

An Extended Pareto Distribution

M. E. Mead

Department of Statistics and Insurance

Faculty of Commerce, Zagazig University, Egypt

Mead9990@gmail.com

Abstract

For the first time, a new continuous distribution, called the generalized beta exponentiated Pareto type I (GBEP) [McDonald exponentiated Pareto] distribution, is defined and investigated. The new distribution contains as special sub-models some well-known and not known distributions, such as the generalized beta Pareto (GBP) [McDonald Pareto], the Kumaraswamy exponentiated Pareto (KEP), Kumaraswamy Pareto (KP), beta exponentiated Pareto (BEP), beta Pareto (BP), exponentiated Pareto (EP) and Pareto, among several others. Various structural properties of the new distribution are derived, including explicit expressions for the moments, moment generating function, incomplete moments, quantile function, mean deviations and Rényi entropy. Lorenz, Bonferroni and Zenga curves are derived. The method of maximum likelihood is proposed for estimating the model parameters. We obtain the observed information matrix. The usefulness of the new model is illustrated by means of two real data sets. We hope that this generalization may attract wider applications in reliability, biology and lifetime data analysis.

Keywords: Beta-Generated class; Pareto type I distribution; Lorenz, Bonferroni and Zenga curves; Rényi entropy; Maximum likelihood estimation.

1. Introduction

The Pareto distribution named after the Italian economist Vilfredo Pareto (1848-1923) is a power law probability distribution that coincides with social, scientific, geophysical, actuarial, and many other types of observable phenomena. Outside the field of economics it is at times referred to as the Bradford distribution. Burroughs and Tebbens (2001) discussed applications of the Pareto distribution in modeling earthquakes, forest fire areas and oil and gas field sizes and Schroeder et al. (2010) presented an application of the Pareto distribution in modeling disk drive sector errors. To add flexibility to the Pareto distribution, various generalizations of the distribution have been derived, the beta Pareto distribution discussed by Akinsete et al. (2008), the Kumaraswamy Pareto distribution introduced by Bourguignon et al. (2013), the beta generalized Pareto defined by Nassar and Nada (2011) and Mahmoudi (2011), the beta exponentiated Pareto distribution presented by Zea et al. (2012), the gamma Pareto distribution introduced by Alzaatreh et al. (2012) and recently, Elbatal (2013) studied the Kumaraswamy exponentiated Pareto distribution.

The cdf of the exponentiated Pareto type I distribution with parameters λ , k and d is given by

$$G(x; d, k, \lambda) = \left[1 - (d/x)^k\right]^\lambda, \quad (1)$$

where $\lambda > 0$, $d > 0$, $k > 0$ and $x \geq d$. The corresponding pdf is given by

$$g(x; d, k, \lambda) = \lambda k d^k x^{-(k+1)} \left[1 - (d/x)^k\right]^{\lambda-1}. \quad (2)$$

Eugene et al. (2002) used the beta distribution as a generator to develop the so-called family of beta-generated distributions (BG). The cdf of a beta-generated random variable X is defined as

$$F(x; a, b) = I_{G(x)}(a, b) = \frac{1}{B(a, b)} \int_0^{G(x)} w^{a-1} (1-w)^{b-1} dw, \quad (3)$$

for $a > 0, b > 0$, where $I_y(a, b) = B_y(a, b) / B(a, b)$ denotes the incomplete beta function ratio of type I and $B_y(a, b) = \int_0^y w^{a-1} (1-w)^{b-1} dw$ is the incomplete beta function. The pdf corresponding to (3) can be expressed as

$$f(x; a, b) = \frac{1}{B(a, b)} g(x) G(x)^{a-1} [1 - G(x)]^{b-1}, \quad (4)$$

where $g(x) = \partial G(x) / \partial x$ is the baseline density function.

Eugene et al. (2002) has used the cdf of normal distribution in (4) to construct the beta normal distribution. The generalization given in (4) has been used by number of authors to propose new distributions. This family of distributions is a generalization of the distributions of order statistics for the random variable X with cdf $F(x)$ as pointed out by Eugene et al. (2002) and Jones (2004). Since the paper by Eugene et al. (2002), many beta-generated distributions have been studied in the literature including, beta gamma distribution by Kong et al. (2007), beta Weibull distribution by Famoye et al. (2005), beta exponential distribution by Nadarajah and Kotz (2006) and others.

Cordeiro and de Castro (2011) extended the beta-generated family of distributions by replacing the beta distribution in (3) with the Kumaraswamy distribution (1980), $g(x) = abx^{a-1}(1-x^a)^{b-1}, 0 \leq x \leq 1$. The cdf of the Kumaraswamy generalized distributions (KG) is given by

$$F(x; b, c) = 1 - [1 - G(x)^c]^b \quad (5)$$

and the corresponding pdf is defined as

$$f(x; b, c) = cb g(x) G(x)^{c-1} [1 - G(x)^c]^{b-1}. \quad (6)$$

Several generalized distributions from (6) have been defined and investigated in the literature including the Kumaraswamy Weibull distribution by Cordeiro et al. (2010), the Kumaraswamy generalized gamma distribution by de Castro et al. (2011) and the Kumaraswamy generalized half-normal distribution by Cordeiro et al. (2012).

Recently, Alexander et al. (2012) introduce the generalized beta-generated (GBG) distribution which has as sub-models the classical beta-generated (BG), Kumaraswamy-

generated (KG) and exponentiated distributions. They defined the cdf of the GBG as the form

$$F(x; a, b, c) = \frac{1}{B(a, b)} \int_0^{G(x)^c} w^{a-1} (1-w)^{b-1} dw \tag{7}$$

$$= I_{G(x)^c}(a, b)$$

and the corresponding pdf is given by

$$f(x; a, b, c) = \frac{c}{B(a, b)} g(x) G(x)^{a-1} [1 - G(x)^c]^{b-1}. \tag{8}$$

The importance of the density(8) is that it contains as special sub-models, the BG ($c = 1$), the KG ($a = 1$) and the exponentiated ($b = c = 1$) distributions .The generalization given in (8) has been used by number of authors including Marciano et al. (2012), Corderio et al. (2014), Oluyede and Rajasooriya (2013) and Tahir et al. (2014).

