NEW RESULTS FROM A BETA-PARETO CLASS

S. L. Makgai
Department of Statistics, University of Pretoria, South Africa

A. Bekker
Department of Statistics, University of Pretoria, South Africa

J. T. Ferreira [1]
Department of Statistics, University of Pretoria, South Africa
e-mail: *johan.ferreira@up.ac.za*

M. Arashi
Department of Statistics, University of Pretoria, South Africa, and Shahrood University of Technology, Iran

Key words: Generalised beta type, Generator distribution, Kolmogorov-Smirnov measure, Parent distribution, Pareto.

Abstract: In this paper, we propose new results on a beta-Pareto class. This extension evolved by using newly proposed class of beta distributions as generator distributions ($F(x)$) within the generator-parent approach: $H(x) = F(G(x))$ with $G(x)$ the Pareto distribution as the parent distribution. Fundamental properties are studied, including moments and Rényi entropy. We illustrate the added value to the literature with a numerical example using hydrology data, indicating that this newly proposed class is a valid candidate to consider in modelling within the hydrology arena.

1. Introduction

Univariate distributions make up the vast majority of studied and applied distributions in the statistical universe. The literature is saturated with different methods on deriving and studying new univariate distributions. One of these is the *generator approach*, pioneered by Eugene, Lee and Famoye (2002) and Jones (2004). In this paper, this technique is used by proposing several new *beta type-generated* distributions by coupling these distributions with the Pareto distribution, and the genesis for this particular paper originates from Akinsete, Famoye and Lee (2008) and Cardeño, Nagar and Sánchez (2005). The methodology followed to construct the new class is rephrased in Definition 1.

Definition 1 *A distribution is said to be a generated distribution, if its cumulative distribution function (cdf)* $F(G(x))$ is obtained as follows:

[1] Corresponding author

AMS: 60E05, 62E15

$$F(G(x)) = \int_0^{G(x)} h(w)\, dw, \tag{1}$$

where $h(w)$ is the probability density function (pdf) of the generator distribution of a non-negative continuous random variable W defined on $[0, \infty)$ and $G(x)$ is the cdf of the parent distribution. This generated distribution has corresponding pdf $f(x) = \frac{d}{dx}F(G(x))$.

In this paper, $h(w)$ will assume different beta type distributions. The pdf corresponding to beta type-I distribution as the generator distribution with

$$h(w) = \frac{1}{B(\alpha, \beta)} w^{\alpha-1} (1-w)^{\beta-1}, \quad 0 \leq w \leq 1,$$

is given by

$$f(x) = \frac{1}{B(\alpha,\, \beta)} (G(x))^{\alpha-1} (1-(G(x)))^{\beta-1} g(x). \tag{2}$$

It is pointed out by Jones (2004) that this class of distributions is a generalisation of order statistics for the random variable X with cdf $G(x)$.

The first distribution of the beta generated class was the beta-normal distribution, derived by Eugene et al. (2002). In this case, the normal distribution's cdf was taken as the parent distribution $G(x)$ and the beta type-I distribution was taken as the generator distribution with pdf $h(w)$. Changing the parent distribution so that it is no longer normal offers more flexibility. Further research sees Nadarajah and Kotz (2004) introducing the beta-Gumbel distribution, Nadarajah and Gupta (2004) proposing the beta-Fréchet distribution and Nadarajah and Kotz (2006) defining the beta-exponential distribution. The refreshing contributions made by the authors Balakrishnan and Ristic (2016), Barreto-Souza, Santos and Cordeiro (2010), Bourguignon, Silva, Zea and Cordeiro (2012), Cordeiro, Nobre, Pescim and Ortega (2012), Ristic and Balakrishnan (2012) and Zografos and Balakrishnan (2009) should also be acknowledged. Akinsete et al. (2008) followed the work of Eugene et al. (2002) with the Pareto distribution as the parent distribution.

Our aim is to introduce a class based on the generator approach as described in (1), which contains more flexible distributions for modelling data.

First, we propose a new beta type distribution (named the *generalised beta type*) as the generator $h(w)$. The pdf of this generator, emanating from (5) and the model proposed by Cardeño et al. (2005), is given as

$$h(w) = \frac{(1+K)^{\alpha} d}{B(\alpha,\, \beta)}\, w^{\alpha d-1} \left(1-w^d\right)^{\beta-1} \left(1+Kw^d\right)^{-(\alpha+\beta)}, \quad 0 \leq w \leq 1, \tag{3}$$

for a random variable W, shape parameters $\alpha > 0$, $\beta > 0$, $d > 0$, $K > 0$ and where $B(\cdot,\cdot)$ is the beta function. This random variable W is denoted by $W \sim Genbeta(\alpha,\, \beta,\, d,\, K)$.

If $K = 1$ in (3), then a new *generalised beta type III model* follows with pdf

$$h(w) = \frac{2^{\alpha} d}{B(\alpha,\, \beta)}\, w^{\alpha d-1} \left(1-w^d\right)^{\beta-1} \left(1+w^d\right)^{-(\alpha+\beta)}, \quad 0 \leq w \leq 1,$$

with random variable W denoted by $W \sim GenbetaIII(\alpha,\, \beta,\, d)$.

If $d = 1$ in (3), then an *extended beta type III model* of Cardeño et al. (2005) follows with pdf

$$h(w) = \frac{(1+K)^{\alpha}}{B(\alpha,\beta)} w^{\alpha-1} (1-w)^{\beta-1} (1+Kw)^{-(\alpha+\beta)}, \qquad 0 \leq w \leq 1, \tag{4}$$

with random variable W denoted by $W \sim Extended\text{-}betaIII(\alpha,\beta,K)$.

If $K = 0$ in (3), then the *generalised beta type-I distribution* follows from McDonald (1984) with pdf:

$$h(w) = \frac{d}{B(\alpha,\beta)} w^{\alpha d-1} \left(1-w^{d}\right)^{\beta-1}, \quad 0 \leq w \leq 1, \tag{5}$$

for random variable W and shape parameters $\alpha > 0$, $\beta > 0$ and $d > 0$. Two well known distributions, namely the beta type I and Kumaraswamy distributions (Kumaraswamy, 1980), stem from this generalised beta type-I distribution.

The following figure gives a broader understanding and overview of the relationships between the different beta generators discussed in this paper. Note that the constant C in Figure 1 differs for each model (see (4) and (5)).

