Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/060201

The Beta Transmuted Pareto Distribution: Theory and Applications

Sher B. Chhetri[1], *Hongwei Long*[1] *and Gokarna Aryal*[2,*]

[1] Department of Mathematical Sciences, Florida Atlantic University, Boca Raton, FL, USA
[2] Department of Mathematics, Statistics and Computer Science, Purdue University Northwest, USA

Received: 12 Jan. 2017, Revised: 6 May 2017, Accepted: 12 May 2017
Published online: 1 Jul. 2017

Abstract: In this paper a composite generalizer of the Pareto distribution is proposed and studied. The genesis of the beta distribution and transmuted map is used to develop the so-called beta transmuted Pareto (BTP) distribution. Several mathematical properties including moments, mean deviation, probability weighted moments, residual life, distribution of order statistics and the reliability analysis are discussed. The method of maximum likelihood is proposed to estimate the parameters of the distribution. We illustrate the usefulness of the proposed distribution by presenting its application to model real-life data sets.

Keywords: Pareto distribution, beta Pareto distribution, transmuted distribution, parameter estimation

1 Introduction

The Pareto distribution is named after economist Vilfredo Pareto who revealed it while modeling income data. Because of its heavy tail properties it is widely used in modeling data from reliability, finance and actuarial sciences, economics, among others. Burroughs and Tebbens [4] discussed applications of the Pareto distribution in modeling earthquakes, forest fire areas and oil and gas field sizes. Newman [17] also provided many other quantities measured in the physical and biological systems where the Pareto distribution has applications. The probability density function (pdf) and cumulative distribution function (cdf) of a Pareto distribution is given by

$$g_1(x;\alpha,x_0) = \frac{\alpha x_0^\alpha}{x^{\alpha+1}}, \tag{1}$$

and

$$G_1(x;\alpha,x_0) = 1 - \left(\frac{x_0}{x}\right)^\alpha, \tag{2}$$

respectively, where $x \in (x_0,\infty)$, $\alpha > 0$ is the shape parameter and $x_0 > 0$ is the scale parameter.

A hierarchy of the Pareto distributions has been established starting from the classical Pareto (I) to Pareto (IV) distributions with subsequent additional parameters related to location, shape and inequality. To add more flexibility to the Pareto distribution many of its generalizations have appeared in the literature in last few years. To name a few, Alzaatreh, Famoye and Lee [2] proposed the Weibull-Pareto distribution, Bourguignon et al. [5] introduced the Kumaraswamy-Pareto distribution, Alzaatreh, Famoye and Lee [3] introduced the Gamma-Pareto distribution, Akinsete, Famoye and Lee [1] introduced the beta-Pareto distribution, Zea et al. [20] studied the beta exponentiated Pareto distribution, Mahmoudi [13] proposed the beta generalized Pareto distribution, Elbatal [8] studied the Kumarswamy exponentiated Pareto distribution, Tahir et al. [19] studied a new Weibull-Pareto distribution, Merovci and Puka [14] introduced transmuted Pareto distribution. In this article we propose a new generalizer which is obtained by the composition of the genesis of beta distribution and transmutation map. We will execute this generalizer to the Pareto distribution to develop the so-called beta transmuted Pareto distribution. This will be the beta generalizer of the

* Corresponding author e-mail: aryalg@pnw.edu

transmuted Pareto (TP) distribution studied by Merovci and Puka [14]. A random variable X is said to have a transmuted Pareto probability distribution with parameter $x_0 > 0$, $\alpha > 0$ and $|\lambda| \leq 1$, if its pdf is given by

$$g(x;\alpha,x_0,\lambda) = \frac{\alpha x_0^{\alpha}}{x^{\alpha+1}}\left[1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right], x > x_0. \tag{3}$$

The corresponding cdf of the transmuted Pareto distribution is given by

$$G(x;\alpha,x_0,\lambda) = \left[1-\left(\frac{x_0}{x}\right)^{\alpha}\right]\left[1+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right], x > x_0. \tag{4}$$

Eugene, Lee and Famoye [9] used the beta distribution as a generator to develop the so-called a family of beta-generated (BG) distributions based on the following formulation.
Let $G(x)$ be the cumulative distribution function (cdf) of a random variable X. Then the cdf of the beta-G random variable is given by

$$F(x) = I_{G(x)}(a,b) = \frac{1}{B(a,b)}\int_0^{G(x)} w^{a-1}(1-w)^{b-1}dw, \tag{5}$$

where $a > 0$ and $b > 0$ are shape parameters. Note that $I_y(a,b) = \frac{B_y(a,b)}{B(a,b)}$ is the incomplete beta function ratio, $B_y(a,b) = \int_0^y w^{a-1}(1-w)^{b-1}dw$ is the incomplete beta function, $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is the beta function and $\Gamma(.)$ is the gamma function.
The probability density function (pdf) of the Beta-G distribution has the form

$$f(x) = \frac{1}{B(a,b)}[G(x)]^{a-1}\left[1-G(x)\right]^{b-1}g(x). \tag{6}$$

This class of generalized distribution has received considerable attention over the last years and several classical distributions have been generalized using this formulation. We generalize the transmuted Pareto distribution (3) using this formulation in order to construct the beta transmuted Pareto (BTP) distribution. We provide a comprehensive description of mathematical properties of BTP distribution and its application to analyze real data sets. The rest of the paper is unfolded as follows. In Section 2 we define the BTP distribution and discuss some of its sub-models. In Section 3 we present the mixture representation of the BTP distribution. Section 4 discusses some structural and mathematical properties of the BTP distribution such as the quantile, moments, mean deviation, residual life, distribution of the order statistics etc. Parameter estimation procedures using method of maximum likelihood estimates are presented in Section 5. In Section 6 we study the elements of reliability analysis. Application to model real world data is discussed in Section 7. Section 8 provides some concluding remarks.

2 The beta transmuted Pareto distribution

In this section we provide the formulation of the beta transmuted Pareto distribution. By inserting (4) into (5) the cumulative distribution function of the beta-transmuted Pareto distribution with five parameters is given by

$$F(x) = \frac{1}{B(a,b)}\int_0^{\left[1-\left(\frac{x_0}{x}\right)^{\alpha}\right]\left[1+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]} w^{a-1}(1-w)^{b-1}dw, \tag{7}$$

where $x > x_0 > 0$, $|\lambda| \leq 1$, $\alpha > 0$, $a > 0$, $b > 0$.
Using the incomplete beta function and the hyper-geometric confluent function, one can express the cdf in a closed form as below (see Cordeiro and Nadarajah [7]),

$$F(x) = \frac{(G(x;\alpha,x_0,\lambda))^a}{B(a,b)}.\sum_{k=0}^{\infty}\frac{(1-b)_k(G(x;\alpha,x_0,\lambda))^k)}{(a+k)k!}.$$

Hence,

$$F(x) = \frac{(G(x;\alpha,x_0,\lambda))^a\,._2F_1\left(a,1-b,a+1;G(x;\alpha,x_0,\lambda)\right)}{aB(a,b)},$$

where ${}_2F_1(c,d;e;z) = \sum_{k=0}^{\infty} \frac{(c)_k (d)_k}{(e)_k} \cdot \frac{z^k}{k!}$ is the Gaussian hyper-geometric function where $(c)_k$ is the ascending factorial defined by (assuming that $(c)_0 = 1$)

$$(c)_k = \begin{cases} c(c+1)(c+2)\cdots(c+k-1) & k = 1,2,3,\cdots \\ 1 & k = 0 \end{cases}$$

Differentiating (7) with respect to x, we get the probability density function of the BTP distribution given by

$$f(x) = \frac{1}{B(a,b)} \frac{\alpha x_0^{\alpha}}{x^{\alpha+1}} \left[1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right] \left[1-\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a-1} \left[1+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a-1} \left\{\left(\frac{x_0}{x}\right)^{\alpha} \left[1-\lambda+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]\right\}^{b-1}, \quad (8)$$

where $x > x_0 > 0$, $|\lambda| \leq 1$, $\alpha > 0$, $a > 0$ and $b > 0$.

The BTP distribution includes the following distributions as special case:

–for $\lambda = 0$, BTP reduces to BP distribution by Akinsete et al. [1]
–for $a = b = 1$, BTP reduces to TP distribution by Merovci and Puka [14]
–for $\lambda = 0$ and $b = 1$, BTP reduces to exponentiated Pareto by Nadarajah [16]
–for $a = b = 1$ and $\lambda = 0$, BTP reduces to Pareto distribution.

