



# The beta Laplace distribution

Gauss M. Cordeiro<sup>a</sup>, Artur J. Lemonte<sup>b,\*</sup>

<sup>a</sup> Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco, Brazil

<sup>b</sup> Departamento de Estatística, Universidade de São Paulo, Brazil

## ARTICLE INFO

### Article history:

Received 14 August 2010

Received in revised form 7 January 2011

Accepted 22 January 2011

Available online 2 February 2011

### Keywords:

Double exponential distribution

Laplace distribution

Maximum likelihood estimation

Mean deviation

Order statistic

## ABSTRACT

The Laplace distribution is one of the earliest distributions in probability theory. For the first time, based on this distribution, we propose the so-called beta Laplace distribution, which extends the Laplace distribution. Various structural properties of the new distribution are derived, including expansions for its moments, moment generating function, moments of the order statistics, and so forth. We discuss maximum likelihood estimation of the model parameters and derive the observed information matrix. The usefulness of the new model is illustrated by means of a real data set.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

One of the earliest distributions in probability theory was introduced by Laplace in 1774 (Laplace, 1774). A random variable  $Z$  has the Laplace distribution with location parameter  $\mu$  and scale parameter  $\sigma > 0$ , say  $Z \sim L(\mu, \sigma)$ , if its probability density function (pdf) is given by

$$g(z) = \frac{1}{2\sigma} \exp\left\{-\frac{|z - \mu|}{\sigma}\right\}, \quad -\infty < z < \infty.$$

The mean, median and mode are all equal to  $\mu$ . The variance is  $2\sigma^2$  and the skewness and kurtosis are 0 and 6, respectively. The moment generating function (mgf) of  $Z$  is  $M(t) = (1 + \sigma^2 t^2)^{-1} \exp(\mu t)$ . In addition, the cumulative distribution function (cdf) becomes

$$G(z) = \begin{cases} \frac{1}{2} \exp\left(\frac{z - \mu}{\sigma}\right), & z < \mu, \\ 1 - \frac{1}{2} \exp\left(-\frac{z - \mu}{\sigma}\right), & z \geq \mu. \end{cases}$$

If we consider the standardized random variable  $X = (Z - \mu)/\sigma$ , the pdf of  $X$  reduces to  $g(x) = \frac{1}{2} \exp(-|x|)$ ,  $-\infty < x < \infty$ , and the corresponding cdf and mgf are given by

\* Corresponding author.

E-mail address: [arturlemonte@gmail.com](mailto:arturlemonte@gmail.com) (A.J. Lemonte).

$$G(x) = \begin{cases} \frac{1}{2} \exp(x), & x < 0, \\ 1 - \frac{1}{2} \exp(-x), & x \geq 0, \end{cases}$$

and  $M(t) = (1 + t^2)^{-1}$ , respectively. In this case,  $X \sim L(0, 1)$ .

The Laplace distribution, also named the double exponential distribution, and its variants are becoming popular in many areas of science and engineering. This distribution is often used for modeling phenomena with “heavier than normal tails”; see for example, Andrews et al. (1972), Manly (1976), Easterling (1978), Hsu (1979), Bagchi et al. (1983), Hoaglin et al. (1983), Dadi and Marks (1987), Damsleth and El-Shaarawi (1989), Puig and Stephens (2000), Chen (2002), and also Johnson et al. (1995) which contains a detailed list of references. A book-length account of Laplace distributions, discussing in great detail their various properties and applications, is available due to Kotz et al. (2001).

In this article we propose a new model, so-called the beta Laplace (BL) distribution, which contains as a sub-model the Laplace distribution. The BL distribution is convenient for modeling asymmetric data as a competitive model to beta normal and skew-normal distributions. We obtain some mathematical properties, discuss maximum likelihood estimation of the parameters and derive the observed information matrix. The article is outlined as follows. In Section 2, we introduce the BL distribution and provide plots of the density function. We demonstrate that the BL density function can be expressed as an infinite linear combination of Laplace density functions in Section 3. We provide in Section 4 a general expansion for the moments and mgf. Expansions for the quantile function and mean deviations are provided in Section 5. In Section 6, we demonstrate that the density function of the BL order statistics can be written as a linear combination of Laplace densities. We also obtain expansions for the moments of the order statistics. The Rényi and Shannon entropies are derived in Section 7. Maximum likelihood estimation is addressed in Section 8. Section 9 illustrates the importance of the BL distribution through the analysis of a real data set. Finally, Section 10 offers some concluding remarks.

## 2. The beta Laplace distribution

The generalization of the Laplace distribution is motivated by the work of Eugene et al. (2002) who defined a class of generalized beta distributions by

$$F(x) = \frac{1}{B(a, b)} \int_0^{G(x)} \omega^{a-1} (1 - \omega)^{b-1} d\omega = I_{G(x)}(a, b). \quad (1)$$

Here,  $a > 0$  and  $b > 0$  are two additional parameters which control skewness through the relative tail weights,  $I_y(a, b) = B_y(a, b)/B(a, b)$  is the incomplete beta function ratio,  $B_y(a, b) = \int_0^y \omega^{a-1} (1 - \omega)^{b-1} d\omega$  is the incomplete beta function,  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  is the beta function and  $\Gamma(\cdot)$  is the gamma function. This class of generalized distributions has been receiving considerable attention over the last years in particular after the work of Jones (2004). The probability density function (pdf) corresponding to (1) is  $f(x) = g(x)G(x)^{a-1}\{1 - G(x)\}^{b-1}/B(a, b)$ , where  $g(x) = dG(x)/dx$  is the parent density function. The density  $f(x)$  will be most tractable when both functions  $G(x)$  and  $g(x)$  have simple analytic expressions. Except for special choices of these functions,  $f(x)$  will be difficult to deal with some generality.

Eugene et al. (2002), Nadarajah and Gupta (2004), Nadarajah and Kotz (2004, 2006), Lee et al. (2007) and Akinsete et al. (2008) defined the beta normal, beta Fréchet, beta Gumbel, beta exponential, beta Weibull and beta Pareto distributions by taking  $G(x)$  to be the cdf of the normal, Fréchet, Gumbel, exponential, Weibull and Pareto distributions, respectively. More recently, Barreto-Souza et al. (2010), Pescim et al. (2010) and Cordeiro and Lemonte (2011) introduced the beta generalized exponential, the beta generalized half-normal and the beta Birnbaum–Saunders distributions, respectively.