2. Generalized Beta Exponentiated Pareto Distribution

In this note we propose the generalized beta exponentiated Pareto (GBEP) distribution by using the density function (2) in (8). The pdf of GBEP can be expressed as

$$f(x; \zeta) = \frac{c\lambda k d^k x^{-(k+1)}}{B(a, b)} [1 - (d/x)^k]^{\lambda a c - 1} [1 - (1 - (d/x)^k)^{\lambda c}]^{b-1}, \tag{9}$$

where $\zeta = (a, b, \lambda, d, k, c)$ is the vector of the model parameters. Plots of the GBEP for selected parameter values ($c = 1.4, \lambda = 2.3$ & $k = 1.2$ and different values of a, b & d) are given in Figure 1.

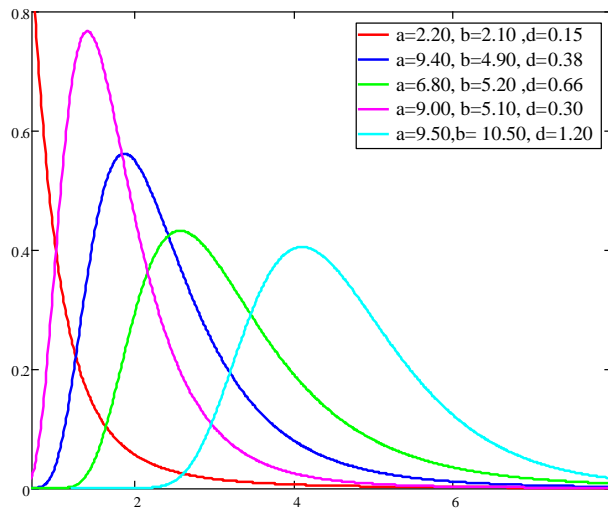


Figure 1: Some possible shapes of the GBEP density function

The cdf corresponding to (9) is

$$F(x; \zeta) = I_{[1 - (d/x)^k]^{\lambda c}}(a, b). \tag{10}$$

Equation (10) can be expressed as follows

$$F(x; \zeta) = \frac{[1 - (d/x)^k]^{\lambda ac}}{a B(a, b)} \left[{}_2F_1\left(a, 1-b; a+1; [1 - (d/x)^k]^{\lambda c}\right) \right], \tag{11}$$

where

$${}_2F_1(a, b; c; z) = \frac{1}{B(b, c-b)} \int_0^1 \frac{t^{b-1} (1-t)^{c-b-1}}{(1-tz)^a} dt,$$

is the well known hypergeometric function (Gradshteyn and Ryzhik, 2007).

For a lifetime random variable t , the survival function $S(t)$, hazard rate function $h(t)$, reversed hazard rate function $r(t)$ and the cumulative hazard rate function $H(t)$ of GBEP distribution are given by

$$S(t) = 1 - F(t) = 1 - I_{[1 - (d/t)^k]^{c\lambda}}(a, b),$$

$$h(t) = \frac{f(t)}{S(t)} = \frac{c\lambda k d^k t^{-(k+1)} [1 - (d/t)^k]^{\lambda ac - 1} [1 - (1 - (d/t)^k)^{\lambda c}]^{b-1}}{[B(a, b) - B_{[1 - (d/t)^k]^{c\lambda}}(a, b)]},$$

$$r(t) = \frac{f(t)}{F(t)} = \frac{c\lambda k d^k t^{-(k+1)} [1 - (d/t)^k]^{\lambda ac - 1} [1 - (1 - (d/t)^k)^{\lambda c}]^{b-1}}{B_{[1 - (d/t)^k]^{c\lambda}}(a, b)}$$

and

$$H(t) = -\ln S(t) = -\ln \left[1 - I_{[1 - (d/t)^k]^{c\lambda}}(a, b) \right].$$

Plots of the hazard rate function (HRF) for selected parameter values ($c = 1.4$, $\lambda = 2.3$ & $k = 1.2$ and different values of a, b & d) are given in Figure 2.

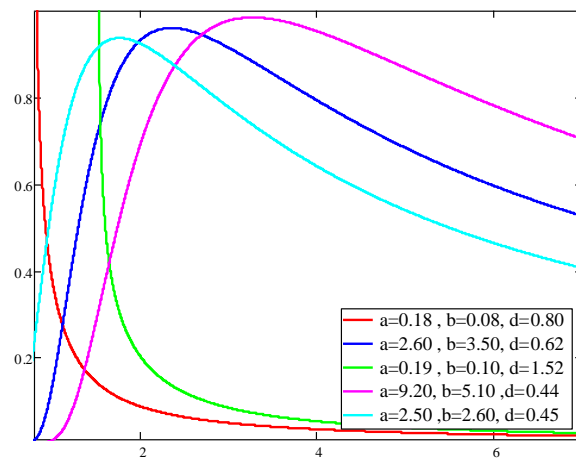


Figure 2: Some possible shapes of the HRF

Note that the GBEP distribution has several well known models as special cases, which make it of distinguishable scientific importance from other distributions.

