Brazilian Journal of Probability and Statistics
2013, Vol. 27, No. 4, 502–543
DOI: 10.1214/11-BJPS179

The beta Burr III model for lifetime data

Antonio E. Gomes[a], **Cibele Q. da-Silva**[a], **Gauss M. Cordeiro**[b] **and Edwin M. M. Ortega**[c]

[a]*Universidade de Brasília*
[b]*Universidade Federal de Pernambuco*
[c]*ESALQ-Universidade de São Paulo*

Abstract. For the first time, the beta Burr III distribution is introduced as an important model for problems in several areas such as actuarial sciences, meteorology, economics, finance, environmental studies, survival analysis and reliability. The new distribution can be expressed as a linear combination of Burr III distributions and then it has tractable properties for the moments, generating and quantile functions, mean deviations, reliability and entropies. The density of its order statistics can be given in terms of an infinite linear combination of Burr III densities. The beta Burr III model is modified for the possibility of long-term survivors. We define a log-beta Burr III regression model to analyze censored data. The estimation of parameters is approached by maximum likelihood and the observed information matrix is derived. The proposed models are applied to three real data sets.

1 Introduction

Burr (1942) introduced a system of distributions which contains the Burr XII (BXII) distribution as the most widely used of these distributions. If a random variable X has the BXII distribution, then X^{-1} has the Burr III (BIII) distribution with cumulative distribution function (c.d.f.) defined (for $x > 0$) by

$$G_{\alpha,\beta,s}(x) = \left[1 + \left(\frac{x}{s}\right)^{-\alpha}\right]^{-\beta} = \left[\frac{(x/s)^{\alpha}}{1 + (x/s)^{\alpha}}\right]^{\beta}, \tag{1.1}$$

where $\alpha > 0$ and $\beta > 0$ are shape parameters and $s > 0$ is a scale parameter. The probability density function (p.d.f.) corresponding to (1.1) is given by

$$g_{\alpha,\beta,s}(x) = \frac{\alpha\beta}{s(x/s)^{\alpha+1}}\left[\frac{(x/s)^{\alpha}}{1 + (x/s)^{\alpha}}\right]^{\beta+1}. \tag{1.2}$$

The BIII distribution has been used in various fields of sciences. In the actuarial literature, it is known as the inverse Burr distribution (see, e.g., Klugman et al., 1998) and as the kappa distribution in the meteorological literature (Mielke, 1973; Mielke and Johnson, 1973). It has also been employed in finance, environmental

Key words and phrases. Beta Burr III distribution, Burr III distribution, exponentiated Burr III distribution, information matrix, maximum likelihood, moment generating function.
Received January 2011; accepted October 2011.

studies, survival analysis and reliability theory (see Sherrick et al., 1996; Lindsay et al., 1996; Al-Dayian, 1999; Shao, 2000; Hose, 2005; Mokhlis, 2005; Gove et al., 2008). Further, Shao et al. (2008) proposed an extended BIII distribution in low-flow frequency analysis where its lower tail is of main interest. A bivariate extension of the BIII distribution was defined by Rodriguez (1980).

The statistics literature is filled with hundreds of continuous univariate distributions. Recent developments focus on new techniques for building meaningful distributions such as the beta generalized class of distributions (Eugene et al., 2002) that has two shape parameters in the generator. Based on this generator, we propose the beta Burr III (BBIII) distribution to accommodate a wide variety of shapes including the BIII distribution. Its two extra shape parameters provide greater flexibility in the form of the generated distribution and, consequently, it is very useful for modeling observed positive data. If G denotes the baseline cumulative function of a random variable, the beta-G distribution (Eugene et al., 2002) is defined by

$$F(x) = I_{G(x)}(a,b) = \frac{1}{B(a,b)} \int_0^{G(x)} w^{a-1}(1-w)^{b-1}\,dw, \tag{1.3}$$

where $a > 0$ and $b > 0$ are two extra shape parameters to the G distribution, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function, $\Gamma(\cdot)$ is the gamma function, $B_y(a,b) = \int_0^y w^{a-1}(1-w)^{b-1}\,dw$ is the incomplete beta function and $I_y(a,b) = B_y(a,b)/B(a,b)$ is the incomplete beta function ratio.

The class of generalized distributions (1.3) has been receiving considerable attention over the last years, in particular after the works of Eugene et al. (2002) and Jones (2004). In fact, following this idea, Eugene et al. (2002), Nadarajah and Kotz (2004), Nadarajah and Gupta (2004), Nadarajah and Kotz (2006), Lee et al. (2007), Akinsete et al. (2008), Barreto-Souza et al. (2010), Fischer and Vaughan (2010), Pescim et al. (2010), Silva et al. (2010), Khan (2010), Paranaíba et al. (2011) and Cordeiro and Lemonte (2011a, 2011b) proposed the beta normal, beta Gumbel, beta Fréchet, beta exponential (BE), beta Weibull, beta Pareto, beta exponentiated exponential (BEE), beta hyperbolic secant, beta generalized half-normal, beta modified Weibull, beta inverse Weibull, beta Burr XII (BBXII), beta Birnbaum–Saunders and beta Laplace distributions by taking $G(x)$ in (1.3) to be the c.d.f. of the normal, Gumbel, Fréchet, exponential, modified Weibull, Pareto, exponentiated exponential (EE), hyperbolic secant, generalized half-normal, modified Weibull, inverse Weibull, BXII, Birnbaum–Saunders and Laplace distributions, respectively. The cumulative function (1.3) can be expressed as

$$F(x) = \frac{G(x)^a}{aB(a,b)} {}_2F_1\bigl(a, 1-b; a+1; G(x)\bigr),$$

where

$${}_2F_1(p,q;r;y) = \sum_{j=0}^{\infty} \frac{(p)_j (q)_j}{(r)_j} \frac{y^j}{j!}$$

is the hypergeometric function and $(p)_j$ is the Pochhammer symbol defined as $(p)_j = p(p+1)\cdots(p+j-1)$. Thus, for any parent $G(x)$, the properties of $F(x)$ could, in principle, be obtained from the well established properties of the hypergeometric function (Gradshteyn and Ryzhik, 2000, Section 9.1).

The density function corresponding to (1.3) has the form

$$f(x) = \frac{1}{B(a,b)} G(x)^{a-1}[1-G(x)]^{b-1} g(x), \tag{1.4}$$

which will be most tractable when $G(x)$ and $g(x) = dG(x)/dx$ have simple analytic expressions. Except for some special choices for $G(x)$ in (1.4), such as the case given by (1.1), the density function $f(x)$ will be difficult to deal with in generality.

The five parameter BBIII distribution is defined from (1.3) by taking $G(x)$ to be the c.d.f. (1.1). Its cumulative distribution becomes

$$F(x) = I_{[1+(x/s)^{-\alpha}]^{-\beta}}(a,b). \tag{1.5}$$

Here, we have four positive shape parameters α, β, a and b and a positive scale parameter s. The p.d.f. and the hazard rate function corresponding to (1.5) (for $x > 0$) are

$$f(x) = \frac{\alpha\beta}{s(x/s)^{\alpha+1}B(a,b)} \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta a+1} \left\{1 - \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^\beta\right\}^{b-1} \tag{1.6}$$

and

$$\begin{aligned} h(x) = {} & \frac{\beta\alpha[s(x/s)^{-\alpha-1}]}{B(a,b) I_{1-[1+(x/s)^{-\alpha}]^{-\beta}}(b,a)} \\ & \times \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta a+1} \left\{1 - \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^\beta\right\}^{b-1}, \end{aligned} \tag{1.7}$$

respectively. The BBIII density function (1.6) allows for greater flexibility of its tails with a continuous crossover towards distributions with different shapes (e.g., a particular combination of skewness and kurtosis). It includes three important submodels: the BIII distribution arises immediately for $a = b = 1$, the exponentiated Burr III (EBIII) distribution for $b = 1$, and the Lehmann type II Burr III (LeBIII), which is also the Kumaraswamy Burr III (KwBIII) distribution for $a = 1$. The EBIII and LeBIII models have not been studied in the literature yet. Moreover, while the transformation (1.3) is not analytically tractable in the general case, the formulas related with the BBIII distribution turn out manageable as it is shown in the rest of the article.

Plots of the density function (1.6) for selected parameter values are given in Figures 1–4, respectively. Figure 5 shows that the BBIII failure rate function can be bathtub shaped, monotonically decreasing or increasing and upside-down bathtub depending on its parameter values.

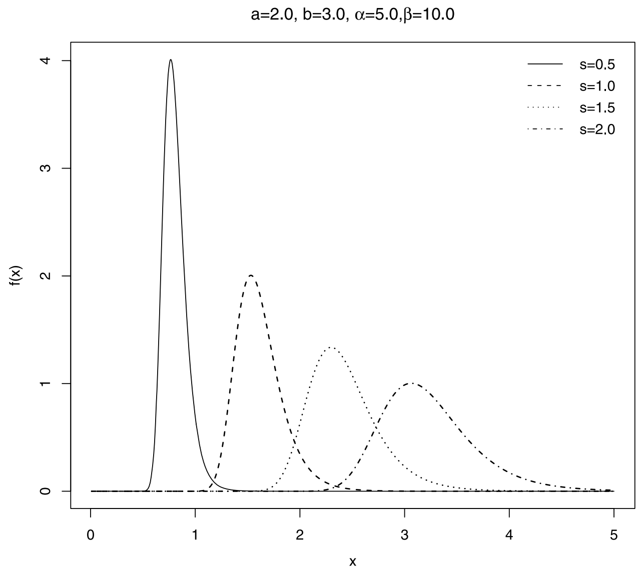

Figure 1 *Plots of the BBIII density function for some values of the scale parameter* s.

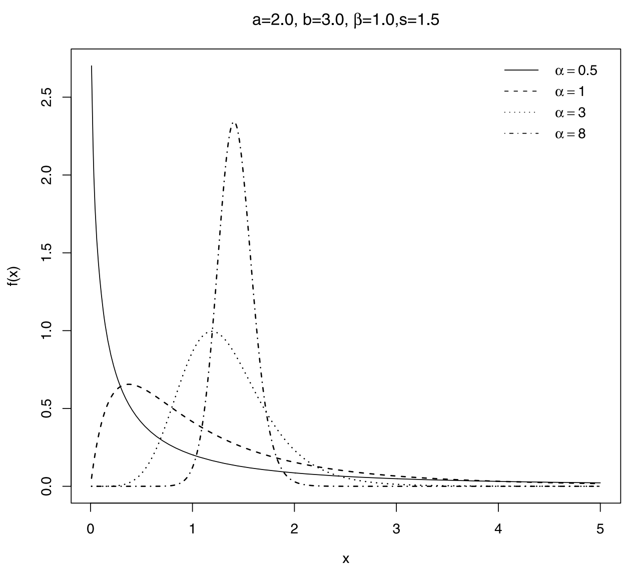

Figure 2 *Plots of the BBIII density function for some values of the shape parameter* α.

The rest of the article is organized as follows. In Section 2, we demonstrate that the new density function can be expressed as a linear combination of BIII densities. This result is important to derive some BBIII mathematical quantities immediately from those quantities of the BIII distribution. A range of the properties is

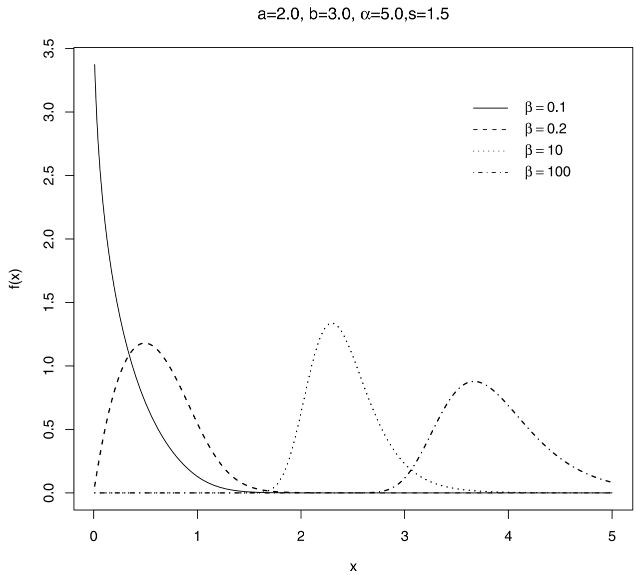

Figure 3 *Plots of the BBIII density function for some values of the shape parameter* β.

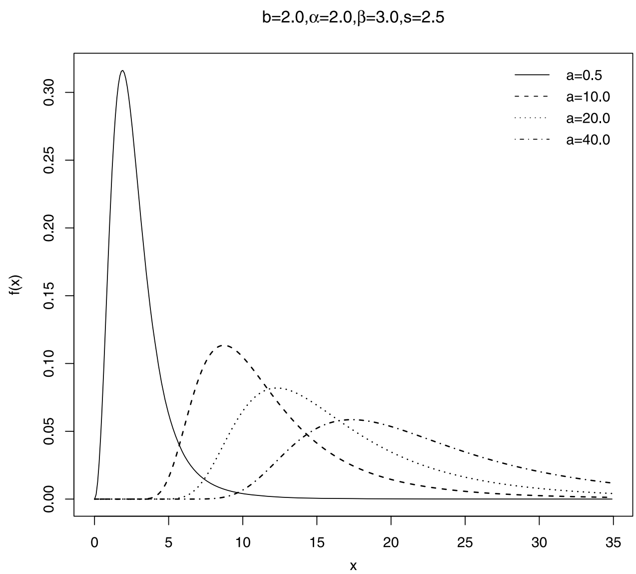

Figure 4 *Plots of the BBIII density function for some values of the shape parameter* a.

considered in Sections 3–6. These include generating and quantile functions, simulation, Bowley skewness and Moors kurtosis, mean deviations and Bonferroni and Lorenz curves. In Section 7, we demonstrate that the density function of the BBIII

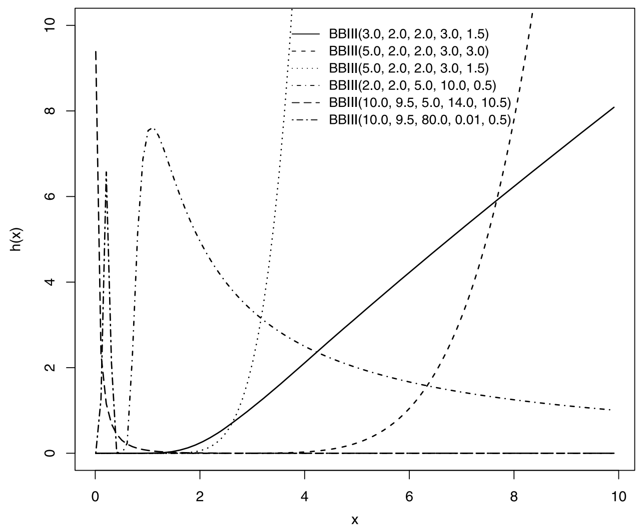

Figure 5 *Plots of the BBIII hazard rate function.*

order statistics can be expressed as a linear combination of BIII densities. Explicit formulae for the moments of the order statistics and L-moments are obtained in Section 8. The reliability and the Rényi and Shannon entropies are determined in Sections 9 and 10, respectively. Maximum likelihood estimation is investigated in Section 11. In Section 12, we define a BBIII model for survival data with long-term survivors. An useful log-beta Burr III regression model for lifetime analysis in proposed in Section 13. Applications of the proposed models to three real life data sets are given in Section 14. Finally, some conclusions are noted in Section 15.