Figure 1 illustrates the graphical behavior of the pdf and the cdf of BTP distribution for selected values of the parameters a, b, λ and α with $x_0 = 0.1$. As we shall see in the sequel, this is a rather flexible distribution and could be useful to model different phenomena exhibited by real world data.

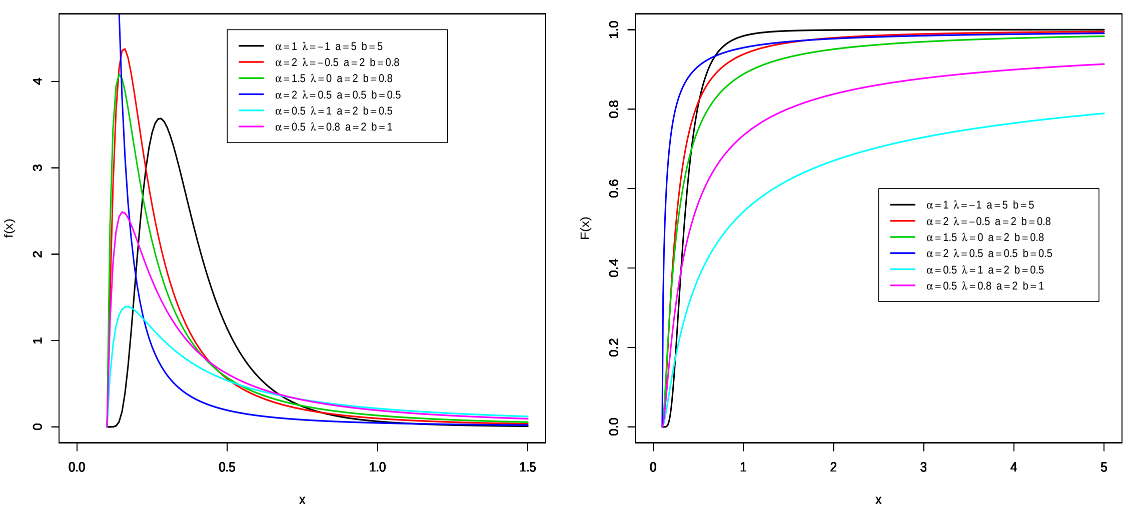

Fig. 1: pdf (left) and cdf (right) of BTP distribution for selected values of the parameters.

3 Mixture representation

In this section we find the series representations of the cdf and the pdf of the BTP distribution which will be useful to study its mathematical characteristics. As we shall see both pdf and cdf of BTP distribution can be expressed in terms of the Pareto distribution. By using (5) and the power series expansion of $(1-w)^{b-1}$, we get

$$\frac{1}{B(a,b)} \int_0^{G(x)} w^{a-1}(1-w)^{b-1} dt = \frac{1}{B(a,b)} \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{[G(x)]^{a+i}}{a+i}$$

with the binomial term $\binom{b-1}{i} = \frac{\Gamma(b)}{\Gamma(b-i)i!}$ defined for any real b. Hence, (7) reduces to

$$F(x) = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} \frac{\left[1-\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a+i}\left[1+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a+i}}{B(a,b)(a+i)}, x > x_0. \tag{9}$$

Again, using the binomial expansion of $\left[1-\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a+i}$ and $\left[1+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a+i}$, we have

$$\begin{aligned} F(x) &= \sum_{i,k,l=0}^{\infty} (-1)^{i+k} \binom{b-1}{i}\binom{a+i}{k}\binom{a+i}{l} \frac{\lambda^l \left(\frac{x_0}{x}\right)^{\alpha(k+l)}}{B(a,b)(a+i)} \\ &= \sum_{i,k,l=0}^{\infty} (-1)^{i+k} \binom{b-1}{i}\binom{a+i}{k}\binom{a+i}{l} \frac{\lambda^l (1-G_1(x;\alpha(k+l),x_0))}{B(a,b)(a+i)} \\ &= \sum_{k,l=0}^{\infty} w_{kl} \left(\frac{x_0}{x}\right)^{\alpha(k+l)}, \end{aligned} \tag{10}$$

where $G_1(x;\alpha(k+l),x_0)$ is the Pareto cdf with shape parameter $\alpha(k+l)$ and the scale parameter x_0 and

$$w_{kl} = \sum_{i=0}^{\infty} (-1)^{i+k} \binom{b-1}{i}\binom{a+i}{k}\binom{a+i}{l} \frac{\lambda^l}{B(a,b)(a+i)}. \tag{11}$$

Using the power series expansion we may express the pdf (8) as below:

$$\begin{aligned} f(x) &= \frac{1}{B(a,b)} \frac{\alpha x_0^{\alpha}}{x^{\alpha+1}} \left[1-\lambda+2\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]\left[1-\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a-1}\left[1+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]^{a-1}\left\{\left(\frac{x_0}{x}\right)^{\alpha}\left[1-\lambda+\lambda\left(\frac{x_0}{x}\right)^{\alpha}\right]\right\}^{b-1} \\ &= \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{\alpha(-1)^i}{B(a,b)} \binom{a-1}{i}\binom{a-1}{k}\binom{b-1}{l} \lambda^{k+l}(1-\lambda)^{b-1-l}\left\{(1-\lambda)x^{-1}+2\lambda\left(\frac{x_0}{x}\right)^{\alpha}x^{-1}\right\}\left(\frac{x_0}{x}\right)^{\alpha(i+k+l+b)} \end{aligned} \tag{12}$$

Further letting $m = i+k+l$, $f(x)$ reduces to

$$\begin{aligned} f(x) &= \sum_{m=0}^{\infty}\sum_{i=0}^{m}\sum_{k=0}^{m-i} \frac{\alpha(-1)^i}{B(a,b)} \binom{a-1}{i}\binom{a-1}{k}\binom{b-1}{m-i-k} \lambda^{m-i}(1-\lambda)^{b-1-m+i+k} \\ &\quad \times \left(\frac{x_0}{x}\right)^{\alpha(m+b)} \left[(1-\lambda)x^{-1}+2\lambda\left(\frac{x_0}{x}\right)^{\alpha}x^{-1}\right] \\ &= \left[(1-\lambda)x^{-1}+2\lambda\left(\frac{x_0}{x}\right)^{\alpha}x^{-1}\right]\sum_{m=0}^{\infty} c_m \left(\frac{x_0}{x}\right)^{\alpha(m+b)}, \end{aligned} \tag{13}$$

where

$$c_m = \sum_{i=0}^{m}\sum_{k=0}^{m-i} \frac{\alpha(-1)^i}{B(a,b)} \binom{a-1}{i}\binom{a-1}{k}\binom{b-1}{m-i-k} \lambda^{m-i}(1-\lambda)^{b-1-m+i+k}. \tag{14}$$

4 Mathematical characterizations

In this section we provide some mathematical properties of the BTP distribution including the moments, quantiles, mean deviations, probability weighted moments, residual life, distribution of order statistic etc.

4.1 Moments and moment generating function

Moments are necessary and important in any statistical analysis, especially in applications. It can be used to study the most important features and characteristics of a distribution (e.g., tendency, dispersion, skewness and kurtosis). Using the mixture representation described in section 3, the r-th moment of the BTP random variable X is given by

$$
\begin{aligned}
E(X^r) &= \int_{x_0}^{\infty} x^r f(x)dx \\
&= \int_{x_0}^{\infty} x^r \left[(1-\lambda)x^{-1} + 2\lambda\left(\frac{x_0}{x}\right)^{\alpha} x^{-1}\right] \sum_{m=0}^{\infty} c_m \left(\frac{x_0}{x}\right)^{\alpha(m+b)} dx \\
&= \sum_{m=0}^{\infty} c_m \left[(1-\lambda)x_0^{\alpha m+\alpha b}\int_{x_0}^{\infty} x^{r-1-\alpha m-\alpha b}dx + 2\lambda x_0^{\alpha m+\alpha b+\alpha}\int_{x_0}^{\infty} x^{r-1-\alpha m-\alpha b-\alpha}dx\right] \\
&= \sum_{m=0}^{\infty} c_m \left[(1-\lambda)x_0^{\alpha m+\alpha b}\frac{x_0^{r-\alpha m-\alpha b}}{\alpha m+\alpha b-r} + 2\lambda x_0^{\alpha m+\alpha b+\alpha}\frac{x_0^{r-\alpha m-\alpha b-\alpha}}{\alpha m+\alpha b+\alpha-r}\right] \\
&= x_0^r \sum_{m=0}^{\infty} c_m \left(\frac{1-\lambda}{\alpha m+\alpha b-r} + \frac{2\lambda}{\alpha m+\alpha b+\alpha-r}\right) \qquad (15)
\end{aligned}
$$

if $r < \alpha b$.