1. The study of the new density (9) is important, since it includes as special sub-models some distributions not previously considered in the literature. Setting $\lambda = 1$, the density (9) gives the generalized beta Pareto (GBP) [also known as the McDonald Pareto type I] distribution.
2. If $a = 1$, equation (9) reduces to Kumaraswamy exponentiated Pareto (KEP) distribution defined by Elbatal (2013).
3. Setting $c = 1$, the density (9) gives the beta exponentiated Pareto (BEP) distribution presented by Zea et al. (2012) and the beta generalized Pareto defined by Nassar and Nada (2011).
4. If $a = \lambda = 1$, equation (9) corresponds to the Kumaraswamy Pareto (KP) distribution introduced by Bourguignon et al. (2013).
5. Setting $c = \lambda = 1$, GBEP becomes the beta Pareto (BP) distribution discussed by Akinsete et al. (2008).
6. When $a = b = c = 1$, the density (9) corresponds to the exponentiated Pareto (EP) [also known as Stoppa or the generalized Pareto type I] distribution defined by Gupta et al. (1998).
7. If we take $a = b = c = \lambda = 1$, equation (9) becomes the Pareto (P) distribution.

3. Some Statistical Properties

We give a mathematical treatment of the new distribution including expansions of the GBEP cdf and pdf, moments, incomplete moments, generating and quantile functions, mean deviations, mean residual life, Lorenz, Bonferroni and Zenga curves and Rényi entropy.

3.1 Expansion of Distribution

We now present a series expansion of the GBEP cdf and pdf. For any positive real number b , and for $|z| < 1$, a generalized binomial expansion holds

$$(1 - z)^{b-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{j! \Gamma(b-j)} z^j.$$

Therefore, the cdf of GBEP can be expanded to obtain

$$\begin{aligned}
 F(x; \zeta) &= \frac{1}{B(a, b)} \int_0^{\left[1 - (d/x)^k\right]^{c\lambda}} w^{a+j-1} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{j! \Gamma(b-j)} dw \\
 &= \sum_{j=0}^{\infty} p_j G(x; d, k, \lambda c(a+j)),
 \end{aligned} \tag{12}$$

where

$$p_j = \frac{(-1)^j \Gamma(a+b)}{\Gamma(a) j! \Gamma(b-j) (a+j)}$$

and $G(x; d, k, \lambda c(a + j))$ denotes the cdf of EP with parameters d , k and $\lambda c(a + j)$. Similarly, we can write the pdf (9) as

$$\begin{aligned}
 f(x; \zeta) &= \sum_{j=0}^{\infty} \frac{c\lambda k d^k \Gamma(a+b)(-1)^j}{\Gamma(a)j!\Gamma(b-j)x^{(k+1)}} \left[1 - (d/x)^k\right]^{\lambda c(a+j)-1} \\
 &= \sum_{j=0}^{\infty} p_j H(x; d, k, \lambda c(a + j)),
 \end{aligned}
 \tag{13}$$

where $H(x; d, k, \lambda c(a + j))$ denotes the EP density function with parameters d , k and $\lambda c(a + j)$.

Again, by using binomial expansion in equation (13), we obtain

$$\begin{aligned}
 f(x; \zeta) &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{c\lambda \Gamma(a+b)(-1)^{j+i} \Gamma(\lambda c(a + j))}{\Gamma(a)j!i!\Gamma(b-j)\Gamma(\lambda c(a + j) - i)} k d^{k(i+1)} x^{-k(i+1)-1} \\
 &= \sum_{i=0}^{\infty} e_i h(x; d, k(i+1)),
 \end{aligned}
 \tag{14}$$

where

$$e_i = \sum_{j=0}^{\infty} \frac{c\lambda(-1)^{j+i} \Gamma(\lambda c(a + j))\Gamma(a + b)}{\Gamma(a)(i+1)!j!\Gamma(b-j)\Gamma(\lambda c(a + j) - i)}$$

and $h(x; d, k(i+1))$ denotes the Pareto density with parameters d and $k(i+1)$. Thus, the GBEP density function can be expressed as an infinite linear combination of Pareto densities. Thus, some of its mathematical properties can be obtained directly from those properties of the Pareto distribution. If b is an integer, then the summation in equations (12), (13) and (14) stops at $b-1$.

3.2 Moments

As with any other distribution, many of the interesting characteristics and features of the GBEP distribution can be studied through the moments. If we assume that Y is a Pareto distributed random variable, then the s th moment of Y is given as

$$E(Y^s) = \frac{kd^s}{k-s}, \quad s < k.$$

Let X be a random variable having the GBEP distribution (9). Using equation (14), it is easy to obtain the s th moment of X as the following form

$$E(X^s) = \sum_{i=0}^{\infty} \frac{k(i+1)d^s e_i}{k(i+1)-s}.
 \tag{15}$$

The mean, variance, skewness and kurtosis can be obtained from (15). If $b > 0$ is integer and $s < k$, the sum stops at $b-1$.

An alternative form for the s th ordinary moment of X from equation (13) as

$$\begin{aligned}
 E(X^s) &= \int_d^\infty x^s f(x; \zeta) dx \\
 &= \sum_{j=0}^\infty p_j \lambda c(a+j) k d^k \int_d^\infty x^{s-k-1} \left[1 - (d/x)^k\right]^{\lambda c(a+j)-1} dx \\
 &= \sum_{j=0}^\infty p_j c \lambda(a+j) d^s B(1-s/k, \lambda c(a+j)).
 \end{aligned} \tag{16}$$

If $b > 0$ is integer and $s < k$, the sum stops at $b-1$.