The cdf of the BL distribution can be written as

$$F(x) = \begin{cases} I_{\exp(x)/2}(a, b), & x < 0, \\ I_{1-\exp(-x)/2}(a, b), & x \geq 0. \end{cases} \quad (2)$$

The density function corresponding to (2) is given by

$$f(x) = \begin{cases} \{2^a B(a, b)\}^{-1} \exp(-|x|) \exp\{-|x|(a-1)\} \{1 - \exp(-|x|)/2\}^{b-1}, & x < 0, \\ \{2^b B(a, b)\}^{-1} \exp(-|x|) \exp\{-|x|(b-1)\} \{1 - \exp(-|x|)/2\}^{a-1}, & x \geq 0. \end{cases} \quad (3)$$

We note that the case  $x < 0$  can be obtained from the case  $x \geq 0$  by replacing  $x$  for  $-x$  and interchanging  $a$  and  $b$ . Clearly, for  $a = b = 1$ , Eq. (3) reduces to the standard Laplace density function. If  $X$  follows (3), we write  $X \sim \text{BL}(a, b)$ . Plots of the  $\text{BL}(a, b)$  distribution are illustrated in Fig. 1 for selected parameter values, including the special case of the standard Laplace distribution. It is evident that the BL distribution is much more flexible than the Laplace distribution.

The properties of a random variable  $Z$  having the BL distribution with location parameter  $\mu$  and dispersion parameter  $\sigma$ , say  $Z \sim \text{BL}(\mu, \sigma, a, b)$ , can be determined directly from those properties of  $X$  using the linear transformation  $Z = \mu + \sigma X$ .

## 3. Expansions

First, if  $|z| < 1$  and  $b > 0$  is real non-integer, we have the power series expansion  $(1 - z)^{b-1} = \sum_{j=0}^{\infty} \binom{b-1}{j} (-1)^j z^j$ . Applying this expansion for  $x < 0$ , we obtain

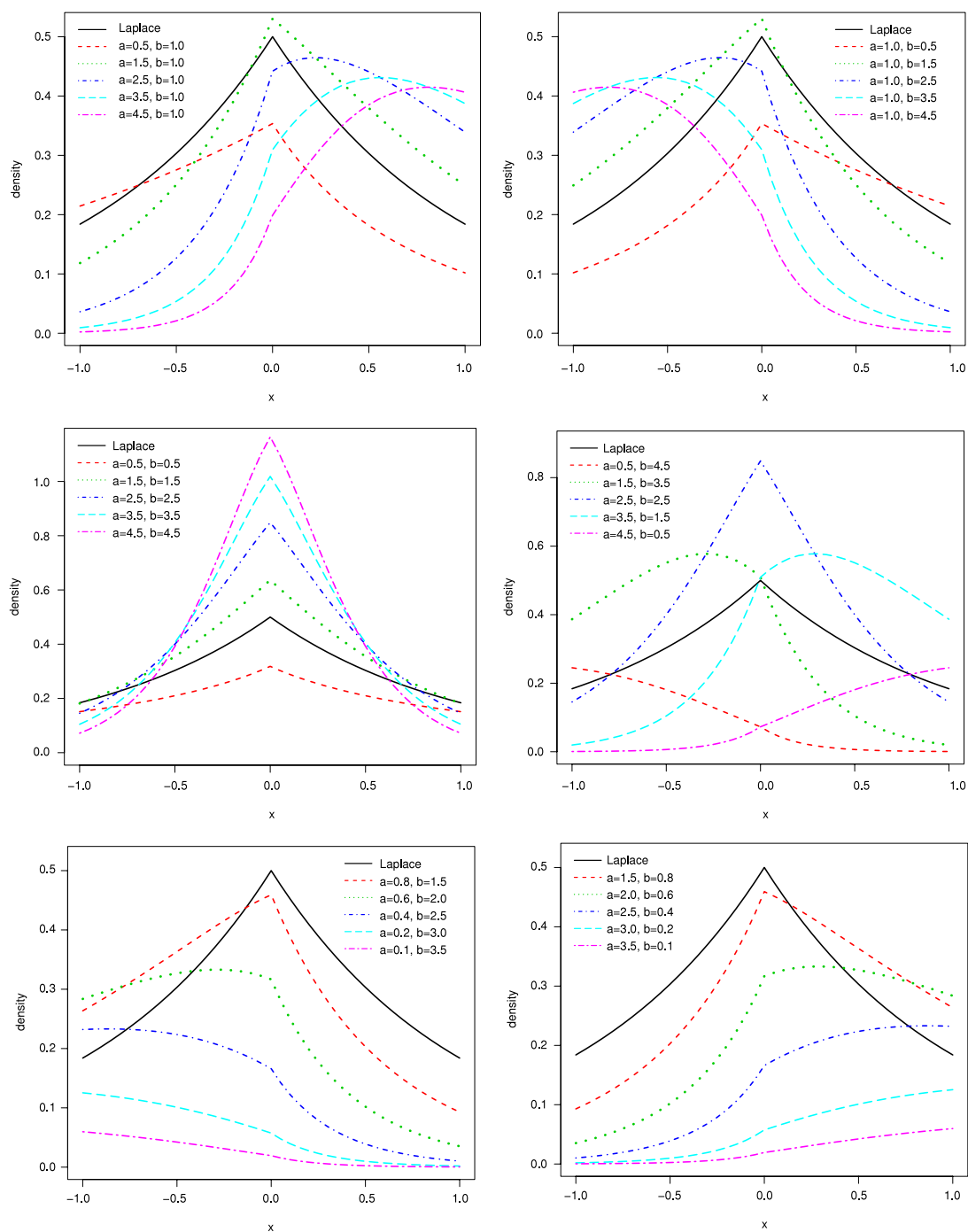


Fig. 1. Plots of the density function (3) for some parameter values.

$$f(x) = \frac{1}{2^a B(a, b)} \sum_{j=0}^{\infty} \binom{b-1}{j} \left(-\frac{1}{2}\right)^j \exp\{-(a+j)|x|\},$$

and then it follows the linear combination form

$$f(x) = \sum_{j=0}^{\infty} w_j g_{a+j}(x).$$

(4)

Here, the coefficients are  $w_j = w_j(a, b) = (-1)^j \binom{b-1}{j} / \{2^{a+j-1}(a+j)B(a, b)\}$  and from now on  $g_\phi(x) = \phi \exp(-\phi|x|)/2$  denotes the Laplace distribution with zero mean and precision parameter  $\phi$  (the variance is  $2/\phi^2$ ). Analogously, for  $x \geq 0$ , we obtain

$$f(x) = \sum_{j=0}^{\infty} v_j g_{b+j}(x), \quad (5)$$

where the coefficients are  $v_j = v_j(a, b) = (-1)^j \binom{a-1}{j} / \{2^{b+j-1}(b+j)B(a, b)\}$ . Some mathematical properties of the BL distribution can be obtained from expansions (4) and (5) and those properties of the Laplace distribution. The cdf corresponding to  $g_\phi(x)$  is  $G_\phi(x) = e^{\phi x}/2$  if  $x < 0$  and  $G_\phi(x) = 1 - e^{-\phi x}/2$  if  $x \geq 0$ . Then, the cdfs corresponding to (4) and (5) are  $F(x) = \sum_{j=0}^{\infty} w_j G_{a+j}(x)$  and  $F(x) = \sum_{j=0}^{\infty} v_j G_{b+j}(x)$ , respectively.

