$$

The graphs of the hazard function are given by Figures 1 and 2. The graphs of the survival function are also given by Figures 3 and 4. An examination of the hazard in Figure 1 and 2 clearly shows that the function possesses an upside-down bathtub shape. This is very rare in many survival and reliability models. The situation reveals that the TKPD can be effectively used to model processes with an initial increasing failure rate before a decrease occurs as with many engineering devices subjected to *work-hardening* or some biological system which obtains *immunity* overtime.

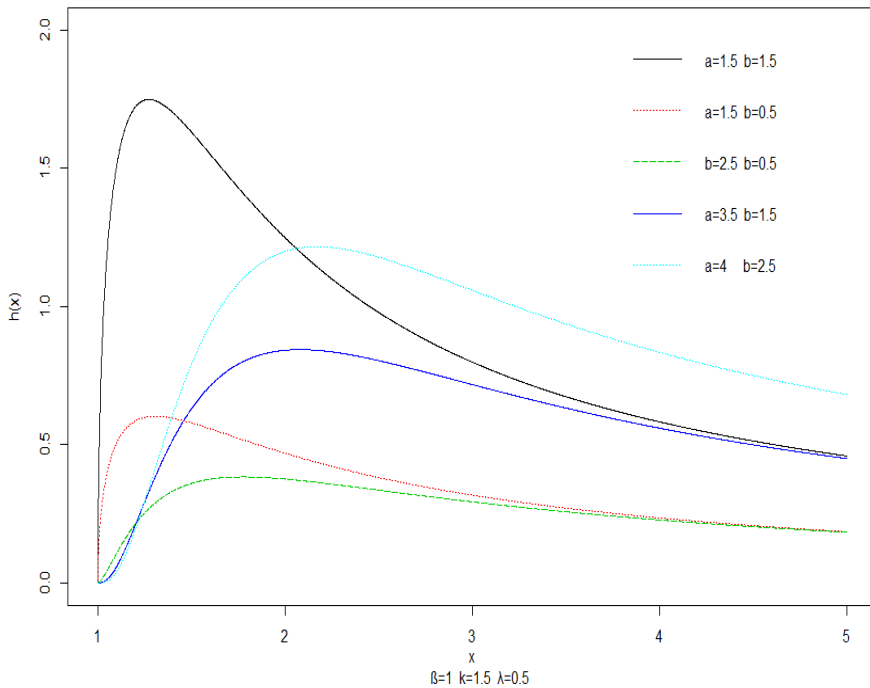


Figure 1. Hazard plot of the TKPD.

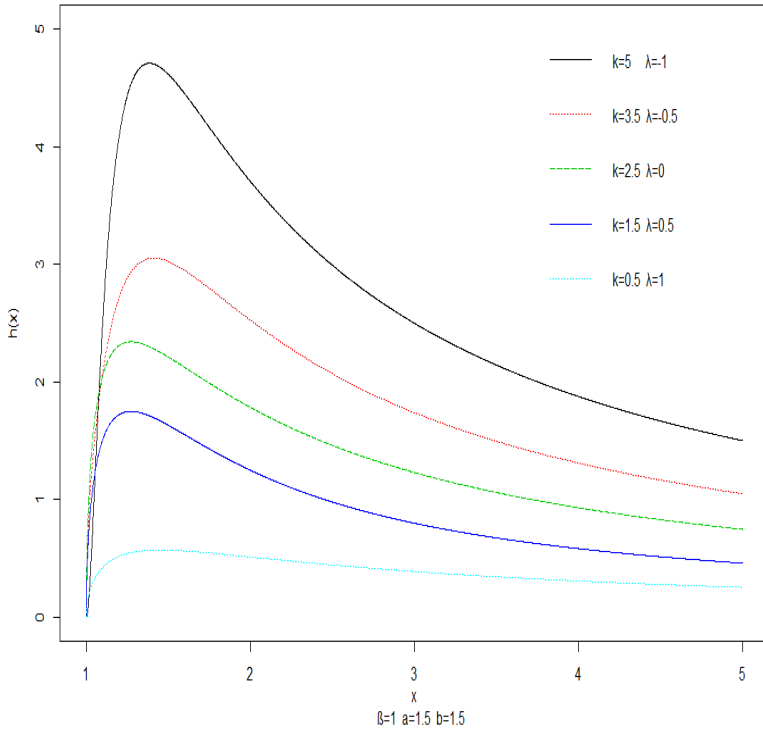


Figure 2. Hazard plot of the TKPD.

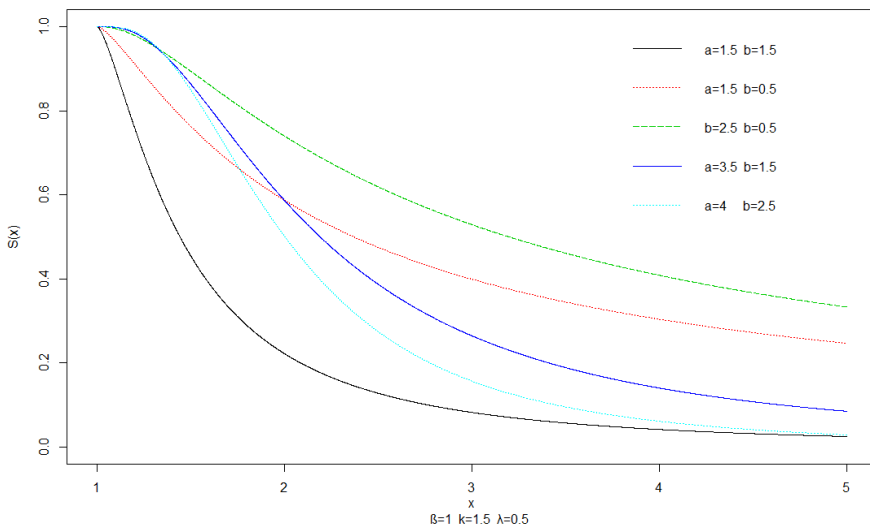


Figure 3. Survival plot of the TKPD.

