



ELSEVIER

Contents lists available at ScienceDirect

## Statistical Methodology

journal homepage: [www.elsevier.com/locate/stamet](http://www.elsevier.com/locate/stamet)

# On families of beta- and generalized gamma-generated distributions and associated inference

K. Zografos<sup>a,\*</sup>, N. Balakrishnan<sup>b</sup>

<sup>a</sup> University of Ioannina, Ioannina, Greece

<sup>b</sup> McMaster University, Hamilton, Ontario, Canada

## ARTICLE INFO

### Article history:

Received 25 September 2008

Received in revised form

5 December 2008

Accepted 8 December 2008

### Keywords:

Beta-generated distributions

Stacy's generalized gamma distribution

Beta-exponential distribution

Discrimination between distributions

Maximum entropy principle

Shannon entropy

## ABSTRACT

A general family of univariate distributions generated by beta random variables, proposed by Jones, has been discussed recently in the literature. This family of distributions possesses great flexibility while fitting symmetric as well as skewed models with varying tail weights. In a similar vein, we define here a family of univariate distributions generated by Stacy's generalized gamma variables. For these two families of univariate distributions, we discuss maximum entropy characterizations under suitable constraints. Based on these characterizations, an expected ratio of quantile densities is proposed for the discrimination of members of these two broad families of distributions. Several special cases of these results are then highlighted. An alternative to the usual method of moments is also proposed for the estimation of the parameters, and the form of these estimators is particularly amenable to these two families of distributions.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

Recently, attempts have been made to define new families of probability distributions that extend well-known families of distributions and at the same time provide great flexibility in modelling data in practice. One such example is a broad family of univariate distributions generated from the beta distribution, proposed by Jones [1] (see also [2]), which extends the original beta family of distributions with the incorporation of two additional parameters. These parameters control the skewness and the tail weight. Earlier, with a similar goal in mind, Eugene et al. [3] defined the family of beta-normal distributions and discussed its properties.

\* Corresponding address: University of Ioannina, Department of Mathematics, 45110 Ioannina, Greece.  
E-mail address: [kzograf@uoi.gr](mailto:kzograf@uoi.gr) (K. Zografos).

Following the notation of Jones [1], the class of “beta-generated distributions” is defined as follows. Consider a continuous distribution function  $F$  with density function  $f$ . Then, the univariate family of distributions generated by  $F$ , and the parameters  $\alpha, \beta > 0$ , has its pdf as [1]

$$g_F^{(B)}(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} f(x) \{F(x)\}^{\alpha-1} \{1 - F(x)\}^{\beta-1}, \quad \alpha > 0 \text{ and } \beta > 0, \quad (1)$$

where  $B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt$  is the complete beta function. Thus, this family of distributions has its cdf as

$$G_F^{(B)}(x) = I_{F(x)}(\alpha, \beta), \quad \alpha > 0 \text{ and } \beta > 0, \quad (2)$$

where the function  $I_{F(x)}$  denotes the incomplete beta ratio defined by

$$I_y(\alpha, \beta) = \frac{B_y(\alpha, \beta)}{B(\alpha, \beta)},$$

where

$$B_y(\alpha, \beta) = \int_0^y t^{\alpha-1} (1-t)^{\beta-1} dt, \quad 0 < y < 1,$$

is the incomplete beta function.

Following the terminology of Arnold in the discussion of Jones' [1] paper, the distribution  $F$  will be referred to as the “parent distribution” in what follows. Based on Jones and Larsen [4], the attractiveness of (1) is that from a symmetric  $f$  as parent pdf (corresponding to  $\alpha = \beta = 1$ ), a large family of distributions can be generated with the parameters  $\alpha$  and  $\beta$  controlling the skewness and the tail weight. The expression in (2) reveals that it is quite easy to simulate observations from  $X \sim G_F^{(B)}$ , as shown by Jones [1], through the relationship  $X = F^{-1}(B)$ , where  $B \sim \text{Beta}(\alpha, \beta)$ . The case  $\alpha = \beta = 1$  corresponds to the well-known quantile function representation  $X = F^{-1}(U)$ , where  $U \sim U(0, 1)$ , which is used in order to generate data from the distribution  $F$ . Finally, in the case when  $\alpha$  and  $\beta$  are positive integers, the beta-generated model in (1) is the distribution of the  $i$ th order statistic in a random sample of size  $n$  from distribution  $F$ , where  $i = \alpha$  and  $n = \alpha + \beta - 1$ .

The idea of beta-generated family of distributions  $g_F$  stemmed from the paper of Eugene et al. [3], wherein the beta-normal distribution was introduced and its properties were studied. Specifically, if  $\phi$  denotes the density of the normal distribution and  $\Phi$  the corresponding distribution function, then  $g_\phi$  is the beta-normal distribution considered by Eugene et al. [3]. Some other beta-generated families of distributions have also been discussed in the literature. For example, the beta-exponential distribution has been defined and studied by Nadarajah and Kotz [5]. Similarly, the beta-logistic distribution can also be generated through the beta variable, but it has been known as a tractable set of statistical models based on the logarithm of a  $F$ -variate; see [6]. The beta-logistic distribution has been reviewed by Jones [1] who has also discussed the skew- $t$  distributions. All these members of the beta-generated family in (1) and some others have been reobtained by means of the maximum entropy principle by Zografos [7] who has also considered the beta-Weibull distribution through this principle. Jones' family of beta-generated distributions in (1) has received great attention recently. Arnold et al. [8] introduced and studied a multivariate version of this family, while Ferreira and Steel [9,10] used it as a skewing mechanism for constructing skewed distributions.

The aim of this paper is two-fold. In the first part, we will concentrate on the family in (1). Our main concern will then be to construct procedures to discriminate between members of this family. In other words, our aim will be to derive a test which would enable us to decide if a random sample from (1) is coming from a specific parent distribution  $F$ . The proposed procedures will be based on information theoretic methods and, in particular, on the maximum entropy principle. In the second part, we will define a broad family of univariate distributions, in the same vein as Jones' family, through Stacy's generalized gamma density generated by the parent distribution  $F$ .

In Section 2, we present suitable constraints leading to the maximum entropy characterization of the family in (1). Section 3 is devoted to the definition of a new family of distributions through Stacy's generalized gamma density generated by the parent distribution  $F$ , and the derivation of constraints

leading to its maximum entropy characterization. In Section 4, an alternative to the method of moments is discussed for the estimation of the parameters of beta- and generalized gamma-generated distributions with a parent distribution  $F$ . The constraints needed to obtain the maximum entropy characterization of these families of distributions enable us to introduce in Section 5 test statistics for the discrimination between members within these two families. Several univariate distributions, generated by beta and gamma models, will be presented in illustrative examples and their moments and Shannon entropies will be derived in a closed form.

**2. Jones’ distribution and maximum entropy identification**

The notion of entropy is of fundamental importance in different areas such as physics, probability and statistics, communication theory, and economics. Since Shannon’s [11] pioneering work on the mathematical theory of communication, Shannon entropy of a continuous distribution with density, say  $g_F$ , defined by

$$\mathcal{H}_{Sh}(g_F) = - \int_{-\infty}^{\infty} g_F(x) \ln g_F(x) dx, \tag{3}$$

has become a major tool in information theory and in almost every branch of science and engineering. Closely related to the Shannon entropy is the maximum entropy method for the identification of a probabilistic model. This method considers a class of density functions

$$\mathcal{F} = \{g_F(x) : E_{g_F}[T_i(X)] = \alpha_i, i = 0, 1, \dots, m\},$$

where  $T_i, i = 0, 1, \dots, m$ , are absolutely integrable functions with respect to  $g_F$ , and  $T_0(X) = \alpha_0 = 1$ . In the continuous case, the maximum entropy principle suggests to derive the unknown density function of the random variable  $X$  by the model that maximizes the Shannon entropy in (3), subject to the information constraints defined in the class  $\mathcal{F}$ . This method, introduced by Jaynes [12], as a general method of inference, has been treated axiomatically by Shore and Johnson [13]. It has been successfully applied in a wide variety of fields and has also been used for the characterization of several standard probability distributions; see, for example, [14,15,7], and the references contained therein.

The maximum entropy distribution, denoted by the density  $g_F^{ME}$  of the class  $\mathcal{F}$ , is obtained as the solution of the optimization problem

$$g_F^{ME}(x) = \arg \max_{g_F \in \mathcal{F}} \mathcal{H}_{Sh}(g_F).$$

Jaynes [12, p. 623], stated that the maximum entropy distribution  $g_F^{ME}$ , obtained by the above constrained maximization problem, “is the only unbiased assignment we can make; to use any other would amount to arbitrary assumption of information which by hypothesis we do not have”. It is the distribution which should not incorporate any additional exterior information other than what is specified by the constraints.

In order to provide a maximum entropy characterization of Jones’ family in (1), we need to derive suitable constraints that define the above class  $\mathcal{F}$ . For this purpose, the next two lemmas play a crucial role.

**Lemma 1.** *If  $F$  and  $f$  are the parent cdf and pdf, respectively, and  $g_F^{(B)}$  is the corresponding density of Jones’ family in (1), then*

- (a)  $E_{g_F^{(B)}}[\ln F(X)] = \Psi(\alpha) - \Psi(\alpha + \beta)$ ;
- (b)  $E_{g_F^{(B)}}[\ln(1 - F(X))] = \Psi(\beta) - \Psi(\alpha + \beta)$ ;
- (c)  $E_{g_F^{(B)}}[\ln f(X)] = E_Y[\ln f(F^{-1}(Y))]$ ,

where  $Y \sim \text{Beta}(\alpha, \beta)$  and  $\Psi$  denotes the digamma function.