Similarly, the moment generating function of X may be obtained as below:

$$
\begin{aligned}
M_X(t) &= \int_{x_0}^{\infty} e^{tx} f(x)dx \\
&= \int_{x_0}^{\infty} e^{tx} \left[(1-\lambda)x^{-1} + 2\lambda\left(\frac{x_0}{x}\right)^{\alpha} x^{-1}\right] \sum_{m=0}^{\infty} c_m \left(\frac{x_0}{x}\right)^{\alpha(m+b)} dx \\
&= \sum_{m=0}^{\infty} c_m \left[(1-\lambda)x_0^{\alpha m+\alpha b}\int_{x_0}^{\infty} e^{tx}x^{-1-\alpha m-\alpha b}dx + 2\lambda x_0^{\alpha m+\alpha b+\alpha}\int_{x_0}^{\infty} e^{tx}x^{-1-\alpha m-\alpha b-\alpha}dx\right] \\
&= \sum_{m=0}^{\infty} c_m(-tx_0)^{\alpha(m+b)}\left[(1-\lambda)\Gamma(-\alpha b-\alpha m,-tx_0) + 2\lambda(-tx_0)^{\alpha}\Gamma(-\alpha b-\alpha m-\alpha,-tx_0)\right], \qquad (16)
\end{aligned}
$$

where $t < 0$ and $\Gamma(.,.)$ denotes the upper incomplete gamma function, i.e. $\Gamma(s,x) = \int_x^{\infty} e^{-t}t^{s-1}dt$.

4.2 Quantiles and random number generator

Quantiles are the points in a distribution that relates to the rank order of values. The quantile function of a distribution is the real solution of $F(x_q) = q$ for $0 \le q \le 1$. The quantiles of BTP are obtained from (7) as

$$x_q = x_0\left(\frac{\lambda-1+\sqrt{(1+\lambda)^2-4\lambda\left(I_q^{-1}(a,b)\right)}}{2\lambda}\right)^{-\frac{1}{\alpha}}, \qquad (17)$$

where $I_q^{-1}(a,b)$ is the inverse of the incomplete beta function with parameters a and b. As shown in Zea et al. [20], the function $I_q^{-1}(a,b)$ can be expressed as a power series

$$I_q^{-1}(a,b) = \sum_{i=1}^{\infty} q_i[aB(a,b)q]^{\frac{i}{a}},$$

where $q_1 = 1$ and the remaining coefficients satisfy the following recursion

$$q_i = \frac{1}{i^2+(a-2)i+(1-a)}\left\{\sum_{r=2}^{i-1}(1-\delta_{i,2})q_r q_{i+1-r}\left[r(1-a)(i-r)-r(r-1)\right]\right.$$

$$\left.+\sum_{r=1}^{i-1}\sum_{s=1}^{i-r} q_r q_{i+1-r-s}\left[r(r-a)+s(a+b-2)(i+1-r-s)\right]\right\},$$

where $\delta_{i,2} = 1$ if $i = 2$, $\delta_{i,2} = 0$ if $i \neq 2$. We can also find the expression of the inverse incomplete beta function on the website: http://functions.wolfram.com/06.23.06.0004.01, which is also mentioned by Pal and Tiensuwan [18].
We use the inverse transformation method to generate random numbers from the beta transmuted Pareto distribution as $F(x) = u$, where $u \sim U(0,1)$. Solving the expression $F(x) = u$ gives

$$x = x_0 \left(\frac{\lambda - 1 + \sqrt{(1+\lambda)^2 - 4\lambda \left(I_u^{-1}(a,b)\right)}}{2\lambda} \right)^{-\frac{1}{\alpha}}, \qquad 0 < u < 1 \tag{18}$$

where $I_u^{-1}(a,b)$ is the inverse of the incomplete beta function. The shortcomings of the classical skewness and kurtosis are well-known in the literature. To illustrate the effect of the parameters λ, a and b on skewness and kurtosis we consider measures based on quantiles. The Bowley's skewness by Kenney and Keeping [11] is one of the earliest skewness measures defined in terms of the quantiles as below

$$B = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} = \frac{Q_{0.75} - 2Q_{0.5} + Q_{0.25}}{Q_{0.75} - Q_{0.25}}$$

and the Moors kurtosis by Moors [15] is defined as

$$M = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2} = \frac{Q_{0.875} - Q_{0.625} + Q_{0.375} - Q_{0.125}}{Q_{0.75} - Q_{0.25}},$$

where $Q(.)$ represents the quantiles and E(.) represents the octiles. Figure 2 displays the Bowley (B) skewness and Moors (M) kurtosis as a function of the parameter λ for $a = b = 5, \alpha = 2$ and $x_0 = 0.1$.

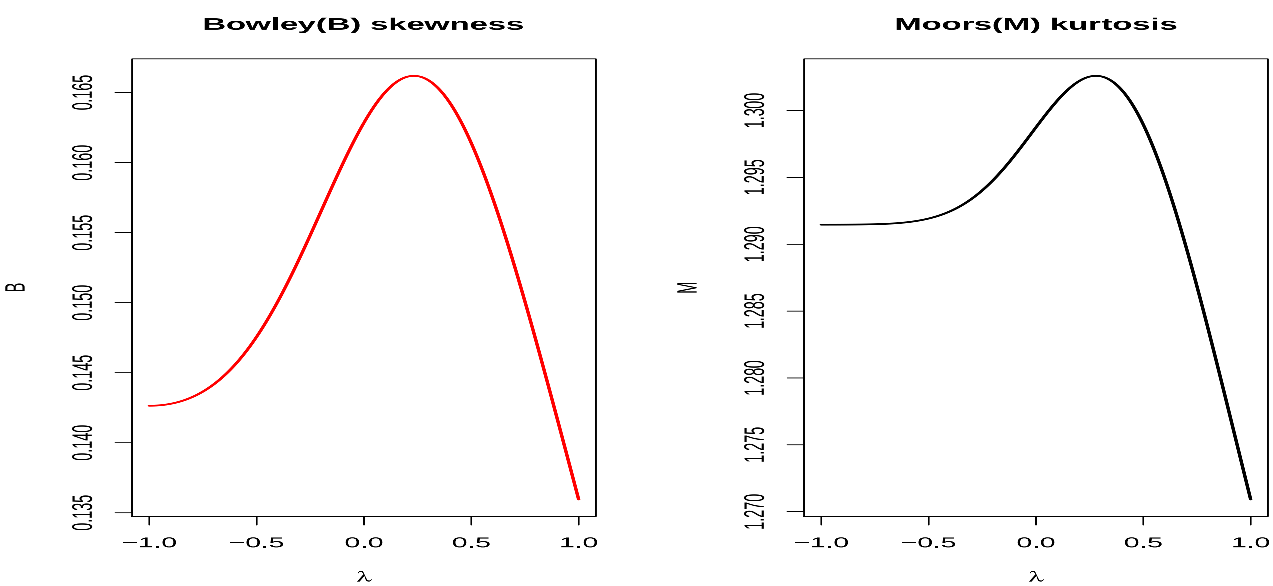

Fig. 2: Bowley skewness (left) and Moors kurtosis (right) of BTP distribution as a function of λ.

Similarly, Figure 3 displays the Bowley (B) skewness and Moors (M) kurtosis for different values of a and b with $\lambda = 0.2, \alpha = 2$ and $x_0 = 0.1$.