3.3 Moment Generating Function

The moment generating function (mgf) $\mu_Y(t)$ corresponding to a random variable Y with Pareto distribution with parameters d and k is only defined for negative values of t (See Zea et al., 2012). It is given by

$$\mu_Y(t) = \int_d^\infty e^{ty} k d^k y^{-(k+1)} dy = k(-dt)^k \Gamma(-k, -dt),$$

where $\Gamma(\alpha, z) = \int_z^\infty t^{\alpha-1} e^{-t} dt$ denotes the incomplete gamma function. From equation (14), the mgf of X is obtained by

$$\mu_X(t) = \sum_{i=0}^\infty k(i+1)(-dt)^{k(i+1)} \Gamma(-k(i+1), -dt) e_i, \quad \text{for } t < 0.$$

3.4 Incomplete Moments

If Y is a random variable with Pareto distribution with parameters d and k , the s th incomplete moment of Y , for $s < k$, is given by

$$M_s(z) = \int_d^z y^s g(y; d, k) dy = \int_d^z y^s k d^k y^{-(k+1)} dy = \frac{kd^s}{k-s} \left[1 - \left(\frac{d}{z}\right)^{k-s}\right].$$

From this equation, we note that $M_s(z) \rightarrow E(Y^s)$ when $z \rightarrow \infty$, whenever $s < k$. Let X be a random variable having the GBEP distribution (9). Using equation (14), the s th incomplete moment of X is then equal to

$$M_s(z) = \sum_{i=0}^\infty e_i \int_d^z x^s h(x; d, k(i+1)) dy = \sum_{i=0}^\infty \frac{k(i+1)d^s e_i}{k(i+1)-s} \left[1 - \left(\frac{d}{z}\right)^{k(i+1)-s}\right], \quad s < k. \tag{17}$$

An alternative expression for the s th incomplete moment of X can be obtained from equation (13) as

$$\begin{aligned}
 E(X^s) &= \int_d^z x^s f(x; \zeta) dx \\
 &= \sum_{j=0}^\infty p_j \lambda c(j+a) k d^k \int_d^z x^{s-k-1} \left[1 - (d/x)^k\right]^{\lambda c(j+a)-1} dx
 \end{aligned}$$

$$= \sum_{j=0}^{\infty} p_j c \lambda (j+a) d^s B^{(d/z)^k} (1-s/k, \lambda c (j+a)), \quad s < k. \quad (18)$$

where $B^y(a, b) = \int_y^1 w^{a-1} (1-w)^{b-1} dw$, is the upper incomplete beta type I function.

3.5 Mean Deviations

The mean deviations about the mean and the median can be used as measures of spread in a population. Let $\mu = E(X)$ and γ be the mean and the median of the GBEP distribution, respectively. The mean deviations about the mean and about the median of X can be calculated as

$$D(\mu) = E|X - \mu| = \int_d^{\infty} |x - \mu| f(x) dx = 2\mu F(\mu) - 2m_1(\mu)$$

and

$$D(\gamma) = E|X - \gamma| = \int_d^{\infty} |x - \gamma| f(x) dx = \mu - 2m_1(\gamma),$$

respectively, where $m_1(\mu)$ denotes the first incomplete moment and $F(\mu)$ follows from (10).

3.6 Quantile Function

Let $Q_{a,b}(u)$ be the beta quantile function with parameters a and b . The quantile function of the GBEP distribution, say $x = Q(u)$, can be easily obtained as

$$x = Q(u) = d \left[1 - (Q_{a,b}(u))^{1/(c\lambda)} \right]^{-1/k}, \quad u \in (0,1). \quad (19)$$

This scheme is useful to generate GBEP random variates because of the existence of fast generators for beta random variables in most statistical packages, i.e. if V is a beta random variable with parameters a and b , then $X = d \left[1 - V^{1/(c\lambda)} \right]^{-1/k}$ follows the GBEP distribution. From equation (19) we conclude that the median m of X is $m = Q(1/2)$. The Bowley skewness SK measure and Moors kurtosis KR (based on octiles) of the GBEP distribution can be calculated using the formulae given below

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

and

$$KR = \frac{[Q(7/8) - Q(5/8)] + [Q(3/8) - Q(1/8)]}{[Q(6/8) - Q(2/8)]}.$$

3.7 Mean Residual Life and Mean Waiting Time

The mean residual life function (MRL) at a given time t measures the expected remaining lifetime of an individual of age t . It is denoted by $m(t)$. The MRL or life expectancy of GBEP is defined as

$$m(t) = \frac{1}{S(t)} \left[E(t) - \int_d^t t f(t) dt \right] - t,$$

$$= \frac{c\lambda d \sum_{j=0}^{\infty} p_j(j+a) \left[B(1-1/k, \lambda c(j+a)) - B^{(d/t)^k}(1-1/k, \lambda c(j+a)) \right]}{1 - I_{[1-(d/t)^k]^{c\lambda}}(a,b)} - t, \quad k > 1,$$

where

$$\int_d^t t f(t) dt = \sum_{j=0}^{\infty} p_j c\lambda(j+a) dB^{(d/t)^k}(1-1/k, \lambda c(j+a)).$$

The mean waiting time (MWT) of an item failed in a interval $[d, t]$ for GBEP is defined as

$$\bar{\mu}(t, \theta) = t - \frac{1}{F(t)} \int_d^t t f(t) dt$$

$$= t - \frac{\sum_{j=0}^{\infty} p_j c\lambda(j+a) dB^{(d/t)^k}(1-1/k, \lambda c(j+a))}{I_{[1-(d/t)^k]^{c\lambda}}(a,b)}, \quad k > 1.$$

3.8 Lorenz, Bonferroni and Zenga Curves

Lorenz and Bonferroni curves have been applied in many fields such as economics, reliability, demography, insurance and medicine, (See Kleiber and Kotz, (2003) for additional details). Zenga curve was presented by Zenga (2007). The Lorenz $L_F(x)$, Bonferroni $B(F(x))$ and Zenga $A(x)$ curves are defined by Oluyede and Rajasooriya (2013) as the following

$$L_F(x) = \frac{\int_d^x t f(t) dt}{E(X)}, \quad B(F(x)) = \frac{\int_d^x t f(t) dt}{F(x)E(X)} = \frac{L_F(x)}{F(x)} \quad \text{and} \quad A(x) = 1 - \frac{\bar{M}(x)}{M(x)},$$

respectively, where

$$\bar{M}(x) = \frac{\int_d^x t f(t) dt}{F(x)} \quad \text{and} \quad M(x) = \frac{\int_x^{\infty} t f(t) dt}{1 - F(x)}$$

are the lower and upper means respectively. For the GBEP distribution, these quantities are derived below

