Transmuted Generalized Pareto Distribution

M. E. Habib, [1, 2] A. M. Abd Elfattah[3,4] and Abdelhamid Eissa [5]
[1]Professor of Statistics, Department of Statistic,
Faculty of Commerce, Al-Azhar University
[2]m.habib2000@yahoo.com [3]a_afattah@hotmail.com
[4]Professor of Statistics,
Institute of Statistical studies and research
[5]Abdo.easa2012@gmail.com

Abstract: In this article, we generalize the generalized Pareto Distribution using the quadratic transmuted map studied by Shaw et al. (2009). The properties of this distribution are derived and the estimation of the model parameters is discussed, estimating parameters by the method of maximum likelihood and provide the observed information matrix. The flexibility of the new model is illustrated with Monte Carlo simulation.

1- Introduction

The quality of the procedures used in a statistical analysis depends heavily on the assumed probability model or distributions. Because of this, considerable effort has been expended in the development of large classes of standard probability distributions along with relevant statistical methodologies. However, there still remain many important problems where the real data does not follow any of the classical or standard probability models.

The four parameters Pareto (generalized Pareto) distribution was introduced by (Abdul Fattah et el [2007]). The cumulative distribution function cdf of the four parameters Pareto distribution (generalized Pareto distribution) is

$$F(x;\alpha,\beta,\eta,\delta)=1-\left(1+\left(\frac{x-\eta}{\beta}\right)^{\delta}\right)^{-\alpha} \qquad x>\eta \qquad (1)$$

A random variable X is said to follow the Pareto distribution with four parameters (generalized Pareto distribution) if the probability density function pdf of X is as follows:

$$f(x;\alpha,\beta,\eta,\delta)=\frac{\delta\alpha}{\beta}\left(\frac{x-\eta}{\beta}\right)^{\delta-1}\left(1+\left(\frac{x-\eta}{\beta}\right)^{\delta}\right)^{-(\alpha+1)} \tag{2}$$

For $x>\eta$, $(\alpha,\beta,\delta)>0$, $\eta\geq 0$ where η is the location parameter, β is scale parameter and α *and* δ are the shape parameter

In this article we use transmutation map approach suggested by Shaw et al. (2009) to define a new model which generalizes the exponentiated exponential model. We will call Generalized Pareto distribution as the Transmuted Generalized Pareto distribution and (TGP) distribution for simplicity. According to the quadratic rank transmutation map (QRTM), Approach the cumulative distribution function (cdf) satisfy the relationship

$$F_T(x)=(1+\lambda)G(x)-\lambda G^2(x), \quad |\lambda|\leq 1 \tag{3}$$

which on differentiation yields,

$$f_T(x)=g(x)\left[1+\lambda-2\lambda G(x)\right], \quad |\lambda|\leq 1 \tag{4}$$

where $G(x)$ is the cdf of the base distribution, Observe that at $\lambda=0$; we have the distribution of the base random variable. Aryal et al. (2009) studied the transmuted Gumbel distribution and it has been observed that transmuted Gumbel distribution can be used to model climate data. In the present study we will provide mathematical formulations of the (TGP) distribution and also some of its properties.

2. Transmuted Generalized Pareto Distribution

A random variable X is said to have the (TGP) distribution with parameter $\underline{\tau} = (\alpha, \beta, \lambda, \delta)^T$ *and* $\eta = 0$, if its probability density is defined as:

$$f_{TGP}(x) = \frac{\delta\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\delta-1}\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-(\alpha+1)}\left\{1-\lambda+2\lambda\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right\} \quad x>0 \qquad (5)$$

and its cdf is

$$F_{TGP}(x) = (1+\lambda)\left(1-\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right)-\lambda\left(1-\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right)^{2} \quad x>0 \qquad (6)$$

For $x>0$, $(\alpha,\beta,\delta)>0$, $|\lambda|\leq 1$ where λ is the trunsmuted parameter, β is scale parameter and α *and* δ are the shape parameter