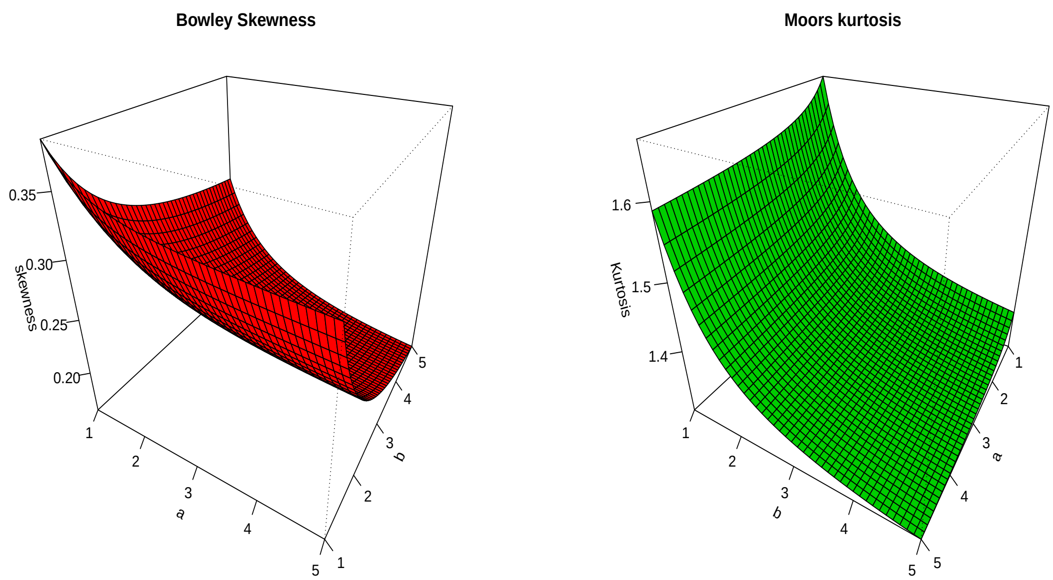

Fig. 3: Bowley skewness (left) and Moors kurtosis (right) of BTP distribution as a function of the parameters a and b.

From Figures 2 and 3 it is evident that both Bowley's skewness and Moors kurtosis depend on the choice of the parameters.

4.3 Mean Deviation

Let X be a beta transmuted Pareto random variable with mean $\mu = E(X)$ and median M. Note that we can find the mean μ by substituting $r = 1$ in equation (15) provided that $\alpha b > 1$. The mean deviation from the mean (μ) and the mean deviation from the median (M) can be expressed as

$$\delta_1(x) = \int_{x_0}^{\infty} |x-\mu| f(x)dx = \int_{x_0}^{\mu} (\mu - x) f(x)dx + \int_{\mu}^{\infty} (x-\mu) f(x)dx = 2[\mu F(\mu) - J(\mu)], \tag{19}$$

$$\delta_2(x) = \int_{x_0}^{\infty} |x-M| f(x)dx = \int_{x_0}^{M} (M - x) f(x)dx + \int_{M}^{\infty} (x-M) f(x)dx = \mu - 2J(M), \tag{20}$$

where $F(.)$ is the cdf of the beta transmuted Pareto distribution and $J(t) = \int_{x_0}^{t} x f(x)dx$.
Similar to (15), we can compute $J(t)$ as below:

$$\begin{aligned} J(t) &= \int_{x_0}^{t} x f(x)dx \\ &= \int_{x_0}^{t} x \times \left[(1-\lambda)x^{-1} + 2\lambda \left(\frac{x_0}{x}\right)^{\alpha} x^{-1}\right] \sum_{m=0}^{\infty} c_m \left(\frac{x_0}{x}\right)^{\alpha(m+b)} dx \\ &= \sum_{m=0}^{\infty} c_m x_0^{\alpha(m+b)} \left[(1-\lambda)\int_{x_0}^{t} x^{-\alpha(m+b)}dx + 2\lambda x_0^{\alpha} \int_{x_0}^{t} x^{-\alpha(m+b+1)}dx\right] \\ &= \sum_{m=0}^{\infty} c_m \left[\frac{(1-\lambda)t(\frac{x_0}{t})^{\alpha(m+b)}}{1-\alpha m-\alpha b} + \frac{(1-\lambda)x_0}{\alpha m+\alpha b-1} + \frac{2\lambda t(\frac{x_0}{t})^{\alpha(m+b+1)}}{1-\alpha m-\alpha b-\alpha} + \frac{2\lambda x_0}{\alpha m+\alpha b+\alpha-1}\right]. \end{aligned} \tag{21}$$

4.4 Probability weighted moments

The probability weighted moments are the expectations of certain functions of a random variable and can be defined for any random variable whose ordinary moments exist. The (s,r)th probability weighted moments of a random variable X which follows BTP, say $\rho_{s,r}$, is defined by

$$\rho_{s,r} = E\{X^s F(X)^r\} = \int_{x_0}^{\infty} x^s F(x)^r f(x)dx. \tag{22}$$

From (10), we have

$$F(x) = \sum_{k,l=0}^{\infty} w_{kl} v^{k+l} = \sum_{p=0}^{\infty} w_p^* v^p, \tag{23}$$

where $v = \left(\frac{x_0}{x}\right)^{\alpha}$ and $w_p^* = \sum_{k,l:k+l=p} w_{kl}$. As mentioned in section 0.314 of Gradshteyn and Ryzhik [10], for any positive integer r,

$$\left(\sum_{p=0}^{\infty} w_p^* v^p\right)^r = \sum_{p=0}^{\infty} d_{r,p} v^p, \tag{24}$$

where the coefficients $d_{r,p}$ for $p = 1,2,3,....$ can be determined from the recurrence equation

$$d_{r,p} = (p w_0^*)^{-1} \sum_{q=1}^{p} \{q(r+1)-p\} w_q^* d_{r,p-q} \tag{25}$$

and $d_{r,0} = (w_0^*)^r$.
Thus we can obtain $d_{r,p}$ from $d_{r,0},\cdots,d_{r,p-1}$ and, therefore, from $w_0^*, w_1^*,\cdots,w_p^*$. Using equations (13), (23), (24) and (25), equation (22) reduces to

$$\begin{aligned}
\rho_{s,r} &= \int_{x_0}^{\infty} x^s \left(\sum_{p=0}^{\infty} w_p^* \left(\frac{x_0}{x}\right)^{\alpha p}\right)^r \sum_{m=0}^{\infty} c_m \left(\frac{x_0}{x}\right)^{\alpha(m+b)} \left[(1-\lambda)x^{-1} + 2\lambda\left(\frac{x_0}{x}\right)^{\alpha} x^{-1}\right] dx \\
&= \int_{x_0}^{\infty} x^s \sum_{p=0}^{\infty} d_{r,p} \left(\frac{x_0}{x}\right)^{\alpha p} \sum_{m=0}^{\infty} c_m \left(\frac{x_0}{x}\right)^{\alpha(m+b)} \left[(1-\lambda)x^{-1} + 2\lambda\left(\frac{x_0}{x}\right)^{\alpha} x^{-1}\right] dx \\
&= \sum_{p=0}^{\infty}\sum_{m=0}^{\infty} d_{r,p} c_m x_0^{\alpha(m+b+p)} \times \left[(1-\lambda)\int_{x_0}^{\infty} x^{-1+s-\alpha m-\alpha b-\alpha p} dx + 2\lambda x_0^{\alpha} \int_{x_0}^{\infty} x^{-1-\alpha-\alpha m-\alpha b-\alpha p+s} dx\right] \\
&= \sum_{p=0}^{\infty}\sum_{m=0}^{\infty} d_{r,p} c_m x_0^{\alpha(m+b+p)} \left[(1-\lambda)\frac{x_0^{s-\alpha m-\alpha b-\alpha p}}{\alpha m+\alpha b+\alpha p-s} + 2\lambda x_0^{\alpha} \frac{x_0^{-\alpha+s-\alpha m-\alpha b-\alpha p}}{\alpha m+\alpha b+\alpha p+\alpha-s}\right] \\
&= \sum_{p=0}^{\infty}\sum_{m=0}^{\infty} d_{r,p} c_m x_0^{s} \left[\frac{1-\lambda}{\alpha m+\alpha b+\alpha p-s} + \frac{2\lambda}{\alpha m+\alpha b+\alpha p+\alpha-s}\right]
\end{aligned} \tag{26}$$

where $s < \alpha b$.