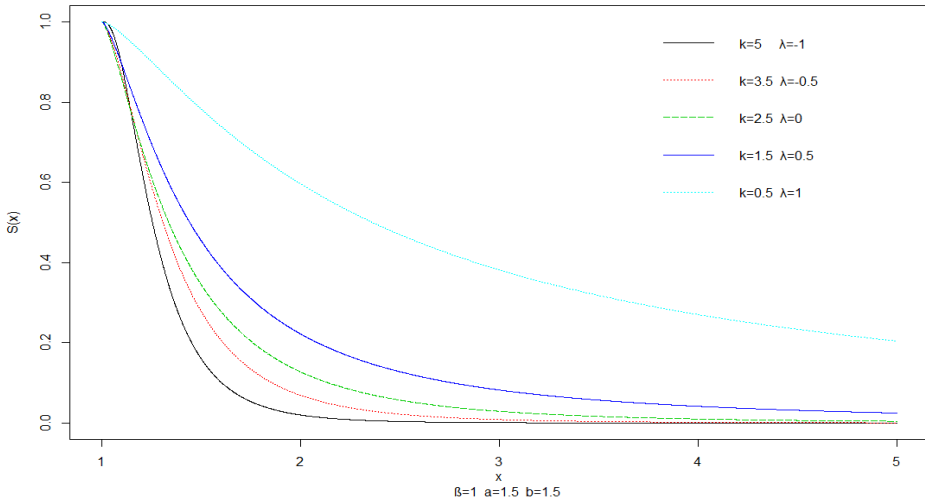


Figure 4. Survival plot of the TKPD.

2.1. Deductions from the TKPD Distribution

We shall deduce the following cases from the above distribution;

1. The TKPD becomes the Kumaraswamy Pareto distribution when the transmutation parameter $\lambda = 0$.

Indeed, consider the cdf of the TKPD in (2.1), expressed as

$$G(x) = (1 + \lambda) \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^{k-1} \right]^a \right\}^b \right\} - \lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\}^2.$$

Taking $\lambda = 0$, then we have

$$G(x) = 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^{k-1} \right]^a \right\}^b,$$

which is the same as the cdf of the Kumaraswamy Pareto distribution.

2. If $x \rightarrow \infty$, then TKPD becomes zero. i.e; $\lim_{x \rightarrow \infty} g(x) = 0$.

Indeed, consider the pdf of the TKPD expressed as

$$g(x) = \frac{abk\beta^k}{x^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} \right\} \\ \times \left[1 - \left(\frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}.$$

Now

$$\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} \frac{abk\beta^k}{x^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} \right\} \\ \times \left[1 - \left(\frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}.$$

Hence,

$$\lim_{x \rightarrow \infty} g(x) = 0. \quad (2.7)$$

3. As $x \rightarrow \beta$ and $a > 1$, then

$$\lim_{x \rightarrow \beta} g(x) = 0.$$

Consider the pdf of the TKPD distribution expressed as

$$g(x) = \frac{abk\beta^k}{x^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b \right\} \right\} \left[1 - \left(\frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}$$

Substituting β for x in the pdf gives

$$g(x) = \frac{abk\beta^k}{\beta^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{\beta} \right)^k \right]^a \right\}^b \right\} \right\} \left[1 - \left(\frac{\beta}{\beta} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{\beta} \right)^k \right]^a \right\}^{b-1}$$

Evaluating the expression gives

$$\lim_{x \rightarrow \beta} g(x) = 0$$

hence our claim has been established.

4. As $x \rightarrow \beta$ and $a = 1$, then

$$\lim_{x \rightarrow \beta} g(x) = \frac{(1 + \lambda)bk}{\beta}.$$

Consider the pdf of the TKPD distribution expressed as

$$g(x) = \frac{abk\beta^k}{x^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^{k\gamma^a} \right]^b \right\} \right\} \right\} \left[1 - \left(\frac{\beta}{x} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}$$

$a = 1$ in the pdf gives

$$g(x) = \frac{bk\beta^k}{x^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^b \right\} \right\} \right\} \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^{b-1}.$$

As $x \rightarrow \beta$

$$\lim_{x \rightarrow \beta} g(x) = \lim_{x \rightarrow \beta} \frac{bk\beta^k}{\beta^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{\beta} \right)^k \right]^b \right\} \right\} \right\} \left\{ 1 - \left[1 - \left(\frac{\beta}{\beta} \right)^k \right]^a \right\}^{b-1}$$

which further reduce to

$$g(x) = \frac{(1 + \lambda)bk}{\beta},$$

And hence the proof is complete.

2.2. Statistical properties of the proposed distribution

2.2.1. Quantile function of the TKPD

$$Q(p) = G^{-1} \left\{ \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda p}}{2\lambda} \right\}, \quad 0 < p < 1, \quad (2.8)$$

where $G^{-1}\{.\}$ is the inverse cdf or quantile function of the distribution $G(x)$.

Proof. The quantile function of the transmuted family of distribution can be obtained by solving the equation

$$G(Q(p)) = p$$

for $Q(p)$. This is done as follows:

$$G(Q(p)) = p$$

By Shaw and Buckley [24] we have

$$(1 + \lambda)F(Q(p)) - \lambda(F(Q(p)))^2 = p$$

$$\lambda(F(Q(p)))^2 - (1 + \lambda)F(Q(p)) + p = 0.$$

Solving the quadratic equation, we have

$$F(Q(p)) = \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda p}}{2\lambda},$$

and taking the inverse of both sides gives

$$Q(p) = F^{-1} \left\{ \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda p}}{2\lambda} \right\},$$

hence the proof is established.

From property 1, we have the following

Random variate vector X can be simulated from the transmuted family of distribution using the expression

$$X = G^{-1} \left\{ \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda U}}{2\lambda} \right\}, \quad (2.9)$$

where U is a uniform random variable defined on the interval $(0,1)$.

From the quantile property of the transmuted family of distribution of Shaw and Buckley, then we deduce the following;

The quantile function of the TKPD distribution is given as

$$Q(p) = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda p}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k} \quad (2.10)$$

$$0 < p < 1.$$

Proof. To prove this, we shall first find expression for $F^{-1}\{.\}$ and evaluate (2.5)

accordingly. Given

$$F(x) = 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b$$

$$F(x) = p$$

implies

$$1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b = p$$

$$\left\{ 1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a \right\}^b = 1 - p$$

$$1 - \left[1 - \left(\frac{\beta}{x} \right)^k \right]^a = (1 - p)^{1/b}$$

$$\left[1 - \left(\frac{\beta}{x} \right)^k \right]^a = 1 - (1 - p)^{1/b}$$

$$1 - \left(\frac{\beta}{x} \right)^k = \left[1 - (1 - p)^{1/b} \right]^{1/a}$$

$$\left(\frac{\beta}{x} \right)^k = 1 - \left[1 - (1 - p)^{1/b} \right]^{1/a}$$

$$\frac{\beta}{x} = \left\{ 1 - \left[1 - (1 - p)^{1/b} \right]^{1/a} \right\}^{1/k}$$

$$\frac{x}{\beta} = \left\{ 1 - \left[1 - (1 - p)^{1/b} \right]^{1/a} \right\}^{-1/k}$$

$$x = \beta \left\{ 1 - \left[1 - (1 - p)^{1/b} \right]^{1/a} \right\}^{-1/k}.$$

Thus

$$F^{-1}\{p\} = \beta \left\{ 1 - \left[1 - (1 - p)^{1/b} \right]^{1/a} \right\}^{-1/k}.$$

Using (2.5) gives the quantile function of the TKPD as

$$Q(p) = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda p}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k} \quad (2.11)$$

And hence the proof is established.

