Shapes of the Transmuted Kumaraswamy Pareto Distribution for Varying Parameter Values

[1] K. U. Urama [2] S.I. Onyeagu [3]F. C. Eze

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

[2, 3] Department of Statistics, Nnamdi-Azikiwe University, Awka, Anambra State, Nigeria

Abstract

In this study, a new generalization of the Pareto distribution is undertaken, by first generalizing the Pareto distribution using the Kumaraswamy method and thereafter transmuting the resulting Kumaraswamy Pareto distribution. A detailed account of the general mathematical properties of the new generalized distribution is presented. The shapes of the Transmuted Kumaraswamy Pareto Density were plotted using R-program. The results show the superiority of the Transmuted Kumaraswamy Pareto distribution over the one parameter Pareto distribution.

Keywords: R-program; Transmutation; Kumaraswamy Distribution; Pareto Distribution

1. Introduction

Probability distribution has numerous applications in the description of real phenomenon. Over the decades, classical and theoretical probability distributions have been developed to model data in several fields of human endeavors such as engineering, medicine, actuarial sciences, finance, demography, and environmental sciences. Several probability distributions have been proposed for modeling lifetime data and environmental data such as Exponential, Lindley, Gamma, Weibull, Pareto, Kumaraswamy, lognormal and their generalizations.

In most practical situations, 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.

Bourguignon et al (2016) asserted that the generated families of probability distributions also called transmuted extended distributions consist of the parent distribution as a special case and other members of the family of the distribution. Most scholars really believed that the composition of two or more distributions enhances their flexibility in modeling various types of real life data especially where the parent model does not provide good fit.

Tahir and Nadarajah (2015) asserted that the addition of a shape parameter to the baseline distribution makes the distribution more flexible especially for studying the tail properties of probability distribution and also helps to improve the goodness of fit of the proposed model. They also observed that generalized distributio n family is capable to fit skewed data better than the base distribution.

Cordeiro and De Castro (2011) on the other hand, described the utility of Kumarasamay Generalized family of distribution as being able to generate effective models for censored data while the Pareto distribution is known for its ability to model long tailed data, or data having outliers or extreme values.

Though there are well known generators like the Marshal-Olkin family (MD-G) by Marshal and Olkin, (1997), Exponentiated family of distribution by Mudholkar and Srivastava(1993), the beta-Generated class of distribution by Eugene et al (2002), McDonald-G, (Mc-G) by Alexander et al (2012), the exponential family of distributions by Mudholkar and Sirivastava (1999), Kumaraswamy-G(Kw-G) by Cordeiro and de Castro(2011), Kumaraswamy's double bounded distribution by Kumaraswamy (1980) was described as a special case of Generalized Beta of first kind (GB1).

Transmuted Generalized family by Shaw and Buckley (2007) seems to be more popular because it compounds two or more distributions.

Within the last two decades, there have been various mechanisms for generalizing a probability distribution in other to allow it capture a wider range of shapes and in the process enhance its flexibility in modeling disparate data types. The works of Eugene et al (2002) is one of the leading works in this regard. Eugene et al. (2002) defined a generalized class of distribution called the beta-generated class and used the approach to develop the beta-normal distribution which can be unimodal and bimodal in shape. The work done by Eugene et al. (2002), attracted several generalizations of Probability distributions, using the beta distribution as a generator as seen in many Statistical literature, such as beta-Gumbel (Nadarajah and Kotz, 2004), beta-Frechet (Nadarajah and Gupta, 2004), beta-Weibull (Famoye et al 2005), beta-Exponential (Nadarajah and Kotz, 2006), beta-gamma (Kong et al 2007),beta-generalized exponential (Barreto-Souza et al 2010), beta-generalized half-normal (Pescim et al 2010), beta-Dagum (Condino and Domma, 2010), beta-Moyal (Cordeiro et al 2012), beta-Cauchy (Alshawarbeh et al 2012) and beta-Lindley (MirMostafaee et al 2015) distributions.

Aryal and Tsokos (2009) 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 its tests they carried out reviews that the data is well described by this distribution.

Aryal (2013) defined and studied the general properties of the transmuted log-logistic distribution. He derived various statistical and structural properties of the distribution including moments, quantiles, mean derivation and also estimated its parameters using the maximum likelihood method. He suggested that the subject distribution can model reliability data as well as lifetime data.