4.5 Residual and reversed residual life

Let X be a BTP random variable and $F(x)$ be its cdf (7). Then the n-th moment of the residual life of X, say, $m_n(t) = E\left[(X-t)^n | X > t\right]$, $n = 1,2,\cdots$ is given by

$$\begin{aligned}
m_n(t) &= \frac{1}{R(t)}\int_t^{\infty} (x-t)^n dF(x) = \frac{1}{R(t)}\int_t^{\infty} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k x^{n-k} t^k f(x) dx \\
&= \frac{1}{R(t)} \sum_{k=0}^{n} \binom{n}{k} (-1)^k t^k \sum_{m=0}^{\infty} c_m x_0^{\alpha(m+b)} \left[(1-\lambda)\int_t^{\infty} x^{-1+n-k-\alpha m-\alpha b} dx + 2\lambda x_0^{\alpha}\int_t^{\infty} x^{-1-\alpha+n-k-\alpha m-\alpha b} dx\right] \\
&= \frac{1}{R(t)} \sum_{k=0}^{n}\sum_{m=0}^{\infty} c_m \binom{n}{k} (-1)^k t^k x_0^{\alpha(m+b)} \left[\frac{(1-\lambda)t^{n-k-\alpha m-\alpha b}}{\alpha m+\alpha b-n+k} + \frac{2\lambda x_0^{\alpha} t^{n-k-\alpha m-\alpha b-\alpha}}{\alpha m+\alpha b+\alpha-n+k}\right] \\
&= \frac{1}{R(t)} \sum_{k=0}^{n}\sum_{m=0}^{\infty} c_m \binom{n}{k} (-1)^k t^n \left(\frac{x_0}{t}\right)^{\alpha(m+b)} \left[\frac{(1-\lambda)}{\alpha m+\alpha b-n+k} + \frac{2\lambda\left(\frac{x_0}{t}\right)^{\alpha}}{\alpha m+\alpha b+\alpha-n+k}\right]
\end{aligned} \tag{27}$$

where $n < \alpha b$ and $R(t) = 1 - F(t)$.

We set $n = 1$ in equation (27) to get mean residual life. The n-th moments of the reversed residual life of X, say, $M_n(t) = E\left[(X-t)^n | X \le t\right]$ for $t > 0$ and $n = 1, 2, \cdots$ uniquely determine $F(x)$. We have

$$M_n(t) = \frac{1}{F(t)} \int_t^{\infty} (t-x)^n dF(x) \tag{28}$$

and follow similar steps that we used to derive the expression (27) to obtain a series expansion formula for $M_n(t)$. We set $n = 1$ in equation (28) to get mean reversed residual life. The mean reversed life is also known as mean inactivity time or mean waiting time.

4.6 Stress-strength model

The stress-strength model is widely used in many applications of physics and engineering such as strength failure and system collapse. In stress-strength modeling, we use $R = P(X_2 < X_1)$ as a measure of reliability of the system with random stress X_2 and strength X_1. Let X_1 and X_2 be two independent random variables with BTP$(x; \alpha, x_0, \lambda, a_1, b_1)$ and BTP$(x; \alpha, x_0, \lambda, a_2, b_2)$ distributions respectively. The reliability can be computed by

$$R = \int_{x_0}^{\infty} f_1(x; \alpha, x_0, \lambda, a_1, b_1) F_2(x; \alpha, x_0, \lambda, a_2, b_2) dx, \tag{29}$$

where $f_1(.)$ and $F_2(.)$ are the pdf and cdf of the BTP random variables X_1 and X_2 respectively. Note that the pdf of X_1 is given by

$$f_1(x; \alpha, x_0, \lambda, a_1, b_1) = \left[(1-\lambda)x^{-1} + 2\lambda \left(\frac{x_0}{x}\right)^{\alpha} x^{-1}\right] \sum_{m=0}^{\infty} c_m^{(1)} \left(\frac{x_0}{x}\right)^{\alpha(m+b_1)}$$

and the cdf of X_2 is given by

$$F_2(x; \alpha, x_0, \lambda, a_2, b_2) = \sum_{k,l=0}^{\infty} w_{kl}^{(2)} \left(\frac{x_0}{x}\right)^{\alpha(k+l)},$$

where $c_m^{(1)}$ and $w_{kl}^{(2)}$ are given by

$$c_m^{(1)} = \sum_{i=0}^{m} \sum_{k=0}^{m-i} \frac{\alpha(-1)^i}{B(a_1, b_1)} \binom{a_1 - 1}{i} \binom{a_1 - 1}{k} \binom{b_1 - 1}{m - i - k} \lambda^{m-i} (1-\lambda)^{b_1 - 1 - m + i + k},$$

and

$$w_{kl}^{(2)} = \sum_{i=0}^{\infty} (-1)^{i+k} \binom{b_2 - 1}{i} \binom{a_2 + i}{k} \binom{a_2 + i}{l} \frac{\lambda^l}{B(a_2, b_2)(a_2 + i)}.$$

Hence, equation (29) reduces to

$$\begin{aligned} R &= \sum_{m,k,l=0}^{\infty} c_m^{(1)} w_{kl}^{(2)} x_0^{\alpha(k+l+m+b_1)} \left[\int_{x_0}^{\infty} (1-\lambda) x^{-1-\alpha m - \alpha b_1 - \alpha k - \alpha l} dx + \int_{x_0}^{\infty} 2\lambda x_0^{\alpha} x^{-1-\alpha-\alpha m - \alpha b_1 - \alpha k - \alpha l} dx\right] \\ &= \sum_{m,k,l=0}^{\infty} c_m^{(1)} w_{kl}^{(2)} x_0^{\alpha(k+l+m+b_1)} \left[\frac{1-\lambda}{\alpha m + \alpha b_1 + \alpha k + \alpha l} + \frac{2\lambda x_0^{\alpha}}{\alpha + \alpha m + \alpha b_1 + \alpha k + \alpha l}\right]. \end{aligned} \tag{30}$$

4.7 Order statistics

Let $X_1, X_2, \cdots, X_n$ be a simple random sample from BTP$(x; \alpha, x_0, \lambda, a, b)$ with cumulative distribution function (7) and probability density function (8) .
Let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ denote the order statistics from this sample. The pdf $f_{(i:n)}(x)$ of ith order statistics is given by

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x) \tag{31}$$

and the cdf is given by

$$F_{i:n}(x) = \sum_{k=i}^{n} \binom{n}{k} [F(x)]^k [1-F(x)]^{n-k} = \int_0^{F(x)} \frac{1}{B(i,n-i+1)} t^{i-1}(1-t)^{n-i} dt. \tag{32}$$

Substituting $v = \left(\frac{x_0}{x}\right)^\alpha$ and using (13), (23), (24), (25) and binomial expansions, equation (31) reduces to

$$\begin{aligned} f_{i:n}(x) &= \frac{1}{B(i,n-i+1)} f(x) \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} F(x)^{i+s-1} \\ &= \frac{\left[(1-\lambda)x^{-1} + 2\lambda v x^{-1}\right]}{B(i,n-i+1)} \sum_{m=0}^{\infty} c_m v^{m+b} \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} \left(\sum_{p=0}^{\infty} w_p^* v^p\right)^{i+s-1} \\ &= \frac{\left[(1-\lambda)x^{-1} + 2\lambda v x^{-1}\right]}{B(i,n-i+1)} \sum_{m=0}^{\infty} c_m v^{m+b} \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} \left(\sum_{p=0}^{\infty} d_{i+s-1,p} v^p\right) \\ &= \frac{\left[(1-\lambda)x^{-1} + 2\lambda v x^{-1}\right]}{B(i,n-i+1)} \sum_{m=0}^{\infty} \sum_{p=0}^{\infty} \left[\sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} c_m d_{i+s-1,p}\right] v^{m+b+p} \\ &= \frac{\left[(1-\lambda) + 2\lambda v\right] v^b}{x} \sum_{m=0}^{\infty} \sum_{p=0}^{\infty} c_i(m,p) v^{m+p}, \end{aligned} \tag{33}$$

where $c_i(m,p) = \frac{1}{B(i,n-i+1)} \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} c_m d_{i+s-1,p}$, $d_{i+s-1,p} = (p w_0^*)^{-1} \sum_{q=1}^{p} \left[q(i+s) - p\right] w_p^* d_{i+s-1,p-q}$ and $d_{i+s-1,0} = (w_0^*)^{i+s-1} = \left(\sum_{j=0}^{\infty} (-1)^j \binom{b-1}{j} \frac{1}{B(a,b)(a+j)}\right)^{i+s-1}$.