Eqs. (4) and (5) (and others expansions in this article) can be computed numerically in software such as MAPLE (Garvan, 2002), MATLAB (Sigmon and Davis, 2002) and MATHEMATICA (Wolfram, 2003). These symbolic software have currently the ability to deal with analytic expressions of formidable size and complexity.

#### 4. Moments, cumulants and generating function

Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness and kurtosis). Let  $I_{r,\phi} = \int_0^\infty x^r g_\phi(x) dx = r!(2\phi)^{-1}$ . The moments of the BL distribution can be obtained using this result. If  $X$  follows the BL distribution (3), the  $r$ th moment about zero can be written as

$$E(X^r) = \sum_{j=0}^{\infty} w_j \int_{-\infty}^0 x^r g_{a+j}(x) dx + \sum_{j=0}^{\infty} v_j I_{r,b+j}.$$

But  $\sum_{j=0}^{\infty} w_j \int_{-\infty}^0 x^r g_{a+j}(x) dx = (-1)^r \sum_{j=0}^{\infty} w_j I_{r,a+j}$  and then

$$\mu'_r = E(X^r) = \frac{r!}{2} \sum_{j=0}^{\infty} \left\{ \frac{(-1)^r w_j}{(a+j)^r} + \frac{v_j}{(b+j)^r} \right\}, \quad (6)$$

where  $w_j$  and  $v_j$  are defined in Section 3. For  $a = b$ , which corresponds to a symmetrical distribution, Eq. (6) implies that  $\mu'_r = 0$  for  $r$  odd, as expected.

The central moments ( $\mu_s$ ) and cumulants ( $\kappa_s$ ) of  $X$  are easily obtained from the ordinary moments by  $\mu_s = \sum_{k=0}^s \binom{s}{k} (-1)^k \mu'_1 \mu'_{s-k}$  and  $\kappa_p = \mu'_p - \sum_{k=1}^{p-1} \binom{p-1}{k-1} \kappa_k \mu'_{p-k}$ , respectively, where  $\kappa_1 = \mu'_1$ . Thus,  $\kappa_2 = \mu'_2 - \mu_1'^2$ ,  $\kappa_3 = \mu'_3 - 3\mu'_2 \mu'_1 + 2\mu_1'^3$ , etc. The  $r$ th descending factorial moment of  $X$  is  $\mu'_{(r)} = E\{X^{(r)}\} = E\{X(X-1) \times \cdots \times (X-r+1)\} = \sum_{k=0}^r s(r, k) \mu'_k$ , where  $s(r, k)$  is the Stirling number of the first kind defined by  $s(r, k) = (k!)^{-1} [d^k x^{(r)} / dx^k]_{x=0}$ . They count the number of ways to permute a list of  $r$  items into  $k$  cycles. The factorial moments of  $X$  can be written as

$$\mu'_{(r)} = \sum_{j=0}^{\infty} \sum_{k=0}^r \frac{k! s(r, k)}{2} \left\{ \frac{(-1)^k w_j}{(a+j)^k} + \frac{v_j}{(b+j)^k} \right\}.$$

The mgf of the BL distribution can be written from (4) and (5) as

$$M(t) = E\{\exp(tX)\} = \sum_{j=0}^{\infty} w_j \int_0^\infty \exp(-tx) g_{a+j}(x) dx + \sum_{j=0}^{\infty} v_j \int_0^\infty \exp(tx) g_{b+j}(x) dx.$$

For  $t < \phi$ , we can write  $\int_0^\infty \exp(tx) g_\phi(x) dx = (\phi/2) \int_0^\infty \exp\{(t-\phi)x\} dx = \phi / \{2(\phi-t)\}$  and then

$$M(t) = \sum_{j=0}^{\infty} \left\{ \frac{(a+j)w_j}{2(a+j+t)} + \frac{(b+j)v_j}{2(b+j-t)} \right\}.$$

#### 5. Quantile function and mean deviations

We can generate a random variable  $X$  having the BL distribution from a standard beta variate  $V$  with parameters  $a > 0$  and  $b > 0$ . We define  $X = \log(2V)$  if  $V < 1/2$  and  $X = -\log[2(1-V)]$  if  $V > 1/2$ . The BL quantile function  $x = Q(u)$  can be obtained by inverting (2). We can express  $x = Q(u)$  in terms of the quantile function  $Q_{a,b}(u)$  of the beta distribution

which can be found in the Wolfram website (<http://functions.wolfram.com/06.23.06.0004.01>)

$$Q_{a,b}(u) = w + \frac{b-1}{a+1}w^2 + \frac{(b-1)(a^2+3ab-a+5b-4)}{2(a+1)^2(a+2)}w^3 \\ + \frac{(b-1)[a^4+(6b-1)a^3+(b+2)(8b-5)a^2]}{3(a+1)^3(a+2)(a+3)}w^4 \\ + \frac{(b-1)[(33b^2-30b+4)a+b(31b-47)+18]}{3(a+1)^3(a+2)(a+3)}w^5 + O(w^{6/a}),$$

where  $w = \{a, B(a, b)u\}^{1/a}$  for  $a > 0$ . For given  $u$ , we can calculate  $Q_{a,b}(u)$  from the above expansion and then obtain  $x = Q(u) = \log\{2Q_{a,b}(u)\}$  if  $Q_{a,b}(u) < 1/2$  and  $x = Q(u) = -\log\{2[1 - 2Q_{a,b}(u)]\}$  if  $Q_{a,b}(u) > 1/2$ .

In what follows we derive the mean deviations. The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. These are known as the mean deviation about the mean and the mean deviation about the median, defined by  $\delta_1 = \int_{-\infty}^{\infty} |x - \mu'_1| dx$  and  $\delta_2 = \int_{-\infty}^{\infty} |x - m| dx$ , respectively, where  $\mu'_1 = E(X)$  and  $m$  is the median of  $X$ . These measures can be expressed as  $\delta_1 = 2\mu'_1 F(\mu'_1) - 2T(\mu'_1)$  and  $\delta_2 = \mu'_1 - 2T(m)$ , where  $T(q) = \int_{-\infty}^q xf(x)dx$ . Clearly, for the BL distribution,  $F(\mu'_1)$  is easily calculated from (2) and the median  $m$  follows  $I_{\frac{1}{2} \exp(m)}(a, b) = 1/2$  if  $m < 0$  and  $I_{1-\frac{1}{2} \exp(-m)}(a, b) = 1/2$  if  $m \geq 0$ .

