ProbStat Forum, Volume 07, January 2014, Pages 1–11
ISSN 0974-3235

ProbStat Forum is an e-journal.
For details please visit www.probstat.org.in

Transmuted Pareto distribution

Faton Merovci[a], Llukan Puka[b]

[a]*Department of Mathematics, Faculty of Natural Science and Mathematics, University of Prishtina, Kosovo*
[b]*Department of Mathematics, Faculty of Natural Science, University of Tirana, Albania*

Abstract. In this article, we generalize the Pareto distribution using the quadratic rank transmutation map studied by Shaw et al. [Shaw, W. T., Buckley, I. R. (2009) The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv preprint arXiv:0901.0434] to develop a transmuted Pareto distribution. We provide a comprehensive description of the mathematical properties of the subject distribution along with its reliability behavior. The usefulness of the transmuted Pareto distribution for modeling data is illustrated using real data.

1. Introduction

The Pareto distribution is named after an Italian-born Swiss professor of economics, Vilfredo Pareto (1848-1923). Pareto's law (Pareto, 1897) dealt with the distribution of income over a population and can be stated as $N = Ax^{-a}$, where N is the number of persons having income greater than x, and A, and a are parameters (a is known both as Pareto's constant and as a shape parameter). It was felt by Pareto that this law was universal and inevitable-regardless of taxation and social and political conditions. "Refutations" of the law have been made by several well-known economists over the past 70 years [e.g., Pigou; Shirras; Hayakawa]. More recently attempts have been made to explain many empirical phenomena using the Pareto distribution or some closely related form.For more detail see Johnson et al. (1995).

In this article we use transmutation map approach suggested by Shaw and Buckley (2009) to define a new model which generalizes the Pareto model. We will call the generalized distribution as the transmuted Pareto distribution. According to the Quadratic Rank Transmutation Map (QRTM), approach the cumulative distribution function(cdf) satisfy the relationship

$$F(x) = (1 + \lambda)G(x) - \lambda G^2(x), \ |\lambda| \leq 1,$$

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 and Tsokos (2009) studied the 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 transmuted Pareto distribution and some of its properties.

Keywords. Pareto distribution, hazard rate function, reliability function, parameter estimation
Received: 10 April 2013; Revised: 25 November 2013; Accepted: 19 December 2013
Email addresses: `fmerovci@yahoo.com` (Faton Merovci), `lpuka2001@yahoo.co.uk` (Llukan Puka)

2. Transmuted Pareto distribution

Definition 2.1. *The pdf of a Pareto distribution is*

$$g(x, a, x_0) = \frac{a x_0^a}{x^{a+1}},$$

and the respective cdf is

$$G(x, a, x_0) = 1 - \left(\frac{x_0}{x}\right)^a,$$

where x_0 is the (necessarily positive) minimum possible value of X, and a is a positive parameter.

The transmuted cdf is

$$F(x, a, x_0, \lambda) = \left[1 - \left(\frac{x_0}{x}\right)^a\right]\left[1 + \lambda\left(\frac{x_0}{x}\right)^a\right], \tag{1}$$

and its pdf

$$f(x, a, x_0, \lambda) = \frac{a x_0^a}{x^{a+1}}\left[1 - \lambda + 2\lambda\left(\frac{x_0}{x}\right)^a\right]. \tag{2}$$

Note that the transmuted Pareto distribution is an extended model to analyze more complex data. The Pareto distribution is clearly a special case for $\lambda = 0$. Figure 1 illustrates some of the possible shapes of the pdf of a transmuted Pareto distribution for selected values of the parameters λ and a.

Lemma 2.2. *The limit of transmuted Pareto density as $x \to \infty$ is 0 and the limit as $x \to x_0$ is $\frac{a(1+\lambda)}{x_0}$.*

Proof. It is straightforward to show the above from the transmuted Pareto density in (2). $\square$

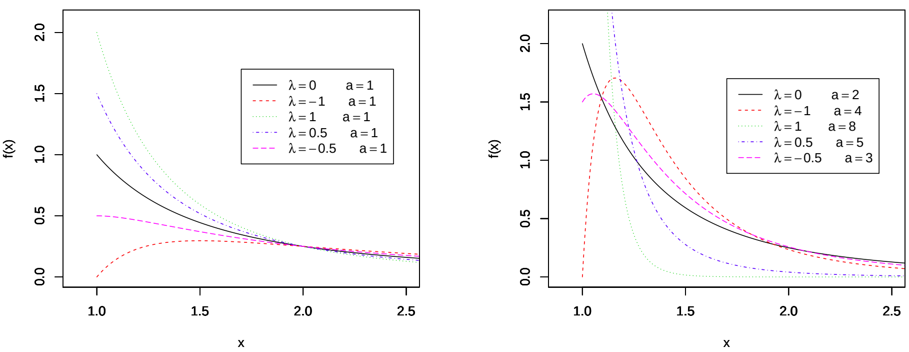

Figure 1: The pdf's of various transmuted Pareto distributions.