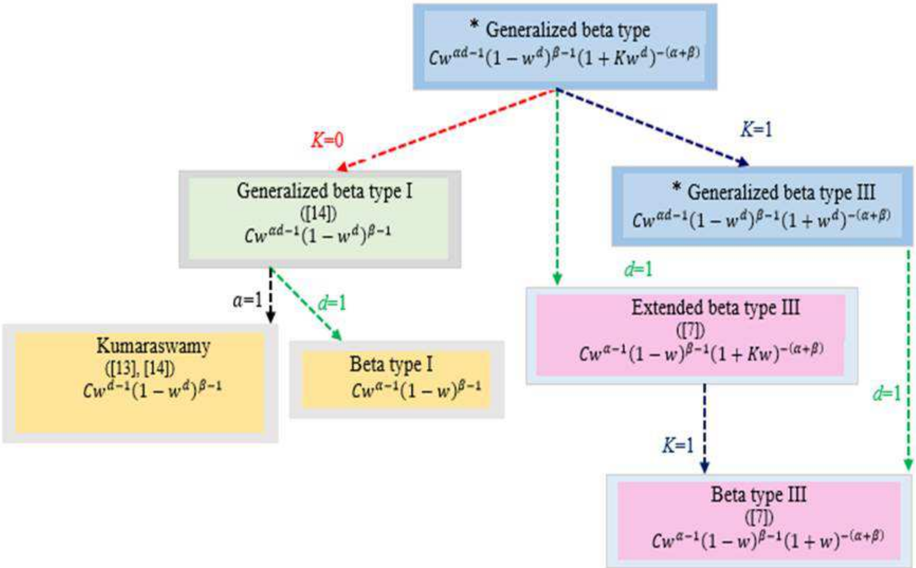

Figure 1: Relationships of the beta-type generator models in this paper.

It can be seen that different models stem from the generalised beta type distribution. For the first time in literature, the following beta type-generated models will be formulated in this paper:

- beta type III-Pareto;
- extended beta type III-Pareto;
- generalised beta type III-Pareto; and

- generalised beta type-Pareto.

The rest of this paper is organized as follows: Section 2 describes properties (cdf, hazard function, and moments) of the generalised beta type model (see (3)). Section 3 describes the beta type-generated class (by coupling with the Pareto distribution) as well as some statistical properties thereof. Section 4 presents an application on a hydrology data set and illustrates the usefulness of this class.

2. Generalised Beta Type Distribution

In this section, some properties of the newly proposed model with pdf (3) are briefly studied. Figure 2 illustrates the effect of each parameter on the pdf (3). Subsequently, Theorem 1 sees the derivation of the cdf of the generalised beta type distribution with pdf (3) after which an expression for the moments of the distribution is derived in Theorem 2.

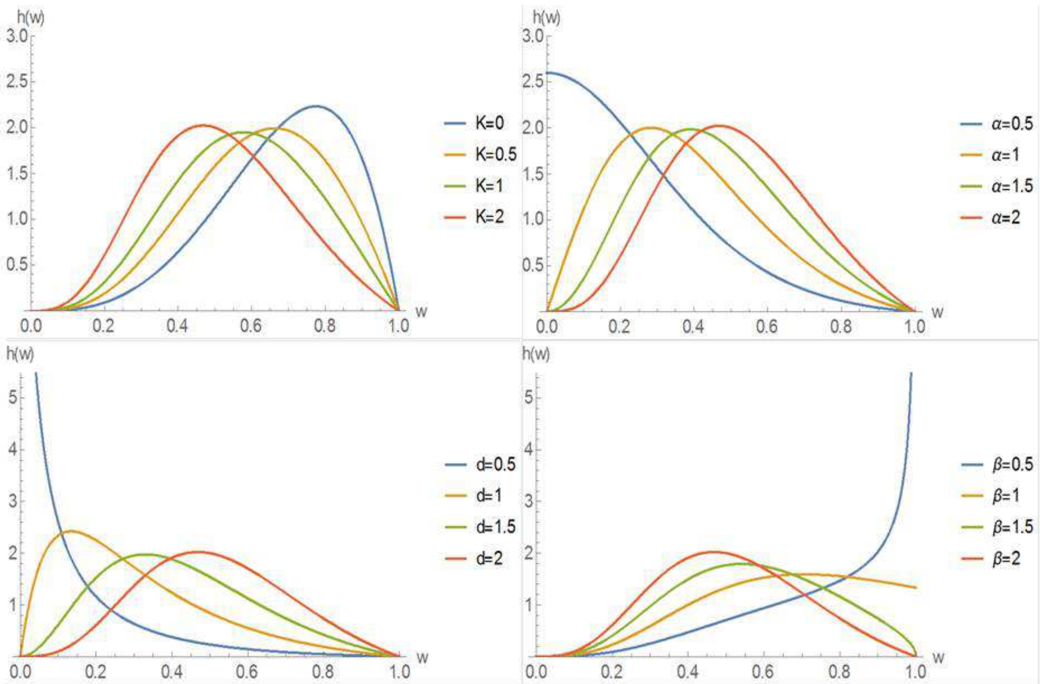

Figure 2: Effect of K (top left when $\alpha = \beta = d = 2$), α (top right when $\beta = d = K = 2$), d (bottom left when $\alpha = \beta = K = 2$), β (bottom right when $\alpha = d = K = 2$) on pdf (3).

Theorem 1 The cdf of the random variable $W \sim Genbeta(\alpha, \beta, d, K)$ with pdf (3) is given by

$$H(w) = \frac{B(w; \alpha,\ \beta)}{B(\alpha,\ \beta)}, \qquad 0 \leq w \leq 1, \tag{6}$$

with shape parameters $\alpha, \beta, d, K > 0$ where $B(z;\beta,\ \alpha)$ is the incomplete beta function (Gradshteyn and Ryzhik, 2007, p. 910, eq. 8.391):

$$\begin{aligned} B(w;\beta,\alpha) &= \frac{w^{\beta}}{\beta}\ {}_2F_1\left(\beta, 1-\alpha;\ 1+\beta;\ w\right) = \frac{w^{\beta}}{\beta}\sum_{i=0}^{\infty}\frac{(\beta)_i\,(1-\alpha)_i}{(\beta+1)_i\ i!}w^i \\ &= w^{\beta}\sum_{i=0}^{\infty}\frac{(1-\alpha)_i}{(\beta+i)\ i!}w^i, \end{aligned}$$

and the Gauss hypergeometric function ${}_2F_1(\cdot)$ (Mathai, 1993, p. 96) with $(\alpha)_i = \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)}$ as the Pochhammer coefficient.