We now obtain  $T(q) = \int_{-\infty}^q xf(x)dx$ . If  $q < 0$ , Eq. (4) gives  $T(q) = \frac{1}{2} \sum_{j=0}^{\infty} (a+j)w_j \int_{-\infty}^q x \exp\{(a+j)x\} dx$  and then

$$T(q) = \frac{1}{2} \sum_{j=0}^{\infty} \frac{w_j}{(a+j)} \{\gamma(-(a+j)q, 2) - 1\},$$

where  $\gamma(a, p) = \int_0^a x^{p-1} e^{-x} dx$  is the incomplete gamma function. On the other hand, for  $q \geq 0$ , we obtain

$$T(q) = \sum_{j=0}^{\infty} \left\{ \frac{w_j}{2(a+j)} + \frac{v_j \gamma((b+j)q, 2)}{2(b+j)} \right\}.$$

The equation for  $T(q)$  can be used to determine Bonferroni and Lorenz curves which have applications in some fields such as economics, reliability, demography, insurance and medicine. They are defined by  $B(p) = T(q)/(p\mu'_1)$  and  $L(p) = T(q)/\mu'_1$ , respectively, where  $q = Q(p)$  and  $p$  is a given probability.

## 6. Order statistics and moments

Moments of order statistics play an important role in quality control testing and reliability, where a practitioner needs to predict the failure of future items based on the times of a few early failures. These predictors are often based on moments of order statistics. We now derive an explicit expression for the density function of the  $i$ th order statistic  $X_{i:n}$ , say  $f_{i:n}(x)$ , in a random sample of size  $n$  from the BL distribution. We can write

$$f_{i:n}(x) = \sum_{l=0}^{n-i} \frac{(-1)^l \binom{n-i}{l} f(x)}{B(i, n-i+1)} F(x)^{i+l-1}.$$

Let  $(h)_k = \Gamma(h+k)/\Gamma(h)$ . Using the incomplete beta function expansion for  $b$  real non-integer

$$I_x = I_x(a, b) = \frac{x^a}{B(a, b)} \sum_{m=0}^{\infty} \frac{(1-b)_m x^m}{(a+m)m!},$$

we can express  $F(x)$  for  $x < 0$  as  $F(x) = \sum_{m=0}^{\infty} a_m \exp\{-(a+m)|x|\}$ , where  $a_m = (1-b)_m / \{2^{a+m} (a+m)m! B(a, b)\}$ . Then,

$$F(x)^{i+l-1} = \sum_{m_1, \dots, m_{i+l-1}=0}^{\infty} A(m_1, \dots, m_{i+l-1}) \exp\{-(a(i+l-1) + m_+) |x|\},$$

where  $A(m_1, \dots, m_{i+l-1}) = a_{m_1} \times \dots \times a_{m_{i+l-1}}$  and  $m_+ = m_1 + \dots + m_{i+l-1}$ . Hence,  $f_{i:n}(x)$  can be written from (4) as a linear combination form

$$f_{i:n}(x) = \sum_{l=0}^{n-i} \sum_{j=0}^{\infty} \sum_{m_1, \dots, m_{i+l-1}=0}^{\infty} w_{i,l,j}^* g_{a(i+l+j)+m_+}(x), \quad x < 0, \quad (7)$$

where

$$w_{i,l,j}^* = w_{i,l,j}^*(a, b) = \frac{(-1)^l \binom{n-i}{l} (a+j)w_j}{B(i, n-i+1)} \frac{A(m_1, \dots, m_{i+l-1})}{a(i+l+j) + m_+}.$$

In a similar way we can obtain a linear combination for  $f_{i:n}(x)$  when  $x \geq 0$ . We have

$$F(x) = \sum_{r=0}^{\infty} a_r \{1 - \exp(-|x|)\}^{a+r} = \sum_{m=0}^{\infty} b_m \exp(-m|x|),$$

where  $b_m = \frac{2(-1)^m}{m} \sum_{r=0}^{\infty} a_r \binom{a+r}{m}$ . Then,

$$F(x)^{i+l-1} = \sum_{m_1, \dots, m_{i+l-1}=0}^{\infty} B(m_1, \dots, m_{i+l-1}) \exp(-m_+ |x|),$$

where  $B(m_1, \dots, m_{i+l-1}) = b_{m_1} \times \dots \times b_{m_{i+l-1}}$ . Thus, we have

$$f_{i:n}(x) = \sum_{l=0}^{n-i} \sum_{j=0}^{\infty} \sum_{m_1, \dots, m_{i+l-1}=0}^{\infty} v_{i,l,j}^* g_{b+j+m_+}(x), \quad x \geq 0, \quad (8)$$

where

$$v_{i,l,j}^* = v_{i,l,j}^*(a, b) = \frac{(-1)^l \binom{n-i}{l} (b+j) v_j B(m_1, \dots, m_{i+l-1})}{B(i, n-i+1) (b+j+m_+)}.$$

We can provide several mathematical properties of the BL order statistics such as the ordinary moments, inverse and factorial moments, mgf, mean deviations (and so forth) by combining Eqs. (7) and (8). The  $p$ th moment of the  $i$ th order statistic is determined from these equations as

$$E(X_{i:n}^p) = \sum_{l=0}^{n-i} \sum_{j=0}^{\infty} \sum_{m_1, \dots, m_{i+l-1}=0}^{\infty} \{(-1)^p w_{i,l,j}^* I_{p,a(i+l+j)+m_+} + v_{i,l,j}^* I_{p,b+j+m_+}\}.$$

The mgf of the BL order statistics can be calculated in a similar manner.