3. Moments and associated measures

Theorem 3.1. *Let X have a transmuted Pareto distribution. Then the r^{th} moment of X, say $E[X^r]$, is*

$$E[X^r] = \frac{a x_0^r \Big(2a - r(1+\lambda)\Big)}{(a-r)(2a-r)}, \quad a > r.$$

Especially we have

$$\mu = E(X) = \frac{ax_0(2a - 1\lambda)}{(a-1)(2a-1)},$$

$$\sigma^2 = var(X) = \frac{ax_0^2}{a-1}\Big[\frac{2a-1-\lambda}{2a-1} - \frac{ax_0^2(a-1-\lambda)^2}{(a-2)^2}\Big], a > 2.$$

Proof. The r^{th} order moment is given by

$$\begin{aligned}
E[X^r] &= \int_{x_0}^{\infty} x^r f(x)dx = \int_{x_0}^{\infty} \frac{ax_0^a}{x^{a-r+1}}\Big[1 - \lambda + 2\lambda\Big(\frac{x_0}{x}\Big)^a\Big]dx \\
&= ax_0^a(1-\lambda)\int_{x_0}^{\infty}\frac{dx}{x^{a-r+1}} + 2\lambda a x_0^{2a}\int_{x_0}^{\infty}\frac{dx}{x^{2a-r+1}} \\
&= ax_0^a(1-\lambda)\Bigg[-\frac{1}{(a-r)x^{a-r}}\Bigg|_{x_0}^{\infty}\Bigg] + 2\lambda a x_0^{2a}\Bigg[-\frac{1}{(2a-r)x^{2a-r}}\Bigg|_{x_0}^{\infty}\Bigg] \\
&= ax_0^a(1-\lambda)\Big[\frac{1}{(a-r)x_0^{a-r}}\Big] + 2\lambda a x_0^{2a}\Big[\frac{1}{(2a-r)x_0^{2a-r}}\Big] \\
&= \frac{ax_0^r(1-\lambda)}{a-r} + \frac{2\lambda a x_0^r}{2a-r} \\
&= \frac{ax_0^r\Big(2a - r(1+\lambda)\Big)}{(a-r)(2a-r)}.
\end{aligned}$$

If $0 < a \leq r$, then $a - r < 0$ and

$$-\frac{1}{(a-r)x^{a-r}} = \frac{1}{r-a}x^{r-a}, \quad -\frac{1}{(2a-r)x^{2a-r}} = \frac{1}{r-2a}x^{r-2a}$$

goes to infinity as $x \to \infty$. Moreover, as $a \to r$ these expressions approaches infinity for all positive x. Therefore, $E[X^r]$ does not exist for $0 < a \leq r$.

The second moment is

$$E(X^2) = \frac{ax_0^2(a-1-\lambda)}{(a-1)(a-2)},$$

and the variance is

$$\sigma^2 = \frac{ax_0^2}{a-1}\Big[\frac{2a-1-\lambda}{2a-1} - \frac{ax_0^2(a-1-\lambda)^2}{(a-2)^2}\Big],\ a > 2. \ \square$$

The skewness and kurtosis measures can be obtained from the expressions

$$Skewness(X) = \frac{E(X^3) - 3E(X^2)\mu + 2\mu^3}{\sigma^3},$$

$$Kurtosis(X) = \frac{E(X^4) - 4E(X^3)\mu + 6E(X^2)\mu^2 - 3\mu^4}{\sigma^4},$$

upon substituting for the raw moments.