The first three quartiles of the TKPD distribution are given by

$$Q\left(\frac{1}{4}\right) = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - \lambda}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k}$$

$$Q\left(\frac{1}{2}\right) = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 2\lambda}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k}$$

$$Q\left(\frac{3}{4}\right) = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 3\lambda}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k} .$$

The second quartile $Q(1/2)$ corresponds to the median of the distribution. Thus

$$Q\left(\frac{1}{2}\right) = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 2\lambda}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k} . \quad (2.12)$$

Hence, The Random variate vector X can be simulated from the TKPD using the expression

$$X = \beta \left\{ 1 - \left[1 - \left(1 - \frac{(1 + \lambda) \pm \sqrt{(1 + \lambda)^2 - 4\lambda U}}{2\lambda} \right)^{1/b} \right]^{1/a} \right\}^{-1/k} , \quad (2.13)$$

where U is a uniform random variable defined on the interval $(0,1)$.

2.2.2. Moments of the TKPD distribution

The r^{th} non-central moments of the TKPD random variable X is given by

$$\mu'_r = E(X^r) = \int_{\beta}^{\infty} x^r g(x) dx \text{ for } x \in [\beta, \infty), \tag{2.14}$$

The r^{th} non-central moments can be expressed as

$$\mu'_r = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \int_{\beta}^{\infty} x^r k j \frac{\beta^{kj}}{x^{kj-1}} dx. \tag{2.15}$$

Observe that the integral represents the r^{th} non-central moments of the Pareto distribution with location parameter β and shape parameter kj . Thus the r^{th} non-central moments of the TKPD distribution is an infinite linear combination of the r^{th} non-central moments of the Pareto distribution. This is another useful and major result of this study.

Now if Y is a Pareto random variable, the r^{th} non-central moments of Y was given by Bourguignon et al. [8] as

$$\mu'_{Yr} = \int_{\beta}^{\infty} x^r k \frac{\beta^k}{x^{k+1}} dx = \begin{cases} \frac{k\beta^r}{k-r}; & \text{if } k > r \\ \infty; & \text{if } k = r \\ -\frac{k\beta^r}{k-r}; & \text{if } k < r \end{cases} \tag{2.16}$$

It follows that the r^{th} non-central moments of the TKPD is given by

$$\mu'_r = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \frac{kj\beta^r}{kj-r}, \quad kj > r. \tag{2.17}$$

The first four moment of the TKPD are given respectively by

$$\mu = \mu'_1 = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \frac{kj\beta}{kj-1}, \quad kj > 1, \tag{2.18}$$

$$\mu'_2 = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \frac{kj\beta^2}{kj-2}, \quad kj > 2, \tag{2.19}$$

$$\mu'_3 = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \frac{kj\beta^3}{kj-3}, \quad kj > 3, \tag{2.20}$$

$$\mu'_4 = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \frac{kj\beta^4}{kj-4}, \quad kj > 4. \quad (2.21)$$

The r^{th} central moments (μ_r) and cummulants (κ_r) of the TKPD can be obtained from (2.15) and expressed respectively as

$$\mu_r = \sum_{n=0}^r (-1)^n \binom{r}{n} \mu_1^n \mu'_{r-n}, \quad (2.22)$$

$$\kappa_r = \mu'_r - \sum_{n=1}^{r-1} \binom{r-1}{n-1} \kappa_n \mu'_{r-n}, \quad (2.23)$$

where $\kappa_1 = \mu'_1$.

The mean (μ), variance (σ^2), skewness (S) and kurtosis (K) of the TKPD is obtained from (2.21) and (2.22) and expressed respectively as

$$\mu = \mu'_1, \quad (2.24)$$

$$\sigma^2 = \mu^2 + \mu'_2 - 2\mu^2, \quad (2.25)$$

$$S = \frac{\mu'_3 - 3\mu\mu'_2 + 2\mu^2}{(\mu'_2 - \mu^2)^{\frac{3}{2}}}, \quad (2.26)$$

$$Ku = \frac{\mu'_4 - 4\mu\mu'_3 + 6\mu^2\mu'_2 - 3\mu^4}{(\mu'_2 - \mu^2)^2}. \quad (2.27)$$

The quantile function can also be used in evaluating skewness and kurtosis of a distribution particularly when the quantile function of the distribution exists in closed form. Galton [16] proposed a quantile measure based approach for evaluating skewness. while Moor [21] did the same for Kurtosis. Galton's skewness and Moor's kurtosis are evaluated using the relations

$$S = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)}, \quad (2.28)$$

$$K = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)}. \quad (2.29)$$

Since the Quantile function of the TKPD exists in closed form as given in (2.11), the above expressions can also be used in evaluating the skewness and kurtosis.

2.2.3. Moment generating function of the TKPD distribution

The moment generating function (mgf) of the TKPD from definition can be expressed as

$$M_X(t) = E(e^{tX}) = \int_{\beta}^{\infty} e^{tx} g(x) dx. \tag{2.30}$$

It follows that

$$M_X(t) = E(e^{tX}) = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} \int_{\beta}^{\infty} e^{tx} k j \frac{\beta^{kj}}{x^{kj+1}} dx. \tag{2.31}$$

Again the mgf of the TKPD is an infinite linear combination of the mgf of the Pareto distribution.

Now if Y is a Pareto random variable, the mgf of Y was given by Bourguignon et al. [9] as

$$M_X(t) = \int_{\beta}^{\infty} e^{tx} k \frac{\beta^k}{x^{k+1}} dx = k(-\beta t)^k \Gamma(-k, -\beta), \quad t \leq 0 \tag{2.32}$$

where $\Gamma(.,.)$ is the incomplete gamma function.