## 7. Rényi and Shannon entropies

The entropy of a random variable is a measure of variation of the uncertainty. Entropy has been used in various situations in science and engineering, and numerous measures of entropy have been studied and compared in the literature. For the BL distribution (3), the Rényi entropy is defined by  $\mathcal{J}_R(\delta) = (1 - \delta)^{-1} \log\{J(\delta)\}$ , where  $J(\delta) = \int f^\delta(x) dx$ ,  $\delta > 0$  and  $\delta \neq 1$ . The quantity  $J(\delta)$  follows from (3) as

$$J(\delta) = \int_{-\infty}^0 \frac{\exp(-\delta|x|)}{2^{a\delta} B(a, b)^\delta} \exp\{-|x|(a-1)\delta\} \left\{1 - \frac{1}{2} \exp(-|x|)\right\}^{(b-1)\delta} dx \\ + \int_0^{\infty} \frac{\exp(-\delta|x|)}{2^{b\delta} B(a, b)^\delta} \left\{1 - \frac{1}{2} \exp(-|x|)\right\}^{(a-1)\delta} \exp\{-|x|(b-1)\delta\} dx.$$

We have only to calculate the first integral since the second one follows by replacing  $a$  by  $b$ . The first integral, say  $J_1(\delta)$ , is immediately obtained by expanding the binomial

$$J_1(\delta) = \frac{1}{2^{a\delta} B(a, b)^\delta} \sum_{k=0}^{\infty} \binom{(b-1)\delta}{k} \frac{(-1)^k}{2^k (a\delta + k)}.$$

Hence, the Rényi entropy reduces to

$$\mathcal{J}_R(\delta) = \frac{1}{1-\delta} \log \left\{ \sum_{k=0}^{\infty} \left[ \frac{(-1)^k \binom{(b-1)\delta}{k}}{2^{a\delta+k} B(a, b)^\delta (a\delta + k)} + \frac{(-1)^k \binom{(a-1)\delta}{k}}{2^{b\delta+k} B(a, b)^\delta (b\delta + k)} \right] \right\}.$$

The Shannon entropy is defined by  $\mathcal{J}_S = E[-\log f(X)]$ . This is a special case derived from  $\lim_{\delta \rightarrow 1} \mathcal{J}_R(\delta)$ . For  $x < 0$ , it follows from (3) that

$$E[-\log f(X)] = a \log 2 + \log B(a, b) + \frac{a}{2} \sum_{j=0}^{\infty} \frac{w_j}{a+j} + \frac{(b-1)}{2B(a, b)} \sum_{j=1}^{\infty} j^{-1} B(a+j, b).$$

For  $x > 0$ , we obtain

$$E[-\log f(X)] = b \log 2 + \log B(a, b) + \frac{b}{2} \sum_{j=0}^{\infty} \frac{v_j}{b+j} + \frac{(a-1)}{2B(a, b)} \sum_{j=1}^{\infty} j^{-1} B(a, b+j).$$

Hence, the Shannon entropy can be written as

$$\begin{aligned} J_S &= (a+b) \log 2 + 2 \log B(a, b) + \frac{1}{2} \sum_{j=0}^{\infty} \left\{ \frac{aw_j}{a+j} + \frac{bv_j}{b+j} \right\} \\ &+ \frac{1}{2B(a, b)} \sum_{j=1}^{\infty} j^{-1} \{ (b-1)B(a+j, b) + (a-1)B(b, a+j) \}. \end{aligned}$$

## 8. Estimation

The parameters of the BL distribution are estimated by maximum likelihood. First, we shall consider the BL distribution with  $\mu = 0$ . Let  $\mathbf{z} = (z_1, \dots, z_n)^\top$  denote a random sample of the BL distribution with unknown parameter vector  $\boldsymbol{\theta} = (\sigma, a, b)^\top$ . The log-likelihood function for  $\boldsymbol{\theta}$  can be written as

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= - \sum_{i=1}^n \delta_i \log \{ \sigma 2^a B(a, b) \} + \sum_{i=1}^n \delta_i \left\{ \frac{az_i}{\sigma} + (b-1)d_1(z_i; \sigma) \right\} \\ &- \sum_{i=1}^n (1-\delta_i) \log \{ \sigma 2^b B(a, b) \} + \sum_{i=1}^n (1-\delta_i) \left\{ -\frac{bz_i}{\sigma} + (a-1)d_1(-z_i; \sigma) \right\}, \end{aligned}$$

where  $d_1(y; \sigma) = \log \{ 1 - \exp(y/\sigma)/2 \}$  and  $\delta_i = \{1: \text{if } z_i < 0; 0: \text{if } z_i \geq 0\}$ , for  $i = 1, \dots, n$ . The components of the score vector  $\mathbf{U}(\boldsymbol{\theta}) = (U_\sigma, U_a, U_b)^\top$  are

$$\begin{aligned} U_\sigma &= -\frac{n}{\sigma} + \sum_{i=1}^n \delta_i \left\{ -\frac{az_i}{\sigma^2} + \frac{(b-1)}{2} d_2(z_i; \sigma) \right\} + \sum_{i=1}^n (1-\delta_i) \left\{ \frac{bz_i}{\sigma^2} + \frac{(a-1)}{2} d_2(-z_i; \sigma) \right\}, \\ U_a &= n\{\psi(a+b) - \psi(a)\} + \sum_{i=1}^n \delta_i \left( -\log 2 + \frac{z_i}{\sigma} \right) + \sum_{i=1}^n (1-\delta_i) d_1(-z_i; \sigma), \\ U_b &= n\{\psi(a+b) - \psi(b)\} + \sum_{i=1}^n \delta_i d_1(z_i; \sigma) + \sum_{i=1}^n (1-\delta_i) \left( -\log 2 - \frac{z_i}{\sigma} \right), \end{aligned}$$

where  $d_2(y; \sigma) = y \exp(y/\sigma) / [\sigma^2 \{1 - \exp(y/\sigma)/2\}]$  and  $\psi(\cdot)$  is the digamma function.

Setting these equations to zero,  $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$ , and solving them simultaneously yields the maximum likelihood estimate (MLE)  $\hat{\boldsymbol{\theta}} = (\hat{\sigma}, \hat{a}, \hat{b})^\top$  of  $\boldsymbol{\theta} = (\sigma, a, b)^\top$ . These equations can be solved numerically via iterative methods. The BFGS method with analytical derivatives has been used for maximizing the log-likelihood function  $\ell(\boldsymbol{\theta})$ . For further details about the BFGS method the reader is referred to [Nocedal and Wright \(1999, Section 8.1\)](#) and [Press et al. \(2007, Section 10.7\)](#).