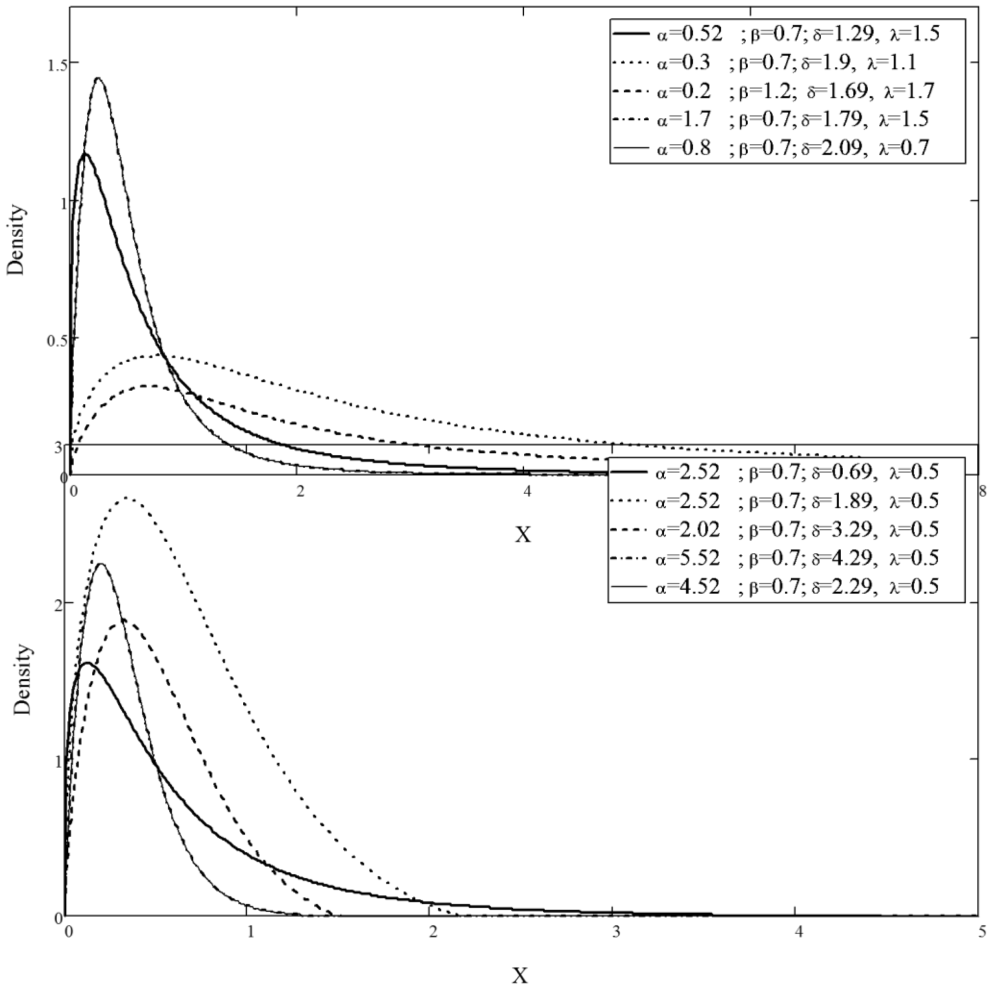
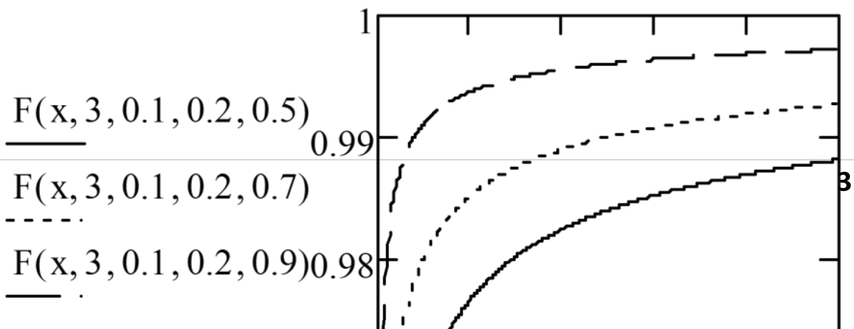


Figure 1: Plots of the TGP density function for some parameter values of transmuted and shape parameters.