**Proof.** (a) Using the transformation  $y = F(x)$ ,

$$E_{g_F^{(B)}}[\ln F(X)] = \frac{1}{B(\alpha, \beta)} \int_0^1 (\ln y) y^{\alpha-1} (1 - y)^{\beta-1} dy. \tag{4}$$

It is easy to see that

$$\frac{\partial}{\partial \alpha} B(\alpha, \beta) = \int_0^1 (\ln y) y^{\alpha-1} (1-y)^{\beta-1} dy, \tag{5}$$

while

$$\frac{\partial}{\partial \alpha} B(\alpha, \beta) = \frac{\partial}{\partial \alpha} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} = B(\alpha, \beta) \{ \Psi(\alpha) - \Psi(\alpha + \beta) \}. \tag{6}$$

Upon using Eqs. (4)–(6), we arrive at the desired result.

(b) The proof is similar to that of Part (a) and is therefore omitted.

(c) In order to prove this part, upon using the transformation  $y = F(x)$ , we obtain

$$\begin{aligned} E_{g_F^{(B)}}[\ln f(X)] &= \int_0^1 \{ \ln f(F^{-1}(y)) \} \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1} dy \\ &= E_Y[\ln f(F^{-1}(Y))], \end{aligned}$$

where  $Y \sim \text{Beta}(\alpha, \beta)$ . ▲

It is well known that if  $X$  is a continuous random variable with distribution function  $F$ , then  $Y = F(X)$  follows a uniform distribution  $U(0, 1)$ . The next lemma states an extension of this result, in the sense that if the random variable  $X$  has Jones' density in (1) with parent distribution  $F$ , then the random variable  $Y = F(X)$  follows a beta distribution  $\text{Beta}(\alpha, \beta)$ .

**Lemma 2.** *Let the random variable  $X$  be described by Jones' distribution  $G_F^{(B)}$  given in (2). Then, the random variable  $Y = F(X)$  has a beta  $\text{Beta}(\alpha, \beta)$  distribution.*

**Proof.** It can be easily shown from (2) that the distribution function  $F_Y$  of  $Y$  is given by

$$F_Y(y) = P(X \leq F^{-1}(y)) = I_y(\alpha, \beta),$$

which is the required result. ▲

The next proposition shows that Jones' distribution has maximum entropy in the class of all probability distributions specified by the constraints stated therein.

**Proposition 1.** *The density  $g_F^{(B)}$  of the random variable  $X$ , given by (1), is the unique solution of the optimization problem*

$$g_F^{(B)}(x) = \arg \max_h \mathcal{H}_{Sh}(h),$$

under the constraints

$$E_h[\ln F(X)] = \Psi(\alpha) - \Psi(\alpha + \beta), \tag{C1}$$

$$E_h[\ln(1 - F(X))] = \Psi(\beta) - \Psi(\alpha + \beta), \tag{C2}$$

$$E_h[\ln f(X)] = E_Y[\ln f(F^{-1}(Y))], \quad \text{where } Y \sim \text{Beta}(\alpha, \beta). \tag{C3}$$

**Proof.** Let  $h$  be a density satisfying the constraints (C1)–(C3). Now consider the well-known Kullback–Leibler divergence between  $h$  and  $g_F$  (see [16]) given by

$$D(h, g_F^{(B)}) = \int h(x) \ln \left( \frac{h(x)}{g_F^{(B)}(x)} \right) dx.$$

Then, following [17], we have

$$\begin{aligned} 0 \leq D(h, g_F^{(B)}) &= \int h(x) \ln h(x) dx - \int h(x) \ln g_F^{(B)}(x) dx \\ &= -\mathcal{H}_{Sh}(h) - \int h(x) \ln g_F^{(B)}(x) dx. \end{aligned} \tag{7}$$

Using the definition of the density  $g_F^{(B)}$ , as given in (1), and based on the constraints (C1)–(C3), we take

$$\int h(x) \ln g_F^{(B)}(x) dx = -\ln B(\alpha, \beta) + E_h[\ln f(X)] + (\alpha - 1)[\Psi(\alpha) - \Psi(\alpha + \beta)] + (\beta - 1)[\Psi(\beta) - \Psi(\alpha + \beta)]. \tag{8}$$

Now utilizing Lemma 1 and Eqs. (1) and (8), we conclude that

$$\int h(x) \ln g_F^{(B)}(x) dx = \int g_F^{(B)}(x) \ln g_F^{(B)}(x) dx = -\mathcal{H}_{Sh}(g_F^{(B)}). \tag{9}$$

Hence, from Eq. (7), we obtain

$$0 \leq -\mathcal{H}_{Sh}(h) + \mathcal{H}_{Sh}(g_F^{(B)}),$$

or equivalently

$$\mathcal{H}_{Sh}(h) \leq \mathcal{H}_{Sh}(g_F^{(B)}),$$

with equality if and only if  $D(h, g_F^{(B)}) = 0$ , i.e., if and only if  $h = g_F^{(B)}$ , a.e. This completes the proof of the proposition. ▲

The intermediate steps in the above proof in fact give the following explicit expression for the Shannon entropy of Jones' distribution in (1).

**Corollary 1.** *The Shannon entropy of Jones' distribution, with density in (1), is given by*

$$\mathcal{H}_{Sh}(g_F^{(B)}) = \ln B(\alpha, \beta) - (\alpha - 1)[\Psi(\alpha) - \Psi(\alpha + \beta)] - (\beta - 1)[\Psi(\beta) - \Psi(\alpha + \beta)] - E_Y[\ln f(F^{-1}(Y))],$$

where  $Y \sim \text{Beta}(\alpha, \beta)$ .

**Proof.** It is readily obtained from Eqs. (8) and (9) and constraint (C3). ▲

It is known that if  $F$  is symmetric, then the parameters  $\alpha$  and  $\beta$  alone control the degree of skewness. Also, it is evident that if the random variable  $X$  has density  $g_X = g_F^{(B)}(x; \alpha, \beta)$  given in (1), then  $Y = -X$  has density  $g_Y = g_F^{(B)}(-x; \beta, \alpha)$ . So, it can be easily shown in this case that the Shannon entropy remains invariant, i.e.,  $\mathcal{H}_{Sh}(g_X) = \mathcal{H}_{Sh}(g_Y)$ .

**Example 1.** As a first example, suppose the parent distribution is uniform in the interval  $(0, \theta)$ ,  $\theta > 0$ . Then,  $f(x) = 1/\theta$ ,  $0 < x < \theta$ , and  $F(x) = x/\theta$ ,  $0 < x < \theta$ . In this case, the beta density generated by the uniform distribution is given by

$$\frac{1}{B(\alpha, \beta)} \frac{1}{\theta} \left(\frac{x}{\theta}\right)^{\alpha-1} \left(1 - \frac{x}{\theta}\right)^{\beta-1}, \quad 0 < x < \theta, \alpha, \beta > 0.$$

It can be easily shown that

$$E_Y[\ln f(F^{-1}(Y))] = -\ln \theta,$$

which also leads to the Shannon entropy of the beta-uniform distribution by a simple application of Corollary 1. ▲

**Example 2.** As a second example, let us consider the beta-exponential distribution (see [5]) with density  $\frac{\lambda}{B(\alpha, \beta)} e^{-\beta\lambda x} (1 - e^{-\lambda x})^{\alpha-1}$ ,  $x > 0$ . This is a member of (1) when the parent density  $f$  and the cdf  $F$  are that of the exponential distribution with parameter  $\lambda > 0$ , i.e.,  $f(x) = \lambda e^{-\lambda x}$ ,  $x > 0$ , and  $F(x) = 1 - e^{-\lambda x}$ ,  $x > 0$ . Simple algebraic manipulations lead to

$$E_Y[\ln f(F^{-1}(Y))] = \ln \lambda + \Psi(\beta) - \Psi(\alpha + \beta),$$

and then to the Shannon entropy of the beta-exponential distribution by a simple application of [Corollary 1](#). It should be noted that the beta-exponential distribution extends the exponentiated exponential distribution of Gupta and Kundu [18]. ▲

**Example 3.** Consider the beta-logistic or log- $F$  distribution (see [6,1]) with density  $\frac{\eta}{B(\alpha, \beta)} \frac{e^{-\eta\beta x}}{(1+e^{-\eta x})^{\alpha+\beta}}$ ,  $x \in R$ ,  $\eta > 0$ . This is obtained from (1) with parent density and cdf  $f(x) = \frac{\eta e^{-\eta x}}{(1+e^{-\eta x})^2}$  and  $F(x) = \frac{1}{1+e^{-\eta x}}$  for  $x \in R$ ,  $\eta > 0$ , respectively. After some algebraic manipulations, it can be shown that

$$E_Y[\ln f(F^{-1}(Y))] = \ln \eta + \Psi(\alpha) + \Psi(\beta) - 2\Psi(\alpha + \beta),$$

and then to the Shannon entropy by an application of [Corollary 1](#). ▲

**Example 4.** Next, let us consider as parent distribution the Pareto with density and cdf as  $f(x) = \frac{k\theta^k}{x^{k+1}}$ ,  $x \geq \theta > 0$ , and  $F(x) = 1 - (\frac{\theta}{x})^k$ ,  $x \geq \theta > 0$  and  $k > 0$ , respectively. Then, the beta-Pareto distribution has density

$$\frac{1}{B(\alpha, \beta)} \frac{k\theta^{k\beta}}{x^{k\beta+1}} \left\{ 1 - \left( \frac{\theta}{x} \right)^k \right\}^{\alpha-1}, \quad x \geq \theta > 0,$$

and

$$E_Y[\ln f(F^{-1}(Y))] = \ln \frac{k}{\theta} + \frac{k+1}{k} [\Psi(\beta) - \Psi(\alpha + \beta)],$$

which can be used to obtain the Shannon entropy by an application of [Corollary 1](#). ▲