The normal approximation of the MLE of  $\boldsymbol{\theta}$  can be used for constructing approximate confidence intervals and for testing hypotheses on the parameters  $\sigma$ ,  $a$  and  $b$ . Under conditions that are fulfilled for the parameters in the interior of the parameter space, we have that  $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{A}{\sim} \mathcal{N}_3(\mathbf{0}, \mathbf{K}(\boldsymbol{\theta})^{-1})$ , where  $\overset{A}{\sim}$  denotes approximately distributed and  $\mathbf{K}(\boldsymbol{\theta})$  is the expected information matrix. This approximation holds if  $\mathbf{K}(\boldsymbol{\theta})$  is replaced by  $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$ , where  $\boldsymbol{\Sigma}(\boldsymbol{\theta})$  is the observed information matrix, given by

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = - \begin{pmatrix} U_{\sigma\sigma} & U_{\sigma a} & U_{\sigma b} \\ \cdot & U_{aa} & U_{ab} \\ \cdot & \cdot & U_{bb} \end{pmatrix},$$

whose elements are

$$\begin{aligned} U_{\sigma\sigma} &= \frac{n}{\sigma^2} + \sum_{i=1}^n \delta_i \left\{ \frac{2az_i}{\sigma^3} - \frac{(b-1)}{\sigma} d_2(z_i; \sigma) - \frac{(b-1)}{2\sigma^2} z_i d_2(z_i; \sigma) - \frac{(b-1)}{4} d_2(z_i; \sigma)^2 \right\} \\ &+ \sum_{i=1}^n (1-\delta_i) \left\{ -\frac{2bz_i}{\sigma^3} - \frac{(a-1)}{\sigma} d_2(-z_i; \sigma) + \frac{(a-1)}{2\sigma^2} z_i d_2(-z_i; \sigma) - \frac{(a-1)}{4} d_2(-z_i; \sigma)^2 \right\}, \\ U_{\sigma a} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \delta_i z_i + \frac{1}{2} \sum_{i=1}^n (1-\delta_i) d_2(-z_i; \sigma), \quad U_{\sigma b} = \frac{1}{2} \sum_{i=1}^n \delta_i d_2(z_i; \sigma) + \frac{1}{\sigma^2} \sum_{i=1}^n (1-\delta_i) z_i, \\ U_{aa} &= n\{\psi'(a+b) - \psi'(a)\}, \quad U_{ab} = n\psi'(a+b), \quad U_{bb} = n\{\psi'(a+b) - \psi'(b)\}, \end{aligned}$$

where  $\psi'(\cdot)$  is the trigamma function. Thus, the multivariate normal  $\mathcal{N}_3(\mathbf{0}, \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})^{-1})$  distribution can be used to construct approximate confidence intervals for the parameters  $\sigma$ ,  $a$  and  $b$ , which are given, respectively, by  $\hat{\sigma} \pm z_{\eta/2} \times [\widehat{\text{var}}(\hat{\sigma})]^{1/2}$ ,  $\hat{a} \pm z_{\eta/2} \times [\widehat{\text{var}}(\hat{a})]^{1/2}$ ,  $\hat{b} \pm z_{\eta/2} \times [\widehat{\text{var}}(\hat{b})]^{1/2}$ , where  $\widehat{\text{var}}(\cdot)$  is the diagonal element of  $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})^{-1}$  corresponding to each parameter, and  $z_{\eta/2}$  is the quantile  $100(1 - \eta/2)\%$  of the standard normal distribution.

Let  $\theta = (\mu, \sigma, a, b)^\top$  be the unknown parameter vector of the BL distribution. In what follows we briefly discuss how to obtain the MLE of  $\theta$ . Let  $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(n)}$  be the corresponding order statistics for the random sample  $z_1, z_2, \dots, z_n$ . The MLE of  $\mu$  is the one of  $z_{(1)}, z_{(2)}, \dots, z_{(n)}$ , say  $\hat{\mu} = z_{(m)}$ , that maximizes

$$L(\theta) = C \exp \left\{ a \sum_{i=1}^{m-1} \frac{(z_{(i)} - \mu)}{\sigma} - b \sum_{i=m}^n \frac{(z_{(i)} - \mu)}{\sigma} \right\} \left[ \prod_{i=1}^{m-1} \left( 1 - \frac{1}{2} \exp \left\{ \frac{(z_{(i)} - \mu)}{\sigma} \right\} \right) \right]^{b-1} \\ \times \left[ \prod_{i=m}^n \left( 1 - \frac{1}{2} \exp \left\{ -\frac{(z_{(i)} - \mu)}{\sigma} \right\} \right) \right]^{a-1},$$

where  $C = C(\sigma, a, b) = [2^{a(m-1)+b(n-m+1)} \sigma^n B(a, b)^n]^{-1}$  and the MLEs of  $\sigma$ ,  $a$  and  $b$  are obtained from the simultaneous solutions of the equations

$$\frac{n}{\hat{\sigma}} + \hat{a} \sum_{i=1}^{m-1} \frac{(z_{(i)} - \hat{\mu})}{\hat{\sigma}^2} = \hat{b} \sum_{i=m}^n \frac{(z_{(i)} - \hat{\mu})}{\hat{\sigma}^2} + \frac{(\hat{b} - 1)}{2} \sum_{i=1}^{m-1} d_2(z_{(i)} - \hat{\mu}; \hat{\sigma}) + \frac{(\hat{a} - 1)}{2} \sum_{i=m}^n d_2(\hat{\mu} - z_{(i)}; \hat{\sigma}), \\ (m-1) \log(2) - n\{\psi(\hat{a} + \hat{b}) - \psi(\hat{a})\} = \sum_{i=1}^{m-1} \frac{(z_{(i)} - \hat{\mu})}{\hat{\sigma}} + \sum_{i=m}^n d_1(\hat{\mu} - z_{(i)}; \hat{\sigma})$$

and

$$(n-m+1) \log(2) - n\{\psi(\hat{a} + \hat{b}) - \psi(\hat{b})\} = - \sum_{i=m}^n \frac{(z_{(i)} - \hat{\mu})}{\hat{\sigma}} + \sum_{i=1}^{m-1} d_1(z_{(i)} - \hat{\mu}; \hat{\sigma}).$$

It should be noticed that the usual regularity conditions stated in the theorems concerning the asymptotic normality of the MLE do not hold for the BL distribution when  $\mu \neq 0$ , since the likelihood function is not differentiable with respect to this parameter. Hence, it may not be reasonable to adopt the observed information matrix to obtain the asymptotic variances of the MLEs.

## 9. Application

In this section we shall compare the fits of the BL, Laplace, beta normal (BN), beta standard normal (BStdN) and normal distributions to a real data set. We shall consider the national index of consumer prices (INPC) of Brazil corresponding to health and personal care, produced by the IBGE ([www.ibge.gov.br](http://www.ibge.gov.br)). The period of collection extends from the day 01 to 30 of the reference month. The INPC measures the cost of living of households with heads as employees. The search is done in the metropolitan regions of Rio de Janeiro, Porto Alegre, Belo Horizonte, Recife, São Paulo, Belém, Fortaleza, Salvador and Curitiba, in addition to Brasília and the city of Goiânia. The data (in percentages) consisting of 106 observations can be obtained from <http://www.ibge.gov.br>. All the computations were done using the Ox matrix programming language (Doornik, 2006). Ox is freely distributed for academic purposes and available at <http://www.doornik.com>.