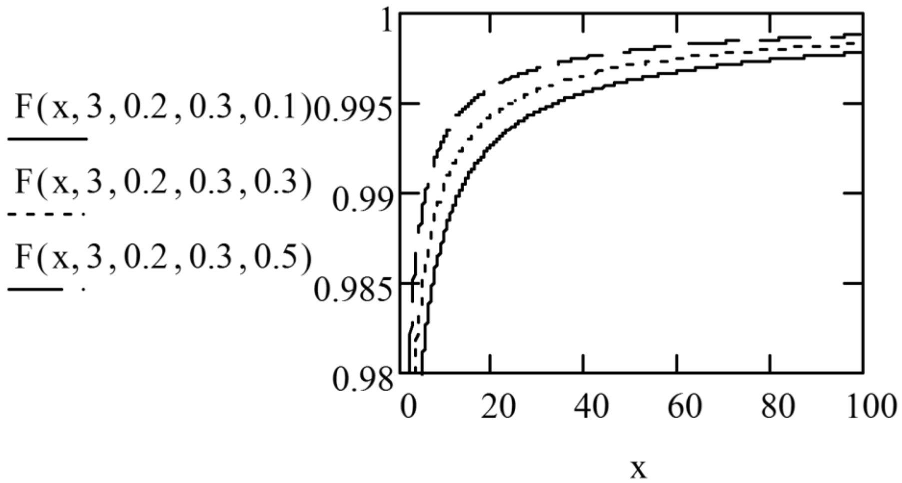


Figure 2: Plots of the TGP cumulative function for some parameter values of transmuted and shape parameters.

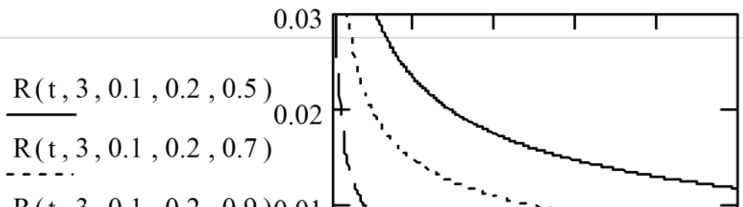
3- RELIABILITY ANALYSIS

The transmuted Generalized Pareto Distribution "TGPD" can be a useful characterization of life time data analysis. The reliability function of the TGPD is denoted by R_{TGP} also known as the survivor function and is defined as:

$$R_{TGP}(x;\underline{\tau}) = 1 - F_{TGP}(x;\underline{\tau}) = 1 - \left\{(1+\lambda)\left(1-\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right) - \lambda\left(1-\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right)^{2}\right\} \quad (7)$$

It is important to note that $R_{TGP} + F_{TGP}(x;\underline{\tau}) = 1$, One of the characteristic in reliability analysis is the hazard rate function (HRF) defined by

$$h_{TGP}(x;\underline{\tau}) = \frac{\frac{\delta\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\delta-1}\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-(\alpha+1)}\left\{1-\lambda+2\lambda\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right\}}{1-\left\{(1+\lambda)\left(1-\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right)-\lambda\left(1-\left(1+\left(\frac{x}{\beta}\right)^{\delta}\right)^{-\alpha}\right)^{2}\right\}} \quad (8)$$



`

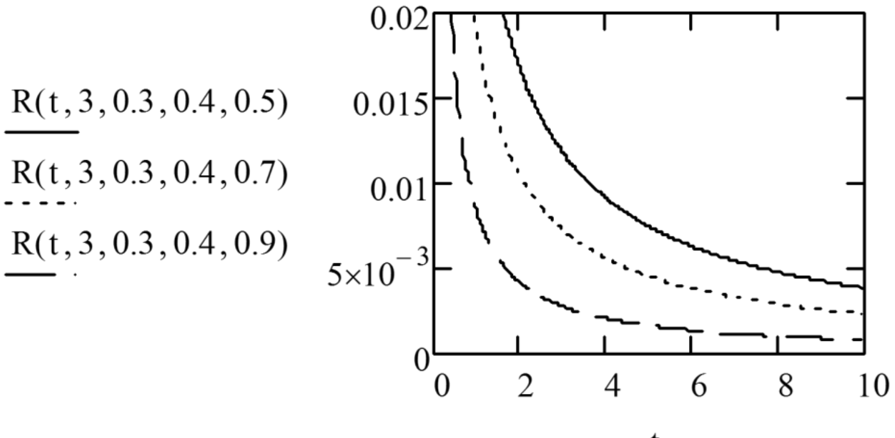


Figure 3: Plots of the TGP Reliability function for some parameter values of transmuted and shape parameters.