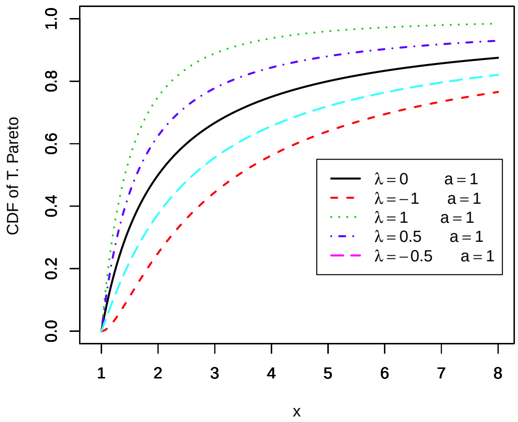

Figure 2: The cdf's of various transmuted Pareto distributions.

Theorem 3.2. *Let X have a transmuted Pareto distribution. Then the moment generating function of X, say $M_X(t)$, is*

$$M_X(t) = \sum_{i=0}^{\infty} \frac{t^i}{i!} \frac{a x_0^i \Big(2a - i(1+\lambda)\Big)}{(a-i)(2a-i)}.$$

Proof. The moment generating function of the random variable X is given by

$$\begin{aligned} M_X(t) = E[e^{tx}] &= \int_0^{\infty} e^{tx} f(x)\, dx \\ &= \int_0^{\infty} \left(1 + tx + \frac{t^2x^2}{2!} + \cdots + \frac{t^n x^n}{n!} + \cdots\right) f(x)\, dx \\ &= \sum_{i=0}^{\infty} \frac{t^i E(X^i)}{i!} \\ &= \sum_{i=0}^{\infty} \frac{t^i}{i!} \frac{a x_0^i \Big(2a - i(1+\lambda)\Big)}{(a-i)(2a-i)}. \quad \square \end{aligned}$$

The p^{th} quantile x_p of the transmuted Pareto distribution can be obtained from (1) as

$$x_p = x_0 \sqrt[a]{\frac{2\lambda}{\lambda - 1 + \sqrt{(1-\lambda)^2 - 4\lambda(p-1)}}}.$$

In particular, the distribution median is

$$x_{0.5} = x_0 \sqrt[a]{\frac{2\lambda}{\lambda - 1 + \sqrt{1+\lambda^2}}}.$$

4. Random number generation and parameter estimation

Using the method of inversion we can generate random numbers from the transmuted Pareto distribution as

$$1-\lambda\Big(\frac{x_0}{x}\Big)^{2a}+(\lambda-1)\Big(\frac{x_0}{x}\Big)^{a}=u,$$

where $u \sim U(0,1)$. After simple calculation this yields

$$x=x_0\left[\frac{\lambda-1+\sqrt{(1-\lambda)^2-4\lambda(u-1)}}{2\lambda}\right]^{-\frac{1}{a}}. \tag{3}$$

One can use equation (3) to generate random numbers when the parameters a and λ are known. The maximum likelihood estimates, MLEs, of the parameters that are inherent within the transmuted Pareto probability distribution function is given by the following: Let $X_1, X_2, \cdots, X_n$ be a sample of size n from a transmuted Pareto distribution. Then the likelihood function is given by

$$L=\prod_{i=1}^{n} f(x_i,a,x_0,\lambda)=\frac{a^n x_0^{an}}{\prod\limits_{i=1}^{n} x_i^{a+1}}\prod_{i=1}^{n}\Big[1-\lambda+2\lambda\Big(\frac{x_0}{x}\Big)^{a}\Big], \tag{4}$$

so, the log-likelihood function is:

$$LL=\ln L=n\ln(a)+na\ln(x_0)-(a+1)\sum_{i=1}^{n}\ln(x_i)+\sum_{i=1}^{n}\ln\Big[1-\lambda+2\lambda\Big(\frac{x_0}{x}\Big)^{a}\Big].$$

Now setting $LL_a=0$ and $LL_\lambda=0$, we have

$$0=\frac{n}{a}+n\ln(x_0)-\sum_{i=1}^{n}\ln(x_i)+\sum_{i=1}^{n}\frac{2\lambda\Big(\frac{x_0}{x}\Big)^{a}\ln\Big(\frac{x_0}{x}\Big)}{1-\lambda+2\lambda\Big(\frac{x_0}{x}\Big)^{a}},$$

$$0=\sum_{i=1}^{n}\frac{2\Big(\frac{x_0}{x}\Big)^{a}-1}{1-\lambda+2\lambda\Big(\frac{x_0}{x}\Big)^{a}}.$$