It follows that the mgf of the TKPD is given by

$$M_X(t) = (1 - \lambda) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} (-1)^{i+j} \binom{b}{i} \binom{ai}{j} k j (-\beta t)^{kj} \Gamma(-kj, -\beta), \quad t \leq 0. \tag{2.33}$$

2.2.4. Mode of the TKPD

The mode of the TKPD is obtained as the solution of

$$g'(x) = 0 \tag{2.34}$$

From (2.2), we realize that

$$g'(x) = (1 + \lambda)f'(x) - 2\lambda\{f(x)^2 + f'(x)F(x)\}. \tag{2.35}$$

Thus

$$g'(x) = 0$$

implies that

$$(1 + \lambda)f'(x) = 2\lambda\{f(x)^2 + f'(x)F(x)\}$$

$$\begin{aligned}
 f'(x) + \lambda f'(x) &= 2\lambda f(x)^2 + 2\lambda f'(x)F(x) \\
 f'(x) + \lambda f'(x) - 2\lambda f'(x)F(x) &= 2\lambda f(x)^2 \\
 (1 + \lambda - 2\lambda F(x))f'(x) &= 2\lambda f(x)^2
 \end{aligned} \tag{2.36}$$

Solving the above expression for x gives the mode of The TKPD and the TKPD as one mode hence it is a unimodal distribution. The above expression does not have an analytic form for the mode but the mode is determined numerically.

2.2.5. Mean deviation of the TKPD

The dispersion and the spread in a population from the center are often measured by the deviation from the mean, and the deviation from the median. The absolute mean deviation about the mean, $D(\mu)$, and the absolute mean deviation about the median, $D(M)$, for the TKPD are defined as

$$D(\mu) = \int_{-\infty}^{\infty} |x - \mu|g(x)dx, \tag{2.37}$$

and

$$D(M) = \int_{-\infty}^{\infty} |x - M|g(x)dx, \tag{2.38}$$

respectively, where $\mu = E(X)$ and $M = Q(1/2)$. Consequently

$$\begin{aligned}
 D(\mu) &= \int_{-\infty}^{\infty} |x - \mu|g(x)dx = \int_{-\infty}^{\mu} (\mu - x)g(x)dx + \int_{\mu}^{\infty} (x - \mu)g(x)dx \\
 &= \mu \int_{-\infty}^{\mu} g(x)dx - \int_{-\infty}^{\mu} xg(x)dx + \int_{\mu}^{\infty} xg(x)dx - \mu \int_{\mu}^{\infty} g(x)dx \\
 &= \mu G(\mu) - \int_{-\infty}^{\mu} xg(x)dx + \int_{\mu}^{\infty} xg(x)dx - \mu(1 - G(\mu)) \\
 &= \mu G(\mu) - \left(\int_{-\infty}^{\infty} xg(x)dx - \int_{\mu}^{\infty} xg(x)dx \right) + \int_{\mu}^{\infty} xg(x)dx - \mu(1 - G(\mu)) \\
 &= \mu G(\mu) - \int_{-\infty}^{\infty} xg(x)dx + \int_{\mu}^{\infty} xg(x)dx + \int_{\mu}^{\infty} xg(x)dx - \mu(1 - G(\mu)) \\
 &= \mu G(\mu) - \mu + 2 \int_{\mu}^{\infty} xg(x)dx - \mu(1 - G(\mu))
 \end{aligned}$$

$$= \mu G(\mu) - \mu + \mu G(\mu) - \mu + 2 \int_{\mu}^{\infty} xg(x)dx.$$

Thus,

$$D(\mu) = 2\mu G(\mu) - 2\mu + 2 \int_{\mu}^{\infty} xg(x)dx. \tag{2.39}$$

Also,

$$\begin{aligned} D(M) &= \int_{-\infty}^{\infty} |x - M|g(x)dx = \int_{-\infty}^M (M - x)g(x)dx + \int_M^{\infty} (x - M)g(x)dx \\ &= M \int_{-\infty}^M g(x)dx - \int_{-\infty}^M xg(x)dx + \int_M^{\infty} xg(x)dx - M \int_M^{\infty} g(x)dx \\ &= MG(M) - \int_{-\infty}^M xg(x)dx + \int_M^{\infty} xg(x)dx - M(1 - G(M)) \\ &= MG(M) - \left(\int_{-\infty}^{\infty} xg(x)dx - \int_M^{\infty} xg(x)dx \right) + \int_M^{\infty} xg(x)dx - M(1 - G(M)) \\ &= MG(M) - \int_{-\infty}^{\infty} xg(x)dx + \int_M^{\infty} xg(x)dx + \int_M^{\infty} xg(x)dx - M(1 - G(M)) \\ &= MG(M) - M + 2 \int_M^{\infty} xg(x)dx - M(1 - G(M)) \\ &= MG(M) - M + MG(M) - M + 2 \int_M^{\infty} xg(x)dx \\ &= 2MG(M) - M - M + 2 \int_M^{\infty} xg(x)dx \\ &= 2M \left(\frac{1}{2} \right) - M - M + 2 \int_M^{\infty} xg(x)dx. \end{aligned}$$

Thus

$$D(M) = -\mu + 2 \int_M^{\infty} xg(x)dx. \tag{2.40}$$

2.2.6. Order Statistics of the TKPD

Let X_1, X_2, \dots, X_n be a random sample of size n from the TKPD and suppose $X_{1:n} < X_{2:n} < \dots < X_{n:n}$ denote the corresponding order statistic. The pdf of the k^{th} order

statistic can be expressed as

$$f_{k:n}(x) = \frac{1}{B(k, n-k+1)} g(x)[G(x)]^{k-1}[1-G(x)]^{n-k}, \quad (2.41)$$

where, $B(.,.)$ is the beta function.

Using the binomial expansion, we have

$$[1-G(x)]^{n-k} = \sum_{i=0}^{n-k} (-1)^i \binom{n-k}{i} G(x)^i, \quad (2.42)$$

Thus

$$f_{k:n}(x) = \sum_{i=0}^{n-k} \frac{(-1)^i}{B(k, n-k+1)} \binom{n-k}{i} g(x)G(x)^{k+i-1}. \quad (2.43)$$

2.2.7. Shannon entropy of the TKPD

Shannon [23] defined the entropy of a random variable X as $E\{-\log(g(X))\}$, where $g(X)$ is the pdf of the random variable. The entropy of the random variable X measures variation of uncertainty. Let η_X be the entropy of a random variable X following the TKPD, we have

$$\eta_X = E\{-\log(g(X))\}. \quad (2.44)$$

$$\begin{aligned} \log(g(X)) &= \log\left(\frac{abk\beta^k}{X^{k+1}}\right) + \log\left(\left\{1 + \lambda - 2\lambda \left\{1 - \left\{1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right\}^b\right\}\right\}\right) \\ &+ (a-1)\log\left(1 - \left(\frac{\beta}{X}\right)^k\right) + (b-1)\log\left(1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right), \end{aligned} \quad (2.45)$$

$$\begin{aligned} \log(g(X)) &= \log(ab) + \log\left(\frac{k\beta^k}{X^{k+1}}\right) \\ &+ \log\left(\left\{1 + \lambda - 2\lambda \left\{1 - \left\{1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right\}^b\right\}\right\}\right) \\ &+ (a-1)\log\left(1 - \left(\frac{\beta}{X}\right)^k\right) + (b-1)\log\left(1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right). \end{aligned} \quad (2.46)$$