Proof. Using generator (3), it follows that

$$H(w) = \frac{(1+K)^{\alpha}\,d}{B(\alpha,\ \beta)}\int_0^w t^{\alpha d-1}\ \left(1-t^d\right)^{\beta-1}\ \left(1+Kt^d\right)^{-(\alpha+\beta)}\ dt.$$

Let $y = t^d$, then $t = y^{\frac{1}{d}}$ and $\frac{dt}{dy} = \frac{1}{d}y^{\frac{1}{d}-1}$. Then from Gradshteyn and Ryzhik (2007, p. 25, eq. 1.110), and Gasper and Rahman (2004, p. 8, eq. 1.3.1), the result follows. ■

Note that the parameters K and d simplify out of expression (6).

Using (3), (6), and the definition from Bain and Engelhardt (1992), the hazard function is given by

$$q(w) = \frac{h(w)}{1-H(w)} = \frac{\frac{(1+K)^{\alpha}d}{B(\alpha,\ \beta)}\ w^{\alpha d-1}\ \left(1-w^d\right)^{\beta-1}\ \left(1+Kw^d\right)^{-(\alpha+\beta)}}{1-\frac{B(w;\alpha,\ \beta)}{B(\alpha,\ \beta)}}.$$

Next, the moments are derived for $W \sim Genbeta(\alpha,\ \beta,\ d,\ K)$ with pdf (3).

Theorem 2 If $W \sim Genbeta(\alpha,\ \beta,\ d,\ K)$ with pdf (3), then the s^{th} moment of W is given by

$$E\left[W^s\right] = \frac{B\left(\alpha+\frac{s}{d},\ \beta\right)}{B(\alpha,\ \beta)}, \tag{7}$$

where $\alpha,\ \beta,\ d,\ K > 0,\ s > 0$.

Proof.

$$E\left[W^s\right] = \frac{(1+K)^{\alpha}\,d}{B(\alpha,\ \beta)}\int_0^1\ w^{\alpha d+s-1}\ \left(1-w^d\right)^{\beta-1}\ \left(1+Kw^d\right)^{-(\alpha+\beta)}dw.$$

Let $y = w^d$, then $w = y^{\frac{1}{d}}$ and $\frac{dw}{dy} = \frac{1}{d}y^{\frac{1}{d}-1}$. Using Gradshteyn and Ryzhik (2007, p. 25, eq. 1.1.10) and Gasper and Rahman (2004, p. 8, eq. 1.3.1) the result follows. ■

Table 1 summarizes the expressions for the first four moments of (3).

Table 1: A summary of moment expressions of the generalised beta type distribution with pdf (3).

Moments	**Expression**
Mean $E(W)$ (μ)	$\dfrac{B\left(\alpha+\frac{1}{d},\beta\right)}{B(\alpha,\beta)}$
Variance $Var(W)$ (σ^2)	$\dfrac{B\left(\alpha+\frac{2}{d},\beta\right)}{B(\alpha,\beta)}-\mu^2$
Skewness (α_3)	$\dfrac{\frac{B\left(\alpha+\frac{3}{d},\beta\right)}{B(\alpha,\beta)}-3\mu\sigma^2-\mu^3}{\sigma^3}$
Kurtosis (α_4)	$\dfrac{B\left(\alpha+\frac{4}{d},\beta\right)}{B(\alpha,\beta)\,\sigma^4}$

Note that the parameter K simplifies out of expression (7), with the first few moments being K invariant.

Figures 3 and 4 illustrate the skewness and kurtosis of $W \sim Genbeta(\alpha,\beta,d,K)$ with pdf (3) for certain combinations of the parameters.

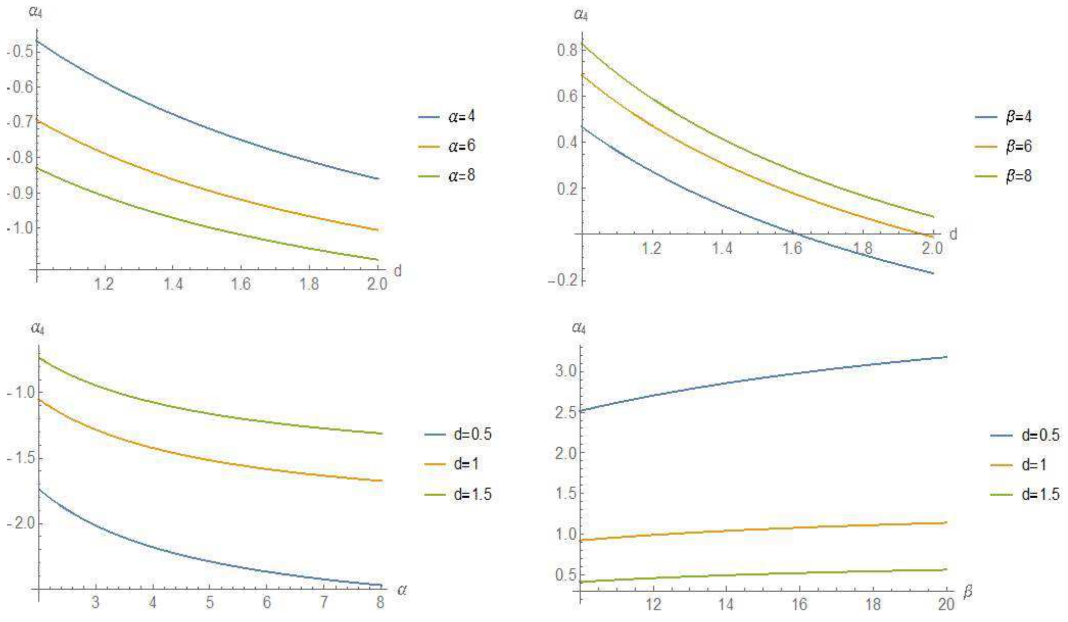

Figure 3: Skewness of $W \sim Genbeta(\alpha, \beta, d, K)$ with pdf (3). Effect of α (top left when $\beta = k = 2$), β (top right when $\alpha = k = 2$), and the effect of d (bottom left and right over α and β).

As the parameter values increase, the skewness decreases significantly, suggesting a negatively skewed distribution. This is seen in the top left and right graphs.