2 Expansion for the density function

Equations (1.5) and (1.6) are straightforward to compute with the use of modern computer resources with analytic and numerical capabilities. However, we obtain expansions for $F(x)$ and $f(x)$ in terms of an infinite (or finite) weighted sums of c.d.f.'s and p.d.f.'s of BIII distributions, respectively. Here and henceforth, denote by X a random variable with density function (1.6), say $X \sim \mathrm{BBIII}(a, b, \alpha, \beta, s)$. First, for $b > 0$ real noninteger, we can write the density function of X as

$$\int_0^x w^{a-1}(1-w)^{b-1}\,dw = \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b-1}{j}}{(a+j)} x^{a+j},$$

where the binomial term $\binom{b-1}{j} = \Gamma(b)/[\Gamma(b-j)j!]$ is defined for any real. From (1.5), after some algebra, we have

$$F(x) = \sum_{j=0}^{\infty} w_j G_{\alpha,\beta(a+j),s}(x), \tag{2.1}$$

where

$$w_j = \frac{(-1)^j \binom{b-1}{j}}{(a+j)B(a,b)}.$$

By differentiating (2.1), the density function of X can be expressed as

$$f(x) = \sum_{j=0}^{\infty} w_j g_{\alpha,\beta(a+j),s}(x), \tag{2.2}$$

which holds for any parameter values. If $b > 0$ is an integer, the index j in the sum stops at $b-1$. From the linear combination (2.2), we can obtain some BBIII structural properties. For example, the ordinary, central, inverse and factorial moments of X can be expressed as linear functions of those BIII quantities. In fact, the rth moment of the BIII distribution (with parameters α, β and s) is given by $\mu_r' = \beta s^r B(\beta + r\alpha^{-1}, 1 - r\alpha^{-1})$, and the rth moment of X follows from (2.2) as

$$E(X^r) = \beta s^r \sum_{j=0}^{\infty} (a+j) w_j B\big(\beta(a+j) + r\alpha^{-1}, 1 - r\alpha^{-1}\big). \tag{2.3}$$

For $a = b = 1$, equation (2.3) yields the rth moment of the BIII distribution. From this formula, we can obtain the skewness and kurtosis of X using well-known relationships. Figures 6–9 show great flexibility for the values of these measures.

3 Quantile function

The BBIII quantile function, say $Q(u) = F^{-1}(u)$, is straightforward to be computed from the beta quantile function ($Q_{\beta(a,b)}(u)$) by inverting (1.5). We have

$$x = Q(u) = F^{-1}(u) = s\big(Q_{\beta(a,b)}(u)^{-1/\beta} - 1\big)^{-1/\alpha}. \tag{3.1}$$

So, the simulation of the BBIII random variables is straightforward from (3.1) as

$$X = s\big(V^{-1/\beta} - 1\big)^{-1/\alpha},$$

where V is a beta variate with shape parameters a and b.

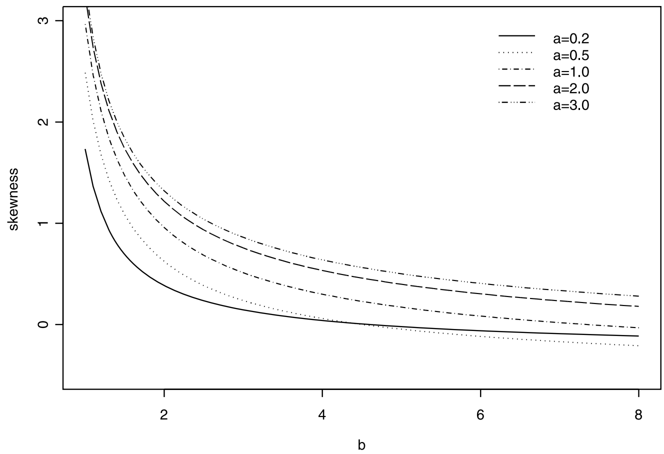

Figure 6 *Skewness of the BBIII distribution as a function of b for some values of a and* $\alpha = 5$, $\beta = 2$ *and* $s = 1$.

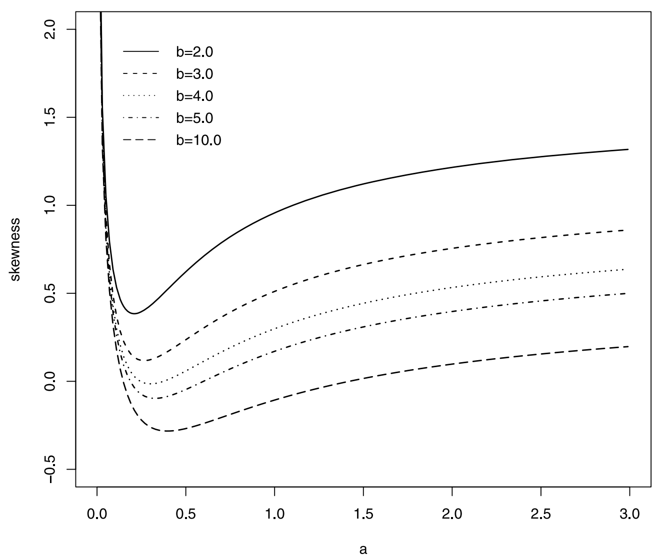

Figure 7 *Skewness of the BBIII distribution as a function of a for some values of b and* $\alpha = 5$, $\beta = 2$ *and* $s = 1$.

4 Quantile measures

The effect of the shape parameters a and b on the skewness and kurtosis of the new distribution can be considered based on quantile measures. The shortcomings of the classical skewness and kurtosis measures are well known. One of the earliest skewness measures to be suggested is the Bowley skewness (Kenney and Keeping,

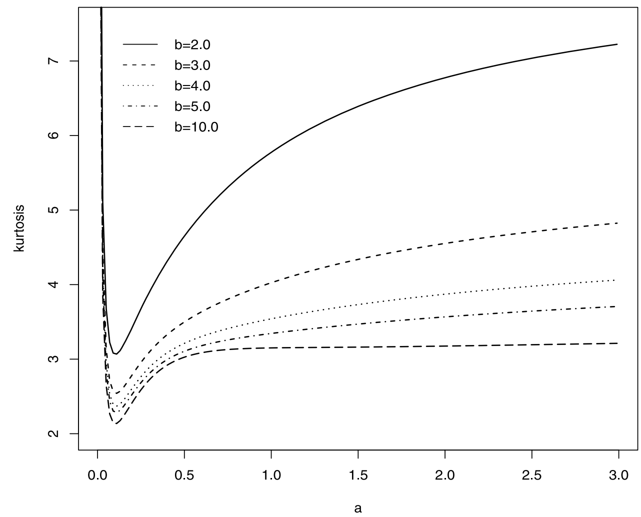

Figure 8 *Kurtosis of the BBIII distribution as a function of a for some values of b and* $\alpha = 5$, $\beta = 2$ *and* $s = 1$.

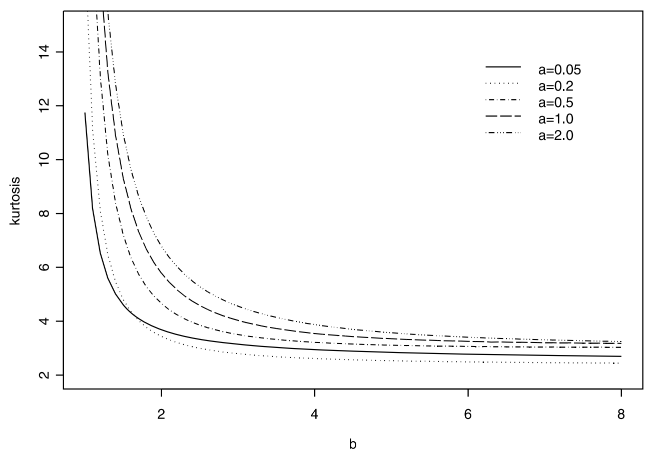

Figure 9 *Kurtosis of the BBIII distribution as a function of b for some values of a and* $\alpha = 5$, $\beta = 2$ *and* $s = 1$.

1962) defined by the average of the quartiles minus the median, divided by half the interquartile range, namely

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

Since only the middle two quartiles are considered and the outer two quartiles are ignored, this adds robustness to the measure.

The Moors kurtosis is based on octiles

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

The measures B and M are less sensitive to outliers and they exist even for distributions without moments. For symmetric unimodal distributions, positive kurtosis indicates heavy tails and peakedness relative to the normal distribution, whereas negative kurtosis indicates light tails and flatness. Because M is based on the octiles, it is not sensitive to variations of the values in the tails or to variations of the values around the median.

The basic justification of M as an alternative measure of kurtosis is the following: keeping $Q(2/8)$ and $Q(6/8)$ fixed, M clearly decreases as $Q(3/8) - Q(1/8)$ and $Q(7/8) - Q(5/8)$ decrease. So, if $Q(3/8) - Q(1/8) \to 0$ and $Q(7/8) - Q(5/8) \to 0$, then $M \to 0$ and half of the total probability mass is concentrated in the neighborhoods of the octiles $Q(2/8)$ and $Q(6/8)$. Clearly, $M > 0$ and there is a good agreement with the usual kurtosis measures for some distributions. For the normal distribution, $B = M = 0$.

In Figures 10 and 11, we plot the measures B and M for the BBIII distribution as functions of a and b for some values of the other parameter, respectively. These

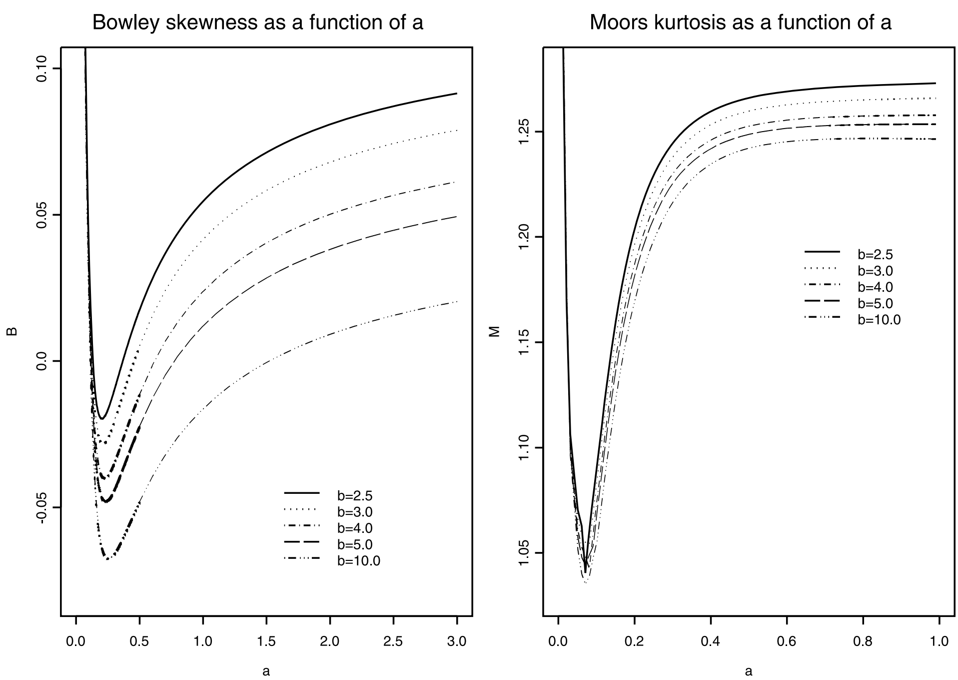

Figure 10 *Bowley skewness and Moors kurtosis of the BBIII distribution as function of a for some values of b and $\alpha = 5$, $\beta = 2$ and $s = 1$.*

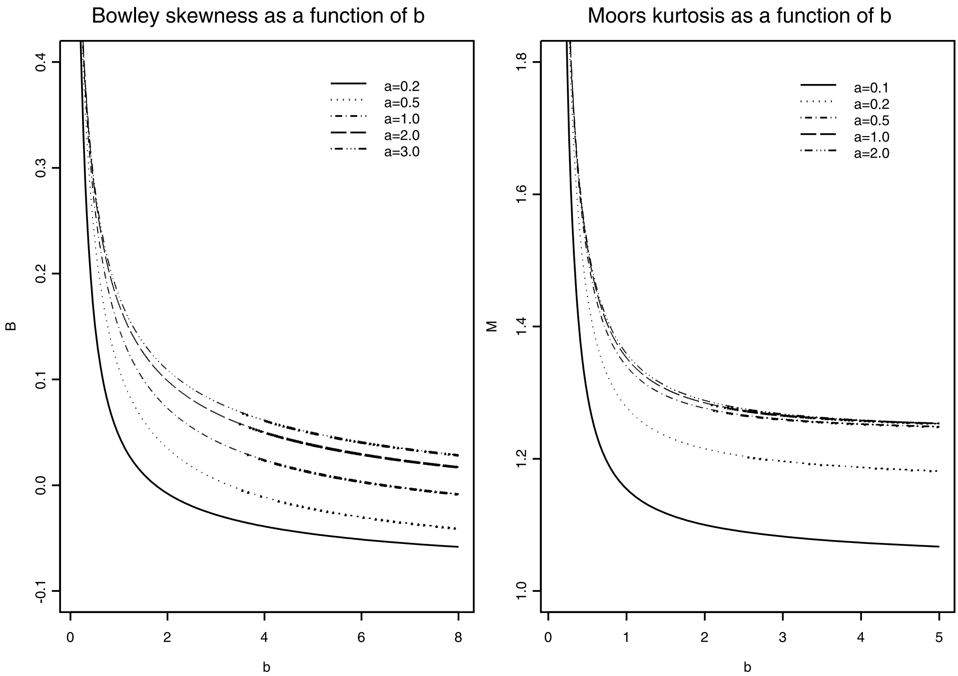

Figure 11 *Bowley skewness and Moors kurtosis of the BBIII distribution as function of b for some values of a and* $\alpha = 5$, $\beta = 2$ *and* $s = 1$.

Table 1 *Values of a for which the Bowley skewness vanishes and achieves its lowest value for different values of b*

b	a_1	a_2	Minimum point
2.5	0.1325	0.3679	0.2038
3.0	0.1265	0.4551	0.2122
4.0	0.1206	0.6186	0.2246
5.0	0.1175	0.7750	0.2332
10.0	0.1119	1.5212	0.2545

plots show that both measures B and M can be very sensitive to these shape parameters, thus indicating the importance of the model (1.6). For fixed b, the Bowley skewness decreases and then increases sharply when $a \to 0$. The value of a corresponding to its minimum value depends on b. For fixed a, this skewness increases when b decreases. The same conclusions apply to the Moors kurtosis. Table 1 lists the values of a in Figure 10, say a_1 and a_2, for which the Bowley skewness vanishes and also the value of a which gives its minimum for different values of b. On the other hand, when $b \to 0$ (a fixed), this skewness increases rapidly. For fixed a, the Moors kurtosis decreases when b increases. When $b \to \infty$, this kurtosis tends to an asymptotic level which depends on the value of a. For fixed b, the Moors kurtosis increases when a increases. The values of (a, b) for which the Bowley skewness vanishes can be obtained from Figure 11 as: (0.2, 1.7544), (0.5, 3.2680), (1.0, 6.4805), (2.0, 13.2837) and (3.0, 20.1986).