It follows that

$$\begin{aligned}
 -\log(g(X)) &= -\log(ab) - \log\left(\frac{k\beta^k}{X^{k+1}}\right) \\
 &\quad - \log\left(\left\{1 + \lambda - 2\lambda \left\{1 - \left\{1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right\}^b\right\}\right\}\right) \\
 &\quad - (a - 1)\log\left(1 - \left(\frac{\beta}{X}\right)^k\right) - (b - 1)\log\left(1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right). \tag{2.47}
 \end{aligned}$$

Hence

$$\begin{aligned}
 \eta_X &= E\{-\log(g(X))\} \\
 &= E\left\{-\log\left(\frac{k\beta^k}{X^{k+1}}\right)\right\} - \log(ab) \\
 &\quad + E\left\{-\log\left(\left\{1 + \lambda - 2\lambda \left\{1 - \left\{1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right\}^b\right\}\right\}\right)\right\} \\
 &\quad - (a - 1)E\left\{\log\left(1 - \left(\frac{\beta}{X}\right)^k\right)\right\} - (b - 1)E\left\{\log\left(1 - \left[1 - \left(\frac{\beta}{X}\right)^k\right]^a\right)\right\}. \tag{2.48}
 \end{aligned}$$

Observe that

$$E\left\{-\log\left(\frac{k\beta^k}{X^{k+1}}\right)\right\}.$$

This is the Shannon entropy of the Pareto distribution and was given by Bourguignon et al. [9] as

$$E\left\{-\log\left(\frac{k\beta^k}{X^{k+1}}\right)\right\} = \log\left(\left(\frac{\beta}{k}\right) e^{1+\frac{1}{k}}\right). \tag{2.49}$$

Thus the Shannon entropy of the of the TKPD is given by

$$\eta_X = \log\left(\left(\frac{\beta}{k}\right) e^{1+\frac{1}{k}}\right) - \log(ab) - (a - 1)E\left\{\log\left(1 - \left(\frac{\beta}{X}\right)^k\right)\right\}$$

$$\begin{aligned}
 &+E \left\{ -\log \left(\left(1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{X} \right)^k \right]^a \right\}^b \right\} \right) \right) \right\} \\
 &\quad - (b - 1)E \left\{ \log \left(1 - \left[1 - \left(\frac{\beta}{X} \right)^k \right]^a \right) \right\}, \tag{2.50}
 \end{aligned}$$

where the expectations are evaluated using standard procedure.

2.2.8. Maximum likelihood estimation

Suppose δ is a $p \times 1$ vector containing all the parameters of the TKPD, for a random sample, x_1, x_2, \dots, x_n of size n from the TKPD the total log-likelihood function is given by the TKPD, for a complete random sample x_1, x_2, \dots, x_n of size n from the TKPD the total log-likelihood function is given by

$$L = \sum_{i=1}^n \log(g(x_i)) \tag{2.51}$$

$$= \sum_{i=1}^n \log \left(\frac{abk\beta^k}{x_i^{k+1}} \left\{ 1 + \lambda - 2\lambda \left\{ 1 - \left\{ 1 - \left[1 - \left(\frac{\beta}{x_i} \right)^k \right]^a \right\}^b \right\} \right\} \left[1 - \left(\frac{\beta}{x_i} \right)^k \right]^{a-1} \left\{ 1 - \left[1 - \left(\frac{\beta}{x_i} \right)^k \right]^a \right\}^{b-1} \right). \tag{2.52}$$

Let

$$z_i = 1 - \left(\frac{\beta}{x_i} \right)^k. \tag{2.53}$$

It follows that

$$L = \sum_{i=1}^n \log \left(\frac{abk\beta^k}{x_i^{k+1}} \left\{ 1 + \lambda - 2\lambda \{ 1 - \{ 1 - z_i^a \}^b \} \right\} z_i^{a-1} \{ 1 - z_i^a \}^{b-1} \right). \tag{2.54}$$

Thus

$$\begin{aligned}
 L &= n\log(a) + n\log(b) + n\log(k) + nk\log(\beta) - (k + 1) \sum_{i=1}^n \log(x_i) \\
 &\quad + \sum_{i=1}^n \log(1 + \lambda - 2\lambda \{ 1 - \{ 1 - z_i^a \}^b \}) \\
 &\quad + (a - 1) \sum_{i=1}^n \log(z_i) + (b - 1) \sum_{i=1}^n \log(1 - z_i^a). \tag{2.55}
 \end{aligned}$$

Observe that the estimate of the parameter β corresponds to the first order statistic of the TKPD. That is $\beta = \min(X)$. We would only need to find the maximum likelihood estimates for the parameters a, b, k and λ .

Let $\Theta = (a \ b \ k \ \lambda)^T$ be the unknown parameter vector of TKPD. Then the associated score function is given by

$$U(\Theta) = \left(\frac{\partial L}{\partial a} \ \frac{\partial L}{\partial b} \ \frac{\partial L}{\partial k} \ \frac{\partial L}{\partial \lambda} \right)^T, \tag{2.56}$$

where $\frac{\partial L}{\partial a} \ \frac{\partial L}{\partial b} \ \frac{\partial L}{\partial k}$ and $\frac{\partial L}{\partial \lambda}$ are given by

$$\begin{aligned} \frac{\partial L}{\partial a} = & \frac{n}{a} - \sum_{i=1}^n \frac{2\lambda b \{1 - z_i^a\}^{b-1} z_i^a \log(z_i)}{1 + \lambda - 2\lambda \{1 - \{1 - z_i^a\}^b\}} \\ & + \sum_{i=1}^n \log(z_i) - (b - 1) \sum_{i=1}^n \frac{z_i^a \log(z_i)}{1 - z_i^a}, \end{aligned} \tag{2.57}$$

$$\frac{\partial L}{\partial b} = \frac{n}{b} - \sum_{i=1}^n \frac{2\lambda \{1 - z_i^a\}^b \log(1 - z_i^a)}{1 + \lambda - 2\lambda \{1 - \{1 - z_i^a\}^b\}} + \sum_{i=1}^n \log(1 - z_i^a), \tag{2.58}$$

$$\begin{aligned} \frac{\partial L}{\partial k} = & \frac{n}{k} + n \log(\beta) - \sum_{i=1}^n \log(x_i) + \sum_{i=1}^n \frac{2\lambda a b z_i^{a-1} \{1 - z_i^a\}^{b-1} \left(\frac{\beta}{x_i}\right)^k \log\left(\frac{\beta}{x_i}\right)}{1 + \lambda - 2\lambda \{1 - \{1 - z_i^a\}^b\}} \\ & - (a - 1) \sum_{i=1}^n \frac{\left(\frac{\beta}{x_i}\right)^k \log\left(\frac{\beta}{x_i}\right)}{z_i} - (b - 1) \sum_{i=1}^n \frac{a z_i^{a-1} \left(\frac{\beta}{x_i}\right)^k \log\left(\frac{\beta}{x_i}\right)}{1 - z_i^a}, \end{aligned} \tag{2.59}$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^n \frac{1 - 2\{1 - \{1 - z_i^a\}^b\}}{1 + \lambda - 2\lambda \{1 - \{1 - z_i^a\}^b\}}. \tag{2.60}$$