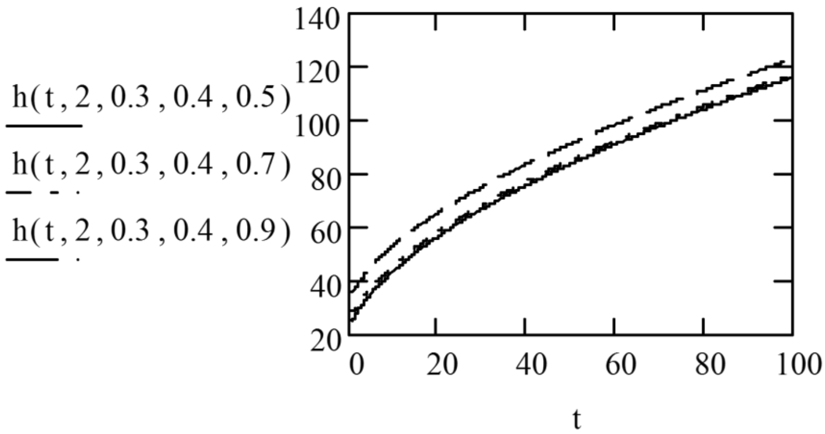


Figure 4: Plots of the TGP Hazard Rate function for some parameter values of transmuted and shape parameters.

4. STATISTICAL PROPERTIES

In this section we discuss the statistical properties of the transmuted Generalized Pareto Distribution "TGPD", specifically Quantile function, median, moments, and moment generating function.

4.1 Moments

$$E(x)^r = -\frac{\delta\alpha}{\beta}\int_0^1 \left[\beta\left[u^{\frac{-1}{\alpha}}-1\right]^{\frac{1}{\delta}}\right]^r [u]^{1+\frac{1}{\alpha}}[(1+\lambda)-2\lambda[1-u]]dx$$

Put $u = \left[1+\left(\frac{x}{\beta}\right)^{\delta}\right]^{-\alpha}$ $dx = \frac{1}{\mathrm{c}}\,\beta\left[u^{\frac{-1}{\alpha}}-1\right]^{\frac{1}{\delta}-1}\left(\frac{-1}{\alpha}\right)u^{-\left(\frac{1}{\alpha}+1\right)}du$

$$E(x)^r = \beta^r\int_0^1\left[u^{\frac{-1}{\alpha}}-1\right]^{\frac{r}{\delta}}[2\lambda[1-u]-(1+\lambda)]du$$

Using the binomial theorem will be

$$\left[u^{\frac{-1}{\alpha}}-1\right]^{\frac{r}{\delta}} = \sum_{i=0}^{\mathrm{r}}(-1)^i\binom{\frac{r}{\delta}}{\mathrm{i}}\left(u^{\frac{-1}{\alpha}}\right)^{\mathrm{i}}$$

$$E(x)^r = \beta^r\sum_{i=0}^{\mathrm{r}}(-1)\binom{\frac{r}{\delta}}{\mathrm{i}}\int_0^1 u^{\frac{-i}{\alpha}}[2\lambda[1-u]-(1+\lambda)]du$$

$$E(x)^r = \beta^r\sum_{i=0}^{\mathrm{r}}(-1)\binom{\frac{r}{\delta}}{\mathrm{i}}\left[\int_0^1 2\lambda u^{\frac{-i}{\alpha}}[1-u]\,du-\int_0^1(1+\lambda)u^{\frac{-i}{\alpha}}\,du\right]$$

$$E(x)^r = \beta^r\sum_{i=0}^{\mathrm{r}}(-1)^i\binom{\frac{r}{\delta}}{\mathrm{i}}\left[\left[2\lambda\mathrm{B}\left[\left(1-\frac{i}{\alpha}\right),2\right]\right]-\left[(1+\lambda)\mathrm{B}\left[\left(1-\frac{i}{\alpha}\right),1\right]\right]\right] \tag{9}$$

$$E(x) = \frac{\beta}{\delta}\left[(1+\lambda)\mathrm{B}\left[\left(1-\frac{1}{\alpha}\right),1\right]-2\lambda\mathrm{B}\left[\left(1-\frac{1}{\alpha}\right),2\right]\right]-1$$