Since $x \geq x_0$, the maximum likelihood estimate of x_0 is the first-order statistic $x_{(1)}$. The maximum likelihood estimator $\hat{\theta}=(\hat{a},\hat{\lambda})'$ of $\theta=(a,\lambda)'$ is obtained by solving this nonlinear system of equations. It is usually more convenient to use nonlinear optimization algorithms such as the quasi-Newton algorithm to numerically maximize the log-likelihood function given in (4). Applying the usual large sample approximation (assuming x_0 to be known), the maximum likelihood estimators of θ can be treated as being approximately bivariate normal with mean θ and variance-covariance matrix equal to the inverse of the expected information matrix. That is,

$$\sqrt{n}(\hat{\theta}-\theta)\rightarrow N(0,I^{-1}(\hat{\theta})),$$

where $I^{-1}(\hat{\theta})$ is the variance-covariance matrix of the unknown parameters $\theta = (\sigma, \lambda)$. The elements of the 2×2 matrix I^{-1}, $\quad I_{ij}(\hat{\theta}), i, j = 1, 2$, can be approximated by $I_{ij}(\hat{\theta})$, where $I_{ij}(\hat{\theta}) = -LL_{\theta_i\theta_j}\big|_{\theta=\hat{\theta}}$. Also

$$\begin{aligned}
I_{11} &= -\frac{\partial^2 \ln L}{\partial a^2} = \frac{n}{a^2} - 2\lambda(1-\lambda)\sum_{i=1}^{n} \frac{\left(\frac{x_0}{x}\right)^a \ln^2\left(\frac{x_0}{x}\right)}{\left[1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^a\right]^2}, \\
I_{12} &= -\frac{\partial^2 \ln L}{\partial a \partial \lambda} = 2\lambda(\lambda-1)\sum_{i=1}^{n} \frac{\left(\frac{x_0}{x}\right)^a \ln\left(\frac{x_0}{x}\right)\left(1-2\left(\frac{x_0}{x}\right)^a\right)}{\left[1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^a\right]^2}, \\
I_{22} &= -\frac{\partial^2 \ln L}{\partial \lambda^2} == \sum_{i=1}^{n}\left[\frac{2\left(\frac{x_0}{x}\right)^a - 1}{1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^a}\right]^2.
\end{aligned}$$

Approximate $100(1-\alpha)\%$ two sided confidence intervals for a and λ are, respectively, given by

$$\hat{a} \pm z_{\alpha/2}\sqrt{I_{11}^{-1}(\hat{\theta})}, \text{ and } \hat{\lambda} \pm z_{\alpha/2}\sqrt{I_{22}^{-1}(\hat{\theta})},$$

where z_α is the upper $\alpha-$th percentiles of the standard normal distribution. Using R we can easily compute the Hessian matrix and its inverse and hence the values of the standard error and asymptotic confidence intervals.

We can compute the maximized unrestricted and restricted log-likelihoods to construct the likelihood ratio (LR) statistics for testing some transmuted Pareto sub-models. For example, we can use LR statistics to check whether the fitted transmuted Pareto distribution for a given data set is statistically "superior" to the fitted Pareto distribution. In any case, hypothesis tests of the type $H_0 : \Theta = \Theta_0$ versus $H_0 : \Theta \neq \Theta_0$ can be performed using LR statistics.In this case, the LR statistic for testing H_0 versus H_1 is $\omega = 2(L(\hat{\Theta}) - L(\hat{\Theta}_0))$, where $\hat{\Theta}$ and $\hat{\Theta}_0$ are the MLEs under H_1 and H_0. The statistic ω is asymptotically(as $n \to \infty$) distributed as χ^2_k, where k is the dimension of the subset Ω of interest. The LR test rejects H_0 if $\omega > \xi_\gamma$, where ξ_γ denotes the upper $100\gamma\%$ point of the χ^2_k distribution.