5 Generating function

The moment generating function (m.g.f.) of X, say $M_X(t) = E(\mathrm{e}^{tX})$, can be written as

$$M_X(t) = \sum_{k=0}^{\infty} \frac{E(X^k)}{k!} t^k.$$

Another representation for $M_X(t)$ as an infinite weighted sum can be obtained from (2.2) as

$$M_X(t) = \sum_{j=0}^{\infty} w_j M_j(t), \tag{5.1}$$

where $M_j(t)$ denotes the m.g.f. of the BIII$(\alpha, \beta(a+j), s)$ distribution. Now, we provide a simple representation for the m.g.f. of the BIII(α, β, s) distribution, namely

$$M_{\mathrm{BIII}}(t) = s\alpha\beta \int_0^{\infty} \exp(syt) y^{\beta\alpha-1} (1 + y^{\alpha})^{-(\beta+1)} \, \mathrm{d}y.$$

First, we require the Meijer G-function defined by

$$G_{p,q}^{m,n}\left(x \middle| \begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_p \end{matrix}\right) = \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^{m} \Gamma(b_j + t) \prod_{j=1}^{n} \Gamma(1 - a_j - t)}{\prod_{j=n+1}^{p} \Gamma(a_j + t) \prod_{j=m+1}^{p} \Gamma(1 - b_j - t)} x^{-t} \, \mathrm{d}t,$$

where $i = \sqrt{-1}$ is the complex unit and L denotes an integration path; see Gradshteyn and Ryzhik (2000, Section 9.3) for a description of this path. The Meijer G-function contains as particular cases many integrals with elementary and special functions (Prudnikov et al., 1986).

Further, we assume that $\alpha = m/k$, where m and k are positive integers. This condition is not restrictive since every positive real number can be approximated by a rational number. For $t < 0$, using the integral (A.1) given in Appendix A, we obtain

$$M_{\mathrm{BIII}}(t) = \frac{\beta s m}{k} I'\left(-st, \beta\frac{m}{k} - 1, \frac{m}{k}, -\beta - 1\right).$$

Hence, from equation (5.1), we can write $M_X(t)$ for $t < 0$ as

$$M_X(t) = \frac{\beta s m}{k} \sum_{j=0}^{\infty} w_j I'\left(-st, \beta(a+j)\frac{m}{k} - 1, \frac{m}{k}, -\beta(a+j) - 1\right). \tag{5.2}$$

Equation (5.2) is the main result of this section.

6 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. If X has the BBIII distribution, we can derive the mean deviations about the mean $\mu = E(X)$ and about the median M from

$$\delta_1 = \int_0^\infty |x-\mu| f(x)\,dx \quad \text{and} \quad \delta_2 = \int_0^\infty |x-M| f(x)\,dx,$$

respectively. The mean μ is obtained from (2.3) with $r=1$ and the median M is the solution of the nonlinear equation $I_{1-[1+(M/s)^{-\alpha}]^{-\beta}}(a,b) = 1/2$.

These measures can be calculated by the following relationships

$$\delta_1 = 2\mu F(\mu) - 2\mu + 2T(\mu) \quad \text{and} \quad \delta_2 = 2T(M) - \mu, \tag{6.1}$$

where $T(a) = \int_a^\infty x f(x)\,dx$ follows from (2.2) as

$$T(q) = \frac{1}{B(a,b)} \sum_{j=0}^\infty w_j \left[E(X_j) - \int_0^q x g_{\alpha,\beta(a+j),s}(x)\,dx \right], \tag{6.2}$$

where X_j has the BIII($\alpha, \beta(a+j), s$) distribution. Let ${}_2F_1$ be the hypergeometric function defined in Section 1. Setting $u = (x/s)^{-\alpha}$, we obtain using Maple

$$\begin{aligned}\int_0^q x g_{\alpha,\beta(a+j),s}(x)\,dx &= s\beta(a+j) \int_{(q/s)^{-\alpha}}^\infty u^{-1/\alpha}(1+u)^{-[\beta(a+j)+1]}\,du \\ &= s\beta(a+j) U_j,\end{aligned}$$

where $U_j = U_j(q/s, \alpha, \beta)$ is given by

$$\begin{aligned}U_j = &-\frac{{}_2F_1[\beta(a+j)+1, -\alpha^{-1}+1; (2-\alpha^{-1}); -(q/s)^{-\alpha}](q/s)^{1-\alpha}}{1-\alpha^{-1}} \\ &- \frac{\Gamma[\beta(a+j)+\alpha^{-1}]\pi \csc(-\pi/\alpha)}{\Gamma[\beta(a+j)+1]\Gamma(\alpha^{-1})}.\end{aligned}$$

Hence,

$$T(q) = \frac{s\beta}{B(a,b)} \sum_{j=0}^\infty (-1)^j \binom{b-1}{j} \left\{ B\left[\beta(a+j)+\alpha^{-1}, 1-\alpha^{-1}\right] + U_j \right\}.$$

The quantity $T(q)$ can also be used to determine Bonferroni and Lorenz curves which have applications in economics, demography, income, poverty, reliability, insurance and medicine. They are defined by

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x)\,dx \quad \text{and} \quad L(p) = \frac{1}{\mu} \int_0^q x f(x)\,dx,$$

respectively, where $q = Q(p) = F^{-1}(p)$ is the quantile function (given in Section 3) for a given probability p. From $\int_0^q x f(x) = \mu - T(q)$, we have

$$B(p) = \frac{1}{p} - \frac{T(q)}{p\mu} \quad \text{and} \quad L(p) = 1 - \frac{T(q)}{\mu}.$$

7 Order statistics

The density function $f_{i:n}(x)$ of the ith order statistic for $i = 1, \ldots, n$ from data values $X_1, \ldots, X_n$ having the BBIII distribution can be expressed as

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

where $F(\cdot)$ is the c.d.f. (1.5) and $f(\cdot)$ is the p.d.f. (1.6). Then

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} \frac{\beta\alpha}{s B(a,b)} \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta a+1} \left\{1 - \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^\beta\right\}^{b-1}$$
$$\times \frac{1}{(x/s)^{\alpha+1}} \sum_{l=0}^{n-i} (-1)^l \binom{n-i}{l} \left\{\sum_{j=0}^{\infty} w_j \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta(a+j)}\right\}^{i+l-1}.$$

We use an equation of Gradshteyn and Ryzhik (2000, Section 0.314) for a power series raised to a positive integer k given by

$$\left(\sum_{j=0}^{\infty} a_j u^j\right)^k = \sum_{j=0}^{\infty} c_{k,j} u^j, \tag{7.1}$$

where the coefficients $c_{k,j}$ (for $k = 1, 2, \ldots$) can be determined from the recurrence equation

$$c_{k,j} = (j a_0)^{-1} \sum_{m=1}^{j} [m(k+1) - j] a_m c_{k,j-m} \tag{7.2}$$

and $c_{k,0} = a_0^k$. Hence, $c_{k,j}$ follows directly from $c_{k,0}, \ldots, c_{k,j-1}$ and, therefore, from $a_0, \ldots, a_k$. We can obtain from (7.1)

$$\left\{\sum_{j=0}^{\infty} w_j \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta(a+j)}\right\}^{i+l-1}$$
$$= \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta a(i+l-1)} \sum_{j=0}^{\infty} c_{i+l-1,j} \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta j},$$

where $c_{i+l-1,0} = w_0^{i+l-1}$ and

$$c_{i+l-1,j} = \frac{a\Gamma(b)}{j} \sum_{m=1}^{j} \frac{(-1)^m[(i+l)m - j]c_{i+l-1,j-m}}{(a+m)m!\Gamma(b-m)}.$$

Thus, setting $a^*_{i,j,l} = a(i+l) + j$ and after some algebra, we have

$$\begin{aligned} f_{i:n}(x) = {} & \frac{1}{B(a,b)B(i,n-i+1)} \\ & \times \sum_{l=0}^{n-i} (-1)^l \binom{n-i}{l} \sum_{j=0}^{\infty} c_{i+l-1,j} B(a^*_{i,j,l}, b) f_{\theta_{i,j,l}}(x), \end{aligned} \tag{7.3}$$

where $f_{\theta_{i,j,l}}(x)$ denotes the BBIII density function defined by the parameter vector $\theta_{i,j,l} = (a^*_{i,j,l}, b, \alpha, \beta, s)^T$. Equation (7.3) is an important result in applications since it gives the density function of the BBIII order statistics as a linear combination of BBIII density functions. Several mathematical properties for the BBIII order statistics (m.g.f., ordinary, inverse and factorial moments) can be derived from this representation form.

8 Moments of order statistics and L-moments

The moments of the BBIII order statistics can be written directly in terms of the moments of BIII distributions from (2.3) and (7.3). We have

$$\begin{aligned} E(X^r_{i:n}) = {} & \frac{\beta s^r \Gamma(a+b)}{B(i,n-i+1)\Gamma(a)} \\ & \times \sum_{l=0}^{n-i} (-1)^l \binom{n-i}{l} \sum_{j,k=0}^{\infty} \frac{(-1)^k c_{i+l-1,j}}{\Gamma(b-k)k!} \\ & \times B\big[\beta\big(a^*_{i,j,l} + k\big) + r/\alpha, 1 - r/\alpha\big], \end{aligned} \tag{8.1}$$

where $a^*_{i,k,l}$ is defined in Section 7.

L-moments (Hoskings, 1990) are summary statistics for probability distributions and data samples but have several advantages over ordinary moments. For example, they apply for any distribution having finite mean and no higher-order moments need be finite. The rth L-moment is computed from linear combinations of the ordered data values by

$$\lambda_r = \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} \binom{r-1+j}{j} \beta_j,$$

where $\beta_j = E\{XF(X)^j\}$. In particular, $\lambda_1 = \beta_0$, $\lambda_2 = 2\beta_1 - \beta_0$, $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$ and $\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$. In general, $\beta_r = (r+1)^{-1}E(X_{r+1:r+1})$, so λ_r can be computed from equation (8.1).

9 Reliability

In the context of reliability, the stress-strength model describes the life of a component which has a random strength X_1 that is subjected to a random stress X_2. The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $X_1 > X_2$. Hence, $R = P(X_2 < X_1)$ is a measure of component reliability (see Kotz et al., 2003). It has many applications especially in the area of engineering. We derive the reliability R when X_1 and X_2 have independent BBIII$(a_1, b_1, \alpha, \beta_1, s)$ and BBIII$(a_2, b_2, \alpha, \beta_2, s)$ distributions with the same shape parameter α and scale parameter s. From equations (1.6) and (2.1), the reliability reduces to

$$
\begin{aligned}
R = P(X_1 > X_2) &= \int_0^\infty f_1(x) F_2(x)\, dx \\
&= \frac{\beta_1 \alpha}{s B(a_1, b_1) B(a_2, b_2)} \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b_2-1}{j}}{a_2 + j} \\
&\quad \times \int_0^\infty \left(\frac{x}{s}\right)^{-(\alpha+1)} \left[1 + \left(\frac{x}{s}\right)^{-\alpha}\right]^{-[\beta_1 a_1 + \beta_2 (a_2 + j) + 1]} \\
&\quad \times \left\{1 - \left[1 + \left(\frac{x}{s}\right)^{-\alpha}\right]^{-\beta_1}\right\}^{b_1 - 1} dx.
\end{aligned}
$$

Setting $z = [1 + (x/s)^{-\alpha}]^{-\beta_1}$, we obtain

$$
R = \frac{1}{B(a_1, b_1) B(a_2, b_2)} \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b_2-1}{j}}{a_2 + j} B\big[a_1 + \beta_2 \beta_1^{-1} (a_2 + j), b_1\big].
$$

From equations (2.1) and (2.2), an alternative expression for R follows as

$$
\begin{aligned}
R &= \sum_{j,k=0}^{\infty} w_j^{(1)} w_k^{(2)} \int_0^\infty \frac{\beta_1 (a_1 + j) \alpha}{s (x/s)^{\alpha+1}} \left[\frac{(x/s)^\alpha}{1 + (x/s)^\alpha}\right]^{\beta_1 (a_1 + j) + \beta_2 (a_2 + k) + 1} dx \\
&= \frac{\beta_1}{B(a_1, b_1) B(a_2, b_2)} \sum_{j,k=0}^{\infty} \frac{(-1)^{j+k} \binom{b_1-1}{j} \binom{b_2-1}{k}}{(a_2 + k)[\beta_1 (a_1 + j) + \beta_2 (a_2 + k)]},
\end{aligned}
$$

where $w_m^{(l)} = (-1)^m \binom{b_l-1}{m} [B(a_l, b_l)(a_l + m)]^{-1}$, for $l = 1, 2$.

10 Entropies

The entropy of a random variable X with density $f(x)$ is a measure of variation of the uncertainty. A large value of the entropy indicates the greater uncertainty in the data. The Rényi entropy is defined by

$$I_R(\rho) = \frac{1}{1-\rho}\log\left(\int f(x)^\rho\,dx\right),$$

where $\rho > 0$ and $\rho \neq 1$. Setting $y = 1/[1 + (x/s)^\alpha]$, the integral L in $I_R(\rho)$ for the BBIII distribution becomes

$$\begin{aligned}
L &= \left(\frac{\beta\alpha}{sB(a,b)}\right)^\rho \int_0^\infty \frac{1}{(x/s)^{\rho(\alpha+1)}}\left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\rho(\beta a+1)} \\
&\qquad \times \left\{1 - \left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^\beta\right\}^{\rho(b-1)} dx \\
&= \left(\frac{\alpha}{s}\right)^{\rho-1}\left(\frac{\beta}{B(a,b)}\right)^\rho \int_0^1 y^{\rho(1+1/\alpha)-1/\alpha-1} \\
&\qquad \times (1-y)^{\rho(\beta a-1/\alpha)+1/\alpha-1}\big[1-(1-y)^\beta\big]^{\rho(b-1)}\,dy \\
&= \frac{s}{\alpha}\left(\frac{\alpha\beta}{sB(a,b)}\right)^\rho \sum_{j=0}^\infty (-1)^j \binom{\rho(b-1)}{j} \\
&\qquad \times B\big[\rho(1+\alpha^{-1}) - \alpha^{-1}, \rho(\beta a - \alpha^{-1}) + \alpha^{-1} + j\beta\big].
\end{aligned}$$

Thus,

$$\begin{aligned}
I_R(\rho) &= \frac{1}{1-\rho}\log\left\{C\sum_{j=0}^\infty (-1)^j \binom{\rho(b-1)}{j}\right. \\
&\qquad \left.\times B\big[\rho(1+\alpha^{-1}) - \alpha^{-1}\rho(\beta a - \alpha^{-1}) + \alpha^{-1} + j\beta\big]\right\},
\end{aligned}$$

where

$$C = \frac{s}{\alpha}\left(\frac{\beta\alpha}{sB(a,b)}\right)^\rho.$$

The Shannon entropy is given by

$$\begin{aligned}
&E\{-\log[f(X)]\} \\
&\quad = -\log\big[\Gamma(a+b)\big] + \log\big[\Gamma(a)\big] + \log\big[\Gamma(b)\big] - \log(\beta) - \log(\alpha) \\
&\qquad - (\beta\alpha a - 1)E\big[\log(X)\big] + \beta\alpha a\log(s) - (\beta a + 1)E\big\{\log[1+(X/s)^\alpha]\big\} \\
&\qquad - (b-1)E\left\{\log\left[1 - \left(\frac{(X/s)^\alpha}{1+(X/s)^\alpha}\right)^\beta\right]\right\}.
\end{aligned}$$