**Example 5.** Let us consider the power function distribution with pdf and cdf  $f(x) = k\theta^k x^{k-1}$ ,  $0 < x < 1/\theta$ , and  $F(x) = (\theta x)^k$ ,  $0 < x < 1/\theta$  and  $k > 0$  as the parent. In this case, the beta-power function distribution has density

$$\frac{1}{B(\alpha, \beta)} k\theta^{k\alpha} x^{k\alpha-1} \left\{ 1 - (\theta x)^k \right\}^{\beta-1}, \quad 0 < x < \frac{1}{\theta},$$

and

$$E_Y[\ln f(F^{-1}(Y))] = \ln(k\theta) + \frac{k-1}{k} [\Psi(\alpha) - \Psi(\alpha + \beta)]. \quad \blacktriangle$$

**Example 6.** The density and the cdf of half-logistic distribution is  $f(x) = \frac{2e^{-x}}{(1+e^{-x})^2}$ ,  $x > 0$ , and  $F(x) = \frac{1-e^{-x}}{1+e^{-x}}$ ,  $x > 0$ , respectively; see [19]. In this case, the beta-half-logistic distribution has density

$$\frac{2^\beta}{B(\alpha, \beta)} \frac{e^{-\beta x} (1 - e^{-x})^{\alpha-1}}{(1 + e^{-x})^{\alpha+\beta}}, \quad 0 < x < \infty,$$

and

$$\begin{aligned} E_Y[\ln f(F^{-1}(Y))] &= \ln 2 + \Psi(\beta) - \Psi(\alpha + \beta) + \frac{1}{B(\alpha, \beta)} \int_0^1 \ln(1+y) y^{\alpha-1} (1-y)^{\beta-1} dy \\ &= \ln 2 + \Psi(\beta) - \Psi(\alpha + \beta) + \sum_{\ell=1}^{\infty} (-1)^{\ell-1} \frac{1}{\ell} \frac{(\alpha + \ell - 1)^{(\ell)}}{(\alpha + \beta + \ell - 1)^{(\ell)}}, \end{aligned}$$

where  $a^{(\ell)} = a(a-1)\dots(a-\ell+1)$  is the  $\ell$ th downward factorial. ▲

**Example 7.** This final example considers the Weibull distribution with density and cdf as  $f(x) = \frac{c}{\theta^c} x^{c-1} e^{-(x/\theta)^c}$ ,  $x > 0$ , and  $F(x) = 1 - e^{-(x/\theta)^c}$  for  $x > 0$  and  $c, \theta > 0$ , respectively. Then, the beta-Weibull distribution, shown by Zografos [7] as the distribution which maximizes the Shannon

entropy, has its density as

$$\frac{c}{\theta^c B(\alpha, \beta)} x^{c-1} \exp \left\{ -\beta \left( \frac{x}{\theta} \right)^c \right\} \left\{ 1 - \exp \left( -\left( \frac{x}{\theta} \right)^c \right) \right\}^{\alpha-1}, \quad x > 0.$$

When  $\beta = 1$ , it is the Exponentiated-Weibull family introduced by Mudholkar and Srivastava [20]. For the beta-Weibull distribution, we also have

$$E_Y[\ln f(F^{-1}(Y))] = \ln \frac{c}{\theta} + [\Psi(\beta) - \Psi(\alpha + \beta)] + \frac{c-1}{c} E_Y[\ln(-\ln(1-Y))],$$

where  $Y \sim \text{Beta}(\alpha, \beta)$ . ▲

**Remark 1.** Looking at Corollary 1, we observe that the Shannon entropy of the beta-generated family of distributions in (1) is factored into two parts. The first part is related to the parameters  $\alpha$  and  $\beta$  of the beta distribution which is the first ingredient in the construction of the family in (1), while the second part is purely related to the “parent” distribution  $F$  which is the second ingredient in the construction of the family in (1). Moreover, all members of the family in (1) have the first part in common and they are discriminated between each other only by means of the second part, viz., the term  $E_Y[\ln f(F^{-1}(Y))]$ , where  $Y \sim \text{Beta}(\alpha, \beta)$ , which depends on the “parent” distribution of the model. If the distribution function of a random variable  $X$  belongs to the family in (1), then the maximum entropy principle, as already mentioned, suggests to choose the model in (1) which maximizes the Shannon entropy, as the most appropriate model to describe the unknown density function of  $X$ . Based on the observation above, the term  $E_Y[\ln f(F^{-1}(Y))]$ , where  $Y \sim \text{Beta}(\alpha, \beta)$ , plays the key role in specifying the particular member of (1) that maximizes the entropy  $\mathcal{H}_{Sh}(g_F^{(B)})$ , and so this term can be used to discriminate between the members of the family in (1). Section 5 concentrates on some inferential issues in this direction.

### 3. Generalized gamma-generated by parent distribution $F$

A family of univariate continuous distributions will be introduced in this section through a particular case of Stacy’s generalized gamma distribution, in the same spirit as Jones’ family defined through the beta distribution. Some of the properties of this family will be explored and a maximum entropy characterization will be obtained which will be exploited later in Section 5 in order to construct a statistical test for the discrimination between members of this family.

Consider a continuous distribution  $F$  with density  $f$ , and further Stacy’s generalized gamma density  $\frac{\gamma}{\Gamma(\delta)} u^{\gamma\delta-1} e^{-u^\gamma}$  for  $u > 0$  and positive values of the parameters  $\gamma$  and  $\delta$ . Based on this density, by replacing  $u$  by  $-\ln(1 - F(x))$ , we introduce the family with pdf

$$g_F^{(S)}(x; \gamma, \delta) = \frac{\gamma}{\Gamma(\delta)} \{-\ln(1 - F(x))\}^{\gamma\delta-1} e^{-\{-\ln(1-F(x))\}^\gamma} \frac{f(x)}{1 - F(x)}, \quad x \in R, \gamma, \delta > 0. \quad (10)$$

This is the generalized gamma (or Stacy)-generated distribution by parent  $F$ . If  $\gamma = 1$  in (10), it corresponds to the gamma-generated distribution by parent  $F$ . This family of distributions has its pdf as

$$\begin{aligned} g_F^{(G)}(x; \delta) &= \frac{1}{\Gamma(\delta)} \{-\ln(1 - F(x))\}^{\delta-1} e^{-\{-\ln(1-F(x))\}} \frac{f(x)}{1 - F(x)} \\ &= \frac{1}{\Gamma(\delta)} \{-\ln(1 - F(x))\}^{\delta-1} f(x), \quad x \in R, \delta > 0. \end{aligned} \quad (11)$$

**Remark 2.** Suppose  $X_{U(1)}, X_{U(2)}, \dots, X_{U(n)}, \dots$  are upper record values arising from a sequence of i.i.d. continuous random variables from a population with cdf  $G(x)$  and pdf  $g(x)$ . Then, the pdf of the  $n$ th upper record value,  $X_{U(n)}$ , is (see [21])

$$g_{X_{U(n)}}(x) = \frac{\{-\ln(1 - G(x))\}^{n-1}}{(n-1)!} g(x), \quad -\infty < x < \infty,$$

for  $n = 1, 2, \dots$  Upon converting the positive integral parameter  $n$  to a positive real parameter  $\delta$ , we readily obtain the family of densities with pdf

$$f_\delta(x) = \frac{\{-\ln(1 - G(x))\}^{\delta-1}}{\Gamma(\delta)}g(x), \quad -\infty < x < \infty, \delta > 0.$$

Observe that this is precisely the class of gamma-generated densities. Hence, just as Jones' family of distributions is generated by order statistics densities, the class in (10), for the case  $\gamma = 1$ , is generated by record value densities. This explicitly implies that the model in (11) is a direct record-analog.

In what follows, we will focus on the family of gamma-generated distributions with parent  $F$ , which is defined by (11). It can be easily shown that the distribution function of the family (11) is given by

$$G_F^{(G)}(x) = \frac{1}{\Gamma(\delta)}I_{-\ln(1-F(x))}(\delta), \quad \delta > 0,$$

where  $I_\omega(\delta) = \int_0^\omega e^{-z}z^{\delta-1}dz$  is the incomplete gamma function.

A logarithmic transformation of the parent distribution  $F$  transforms the random variable  $X$  with density (11) to a gamma distribution. That is, if  $X$  has a density of the form (11), then the random variable  $Z = -\ln(1 - F(X))$  has a gamma distribution  $G(\delta, 1)$  with density  $\frac{1}{\Gamma(\delta)}e^{-z}z^{\delta-1}, z > 0$ . The opposite is also true, that is, if  $Z$  has a gamma distribution  $G(\delta, 1)$  with density  $\frac{1}{\Gamma(\delta)}e^{-z}z^{\delta-1}, z > 0$ , then the random variable  $X = F^{-1}(1 - e^{-z})$  has a gamma distribution generated by the parent distribution  $F$ , with density (11). This helps to generate random numbers from the density  $g_F^{(G)}$  simply from observations from a gamma distribution  $G(\delta, 1), \delta > 0$ .