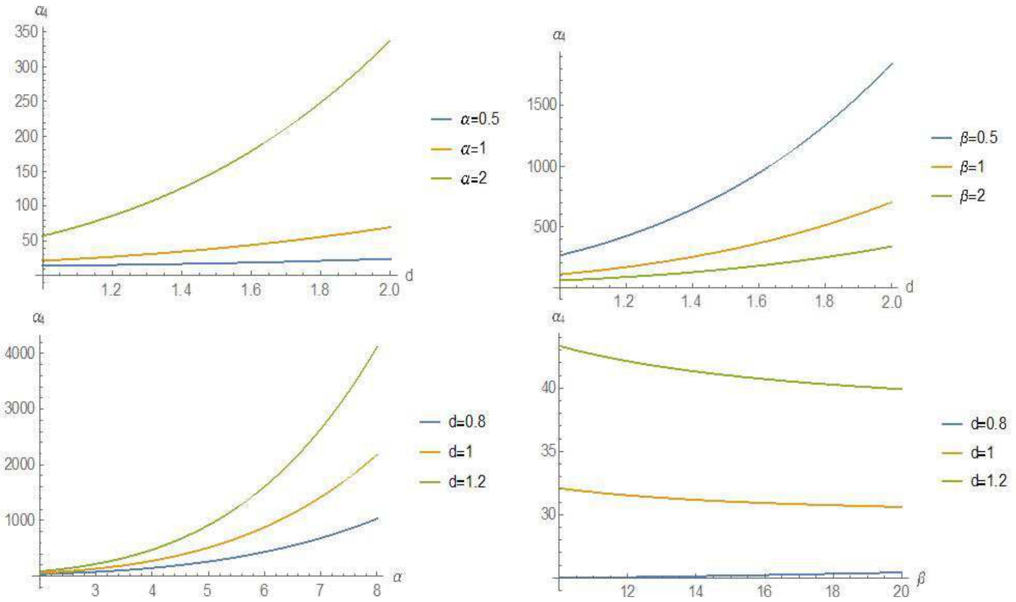

Figure 4: Kurtosis of $W \sim Genbeta(\alpha, \beta, d, K)$ with pdf (3). Effect of α (top left when $\beta = k = 2$), β (top right when $\alpha = k = 2$), and the effect of d (bottom left and right over α and β).

It is clear that the model with pdf (3) exhibits high curves as kurtosis is fairly high. This means that the values tend to have a distinct peak near the mean and decline rather rapidly, indicating the presence of heavy tails.

3. The Beta Type-Generated Class

3.1. Probability density function

In this section, the pdf and cdf of the *beta type-generated* distribution are derived (see (1)), with the generalised beta-type model (3) as the generator (see Figure 1), and the cdf of the Pareto as the parent distribution $G(x)$. The cdf of the Pareto distribution is given by

$$G(x) = 1 - \left(\frac{x}{\theta}\right)^{-k}, \quad x \geq \theta > 0, \tag{8}$$

with pdf

$$g(x) = \frac{k\theta^k}{x^{k+1}}, \quad x \geq \theta > 0,$$

for random variable X, with shape parameter $k > 0$ and scale parameter $\theta > 0$.

Therefore, the pdf of the generalised beta type-Pareto follows from (2) as

$$f(x) = \frac{d\,(1+K)^{\alpha}}{B(\alpha,\,\beta)}\,(G(x))^{\alpha d-1}\left(1-(G(x))^{d}\right)^{\beta-1}\left(1+K\,(G(x))^{d}\right)^{-(\alpha+\beta)} g(x) \tag{9}$$

$$= \frac{d\,(1+K)^{\alpha}}{B(\alpha,\,\beta)}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha d-1}\left(1-\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{\beta-1} \times\left(1+K\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{-(\alpha+\beta)}\frac{k\theta^{k}}{x^{k+1}} \tag{10}$$

$$= \frac{(1+K)^{\alpha}\,dk}{\theta B(\alpha,\,\beta)}\left(\frac{x}{\theta}\right)^{-k-1}\sum_{i=0}^{\infty}\sum_{r=0}^{\infty}(-1)^{r}\binom{1-\beta}{i}\frac{(\alpha+\beta)_{r}}{r!} \times K^{r}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d(\alpha+i+r)-1}, \tag{11}$$

with $x \geq \theta > 0$, α, β, k, d, $K > 0$, and is denoted by $X \sim \mathit{Genbeta\text{-}Pareto}(\alpha,\beta,\theta,k,d,K)$. By expanding the binomial function in (9), with the Pareto distribution as the chosen parent, $G(x)$ (see (8)), the latter equation (11) can be described as a mixture of exponentiated Pareto distributions. This generated distribution (11) has cdf

$$F(x) = 1-\frac{(1+K)^{\alpha}\,d}{B(\alpha,\,\beta)}\sum_{i=0}^{\infty}\sum_{r=0}^{\infty}(-1)^{r}\,\frac{(1-\beta)_{i}}{i!}\,\frac{(\alpha+\beta)_{r}}{r!}\,K^{r}\,B(z;1,\,d\,(\alpha+i+r)),$$

where $z = \left(\frac{x}{\theta}\right)^{-k}$.

Figure 5 illustrates the generalised beta type-Pareto distribution's pdf (11) for different arbitrary parameter values, of K, k, d, when $\alpha = \beta = 2$ and $\theta = 3$.

The hazard function for this case is given by

$$q(x) = \frac{\frac{k}{\theta}\,z^{-1}\,(1-z)^{\alpha d-1}\left(1-(1-z)^{d}\right)^{\beta-1}\left(1+K\,(1-z)^{d}\right)^{-(\alpha+\beta)}}{\sum_{i=0}^{\infty}\sum_{r=0}^{\infty}(-1)^{r}\,\frac{(1-\beta)_{i}}{i!}\,\frac{(\alpha+\beta)_{r}}{r!}\,K^{r}\,B(z;1,d\,(\alpha+i+r))},$$

with $x \geq \theta > 0$, α, β, k, d, $K > 0$ and $z = \left(\frac{x}{\theta}\right)^{-k}$.

The derivations for the other beta type-generators in Figure 1 is similar and is summarised in Table 2.