Setting $y = (x/s)^\alpha[1 + (x/s)^\alpha]^{-1}$, we have

$$E\big[\log(X)\big] = \frac{\beta}{B(a,b)} \int_0^1 \left[\log(s) + \frac{1}{\alpha}\log\left(\frac{y}{1-y}\right)\right] y^{\beta a+1}(1-y^\beta)^{b-1}\,dy$$

$$= \frac{\beta}{B(a,b)} \int_0^1 \left[\log(s) + \frac{1}{\alpha}\log\left(\frac{y}{1-y}\right)\right] y^{\beta a+1} \times \sum_{j=0}^{\infty}(-1)^j \binom{b-1}{j} y^{\beta j}\,dy$$

$$= \frac{\beta}{B(a,b)} \Bigg\{ \log(s) \sum_{j=0}^{\infty}(-1)^j \binom{b-1}{j} \int_0^1 y^{\beta(a+j)+1}\,dy + \frac{1}{\alpha}\sum_{j=0}^{\infty}(-1)^j \binom{b-1}{j} \times \left[\int_0^1 \log(y) y^{\beta(a+j)+1}\,dy - \int_0^1 \log(1-y) y^{\beta(a+j)+1}\,dy\right]\Bigg\}.$$

The last three integrals are easily calculated. The first one is equal to $[\beta(a+j)+2]^{-1}$. Setting $z = -\log(y)$, the second integral becomes

$$\int_0^1 \log(y) y^{\beta(a+j)+1}\,dy = -\frac{1}{[\beta(a+j)+2]^2},$$

and using the expansion $\log(1-y) = -\sum_{k=1}^{\infty} y^k/k$, the last integral reduces to

$$\int_0^1 \log(1-y) y^{\beta(a+j)+1}\,dy = -\sum_{k=1}^{\infty} \frac{1}{k[\beta(a+j)+k+2]}.$$

Since the expected values of the score functions vanish, equations (11.1) and (11.2) (given in the next section) yield

$$E\big\{\log\big[1 + (X/s)^\alpha\big]\big\} = \frac{\psi(a+b) - \psi(a)}{\beta} + \alpha\big\{E\big[\log(X)\big] - \log(s)\big\}$$

and

$$E\left\{\log\left[1 - \left(\frac{(X/s)^\alpha}{1+(X/s)^\alpha}\right)^\beta\right]\right\} = \psi(b) - \psi(a+b),$$

where $\psi(z) = \frac{d\log[\Gamma(z)]}{dz} = \frac{\Gamma'(z)}{\Gamma(z)}$ is the digamma function.

Hence,

$$
\begin{aligned}
E\{-\log[f(X)]\}
&= -\log\big[\Gamma(a+b)\big] + \log\big[\Gamma(a)\big] + \log\big[\Gamma(b)\big] \\
&\quad - \log(\beta) - \log(\alpha) + \alpha(a\beta+1)\log(s) \\
&\quad - \frac{\psi(a+b)-\psi(a)}{\beta} - (b-1)\big[\psi(b)-\psi(a+b)\big] \\
&\quad - \frac{\beta[\alpha(a\beta+1)-1]}{B(a,b)} \\
&\quad \times \Bigg\{\log(s)\sum_{j=0}^{\infty}\frac{(-1)^j\binom{b-1}{j}}{\beta(a+j)+2} \\
&\quad - \frac{1}{\alpha}\Bigg[\sum_{j=0}^{\infty}\frac{(-1)^j\binom{b-1}{j}}{[\beta(a+j)+2]^2} \\
&\quad - \sum_{j=0}^{\infty}\sum_{k=1}^{\infty}\frac{(-1)^j\binom{b-1}{j}}{k[\beta(a+j)+k+2]}\Bigg]\Bigg\}.
\end{aligned}
$$

11 Estimation

Let $\boldsymbol{\theta} = (a, b, \alpha, \beta, s)^T$ be the parameter vector of the BBIII distribution (1.6). We consider the method of maximum likelihood to estimate $\boldsymbol{\theta}$. The log-likelihood function for the five parameters from a single observation $x > 0$, say $\ell = \ell(\boldsymbol{\theta})$, is

$$
\begin{aligned}
\ell(\boldsymbol{\theta}) &= \log\big[\Gamma(a+b)\big] - \log\big[\Gamma(a)\big] - \log\big[\Gamma(b)\big] + \log(\beta) + \log(\alpha) \\
&\quad + (a\alpha\beta-1)\log(x) - a\alpha\beta\log(s) - (a\beta+1)\log\big[1+(x/s)^{\alpha}\big] \\
&\quad + (b-1)\log\Bigg\{1-\left[\frac{(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right]^{\beta}\Bigg\}.
\end{aligned}
$$

The components of the unit score vector $\mathbf{U} = (\frac{\partial\ell}{\partial a}, \frac{\partial\ell}{\partial b}, \frac{\partial\ell}{\partial\alpha}, \frac{\partial\ell}{\partial\beta}, \frac{\partial\ell}{\partial s})^T$ are

$$
\frac{\partial\ell}{\partial a} = \psi(a+b) - \psi(a) + \beta\big\{\alpha\log(x/s) - \log\big[1+(x/s)^{\alpha}\big]\big\}, \tag{11.1}
$$

$$
\frac{\partial\ell}{\partial b} = \psi(a+b) - \psi(b) + \log\Bigg\{1-\left[\frac{(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right]^{\beta}\Bigg\}, \tag{11.2}
$$

$$
\frac{\partial\ell}{\partial\alpha} = \frac{1}{\alpha} + \log(x/s)\left[\frac{\beta a-(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right] - \frac{\beta(b-1)(x/s)^{\alpha\beta}\log(x/s)}{[1+(x/s)^{\alpha}]\{[1+(x/s)^{\alpha}]^{\beta}-(x/s)^{\alpha\beta}\}},
$$

$$\frac{\partial \ell}{\partial \beta} = \frac{1}{\beta} + \log\left[\frac{(x/s)^{\alpha}}{1+(x/s)^{\alpha}}\right]\left\{a - \frac{(b-1)(x/s)^{\alpha\beta}}{[1+(x/s)^{\alpha}]^{\beta} - (x/s)^{\alpha\beta}}\right\} \quad \text{and}$$

$$\frac{\partial \ell}{\partial s} = -\frac{\alpha\beta a}{s} + \frac{\alpha(\beta a+1)(x/s)^{\alpha}}{s[1+(x/s)^{\alpha}]} + \frac{\beta\alpha(b-1)(x/s)^{\alpha\beta}}{s[1+(x/s)^{\alpha}]\{[1+(x/s)^{\alpha}]^{\beta} - (x/s)^{\alpha\beta}\}}.$$

For a random sample $(x_1, \ldots, x_n)$ of size n from X, the total log-likelihood is $\ell_n = \ell_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell^{(i)}$, where $\ell^{(i)}$ is the log-likelihood for the ith observation $(i = 1, \ldots, n)$. The total score function is $\mathbf{U}_n = \sum_{i=1}^{n} \mathbf{U}^{(i)}$, where $\mathbf{U}^{(i)}$ has the form given before for $i = 1, \ldots, n$. The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained numerically from the nonlinear equations $\mathbf{U}_n = \mathbf{0}$. For interval estimation and tests of hypotheses on the parameters in $\boldsymbol{\theta}$, we require the 5×5 observed information matrix $J = J(\boldsymbol{\theta}) = \{-\mathbf{L}_{p,q}\}$, where the entries $\mathbf{L}_{p,q}$, for $p, q = a, b, \alpha, \beta, s$, are given in Appendix B.

Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the approximate multivariate normal $N_5(\mathbf{0}, J(\widehat{\boldsymbol{\theta}})^{-1})$ distribution of $\widehat{\boldsymbol{\theta}}$ can be used to construct confidence intervals for the parameters. In fact, an approximate confidence interval with significance level γ for each parameter θ_p is given by

$$\text{ACI}(\theta_p, 100(1-\gamma)\%) = \left(\widehat{\theta}_p - z_{\gamma/2}\sqrt{\widehat{J}^{\theta_p,\theta_p}}, \widehat{\theta}_p + z_{\gamma/2}\sqrt{\widehat{J}^{\theta_p,\theta_p}}\right),$$

where $\widehat{J}^{\theta_p,\theta_p}$ is the diagonal element of $J(\widehat{\boldsymbol{\theta}})^{-1}$ corresponding to each parameter $(p = a, b, \alpha, \beta, s)$ and $z_{\gamma/2}$ is the quantile $(1-\gamma/2)$ of the standard normal distribution.

The likelihood ratio (LR) statistic can be used for comparing the BBIII distribution with some of its special sub-models. Considering the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$, tests of hypotheses of the type $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ versus $H_A : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$ can be performed using LR statistics given by $w = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\}$, where $\widehat{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}$ are the MLEs of $\boldsymbol{\theta}$ under H_A and H_0, respectively. Under the null hypothesis, $w \xrightarrow{d} \chi_q^2$, where q is the dimension of the vector $\boldsymbol{\theta}_1$ of interest. The LR test rejects H_0 if $w > \xi_\gamma$, where ξ_γ denotes the upper $100\gamma\%$ point of the χ_q^2 distribution. For example, we can verify if the fit using the BBIII distribution is statistically "superior" to a fit using the EBIII distribution (for a given data set) by testing $H_0 : b = 1$ versus $H_A : b \neq 1$.

12 A BBIII model for survival data with long-term survivors

In population based cancer studies, cure is said to occur when the mortality in the group of cancer patients returns to the same level as that expected in the general population. The cure fraction is a useful measure of interest when analyzing trends in cancer patient survival. Models for survival analysis typically assume that every subject in the study population is susceptible to the event under study and will eventually experience such event if the follow-up is sufficiently long. However, there are situations when a fraction of individuals are not expected to experience the event of interest, that is, those individuals are cured or not susceptible. For example, researchers may be interested in analyzing the recurrence of a disease. Many individuals may never experience a recurrence and, therefore, a cured fraction of the population exists. Cure rate models have been used for modeling time-to-event data for various types of cancers and to estimate the cured fraction of patients. These models are survival models which allow for a cured fraction of individuals and extend the understanding of time-to-event data by allowing for the formulation of more accurate and informative conclusions. These conclusions are otherwise unobtainable from an analysis which fails to account for a cured or insusceptible fraction of the population. If a cured component is not present, the analysis reduces to standard approaches of survival analysis.

Perhaps the most popular type of cure rate models is the mixture model (Berkson and Gage, 1952; Maller and Zhou, 1996), where the population is divided into two sub-populations so that an individual either is cured with probability p, or has a proper survival function $S(x)$ with probability $1-p$. This leads to an improper population survivor function $S^*(x)$ in the mixture form, namely

$$S^*(x)=p+(1-p)S(x), \qquad S(0)=0, \qquad S^*(\infty)=p. \tag{12.1}$$

Common choices for $S(x)$ in (12.1) are the exponential and Weibull distributions. Mixture models involving these distributions have been studied by several authors (Farewell, 1982; Sy and Taylor, 2000 and Ortega et al., 2009). The book by Maller and Zhou (1996) provides a wide range of applications of the long-term survivor mixture model. The use of survival models with a cure fraction has become more and more frequent because traditional survival analysis does not allow for modeling data in which nonhomogeneous parts of the population do not represent the event of interest even after a long follow-up.

Here, we adopt the BBIII distribution to compose a mixture model for cure rate. Consider a sample $x_1,\dots,x_n$, where x_i is either the observed lifetime or censoring time for the ith individual. Let a binary random variable q_i, for $i=1,\dots,n$, indicating that the ith individual in a population is at risk or not with respect to a certain type of failure, i.e., $q_i=1$ indicates that the ith individual will eventually experience a failure event (uncured) and $q_i=0$ indicates that this individual will never experience such event (cured).

For an individual i, the proportion of uncured $1-p$ individuals can be specified such that the conditional distribution of q_i is given by $\Pr(q_i=1)=1-p$. The cure probability varies over the individuals, so that the probability that individual i is cured is modeled by p. Suppose that the X_i's are independent and identically distributed random variables having the BBIII distribution (1.6). Thus, the contribution of an individual that failed at x_i to the likelihood function becomes

$$\frac{(1-p)\alpha\beta}{s(x/s)^{\alpha+1}B(a,b)}\left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta a+1}\left\{1-\left[\frac{(x/s)^\alpha}{1+(x/s)^\alpha}\right]^{\beta}\right\}^{b-1},$$

and the contribution of an individual that is at risk at time x_i is

$$p + (1-p)\{1 - I_{[1+(x/s)^{-\alpha}]^{-\beta}}(a,b)\},$$

where $I.(\cdot,\cdot)$ is the incomplete beta function ratio (Section 1). The model defined from the last two equations is referred to as the BBIII mixture model with long-term survivors. For $a = b = 1$, we obtain the BIII mixture model (a new model) with long-term survivors. For $b = 1$, we have the exponentiated Burr III (EBIII) mixture model and, for $a = 1$ the Lehmann type II Burr III (LeBIII) mixture model. Thus, the log-likelihood function for the parameter vector $\boldsymbol{\theta} = (p, a, b, \alpha, \beta, s)^T$ can be expressed as

$$\begin{aligned} l(\boldsymbol{\theta}) = {} & r \log\left[\frac{(1-p)\alpha\beta}{sB(a,b)}\right] - (\alpha+1)\sum_{i\in F}\log\left(\frac{x_i}{s}\right) \\ & + (\beta a + 1)\sum_{i\in F}\log\left[\frac{(x_i/s)^\alpha}{1+(x_i/s)^\alpha}\right] \\ & + (b-1)\sum_{i\in F}\log\left\{1 - \left[\frac{(x_i/s)^\alpha}{1+(x_i/s)^\alpha}\right]^\beta\right\} \\ & + \sum_{i\in C}\log\{p + (1-p)[1 - I_{[1+(x_i/s)^{-\alpha}]^{-\beta}}(a,b)]\}, \end{aligned}$$

where F and C denote the sets of individuals corresponding to lifetime observations and censoring times, respectively, and r is the number of uncensored observations (failures). The score functions for the parameters p, a, b, α, β and s are

$$\begin{aligned} U_p(\boldsymbol{\theta}) = {} & \frac{r}{(1-p)} + \sum_{i\in C}\frac{I_{q_i}(a,b)}{p + (1-p)[1 - I_{q_i}(a,b)]}, \\ U_a(\boldsymbol{\theta}) = {} & r[\psi(a+b) - \psi(a)] \\ & + \sum_{i\in F}\{\beta\{\alpha\log(x_i/s) - \log[1+(x_i/s)^\alpha]\}\} \\ & - \sum_{i\in C}\frac{(1-p)[\dot{I}_{q_i}(a,b)]_a}{p + (1-p)[1 - I_{q_i}(a,b)]}, \\ U_b(\boldsymbol{\theta}) = {} & r[\psi(a+b) - \psi(b)] \\ & + \sum_{i\in F}\log\left\{1 - \left[\frac{(x_i/s)^\alpha}{1+(x_i/s)^\alpha}\right]^\beta\right\} \\ & - \sum_{i\in C}\frac{(1-p)[\dot{I}_{q_i}(a,b)]_b}{p + (1-p)[1 - I_{q_i}(a,b)]}, \end{aligned}$$

$$U_\alpha(\boldsymbol{\theta}) = \frac{r}{\alpha} + \sum_{i\in F}\log(x_i/s)\left[\frac{\beta a - (x_i/s)^\alpha}{1+(x_i/s)^\alpha}\right]$$
$$- \sum_{i\in F}\frac{\beta(b-1)(x_i/s)^{\alpha\beta}\log(x_i/s)}{[1+(x_i/s)^\alpha]\{[1+(x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}}$$
$$- \sum_{i\in C}\frac{(1-p)[\dot{I}_{q_i}(a,b)]_\alpha}{p+(1-p)[1-I_{q_i}(a,b)]},$$

$$U_\beta(\boldsymbol{\theta}) = \frac{r}{\beta} + \sum_{i\in F}\log\left[\frac{(x_i/s)^\alpha}{1+(x_i/s)^\alpha}\right]\left\{a - \frac{(b-1)(x_i/s)^{\alpha\beta}}{[1+(x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}}\right\}$$
$$- \sum_{i\in C}\frac{(1-p)[\dot{I}_{q_i}(a,b)]_\beta}{p+(1-p)[1-I_{q_i}(a,b)]},$$

$$U_s(\boldsymbol{\theta}) = -\frac{r\alpha\beta a}{s} + \sum_{i\in F}\frac{\alpha(\beta a+1)(x_i/s)^\alpha}{s[1+(x_i/s)^\alpha]}$$
$$+ \sum_{i\in F}\frac{\beta\alpha(b-1)(x_i/s)^{\alpha\beta}}{s[1+(x_i/s)^\alpha]\{[1+(x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}}$$
$$- \sum_{i\in C}\frac{(1-p)[\dot{I}_{q_i}(a,b)]_s}{p+(1-p)[1-I_{q_i}(a,b)]}.$$