A general expression for the moments of the gamma density generated by the distribution  $F$  is presented now. Indeed, the  $\nu$ th moment of  $g_F^{(G)}$  defined by (11) is

$$E_{g_F^{(G)}}[X^\nu] = \int_{-\infty}^{+\infty} x^\nu \frac{1}{\Gamma(\delta)} \{-\ln(1 - F(x))\}^{\delta-1} f(x)dx.$$

Using the transformation  $y = F(x)$ , we obtain

$$E_{g_F^{(G)}}[X^\nu] = \int_0^1 [F^{-1}(y)]^\nu \frac{1}{\Gamma(\delta)} \{-\ln(1 - F(x))\}^{\delta-1} dy, \tag{12}$$

a formula quite similar to that of Arnold et al. [21, p. 30]. It is possible, in some cases, to obtain an analytic form for the moments  $E_{g_F^{(G)}}[X^\nu]$  of  $g_F^{(G)}$ , given by (12), and these are given in the examples of this section for some specific choices of  $g_F^{(G)}$ .

The next result presents the expected values of some special transformations of the parent density and distribution functions  $f$  and  $F$ , which will be used later for the maximum entropy characterization of  $g_F^{(G)}$ .

**Lemma 3.** For the gamma density  $g_F^{(G)}$ , generated by the parent distribution  $F$ , we have

- (a)  $E_{g_F^{(G)}}\{\ln[-\ln(1 - F(X))]\} = \Psi(\delta)$ ,
- (b)  $E_{g_F^{(G)}}[\ln f(X)] = E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\}$ ,

where  $Z$  has gamma distribution  $G(\delta, 1)$  with density  $\frac{1}{\Gamma(\delta)}e^{-z}z^{\delta-1}, z > 0$ , and  $\Psi(\cdot)$  denotes the digamma function.

**Proof.** Both parts of the lemma are proved by using first the transformation  $y = F(x)$  in the integrals and then the transformation  $z = -\ln(1 - y)$ . We will outline only the proof of Part (a). Using the above transformations,

$$E_{g_F^{(G)}}\{\ln[-\ln(1 - F(X))]\} = \frac{1}{\Gamma(\delta)} \int_0^\infty (\ln z)e^{-z}z^{\delta-1}dz.$$

The proof is completed by observing that the last integral is expressed as the derivative  $\frac{d}{d\delta}\Gamma(\delta)$ .  $\blacktriangle$



**Remark 3.** A series expansion of integrals of the form  $\int_0^\infty (\ln x)e^{-x}x^{\delta-1}dx$  is given in Lemma 1 by Baratpour et al. [22], for  $\delta$  a natural number. Similar formulas for the expected values of the lemma have been derived in Eqs. (7) and (8) of Baratpour et al. [22] for the case of natural  $\delta$ .

The next proposition shows that the gamma distribution generated by  $F$  has maximum entropy in the class of all probability distributions specified by the constraints (G1) and (G2) stated in the proposition. The proof follows exactly along the same lines as those in the proof of the similar proposition for Jones' family, and is therefore omitted.

**Proposition 2.** The density  $g_F^{(G)}$  of the random variable  $X$ , given by (11), is the unique solution of the optimization problem

$$g_F^{(G)}(x) = \arg \max_h \mathcal{H}_{Sh}(h),$$

under the constraints

$$E_h\{\ln[-\ln(1 - F(X))]\} = \Psi(\delta), \tag{G1}$$

and

$$E_h\{\ln f(X)\} = E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\}, \tag{G2}$$

where  $Z$  is a gamma random variable with density  $\frac{1}{\Gamma(\delta)}e^{-z}z^{\delta-1}, z > 0$ .

Direct use of the constraints (G1) and (G2) leads to the Shannon entropy of the gamma density  $g_F^{(G)}$ , generated by  $F$ .

**Corollary 2.** The expression of Shannon entropy of gamma density  $g_F^{(G)}$  is given by

$$\mathcal{H}_{Sh}(g_F^{(G)}) = \ln \Gamma(\delta) - (\delta - 1)\Psi(\delta) - E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\},$$

with  $Z$  being the gamma random variable of the above proposition.

A similar expression for Shannon entropy of  $g_F^{(G)}$  has been derived in [22] for the case of natural  $\delta$ . Based on the above corollary, all the members of the family (11) are discriminated between each other by means of the expected value  $E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\}$ , which depends on the “parent” distribution  $F$  of the model. Hence, in a manner similar to the Jones' family, the term  $E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\}$ , where  $Z$  has gamma distribution  $G(\delta, 1)$  with density  $\frac{1}{\Gamma(\delta)}e^{-z}z^{\delta-1}, z > 0$ , now plays the key role in discriminating between the members of the family (11). Section 5 concentrates on some specific problems in this direction, and in particular to the derivation of a statistical test for discriminating between members of the family  $g_F^{(G)}$ .

It is possible, in some cases, to obtain an analytic form for the expected value of Part (b) of Lemma 3, as shown in the following examples. The same is also true for the moments  $E_{g_F^{(G)}}[X^\nu]$  of  $g_F^{(G)}$ , which are given by (12).

**Example 8.** Suppose that the parent distribution is uniform in the interval  $(0, \theta), \theta > 0$ . Then,  $f(x) = 1/\theta, 0 < x < \theta$ , and  $F(x) = x/\theta$ . In this case,  $g_F^{(G)}$  is the gamma density generated by the uniform distribution, with density function

$$g_F^{(G)}(x; \delta, \theta) = \frac{1}{\theta\Gamma(\delta)}\{-\ln(1 - (x/\theta))\}^{\delta-1}, \quad x \in R, \delta, \theta > 0.$$

It can be easily shown that for  $Z$ , a gamma random variable with density  $\frac{1}{\Gamma(\delta)}e^{-z}z^{\delta-1}, z > 0$ ,

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = -\ln \theta.$$

This formula can be used to obtain the Shannon entropy by an application of [Corollary 2](#). The moments of the above density  $g_F^{(G)}$  are given by

$$E_{g_F^{(G)}}[X^v] = \theta^v \sum_{r=0}^v \binom{v}{r} \frac{(-1)^r}{(1+r)^\delta}, \quad v = 1, 2, \dots,$$

for the case of natural  $\delta > 1$ . ▲

**Example 9.** Consider the gamma-exponential distribution with density

$$g_F^{(G)}(x; \delta, \lambda) = \frac{\lambda^\delta}{\Gamma(\delta)} x^{\delta-1} e^{-\lambda x}, \quad x \in R,$$

for  $\delta, \lambda > 0$ . It is simply the gamma distribution with shape parameter  $\delta$  and scale parameter  $\lambda$ . This is obtained from [\(11\)](#) if the parent density  $f$  and the parent distribution function  $F$  are those of the exponential distribution with parameter  $\lambda > 0$ . In this case,  $f(x) = \lambda e^{-\lambda x}$  and  $F(x) = 1 - e^{-\lambda x}$ ,  $x > 0$ . Simple algebraic manipulations then lead to

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln \lambda - \delta,$$

for  $Z$ , a gamma random variable. This formula and [Corollary 2](#) can be used to obtain the Shannon entropy of the gamma-exponential distribution. The moments of the gamma-exponential distribution  $g_F^{(G)}$  are given by

$$E_{g_F^{(G)}}[X^v] = \frac{\Gamma(v + \delta)}{\lambda^v \Gamma(\delta)}, \quad v = 1, 2, \dots \quad \blacktriangle$$

**Example 10.** The gamma-Pareto distribution is obtained from [\(11\)](#) for  $f(x) = \frac{k\theta^k}{x^{k+1}}$ ,  $x \geq \theta > 0$ , and  $F(x) = 1 - \left(\frac{\theta}{x}\right)^k$ ,  $k > 0$ , and has density as

$$g_F^{(G)}(x; \delta, k, \theta) = \frac{(-1)^{\delta-1} k^\delta \theta^k}{\Gamma(\delta) x^{k+1}} \left(\ln \frac{\theta}{x}\right)^{\delta-1}, \quad x \in R,$$

for  $\delta, k, \theta > 0$ . Algebraic manipulation then leads to

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln \frac{k}{\theta} - \frac{k+1}{k} \delta,$$

which can be used to obtain the Shannon entropy of the gamma-Pareto distribution. Based on [\(12\)](#), the moments of this distribution are

$$E_{g_F^{(G)}}[X^v] = \theta^v \left(\frac{k}{k-v}\right)^\delta, \quad \text{for } v = 1, 2, \dots \quad \blacktriangle$$

**Example 11.** Consider the power function distribution with density and distribution function as  $f(x) = k\theta^k x^{k-1}$ ,  $0 < x < 1/\theta$ , and  $F(x) = (\theta x)^k$ , for  $k > 0$ . Then, the gamma density generated by the power function distribution is

$$g_F^{(G)}(x; \delta, k, \theta) = \frac{k\theta^k}{\Gamma(\delta)} x^{k-1} \{-\ln(1 - (\theta x)^k)\}^{\delta-1}, \quad x \in R, \delta, k, \theta > 0.$$

Moreover, by using the transformation  $y = e^{-z}$ ,  $z > 0$ , we obtain

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln(k\theta) + \frac{(-1)^{\delta-1}(k-1)}{k\Gamma(\delta)} \int_0^1 (\ln y)^{\delta-1} \ln(1-y) dy.$$

Taking into account that  $-\ln(1 - y) = y + (y^2/2) + (y^3/3) + \dots$ , for  $0 < y < 1$ , after some algebraic manipulations, we obtain for a natural  $\delta > 1$  that

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln(k\theta) - \frac{k - 1}{k} \sum_{m=1}^{\infty} \frac{1}{m(m + 1)^\delta}.$$