1. Lorenz curve:

$$L_{FG}(x; \zeta) = \frac{\sum_{j=0}^{\infty} p_j c \lambda(j+a) dB^{(d/x)^k} (1-1/k, \lambda c(j+a))}{\sum_{j=0}^{\infty} p_j c \lambda(j+a) dB(1-1/k, \lambda c(j+a))}.$$

2. Bonferroni curve:

$$B(F_G(x; \zeta)) = \frac{\sum_{j=0}^{\infty} p_j c \lambda(j+a) dB^{(d/x)^k} (1-1/k, \lambda c(j+a))}{\sum_{j=0}^{\infty} p_j G(x; d, k, \lambda c(a+j)) \sum_{j=0}^{\infty} p_j c \lambda(j+a) dB(1-1/k, \lambda c(j+a))}.$$

3. Zenga curve

$$A(x; \zeta) = 1 - \frac{[1-F(x)] \int_d^x t f(t) dt}{F(x) \int_x^{\infty} t f(t) dt},$$

$$= 1 - \frac{\left[1 - \sum_{j=0}^{\infty} p_j G(x; d, k, \lambda c(a+j)) \right] \left[\sum_{j=0}^{\infty} p_j c \lambda(j+a) dB^{(d/x)^k} (1-1/k, \lambda c(j+a)) \right]}{\left[\sum_{j=0}^{\infty} p_j G(x; d, k, \lambda c(a+j)) \right] \left[c \lambda \sum_{j=0}^{\infty} p_j (j+a) [B(1-1/k, \lambda c(j+a)) - B^{(d/x)^k} (1-1/k, \lambda c(j+a))] \right]}.$$

3.9 Rényi Entropy

The entropy of a random variable X is a measure of uncertainty variation. The Rényi entropy is defined as

$$I_R(\delta) = \frac{1}{1-\delta} [\log I(\delta)],$$

where $I(\delta) = \int f^\delta(x) dx$ $\delta > 0$ and $\delta \neq 1$. Using equation (9) we obtain

$$I(\delta) = \frac{c^\delta \lambda^\delta k^\delta d^{k\delta}}{B^\delta(a, b)} \int_d^\infty x^{-\delta(k+1)} \left[1 - (d/x)^k \right]^{\delta(\lambda a c - 1)} \left[1 - \left(1 - (d/x)^k \right)^{\lambda c} \right]^{\delta(b-1)} dx.$$

Based on the binomial expansion to the last factor in the above integrand yields

$$I(\delta) = \frac{c^\delta \lambda^\delta k^\delta d^{k\delta}}{B^\delta(a, b)} \sum_{j=1}^{\infty} \binom{\delta(b-1)}{j} (-1)^j \int_d^\infty x^{-\delta(k+1)} \left[1 - (d/x)^k \right]^{\delta(\lambda a c - 1) + \lambda c j} dx.$$

Using the transformation $y = (\beta/x)^k$ in above expression and simplifying,

$$I(\delta) = \frac{c^\delta \lambda^\delta k^{\delta-1} d^{1-\delta}}{B^\delta(a, b)} \sum_{j=1}^{\infty} \binom{\delta(b-1)}{j} (-1)^j B\left(\frac{(\delta-1)(k+1)}{k} + 1, \delta(\lambda a c - 1) + \lambda c j + 1\right).$$

Hence, the Rényi entropy reduces to

$$I_R(\delta) = \left(\frac{1}{\delta-1} \right) \left[\delta \log \left(\frac{c\lambda}{B(a,b)} \right) + \log \sum_{j=1}^{\infty} \binom{\delta(b-1)}{j} (-1)^j B \left(\frac{(\delta-1)(k+1)}{k} + 1, \delta(\lambda a c - 1) + \lambda c j + 1 \right) \right] + \log \left(\frac{d}{k} \right).$$

4. Estimation of Parameters

The maximum likelihood estimation (MLE) is one of the most widely used estimation method for finding the unknown parameters. Let x_1, x_2, \dots, x_n be an independent random sample from GBEP. The total log-likelihood is given by

$$\ell = n \ln(c) + n \ln(\lambda) + n \ln(k) + n k \ln(d) - n \ln B(a,b) - (k+1) \sum_{i=1}^n \ln(x_i) + (\lambda a c - 1) \sum_{i=1}^n \ln(D_i) + (b-1) \sum_{i=1}^n \ln[1 - (D_i)^{\lambda c}], \quad (20)$$

where $Z_i = (d/x_i)$, $W_i = (Z_i)^k$ and $D_i = 1 - W_i$.

The score vector $\nabla \ell = (\frac{\partial \ell}{\partial a}, \frac{\partial \ell}{\partial b}, \frac{\partial \ell}{\partial c}, \frac{\partial \ell}{\partial \lambda}, \frac{\partial \ell}{\partial k})$ has components

$$\begin{aligned} \frac{\partial \ell}{\partial a} &= n(\psi(a+b) - \psi(a)) + \lambda c \sum_{i=1}^n \ln(D_i), \\ \frac{\partial \ell}{\partial b} &= n(\psi(a+b) - \psi(b)) + \sum_{i=1}^n \ln[1 - (D_i)^{\lambda c}], \\ \frac{\partial \ell}{\partial c} &= \frac{n}{c} + \lambda a \sum_{i=1}^n \ln(D_i) - \lambda(b-1) \sum_{i=1}^n [1 - (D_i)^{\lambda c}]^{-1} (D_i)^{\lambda c} \ln(D_i), \\ \frac{\partial \ell}{\partial \lambda} &= \frac{n}{\lambda} + a c \sum_{i=1}^n \ln(D_i) - c(b-1) \sum_{i=1}^n [1 - (D_i)^{\lambda c}]^{-1} (D_i)^{\lambda c} \ln(D_i) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \ell}{\partial k} &= \frac{n}{k} + n \ln(d) + \lambda c(b-1) \sum_{i=1}^n [1 - (D_i)^{\lambda c}]^{-1} (D_i)^{\lambda c - 1} (W_i) \ln(Z_i) \\ &\quad - \sum_{i=1}^n \ln(x_i) - (\lambda a c - 1) \sum_{i=1}^n (D_i)^{-1} (W_i) \ln(Z_i), \end{aligned}$$