Here, $q_i = [1 + (x_i/s)^{-\alpha}]^{-\beta}$, $[\dot{I}_{q_i}(a,b)]_p = \partial I_{q_i}(a,b)/\partial p$, $[\dot{I}_{q_i}(a,b)]_a = \partial I_{q_i}(a,b)/\partial a$, $[\dot{I}_{q_i}(a,b)]_b = \partial I_{q_i}(a,b)/\partial b$, $[\dot{I}_{q_i}(a,b)]_\alpha = \partial I_{q_i}(a,b)/\partial\alpha$, $[\dot{I}_{q_i}(a,b)]_\beta = \partial I_{q_i}(a,b)/\partial\beta$, $[\dot{I}_{q_i}(a,b)]_s = \partial I_{q_i}(a,b)/\partial s$. The MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by solving the nonlinear likelihood equations $U_p(\boldsymbol{\theta}) = 0$, $U_a(\boldsymbol{\theta}) = 0$, $U_b(\boldsymbol{\theta}) = 0$, $U_\alpha(\boldsymbol{\theta}) = 0$, $U_\beta(\boldsymbol{\theta}) = 0$ and $U_s(\boldsymbol{\theta}) = 0$. They cannot be solved analytically and statistical software can be used to solve them numerically. We can use iterative techniques such as a Newton–Raphson type algorithm to calculate $\widehat{\boldsymbol{\theta}}$. For interval estimation of $(p, a, b, \alpha, \beta, s)$, we can use the 6×6 observed information matrix $J(\boldsymbol{\theta}) = \{-\mathbf{L}_{i,j}\}$, where the entries $\mathbf{L}_{i,j}$ are given in Appendix C.

13 The log-beta Burr III regression model

Let X be a random variable having the BBIII density function (1.6). The random variable $Y = \log(X)$ has a log-beta Burr III (LBBIII) distribution, whose density function (parameterized in terms of $\sigma = \alpha + 1$ and $\mu = -\log(s)$) can be expressed as

$$f(y) = \frac{\beta}{\sigma B(a,b)} \exp\left\{-\left(\frac{y-\mu}{\sigma}\right)\right\}$$
$$\times \left[\frac{\exp((y-\mu)/\sigma)}{1+\exp((y-\mu)/\sigma)}\right]^{\beta a-1} \left\{1 - \left[\frac{\exp((y-\mu)/\sigma)}{1+\exp((y-\mu)/\sigma)}\right]^{\beta}\right\}^{b-1}, \tag{13.1}$$

where $-\infty < y < \infty$, $\sigma > 0$ and $-\infty < \mu < \infty$.

We refer to equation (13.1) as the (new) LBBIII distribution, say $Y \sim \text{LBBIII}(a, b, \beta, \sigma, \mu)$, where μ is a location parameter, σ is a dispersion parameter and a and b are shape parameters. Thus,

$$\text{if } X \sim \text{BBIII}(a, b, \beta, \alpha, s) \qquad \text{then } Y = \log(T) \sim \text{LBBIII}(a, b, \beta, \sigma, \mu).$$

The plots of (13.1) in Figure 12 for selected parameter values show great flexibility of the density function in terms of the parameters a and b. In Figure 12(a), $b = 0.5$ and in Figure 12(b), $a = 1.5$. The survival function corresponding to (13.1) becomes

$$S(y) = 1 - I_{[\exp((y-\mu)/\sigma)/(1+\exp((y-\mu)/\sigma))]^{\beta}}(a, b). \tag{13.2}$$

We define the standardized random variable $Z = (Y - \mu)/\sigma$ with density function

$$\pi(z; a, b) = \frac{\beta \exp(-z)}{B(a,b)} \left[\frac{\exp(z)}{1+\exp(z)}\right]^{\beta a-1} \left\{1 - \left[\frac{\exp(z)}{1+\exp(z)}\right]^{\beta}\right\}^{b-1},$$
$$-\infty < z < \infty. \tag{13.3}$$

The special case $a = b = 1$ leads to the standard log-Burr III (LBIII) distribution. For $b = 1$ and $a = 1$, we obtain the log-exponentiated Burr III (LEBIII) and log-Lehmann type II Burr III (LLeBIII) distributions, respectively.

In many practical applications, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric regression models to estimate univariate survival functions for censored data are widely used. A parametric model that provides a good fit to lifetime data tends to yield more precise estimates of the quantities of interest. Based on the LBBIII density function, we propose a linear location-scale regression model for censored data linking the response variable y_i and the explanatory variable vector $\mathbf{v}_i^T = (v_{i1}, \ldots, v_{ip})$ as follows

$$y_i = \mathbf{v}_i^T \boldsymbol{\gamma} + \sigma z_i, \qquad i = 1, \ldots, n, \tag{13.4}$$

where the random error z_i has density function (13.3), $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)^T$, $\sigma > 0$, $a > 0$ and $b > 0$ are unknown parameters. The parameter $\mu_i = \mathbf{v}_i^T \boldsymbol{\gamma}$ is the location of y_i. The location parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^T$ is given by a linear model $\boldsymbol{\mu} = \mathbf{V}\boldsymbol{\gamma}$, where $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)^T$ is a known model matrix. The LBBIII regression model (13.4) opens new possibilities for fitting many different types of

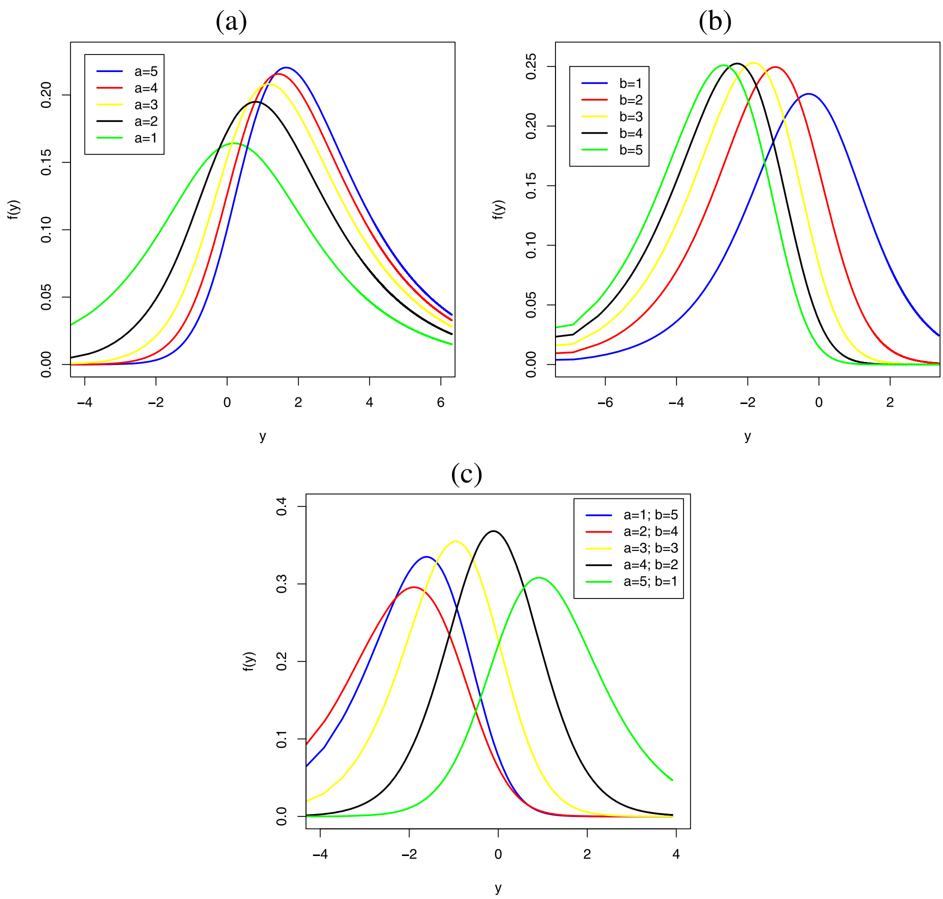

Figure 12 *Plots of the LBBIII density for some parameter values. Figures* (a), (b) *and* (c) $\beta = 0.5$, $\mu = 0$ *and* $\sigma = 1$.

censored data. It is an extension of an accelerated failure time model using the BBIII distribution for censored data.

Consider a sample $(y_1, \mathbf{v}_1), \ldots, (y_n, \mathbf{v}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(x_i), \log(c_i)\}$. We assume non-informative censoring such that the observed lifetimes and censoring times are independent. Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. Conventional likelihood estimation techniques can be applied here. The log-likelihood function for the vector of parameters $\boldsymbol{\tau} = (a, b, \sigma, \beta, \boldsymbol{\gamma}^T)^T$ from model (13.4) has the form $l(\boldsymbol{\tau}) = \sum_{i \in F} l_i(\boldsymbol{\tau}) + \sum_{i \in C} l_i^{(c)}(\boldsymbol{\tau})$, where $l_i(\boldsymbol{\tau}) = \log[f(y_i|\mathbf{v}_i)]$, $l_i^{(c)}(\boldsymbol{\tau}) = \log[S(y_i|\mathbf{v}_i)]$, $f(y_i|\mathbf{v}_i)$ is the density (13.1) and $S(y_i|\mathbf{v}_i)$ is the survival function (13.2) of Y_i. The total log-likelihood function for $\boldsymbol{\tau}$ reduces to

$$l(\boldsymbol{\tau}) = r\log\left[\frac{\beta}{\sigma B(a,b)}\right] - \sum_{i\in F} z_i + (\beta a - 1)\sum_{i\in F}\log\left[\frac{\exp(z_i)}{1+\exp(z_i)}\right]$$
$$+ (b-1)\sum_{i\in F}\log\left\{1 - \left[\frac{\exp(z_i)}{1+\exp(z_i)}\right]^{\beta}\right\} \tag{13.5}$$
$$+ \sum_{i\in C}\log\{1 - I_{[\exp(z_i)/(1+\exp(z_i))]^{\beta}}(a,b)\},$$

where $z_i = (y_i - \mathbf{v}_i^T\boldsymbol{\gamma})/\sigma$ and r is the number of uncensored observations (failures). The score functions for the parameters a, b, β, σ and $\boldsymbol{\gamma}$ are given by

$$U_a(\boldsymbol{\tau}) = -\frac{r}{\sigma}\big[\psi(a) - \psi(a+b)\big] + \beta\sum_{i\in F}\log(u_i) - \sum_{i\in C}\frac{[\dot{I}_{u_i^{\beta}}(a,b)]_a}{1 - I_{u_i^{\beta}}(a,b)},$$

$$U_b(\boldsymbol{\tau}) = -\frac{r}{\sigma}\big[\psi(a) - \psi(a+b)\big] + \sum_{i\in F}\log\big(1 - u_i^{\beta}\big) - \sum_{i\in C}\frac{[\dot{I}_{u_i^{\beta}}(a,b)]_b}{1 - I_{u_i^{\beta}}(a,b)},$$

$$U_\beta(\boldsymbol{\tau}) = \frac{r}{\beta} + a\sum_{i\in F}\log(u_i) - (b-1)\sum_{i\in F}\frac{u_i^{\beta}\log(u_i)}{1-u_i^{\beta}} - \sum_{i\in C}\frac{[\dot{I}_{u_i^{\beta}}(a,b)]_\beta}{1 - I_{u_i^{\beta}}(a,b)},$$

$$U_\sigma(\boldsymbol{\tau}) = \frac{-r}{\sigma} + \frac{1}{\sigma}\sum_{i\in F}(z_i) - \frac{(a\beta-1)}{\sigma}\sum_{i\in F}\frac{z_i u_i}{\exp(z_i)}$$
$$- \frac{(b-1)\beta}{\sigma}\sum_{i\in F}\frac{u_i^{\beta+1} z_i}{(1-u_i^{\beta})(\exp(z_i))} - \sum_{i\in C}\frac{[\dot{I}_{u_i^{\beta}}(a,b)]_\sigma}{1 - I_{u_i^{\beta}}(a,b)},$$

$$U_{\gamma_j}(\boldsymbol{\tau}) = \frac{1}{\sigma}\sum_{i\in F}v_{ij} - \frac{(a\beta-1)}{\sigma}\sum_{i\in F}\frac{v_{ij}u_i}{\exp(z_i)}$$
$$+ \frac{(b-1)\beta}{\sigma}\sum_{i\in F}\frac{u_i^{\beta+1} v_{ij}}{(1-u_i^{\beta})(\exp(z_i))} - \sum_{i\in C}\frac{[\dot{I}_{u_i^{\beta}}(a,b)]_{\gamma_j}}{1 - I_{u_i^{\beta}}(a,b)}.$$

Here, $[\dot{I}_{u_i^{\beta}}(a,b)]_a = \partial I_{u_i^{\beta}}(a,b)/\partial a$, $[\dot{I}_{u_i^{\beta}}(a,b)]_b = \partial I_{u_i^{\beta}}(a,b)/\partial b$, $[\dot{I}_{u_i^{\beta}}(a,b)]_\beta = \partial I_{u_i^{\beta}}(a,b)/\partial\beta$, $[\dot{I}_{u_i^{\beta}}(a,b)]_\sigma = \partial I_{u_i^{\beta}}(a,b)/\partial\sigma$, $[\dot{I}_{u_i^{\beta}}(a,b)]_{\gamma_j} = \partial I_{u_i^{\beta}}(a,b)/\partial\gamma_j$, $u_i = \exp(z_i)/1 + \exp(z_i)$ and $j = 1, \ldots, p$. The MLE $\widehat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ is obtained by solving the nonlinear equations $U_a(\boldsymbol{\tau}) = 0$, $U_b(\boldsymbol{\tau}) = 0$, $U_\beta(\boldsymbol{\tau}) = 0$, $U_\sigma(\boldsymbol{\tau}) = 0$ and $U_{\gamma_j}(\boldsymbol{\tau}) = 0$. These equations cannot be solved analytically and statistical software can be used to solve them numerically. We can use iterative techniques such as a Newton–Raphson type algorithm to calculate the estimate $\widehat{\boldsymbol{\tau}}$. The elements of the observed information matrix corresponding to (13.5) can be by obtained from the authors upon request.