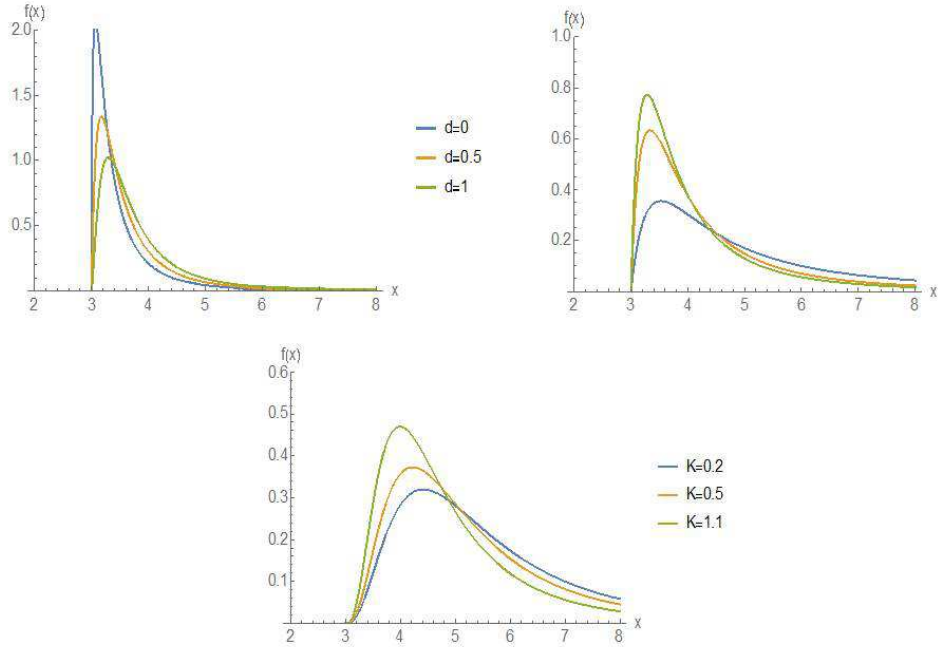

Figure 5: Effect of d (top left when $\alpha=\beta=k=K=2$), k (top right when $\alpha=\beta=d=K=2$), and K (bottom/middle when $\alpha=\beta=k=d=2$) and $\theta=3$ on pdf (11).

Table 2: Pdfs of beta-type generated models.

Beta type generator	**Pdf**
1. *generalised beta type-Pareto*	$\frac{(1+K)^{\alpha}dk}{\theta B(\alpha,\beta)}\left(\frac{x}{\theta}\right)^{-k-1}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha d-1}\left(1-\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{\beta-1}$ $\times\left(1+K\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{-(\alpha+\beta)}$
2. *generalised beta type III-Pareto*	$\frac{2^{\alpha}dk}{\theta B(\alpha,\beta)}\left(\frac{x}{\theta}\right)^{-k-1}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha d-1}\left(1-\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{\beta-1}$ $\times\left(1+\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{-(\alpha+\beta)}$
3. *extended beta type III-Pareto*	$\frac{k(1+K)^{\alpha}}{\theta B(\alpha,\beta)}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha-1}\left(\frac{x}{\theta}\right)^{-k\beta-1}\left(1+K\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)\right)^{-(\alpha+\beta)}$
4. *beta type III-Pareto*	$\sum_{i=0}^{\infty}\frac{(\beta)_i}{2^{(\beta+i)}i!}\frac{k}{\theta B(\alpha,\beta+i)}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha-1}\left(\frac{x}{\theta}\right)^{-k(\beta+i)-1}$
5. *generalised beta type I-Pareto*	$\frac{dk}{\theta B(\alpha,\beta)}\left(\frac{x}{\theta}\right)^{-k-1}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha d-1}\left(1-\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{\beta-1}$
6. *beta type I-Pareto*	$\frac{k}{\theta B(\alpha,\beta)}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha-1}\left(\frac{x}{\theta}\right)^{-k\beta-1}$
7. *Kumaraswamy-Pareto*	$\frac{dk}{\theta B(1,\beta)}\left(\frac{x}{\theta}\right)^{-k-1}\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d-1}\left(1-\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{d}\right)^{\beta-1}$

3.2. Moments

In this section, an expression for the moments of the generalised beta type-Pareto distribution with pdf (10) is derived. It is used thereafter for investigating the skewness and kurtosis of the distribution.

Theorem 3 If $X \sim$ *Genbeta-Pareto*$(\alpha,\ \beta,\ \theta,\ k,\ d,\ K)$ with pdf (11), then the s^{th} moment of X is given by

$$E\left[X^s\right] = \theta^s\ \frac{(1+K)^{-\beta}}{B(\alpha,\beta)}\ \sum_{i=0}^{\infty} \frac{(\frac{s}{k})_i}{i!}\ B\left(\alpha+\frac{i}{d},\beta\right)\ {}_2F_1\left(\alpha+\beta,\ \beta;\ \alpha+\frac{i}{d}+\beta;\ \frac{K}{K+1}\right)$$

where $\alpha,\ \beta,\ d,\ K,\ \theta,\ s>0$.

Proof. Consider the following from (10):

$$\begin{aligned} E\left[\left(\frac{X}{\theta}\right)^s\right] &= \int_{-\infty}^{\infty} \left(\frac{x}{\theta}\right)^s \frac{(1+K)^{\alpha} d}{B(\alpha,\ \beta)} \left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^{\alpha d-1} \left(1-\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^d\right)^{\beta-1} \\ &\quad \times \left(1+K\left(1-\left(\frac{x}{\theta}\right)^{-k}\right)^d\right)^{-(\alpha+\beta)} k\theta^k \frac{1}{x^{k+1}} dx. \end{aligned}$$

Let $y=\left(\frac{x}{\theta}\right)^{-k}$, then $x=\theta y^{-\frac{1}{k}}$ and $\frac{dx}{dy}=-\frac{\theta}{k}y^{-\frac{1}{k}-1}$. Then

$$E\left[\left(\frac{X}{\theta}\right)^s\right] = \int_0^1 \frac{(1+K)^{\alpha} d}{B(\alpha,\beta)}\ y^{-\frac{s}{k}}\ (1-y)^{\alpha d-1}\ \left(1-(1-y)^d\right)^{\beta-1}\ \left(1+K(1-y)^d\right)^{-(\alpha+\beta)}\ dy.$$

Let $z=(1-y)^d$, then $y=1-z^{\frac{1}{d}}$ and $\frac{dy}{dz}=-\frac{1}{d}z^{\frac{1}{d}-1}$. Then from Gradshteyn and Ryzhik (2007, p. 317, eq. 3.197.3),

$$\begin{aligned} E\left[\left(\frac{X}{\theta}\right)^s\right] &= \int_0^1 \frac{(1+K)^{\alpha} d}{B(\alpha,\beta)}\ \left(1-z^{\frac{1}{d}}\right)^{-\frac{s}{k}}\ z^{\alpha-\frac{1}{d}}\ (1-z)^{\beta-1}\ (1+Kz)^{-(\alpha+\beta)}\ \frac{1}{d}\ z^{\frac{1}{d}-1}\ dz \\ &= \frac{(1+K)^{-\beta}}{B(\alpha,\beta)}\ \sum_{i=0}^{\infty} \frac{(\frac{s}{k})_i}{i!}\ B\left(\alpha+\frac{i}{d},\beta\right)\ {}_2F_1\left(\alpha+\beta,\ \beta;\ \alpha+\frac{i}{d}+\beta;\ \frac{K}{K+1}\right). \end{aligned}$$

From Gradshteyn and Ryzhik (2007, p. 8, eq. 1.3.1) the result follows. ■

Table 3 gives a summary of some moment expressions for the generalised beta type-Pareto distribution with pdf (11). This includes, respectively, the mean $(E(X))$, variance $(Var(X))$, skewness coefficient (α_3) and kurtosis (α_4). Further simplifications prove that the third and fourth moments α_3, α_4 are θ invariant.