5. Reliability analysis

The reliability function $R(t)$, which is the probability of an item not failing prior to some time t, is defined by $R(t) = 1 - F(t)$. The reliability function of a transmuted Pareto distribution is given by

$$R(t) = \left(\frac{x_0}{t}\right)^a \left(\lambda\left(\frac{x_0}{t}\right)^a - \lambda + 1\right).$$

The other characteristic of interest of a random variable is the hazard rate function defined by

$$h(t) = \frac{f(t)}{1 - F(t)},$$

which is an important quantity characterizing life phenomenon. It can be loosely interpreted as the conditional probability of failure, given it has survived to the time t. The hazard rate function for a transmuted Pareto random variable is given by

$$h(t) = \frac{a\left(1-\lambda+2\lambda\left(\frac{x_0}{t}\right)^a\right)}{t\left(1-\lambda+\lambda\left(\frac{x_0}{t}\right)^a\right)}.$$

Lemma 5.1. *The limit of transmuted Pareto hazard function as $t \to \infty$ is 0 and the limit as $t \to x_0$ is* $\frac{a(1+\lambda)}{x_0}$.

Proof. It is straightforward to show the results of Lemma 2.2 by taking the limit of transmuted Pareto hazard function. □

Figure 3 illustrates the reliability function of a transmuted Pareto distribution for different combinations of parameters a and λ.

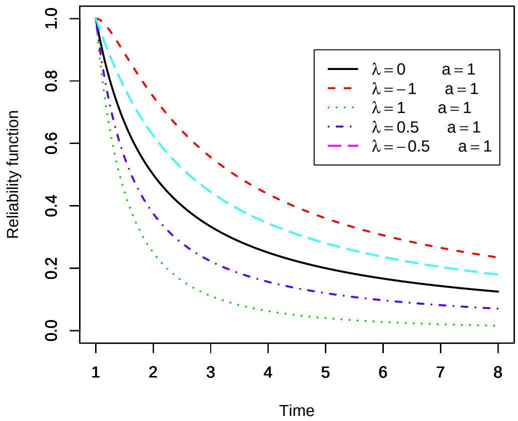

Figure 3: The reliability function of a transmuted Pareto distribution

6. Order statistics

In statistics, the k^{th} order statistic of a statistical sample is equal to its k^{th} smallest value. Together with rank statistics, order statistics are among the most fundamental tools in non-parametric statistics and inference. For a sample of size n, the n^{th} order statistic (or largest order statistic) is the maximum, that is $X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$. The sample range is the difference between the maximum and minimum. It is clearly a function of the order statistics $\text{Range}\{X_1, X_2, \ldots, X_n\} = X_{(n)} - X_{(1)}$. We know that if $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ denotes the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$ then the pdf of $X_{(j)}$ is given by

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{j-1}[1 - F_X(x)]^{n-j}$$

for $j = 1, 2, \ldots, n$. The pdf of the j^{th} order statistic for transmuted Pareto distributions is given by

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} \frac{a x_0^a}{x^{a+1}} \left[1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^a\right] \left[\left[1-\left(\frac{x_0}{x}\right)^a\right]\left[1+\lambda\left(\frac{x_0}{x}\right)^a\right]\right]^{j-1} \left[\left(\frac{x_0}{x}\right)^a\left(\lambda\left(\frac{x_0}{x}\right)^a-\lambda+1\right)\right]^{n-j}.$$

Therefore, the pdf of the largest order statistic $X_{(n)}$ is given by

$$f_{X_{(n)}}(x) = \frac{n a x_0^a}{x^{a+1}} \left[1 - \lambda + 2\lambda\left(\frac{x_0}{x}\right)^a\right] \times \left[\left[1 - \left(\frac{x_0}{x}\right)^a\right]\left[1 + \lambda\left(\frac{x_0}{x}\right)^a\right]\right]^{n-1},$$

and the pdf of the smallest order statistic $X_{(1)}$ is given by

$$f_{X_{(1)}}(x) = \frac{nax_0^a}{x^{a+1}}\left[1-\lambda+2\lambda\Big(\frac{x_0}{x}\Big)^a\right]\left[\Big(\frac{x_0}{x}\Big)^a\left(\lambda\Big(\frac{x_0}{x}\Big)^a-\lambda+1\right)\right]^{n-1}.$$

7. Application of transmuted Pareto distribution

In this section, we use a real data set to show that the transmuted Pareto distribution can be a better model than one based on the Pareto distribution.

Data Set 1. The data are the exceedances of flood peaks (in m^3/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. This data were analyzed by Choulakian and Stephens (2011) and are given in Table 1.

Table 1. Exceedances of Wheaton River flood data.