Table 1 lists the MLEs of the parameters (standard errors in parentheses), the values of the log-likelihood functions and the statistics AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). The BL distribution yields the highest value of the log-likelihood function and smallest values of the AIC and BIC statistics. From the values of these statistics, we can conclude that the BL model is better than the other distributions to fit these data.

Plots of the estimated pdf and cdf of all fitted models are given in Fig. 2. It is evident that the BL model provides a better fit than the other models. In particular, the estimated cdf shows that the BL model provides an excellent fit to these data, whereas the other models provide poor fits.

In what follows we shall apply formal goodness-of-fit tests in order to verify which distribution fits better to these data. We apply the Cramér–von Mises ( $W^*$ ) and Anderson–Darling ( $A^*$ ) test statistics. The test statistics  $W^*$  and  $A^*$  are described in detail in Chen and Balakrishnan (1995). In general, the smaller the values of the statistics  $W^*$  and  $A^*$ , the better the fit to the data. Additionally, from the critical values of the statistics  $W^*$  and  $A^*$  given in Chen and Balakrishnan (1995), it is possible to calculate the  $p$ -values corresponding to each test statistic. The null hypothesis is  $\mathcal{H}_0 : X_1, \dots, X_n$  and comes from a continuous distribution with cdf  $F(x; \theta)$ , where the form of  $F$  is known but  $\theta$  (a  $k$ -dimensional parameter vector, say) is unknown. To test  $\mathcal{H}_0$ , we can proceed as follows:

1. Compute  $v_i = F(x_i; \hat{\theta})$ , where the  $x_i$ 's are in ascending order;
2. Compute  $y_i = \Phi^{-1}(v_i)$ , where  $\Phi(\cdot)$  is the standard normal cdf and  $\Phi^{-1}(\cdot)$  its inverse;
3. Compute  $u_i = \Phi\{(y_i - \bar{y})/s_y\}$ , where  $\bar{y} = (1/n) \sum_{i=1}^n y_i$  and  $s_y^2 = (n-1)^{-1} \sum_{i=1}^n (y_i - \bar{y})^2$ ;
4. Calculate

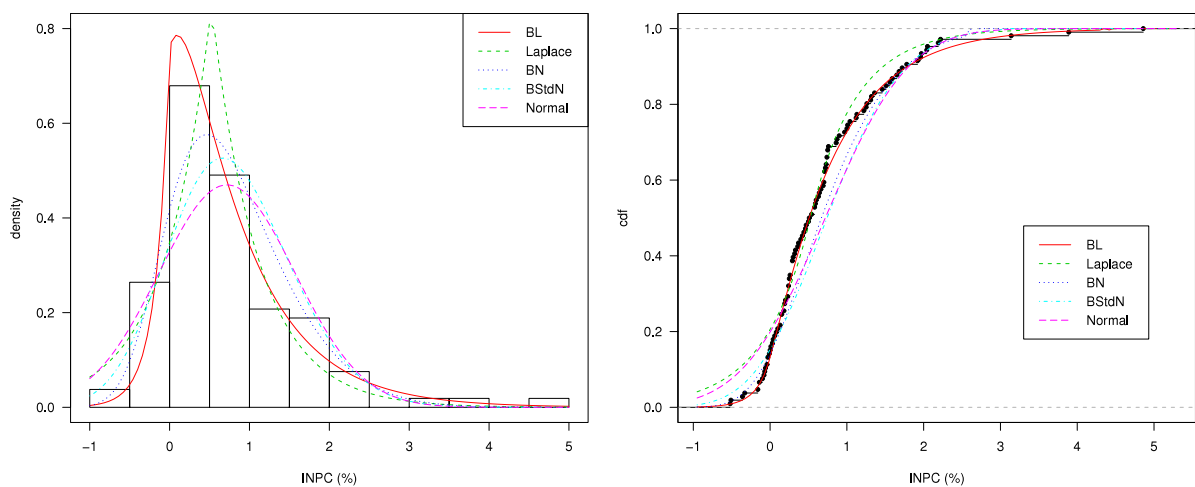
$$W^2 = \sum_{i=1}^n \left\{ u_i - \frac{(2i-1)}{2n} \right\}^2 + \frac{1}{12n}$$



**Table 1**

MLEs of the model parameters, log-likelihood functions and AIC and BIC statistics.

Distribution	Estimates				$\ell(\hat{\theta})$	AIC	BIC
	$\mu$	$\sigma$	$a$	$b$			
BL	0	0.3503 (0.1426)	1.7227 (0.6207)	0.4474 (0.2248)	−109.05	224.10	232.09
Laplace	0.5250 (0.0405)	0.5897 (0.0573)	1	1	−123.50	251.00	256.33
BN	−0.9576 (0.3484)	0.7129 (0.1371)	7.8454 (0.8696)	0.3181 (0.2534)	−116.85	241.70	252.35
BStdN	0	0.6621 (0.0007)	1.5787 (0.0157)	0.4527 (0.0004)	−126.20	258.40	266.39
Normal	0.7186 (0.0698)	0.8490 (0.0494)	1	1	−133.05	270.10	275.43

**Fig. 2.** Estimated pdf and cdf of the BL, Laplace, BN, BStdN and normal distributions for the real data set.**Table 2**

Goodness-of-fit tests.

Distribution	Statistics	
	$W^*$	$A^*$
BL	0.035 ( $p$ -value > 0.50)	0.267 ( $p$ -value > 0.50)
Laplace	0.307 ( $p$ -value < 0.01)	1.862 ( $p$ -value < 0.01)
BN	0.502 ( $p$ -value < 0.01)	3.313 ( $p$ -value < 0.01)
BStdN	0.819 ( $p$ -value < 0.01)	5.057 ( $p$ -value < 0.01)
Normal	0.634 ( $p$ -value < 0.01)	3.661 ( $p$ -value < 0.01)

and

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \{(2i-1) \log(u_i) + (2n+1-2i) \log(1-u_i)\};$$

5. Modify  $W^2$  into  $W^* = W^2(1 + 0.5/n)$  and  $A^2$  into  $A^* = A^2(1 + 0.75/n + 2.25/n^2)$ . Reject  $\mathcal{H}_0$  at the significance level  $\alpha$  if the modified statistics exceed the upper tail significance points given in Table 1 of Chen and Balakrishnan (1995). For further details the reader is referred to these authors.

The values of the statistics  $W^*$  and  $A^*$  ( $p$ -values between parentheses) for all models are given in Table 2. From this table, we conclude that the null hypothesis is not rejected for the BL distribution, whereas the null hypothesis is strongly rejected at any usual significance level for the other models. Thus, according to these goodness-of-fit tests, the BL model fits the current data better than the other models. These results illustrate the potentiality of the BL distribution and the necessity of the additional shape parameters.