We use the subroutine NLMixed in SAS to compute $\widehat{\boldsymbol{\tau}}$. Initial values for β, $\boldsymbol{\gamma}$ and σ are taken from the fit of the LBIII regression model with $a = b = 1$. The fit of the LBBIII model produces the estimated survival function for y_i given by

$$S(y_i; \hat{a}, \hat{b}, \hat{\sigma}, \hat{\beta}, \widehat{\boldsymbol{\gamma}}^T) = 1 - I_{[\exp(\hat{z}_i)/(1+\exp(\hat{z}_i))]^{\hat{\beta}}}(\hat{a}, \hat{b}), \tag{13.6}$$

where $\hat{z}_i = (y_i - \mathbf{v}_i^T \hat{\boldsymbol{\gamma}})/\hat{\sigma}$.

Under standard regularity conditions, the approximate multivariate normal distribution $N_{p+4}(0, J(\boldsymbol{\tau})^{-1})$ for $\widehat{\boldsymbol{\tau}}$ can be used in the classical way to construct confidence intervals for the parameters in $\boldsymbol{\tau}$, where $J(\boldsymbol{\tau})$ is the observed information matrix. Further, we can use LR statistics for comparing some special sub-models with the LBBIII model. We consider the partition $\boldsymbol{\tau} = (\boldsymbol{\tau}_1^T, \boldsymbol{\tau}_2^T)^T$, where $\boldsymbol{\tau}_1$ is a subset of parameters of interest and $\boldsymbol{\tau}_2$ is a subset of the remaining parameters. The LR statistic for testing the null hypothesis $H_0 : \boldsymbol{\tau}_1 = \boldsymbol{\tau}_1^{(0)}$ versus the alternative hypothesis $H_1 : \boldsymbol{\tau}_1 \neq \boldsymbol{\tau}_1^{(0)}$ is $w = 2\{\ell(\widehat{\boldsymbol{\tau}}) - \ell(\widetilde{\boldsymbol{\tau}})\}$, where $\widetilde{\boldsymbol{\tau}}$ and $\widehat{\boldsymbol{\tau}}$ are the estimates under the null and alternative hypotheses, respectively. The statistic w is asymptotically (as $n \to \infty$) distributed as χ_k^2, where k is the dimension of the subset of parameters $\boldsymbol{\tau}_1$ of interest.

14 Applications

In this section, we give three applications using well-known data sets, two of them with censoring, to demonstrate the flexibility and applicability of the proposed models. These data show that it is necessary to have positively skewed distributions with nonnegative support in different fields. These data present different degrees of skewness and kurtosis.

14.1 Acute myelogeneous data

Here, we apply our methods to a survival data set that was analyzed by Feigl and Zelen (1965). The data represent the survival times, in weeks, of 33 patients suffering from acute myelogeneous Leukaemia. The data, that can also be found at library `SMIR` of the R program (http://cran.r-project.org), are the following: 65, 156, 100, 134, 16, 108, 121, 4, 39, 143, 56, 26, 22, 1, 1, 5, 65, 56, 65, 17, 7, 16, 22, 3, 4, 2, 3, 8, 4, 3, 30, 4, 43. We fit the BBIII, EBIII, LeBIII and BIII distributions to these data using the method of maximum likelihood. The MLEs of the parameters (with standard errors) and the Akaike Information Criterion (AIC) for the fitted models are listed in Table 2.

A comparison of the new distribution with three of its sub-models using LR statistics is performed in Table 3. So, considering a significance level of 10%, we reject the null hypotheses in favor of the BBIII distribution in the three tests.

Table 2 *MLEs of the model parameters for the acute myelogeneous data, the corresponding SEs (given in parentheses) and the AIC measure*

Model	a	b	α	β	s	AIC
BBIII	0.0671	68.8649	0.7687	15.6785	40.5867	315.02
	(0.0080)	(20.6863)	(0.0615)	(2.2911)	(11.5609)	
LeBIII	1	4.6233	0.5024	2.8857	17.2698	316.37
	(–)	(1.0142)	(1.1221)	(0.0816)	(16.9161)	
EBIII	9.3906	1	0.8303	0.3084	3.0707	319.17
	(8.9829)	(–)	(0.0404)	(0.2951)	(1.4018)	
BIII	1	1	0.8478	2.5633	3.7879	317.12
	(–)	(–)	(0.0471)	(0.7354)	(1.8414)	

Table 3 *LR statistics for the acute myelogeneous data*

Model	Hypotheses	Statistics w	p-value
BBIII vs EBIII	$H_0: b = 1$ vs $H_1: H_0$ is false	6.1515	0.01312
BBIII vs LeBIII	$H_0: a = 1$ vs $H_1: H_0$ is false	3.3500	0.06720
BBIII vs BIII	$H_0: a = b = 1$ vs $H_1: H_0$ is false	6.0972	0.04742

The plots of the fitted BBIII, EBIII, BIII and LeBIII densities are given in Figure 13. They show that the new distribution provides a better fit than the other three sub-models. The required numerical evaluations were implemented by using an R program (sub-routine `nlminb` that can be found at http://cran.r-project.org).

Chen and Balakrishnan (1995) proposed a general approximate goodness-of-fit test for the hypothesis $H_0: X_1, \ldots, X_n$ with X_i following $F(x, \theta)$, that is, under H_0, $X_1, \ldots, X_n$ is a random sample from a continuous distribution with cumulative distribution $F(x, \theta)$, where the form of F is known but the p-vector θ is unknown. The method is based on the Cramér–von Mises and Anderson–Darling statistics and, in general, the smaller the values of those statistics, the better the fit. Next, we apply such methodology in order to provide goodness-of-fit tests for the distributions in study.

Table 4 gives the values of the Cramér–von Mises and Anderson–Darling statistics for the acute myelogeneous data. According to the critical points given in Table 1 of Chen and Balakrishnan (1995), when the Anderson–Darling statistic is used, the null hypotheses are rejected at the significance level of 5% for the BIII and EBIII models and at 10% for the LeIII model. The null hypothesis is not rejected for the BBIII model. The same conclusions are obtained when the Cramér–von Mises statistic is used.

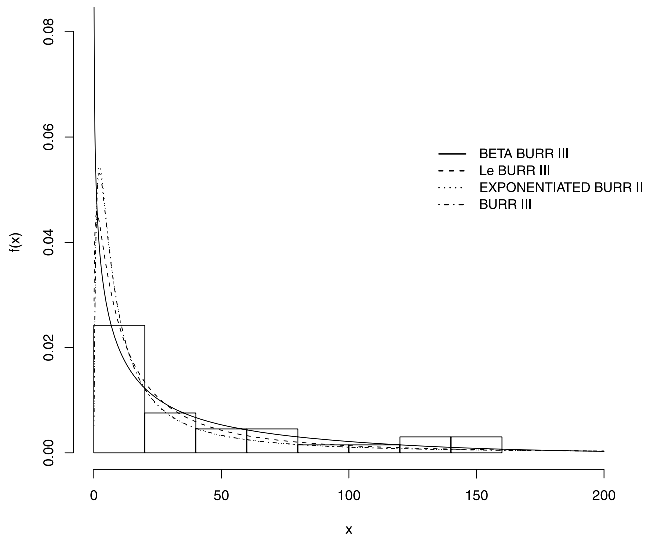

Figure 13 *Fitted BBIII, EBIII, BIII and LeBIII densities for the acute myelogeneous data.*

Table 4 *Goodness-of-fit statistics for the acute myelogeneous data*

Model	Anderson–Darling	Cramér–von Mises
BIII	0.8922	0.1498
EBIII	0.8984	0.1512
LeBIII	0.7125	0.1128
BBIII	0.5657	0.0832

14.2 Melanoma data with long-term survivors

Here, we apply the BBIII model for survival data with long-term survivors to predict cancer recurrence. The data are part of a study on cutaneous melanoma (a type of malignant cancer) for the evaluation of postoperative treatment performance with a high dose of a certain drug (interferon alfa-2b) in order to prevent recurrence. Patients were included in the study from 1991 to 1995, and follow-up was conducted until 1998. The data were collected by Ibrahim et al. (2001). The survival time X is defined as the time until the patient's death. The original sample size was $n = 427$ patients, 10 of whom did not present a value for the explanatory variable *tumor thickness*. When such cases are removed, a sample of size $n = 417$ patients was retained. The percentage of censored observations was 56%.

We start the analysis of the data considering only failure (x_i) and censoring ($cens_i$) observations. An appropriate model for fitting such data could be the BBIII distribution. Table 5 lists the MLEs (and the corresponding standard errors in parentheses) of the model parameters and the AIC statistic for the fitted models.

Table 5 *Estimates of the model parameters for the melanoma data, the corresponding SEs (given in parentheses) and the statistics AIC, CAIC and BIC*

Model	a	b	α	β	s	AIC
BBIII	0.0348	0.0536	1.2575	99.3570	0.6363	1061.6
	(0.0194)	(0.0326)	(0.6433)	(25.87)	(0.9928)	
EBIII	5.5084	1	0.6617	100.12	0.0002	1066.1
	(1.6584)	–	(0.0384)	(0.0144)	(0.00001)	
LeBIII	1	0.7719	0.7440	142.47	0.0027	1065.8
	–	(0.3624)	(0.1604)	(0.0004)	(0.0002)	
BIII	1	1	0.6672	99.4551	0.0031	1064.1
	–	–	(0.0372)	(2.712E−6)	(0.0011)	
	α_1	γ_1				
Weibull	6.9437	1.0509				1102.1
	(0.5611)	(0.0691)				

Table 6 *LR statistics for the melanoma data*

Model	Hypotheses	Statistics w	p-value
BBIII vs EBIII	$H_0: b = 1$ vs $H_1: H_0$ is false	6.40	0.0114
BBIII vs LeBIII	$H_0: a = 1$ vs $H_1: H_0$ is false	6.20	0.0127
BBIII vs BIII	$H_0: a = b = 1$ vs $H_1: H_0$ is false	6.50	0.0388

The computations were performed using the subroutine NLMixed in SAS. These results indicate that the BBIII model has the lowest AIC value among those values of the fitted models, and therefore it could be chosen as the best model.

Note that α_1 and γ_1 are the scale and shape parameters of the Weibull distribution, respectively.

A comparison of the new distribution with three of its sub-models using LR statistics is performed in Table 6.

From the values of these statistics in Table 6, we conclude that the BBIII distribution provides a good fit for these data. In Figure 14, we plot the empirical survival function and the estimated survival functions of the BBIII, BIII and Weibull distributions. These plots indicate that the BBIII model gives the best fit to these data. Next, we present results by fitting the BBIII mixture model. The MLEs (approximate standard errors in parentheses) are: $\hat{a} = 0.8459$ (0.0472), $\hat{b} = 7.5865$ (0.2401), $\hat{p} = 0.4822$ (0.0408), $\hat{\beta} = 5.5316$ (2.9311), $\hat{\alpha} = 0.7596$ (0.2210) and $\hat{s} = 0.9516$ (1.5506). The proportion of cured individuals estimated by the BBIII mixture model is $\widehat{p}_{\text{BBIII}} = 0.4822$. In Figure 15, we plot the empirical survival function and the estimated survival function for the BBIII mixture model which indicates an appropriate fit to the current data.

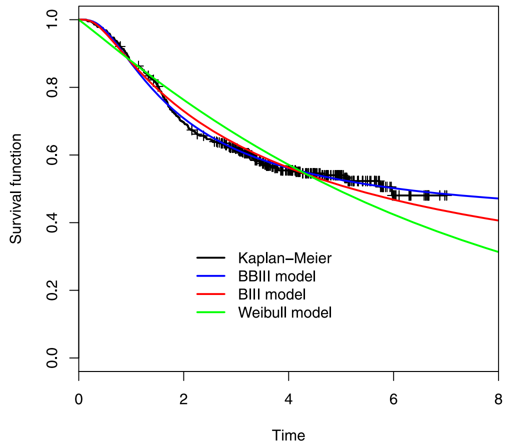

Figure 14 *Estimated survival functions and the empirical survival for melanoma data.*

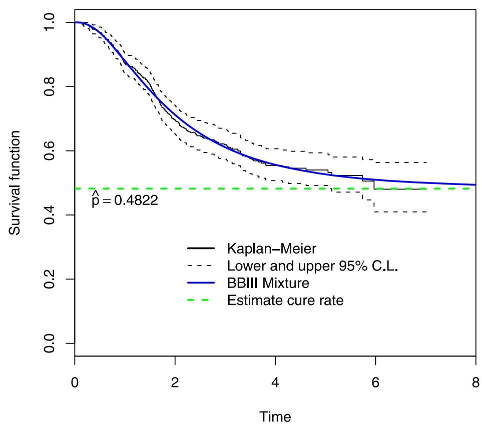

Figure 15 *Estimated survival function for the BBIII mixture model and the empirical survival for melanoma data.*

14.3 Ovarian carcinoma data

Fleming et al. (1980) reports an study, performed at the Clinic Mayo, of patients having limited Stage II or III ovarian carcinoma. The main goal was to determine whether or not the grade of the disease was associated with the time to progression of the disease. The sample size is $n = 35$ and the percentage of censored observations was 34%. The variables involved in the study are:

- t_i—survival times (in days);
- $cens_i$—censoring indicator ($0 =$ censoring, $1 =$ lifetime observed);
- x_{i1}—grade of disease ($0 =$ patients with low-grade or well-differentiated cancer, $1 =$ patients with high-grade or undifferentiated cancer).

Table 7 *MLEs of the parameters from the LBBIII and LBIII regression models fitted to the ovarian carcinoma data, the corresponding SEs (given in parentheses), p-value in* [·] *and the statistic AIC*

Model	a	b	β	σ	β_0	β_1	AIC
LBBIII	0.0072	0.0166	10.7400	0.0520	6.1856	−0.1403	86.7
	(0.0050)	(0.0140)	(4.9442)	(0.0343)	(0.0942)	(0.1313)	
					[<0.001]	[0.2928]	
LEBIII	0.4396	1	0.7916	0.2954	7.1673	−0.8968	94.6
	(0.1019)	–	(0.1944)	(0.0151)	(0.4636)	(0.4046)	
					[<0.001]	[0.0333]	
LLeBIII	1	81.0385	2.2209	1.5580	9.6877	−0.7698	96.3
	–	(2.5643)	(0.6492)	(0.1971)	(1.1108)	(0.3746)	
					[<0.001]	[0.0474]	
LBIII	1	1	0.3480	0.2954	7.1673	−0.8968	92.6
			(0.2236)	(0.1498)	(0.4605)	(0.4025)	
					[<0.001]	[0.0324]	

Table 8 *LR statistics for the ovarian carcinoma data*

Model	Hypotheses	Statistics w	p-value
LBBIII vs LEBIII	$H_0 : b = 1$ vs $H_1 : H_0$ is false	9.8	0.0017
LBBIII vs LLeBIII	$H_0 : a = 1$ vs $H_1 : H_0$ is false	11.6	0.0007
LBBIII vs LBIII	$H_0 : a = b = 1$ vs $H_1 : H_0$ is false	9.9	0.0071

Now, we present results by fitting the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \sigma z_i,$$

where the random variable z_i follows the LBBIII distribution (13.1) for $i = 1, \ldots, 35$. The MLEs of the model parameters are calculated using the procedure NLMixed in SAS. Iterative maximization of the logarithm of the likelihood function (13.5) starts with initial values for $\boldsymbol{\beta}$ and σ taken from the fit of the LBIII regression model with $a = b = 1$. Table 7 lists the MLEs of the model parameters. The value of the AIC statistic is smaller for the LBBIII regression model when compared to the value of the LBIII regression model.