The maximum likelihood estimate of $\Theta, \hat{\Theta}$, can be obtained by solving the non-linear systems of equations, $\mathbf{U}(\Theta) = 0$. Since the resulting systems of equations are not in closed form, (An equation is in a closed form if it can be expressed in a simple algebraic form) the solutions can be found by numerical method using iterative scheme such as the Newton-Raphson type algorithms. The R Software was used to derive the maximum likelihood parameter estimates of the real-life dataset in Table 3.1 and displayed in Table 3.2.

For interval estimation of the parameters of theTKPD, one would require the Fisher information matrix (FIM) given by the 4×4 symmetric matrix

$$I_{i,j}(\hat{\Theta}) = -E_{\Theta} \begin{vmatrix} I_{aa} & I_{ab} & I_{ak} & I_{a\lambda} \\ I_{ba} & I_{bb} & I_{bk} & I_{b\lambda} \\ I_{ka} & I_{kb} & I_{kk} & I_{k\lambda} \\ I_{\lambda a} & I_{\lambda b} & I_{\lambda k} & I_{\lambda\lambda} \end{vmatrix}$$

where the elements $I_{i,j}(\Theta) = \left[\frac{\partial^2 \mathcal{L}}{\partial \theta_i \partial \theta_j} \right]$. Thus, the elements of the FIM can be obtained by considering the second order partial derivatives of the log-likelihood function w.r.t. to the parameters. These elements can be numerically obtained by using the R software. The total FIM, $I_{i,j}(\Theta)$, can be approximated by

$$I_{i,j}(\hat{\Theta}) \approx \left[- \frac{\partial^2 \mathcal{L}}{\partial \theta_i \partial \theta_j} \Big|_{\Theta = \hat{\Theta}} \right]_{4 \times 4}.$$

For real data, $I_{i,j}(\hat{\Theta})$ can be obtained after the maximum likelihood estimate of Θ is gotten, which implies the convergence of the iterative numerical procedure involved in finding such estimate.

Suppose $\hat{\Theta}$ is the maximum likelihood estimate of Θ . Under the usual regularity conditions and that the parameters are in the interior of the parameter space, but not on the boundary, we have: $\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{d} N_4(\mathbf{0}, I^{-1}(\Theta))$, where $I^{-1}(\Theta)$ is the inverse of the expected FIM, which also corresponds to the variance-covariance matrix of the parameters. The asymptotic behavior is still valid if $I^{-1}(\Theta)$ is replaced by the inverse of the observed information matrix evaluated at $\hat{\Theta}$, that is $I^{-1}(\hat{\Theta})$. The multivariate normal distribution with mean vector $\mathbf{0} = (0 \ 0 \ 0 \ 0)^T$ and covariance matrix $I^{-1}(\Theta)$ can be used to construct confidence intervals for the TKPD parameters. The approximate $100(1 - \theta)\%$ two-sided confidence interval for the parameters a, b, k, λ and s are given by

$$\hat{a} \pm Z_{\theta/2} \sqrt{I_{aa}^{-1}(\hat{\Theta})},$$

$$\hat{b} \pm Z_{\theta/2} \sqrt{I_{bb}^{-1}(\hat{\Theta})},$$

$$\hat{k} \pm Z_{\theta/2} \sqrt{I_{kk}^{-1}(\hat{\Theta})},$$

$$\hat{\lambda} \pm Z_{\theta/2} \sqrt{I_{\lambda\lambda}^{-1}(\hat{\theta})},$$

respectively, where $I_{aa}^{-1}(\hat{\theta})$, $I_{bb}^{-1}(\hat{\theta})$, $I_{kk}^{-1}(\hat{\theta})$ and $I_{\lambda\lambda}^{-1}(\hat{\theta})$ are diagonal elements of $I^{-1}(\hat{\theta})$ and $Z_{\theta/2}$ is the upper $(\theta/2)^{th}$ percentile of a standard normal distribution.

3. Application

For the application, we shall apply the proposed TKPD to a real life data. The data correspond to the exceedances of flood peaks (in m³/s) of the Wheaton River near Carcross in Yukon Territory, Canada. The data consist of 72 exceedances for the years 1958-1984, rounded to one decimal place. They were analyzed by Choulakian and Stephens [10] and are listed in Table 3.1. To aid the fitting of the data, the data was scaled by 100 and since scaling does not change the shape of a distribution, the result is as same as using the unscaled data. The purpose of scaling here is to improve the convergence rate of the numerical procedure used in obtaining the maximum likelihood estimates of the distribution. The density plots and the histogram of the fitted distributions in Figure 7 showed that the data is highly skewed to the right. We shall fit the TKPD to the data and compare the results with its sub-models namely: Pareto (P) and Kumaraswamy Pareto (KP) distributions. The computation of the estimates for the parameters of all the distributions was carried out using the R software.

Table 3.1. Exceedances of Wheaton River flood data.

1.7, 2.2, 14.4, 1.1, 0.4, 20.6, 5.3, 0.7, 1.9, 13.0, 12.0, 9.3, 1.4, 18.7, 8.5,
25.5, 11.6, 14.1, 22.1, 1.1, 2.5, 14.4, 1.7, 37.6, 0.6, 2.2, 39.0, 0.3, 15.0,
11.0, 7.3, 22.9, 1.7, 0.1, 1.1, 0.6, 9.0, 1.7, 7.0, 20.1, 0.4, 2.8, 14.1, 9.9, 10.4,
10.7, 30.0, 3.6, 5.6, 30.8, 13.3, 4.2, 25.5, 3.4, 11.9, 21.5, 27.6, 36.4, 2.7, 64.0,
1.5, 2.5, 27.4, 1.0, 27.1, 20.2, 16.8, 5.3, 9.7, 27.5, 2.5, 27.0

A descriptive statistics of the data set are given in Table 3.2.