1.7	2.2	14.4	1.1	0.4	20.6	5.3	0.7
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
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	1.9	2.8

Model	Parameter Estimates	Standard Error	-LL
Transmuted	$\hat{a} = 0.349$	0.031	286.201
Pareto	$\hat{\lambda} = -0.952$	0.047	
	$x_0 = \min(x) = 0.1$		
Pareto	$\hat{a}$=0.2438634	0.028	303.064
	$x_0 = \min(x) = 0.1$		
Generalized Pareto	$\hat{a}$=12.192	2.294	252.128
	$\hat{k} = -0.932 \cdot 10^{-3}$	0.146	
Exponentiated Weibull	$\hat{\alpha} = 0.0502$	0.021	251.025
	$\hat{\lambda} = 1.386$	0.614	
	$\hat{\theta} = 0.518$	0.324	

Table 2. Estimated parameters of the Pareto and transmuted Pareto distribution for exceedances of Wheaton River flood data by assuming $x_0 = \min(x)$.

The hessian matrix of transmuted Pareto($\hat{a} = 0.349, \hat{\lambda} = -0.952$) by assuming $x_0 = \min(x) = 0.1$ is computed as

$$\begin{pmatrix} 1047.543 & 71.888 \\ 71.888 & 448.564 \end{pmatrix}$$

and the variance covariance matrix

$$I(\hat{\theta})^{-1} = \begin{pmatrix} 0.965 \times 10^{-3} & -0.546 \times 10^{-3} \\ -0.546 \times 10^{-3} & 0.225 \times 10^{-2} \end{pmatrix}.$$

Thus, the variances of the MLE of a and λ become $Var(\hat{a}) = 0.965 \times 10^{-3}$ and $Var(\hat{\lambda}) = 0.225 \times 10^{-2}$. Therefore, the 95% C.I of a and λ, respectively, are $[0.289, 0.410]$ and $[-0.859, -1]$.

In order to compare the distributions, we consider some other criterion like K-S (Kolmogorow-Smirnov), $-2\log(L)$, AIC (Akaike Information Criterion), AICC (Akaike Information Criterion Corrected) and BIC

(Bayesian information criterion) for the real data set. The best distribution corresponds to lower K-S, $-2\log(L)$, AIC, AICC and BIC values:

$$KS = \max_{1\le i\le n}\left(F(X_i) - \frac{i-1}{n}, \frac{i}{n} - F(X_i)\right),$$

$$AIC = 2k - 2\log(L), \quad AICC = AIC + \frac{2k(k+1)}{n-k-1}, \quad \text{and} \quad BIC = k\log(n) - 2\log L,$$

where k is the number of parameters in the statistical model, n the sample size and L is the maximized value of the likelihood function for the estimated model. Also, here for calculating the values of K-S we use the sample estimates of λ and a. Table 1 shows parameter MLEs to each one of the two fitted distributions, table 2 shows the values of K-S, $-2\log(L)$, AIC, AICC and BIC values. The values in Table 2 indicate that the transmuted Pareto distribution leads to a better fit than the Pareto distribution.

Model	K-S	-2LL	AIC	AICC	BIC
Pareto.	0.456	606.128	610.128	610.302	610.405
T. Pareto	0.389	572.401	578.402	578.755	580.955

Table 3.Criteria for comparison.

The LR statistics to test the hypotheses $H_0 : \lambda = 0$ versus $H_1 : \lambda \neq 0 : \omega = 33.7265 > 3.841 = \chi^2_1(\alpha = 0.05)$, so we reject the null hypothesis.

Data Set 2. The second flood data is for the Floyd River located in James, Iowa, USA. The Floyd River flood rates for the years 1935–1973 are provided in Table 4. For more details and the source of the data, see Mudholkar and Hutson (1996).

Table 4. Annual flood discharge rates of the Floyd River.

1935 - 1944	1460	4050	3570	2060	1300	1390	1720	6280	1360	7440
1945 - 1954	5320	1400	3240	2710	4520	4840	8320	13900	71500	6250
1955 - 1964	2260	318	1330	970	1920	15100	2870	20600	3810	726
1965 - 1973	7500	7170	2000	829	17300	4740	13400	2940	5660	

Model	Parameter Estimates	Standard Error	-LL
Pareto	$\hat{a}$=0.412	0.066	392.810
Transmuted	$\hat{a} = 0.585$	0.072	385.349
Pareto	$\hat{\lambda} = -0.910$	0.089	
Generelazied Pareto	$\hat{a}$=4582.526	1061.214	379.543
	$\hat{k}$=-0.301	0.174	
Exponentiated Weibull	$\hat{\alpha} = 0.092$	0.214	376.362
	$\hat{\lambda} = 0.254$	0.060	
	$\hat{\theta} = 52.127$	57.525	

Table 5. Estimated parameters of the Pareto and transmuted Pareto distribution for the annual flood discharge rates of the Floyd River by assuming $x_0 = \min(x)$.