A comparison of the new distribution with three of its sub-models using LR statistics is performed in Table 8. From the values of these statistics, we conclude that the LBBIII distribution provides a good fit for these data.

We note from the fitted LBBIII regression model that x_1 is not significant at 5% and that there is not a significant difference between the patients with low-grade or high-grade for the survival times.

15 Conclusion

In this article, we study several mathematical properties of the beta Burr III (BBIII) distribution, which represents a generalization of the Burr III (BIII) distribution. The current generalization is important because of the wide usage of the baseline distribution and the following two facts: it is quite flexible to analyze positive data and it is an important alternative model to the exponentiated Burr III (EBIII), Lehmann type II Burr III (LeBIII) and Burr III (BIII) sub-models. The new density function can be expressed as a linear combination of BIII densities, which provide some expansions for the ordinary moments, generating function, mean deviations, Bonferroni and Lorenz curves, reliability and two measures of entropy. The density function of the BBIII order statistics can also be expressed as a linear combination of BIII densities. We provide a general formula for the moments of the order statistics. The estimation of parameters is approached by the method of maximum likelihood and the observed information matrix is derived. We adopt the BBIII distribution to compose a mixture model for cure rate and propose a log-BBIII regression model for censored data. The usefulness of the proposed models is illustrated in three applications to real data sets.

Appendix A

We have the following result which holds for m and k positive integers, $\mu > -1$ and $p > 0$ (Prudnikov et al., 1986, page 21):

$$\begin{aligned} I'\left(p, \mu, \frac{m}{k}, \nu\right) &= \int_0^\infty x^\mu \exp(-px)\left(1 + x^{m/k}\right)^\nu dx \qquad \text{(A.1)} \\ &= \frac{k^{-\nu} m^{\mu+1/2}}{(2\pi)^{(m-1)/2}\Gamma(-\nu) p^{\mu+1}} G^{k,k+m}_{k+m,k}\left(\frac{m^m}{p^m} \,\middle|\, \begin{matrix} \Delta(m, -\mu), \Delta(k, \nu+1) \\ \Delta(k, 0) \end{matrix}\right), \end{aligned}$$

where $\Delta(k, a) = \frac{a}{k}, \frac{a+1}{k}, \dots, \frac{a+k}{k}$.

Appendix B

The elements of the observed information matrix $J(\boldsymbol{\theta})$ for $\boldsymbol{\theta} = (a, b, \alpha, \beta, s)^T$ are:

$$\mathbf{L}_{a,a} = n\left[\psi'(a+b) - \psi'(a)\right],$$

$$\mathbf{L}_{a,b} = n\psi'(a+b),$$

$$\mathbf{L}_{a,\alpha} = \beta \sum_{i=1}^{n} \left[\log(x_i/s) - \frac{(x_i/s)^\alpha \log(x_i/s)}{1 + (x_i/s)^\alpha}\right],$$

$$\mathbf{L}_{a,\beta} = \alpha \sum_{i=1}^{n} \log(x_i/s) - \sum_{i=1}^{n} \log\bigl[1 + (x_i/s)^{\alpha}\bigr],$$

$$\mathbf{L}_{a,s} = -\frac{\beta\alpha}{s} \sum_{i=1}^{n} \frac{1}{1 + (x_i/s)^{\alpha}},$$

$$\mathbf{L}_{b,b} = n\bigl[\psi'(a+b) - \psi'(b)\bigr],$$

$$\mathbf{L}_{b,\alpha} = -\beta \sum_{i=1}^{n} \frac{(x_i/s)^{\alpha\beta} \log(x_i/s)}{[1 + (x_i/s)^{\alpha}]\{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}\}},$$

$$\mathbf{L}_{b,\beta} = -\sum_{i=1}^{n} \frac{(x_i/s)^{\alpha\beta}[\alpha \log(x_i/s) - \log(1 + (x_i/s)^{\alpha})]}{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}},$$

$$\mathbf{L}_{b,s} = \frac{\beta\alpha}{s} \sum_{i=1}^{n} \frac{(x_i/s)^{\alpha\beta}}{[1 + (x_i/s)^{\alpha}]\{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}\}},$$

$$\begin{aligned}\mathbf{L}_{\alpha,\alpha} = {} & -\frac{n}{\alpha^2} - (\beta a + 1) \sum_{i=1}^{n} \frac{(x_i/s)^{\alpha}[\log(x_i/s)]^2}{[1 + (x_i/s)^{\alpha}]^2} \\ & - \beta(b-1) \\ & \quad \times \sum_{i=1}^{n} (x_i/s)^{\alpha\beta}\bigl[\log(x_i/s)\bigr]^2 \\ & \qquad \times \frac{\{\beta[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha}\{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}\}\}}{[1 + (x_i/s)^{\alpha}]^2\{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}\}^2},\end{aligned}$$

$$\begin{aligned}\mathbf{L}_{\alpha,\beta} = {} & a \sum_{i=1}^{n} \frac{\log(x_i/s)}{1 + (x_i/s)^{\alpha}} \\ & - (b-1) \sum_{i=1}^{n} (x_i/s)^{\alpha\beta} \log(x_i/s) \\ & \times \left\{ \frac{[1 + (x_i/s)^{\alpha}]^{\beta}\{1 + \alpha\beta \log(x_i/s) - \beta \log[1 + (x_i/s)^{\alpha}]\} - (x_i/s)^{\alpha\beta}}{[1 + (x_i/s)^{\alpha}]\{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}\}^2} \right\},\end{aligned}$$

$$\begin{aligned}\mathbf{L}_{\alpha,s} = {} & -\frac{a\beta}{s} \sum_{i=1}^{n} \frac{1}{[1 + (x_i/s)^{\alpha}]} + \frac{1}{s} \sum_{i=1}^{n} \frac{(x_i/s)^{\alpha}}{[1 + (x_i/s)^{\alpha}]} \\ & + \frac{(a\beta + 1)\alpha}{s} \sum_{i=1}^{n} \frac{(x_i/s)^{\alpha} \log(x_i/s)}{[1 + (x_i/s)^{\alpha}]^2} \\ & + \frac{\beta(b-1)}{s} \sum_{i=1}^{n} \left\{ \frac{[1 + \alpha\beta \log(x_i/s)](x_i/s)^{\alpha\beta}}{[1 + (x_i/s)^{\alpha}]\{[1 + (x_i/s)^{\alpha}]^{\beta} - (x_i/s)^{\alpha\beta}\}} \right.\end{aligned}$$

$$+\frac{\beta\alpha(x_i/s)^{\alpha\beta}\log(x_i/s)\{(x_i/s)^{\alpha\beta}-(x_i/s)^{\alpha}[1+(x_i/s)^{\alpha}]^{\beta-1}\}}{[1+(x_i/s)^{\alpha}]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}^2}$$
$$-\left.\frac{\alpha(x_i/s)^{\alpha(\beta+1)}\log(x_i/s)}{[1+(x_i/s)^{\alpha}]^2\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}}\right\},$$

$$\mathbf{L}_{\beta,\beta}=-\frac{n}{\beta^2}$$
$$-(b-1)\sum_{i=1}^{n}\frac{[1+(x_i/s)^{\alpha}]^{\beta}(x_i/s)^{\alpha\beta}[\alpha\log(x_i/s)-\log(1+(x_i/s)^{\alpha})]^2}{\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}^2},$$

$$\mathbf{L}_{\beta,s}=\frac{\alpha(b-1)}{s}\sum_{i=1}^{n}\frac{1}{[1+(x_i/s)^{\alpha}]}$$
$$\times\left\{\frac{(x_i/s)^{\alpha\beta}}{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}}-a\right.$$
$$\left.+\beta\log\left[\frac{(x_i/s)^{\alpha}}{1+(x_i/s)^{\alpha}}\right]\frac{[1+(x_i/s)^{\alpha}]^{\beta}}{\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}^2}\right\},$$

$$\mathbf{L}_{s,s}=\frac{na\alpha\beta}{s^2}+\frac{\alpha(a\beta+1)}{s^2}\sum_{i=1}^{n}\frac{\{\alpha-(1+\alpha)[1+(x_i/s)^{\alpha}]\}}{\{[1+(x_i/s)^{\alpha}]\}^2}$$
$$+\frac{(\alpha\beta)^2(b-1)}{s^2}\sum_{i=1}^{n}\frac{(x_i/s)^{\alpha\beta}}{[1+(x_i/s)^{\alpha}]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}}$$
$$-\frac{\beta\alpha(b-1)}{s^2}$$
$$\times\sum_{i=1}^{n}\frac{(x_i/s)^{\alpha\beta}[(x_i/s)^{\alpha}(1-\alpha)+1]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}}{\{[1+(x_i/s)^{\alpha}]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}\}^2}$$
$$-\frac{\alpha\beta^2(b-1)a}{s^2}\sum_{i=1}^{n}\frac{[1+(x_i/s)^{\alpha}]\{(x_i/s)^{\alpha\beta}-(x_i/s)^{\alpha}[1+(x_i/s)^{\alpha}]^{\beta-1}\}}{\{[1+(x_i/s)^{\alpha}]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}\}^2}.$$

Appendix C

The elements of the observed information matrix $J(\boldsymbol{\tau})$ for the parameters (p,a,b,α,β,s) are:

$$\mathbf{L}_{p,p}=\frac{r}{(1-p)^2}-\sum_{i\in C}\frac{[I_{q_i}(a,b)]^2}{\{p+(1-p)[1-I_{q_i}(a,b)]\}^2},$$

$$\mathbf{L}_{p,a} = \sum_{i \in C} \frac{[\dot{I}_{q_i}(a,b)]_a [p + 2(1-p) I_{q_i}(a,b)]}{\{p + (1-p)[1 - I_{q_i}(a,b)]\}^2},$$

$$\mathbf{L}_{p,b} = \sum_{i \in C} \frac{[\dot{I}_{q_i}(a,b)]_b [p + 2(1-p) I_{q_i}(a,b)]}{\{p + (1-p)[1 - I_{q_i}(a,b)]\}^2},$$

$$\mathbf{L}_{p,\alpha} = \sum_{i \in C} \frac{[\dot{I}_{q_i}(a,b)]_\alpha [p + 2(1-p) I_{q_i}(a,b)]}{\{p + (1-p)[1 - I_{q_i}(a,b)]\}^2},$$

$$\mathbf{L}_{p,\beta} = \sum_{i \in C} \frac{[\dot{I}_{q_i}(a,b)]_\beta [p + 2(1-p) I_{q_i}(a,b)]}{\{p + (1-p)[1 - I_{q_i}(a,b)]\}^2},$$

$$\mathbf{L}_{p,s} = \sum_{i \in C} \frac{[\dot{I}_{q_i}(a,b)]_s [p + 2(1-p) I_{q_i}(a,b)]}{\{p + (1-p)[1 - I_{q_i}(a,b)]\}^2},$$

$$\begin{aligned}\mathbf{L}_{a,a} = {} & r\big[\psi'(a+b) - \psi'(a)\big] \\ & - \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{aa} \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\} \\ & \qquad + (1-p)^2 \big[\dot{I}_{q_i}(a,b)\big]_a^2\big) / \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2,\end{aligned}$$

$$\begin{aligned}\mathbf{L}_{a,b} = {} & r\psi'(a+b) \\ & - \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{ab} \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\} \\ & \qquad + (1-p)^2 \big[\dot{I}_{q_i}(a,b)\big]_a \big[\dot{I}_{q_i}(a,b)\big]_b\big) / \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2,\end{aligned}$$

$$\begin{aligned}\mathbf{L}_{a,\alpha} = {} & \beta \sum_{i \in F} \left[\log(x_i/s) - \frac{(x_i/s)^\alpha \log(x_i/s)}{1 + (x_i/s)^\alpha}\right] \\ & - \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{a\alpha} \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\} \\ & \qquad + (1-p)^2 \big[\dot{I}_{q_i}(a,b)\big]_a \big[\dot{I}_{q_i}(a,b)\big]_\alpha\big) \\ & \qquad / \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2,\end{aligned}$$

$$\begin{aligned}\mathbf{L}_{a,\beta} = {} & \alpha \sum_{i \in F} \log(x_i/s) - \sum_{i=1}^{n} \log\big[1 + (x_i/s)^\alpha\big] \\ & - \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{a\beta} \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\} \\ & \qquad + (1-p)^2 \big[\dot{I}_{q_i}(a,b)\big]_a \big[\dot{I}_{q_i}(a,b)\big]_\beta\big) \\ & \qquad / \{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2,\end{aligned}$$

$$
\begin{aligned}
\mathbf{L}_{a,s} &= -\frac{\beta\alpha}{s}\sum_{i\in F}\frac{1}{1+(x_i/s)^\alpha} \\
&\quad - \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{as}\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\} \\
&\qquad + (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_a\big[\dot{I}_{q_i}(a,b)\big]_s\big) \\
&\qquad /\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\}^2, \\
\mathbf{L}_{b,b} &= r\big[\psi'(a+b)-\psi'(b)\big] \\
&\quad - \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{bb}\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\} \\
&\qquad + (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_b^2\big)/\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\}^2, \\
\mathbf{L}_{b,\alpha} &= \beta\sum_{i\in F}\frac{(x_i/s)^{\alpha\beta}\log(x_i/s)}{[1+(x_i/s)^\alpha]\{[1+(x_i/s)^\alpha]^\beta-(x_i/s)^{\alpha\beta}\}} \\
&\quad - \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{b\alpha}\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\} \\
&\qquad + (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_b\big[\dot{I}_{q_i}(a,b)\big]_\alpha\big)/\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\}^2, \\
\mathbf{L}_{b,\beta} &= -\sum_{i\in F}\frac{(x_i/s)^{\alpha\beta}[\alpha\log(x_i/s)-\log(1+(x_i/s)^\alpha)]}{[1+(x_i/s)^\alpha]^\beta-(x_i/s)^{\alpha\beta}} \\
&\quad - \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{b\beta}\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\} \\
&\qquad + (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_b\big[\dot{I}_{q_i}(a,b)\big]_\beta\big)/\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\}^2, \\
\mathbf{L}_{b,s} &= \frac{\beta\alpha}{s}\sum_{i\in F}\frac{(x_i/s)^{\alpha\beta}}{[1+(x_i/s)^\alpha]\{[1+(x_i/s)^\alpha]^\beta-(x_i/s)^{\alpha\beta}\}} \\
&\quad - \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{bs}\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\} \\
&\qquad + (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_b\big[\dot{I}_{q_i}(a,b)\big]_s\big)/\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\}^2, \\
\mathbf{L}_{\alpha,\alpha} &= -\frac{r}{\alpha^2}-(\beta a+1)\sum_{i\in F}\frac{(x_i/s)^\alpha[\log(x_i/s)]^2}{[1+(x_i/s)^\alpha]^2} \\
&\quad - \beta(b-1)\sum_{i\in F}(x_i/s)^{\alpha\beta}\big[\log(x_i/s)\big]^2 \\
&\qquad\qquad \times\frac{\{\beta[1+(x_i/s)^\alpha]^\beta-(x_i/s)^\alpha\{[1+(x_i/s)^\alpha]^\beta-(x_i/s)^{\alpha\beta}\}\}}{[1+(x_i/s)^\alpha]^2\{[1+(x_i/s)^\alpha]^\beta-(x_i/s)^{\alpha\beta}\}^2}
\end{aligned}
$$