where $\psi(p)$ is the digamma function which is the derivative of $\log \Gamma(\cdot)$. The maximum likelihood estimates (MLEs) of the unknown five parameters can be obtained by solving the system of nonlinear equations $\nabla \ell = 0$, iteratively. Since $x \geq d$, the MLE of d is the first-order statistic $x_{(1)}$. For interval estimation and hypothesis tests on the model parameters, we require the observed information matrix

$$J_n(\theta) = n \begin{bmatrix} J_{aa} & J_{ab} & J_{ac} & J_{a\lambda} & J_{ak} \\ J_{ab} & J_{bb} & J_{bc} & J_{b\lambda} & J_{bk} \\ J_{ac} & J_{bc} & J_{cc} & J_{c\lambda} & J_{ck} \\ J_{a\lambda} & J_{b\lambda} & J_{c\lambda} & J_{\lambda\lambda} & J_{\lambda k} \\ J_{ak} & J_{bk} & J_{ck} & J_{\lambda k} & J_{kk} \end{bmatrix},$$

whose entries are obtained from standard calculations:

$$J_{aa} = n(\psi'(a+b) - \psi'(a)),$$

$$J_{ab} = n(\psi'(a+b)),$$

$$J_{ac} = \lambda \sum_{i=1}^n \ell n(D_i),$$

$$J_{a\lambda} = c \sum_{i=1}^n \ell n(D_i),$$

$$J_{ak} = -\lambda ck \sum_{i=1}^n (D_i)^{-1} (W_i) \ell n(Z_i),$$

$$J_{bb} = n(\psi'(a+b) - \psi'(b)),$$

$$J_{bc} = -\lambda \sum_{i=1}^n [1 - (D_i)^{\lambda c}]^{-1} (D_i)^{\lambda c} \ell n(D_i),$$

$$J_{b\lambda} = -c \sum_{i=1}^n [1 - (D_i)^{\lambda c}]^{-1} (D_i)^{\lambda c} \ell n(D_i),$$

$$J_{bk} = \lambda c \sum_{i=1}^n [1 - (D_i)^{\lambda c}]^{-1} (D_i)^{\lambda c - 1} (W_i) \ell n(Z_i),$$

$$J_{cc} = -\frac{n}{c^2} - \lambda^2 (b-1) \sum_{i=1}^n \left[(1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c} \ell n(D_i) \right]^2 \left[1 + (1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c} \right],$$

$$J_{c\lambda} = a \sum_{i=1}^n \ell n(D_i) - (b-1) \sum_{i=1}^n \left[(1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c} \ell n(D_i) \right] \left[1 + \lambda c \ell n(D_i) \left[1 + (1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c} \right] \right],$$

$$J_{ck} = -\lambda a \sum_{i=1}^n \left[(D_i)^{-1} (W_i) \ell n(Z_i) \right] + \lambda (b-1) \sum_{i=1}^n \left[(1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c - 1} (W_i) \ell n(Z_i) \right] \left[1 + \lambda c \ell n(D_i) \left[1 + (1 - (D_i)^{\lambda c})^{-1} \right] \right],$$

$$J_{\lambda\lambda} = -\frac{n}{\lambda^2} - c^2 (b-1) \sum_{i=1}^n \left[(1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c} \ell n(D_i) \right]^2 \left[1 + (D_i)^{\lambda c} (1 - (D_i)^{\lambda c})^{-1} \right],$$

$$J_{\lambda k} = -ac \sum_{i=1}^n (D_i)^{-1} (W_i) \ell n(Z_i) - c(b-1) \sum_{i=1}^n \left[(1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c - 1} \ell n(D_i) (W_i) \ell n(Z_i) \right] \left[-1 + \lambda c \ell n(D_i) \left[1 - (D_i)^{\lambda c} (1 - (D_i)^{\lambda c})^{-1} \right] \right]$$

and

$$J_{kk} = -\frac{n}{k^2} - (\lambda ac - 1) \sum_{i=1}^n (D_i)^{-1} (W_i) \ell n(Z_i)^2 \left[(W_i) (D_i)^{-1} - 1 \right] + \lambda c (b-1) \sum_{i=1}^n \left[(1 - (D_i)^{\lambda c})^{-1} (D_i)^{\lambda c - 1} (W_i) \ell n(Z_i)^2 \right] \left[1 - (\lambda c - 1) (W_i) (D_i)^{-1} - \lambda c (1 - (D_i)^{\lambda c})^{-1} (W_i) (D_i)^{\lambda c - 1} \right].$$

5. Empirical Illustrations

In this section, we present two applications of the proposed GBEP distribution (and their sub-models: GBP, BEP, KEP, BP, KP, EP and P distributions) in two real data sets to illustrate its potentiality. The first data correspond to 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. These data were analysed by Choulakian and Stephens (2001). Recently, Akinsete et al. (2008) and Bourguignon et al. (2013) studied these data using the BP and KP respectively. The data are as follows: 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. The second real data set represents the actual taxes data set. The revenue in Egypt is divided onto 5 chapters and although the taxes is only one chapter from these 5 chapters, but it records the majority of the income. The data consists of the monthly actual taxes revenue in Egypt from January 2006 to November 2010. The distribution is highly skewed to the right. The data (in 1000 million Egyptian pounds) are: 5.9, 20.4, 14.9, 16.2, 17.2, 7.8, 6.1, 9.2, 10.2, 9.6, 13.3, 8.5, 21.6, 18.5, 5.1, 6.7, 17, 8.6, 9.7, 39.2, 35.7, 15.7, 9.7, 10, 4.1, 36, 8.5, 8, 9.2, 26.2, 21.9, 16.7, 21.3, 35.4, 14.3, 8.5, 10.6, 19.1, 20.5, 7.1, 7.7, 18.1, 16.5, 11.9, 7, 8.6, 12.5, 10.3, 11.2, 6.1, 8.4, 11, 11.6, 11.9, 5.2, 6.8, 8.9, 7.1, 10.8. These data were studied by Nassar and Nada (2011) using beta generalized Pareto distribution.