$$E(x)^2 = \left|\begin{array}{l}\frac{2\beta^2}{\delta}\left\{(1+\lambda)\mathrm{B}\left(\left(1-\frac{1}{\alpha}\right),1\right)-2\lambda\mathrm{B}\left(\left(1-\frac{1}{\alpha}\right),2\right)\right\}+\\ \frac{\beta^2}{\delta}\left(\frac{2}{\delta}\right)\left\{2\lambda\mathrm{B}\left(\left(1-\frac{2}{\alpha}\right),2\right)-(1+\lambda)\mathrm{B}\left(\left(1-\frac{2}{\alpha}\right),1\right)\right\}-\beta^2\end{array}\right.$$

4.2 Variance

Using the appropriate moment expressions, the variance is given as follows:

$$Var(X) = E(X^2)-(E(X))^2 \tag{10}$$

4.3 Quantiles

The q^{th} quantile x_q of the transmuted G-Pareto Distribution "TGPD" can be obtain from (3),

$$x_q = \beta\left\{\left(1+\frac{(1+\lambda)+\sqrt{(1+\lambda)^2-4\lambda q}}{2\lambda}\right)^{-\frac{1}{\alpha}}-1\right\}^{\frac{1}{\delta}} \tag{11}$$

Hence, the distribution median is

$$x_{0.5} = \beta \left\{ \left(1 + \frac{(1+\lambda) + \sqrt{1+\lambda^2}}{2\lambda} \right)^{-\frac{1}{\alpha}} - 1 \right\}^{\frac{1}{\delta}}$$

4.2 Skewness and Kurtosis

The shortcomings of the classical kurtosis measure are well-known. There are many heavy tailed distributions for which this measure is infinite. So, it becomes uninformative precisely when it needs to be. Indeed, our motivation to use quantile-based measures stemmed from the non-existance of classical kurtosis for many of the Trunsmuted distributions; the Bowley's skewness (see Kenney and Keeping) is based on quartiles:

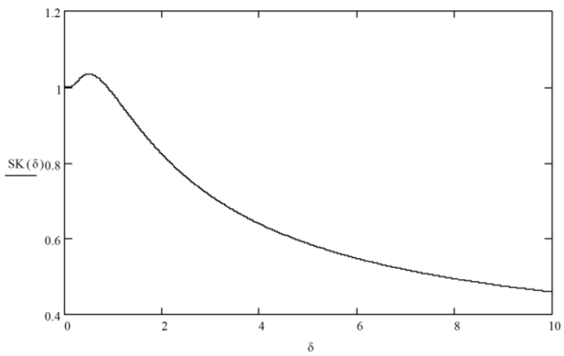
$$S_k = \frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)} \qquad (12)$$

And the Moors' kurtosis (see Moors (28)) is based on octiles:

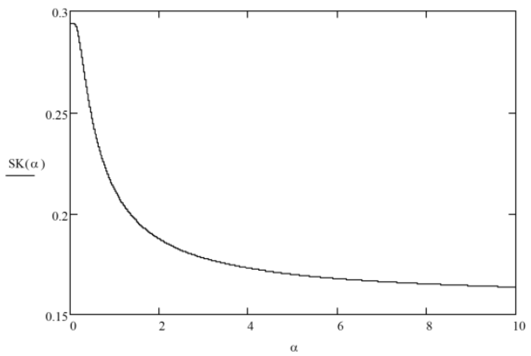
$$K_u = \frac{Q(7/8) - Q(5/8) - Q(3/8) + Q(1/8)}{Q(6/8) - Q(2/8)} \qquad (13)$$

Where $Q(\cdot)$ represents the quantile function

(a)



(b)



(c)