Merovci (2013) generalized the Rayleigh distribution using the transmutation map technique to develop the transmuted Rayleigh distribution. The Mathematical properties of the distribution were studied by the author along with its reliability behaviour. The researcher observed that the subject distribution can handle more complex data than the Rayleigh distribution. The performance of the transmuted Rayleigh distribution for nicotine measurement data was used to compare the two distribution models, applying in performance measure criteria like Kolgomorov Sminorv (K-S), Log Likelihood (-2L), Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC). Results obtained indicated that the transmuted Rayleigh distribution leads to a better fit than the Rayleigh distribution.

Mahmoud and Mandouh (2013) used the quadratic rank transmutation map (QRTM) to define a random variable X following the transmuted Frechet distribution. The authors used Frěchet distribution as the base distribution here. They studied the properties of this distribution and estimated the parameters of the model using maximum likelihood and Bayesian methods. They also applied transmuted Frechet distribution to two examples of uncensored data. The Kolmogorov –Smirnov goodness of fit test they carried out at 5% level of significance for the data, showed that there is no disagreement between the empirical distribution and the theoretical distribution; hence the empirical distribution fits the transmuted Frechet distribution.

Ashour and Eltehiwy (2013a) investigated the transmuted Lomax distribution. The authors also derived explicit expressions for moments, quantiles and mean deviation of the model and estimated the model parameters using the maximum likelihood estimate (mle) procedure.

Afify et al (2015) proposed the generalization of Weibull-Lomax distribution using transmutation map technique and called it Transmuted Weibull-Lomax (TWL) distribution. They derived the various structural and statistical properties of the model like the ordinary and incomplete moments, quantiles and generating functions, probability weighted moments, Renyi, q-entropies and order statistics. The method of maximum likelihood was used to obtain the model parameters. When compared with Weibull Lomax and other related models using some performance measures, the result showed that Transmuted Weibull Lomax could be chosen as the best model. They concluded that the proposed model will attract wider applications in areas such as Engineering, Survival and lifetime data, meteorology, hydrology, Economics (income inequality) and others.

Ashour and Eltehiwy (2013b) proposed and studied a generalization of the Exponentiated Lomax distribution known as the transmuted Exponentiated Lomax distribution using the transmutation techniques. 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.

Ashour and Eltehiwy (2013c) studied the Transmuted Exponentiated Modified Weibull distribution taking the Exponentiated Modified Weibull as the baseline distribution. 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.

Khan et al (2016) extended the Kumaraswamy distribution (KW) to study the Transmuted Kumaraswamy distribution by using the quadratic rank transmutation map technique. They provided two dataset to access the goodness- of- fit of the Transmuted Kumaraswamy distribution. They compared the Kumaraswamy distribution and transmuted Kumaraswamy distribution for analysis of flood data using the Kolmogorov-Smirnov (K-S) test, Cramer-Von Mises and Anderson-Darling goodness- of- fit statistics for the flood data and found that the Transmuted Kumaraswamy distribution provides an adequate fit for the flood data and in the second dataset for HIV/AIDS data; the result also showed that the TKW distribution provided better fit for the HIV/AIDS data. In order to compare the distribution, they considered the Akaike Information Criterion (AIC) for the HIV/AIDS data.

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. However, the performance in modeling datasets exhibiting extreme value properties was remarked by Bourguignon, Silva, Zea, and, Coreiro,(2013). 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.

2. Methodology

In this section, a family of probability distributions known as Transmuted Kumaraswamy Pareto distribution (TKPD) is proposed. In the definition of this Transmuted distribution we adopted Transmutation map technique as for generalizing Pareto distribution. The transmutation approach involves the transformation of a given baseline probability distribution.

The need to introduce a parametric that is both flexible and tractable with a wide application to skewed data spurred Shaw and Buckley (2007) to propose the *quadratic rank transmutation map* (QRTM), used by many researchers in the literature to define new family of Probability distributions. The Transmutation approach considers a random variable X with cdf $F(x)$ and a corresponding parameter λ to define a new family of transmuted Probability distribution by:

$$G(x)=(1+\lambda)F(x)-\lambda\left(F(x)\right)^2, \lambda\le 1 \quad (2.1)$$

where, when λ assumes the value 0,equation (2.1) reduces to the baseline distribution $F(x)$ otherwise, equation (2.1) assumes a quadratic function of the baseline distribution $F(x)$.