This expected value also helps to derive, by means of Corollary 2, the Shannon entropy of the gamma density generated by the power function distribution. For a natural  $\delta > 1$  and  $\nu/k$  a natural number, the moments of  $g_F^{(G)}$  are given by

$$E_{g_F^{(G)}}[X^\nu] = \frac{1}{\theta^\nu} \sum_{r=0}^{\nu/k} \binom{\nu/k}{r} \frac{(-1)^r}{(1 + r)^\delta}, \quad \nu = 1, 2, \dots \quad \blacktriangle$$

**Example 12.** Consider the gamma-logistic distribution with density

$$g_F^{(G)}(x; \delta, \eta) = \frac{1}{\Gamma(\delta)} \left\{ -\ln \frac{e^{-\eta x}}{1 + e^{-\eta x}} \right\}^{\delta-1} \frac{\eta e^{-\eta x}}{(1 + e^{-\eta x})^2}, \quad x \in R,$$

for  $\delta, \eta > 0$ . This is obtained from (11) if the parent density  $f$  and the parent distribution function  $F$  are that of the logistic distribution with parameter  $\eta > 0$ . In this case  $f(x) = \frac{\eta e^{-\eta x}}{(1 + e^{-\eta x})^2}$  and  $F(x) = \frac{1}{1 + e^{-\eta x}}$  for  $x \in R, \eta > 0$ , respectively. Algebraic manipulations then lead to

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln \eta - \delta - \sum_{m=1}^{\infty} \frac{1}{m(m + 1)^\delta}$$

if  $\delta > 1$  is a natural number. Based on (12), the moments of the gamma-logistic distribution are

$$E_{g_F^{(G)}}[X^\nu] = \frac{(-1)^{\nu+\delta-1}}{\eta^\nu \Gamma(\delta)} \int_0^1 (\ln((1 - y)/y))^\nu (\ln(1 - y))^{\delta-1} dy,$$

and they cannot be simplified any further.  $\blacktriangle$

**Example 13.** The gamma-half-logistic distribution is obtained from (11) for  $f(x) = \frac{2e^{-x}}{(1 + e^{-x})^2}, x > 0$ , and  $F(x) = \frac{1 - e^{-x}}{1 + e^{-x}}, x > 0$ , and has density function as

$$g_F^{(G)}(x; \delta) = \frac{1}{\Gamma(\delta)} \left\{ -\ln \frac{2e^{-x}}{1 + e^{-x}} \right\}^{\delta-1} \frac{2e^{-x}}{(1 + e^{-x})^2}, \quad x \in R,$$

and  $\delta > 0$ . If  $\delta > 0$  is a natural number, then

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln(1/2) - \delta + \sum_{m=1}^{\infty} \sum_{r=0}^m \binom{m}{r} \frac{(-1)^{m+r+1}}{m(1 + r)^\delta},$$

which can be used to obtain the Shannon entropy of the gamma-half-logistic distribution, in view of Corollary 2. Based on (12), the moments of this distribution are

$$E_{g_F^{(G)}}[X^\nu] = \frac{(-1)^{\delta-1}}{\Gamma(\delta)} \int_0^1 [\ln((1 + y)/(1 - y))]^\nu (\ln(1 - y))^{\delta-1} dy,$$

and they cannot be simplified any further.  $\blacktriangle$

**Example 14.** The final example concentrates on the gamma-Weibull distribution with density

$$g_F^{(G)}(x; \delta, \theta, c) = \frac{c}{\theta \Gamma(\delta)} \left(\frac{x}{\theta}\right)^{c\delta-1} e^{-(x/\theta)^c}, \quad x \in R,$$

$\delta, c, \theta > 0$ , which is obtained from (11) with parent density and distribution functions  $f(x) = \frac{c}{\theta^c} x^{c-1} e^{-(x/\theta)^c}$ ,  $x > 0$  and  $F(x) = 1 - e^{-(x/\theta)^c}$ , for  $c, \theta > 0$ . For this distribution, we have

$$E_Z\{\ln f(F^{-1}(1 - e^{-Z}))\} = \ln \frac{c}{\theta} - \delta + \frac{c-1}{c} \Psi(\delta).$$

The moments of the gamma-Weibull distribution are

$$E_{g_F^{(G)}}[X^\nu] = \frac{\theta^\nu}{\Gamma(\delta)} \Gamma\left(\frac{\nu}{c} + \delta\right), \quad \nu = 1, 2, \dots$$

The above formulas for  $E_{g_F^{(S)}}[X^\nu]$ ,  $\nu = 1, 2, \dots$ , can be used to obtain the first four moments of all these distributions, thereby giving formulas for the skewness and kurtosis of the gamma-generated densities. ▲

#### 4. An alternative method of moments

An alternative method of moments will be developed in this section for the estimation of the parameters of the beta and gamma families of skewed distributions which are generated by a parametric parent distribution  $F$ . For this purpose, let us suppose that the parent distribution function  $F$  of the skewed models (1) and (11) involves a  $p$ -dimensional real parameter  $\theta$ . We will denote by  $F_\theta$  the parent distribution and by  $g_{F_\theta}^{(B)}$  and  $g_{F_\theta}^{(G)}$  the respective families (1) and (11). In order to estimate the parameters  $\theta$  by using the method of moments, it is necessary to have the moments of  $g_{F_\theta}^{(B)}$  and  $g_{F_\theta}^{(G)}$  in a closed form. It is hard in general to obtain in an analytic form the central moments  $E_{g_{F_\theta}^{(B)}}(X^\nu)$  and  $E_{g_{F_\theta}^{(G)}}(X^\nu)$ ,  $\nu = 1, 2, \dots$ , and is in fact impossible in most of the cases. On the other hand, we observe in Lemmas 1 and 3 that we can obtain a simple analytic form for the expected values of some suitable functions of the parent distribution  $F_\theta$ . This fact motivates us to look at the analytic forms of powers of suitable functions of the parent distribution  $F_\theta$ . These expected values, with respect to  $g_{F_\theta}^{(B)}$  and  $g_{F_\theta}^{(G)}$ , are stated below in Lemmas 4 and 5.

**Lemma 4.** For  $\nu = 1, 2, \dots$ ,

- (a)  $E_{g_{F_\theta}^{(B)}}[(F_\theta(X))^\nu] = \frac{\Gamma(\alpha+\nu)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+\nu)},$
- (b)  $E_{g_{F_\theta}^{(B)}}[(\ln F_\theta(X))^\nu] = \frac{1}{B(\alpha, \beta)} \frac{d^\nu}{d\alpha^\nu} B(\alpha, \beta),$
- (c)  $E_{g_{F_\theta}^{(B)}}[(1 - F_\theta(X))^\nu] = \frac{\Gamma(\alpha+\beta)\Gamma(\beta+\nu)}{\Gamma(\beta)\Gamma(\alpha+\beta+\nu)},$
- (d)  $E_{g_{F_\theta}^{(B)}}[(\ln(1 - F_\theta(X)))^\nu] = \frac{1}{B(\alpha, \beta)} \frac{d^\nu}{d\beta^\nu} B(\alpha, \beta).$

The lemma can be proved easily by using the transformation  $y = F_\theta(x)$ , and is therefore omitted.

This lemma is essential for developing an alternative method of moments for estimating the parameter  $\theta$  of the family  $g_{F_\theta}^{(B)}$ . In this context, consider a random sample  $X_1, X_2, \dots, X_n$  from  $g_{F_\theta}^{(B)}$  with parent distribution  $F_\theta$ . Suppose, without any loss of generality, that the parameters  $\alpha$  and  $\beta$  of  $g_{F_\theta}^{(B)}$  given by (1), are known. Then, instead of using the sample version of the moments  $E_{g_{F_\theta}^{(B)}}(X^\nu)$  to estimate the parameter  $\theta$ , we can use in a similar manner the sample versions of the expected values in the above lemma. In this direction, the estimator of the  $p$ -dimensional parameter  $\theta$ , denoted by  $\hat{\theta}^{AME}$ , can be obtained as the solution of the system of equations

$$\frac{1}{n} \sum_{i=1}^n [F_\theta(X_i)]^\nu = \frac{\Gamma(\alpha + \nu)\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\alpha + \beta + \nu)}, \quad \text{for } \nu = 1, \dots, p.$$

Similarly, solution of the system of equations

$$\frac{1}{n} \sum_{i=1}^n [(\ln(1 - F_\theta(X_i)))^\nu] = \frac{1}{B(\alpha, \beta)} \frac{d^\nu}{d\beta^\nu} B(\alpha, \beta), \quad \text{for } \nu = 1, \dots, p, \tag{13}$$

may be used for estimation purpose as well. Observe that the above lemma gives a great flexibility to construct suitable equations depending on the specific form of the parent distribution  $F_\theta$ .