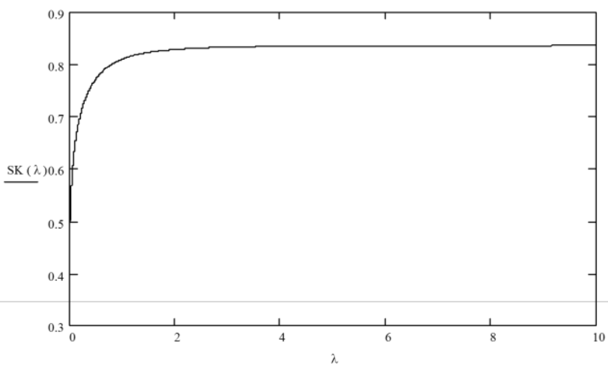
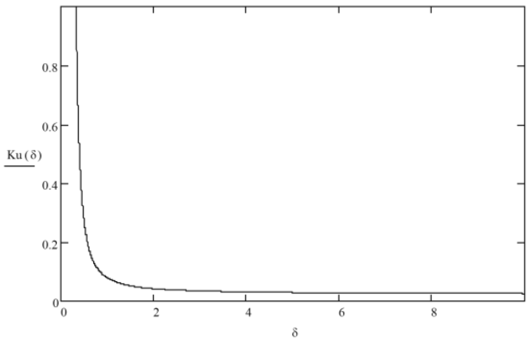


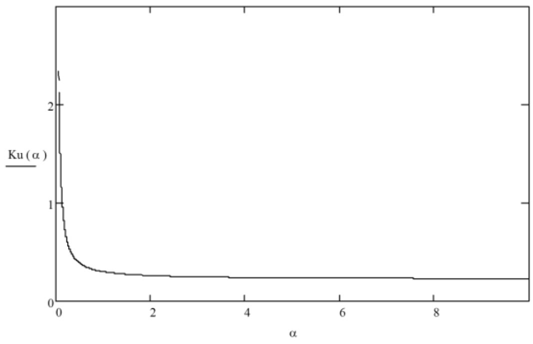
Figure 6:

(a) Plots of the TGP skewness for some parameter values of shape parameters
(a) Plots of the TGP skewness for some parameter values of transmuted parameters.
(b) Plots of the TGP skewness for some parameter values of shape parameters.

(a)



(b)



(c)

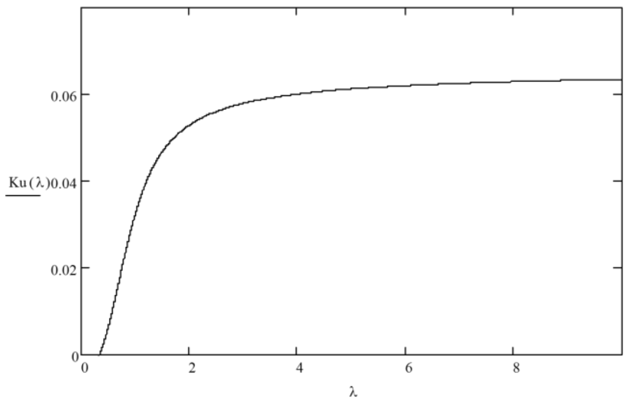


Figure 6:

(a) Plots of the TGP kurtosis for some parameter values of shape parameters.
(b) Plots of the TGP kurtosis for some parameter values of shape parameters.
(c) Plots of the TGP kurtosis for some parameter values of transmuted parameters

4.3 Quantile function and simulation

We present a method for simulating from the "TGP" distribution (5). The quantile function corresponding to (6) is

$$Q(u) = F^{-1}(u) = \beta\left\{\left(1+\frac{(1+\lambda)+\sqrt{(1+\lambda)^2-4\lambda u}}{2\lambda}\right)^{-\frac{1}{\alpha}}-1\right\}^{\frac{1}{\delta}} \tag{14}$$

Simulating the TGP random variable is straightforward. Let U be a uniform variate on the unit interval$(0, 1)$. Thus, by means of the inverse transformation method, we consider the random variable X given by

$$x = \beta\left\{\left(1+\frac{(1+\lambda)+\sqrt{(1+\lambda)^2-4\lambda u}}{2\lambda}\right)^{-\frac{1}{\alpha}}-1\right\}^{\frac{1}{\delta}} \tag{15}$$

which follows (6), i.e. $X \sim TGP(\beta,\eta,\alpha,\delta,\lambda)$, the plots comparing the exact TGP density and histograms from two simulated data sets for some parameter values