The corresponding pdf to (2.1) is obtained by differentiating (2.1) with respect to x to obtain

$$g(x)=(1+\lambda)f(x)-2\lambda f(x)F(x),\lambda\le 1$$

This expression factorizes to

$$g(x)=f(x)\left[1+\lambda-2\lambda F(x)\right],|\lambda|\le 1 \quad (2.2)$$

by differentiation

$$f(x)=F'(x) \quad (2.3)$$

Since the definition of equation (2.1) by Shaw and Buckley (2007), various transmuted family of probability distributions have appeared in the literature by simply replacing $F(x)$ in equation (2.1) with the cdf of another probability distribution.

2.1 The Transmuted Kumaraswamy Pareto distribution

A random variable, X is said to follow the Pareto distribution if it has the cumulative distribution function (cdf) and probability density function (pdf) given respectively by:

$$G(x)=1-\left(\frac{\beta}{x}\right)^{k},x\ge\beta,\beta>0,k>0 \quad (2.4)$$

$$g(x)=\frac{k\beta^{k}}{x^{k+1}},x\ge\beta,\beta>0,k>0 \quad (2.5)$$

Where, x is the minimum possible value of X and β is a positive parameter and shape parameters respectively.

The Kumaraswamy distribution (Kumaraswamy, 1980) has cdf and pdf defined respectively as

$$G(x)=1-(1-x^{a})^{b},0<x<1,a,b,>0 \quad (2.6)$$

$$g(x)=abx^{a-1}(1-x^{a})^{b-1},0<x<1,a,b,>0 \quad (2.7)$$

Where a and b are shape parameters respectively.

Bourguignon et al. (2012) defined the Kumaraswamy Pareto distribution by the cdf and pdf of the forms given respectively by

$$F(x)=1-\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a}\right\}^{b}, \quad (2.8)$$

$$x \geq \beta, a, b, \beta, k > 0$$

$$f(x) = \frac{abk\beta^{k}}{x^{k+1}}\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a-1}\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a}\right\}^{b-1} \quad (2.9)$$

$$x \geq \beta; a, b, \beta, k > 0$$

Where a, b and k controls the shape of the distribution and β a location parameter.

Substituting equation (2.8) into (2.1), to obtain a new family of probability distribution called the transmuted Kumaraswamy Pareto distribution (TKPD) whose cdf is of the form:

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

which can be factorized to get,

$$G(x) = (1+\lambda) - \left\{1-\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a}\right\}^{b}\right\}^{2} \quad (2.10)$$

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

The corresponding pdf of the TKPD is obtained by substituting (2.8) and (2.9) into (2.2) to obtain:

$$g(x) = \frac{abk\beta^{k}}{x^{k+1}}\left\{(1+\lambda) - 2\lambda\left\{1-\left\{1-\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a}\right\}^{b}\right\}\right\}\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a-1} \left\{1-\left[1-\left(\frac{\beta}{x}\right)^{k}\right]^{a}\right\}^{b-1} \quad (2.11)$$

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

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

3 Shapes of the Transmuted Kumaraswamy Pareto Density

In order to see the beauty of the shapes of the Transmuted Kumaraswamy Pareto Density, we used the R programe to plot the shapes with varying parameters. The shapes are displayed in figure 1 and 2.