**Example 15.** As a first example, suppose that the parent distribution is uniform in the interval  $(0, \theta)$ , with cdf  $F(x) = x/\theta$  for  $0 < x < \theta$ . In this case, the beta density generated by the uniform distribution is given by

$$\frac{1}{B(\alpha, \beta)} \frac{1}{\theta} \left(\frac{x}{\theta}\right)^{\alpha-1} \left(1 - \frac{x}{\theta}\right)^{\beta-1}, \quad x \in R, \alpha, \beta > 0.$$

In order to obtain the alternative moment estimator  $\hat{\theta}^{AME}$  of  $\theta$ , let us consider the equation  $\frac{1}{n} \sum_{i=1}^n [F_\theta(X_i)]^\nu = \frac{\Gamma(\alpha+\nu)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+\nu)}$  for  $\nu = 1$ , since  $\theta$  is univariate. Then,

$$\frac{1}{n} \sum_{i=1}^n F_\theta(X_i) = \frac{\bar{X}}{\theta},$$

and the alternative moment estimator  $\hat{\theta}^{AME}$  is

$$\hat{\theta}^{AME} = \bar{X} \frac{\Gamma(\alpha)\Gamma(\alpha + \beta + 1)}{\Gamma(\alpha + 1)\Gamma(\alpha + \beta)}.$$

It can be easily shown that the likelihood equation is

$$-\frac{n(\alpha + \beta - 1)}{\theta} + (\beta - 1) \sum_{i=1}^n \frac{1}{\theta - x_i} = 0,$$

which does not lead to an analytic form for the maximum likelihood estimator of  $\theta$ . ▲

**Example 16.** Consider the beta-exponential distribution obtained from the parent distribution  $F_\theta(X) = 1 - e^{-\theta x}$ ,  $\theta > 0$ . It seems appropriate to use, in this case, Eq. (13), with  $\nu = 1$ , because the parameter  $\theta$  is univariate. We then have

$$\frac{1}{n} \sum_{i=1}^n [\ln(1 - F_\theta(X_i))] = -\theta \bar{X}.$$

Moreover,  $\frac{1}{B(\alpha, \beta)} \frac{d}{d\beta} B(\alpha, \beta) = \Psi(\beta) - \Psi(\alpha + \beta)$ . Hence, by using Eq. (13), with  $\nu = 1$ , we obtain

$$\hat{\theta}^{AME} = \frac{1}{\bar{X}} [\Psi(\alpha + \beta) - \Psi(\beta)].$$

Based on the results of Nadarajah and Kotz [5],  $\hat{\theta}^{AME}$  coincides with the maximum likelihood estimator and the classic moment estimator of  $\theta$ . ▲

**Example 17.** The parent distribution of the beta Pareto is  $F_\theta(X) = 1 - \left(\frac{\theta}{x}\right)^k$ ,  $\theta > 0, k > 0$ . Based once again on Eq. (13) and taking into account that

$$\begin{aligned} \frac{1}{B(\alpha, \beta)} \frac{d}{d\beta} B(\alpha, \beta) &= \Psi(\beta) - \Psi(\alpha + \beta), \\ \frac{1}{B(\alpha, \beta)} \frac{d^2}{d\beta^2} B(\alpha, \beta) &= [\Psi(\beta) - \Psi(\alpha + \beta)]^2 + \Psi'(\beta) - \Psi'(\alpha + \beta), \end{aligned}$$

the alternative moment estimators  $\hat{\theta}^{AME}$  and  $\hat{k}^{AME}$  of the parameters  $\theta$  and  $k$ , respectively, are obtained as the solution of the system of equations

$$k \ln \theta - \frac{1}{n} \sum_{i=1}^n \ln X_i = \Psi(\beta) - \Psi(\alpha + \beta),$$

$$\frac{1}{n} \sum_{i=1}^n k^2 (\ln \theta - \ln X_i)^2 = [\Psi(\beta) - \Psi(\alpha + \beta)]^2 + \Psi'(\beta) - \Psi'(\alpha + \beta).$$

The respective likelihood equations for the parameters  $\theta$  and  $k$  of the beta-Pareto distribution are

$$\frac{nk\beta}{\theta} - k(\alpha - 1)\theta^{k-1} \sum_{i=1}^n \frac{1}{x_i^k - \theta^k} = 0,$$

$$\frac{n}{k} + n\beta \ln \theta - \beta \sum_{i=1}^n \ln x_i - (\alpha - 1)\theta^k \sum_{i=1}^n \frac{\ln \theta - \ln x_i}{x_i^k - \theta^k} = 0.$$

The equations that lead to the alternative moment estimators of the parameters are simpler than the corresponding likelihood equations. ▲

**Example 18.** In a similar manner, the system of equations which leads to the alternative moment estimators  $\hat{\theta}^{AME}$  and  $\hat{c}^{AME}$  of the parameters  $\theta$  and  $c$  of the beta-Weibull distribution with parent distribution  $F_\theta(x) = 1 - e^{-(x/\theta)^c}$ , for  $c, \theta > 0$ , are

$$\frac{1}{n\theta^c} \sum_{i=1}^n \ln X_i^c = \Psi(\alpha + \beta) - \Psi(\beta),$$

$$\frac{1}{n\theta^{2c}} \sum_{i=1}^n \ln X_i^{2c} [\Psi(\beta) - \Psi(\alpha + \beta)]^2 + \Psi'(\beta) - \Psi'(\alpha + \beta).$$

It should be noted that the classical method of moments is not applicable in this case since the moments of the beta-Weibull distribution are, in general, not available in a closed form; see [7]. ▲

Let us now focus on the gamma skewed density  $g_{F_\theta}^{(G)}$  generated by the parametric parent distribution  $F_\theta$ . An alternative method of moments can be developed in a similar manner for  $g_{F_\theta}^{(G)}$ . The next lemma is key for this method of estimation.

**Lemma 5.** For  $\nu = 1, 2, \dots$ ,

- (a)  $E_{g_{F_\theta}^{(G)}}[(F_\theta(X))^\nu] = \sum_{r=0}^\nu \binom{\nu}{r} \frac{(-1)^r}{(1+r)^\delta}$  for a natural  $\delta > 1$ ,
- (b)  $E_{g_{F_\theta}^{(G)}}[(\ln F_\theta(X))^\nu] = \frac{(-1)^{\delta-1}}{\Gamma(\delta)} \frac{\partial^{\nu+\delta-1}}{\partial \alpha^\nu \partial \beta^{\delta-1}} B(\alpha, \beta) \Big|_{\alpha=\beta=1}$  for a natural  $\delta > 1$ ,
- (c)  $E_{g_{F_\theta}^{(G)}}[(1 - F_\theta(X))^\nu] = \frac{1}{(\nu+1)^\delta}$  for a natural  $\delta > 1$ ,
- (d)  $E_{g_{F_\theta}^{(G)}}[(\ln(1 - F_\theta(X)))^\nu] = \frac{(-1)^\nu}{\Gamma(\delta)} \Gamma(\nu + \delta)$  for a natural  $\delta > 0$ .

**Proof.** The lemma is proved by using suitable transformations. We will outline the proof for Parts (b) and (d). First, for the proof of Part (b), using the transformation  $y = F_\theta(x)$ , we get

$$E_{g_{F_\theta}^{(G)}}[(\ln F_\theta(X))^\nu] = \frac{(-1)^{\delta-1}}{\Gamma(\delta)} \int_0^1 (\ln y)^\nu ((\ln(1 - y))^{\delta-1} dy.$$

For a natural  $\delta > 1$  and  $B(\alpha, \beta) = \int_0^1 y^{\alpha-1} (1 - y)^{\beta-1} dy$ , it is easy to see that the last integral is equal to  $\frac{\partial^{\nu+\delta-1}}{\partial \alpha^\nu \partial \beta^{\delta-1}} B(\alpha, \beta) \Big|_{\alpha=\beta=1}$ , which completes the proof of Part (b). For the proof of Part (d), we first use the transformation  $y = F_\theta(x)$  and then the transformation  $z = -\ln(1 - y)$ , to obtain

$$E_{g_{F_\theta}^{(G)}}[(\ln(1 - F_\theta(X)))^\nu] = \frac{(-1)^\nu}{\Gamma(\delta)} \int_0^\infty z^{(\nu+\delta)-1} e^{-z} dz,$$

which completes the proof. ▲

In order to present the alternative method of moments for  $g_{F_\theta}^{(G)}$ , consider a random sample  $X_1, X_2, \dots, X_n$  from  $g_{F_\theta}^{(G)}$  with parent distribution  $F_\theta$ . Suppose the parameter  $\delta$  of  $g_{F_\theta}^{(G)}$ , given by (11), is known. Then, based on the above lemma, an alternative moment estimator of the  $m$ -dimensional parameter  $\theta$ , denoted by  $\hat{\theta}^{AME}$ , can be obtained as the solution of the system of equations

$$\frac{1}{n} \sum_{i=1}^n [(1 - F_\theta(X_i))^v] = \frac{1}{(v + 1)^\delta}, \quad \text{for } v = 1, \dots, m.$$

All other cases of Lemma 5 can be used in a similar manner.

**Example 19.** Let the parent distribution  $F_\theta$  be uniform in the interval  $(0, \theta)$ ,  $\theta > 0$ . Then,  $F_\theta(x) = x/\theta$  and the alternative moment estimator  $\hat{\theta}^{AME}$  of  $\theta$ , is obtained as the solution of the equation  $\frac{1}{n} \sum_{i=1}^n F_\theta(X_i) = \sum_{r=0}^1 \binom{1}{r} \frac{(-1)^r}{(1+r)^\delta}$ , which is given by

$$\hat{\theta}^{AME} = \bar{X} \frac{2^\delta}{2^\delta - 1}.$$

This alternative moment estimator coincides with the classical moment estimator of the gamma-uniform distribution since the first moment of this distribution is  $\theta \left(1 - \frac{1}{2^\delta}\right)$ . If the parameter  $\delta$  is also unknown, then the alternative moment estimators of  $\theta$  and  $\delta$  can be obtained as the solution of the equations

$$\theta = \bar{X} \frac{2^\delta}{2^\delta - 1} \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \frac{X_i^2}{\theta^2} = 1 - \frac{1}{2^{\delta-1}} + \frac{1}{3^\delta}. \quad \blacktriangle$$