To estimate the parameters of the BL model by maximum likelihood, we have considered this model with parameter vector  $\theta = (\sigma, a, b)^T$ , i.e.  $\mu = 0$ . Now, we will estimate the BL model with unknown parameter vector  $\theta = (\mu, \sigma, a, b)^T$ . The MLEs of  $\mu, \sigma, a$  and  $b$  are  $\hat{\mu} = -0.03, \hat{\sigma} = 0.3531, \hat{a} = 1.8692$  and  $\hat{b} = 0.4508$ , respectively. Notice that the MLEs of  $\sigma, a$  and  $b$  are very close to the MLEs of these parameters when  $\mu = 0$  (see Table 1). The value of the log-likelihood ratio statistic for testing the null hypothesis  $\mathcal{H}_0 : \mu = 0$  against  $\mathcal{H}_1 : \mu \neq 0$  is 1.302 ( $p$ -value = 0.254), which implies that  $\mathcal{H}_0$  should not be rejected at any usual significance level. Hence, we conclude that  $\mu = 0$  for the current data set.

## 10. Concluding remarks

We introduce a new distribution, the so-called beta Laplace (BL) distribution, which extends the Laplace distribution, and study some of its general mathematical and statistical properties. We provide a mathematical treatment of the new distribution including expansions for the density function, moments, moment generating function, order statistics, mean deviations and Rényi and Shannon entropies. The BL density function can be expressed as an infinite linear combination of Laplace density functions. The same happens for the density function of BL order statistics. Our formulas related with the BL model are manageable, and with the use of modern computer resources with analytic and numerical capabilities, may turn into adequate tools comprising the arsenal of applied statisticians. The estimation of the model parameters is approached by the method of maximum likelihood and the observed information matrix is derived. An application of the BL distribution to real data is given to show that the new distribution provides consistently better fit than other models available in the literature. Finally, a proper analysis of MLE estimation for the 4-parameter model requires further research and it will be discussed in future research.

## Acknowledgements

The authors are grateful to a referee and the Associate Editor for helpful comments and suggestions. We gratefully acknowledge grants from CNPq and FAPESP (Brazil).

## References

- Akinsete, A., Famoye, F., Lee, C., 2008. The beta-pareto distribution. *Statistics* 42, 547–563.
- Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., Tukey, J.W., 1972. *Robust Estimates of Location*. Princeton University Press, Princeton.
- Bagchi, U., Hayya, J.C., Ord, J.K., 1983. The Hermite distribution as a model of demand during lead time for slow-moving items. *Decision Sciences* 14, 447–466.
- Barreto-Souza, W., Santos, A., Cordeiro, G.M., 2010. The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation* 80, 159–172.
- Chen, C., 2002. Tests for the goodness-of-fit of the Laplace distribution. *Communications in Statistics–Simulation and Computation* 1, 159–174.
- Chen, G., Balakrishnan, N., 1995. A general purpose approximate goodness-of-fit test. *Journal of Quality Technology* 27, 154–161.
- Cordeiro, G.M., Lemonte, A.J., 2011. The  $\beta$ -Birnbau–Saunders distribution: an improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis* 55, 1445–1461.
- Dadi, M.I., Marks, R.J., 1987. Detector relative efficiencies in the presence of Laplace noise. *IEEE Transactions on Aerospace and Electronic Systems* 23, 568–582.
- Damsleth, E., El-Shaarawi, A.H., 1989. ARMA models with double-exponentially distributed noise. *Journal of the Royal Statistical Society B* 51, 61–69.
- Doornik, J.A., 2006. *An Object-Oriented Matrix Language–Ox 4*, 5th ed. Timberlake Consultants Press, London.
- Easterling, R.G., 1978. Exponential responses with double exponential measurement error—a model for steam generator inspection. In: *Proceedings of the DOE Statistical Symposium*, US Department of Energy, pp. 90–110.
- Eugene, N., Lee, C., Famoye, F., 2002. Beta-normal distribution and its applications. *Communications in Statistics–Theory and Methods* 31, 497–512.
- Garvan, F., 2002. *The Maple Book*. Chapman and Hall/CRC, London.
- Hoaglin, D.C., Mosteller, F., Tukey, J.W., 1983. *Understanding Robust and Exploratory Data Analysis*. Wiley, New York.
- Hsu, D.A., 1979. Long-tailed distributions for position errors in navigation. *Journal of the Royal Statistical Society C* 28, 62–72.
- Johnson, N., Kotz, S., Balakrishnan, N., 1995. *Continuous Univariate Distributions–Volume 2*, 2nd ed. Wiley, New York.
- Jones, M.C., 2004. Families of distributions arising from distributions of order statistics. *Test* 13, 1–43.
- Kotz, S., Kozubowski, T.J., Podgórski, K., 2001. *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser, Boston.
- Laplace, P.S., 1774. Mémoire sur las probabilité des causes par les évènements. *Mémoires de Mathématique et de Physique* 6, 621–656.
- Lee, C., Famoye, F., Olumolade, O., 2007. Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* 6, 173–186.
- Manly, B.F.J., 1976. Some examples of double exponential fitness functions. *Heredity* 36, 229–234.
- Nadarajah, S., Gupta, A.K., 2004. The beta Fréchet distribution. *Far East Journal of Theoretical Statistics* 14, 15–24.
- Nadarajah, S., Kotz, S., 2004. The beta Gumbel distribution. *Mathematical Problems in Engineering* 10, 323–332.
- Nadarajah, S., Kotz, S., 2006. The beta exponential distribution. *Reliability Engineering and System Safety* 91, 689–697.
- Nocedal, J., Wright, S.J., 1999. *Numerical Optimization*. Springer, New York.
- Pescim, R.R., Demétrio, C.G.B., Cordeiro, G.M., Ortega, E.M.M., Urbano, M.R., 2010. The beta generalized half-normal distribution. *Computational Statistics and Data Analysis* 54, 945–957.
- Press, W.H., Teulosky, S.A., Vetterling, W.T., Flannery, B.P., 2007. *Numerical Recipes in C: The Art of Scientific Computing*, 3rd ed. Cambridge University Press.
- Puig, P., Stephens, M.A., 2000. Tests of fit for the laplace distribution, with applications. *Technometrics* 42, 417–424.
- Sigmon, K., Davis, T.A., 2002. *MATLAB Primer*, 6th ed. Chapman and Hall/CRC.
- Wolfram, S., 2003. *The Mathematica Book*, 5th ed. Cambridge University Press.