5. Estimation and information matrix

In this section, we discuss maximum likelihood estimation and inference for the T-GP distribution. Let $X_1, X_2, \dots, X_n$ be a random sample from $X \sim TGP(\beta,\eta,\alpha,\delta,\lambda)$ where $(\beta,\eta,\alpha,\delta,\lambda)$be the vector of the model Parameters, hen the likelihood function can be written as

$$L(\alpha,\beta,c,\lambda;x) = \prod_{i=1}^{n} f(x_{\mathrm{i}};\alpha,\beta,c,\lambda)$$

$$L(\alpha,\beta,\delta,\lambda;x) = \prod_{i=1}^{n}\frac{\delta\alpha}{\beta}\left(\frac{x_{\mathrm{i}}}{\beta}\right)^{\delta-1}\left[1+\left(\frac{x_{\mathrm{i}}}{\beta}\right)^{\delta}\right]^{-(\alpha+1)}\left\{(1+\lambda)-2\lambda\left(1-\left[1+\left(\frac{x_{\mathrm{i}}}{\beta}\right)^{\delta}\right]^{-\alpha}\right)\right\} \tag{16}$$

By accumulation taking logarithm of equation (16), and the log-likelihood function can be written as

$$\ln L(\alpha,\beta,\delta,\lambda;x) = \left| \begin{array}{l} n\,ln\left[\frac{\delta\alpha}{\beta}\right]+\left[(\delta-1)\sum_{i=1}^{n} ln\left(\frac{x_{\mathrm{i}}}{\beta}\right)\right]-(\alpha+1)\left[\sum_{i=1}^{n} ln\left[1+\left(\frac{x_{\mathrm{i}}}{\beta}\right)^{\delta}\right]\right]+ \\ \sum_{i=1}^{n} ln\left[(1+\lambda)^n-2\lambda\left[1-\left[1+\left(\frac{x_{\mathrm{i}}}{\beta}\right)^{\delta}\right]^{-\alpha}\right]\right] \end{array}\right.$$

(17)

The score vector $U_{\underline{\tau}} = \left(\frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial \delta}, \frac{\partial L}{\partial \lambda}\right)^T$, where the components corresponding to the parameters in $\underline{\tau}$ are given by differentiating (17) by setting $\upsilon_i = \left[1 + \left(\frac{x_i}{\beta}\right)^{\delta}\right]$.

$$\frac{\partial L}{\partial \alpha} = \sum_{i=1}^{n} \frac{1}{\alpha} - \sum_{i=1}^{n} \ln(\upsilon_i) - \sum_{i=1}^{n} \frac{2\lambda \sum_{i=1}^{n} \ln(\upsilon_i)\,(\upsilon_i)^{-\alpha}}{(1+\lambda) - 2\lambda(1-(\upsilon_i)^{-\alpha})} = 0$$

$$\frac{\partial L}{\partial \beta} = -\sum_{i=1}^{n} \frac{1}{\beta} - \frac{\delta - 1}{\beta} + (\alpha+1) \sum_{i=1}^{n} \frac{\delta \beta^{-\delta-1} x_i{}^{\delta}}{\upsilon_i} + \sum_{i=1}^{n} \frac{2\delta\lambda\alpha(\upsilon_i)^{-\alpha-1}\left(\frac{x_i}{\beta}\right)^{\delta}}{\beta[(1+\lambda) - 2\lambda(1-(\upsilon_i)^{-\alpha})]} = 0$$

$$\frac{\partial L}{\partial \delta} = \sum_{i=1}^{n} \frac{1}{\delta} + \sum_{i=1}^{n} \ln\left(\frac{x_i}{\beta}\right) - (\alpha+1) \sum_{i=1}^{n} \frac{\left(\frac{x_i}{\beta}\right)^{\delta} \ln\left(\frac{x_i}{\beta}\right)}{\upsilon_i} + \sum_{i=1}^{n} \frac{2\lambda \sum_{i=1}^{n} \ln(\upsilon_i)\,(\upsilon_i)^{-\alpha}}{(1+\lambda) - 2\lambda(1-(\upsilon_i)^{-\alpha})} = 0$$

and (18)

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{n} \frac{2(\upsilon_i)^{-\alpha} - 1}{(1+\lambda) - 2\lambda(1-(\upsilon_i)^{-\alpha})} = 0$$

The maximum likelihood estimates (MLEs) of the parameters are the solutions of the nonlinear equations (14), $\nabla \ell = 0$, which are solved iteratively, these solutions will yield the ML estimators for the parameters.