**Example 20.** Consider the gamma-exponential distribution. In this case, the parent distribution is the exponential distribution  $F_\theta(x) = 1 - e^{-\theta x}$ ,  $\theta > 0$ . Then, using Part (d) of Lemma 5, the alternative moment estimator  $\hat{\theta}^{AME}$  of  $\theta$  is obtained as the solution of the equation  $\frac{1}{n} \sum_{i=1}^n \ln(1 - F_\theta(X_i)) = \frac{(-1)}{\Gamma(\delta)} \Gamma(\delta + 1)$ , which yields

$$\hat{\theta}^{AME} = \frac{\delta}{\bar{X}}.$$

This coincides with the moment estimator of  $\theta$ . Taking into account that the gamma-exponential distribution is the gamma distribution  $G(\delta, \frac{1}{\theta})$ , we conclude that  $\hat{\theta}^{AME} = \delta/\bar{X}$  coincides also with the maximum likelihood estimator of  $\theta$ . If the parameter  $\delta$  of the gamma-exponential distribution is also unknown, then the alternative moment estimators of  $\theta$  and  $\delta$  are given by

$$\hat{\theta}^{AME} = \frac{\hat{\delta}^{AME}}{\bar{X}} \quad \text{and} \quad \hat{\delta}^{AME} = \left( \frac{1}{n\bar{X}^2} \sum_{i=1}^n X_i^2 - 1 \right)^{-1}.$$

Note that in the gamma-exponential distribution, the maximum likelihood estimator of  $\delta$  is not available in a closed form.  $\blacktriangle$

**Example 21.** Let the parent distribution be Pareto. In this case, the distribution function is  $F_\theta(x) = 1 - (\theta/x)^k$ , which involves two parameters  $\theta$  and  $k$ . Using again Part (d) of Lemma 5, the alternative moment estimators of the parameters  $\theta$  and  $k$  can be obtained as the solution of the system of equations

$$k \ln \theta - \frac{k}{n} \sum_{i=1}^n \ln X_i = -\delta,$$

$$\frac{1}{n} \sum_{i=1}^n k^2 (\ln X_i - \ln \theta)^2 = \delta(\delta + 1).$$

The likelihood equations leading to the maximum likelihood estimators of  $\theta$  and  $k$  are

$$k \ln \theta - \frac{k}{n} \sum_{i=1}^n \ln X_i = -\delta,$$

$$\frac{nk}{\theta} + (\delta - 1) \sum_{i=1}^n \frac{1}{\theta(\ln \theta - \ln X_i)} = 0. \quad \blacktriangle$$

**Example 22.** Now consider the gamma-Weibull distribution with distribution function  $F(x) = 1 - e^{-(x/\theta)^c}$  for  $c, \theta > 0$ . For known  $\delta$ , the alternative moment estimators of  $c$  and  $\theta$  are obtained as the solution of the system of equations

$$\frac{1}{n\theta^c} \sum_{i=1}^n X_i^c = \delta,$$

$$\frac{1}{n\theta^{2c}} \sum_{i=1}^n X_i^{2c} = \delta(\delta + 1). \quad \blacktriangle$$

### 5. Discriminating between members of beta and gamma distributions generated by the distribution $F$

The problem of testing whether some given observations can be considered as coming from one of two probability distributions is an old problem in statistics; see [23] and the references contained therein. In this framework, consider a random sample  $X_1, X_2, \dots, X_n$  of size  $n$  from the family of Jones' distributions with density (1). Our interest is to identify the specific model of (1) that is most appropriate to describe the data  $X_1, X_2, \dots, X_n$ . We therefore need a way to discriminate between the models of the family in (1). In the spirit of the maximum entropy principle, the most appropriate model to describe  $X_1, X_2, \dots, X_n$  is the model with “parent” distribution function  $F$  with the corresponding Shannon entropy  $\mathcal{H}_{Sh}(g_F^{(B)})$  as large as possible. Hence, between two candidate “parent” models  $F_1$  and  $F_2$ , with respective densities  $f_1$  and  $f_2$ , we have to decide in favour of one of them on the basis of the difference  $\mathcal{H}_{Sh}(g_{F_1}^{(B)}) - \mathcal{H}_{Sh}(g_{F_2}^{(B)})$ .

It is easy to see, in view of Corollary 1, that

$$\mathcal{D}_{1,2}^{(B)} \triangleq \mathcal{H}_{Sh}(g_{F_1}^{(B)}) - \mathcal{H}_{Sh}(g_{F_2}^{(B)}) = E_Y \left\{ \ln \frac{f_2(F_2^{-1}(Y))}{f_1(F_1^{-1}(Y))} \right\}, \quad (14)$$

with  $Y \sim Beta(\alpha, \beta)$ , and if we will use the respective quantile density functions  $q_i(u) = 1/f_i(F_i^{-1}(u))$ ,  $i = 1, 2$  and  $0 < u < 1$ , an equivalent form is given by

$$\mathcal{D}_{1,2}^{(B)} = \mathcal{H}_{Sh}(g_{F_1}^{(B)}) - \mathcal{H}_{Sh}(g_{F_2}^{(B)}) = E_Y \left\{ \ln \frac{q_1(Y)}{q_2(Y)} \right\}, \quad \text{with } Y \sim Beta(\alpha, \beta). \quad (15)$$

Large values of  $\mathcal{D}_{1,2}^{(B)}$  support the “parent” model  $F_1$ , while small values of  $\mathcal{D}_{1,2}^{(B)}$  are in favour of  $F_2$ .

The difference  $\mathcal{D}_{1,2}^{(B)}$  can be estimated as follows. Using the random sample  $X_1, X_2, \dots, X_n$ , we can obtain the random samples  $Y_1^{(j)}, Y_2^{(j)}, \dots, Y_n^{(j)}$ , with  $Y_i^{(j)} = F_j(X_i)$ ,  $i = 1, \dots, n$ , and  $F_j$  a “parent” model, for  $j = 1, 2$ . Then, based on Lemma 2,  $Y_1^{(j)}, Y_2^{(j)}, \dots, Y_n^{(j)}$ ,  $j = 1, 2$ , is a random sample from a beta distribution  $Beta(\alpha, \beta)$  with parameters  $\alpha, \beta$ . Based on (14) and (15), the sample analogs of  $\mathcal{D}_{1,2}^{(B)}$  are

$$\widehat{\mathcal{D}}_{1,2}^{(B)} = \frac{1}{n} \sum_{i=1}^n \ln \frac{f_2(F_2^{-1}(Y_i^{(2)}))}{f_1(F_1^{-1}(Y_i^{(1)}))} \quad (16)$$



and

$$\widehat{\mathcal{D}}_{1,2}^{(B)} = \frac{1}{n} \sum_{i=1}^n \ln \frac{q_1 \left( Y_i^{(1)} \right)}{q_2 \left( Y_i^{(2)} \right)}. \tag{17}$$

This procedure can be used effectively to discriminate between two members of (1) that are based on “parent” distributions  $F_1$  and  $F_2$ . In this context, in order to decide between

$H_0$  : “parent” distribution is  $F_2$  and  $H_1$  : “parent” distribution is  $F_1$ ,

we can exploit the statistic  $\widehat{\mathcal{D}}_{1,2}^{(B)}$  and the null hypothesis is rejected at a level  $\alpha$  if

$$\widehat{\mathcal{D}}_{1,2}^{(B)} \geq d_\alpha,$$

where  $d_\alpha$  is the upper  $100 \times \alpha\%$  point of the distribution of  $\widehat{\mathcal{D}}_{1,2}^{(B)}$ , under the null hypothesis  $H_0$ .

A similar formulation can be stated for the discrimination between members of gamma family in (11). Between two candidate “parent” models  $F_1$  and  $F_2$ , with respective densities  $f_1$  and  $f_2$ , we have to decide in favour of one of them on the basis of the difference  $\mathcal{H}_{Sh}(g_{F_1}^{(G)}) - \mathcal{H}_{Sh}(g_{F_2}^{(G)})$ , which can be easily expressed, through Corollary 2, as

$$\mathcal{D}_{1,2}^{(G)} \triangleq \mathcal{H}_{Sh}(g_{F_1}^{(G)}) - \mathcal{H}_{Sh}(g_{F_2}^{(G)}) = E_Z \left\{ \ln \frac{f_2(F_2^{-1}(1 - e^{-Z}))}{f_1(F_1^{-1}(1 - e^{-Z}))} \right\}, \tag{18}$$

with  $Z$  being a gamma random variable with density  $\frac{1}{\Gamma(\delta)} e^{-z} z^{\delta-1}, z > 0$ .

An equivalent expression of (18) is given by

$$\mathcal{D}_{1,2}^{(G)} = \mathcal{H}_{Sh}(g_{F_1}^{(G)}) - \mathcal{H}_{Sh}(g_{F_2}^{(G)}) = E_Z \left\{ \ln \frac{q_1(1 - e^{-Z})}{q_2(1 - e^{-Z})} \right\}, \tag{19}$$

with  $q_1$  and  $q_2$  being the corresponding quantile densities.

The difference  $\mathcal{D}_{1,2}^{(G)}$  can be estimated as follows. Using the random sample  $X_1, X_2, \dots, X_n$  from (11), we can obtain the random samples  $Z_1^{(j)}, Z_2^{(j)}, \dots, Z_n^{(j)}$  with  $Z_i^{(j)} = -\ln[1 - F_j(X_i)], i = 1, \dots, n$ , and  $F_j$  a “parent” model, for  $j = 1, 2$ . Then,  $Z_1^{(j)}, Z_2^{(j)}, \dots, Z_n^{(j)}, j = 1, 2$ , is a random sample from a gamma distribution with density  $\frac{1}{\Gamma(\delta)} e^{-z} z^{\delta-1}, z > 0$ . Based on (18) and (19), the sample analogs of  $\mathcal{D}_{1,2}^{(G)}$  are

$$\widehat{\mathcal{D}}_{1,2}^{(G)} = \frac{1}{n} \sum_{i=1}^n \ln \frac{f_2 \left( F_2^{-1} \left( 1 - e^{-Z_i^{(2)}} \right) \right)}{f_1 \left( F_1^{-1} \left( 1 - e^{-Z_i^{(1)}} \right) \right)} \tag{20}$$

and

$$\widehat{\mathcal{D}}_{1,2}^{(G)} = \frac{1}{n} \sum_{i=1}^n \ln \frac{q_1 \left( 1 - e^{-Z_i^{(1)}} \right)}{q_2 \left( 1 - e^{-Z_i^{(2)}} \right)}. \tag{21}$$

This formulation can be also used, as before, with the use of  $\widehat{\mathcal{D}}_{1,2}^{(B)}$  to discriminate between two members of (11) that are based on “parent” models  $F_1$  and  $F_2$ .