$$- \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{\alpha\alpha}\{p+(1-p)[1-I_{q_i}(a,b)]\}$$

$$+ (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_{\alpha}^2\big)/\big\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\big\}^2,$$

$$\mathbf{L}_{\alpha,\beta} = a\sum_{i\in F}\frac{\log(x_i/s)}{1+(x_i/s)^{\alpha}}$$

$$-(b-1)\sum_{i\in F}(x_i/s)^{\alpha\beta}\log(x_i/s)$$

$$\times\left\{\frac{[1+(x_i/s)^{\alpha}]^{\beta}\{1+\alpha\beta\log(x_i/s)-\beta\log[1+(x_i/s)^{\alpha}]\}-(x_i/s)^{\alpha\beta}}{[1+(x_i/s)^{\alpha}]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}^2}\right\}$$

$$- \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{\alpha\beta}\{p+(1-p)[1-I_{q_i}(a,b)]\}$$

$$+ (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_{\alpha}\big[\dot{I}_{q_i}(a,b)\big]_{\beta}\big)/\big\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\big\}^2,$$

$$\mathbf{L}_{\alpha,s} = -\frac{\beta a}{s}\sum_{i=1}^{n}\frac{1}{[1+(x_i/s)^{\alpha}]}$$

$$+\frac{1}{s}\sum_{i\in F}\frac{(x_i/s)^{\alpha}}{[1+(x_i/s)^{\alpha}]}+\frac{(\beta a+1)\alpha}{s}\sum_{i\in F}\frac{(x_i/s)^{\alpha}\log(x_i/s)}{[1+(x_i/s)^{\alpha}]^2}$$

$$+\frac{\beta(b-1)}{s}\sum_{i\in F}\left\{\frac{[1+\alpha\beta\log(x_i/s)](x_i/s)^{\alpha\beta}}{[1+(x_i/s)^{\alpha}]\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}}\right.$$

$$+\big(\beta\alpha(x_i/s)^{\alpha\beta}\log(x_i/s)$$

$$\times\big\{(x_i/s)^{\alpha\beta}-(x_i/s)^{\alpha}\big[1+(x_i/s)^{\alpha}\big]^{\beta-1}\big\}\big)$$

$$/\big(\big[1+(x_i/s)^{\alpha}\big]\big\{\big[1+(x_i/s)^{\alpha}\big]^{\beta}-(x_i/s)^{\alpha\beta}\big\}^2\big)$$

$$\left.-\frac{\alpha(x_i/s)^{\alpha(\beta+1)}\log(x_i/s)}{[1+(x_i/s)^{\alpha}]^2\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}}\right\}$$

$$- \sum_{i\in C}\big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{\alpha s}\{p+(1-p)[1-I_{q_i}(a,b)]\}$$

$$+ (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_{\alpha}\big[\dot{I}_{q_i}(a,b)\big]_{s}\big)/\big\{p+(1-p)\big[1-I_{q_i}(a,b)\big]\big\}^2,$$

$$\mathbf{L}_{\beta,\beta} = -\frac{n}{\beta^2}$$

$$-(b-1)\sum_{i\in F}\frac{[1+(x_i/s)^{\alpha}]^{\beta}(x_i/s)^{\alpha\beta}[\alpha\log(x_i/s)-\log(1+(x_i/s)^{\alpha})]^2}{\{[1+(x_i/s)^{\alpha}]^{\beta}-(x_i/s)^{\alpha\beta}\}^2}$$

$$- \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{\beta\beta}\big\{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}$$
$$+ (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_{\beta}^2\big)/\big\{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2,$$

$$\mathbf{L}_{\beta,s} = \frac{\alpha(b-1)}{s} \sum_{i \in F} \frac{1}{[1 + (x_i/s)^\alpha]}$$
$$\times \left\{ \frac{(x_i/s)^{\alpha\beta}}{[1 + (x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}} - a \right.$$
$$\left. + \beta \log\left[\frac{(x_i/s)^\alpha}{1 + (x_i/s)^\alpha}\right] \frac{[1 + (x_i/s)^\alpha]^\beta}{\{[1 + (x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}^2} \right\}$$
$$- \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{\beta s}\big\{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}$$
$$+ (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_{\beta}\big[\dot{I}_{q_i}(a,b)\big]_{s}\big)/\big\{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2,$$

$$\mathbf{L}_{s,s} = \frac{n\beta\alpha a}{s^2} + \frac{\alpha(\beta a + 1)}{s^2} \sum_{i \in F} \frac{\{\alpha - (1+\alpha)[1 + (x_i/s)^\alpha]\}}{\{[1 + (x_i/s)^\alpha]\}^2}$$
$$+ \frac{(\beta\alpha)^2(b-1)}{s^2} \sum_{i \in F} \frac{(x_i/s)^{\alpha\beta}}{[1 + (x_i/s)^\alpha]\{[1 + (x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}}$$
$$- \frac{\beta\alpha(b-1)}{s^2}$$
$$\times \sum_{i \in F} \frac{(x_i/s)^{\alpha\beta}[(x_i/s)^\alpha(1-\alpha) + 1]\{[1 + (x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}}{\{[1 + (x_i/s)^\alpha]\{[1 + (x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}\}^2}$$
$$- \frac{\beta^2\alpha(b-1)a}{s^2} \sum_{i \in F} \frac{[1 + (x_i/s)^\alpha]\{(x_i/s)^{\alpha\beta} - (x_i/s)^\alpha[1 + (x_i/s)^\alpha]^{\beta-1}\}}{\{[1 + (x_i/s)^\alpha]\{[1 + (x_i/s)^\alpha]^\beta - (x_i/s)^{\alpha\beta}\}\}^2}$$
$$- \sum_{i \in C} \big((1-p)\big[\ddot{I}_{q_i}(a,b)\big]_{ss}\big\{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}$$
$$+ (1-p)^2\big[\dot{I}_{q_i}(a,b)\big]_{s}^2\big)/\big\{p + (1-p)\big[1 - I_{q_i}(a,b)\big]\big\}^2.$$

Here, $[\ddot{I}_{q_i}(a,b)]_{aa} = \partial^2 I_{q_i}(a,b)/\partial a^2$, $[\ddot{I}_{q_i}(a,b)]_{ab} = \partial^2 I_{q_i}(a,b)/\partial a\,\partial b$, $[\ddot{I}_{q_i}(a,b)]_{a\alpha} = \partial^2 I_{q_i}(a,b)/\partial a\,\partial\alpha$, $[\ddot{I}_{q_i}(a,b)]_{a\beta} = \partial^2 I_{q_i}(a,b)/\partial a\,\partial\beta$, $[\ddot{I}_{q_i}(a,b)]_{as} = \partial^2 I_{q_i}(a,b)/\partial a\,\partial s$, $[\ddot{I}_{q_i}(a,b)]_{bb} = \partial^2 I_{q_i}(a,b)/\partial b^2$, $[\ddot{I}_{q_i}(a,b)]_{b\alpha} = \partial^2 I_{q_i}(a,b)/\partial b\,\partial\alpha$, $[\ddot{I}_{q_i}(a,b)]_{b\beta} = \partial^2 I_{q_i}(a,b)/\partial b\,\partial\beta$, $[\ddot{I}_{q_i}(a,b)]_{bs} = \partial^2 I_{q_i}(a,b)/\partial b\,\partial s$, $[\ddot{I}_{q_i}(a,b)]_{\alpha\alpha} = \partial^2 I_{q_i}(a,b)/\partial\alpha^2$, $[\ddot{I}_{q_i}(a,b)]_{\alpha\beta} = \partial^2 I_{q_i}(a,b)/\partial\alpha\,\partial\beta$, $[\ddot{I}_{q_i}(a,b)]_{\alpha s} = \partial^2 I_{q_i}(a,b)/\partial\alpha\,\partial s$, $[\ddot{I}_{q_i}(a,b)]_{\beta\beta} = \partial^2 I_{q_i}(a,b)/\partial\beta^2$, $[\ddot{I}_{q_i}(a,b)]_{\beta s} = \partial^2 I_{q_i}(a,b)/\partial\beta\,\partial s$ and $[\ddot{I}_{q_i}(a,b)]_{ss} = \partial^2 I_{q_i}(a,b)/\partial s^2$.

Acknowledgments

The authors would like to thank the anonymous reviewer for his careful reading of the manuscript and for constructive comments which considerably improve the article. The first three authors are deeply indebted to CAPES, Brazil, for financial support (Project PROCAD-NF 2008). The last three authors were supported by CNPq, Brazil.

References

Akinsete, A., Famoye, F. and Lee, C. (2008). The beta-Pareto distribution. *Statistics* **42**, 547–563. MR2465134

Al-Dayian, G. R. (1999). Burr type III distribution: Properties and estimation. *The Egyptian Statistical Journal* **43**, 102–116.

Barreto-Souza, W., Santos, A. H. S. and Cordeiro, G. M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation* **80**, 159–172. MR2603623

Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**, 501–515.

Burr, I. W. (1942). Cumulative frequency distributions. *Annals of Mathematical Statistics* **13**, 215–232. MR0006644

Chen, G. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology* **27**, 154–161.

Cordeiro, G. M. and Lemonte, A. J. (2011a). The β-Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis* **55**, 1445–1461. MR2741426

Cordeiro, G. M. and Lemonte, A. J. (2011b). The beta Laplace distribution. *Statistics and Probability Letters* **81**, 973–982. MR2803732

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communication in Statistics—Theory and Methods* **31**, 497–512. MR1902307

Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics* **38**, 1041–1046.

Feigl, P. and Zelen, M. (1965). Estimation of exponential probabilities with concomitant information. *Biometrics* **21**, 826–838.

Fischer, M. J. and Vaughan, D. (2010). The beta-hyperbolic secant distribution. *Austrian Journal of Statistics* **39**, 245–258.

Fleming, T. R., O'Fallon, J. R., O'Brien, P. C. and Harrington, D. P. (1980). Modified Kolgomorov–Smirnov test procedures with application to arbitrarily right censored data. *Biometrics* **36**, 607–626. MR0672339

Gove, J. H., Ducey M. J., Leak, W. B. and Zhang, L. (2008). Rotated sigmoid structures in managed uneven-aged northern hardwood stands: A look at the Burr type III distribution. *Forestry* **81**, 161–176.

Gradshteyn, I. S. and Ryzhik, I. M. (2000). *Table of Integrals, Series, and Products*. New York: Academic Press. MR1773820

Hose, G. C. (2005). Assessing the need for groundwater quality guidlines for pesticides using the species sensitivity distribution approach. *Human and Ecological Risk Assessment* **11**, 951–966.

Hoskings, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B* **52**, 105–124. MR1049304

Ibrahim, J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis*. New York: Springer. MR1876598

Jones, M. C. (2004). Family of distributions arising from distribution of order statistics. *Test* **13**, 1–43. MR2065642

Kenney, J. F. and Keeping, E. S. (1962). *Mathematics of Statistics, Part 1*. Princeton, NJ: Van Nostrand, pp. 101–102.

Khan, M. S. (2010). The beta inverse Weibull distribution. *International Transactions in Mathematical Sciences and Computer* **3**, 113–119.

Klugman, S. A., Panjer, H. H. and Willmot, G. E. (1998). *Loss Models*. New York: Wiley. MR1490300

Kotz, S., Lumelskii, Y. and Pensky, M. (2003). *The Stress-Strength Model and Its Generalizations: Theory and Applications*. River Edge, NJ: World Scientific. MR1980497

Lee, C., Famoye and F. Olumolade, O. (2007). Beta-Weibull distribution: Some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* **6**, 173–186.

Lindsay, S. R., Wood, G. R. and Woollons R. C. (1996). Modelling the diameter distribution of forest stands using the Burr distribution. *Journal of Applied Statistics* **23**, 609–620.

Maller, R. and Zhou, X. (1996). *Survival Analysis with Long-term Survivors*. New York: Wiley. MR1453117

Mielke, P. W. (1973). Another family of distributions for describing and analyzing precipitation data. *Journal of Applied Meteorology* **12**, 275–280.

Mielke, P. W. and Johnson, E. S. (1973). Three-parameter kappa distribution maximum likelihood estimates and likelihood ratio test. *Monthly Weather Review* **101**, 701–707.

Mokhlis, N. A. (2005). Reliability of a stress-strength model with Burr type III distributions. *Communications in Statistics—Theory and Methods* **34**, 1643–1657. MR2168516

Nadarajah, S. and Gupta, A. K. (2004). The beta Fréchet distribution. *Far East Journal of Theoretical Statistics* **14**, 15–24. MR2108090

Nadarajah, S. and Kotz, S. (2004). The beta Gumbel distribution. *Mathematical Problems in Engineering* **10**, 323–332. MR2109721

Nadarajah, S. and Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering and System Safety* **91**, 689–697.

Ortega, E. M. M., Rizato, H. and Demétrio, C. G. B. (2009). The generalized log-gamma mixture model with covariates: Local influence and residual analysis. *Statistics Methods and Applications* **18**, 305–331. MR2529260

Paranaíba, P. F., Ortega, E. M. M., Cordeiro, G. M. and Pescim, R. R. (2011). The beta Burr XII distribution with application to lifetime data. *Computational Statistics and Data Analysis* **55**, 1118–1136. MR2736499

Pescim, R. R., Demétrio, C. G. B, Cordeiro, G. M., Ortega, E. M. M. and Urbano, M. R. (2010). The beta generalized half-normal distribution. *Computational Statistics and Data Analysis* **54**, 945–957. MR2580929

Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). *Integrals and Series, Vol. 1*. Amsterdam: Gordon and Breach Science Publishers. MR0874986

Rodriguez, R. N. (1980). Multivariate Burr III distributions, Part I. Theoretical properties. Research Publication GMR-3232, General Motors Research Laboratories, Warren, MI.

Shao, Q. (2000). Estimation for hazardous concentrations based on NOEC toxicity data: An alternative approach. *Environmetrics* **11**, 583–595.

Shao, Q., Chen, Y. D. and Zhang, L. (2008). An extension of three-parameter Burr III distribution for low-flow frequency analysis. *Computational Statistics and Data Analysis* **52**, 1304–1314. MR2418567

Sherrick, B. J., Garcia, P. and Tirupattur, V. (1996). Recovering probabilistic information from option markets: Test of distributional assumptions. *The Journal of Future Markets* **16**, 545–560.

Silva, G. O., Ortega, E. M. M. and Cordeiro, G. M. (2010). The beta modified Weibull distribution. *Lifetime Data Analysis* **16**, 409–430. MR2657898

Sy, J. P. and Taylor, M. M. G. (2000). Estimation in a proportional hazards cure model. *Biometrics* **56**, 227–336. MR1767631

A. E. Gomes
C. Q. da-Silva
Departamento de Estatística
Universidade de Brasília
70910-900, Brasília, DF
Brazil
E-mail: aegomes@unb.br
cibeleqs@unb.br

G. M. Cordeiro
Departamento de Estatística
Universidade Federal de Pernambuco
52171-900, Recife, PE
Brazil
E-mail: gausscordeiro@uol.com.br

E. M. M. Ortega
Departamento de Ciências Exatas
ESALQ-Universidade de São Paulo
Piracicaba, SP
Brazil
E-mail: edwin@usp.br