5 Parameter estimation

Several methods for parameter estimation have been proposed in the literature but the maximum likelihood method is the most commonly employed method. The maximum likelihood estimators (MLEs) enjoy desirable properties and can be used for constructing confidence intervals for the model parameters and also for hypothesis testing. So, we consider the estimation of the unknown parameters for this family from complete samples only by maximum likelihood. Let $X_1, X_2, \cdots, X_n$ be a random sample from the BTP distribution with observed values $x_1, x_2, \cdots, x_n$ and $\Theta = (\alpha, \lambda, a, b, x_0)^T$ be parameter vector. The likelihood function for Θ may be expressed as

$$L(\Theta) = \frac{1}{(B(a,b))^n} \prod_{i=1}^{n} \frac{\alpha x_0^\alpha}{x_i^{\alpha+1}} \left[1 - \lambda + 2\lambda \left(\frac{x_0}{x_i}\right)^\alpha\right] \left[1 - \left(\frac{x_0}{x_i}\right)^\alpha\right]^{a-1} \left[1 + \lambda \left(\frac{x_0}{x_i}\right)^\alpha\right]^{a-1} \left\{\left(\frac{x_0}{x_i}\right)^\alpha \left[1 - \lambda + \lambda \left(\frac{x_0}{x_i}\right)^\alpha\right]\right\}^{b-1} \tag{34}$$

Therefore, the log-likelihood function for Θ becomes

$$\begin{aligned} l(\Theta) &= n \ln(\Gamma(a+b)) - n \ln(\Gamma(a)) - n \ln(\Gamma(b)) + n \ln \alpha + n\alpha \ln(x_0) - (\alpha+1) \sum_{i=1}^{n} \ln x_i + \sum_{i=1}^{n} \ln\left(1 - \lambda + 2\lambda \left(\frac{x_0}{x_i}\right)^\alpha\right) \\ &\quad + (a-1)\left[\sum_{i=1}^{n} \ln\left(1 - \left(\frac{x_0}{x_i}\right)^\alpha\right) + \sum_{i=1}^{n} \ln\left(1 + \lambda \left(\frac{x_0}{x_i}\right)^\alpha\right)\right] + (b-1) \sum_{i=1}^{n} \left[\alpha \ln\left(\frac{x_0}{x_i}\right) + \ln\left(1 - \lambda + \lambda \left(\frac{x_0}{x_i}\right)^\alpha\right)\right] \end{aligned} \tag{35}$$

Since $x \in (x_0, \infty)$, the maximum likelihood estimator of x_0 is the first order statistic $X_{(1)}$. Next we discuss the maximum likelihood estimation for α, λ, a, and b. We differentiate (35) with respect to α, λ, a and b respectively to obtain the elements of score vector $\left(\frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial a}, \frac{\partial l}{\partial b}\right)^T$ as below

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + n\ln(x_0) - \sum_{i=1}^{n}\ln(x_i) + \sum_{i=1}^{n}\frac{2\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}\ln\left(\frac{x_0}{x_i}\right)}{1-\lambda+2\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}} + (a-1)\sum_{i=1}^{n}\frac{\left(\frac{x_0}{x_i}\right)^{\alpha}\ln\left(\frac{x_0}{x_i}\right)}{\left(\frac{x_0}{x_i}\right)^{\alpha}-1}$$
$$+(a-1)\sum_{i=1}^{n}\frac{\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}\ln\left(\frac{x_0}{x_i}\right)}{1+\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}} + (b-1)\left[\sum_{i=1}^{n}\ln\left(\frac{x_0}{x_i}\right) + \frac{\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}\ln\left(\frac{x_0}{x_i}\right)}{1-\lambda+\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}}\right], \tag{36}$$

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n}\frac{2\left(\frac{x_0}{x_i}\right)^{\alpha}-1}{1-\lambda+2\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}} + (a-1)\sum_{i=1}^{n}\frac{\left(\frac{x_0}{x_i}\right)^{\alpha}}{1+\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}} + (b-1)\sum_{i=1}^{n}\frac{\left(\frac{x_0}{x_i}\right)^{\alpha}-1}{1-\lambda+\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}}, \tag{37}$$

$$\frac{\partial l}{\partial a} = n\left[\psi(a+b)-\psi(a)\right] + \sum_{i=1}^{n}\ln\left(1-\left(\frac{x_0}{x_i}\right)^{\alpha}\right) + \sum_{i=1}^{n}\ln\left(1+\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}\right), \tag{38}$$

$$\frac{\partial l}{\partial b} = n\left[\psi(a+b)-\psi(b)\right] + \sum_{i=1}^{n}\left[\alpha\ln\left(\frac{x_0}{x_i}\right) + \ln\left(1-\lambda+\lambda\left(\frac{x_0}{x_i}\right)^{\alpha}\right)\right], \tag{39}$$

where $\psi(.)$ is the digamma function, i.e. $\psi(x) = \frac{d}{dx}(\ln\Gamma(x))$.

The maximum likelihood estimators $\hat{\alpha}, \hat{\lambda}, \hat{a}, \hat{b}$ of the unknown parameters α, λ, a, b respectively, can be obtained by setting the score vector to zero and solving the system of nonlinear equations simultaneously. Since there is no closed form solution of these non-linear system of equations, we can use numerical methods such as the quasi-Newton algorithm to numerically optimize the log-likelihood function given in (35) to get the maximum likelihood estimates of the parameters α, λ, a, b. To compute the standard error and the asymptotic confidence interval, we use the usual large sample approximation in which the maximum likelihood estimators for Θ can be treated as being approximately normal.

6 Reliability analysis

The survival function, also known as the reliability function in engineering, of a probability distribution is the characteristic of an explanatory variable that maps a set of events, usually associated with mortality or failure of some system onto time. It is the probability that the system will survive beyond a specified time. The reliability function $R(t)$ is defined by $R(t) = 1 - F(t)$, where $F(.)$ is the cdf of the distribution. The other characteristic of interest of a random variable is its hazard rate function which is also known as instantaneous failure rate of a random variable X which is an important quantity characterizing life phenomenon. The hazard function $h(t)$ is defined as

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

where $F(.)$ and $f(.)$ are, respectively, the cdf and pdf of the given distribution.Using equations (7) and (8), the hazard rate function of the BTP distribution can be expressed as

$$h(t) = \frac{\alpha x_0^{\alpha}}{t^{\alpha+1}}\frac{\left[1-\lambda+2\lambda\left(\frac{x_0}{t}\right)^{\alpha}\right]\left[1-\left(\frac{x_0}{t}\right)^{\alpha}\right]^{a-1}}{B(a,b)I_{\left[\left(\frac{x_0}{t}\right)^{\alpha}\left\{1-\lambda+\lambda\left(\frac{x_0}{t}\right)^{\alpha}\right\}\right]}(b,a)}\left[1+\lambda\left(\frac{x_0}{t}\right)^{\alpha}\right]^{a-1}\left[\left(\frac{x_0}{t}\right)^{\alpha}\left\{1-\lambda+\lambda\left(\frac{x_0}{t}\right)^{\alpha}\right\}\right]^{b-1}. \tag{40}$$

The flexibility of BTP distribution to model reliability data is illustrated by varying shape of reliability function and hazard rate function in Figure 4.

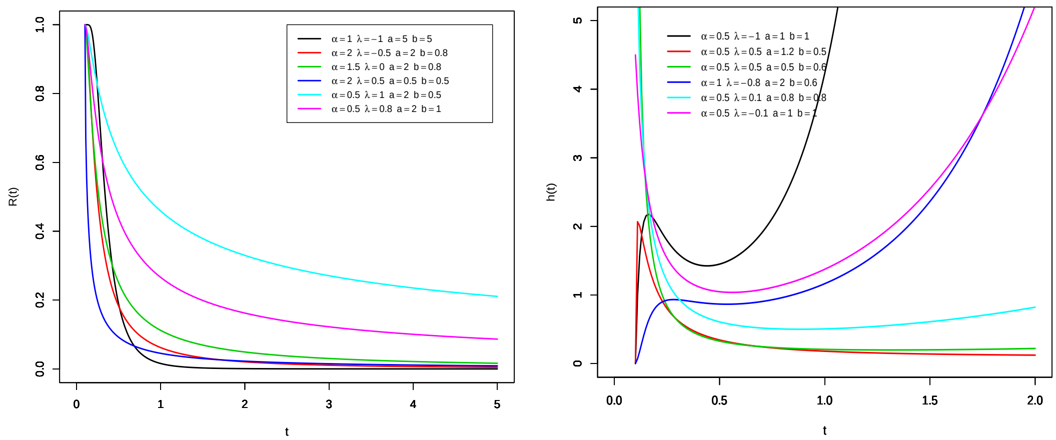

Fig. 4: Reliability function (left) and hazard rate function (right) of BTP distribution.

Lemma 1 below provides the limiting behavior of the hazard rate function.