The next examples give in a closed form the differences  $\mathcal{D}_{1,2}^{(B)}$  and  $\mathcal{D}_{1,2}^{(G)}$  for specific parent distributions  $F_1$  and  $F_2$ .

**Example 23.** Consider the beta-exponential distribution and the beta-Weibull distribution with parent densities  $f_1(x) = (1/\lambda)e^{-x/\lambda}, x > 0, \lambda > 0$ , and  $f_2(x) = \frac{c}{\theta^c} x^{c-1} e^{-(x/\theta)^c}, x > 0$  and  $c, \theta > 0$ . Then, by using (14), we have

$$\mathcal{D}_{1,2}^{(B)} = \ln c - \ln \lambda - \ln \theta + \frac{c - 1}{c} E_Y [\ln(-\ln(1 - Y))], \quad \text{with } Y \sim \text{Beta}(\alpha, \beta), \quad \blacktriangle$$

**Example 24.** In a similar manner, if we consider the gamma-uniform and the gamma-Pareto distributions with parent densities  $f_1(x) = 1/\mu, 0 < x < \mu$ , and  $f_2(x) = k\theta^k/x^{k+1}, x \geq \theta > 0$ , respectively, then use of (18) gives

$$\mathcal{D}_{1,2}^{(G)} = \ln \frac{k}{\theta} - \frac{k+1}{k} \delta + \ln \mu. \blacktriangle$$

**Remark 4.** The procedure for discrimination, introduced in this section, can be extended to the case of more than two, say  $r \geq 2$ , “parent” models. In this framework, consider a random sample  $X_1, X_2, \dots, X_n$  from (1) and suppose the “parent” distribution  $F$  and density  $f$  are unknown, but it is known that they belong to the sets of “parent” distributions  $\{F_1, \dots, F_r\}$  and densities  $\{f_1, \dots, f_r\}$ , respectively. Based on  $X_1, X_2, \dots, X_n$ , it is easy to create random samples  $Y_i^{(j)} = F_j(X_i), i = 1, \dots, n$  and  $j = 1, \dots, r$ . To each pair of “parent” models  $F_j$  and  $f_j$ , there corresponds the expected value  $E_Y[\ln f_j(F_j^{-1}(Y))], j = 1, \dots, r$ , where  $Y \sim \text{Beta}(\alpha, \beta)$ , and its sample analog  $\widehat{\mathcal{J}}_j^{(B)} = \frac{1}{n} \sum_{i=1}^n \ln f_j(F_j^{-1}(Y_i^{(j)})), j = 1, \dots, r$ . Based on the above discussion, it is then reasonable to choose the “parent” models  $F_k$  and  $f_k$  such that  $\widehat{\mathcal{J}}_k^{(B)} = \min_{1 \leq j \leq r} \widehat{\mathcal{J}}_j^{(B)}$ . Exactly the same procedure can also be followed for discriminating between  $r \geq 2$  models obtained from (11) for different “parent” distributions  $F_j$  and densities  $f_j, j = 1, \dots, r$ .

**Remark 5.** Another interesting problem is to consider the parent density and distribution functions  $f$  and  $F$  fixed and to focus on the discrimination between two members of (1) or (11) on the basis of the beta parameters  $\alpha$  and  $\beta$ , or the gamma parameter  $\delta$ . That is the case when we are interested in discriminating between the distributions of two order statistics or between two record densities. The procedures proposed in the papers by Kundu and Gupta can possibly be adopted to study this problem.

**Remark 6.** A parallel problem is the problem of discrimination between beta-generated and gamma-generated distributions. For this case, we need to decide if a random sample  $X_1, X_2, \dots, X_n$  is coming from either  $g_F^{(B)}$  or  $g_F^{(G)}$ . In this regard, we can use for the discrimination between  $g_F^{(B)}$  and  $g_F^{(G)}$  the Kullback–Leibler divergence

$$D_0(g_F^{(G)}, g_F^{(B)}) = \int_{-\infty}^{+\infty} g_F^{(G)}(x) \ln \frac{g_F^{(G)}(x)}{g_F^{(B)}(x)} dx = -\mathcal{H}_{Sh}(g_F^{(G)}) - \int_{-\infty}^{+\infty} g_F^{(G)}(x) \ln g_F^{(B)}(x) dx.$$

It is the only member of the phi-divergence which can be obtained in a closed form for  $g_F^{(B)}$  and  $g_F^{(G)}$ , defined by (1) and (11), respectively. It can be shown, in view of Corollary 2, Lemmas 3 and 5, that

$$D_0(g_F^{(G)}, g_F^{(B)}) = -\ln \Gamma(\delta) + (\delta - 1)\Psi(\delta) + \ln B(\alpha, \beta) + (\beta - 1)\delta + (\alpha - 1) \sum_{m=1}^{\infty} \frac{1}{m(m+1)^\delta}.$$

Observe that  $D_0(g_F^{(G)}, g_F^{(B)})$  does not depend on the parent distribution  $F$ . It is intuitively clear because  $g_F^{(B)}$  and  $g_F^{(G)}$  are based on the same parent  $F$ . The right-hand side of the above equation is the explicit form of the Kullback–Leibler divergence between the density of an order statistic and the density of an upper record.

**Acknowledgement**

A part of this paper has been done while the first author was visiting the Department of Mathematics and Statistics, of the McMaster University, Hamilton, Ontario, Canada in the summer of 2007.

## References

- [1] M.C. Jones, Families of distributions arising from distributions of order statistics, *TEST* 13 (2004) 1–43.
- [2] M.C. Jones, On a class of distributions defined by the relationship between their density and distribution functions, *Comm. Statist. Theory Methods* 36 (2007) 1835–1843.
- [3] N. Eugene, C. Lee, F. Famoye, Beta-normal distribution and its applications, *Comm. Statist. Theory Methods* 31 (2002) 497–512.
- [4] M.C. Jones, P.V. Larsen, Multivariate distributions with support above the diagonal, *Biometrika* 91 (2004) 975–986.
- [5] S. Nadarajah, S. Kotz, The beta exponential distribution, *Reliab. Eng. Syst. Safety* 91 (2006) 689–697.
- [6] B.W. Brown, F.M. Spears, L.B. Levy, The log  $F$ : A distribution for all seasons, *Comput. Statist* 17 (2002) 47–58.
- [7] K. Zografos, On some beta generated distributions and their maximum entropy characterization: The beta-Weibull distribution, in: N.S. Barnett, S.S. Dragomir (Eds.), *Advances in Inequalities from Probability Theory & Statistics*, Nova Science Publishers, New Jersey, 2008, pp. 237–260.
- [8] B.C. Arnold, E. Castillo, J.M. Sarabia, Families of multivariate distributions involving the Rosenblatt construction, *J. Amer. Statist. Assoc.* 101 (2006) 1652–1662.
- [9] J.T.A.S. Ferreira, M.F.J. Steel, A constructive representation of univariate skewed distributions, *J. Amer. Statist. Assoc.* 101 (2006) 823–829.
- [10] J.T.A.S. Ferreira, M.F.J. Steel, Model comparison of coordinate-free multivariate skewed distributions with an application to stochastic frontiers, *J. Econometrics* 137 (2007) 641–673.
- [11] C.E. Shannon, A mathematical theory of communication, *Bell System Tech. J.* 27 (1948) 379–432.
- [12] E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* 106 (1957) 620–630.
- [13] J.E. Shore, R.W. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, *IEEE Trans. Inform Theory* 26 (1980) 26–37.
- [14] J.N. Kapur, *Maximum Entropy Models in Science and Engineering*, John Wiley & Sons, New York, 1989.
- [15] E.S. Soofi, Principal information theoretic approaches, *J. Amer. Statist. Assoc.* 95 (2000) 1349–1353.
- [16] S. Kullback, *Information Theory and Statistics*, John Wiley & Sons, New York, 1959.
- [17] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.
- [18] R.D. Gupta, D. Kundu, Exponentiated exponential family: An alternative to gamma and Weibull distributions, *Biomed. J.* 43 (2001) 117–130.
- [19] N. Balakrishnan, Order statistics from the half logistic distribution, *J. Stat. Comput. Simul.* 20 (1985) 287–309.
- [20] G.S. Mudholkar, D.K. Srivastava, Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Trans. on Reliab.* 42 (1993) 299–302.
- [21] B.C. Arnold, N. Balakrishnan, H.N. Nagaraja, *Records*, John Wiley & Sons, New York, 1998.
- [22] S. Baratpour, J. Ahmadi, N.R. Arghami, Entropy properties of record statistics, *Statist. Papers* 48 (2007) 197–213.
- [23] D. Kundu, R.D. Gupta, A. Manglick, Discriminating between the log-normal and generalized exponential distributions, *J. Statist. Plann. Inference* 127 (2005) 213–227.