Table 3.2. Descriptive statistics of the Exceedances of Wheaton River flood data.

Min	Q1	Q2	Mean	Q3	Max	Variance
0.100	2.125	9.500	12.200	20.120	64.000	151.221

The Maximum likelihood estimates of the parameters of all the fitted distributions alongside the loglikelihood value (loglik), Akaike Information Criterion (AIC) and the Kolmogorov-Smirnov (K-S) statistics are reported in Table 3.3.

Table 3.3. Maximum likelihood estimation of the distributions

Distribution	a	b	k	λ	β	loglik	AIC	K-S
TKPD	25.9190 (10.0747)	6.0721 (2.7302)	0.3054 (0.0568)	-0.5638 (0.0568)	0.1 -	-589.4	1186.8	0.1376
KP	29.5487 (10.2657)	8,6184 (4.6087)	0.2869 (0.0569)	0 -	0.1 -	-589.2	1184.4	0.1445
P	1 -	1 -	0.1149 (0.0135)	0 -	0.1 -	-688.8	1379.7	0.4698

(Standard error of estimates in parenthesis)

Observe from Table 3.3 that the maximum likelihood estimate of the parameter β corresponds to the minimum value of the data. This is so since $x \geq \beta$. Plots of the densities of the fitted distribution distributions alongside the histogram of the data are given by Figure 7. The cdf plots of all the fitted distributions are given by Figure 6 while the P-P plots of all the fitted distributions are given by Figure 7.

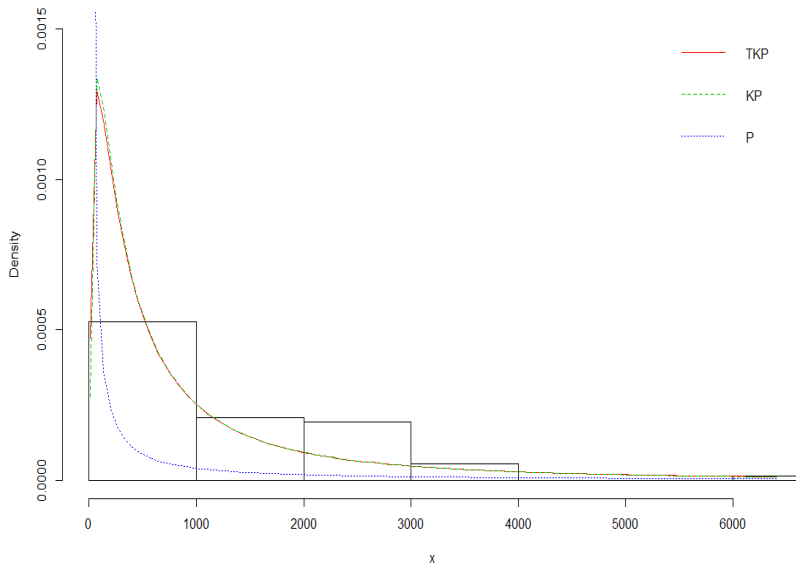


Figure 5. Density plots of the fitted distributions.

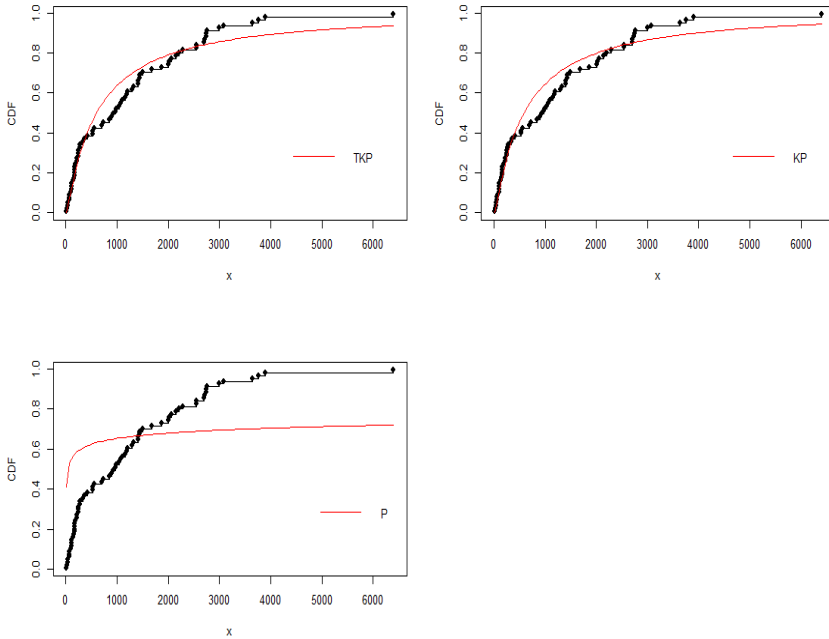


Figure 6. Cdf plots of all the fitted distributions.

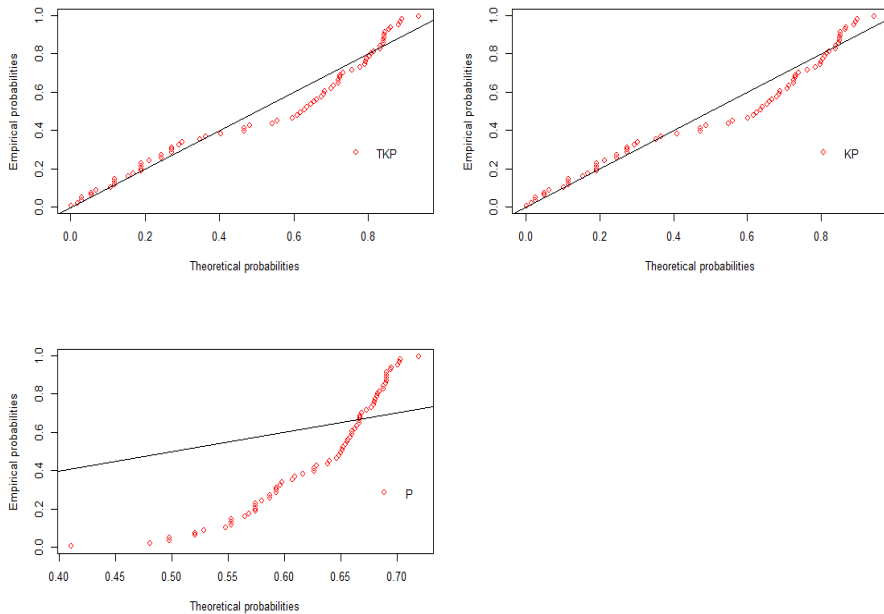


Figure 7. P-P plots of the fitted distributions.