Table 3: A summary of moment expressions of the generalised beta type-Pareto distribution with pdf (10).

Moments	Expression
Mean, (μ) $E(X)$	$\theta\,\frac{(1+K)^{-\beta}}{B(\alpha,\beta)}\sum_{i=0}^{\infty}\frac{(\frac{1}{k})_i}{i!}B\left(\alpha+\frac{i}{d},\beta\right)\,{}_2F_1\left(\alpha+\beta,\beta;\alpha+\frac{i}{d}+\beta;\frac{K}{K+1}\right)$
Variance, (σ^2) $Var(X)$	$\theta^2\frac{(1+K)^{-\beta}}{B(\alpha,\beta)}\sum_{i=0}^{\infty}\frac{(\frac{2}{k})_i}{i!}B\left(\alpha+\frac{i}{d},\beta\right)\,{}_2F_1\left(\alpha+\beta,\ \beta;\alpha+\frac{i}{d}+\beta;\ \frac{K}{K+1}\right)-\mu^2$
Skewness (α_3)	$\dfrac{\theta^3\frac{(1+K)^{-\beta}}{B(\alpha,\beta)}\sum_{i=0}^{\infty}\frac{(\frac{3}{k})_i}{i!}B\left(\alpha+\frac{i}{d},\beta\right)\,{}_2F_1\left(\alpha+\beta,\beta;\ \alpha+\frac{i}{d}+\beta;\frac{K}{K+1}\right)-3\mu\sigma^2-\mu^3}{\sigma^3}$
Kurtosis (α_4)	$\dfrac{\theta^4\frac{(1+K)^{-\beta}}{B(\alpha,\beta)}\sum_{i=0}^{\infty}\frac{(\frac{4}{k})_i}{i!}B\left(\alpha+\frac{i}{d},\beta\right)\,{}_2F_1\left(\alpha+\beta,\beta;\ \alpha+\frac{i}{d}+\beta;\frac{K}{K+1}\right)}{\sigma^4}$

The visual representation of the skewness and kurtosis for the *new* beta-type generated model are given in Figures 6 and 7 for different arbitrary parameter values and specifically when $\alpha=\beta=2$.

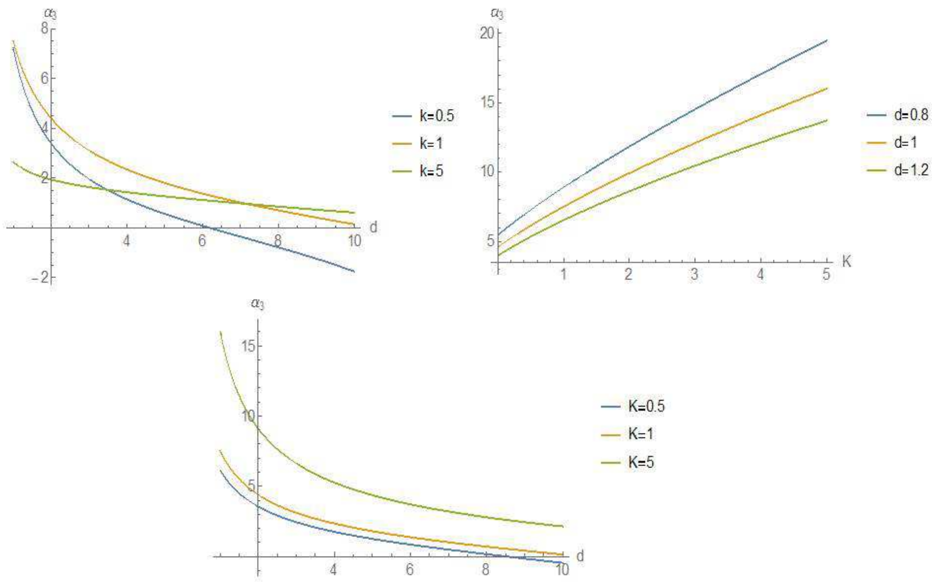

Figure 6: Skewness (α_3) of the generalised beta type-Pareto distribution. Effect of k (top left when $\alpha=\beta=2$, $K=1$) over d, effect of d (top right when $\alpha=\beta=2$, $k=1$) over K, and the effect of K (bottom/middle when $\alpha=\beta=2$, $k=1$) over d, with $\theta=1$ for all graphs in this figure.

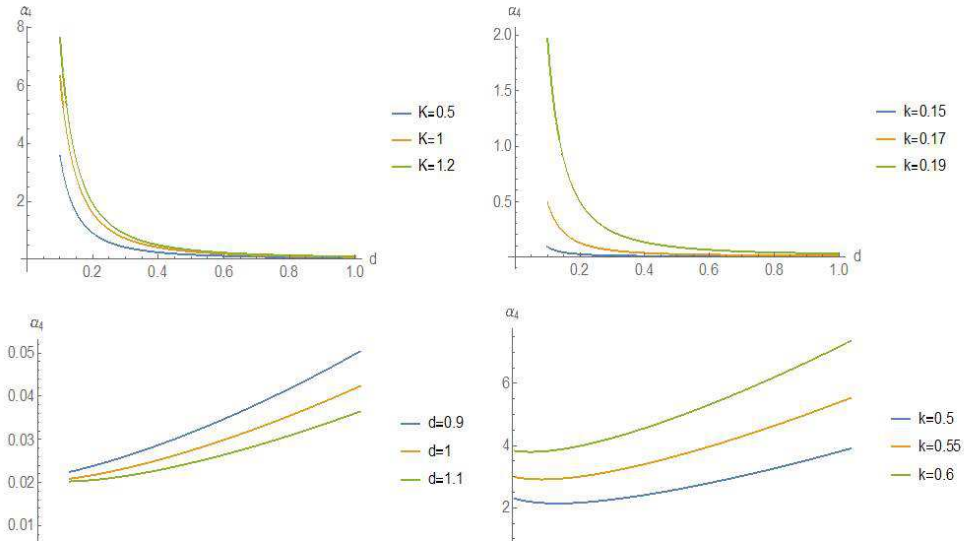

Figure 7: Kurtosis (α_4) of the generalised beta type-Pareto distribution. Effect of K (top left when $\alpha = \beta = 2$, $k = 0.2$) over d, effect of k (top right when $\alpha = \beta = 2$, $K = 0.5$) over d, effect of d (bottom left when $\alpha = \beta = 2$, $k = 0.19$) over K, and again the effect of k (bottom right when $\alpha = \beta = 2$, $d = 6$) over K, with $\theta = 1$ for all graphs in this figure.