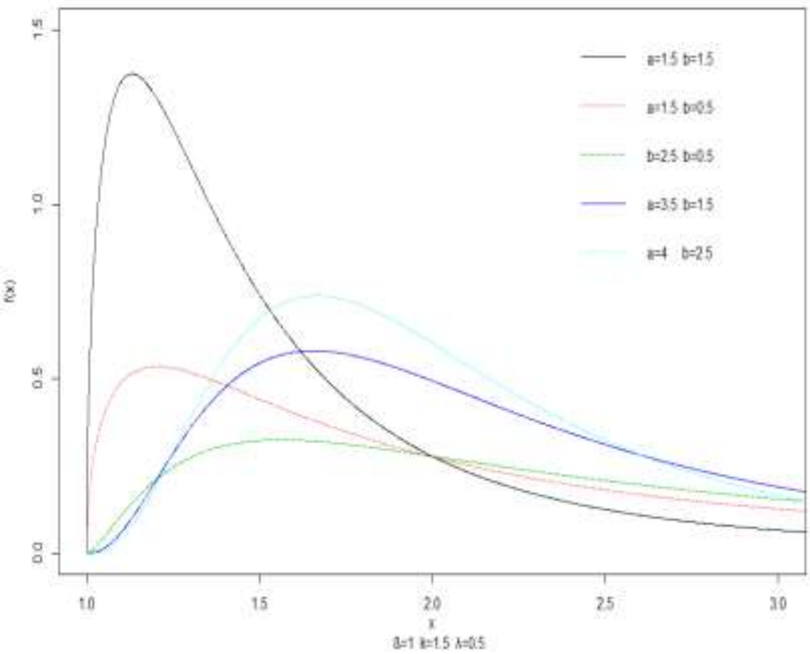

Figure 1: The TKPD density for varying parameters values

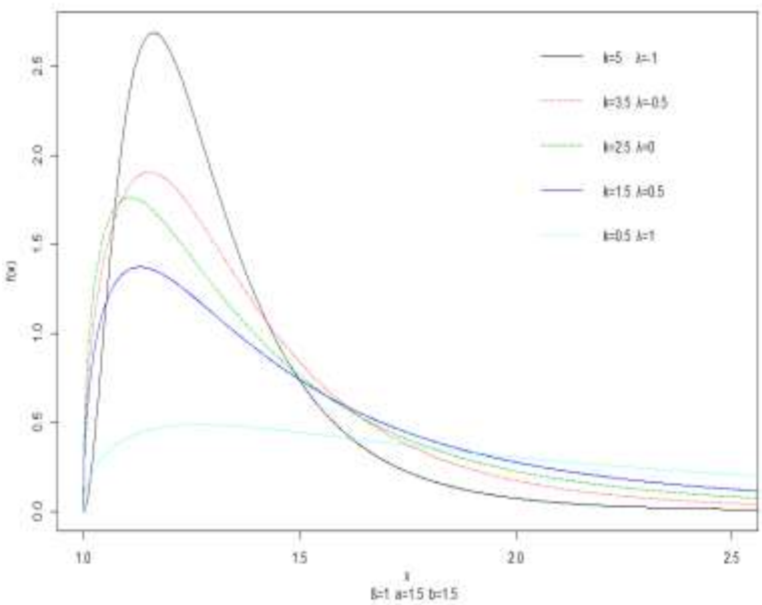

Figure 2: TKPD density for varying parameter values

From Figures 1 and 2, it can be observed that the TKPD is right-skewed and unimodal. Figures 1 and 2 also showed that as the value of the shape parameter k increases, the TKPD distribution becomes highly peaked. Again for increasing values of the parameter a the peak of the distribution flattens. The graphs of figures 1 and 2 showed that when the values of the parameter λ increase the TKPD also flattens in its peaks. This situation points to the flexibility of the TKPD distribution over the Pareto distribution whose shape is only controlled by one shape parameter.

4. Conclusion

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.

The plots of the TKPD density function showed that it is right skewed and unimodal. When the shape parameter value increases the function becomes highly peaked, and when the value of the transmuted parameter increases the peak flattens. This shows the superiority of the compound distribution, the Transmuted Kumaraswamy Pareto Distribution, (TKPD), over the one parameter Pareto distribution whose shape is only controlled by only one parameter.

References

1. Afify A.Z., Nofal Z.M., Yousef H.M., El Gebaly Y.M. and Butt N.S. (2015b). The transmuted Weibull-Lomax distribution: properties and Application. *Pakistan Journal of Statistics and Operation Research*, **XI**, 135 – 152.

2. Alshawarbeh E., Lee C. and Famoye F. (2012).Beta Cauchy distribution.*Journal of Probability and Statistical Science,* **10**, 41 – 57.

3. Aryal G.R. (2013). Transmuted log-logistics distribution,*Journal of Statistics Applications & Probability*, 2, 11 – 20.

4. Aryal G.R. and Tsokos C.P. (2009).On the transmuted extreme value distribution with application,*Nonlinear Analysis,***71,** 1401 - 1407.

5. Ashour S.K. and Eltehiwy M.A. (2013a).Transmuted Lomax distribution,*AmericanJournal of Applied Mathematics and Statistics,***1**, 121 – 127.