Lemma 1. *If $h(t)$ is the hazard function of the beta transmuted Pareto distribution, then*

$$\lim_{t\to x_0} h(t) = \begin{cases} 0, & if\ a > 1 \\ \frac{(1+\lambda)\alpha}{x_0 B(a,b)}, & if\ a = 1 \\ \infty, & if\ a < 1 \end{cases}$$

and

$$\lim_{t\to\infty} h(t) = 0.$$

Proof. First note that $\lim_{t\to x_0}(1-F(t)) = 1$ and

$$\lim_{t\to x_0}\left[1-\left(\frac{x_0}{t}\right)^{\alpha}\right]^{a-1} = \begin{cases} 0, & \text{if } a > 1 \\ 1, & \text{if } a = 1 \\ \infty, & \text{if } a < 1. \end{cases}$$

Then we have

$$\lim_{t\to x_0} h(t) = \lim_{t\to x_0} \frac{f(t)}{1-F(t)} = \lim_{t\to x_0} f(t) = \begin{cases} 0, & \text{if } a > 1 \\ \frac{\alpha(1+\lambda)}{x_0 B(a,b)}, & \text{if } a = 1 \\ \infty, & \text{if } a < 1. \end{cases}$$

Since $\lim_{t\to\infty} f(t) = 0$ and $\lim_{t\to\infty} F(t) = 1$, using L'Hospital's rule we have

$$\lim_{t\to\infty} h(t) = \lim_{t\to\infty} \frac{f(t)}{1-F(t)} = \lim_{t\to\infty} \frac{f'(t)}{-f(t)}$$

Note that $f(t) = \mathscr{O}(t^{-\alpha b-1})$ and $f'(t) = \mathscr{O}(t^{-\alpha b-2})$. We have

$$\lim_{t\to\infty} h(t) = \lim_{t\to\infty} \frac{f(t)}{1-F(t)} = \lim_{t\to\infty} \frac{c_1 t^{-\alpha b-2}}{c_2 t^{-\alpha b-1}} = \lim_{t\to\infty} \frac{c_1}{c_2 t} = 0,$$

where c_1 and c_2 are non-zero constants.

7 Application of beta transmuted Pareto distribution

In this section we illustrate the flexibility of the BTP distribution to model both heavy tailed and approximately symmetric data. We estimate the model parameters and calculate the goodness-of-fit statistics in order to assess the model. Our first data corresponds to the exceedances of flood peaks (in m^3/s) of the Wheaton river near Carcross in Yukon Territory, Canada. of 72 exceedances measures for the years 1958-1984, rounded to one decimal place and are provided below.

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.

These data were analyzed by many authors including Choulakian and Stephenes [6], Akinsete et al. [1], Nadarajsh [16], Merovci and Puka [14], Bourguignon et al. [5], among others. We have chosen the same data in order to compare our results with other models proposed by these authors. We estimate the parameters of the BTP model and compare its appropriateness to model this data with its submodels including beta Pareto (BP), transmuted Pareto (TP), exponentiated Pareto (EP) and Pareto (P) distributions. The required computations use a script *AdequacyModel* of the R-package by Marinho et al. [12]. Table 1 provides the estimated values and corresponding standard errors (in parentheses) of the model parameters.

Table 1: Estimated parameters and their standard errors- Wheaton river data.

model	a	b	λ	α	x_0
BTP	3.9118	17.3874	-0.8518	0.1159	0.1
	(1.8159)	(11.7365)	(0.2588)	(0.0509)	-
BP	3.1473	85.7508	0	0.0088	0.1
	(0.4993)	(0.0001)	-	(0.0015)	-
TP	1	1	-0.952	0.3490	0.1
	-	-	(0.089)	(0.072)	-
EP	2.8797	1	0	0.4241	0.1
	(0.4911)	-	-	(0.0463)	-
P	1	1	0	0.2438	0.1
	-	-	-	(0.0287)	-

The model selection is carried out by measuring the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the consistent Akaike information criteria (CAIC) and the Hannan-Quinn information criterion (HQIC). Note that the smaller the values of goodness-of-fit measures the better the fit of the data. These measures are defined as

$$AIC = -2\ell(\hat{\Theta}) + 2q, \qquad BIC = -2\ell(\hat{\Theta}) + q\ln(n),$$
$$HQIC = -2\ell(\hat{\Theta}) + 2q\ln(\ln(n)), \quad CAIC = -2\ell(\hat{\Theta}) + \frac{2qn}{n-q-1},$$

where $\ell(\hat{\Theta})$ denotes the log-likelihood function evaluated at the maximum likelihood estimates, q is the number of parameters, n is the sample size and Θ denote the parameters. The, $-\ell$, AIC, BIC, HQIC and CAIC statistics for each model is provided in Table 2. We also provide the Kolmogorov-Smirnov (KS) test statistic. It can be seen that BTP distribution leads to a better fit than any of its submodels.

One can perform the Likelihood Ratio (LR) test in order to assess whether BTP is superior than one of its submodels to characterize the subject data. Table 3 provides the results of the LR test (Null hypothesis (H_0) versus Alternative hypothesis (H_a)).

In all cases we rejected the null hypothesis and conclude that BTP is a superior distribution to model this data. Plots comparing the exact BTP distribution with its submodels for Wheaton river data is given in Figure 5. It is evident that the BTP fits better than any of its submodels.

Table 2: The AIC, CAIC, BIC, HQIC and KS test statistic-Wheaton river data.

Model	statistics					
	$-\ell(.,x)$	AIC	CAIC	BIC	HQIC	KS
BTP	256.577	521.154	521.760	530.204	524.753	0.1599
BP	283.700	573.400	573.753	580.230	576.119	0.1747
TP	286.201	576.402	576.575	580.954	578.214	0.2870
EP	287.300	578.600	578.774	583.153	580.413	0.1987
P	303.100	608.200	608.257	610.477	609.106	0.3324

Table 3: Results of likelihood ratio tests.

Model	H_0	H_a	LR- test statistic	df	p-value
BTP vs. BP	$\lambda = 0$	$\lambda \neq 0$	54.246	1	0.000
BTP vs. TP	$a = b = 1$	$a \neq 1 \& b \neq 1$	59.248	2	0.000
BTP vs. EP	$b = 1 \& \lambda = 0$	$b \neq 1 \& \lambda \neq 0$	61.446	2	0.000
BTP vs. P	$a = b = 1 \& \lambda = 0$	$a \neq 1, b \neq 1 \& \lambda \neq 0$	93.046	3	0.000

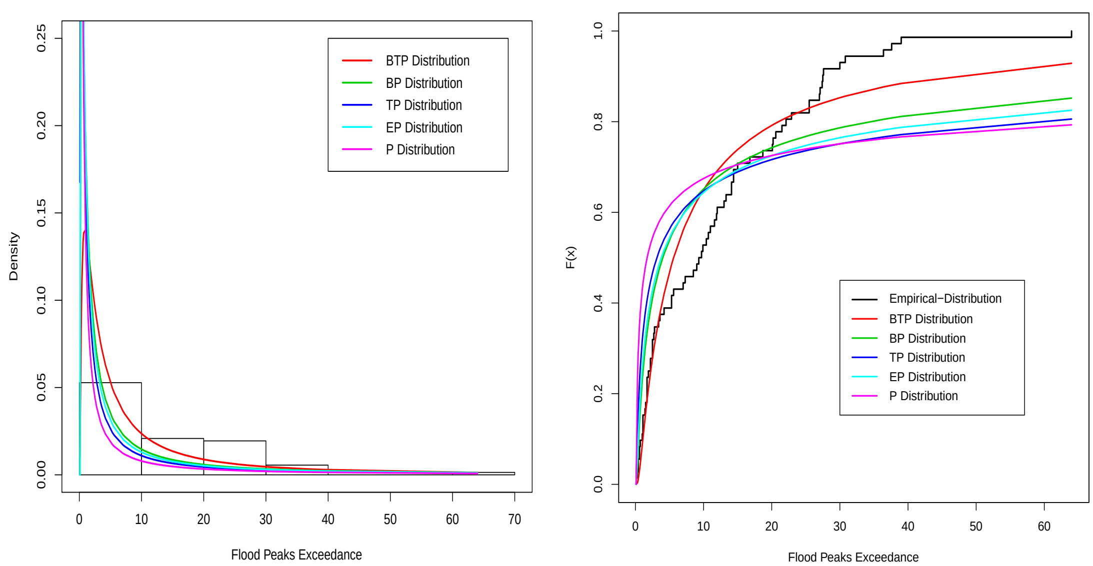

Fig. 5: Fitted pdf (left) and cdf (right) of BTP distribution and its submodels for Wheaton river data.