6 Simulation study

We conducted Monte Carlo simulation studies to assess on the finite sample behavior of the maximum likelihood estimators of α , β, δ, and

λ All results were obtained from 1000 Monte Carlo replications and the simulations were carried out using the statistical software Mathcad , The true parameter values used in the data generating processes are $\alpha = 1.2,\ \beta = 0.2,\ \delta = 0.3$ and $\lambda = 0.4$, Table 1 presents the mean maximum likelihood estimates of the parameters that index the *TGP* distribution along with the respective root mean squared errors (RMSE) and bias for sample sizes $n = 30\ , 50, 80, 100$

Table 1: Mean estimates, root mean squared errors and bias of $\alpha\ , \beta, \delta,$ and λ the maximum likelihood Estimators of the TGP parameters

n	Parameter	estimates	RMSE	bias	μ	$\sigma_{(x)}$
30	α	1.319	0.336	0.319	0.464	1.335
	β	0.29	0.536	0.2		
	δ	0.469	2.84	0.469		
	λ	0.958	6.633	-1.358		
50	α	1.297	0.316	0.297	0.882	1.404
	β	0.26	0.408	0.06		
	δ	0.430	3.016	0.13		
	λ	1.336	6.51	1.336		
100	α	1.257	0.26	0.63	2.737	3.677
	β	0.251	0.46	-2.323		
	δ	0.22	3.878	-1.808		
	λ	0.857	6.15	0.706		

From the results in Table 1, we notice that the biases and root mean squared errors of the maximum likelihood estimators of $\alpha\ , \beta, \delta,\ \beta$ and λ decay toward zero as the sample size increases, as expected. We also note that there is small sample bias in the estimation of the parameters that index the *TGP* distribution. Future research should obtain bias corrections for these estimators.

Conclusion

In the present study, we have introduced a new generalization of generalized Pareto Distribution called the transmuted Generalized Pareto distribution. The subject distribution is generated by using the quadratic rank transmutation map and taking the Generalized Pareto distribution as the base distribution. Some mathematical properties along with estimation issues are addressed. The hazard rate function and reliability behavior of the transmuted Generalized Pareto distribution shows that the subject distribution can be used to model reliability data. We expect that this study will serve as a reference and help to advance future research in the subject area.

REFERENCES

1- Abdul Fattah, A. M., Elsherpiny, E. A. and Hussein, E. A. (2007) “A new generalized Pareto distribution" Interstat Journal, Dec07.

2- Aryall G. R (2013) Transmuted Log-Logistic Distribution. J. Stat. Appl. Pro. 2, No. 1, 11-20.

3- Aryall G. R and Tsokos C. P. (2009). On the transmuted extreme value distribution with applications. Nonlinear Analysis: Theory, Methods and applications, V (71), 1401-1407.

4- Aryall G. R and Tsokos C. P.(2011). Transmuted Weibull distribution: A Generalization of the Weibull Probability Distribution. European Journal of Pure and Applied Mathematics, 4(2), 89-102.

5- F. Merovci, (2014). Transmuted generalized Rayleigh distribution. Journal of Statistics Applications and Probability, 3(1), 9-20.

6- F. Merovci, Transmuted Lindley Distribution. Int. J. Open Problems Compt. Math., 6(2013), (2), 63-72. http://dx.doi.org/10.12816/0006170 http://dx.doi.org/10.12785/jsap/030102

7- Khan M. S and Robert, K (2013). Transmuted Modified Weibull Distribution: A Generalization of the Modified Weibull Probability Distribution European Journal of Pure and Applied Mathematics. 6 (1), 66- 88.

8- M. S. Khan, & R. King, Transmuted Modified Weibull Distribution: A Generalization of the Modified Weibull Probability Distribution. European Journal of Pure and Applied Mathematics, 6(1) (2013), 66-88.

9- M.M. Mansour, M. A. Enayat, S.H. Mohamed and M. M. Salah, A New Transmuted Additive Weibull Distribution Based on a New Method for

Adding a Parameter to a Family of Distributions. International Journal of Applied Mathematical Sciences, (2015).

10- *Tarek M. Shams*, (2013). The Kumaraswamy-Generalized Exponentiated Pareto Distribution, European Journal of Applied Sciences 5 (3): 92-99, 2013