6. Ashour S.K. and Eltehiwy M.A. (2013b). Transmuted exponentiated Lomax distribution, *Australian Journal of Basic and Applied Sciences*, **7**, 658 – 667.

7. Ashour S.K. and Eltehiwy M.A. (2013c). Transmuted Exponentiated modified Weibull distribution,*International Journal of Basic and Applied Sciences*, **2,** 258 – 269.

8. Barreto-Souza W., Santos A.H.S. and Cordeiro G.M.(2010). The Beta generalized exponential distribution, *Journal of Statistical Computation and Simulation*, **80**, 159 - 172.

9. Bourguignon M., Ghosh I. and Cordeiro G.M. (2016).General results for the transmuted family of distributions and new models.*Journal of Probability and Statistics.*http://dx.doi.org/10.1155/2016/7208425

10. Condino F.and Domma F. (2010). The beta Dagum distribution. Presented at the *45th Scientific Meeting of the Italian Statistical Society*, University of Padua, June 15 – June 18.

11. Cordeiro G.M. and Castro De M. (2011). A new family of generalized distributions.*Journal of Statistical Computation and Simulation*, **81**, 883-893.

12. Cordeiro G.M., Nobre J.S., Pescim R.R. and Edwin M.M.(2012). The beta Moyal: A useful skew distribution,*International Journal of Research and Revises in Applied Sciences*, **10**, 171 – 192.

13. Eugene N., Lee C.and Famoye F.(2002).Beta-normal distribution and its applications, *Communications in Statistics - Theory & Methods,* **31**, 497-512.

14. Famoye F., Lee C. and Olumolade O.(2005).The beta-Weibull distribution, *Journal of Statistical Theory and Applications*, **4**, 121-136.

15. Khan M. S., King R. and Hudson I.L.(2016).Transmuted Kumaraswamy distribution. *Statistics In Transition – new series,* **17**, 183 – 210.

16. Kong L., Lee C. and Sepanski J.H. (2007).On the properties of beta-gamma distribution, *Journal of Modern Applied Statistical Methods*, **6**, 187 - 211.

17. Kumaraswamy P. (1980). A generalized probability density functions for double-bounded random processes.*Journal of Hydrology*, **46**, 79-88.

18. Mahmoud M.R. and Mandouh R.M. (2013).On the transmuted Frechet distribution,*Journal of Applied Sciences Research*, **9**, 5553 – 5561.

19. Marshall A.W. and Olkin I. (1997).A new method for adding a parameter to a family of distributions with application to the Exponential and Weibull families. *Biometrika,***84**,641-652.

20. Merovci F. (2013).Transmuted Rayleigh distribution,*Austrian Journal of Statistics,***42**, 21 – 31.

21. MirMostafaee S.M.T.K., Mahdizadeh M. and Nadarajah S. (2015). The beta Lindley distribution, *Journal of Data Science,***13**, 603 - 626.

22. Mudholkar, G.S and Srivastava,D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure rate,IEEE Transact. Reliability, 42 (2), 299-302.

23. Nadarajah S. and Kotz S.(2004).The beta-Gumbel distribution, *Mathematical Problems in Engineering*, **4**, 323-332.

24. Nadarajah S. and Gupta A.K.(2004).The beta-Frechet distribution, *Far East Journal of Theoretical Statistics*, **14**, 15 – 24.

25. Nadarajah S. and Kotz S.(2006).The beta exponential distribution, *Reliability Engineering and System Safety*, **91**, 689-697.

26. Pescim R.R., Demetrio C.B.G., Cordeiro G.M., OrtegaE.M.M. and Urbano M.R. (2010).The beta generalized half-normal distribution. *Computational Statistics and Data Analysis*, **54**, 945 – 957.

27. Shaw W.T. and Buckley I.R. (2007).The alchemy of probability distributions: Beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. http://library.wolfram.com/infocenter/Articles/6670/alchemy.pdf

28. Tahir M.H. and Nadarajah S.(2015). Parameter Induction in Continouos Univariate distributions.Well-established G-Families.Anais da Academia Brasileira de Ciências **87**(2);539-568.(Annals of the Brazilian Academy of Sciences). http://dx.doi.org/10.1590/0001-376520140299www.scielo.br/aabc