Table 1 lists the MLEs (the corresponding standard errors in parentheses) of the parameters of all the models for the first data set (the exceedances of flood peaks data) and the statistics: Akaike information criterion (AIC), Bayesian information criterion (BIC) and consistent Akaike information criterion (CAIC). Table 2 gives the values of the statistics Kolmogorov-Smirnov (K-S) and $-2\ell(\hat{\theta})$ (where $\ell(\hat{\theta})$ denotes the log-likelihood function evaluated at the maximum likelihood estimates) for the first data set. The results of the last four models (BP, KP, EP and P) can be obtained from Bourguignon et al. (2013). Since the KEP distribution has the lowest AIC, BIC, CAIC, $-2\ell(\hat{\theta})$ and K-S values among all the other models, and so it could be chosen as the best model. Additionally, it is evident that the P distribution presents the worst fit to the first data.

Tables 3 and 4 provide the MLEs (the corresponding standard errors in parentheses) of the parameters of all the models and the statistics AIC, BIC, CAIC, $-2\ell(\hat{\theta})$ and K-S for the second data set (actual taxes revenue). Again, the results indicate that the KP model presents the smallest values for the AIC, BIC, CAIC, $-2\ell(\hat{\theta})$ and K-S statistics among the fitted models and therefore it could be chosen as the best model. The required numerical evaluations are implemented using the MATH- CAD PROGRAM.

Table 1: MLEs (standard errors in parentheses) and the statistics AIC, BIC and CAIC; first data set. ($\hat{d} = 0.1$ for all models, Since \hat{d} is the first order statistic).

Model	Estimates					Statistics		
	\hat{a}	\hat{b}	\hat{c}	$\hat{\lambda}$	\hat{k}	AIC	BIC	CAIC
GBEP	67.65386 (33.9215)	4.01038 (2.598)	0.34919 (3.0681)	0.25325 (2.225)	0.20803 (0.1099)	554.5	565.9	555.4
GBP	35.8031 (18.0603)	4.07272 (2.655)	0.16468 (0.3836)	- -	0.20572 (0.1096)	552.5	561.60	553.0
BEP	56.92285 (21.051)	4.02054 (2.604)	- -	0.10476 (0.38005)	0.20765 (0.10971)	552.5	561.6	553.1
KEP	- -	136.0781 (30.198)	1.97714 (13.987)	1.75478 (3.3557)	0.06153 (0.0252)	530.2	539.3	530.8
BP	3.1473 (0.4993)	85.7508 (0.0001)	- -	- -	0.0088 (0.0015)	573.4	580.3	573.8
KP	- -	85.8468 (0.3371)	2.8553 (0.3371)	- -	0.0528 (0.0185)	548.4	555.3	548.8
EP	- -	- -	- -	2.8797 (0.4911)	0.4241 (0.0463)	578.6	583.2	578.8
P	- -	- -	- -	- -	0.2438 (0.0287)	608.2	610.4	608.2

Table 2: K-S and $-2\ell(\hat{\theta})$ statistics; first data set.

Model	GBEP	GBP	BEP	KEP	BP	KP	EP	P
K-S	0.15665	0.16028	0.15663	0.14314	0.1747	0.1700	0.1987	0.3324
$-2\ell(\hat{\theta})$	544.535	544.494	544.529	522.211	567.4	542.4	574.6	606.2

Table 3: MLEs (standard errors in parentheses) and the statistics AIC, BIC and CAIC; second data set. ($\hat{d} = 4.1$ for all models, Since \hat{d} is the first order statistic).

Model	Estimates					Statistics		
	\hat{a}	\hat{b}	\hat{c}	$\hat{\lambda}$	\hat{k}	AIC	BIC	CAIC
GBEP	50.173 (10.683)	1.61244 (0.782)	0.27553 (2.0699)	0.20069 (1.508)	1.09171 (0.406)	403.04	413.427	404.172
GBP	34.7029 (25.352)	1.61671 (0.986)	0.07979 (0.578)	- -	1.08948 (0.5589)	401.029	409.34	401.77
BEP	27.025 (15.333)	1.622 (1.031)	- -	0.10229 (0.566)	1.08673 (0.58)	401.016	409.326	401.757
KEP	- -	75.33529 (39.223)	1.41497 (20.494)	1.46209 (21.177)	0.11359 (0.105)	390.97	399.28	391.71
BP	2.57098 (0.44668)	17.09134 (6.4901)	- -	- -	0.1373 (0.5001)	396.775	403.008	397.212
KP	- -	77.3103 (22.144)	2.06776 (0.2537)	- -	0.11195 (0.1033)	388.956	395.188	389.392
EP	- -	- -	- -	2.54142 (0.4883)	1.58491 (0.205)	397.571	401.726	397.786
P	- -	- -	- -	- -	0.95346 (0.1241)	415.881	417.958	415.951

Table 4: K-S and $-2\ell(\hat{\theta})$ statistics; second data set

Model	GBEP	GBP	BEP	KEP	BP	KP	EP	P
K-S	0.11442	0.11439	0.11435	0.06736	0.1061	0.06726	0.11983	0.25527
$-2\ell(\hat{\theta})$	393.04	393.029	393.016	382.969	390.775	382.956	393.571	413.881