Results in Table 7 clearly revealed the superiority of the proposed TKPD over the other distributions which are its sub-models. This is evident from the fact that the proposed TKPD possessed the smallest K-S statistics value. The graphs of the fitted distributions in Figure 5 also showed that the TKPD density fitted the histogram of the data better than the rest of them. The cdf plot of Figure 6 and P-P plots of Figure 7 further displayed the superiority of the TKPD over its sub-models. This clearly supports the fact that the generalization of the Pareto distribution as contained in the newly proposed distribution is a very useful one and will always come useful when the Pareto distribution fails.

4. Conclusions

The major conclusion that can be drawn from the results is that the combination of two or more distributions to form a compounded distribution function is an effective tool to deal with more real life datasets, especially when the population characteristics are many and requires many parameters in order to describe the pattern and behaviour of some random phenomenon. In this study, transmutation mapping technique was used to generate a univariate continuous probability distribution known as Transmuted Kumaraswamy Pareto Distribution. Actually, the baseline distribution of the proposed distribution is the combination of two classical distributions, the Kumaraswamy distribution and the Pareto distribution. Kumaraswamy Pareto distribution is also seen as a generalized family of the Pareto distribution using the Kumaraswamy distribution as the generator.

References

- [1] A. Akinsete, F. Famoye and C. Lee, The beta-Pareto distribution, *Statistics* 42 (2008), 547- 563. <https://doi.org/10.1080/02331880801983876>
- [2] G.R. Aryal and C.P. Tsokos, On the transmuted extreme value distribution with application, *Nonlinear Analysis* 71 (2009), 1401-1407. <https://doi.org/10.1016/j.na.2009.01.168>
- [3] G.R. Aryal and C.P. Tsokos, Transmuted Weibull distribution: A generalization of the Weibull probability distribution, *European Journal of Pure and Applied Mathematics* 4 (2011), 89-102.
- [4] S.K. Ashour and M.A. Eltehiwy, Transmuted Lomax distribution, *American Journal of Applied Mathematics and Statistics* 1 (2013a), 121-127. <https://doi.org/10.12691/ajams-1-6-3>

-
- [5] S.K. Ashour and M.A. Eltehiwy, Transmuted exponentiated Lomax distribution, *Australian Journal of Basic and Applied Sciences* 7 (2013b), 658-667.
- [6] S.K. Ashour and M.A. Eltehiwy, Transmuted exponentiated modified Weibull distribution, *International Journal of Basic and Applied Sciences* 2 (2013c), 258-269.
- [7] A. Azzalini, A class of distributions which includes the normal ones, *Candinavian Journal of Statistics* 12 (1985), 171-178.
- [8] M. Bourguignon, J. Leão, V. Leiva and M. Santos-Neto, The transmuted Birnbaum-Saunders distribution, *REVSTAT – Statistical Journal* 15 (2016), 601-628.
- [9] M. Bourguignon, R.B. Silva, L.M. Zea and G.M. Cordeiro, The Kumaraswamy Pareto distribution, *Journal of Statistical Theory and Applications* 12(2) (2012), 129-144.
<https://doi.org/10.2991/jsta.2013.12.2.1>
- [10] V. Choulakian and M.A. Stephen, Goodness-of-fit for generalized Pareto distribution *Technometrics* 43 (2011), 478-484. <https://doi.org/10.1198/00401700152672573>
- [11] I. Elbatal, Transmuted modified inverse Weibull distribution: A generalization of the modified inverse Weibull probability distribution, *International Journal of Mathematics Archive* 4 (2013a), 117-129.
- [12] I. Elbatal, Transmuted generalized inverted exponential distribution, *Economics and Quality Control* 28 (2013b), 125-133. <https://doi.org/10.1515/eqc-2013-0020>
- [13] I. Elbatal and G. Aryal, On the transmuted additive Weibull distribution, *Austrian Journal of Statistics* 42 (2013), 117-132. <https://doi.org/10.17713/ajs.v42i2.160>
- [14] I. Elbatal, L.S. Diab and N.A. Abdul Alim, Transmuted generalized linear exponential distribution, *International Journal of Computer Applications* 83 (2013), 29-37.
<https://doi.org/10.5120/14671-2681>
- [15] I. Elbatal and M. Elgarhy, Transmuted quasi-Lindley distribution: a generalization of the quasi-Lindley distribution, *International Journal of Pure and Applied Sciences and Technology* 18 (2013), 59-70
- [16] F. Galton, Colour associations, In F. Galton, *Inquiries into human faculty and its development* (p. 145-154), MacMillan Co., 1883. <https://doi.org/10.1037/14178-018>
- [17] M.S. Khan and R. King, Transmuted modified Weibull distribution, A generalization of the modified Weibull probability distribution, *European Journal of Pure and Applied Mathematics* 6 (2013), 66-88.
- [18] M.S. Khan, R. King and I.L. Hudson, Transmuted Kumaraswamy distribution, *Statistics in Transition New Series* 17 (2016), 183-210.
<https://doi.org/10.21307/stattrans-2016-013>
-

-
- [19] P. Kumaraswamy, A generalized probability density functions for double-bounded random processes, *Journal of Hydrology* 46 (1980), 79-88.
[https://doi.org/10.1016/0022-1694\(80\)90036-0](https://doi.org/10.1016/0022-1694(80)90036-0)
- [20] F. Merovci, Transmuted Rayleigh distribution, *Austrian Journal of Statistics* 42(2013), 21-31. <https://doi.org/10.17713/ajs.v42i1.163>
- [21] I.J. de Moor and M.N. Bruton, Atlas of alien and translocated indigenous aquatic animals in southern Africa, A Report of the Committee for Nature Conservation Research National Programme for Ecosystem Research, South African Scientific Programmes Report No. 144, 310 p., Port Elizabeth, South Africa, 1988.
- [22] P.E. Oguntude, O.S. Babatunde and A.O. Ogumola, Theoretical analysis of the Kumaraswamy-inverse exponential distribution, *Journal of Statistics and Applications* 4(2) (2014), 113-116.
- [23] C.E Shannon, A Mathematical Theory of Communication, *The Bell System Technical Journal* 27 (1948), 379-423, 623-656.
<https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
- [24] W.T. Shaw and I.R. Buckley, The alchemy of probability distributions: Beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map, 2007. <http://library.wolfram.com/infocenter/Articles/6670/alchemy.pdf>
- [25] K.U. Urama, S.I. Onyeagu and F.C. Eze, Shapes of the transmuted Kumaraswamy Pareto distribution for varying parameter values, *Mathlab J.* 5 (2020), 102-109.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted, use, distribution and reproduction in any medium, or format for any purpose, even commercially provided the work is properly cited.