The hessian matrix of transmuted Pareto($\hat{a} = 0.585, \hat{\lambda} = -0.910$) by assuming $x_0 = \min(x) = 318$ is computed as

$$\begin{pmatrix} 197.257 & 23.940 \\ 23.940 & 128.290 \end{pmatrix}$$

and the variance covariance matrix is

$$I(\hat{\theta})^{-1} = \begin{pmatrix} 0.518\times 10^{-2} & -0.967\times 10^{-3} \\ -0.967\times 10^{-3} & 0.797\times 10^{-2} \end{pmatrix}.$$

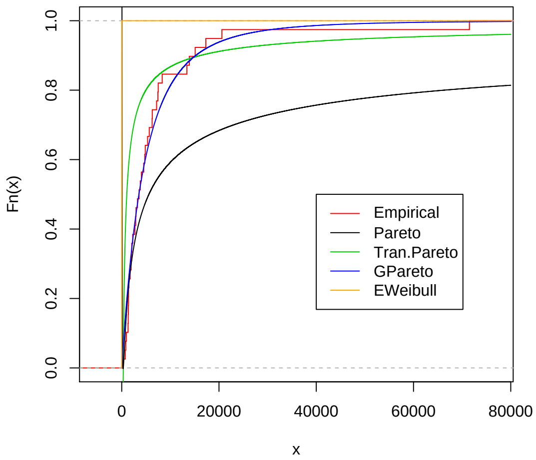

Figure 4: Empirical, fitted Pareto and transmuted Pareto cdf of exceedances of annual flood discharge rates of the Floyd River

Thus, the variances of the MLE of a and λ become $Var(\hat{\sigma}) = 0.518 \times 10^{-2}$ and $Var(\hat{\lambda}) = 0.797 \times 10^{-2}$. Therefore, the 95% C.I of a and λ, respectively, are $[0.444, 0.727]$ and $[-0.735, -1]$.

Model	K-S	-2LL	AIC	AICC	BIC
Pareto.	0.459	785.619	789.619	789.722	789.283
T. Pareto	0.287	770.698	776.698	777.014	778.025

Table 6. Criteria for comparison.

The LR statistics to test the hypotheses $H_0 : \lambda = 0$ versus $H_1 : \lambda \neq 0 : \omega = 14.921 > 3.841 = \chi_1^2(\alpha = 0.05)$, so we reject the null hypothesis.

Table 5 shows parameter MLEs to each one of the two fitted distributions, Table 6 shows the values of K-S, $-2\log(L)$, AIC, AICC and BIC values for data set 2. The values in Table 6 indicate that the transmuted Pareto distribution leads to a better fit than the Pareto distribution.

8. Conclusion

In this article, we propose a new model: the so-called the transmuted Pareto distribution which extends the Pareto distribution in the analysis of data with real support. An obvious reason for generalizing a standard distribution is because the generalized form is that it provides greater flexibility in modeling real data. We derive expansions for the expectation, variance, moments and the moment generating function. The estimation of parameters is approached by the method of maximum likelihood, also the information matrix is derived. We consider the likelihood ratio statistic to compare the model with its baseline model. An application of the transmuted Pareto distribution to real data show that the new distribution can be used quite effectively to provide better fits than the Pareto distribution.

References

Aryal, G. R., Tsokos, C. P. (2009) On the transmuted extreme value distribution with application, Nonlinear Analysis: Theory, Methods & Applications 71(12), e1401–e1407.

Choulakian, V., Stephens, M. A. (2001). Goodness-of-fit tests for the generalized Pareto distribution, Technometrics 43(4), 478–484.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1995) Continuous univariate distributions, Vol. 2, Wiley, New York.

Mudholkar, G. S., Hutson, A. D. (1996) The exponentiated Weibull family: Some properties and a flood data application, Communications in Statistics - Theory and Methods 25(12), 3059–3083.

Shaw, W. T., Buckley, I. R. (2009) The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map, arXiv preprint arXiv:0901.0434.