The kurtosis of $X \sim \textit{Genbeta-Pareto}(\alpha, \beta, \theta, k, d, K)$ is very sensitive to the value of k. The effect of α and β is not illustrated, since it has the same role as the generator distribution, namely the beta type I.

3.3. The Rényi entropy

The Rényi entropy is introduced by Rényi (1961), which is a measure to quantify diversity, uncertainty or randomness of a system. Bromiley, Thacker and Bouhord-Thacker (2010) explained that the entropy formula can be used to measure the "peakedness" of a distribution. Akinsete et al. (2008) further explained that the entropy of a random variable is a measure of variation of the uncertainty. It is defined as

$$I_R(\xi) = \frac{1}{1-\xi} \log \left\{ \int_{-\infty}^{\infty} f^{\xi}(x) dx \right\}, \qquad \xi \neq 0, \xi > 0$$

where $f(x)$ is the pdf of the random variable X.

It follows that the Rényi entropy for $X \sim GenBetatypePareto(\alpha,\ \beta,\ d,\ \theta,\ k,\ K)$ is

$$\begin{aligned} I_R(\xi) &= \frac{\xi}{1-\xi}\log\left((1+K)^{\alpha} d\right) - \log\left(\frac{k}{\theta}\right) \\ &\quad + \frac{1}{1-\xi}\log\left\{\sum_{i=0}^{\infty}\sum_{r=0}^{\infty}(-1)^r \frac{(\xi(1-\beta))_i}{i!}\frac{(\xi(\alpha+\beta))_r}{r!}K^r \right. \\ &\quad \left. \times \frac{B\left(\xi(\alpha d-1)+d(i+r)+1, \xi\left(1+\frac{1}{k}\right)-\frac{1}{k}\right)}{B^{\xi}(\alpha,\beta)}\right\}. \end{aligned} \tag{12}$$

A visual representation of the Rényi entropy of the generalised beta type-Pareto model is given in Figure 8.

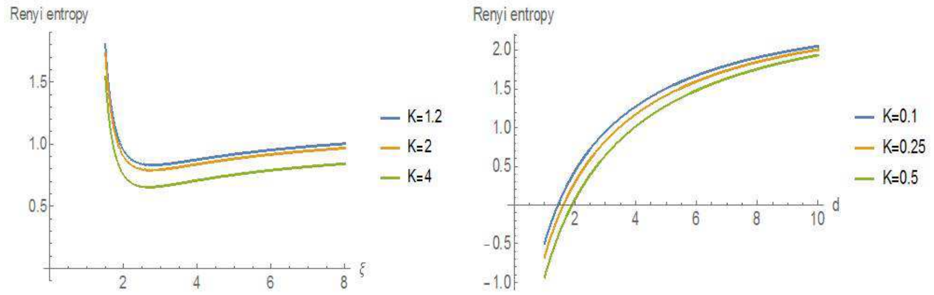

Figure 8: The Rényi entropy (12) for values of (left) $\alpha = \beta = d = 2$, $k = 0.6$, $\theta = 4$ over ξ and (right) $\alpha = \beta = 1$, $d = 0.5$, $k = 0.1$, $\theta = 1$ over d.

For this combination of parameter values, the Rényi entropy decreases as K increases. The above figure illustrates that as d increases, the Rényi entropy becomes asymptotic to the value of 2. More speculative research for different combinations of parameter values shows the same behaviour.

4. Application and Conclusion

In this paper we proposed a new class of distributions by coupling the Pareto distribution with a new beta type class. Here we illustrate the importance of the class in an application to a real dataset. The data corresponds to the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada (Akinsete et al., 2008). Wheaton River is a stream in the region of Yukon Territory, the country of Canada with an average elevation of 656 metres above sea level. The location is sparsely populated with 0 people per km^2. The nearest town larger than 50,000 inhabitants takes about 22 hours by local transportation. Wheaton River can have very strong earthquakes (on average one every 50 years), measuring between $6-7$ on the Richter scale. The damage will be negligible in buildings of good design and construction, but considerable damage may be inflicted on poorly built or badly designed structures. It is with this motivation why the fitting of distributions to the exceedances of these flood peaks are of vital importance. For this application, the parameters are estimated based on the minimisation of the Kolmogorov-Smirnov (KS) distance between the analytical and empirical distribution functions. Weber, Leemis and Kincaid (2006) defined the KS measure as

$$KS = \max\left[D_n^+, D_n^-\right],$$

where

$$D_n^+ = \max\left|\frac{i}{n} - \widehat{G}(x_{(i)})\right| \; i = 1, 2, \cdots, n$$

and

$$D_n^- = \max\left|\widehat{G}(x_{(i)}) - \frac{i-1}{n}\right| \; i = 1, 2, \cdots, n.$$

The KS distance values of the parameter estimates of the Pareto, beta type I-Pareto and generalised beta-type Pareto are given in Table 4 below with $\hat{\theta} = x_{\min} = 0.1$. In Figure 10, the parameter estimated values are also used to graphically demonstrate how the Rényi entropy behaves under the beta type I-Pareto and generalised beta-type Pareto distributions. The authors consider only the generalised beta type-Pareto and not all other developed models, since this is the flexible model that contains all the other cases.

Table 4: Parameter estimates for Wheaton River data.

	Parameter estimates					
Models	α	β	k	d	K	KS measure
Pareto	n/a	n/a	0.2	n/a	n/a	0.29
Beta type I-Pareto	3.35	3.88	0.17	n/a	n/a	0.17
Generalised beta type-Pareto	2.31	9.87	0.19	3.1	0.17	0.11

A visual representation of how the above three distributions as well as the empirical distribution fits to the dataset.