Next, we consider a slightly bigger and approximately symmetric data to present the usefulness of BTP distribution. We compare the results with some of the models generated from Pareto distribution. We consider the data from Mahmoudi [13] which represents the fatigue life of 6061-T6 aluminum coupons cut parallel with the direction of rolling and oscillated at 18 cycles per second. Recently, Alzaatreh et al. [3] also used this data to illustrate the usefulness of gamma-Pareto distribution. The estimated parameters of BTP distribution using this data are: $\hat{a} = 9.2490$, $\hat{b} = 30.5312$, $\hat{\lambda} = -0.7027$, $\hat{\alpha} = 0.8026$ and $\hat{x}_0 = 70$. The values of test statistic to measure the goodness of the BTP distribution are provided in Table 4. Readers are referred to Alzaatreh et al. [3] to compare and contrast the BTP with Pareto, beta Pareto, beta generalized Pareto and the gamma-Pareto distribution. Note that the $-\ell$ statistic for BTP is the lowest of all the models discussed in Alzaatreh et al. [3] but AIC value is slightly higher for BTP model than for gamma-Pareto due to the presence of more parameters in BTP than those in the gamma-Pareto model. The fitted pdf and cdf of the BTP distribution for this data are provided in Figure 6. It is evident that BTP distribution fits very well the fatigue life of 6061-T6 aluminum coupons data. This example suggests that the BTP distribution works well in fitting approximately symmetric data too.

Table 4: The AIC, CAIC, BIC, HQIC and KS test statistic-fatigue life of 6061-T6 aluminum coupons data.

Model	statistics					
	$-\ell(.,x)$	AIC	CAIC	BIC	HQIC	KS
BTP	447.5226	903.0453	903.4663	913.4660	907.2627	0.0984

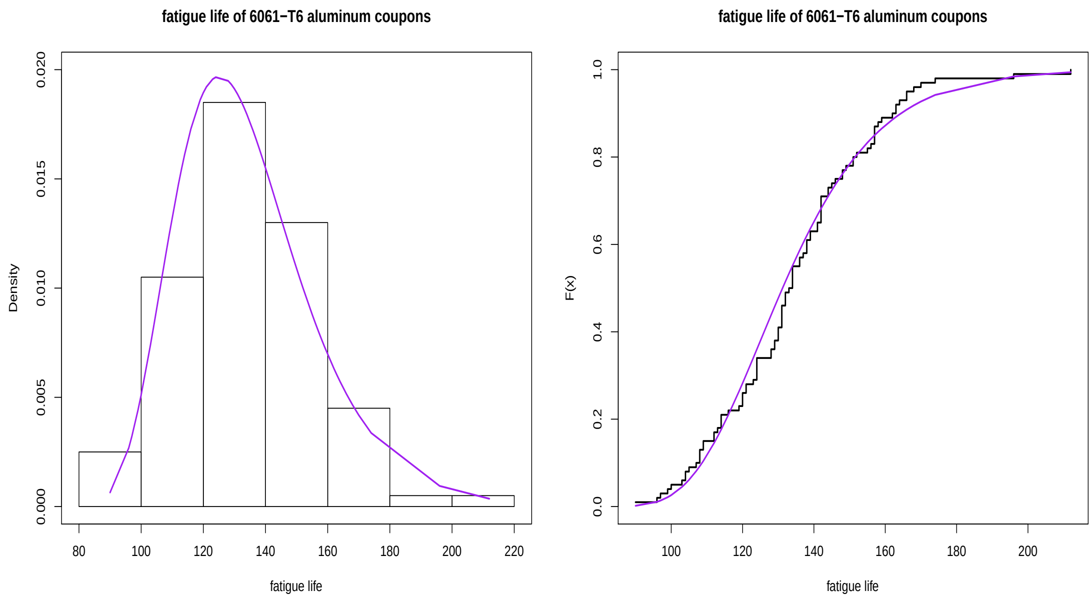

Fig. 6: Fitted pdf (left) and cdf (right) of BTP distribution for fatigue life data.

8 Concluding remarks

In this study, we have introduced the so-called beta transmuted Pareto (BTP) distribution. This is a generalization of the transmuted Pareto distribution using the genesis of the beta distribution. Many distributions including Pareto, beta Pareto, transmuted Pareto and exponentiated Pareto are embedded in this newly developed BTP distribution. Some mathematical properties along with parameter estimation issues of the subject distribution are discussed. We have presented two examples to illustrate the application of the subject distribution to model real world data. We have compared the goodness-of-fit with its competitive models and it has been shown that BTP is superior to model both heavy tailed and approximately symmetric data. We expect this study will serve as a reference and help to advance future research in the subject area.

Acknowledgements

The authors are grateful to the editor and anonymous reviewer for their constructive comments and valuable suggestions which certainly improved the presentation and quality of the paper.

References

[1] A. Akinsete, F. Famoye and C. Lee, The beta-Pareto distribution, Statistics **42**(6), 547-563 (2008).

[2] A. Alzaatreh, F. Famoye and C. Lee, Weibull-Pareto distribution and its applications, Communications in Statistics–Theory and Methods **42**, 1673-1691 (2013).

[3] A. Alzaatreh, F. Famoye and C. Lee, Gamma-Pareto distribution and its applications, Journal of Modern Applied Statistical Methods **11**(1), 78-94 (2012).

[4] S. M. Burroughs and S. F. Tebbens, Upper-truncated power law distributions, Fractals **9**, 209-222 (2001).

[5] M. Bourguignon, M. B. Silva, L. M. Zea and G. M. Cordeiro, The Kumaraswamy Pareto distribution, Journal of Statistical Theory and Applications **12**, 129-144 (2013).

[6] V. Choulakian and M. A. Stephenes, Goodness-of-fit for the generalized Pareto distribution, Technometrics **43**, 478-484 (2001).

[7] G. M. Cordeiro and S. Nadarajah, Closed form expressions of moments of a class of a beta generalized distributions, Brazilian Journal of Probability and Statistics **25**, 14-33 (2011).

[8] I. Elbatal, The Kumaraswamy exponentiated Pareto distribution, Economic Quality Control **28**, 1-9 (2015).

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

[10] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series, and Products. New York: Academic Press, (2000).

[11] J. F. Kenney and E. S. Keeping, Mathematics of Statistics, Princeton, NJ. (1962).

[12] P. R. D. Marinho, M. Bourguignon and C. R. B. Dias, Adequacy of probabilistic models and generation of pseudo-random numbers, R package-AdequacyModel, Available from: *http://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf*

[13] E. Mahmoudi, The beta generalized Pareto distribution with application to lifetime data, Mathematics and Computers in Simulation **81**, 2414-2430 (2011).

[14] F. Merovci and L. Puka, Transmuted Pareto distribution, ProbStat Forum **7**, 1-11 (2014).

[15] J. A. Moors, A quantile alternative for kurtosis, Journal of the Royal Statistical Society, D, **37**, 25-32 (1998).

[16] S. Nadarajah, Exponentiated Pareto distributions, Statistics **39**, 255-260 (2005).

[17] M. E. J. Newman, Power laws, Pareto distributions and Zipfs law, Contemporary Physics **46**(5), 323-351 (2005).

[18] M. Pal and M. Tiensuwan, The beta transmuted Weibull distrirbution, Austrian Journal of Statistics **43**(2), 133-149 (2014).

[19] M. H. Tahir, G. M. Cordeiro, A. Alzaatreh, M. Mansoor and M. Zubair, A new Weibull-Pareto distribution: properties and applications, Communications in Statistics - Simulation and Computation **45**(10), 3548-3567 (2016).

[20] L. M. Zea, R. B. Silva, M. Bourguignon, A. M. Santos and G. M. Cordeiro, The beta exponentiated Pareto distribution with application to bladder cancer susceptibility, International Journal of Statistics and Probability **1**, 8-19 (2012).

Sher Chhetri received his Master's degree from Tribhuvan University, Kathmandu, Nepal. He is now pursuing Ph.D. in the Department of Mathematical Sciences at Florida Atlantic University under the supervision of Dr. Hongwei Long. Sher's research interests include statistical distributions, parametric and non-parametric estimation for stochastic systems, mathematical finance and mathematics/statistics education. He is also co-author of several K-12 textbooks.

Hongwei Long obtained his Ph.D. in Mathematics from University of Warwick, England, in 1998. Now he is a professor in the Department of Mathematical Sciences at Florida Atlantic University. His research interests include parametric and non-parametric estimation for stochastic systems, mathematical finance, filtering theory and its applications, stochastic partial differential equations, and stochastic analysis.

Gokarna Aryal is an associate professor of Statistics at Purdue University Northwest. He obtained his Ph.D. from University of South Florida in 2006. His area of research includes probability distributions and their generalizations, reliability analysis, RFID data analysis, among others. He has published research articles in reputed international journals of Mathematics and Statistics.