6. Conclusion

The new five-parameter model (Since the sixth parameter is the first order statistic) includes as special sub-models the Pareto, exponentiated Pareto (Gupta et al., 1998), beta Pareto (Akinsete et al., 2008), Kumaraswamy Pareto (Bourguignon et al., 2013), beta generalized Pareto (Nassar and Nada, 2011), beta exponentiated Pareto (Zea et al., 2012), Kumaraswamy exponentiated Pareto (Elbatal, 2013) and generalized beta Pareto (new) distributions. We provide a mathematical treatment of this distribution including analytical expressions for the moments, moment generating function, mean deviations, mean residual life, Lorenz, Bonferroni and Zenga curves, quantile function and Rényi entropy. The estimation of the model parameters is approached by maximum likelihood and the observed information matrix is derived. The usefulness of the new model is illustrated in two applications to real data using goodness-of-fit tests.

Acknowledgment

The author thanks the Editor and the Referees for their helpful remarks that improved the original manuscript.

References

1. Akinsete, A., Famoye, F. & Lee, C. (2008). The beta-Pareto distribution. *Statistics*, 42, 547-563.
2. Alexander, C., Cordeiro, G. M., Ortega, E. M. M. & Sarabia, J. M. (2012). Generalized beta generated distributions. *Computational Statistics and Data Analysis*, 56, 1880-1897.
3. Alzaatreh, A., Famoye, F. & Lee, C. (2012). Gamma Pareto distribution and its applications. *Journal of Modern Statistical Methods*, 11, 78-95.
4. Bourguignon, M., Silva, R. B., Zea, L. M. & Cordeiro, G. M. (2013). The Kumaraswamy Pareto distribution. *Journal of Statistical Theory and Applications* 12, 129-144.
5. Burroughs, S. M. & Tebbens, S. F. (2001). Upper-truncated power law distributions. *Fractals*, 9, 209-222.
6. Choulakian, V. & Stephens, M.A. (2001). Goodness-of-fit for the generalized Pareto distribution. *Technometrics*, 43, 478-484.
7. Cordeiro, G.M. & de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81, 883-898.

8. Corderio, G.M., Hashimoto, E.H. & Ortega, E.M.M. (2014). The McDonald Weibull model. *Statistics*, 48, 256-278.
9. Cordeiro, G. M., Ortega, E. M. M. & Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, 347, 1399 -1429.
10. Cordeiro, G.M., Pescim, R.R. & Ortega, E.M.M. (2012). The Kumaraswamy generalized half-normal distribution for skewed positive data. *Journal of Data Science*, 10,195-224.
11. de Castro, M.A.R., Ortega, E.M.M. & Cordeiro, G.M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, 8, 411- 433.
12. Elbatal (2013). The Kumaraswamy exponentiated Pareto distribution. *Economic Quality Control*, 28, 1-9.
13. Eugene, N., Lee, C. & Famoye, F. (2002). The beta-normal distribution and its applications. *Communications in Statistics -Theory and Methods*, 31, 497-512.
14. Famoye, F., Lee, C. & Olumolade, O. (2005). The beta-Weibull distribution. *Journal of Statistical Theory and Applications*, 4, 121-136.
15. Gradshteyn, I. S. & Ryzhik, I. M. (2007). *Table of integrals, series and products*. Seventh Edition, Alan Jeffrey and Daniel Zwillinger (eds.), Academic Press.
16. Gupta, R.C., Gupta, R.D. & Gupta, P.L. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics-Theory and Methods* ,27, 887-904.
17. Jones, M.C. (2004). Families of distributions arising from distributions of order statistics. *Test*, 13, 1-43.
18. Kleiber, C. & Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. Wiley Series in Probability and Statistics. John Wiley & Sons.
19. Kong, L., Lee, C. & Sepanski, J.H. (2007). On the properties of beta-gamma distribution. *Journal of Modern Applied Statistical Methods*, 6, 187-211.
20. Kumaraswamy, P. (1980). Generalized probability density function for double bounded random processes. *Journal of Hydrology*, 462, 79-88.
21. Mahmoudi, E. (2011).The beta generalized Pareto distribution with application to lifetime data. *Mathematics and Computers in Simulation*, 81, 2414-2430.
22. Marciano, F.W.P., Nascimento, A.D.C., Santos-Neto, M. & Corderio, G. M. (2012). The Mc- Γ distribution and its statistical properties: an application to reliability data. *International Journal of Statistics and Probability*, 1, 53-71.
23. Nadarajah, S. & Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering & System Safety*, 91, 689-697.
24. Nassar, M. M. & Nada, N. K. (2011).The beta generalized Pareto distribution. *Journal of Statistics: Advances in Theory and Applications*, 6, 1-17.

25. Oluyede, B. O. & Rajasooriya, S. (2013). The Mc-Dagum distribution and its statistical properties with applications. *Asian Journal of Mathematics and Applications*, 44, 1-16.
26. Schroeder, B., Damouras, S. & Gill, P. (2010). Understanding latent sector error and how to protect against them. *ACM Transactions on Storage (TOS)*, 6(3), Article 8.
27. Tahir, M. H., Mansoor, M., Zubair, M. & Hamedani, G. G. (2014). McDonald log-logistic distribution with an application to breast cancer data. *Journal of Statistical Theory and Applications*, 13, 65-82.
28. Zea, L. M., Silva, R.B., Bourguignon, M., Santos, A., M. & Cordeiro, G. M. (2012). The beta exponentiated Pareto distribution with application to bladder cancer susceptibility. *International Journal of Statistics and Probability*, 2, 8 -19.
29. Zenga, M. (2007). Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. *Statistica & Applicazioni*, 1, 3-27.