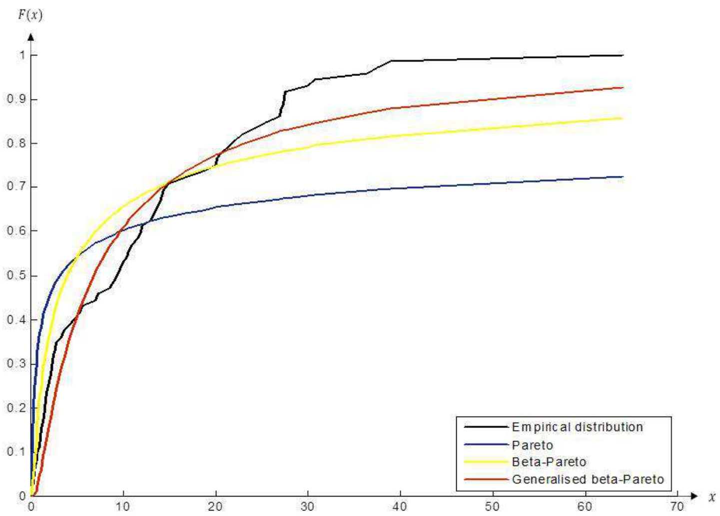

Figure 9: Distribution fits to the Wheaton River data, using the empirical cdf, Pareto cdf, beta type I-Pareto cdf, generalised beta type-Pareto cdf.

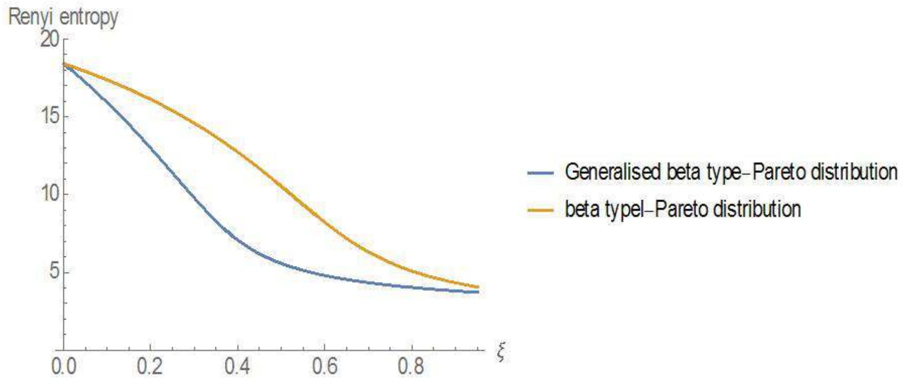

Figure 10: Rényi entropy of the beta type I-Pareto and generalised beta type-Pareto distributions.

Figure 9 illustrates the model fit for the Pareto, beta type I-Pareto-, and the generalised beta type-Pareto together with the empirical distribution function of the Wheaton river data. As suggested by the KS measure, the generalised beta type-Pareto provides an improved fit. This observation demonstrates the noteworthy contribution of this broader beta type-Pareto class to literature.

Furthermore, Figure 10 illustrates the Rényi entropy for this dataset under the beta type I-Pareto- and the generalised beta type-Pareto distributions. The observations are favourable for the generalised beta type-Pareto distribution as it indicates lower Rényi entropy for this dataset.

Acknowledgements

The authors acknowledge the support of the StatDisT group, and are also thankful for the reviewer's constructive comments. This work is based upon research supported by the National Research Foundation, South Africa (ref. CPRR13090132066 grant nr: 91497).

References

AKINSETE, A., FAMOYE, F., AND LEE, C. (2008). The beta-Pareto distribution. *Statistics*, **42** (6), 547–563.

BAIN, L. J. AND ENGELHARDT, M. (1992). *Introduction to Probability and Mathematical Statistics*. Second edition. Duxbury Press Belmont: California.

BALAKRISHNAN, N. AND RISTIC, M. (2016). Multivariate families of gamma-generated distributions with finite or infinite support above or below the diagonal. *Journal of Multivariate Analysis*, **143**, 194–207.

BARRETO-SOUZA, W., SANTOS, A. H. S., AND CORDEIRO, G. M. (2010). The beta generalized-exponential distribution. *Journal of Statistical Computation and Simulation*, **80** (2), 159–172.

BOURGUIGNON, M., SILVA, R. B., ZEA, L. M., AND CORDEIRO, G. M. (2012). The Kumaraswamy-Pareto distribution. *Journal of Statistical Theory and Applications*, **12** (6), 129–144.

BROMILEY, P. A., THACKER, N. A., AND BOUHORD-THACKER, E. (2010). Shannon entropy, Renyi entropy, and information. Tina memo 2008-001, University of Manchester.

CARDEÑO, L., NAGAR, D. K., AND SÁNCHEZ, L. E. (2005). Beta type 3 distribution and its multivariate generalisation. *Tamsui Oxford Journal of Mathematical Sciences*, **21** (2), 225–241.

CORDEIRO, G. M., NOBRE, J. S., PESCIM, R. R., AND ORTEGA, E. M. M. (2012). The Beta Moyal: A useful-skew distribution. *International Journal of Research and Reviews in Applied Sciences*, **10**, 171–192.

EUGENE, N., LEE, C., AND FAMOYE, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, **31** (4), 497–512.

GASPER, G. AND RAHMAN, M. (2004). *Basic Hypergeometric Series*. Cambridge University Press: Cambridge.

GRADSHTEYN, I. S. AND RYZHIK, I. M. (2007). *Table of Integrals, Series and Products*. Seventh edition. Academic Press: San Diego.

JONES, M. C. (2004). Families of distributions arising from distributions of order statistics. *Test*, **13** (1), 1–43.

KUMARASWAMY, P. (1980). Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology*, **46**, 79–88.

MATHAI, A. M. (1993). *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*. Clarendon Press: Oxford.

MCDONALD, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica,*, **52**, 647–663.

NADARAJAH, S. AND GUPTA, A. K. (2004). The beta-Fréchet distribution. *Far East Journal of Theoretical Statistics*, **14**, 15–24.

NADARAJAH, S. AND KOTZ, S. (2004). The beta-Gumbel distribution. *Mathematical Problems in Engineering*, **10** (4), 323–332.

NADARAJAH, S. AND KOTZ, S. (2006). The beta-exponential distribution. *Reliability Engineering and System Safety*, **91** (6), 689–697.

RÉNYI, A. (1961). On measures of entropy and information. *Fourth Berkeley symposium on mathematical statistics and probability*, **1**, 547–561.

RISTIC, M. AND BALAKRISHNAN, N. (2012). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, **82** (8), 1191–1206.

WEBER, M. D., LEEMIS, L. M., AND KINCAID, R. K. (2006). Minimum Kolmogorov-Smirnov test statistic parameter estimates. *Journal of Statistical Computation and Simulation*, **76** (3), 195–205.

ZOGRAFOS, K. AND BALAKRISHNAN, N. (2009). On families of beta and generalized gamma generated distributions and associated inference. *Statistical Methodology*, **6** (4), 344–362.

Manuscript received, 2016-06-20, revised, 2016-09-09, accepted, 2016-09-12.