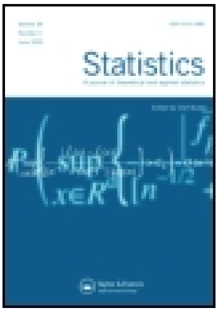

Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gsta20

An extended fatigue life distribution

Gauss M. Cordeiro [a], Artur J. Lemonte [b] & Edwin M.M. Ortega [c]

[a] Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil

[b] Departamento de Estatística, Universidade de São Paulo, São Paulo, Brazil

[c] Departamento de Ciências Exatas, Universidade de São Paulo, São Paulo, Brazil

Available online: 26 Sep 2011

To cite this article: Gauss M. Cordeiro, Artur J. Lemonte & Edwin M.M. Ortega (2011): An extended fatigue life distribution, Statistics, DOI:10.1080/02331888.2011.617447

To link to this article: http://dx.doi.org/10.1080/02331888.2011.617447

Statistics, iFirst, 2011, 1–28

An extended fatigue life distribution

Gauss M. Cordeiro[a], Artur J. Lemonte[b]* and Edwin M.M. Ortega[c]

[a]*Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil;* [b]*Departamento de Estatística, Universidade de São Paulo, São Paulo, Brazil;* [c]*Departamento de Ciências Exatas, Universidade de São Paulo, São Paulo, Brazil*

(*Received 27 December 2010; final version received 22 August 2011*)

A five-parameter extended fatigue life model called the McDonald–Birnbaum–Saunders (McBS) distribution is proposed. It extends the Birnbaum–Saunders and beta Birnbaum–Saunders [G.M. Cordeiro and A.J. Lemonte, *The β-Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling*. Comput. Statist. Data Anal. 55 (2011), pp. 1445–1461] distributions and also the new Kumaraswamy–Birnbaum–Saunders distribution. We obtain the ordinary moments, generating function, mean deviations and quantile function. The method of maximum likelihood is used to estimate the model parameters and its potentiality is illustrated with an application to a real fatigue data set. Further, we propose a new extended regression model based on the logarithm of the McBS distribution. This model can be very useful to the analysis of real data and could give more realistic fits than other special regression models.

Keywords: Birnbaum–Saunders distribution; Kumaraswamy distribution; maximum-likelihood estimation; McDonald distribution; regression model

1. Introduction

The statistics literature is filled with hundreds of continuous univariate distributions which have been extensively used over the past decades for modelling data in several fields such as environmental and medical sciences, engineering, demography, biological studies, actuarial, economics, finance and insurance. However, in many applied areas such as lifetime analysis, finance and insurance, there is a clear need for extended forms of these distributions. The Birnbaum–Saunders (BS) [1,2] distribution, also known as the fatigue life distribution, is a very popular model which has been extensively used for modelling failure times of fatiguing materials and lifetime data in the fields cited above. A random variable T having the $\mathrm{BS}(\alpha, \beta)$ distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ is defined by

$$T = \beta\left[\frac{\alpha Z}{2} + \left\{\left(\frac{\alpha Z}{2}\right)^2 + 1\right\}^{1/2}\right]^2,$$

*Corresponding author. Email: arturlemonte@gmail.com

ISSN 0233-1888 print/ISSN 1029-4910 online

http://dx.doi.org/10.1080/02331888.2011.617447
http://www.tandfonline.com

where Z is a standard normal random variable. The cumulative distribution function (cdf) of T is given by

$$G_{\alpha,\beta}(t) = \Phi(v), \quad t > 0, \tag{1}$$

where $v = \alpha^{-1}\rho(t/\beta)$, $\rho(z) = z^{1/2} - z^{-1/2}$ and $\Phi(\cdot)$ is the standard normal cumulative function. The parameter β is the median of the distribution: $G(\beta) = \Phi(0) = 1/2$. For any $k > 0$, $kT \sim \mathrm{BS}(\alpha, k\beta)$. Kundu *et al.* [3] discussed the shape of its hazard function. The probability density function (pdf) corresponding to Equation (1) is

$$g_{\alpha,\beta}(t) = \kappa(\alpha,\beta)t^{-3/2}(t+\beta)\exp\left\{-\frac{\tau(t/\beta)}{2\alpha^2}\right\}, \quad t > 0, \tag{2}$$

where $\kappa(\alpha,\beta) = \exp(\alpha^{-2})/(2\alpha\sqrt{2\pi\beta})$ and $\tau(z) = z + z^{-1}$. Results on improved statistical inference for the BS model are discussed by Wu and Wong [4] and Lemonte *et al.* [5,6]. The moments of T are [7]

$$E(T^p) = \beta^p I(p,\alpha), \tag{3}$$

where

$$I(p,\alpha) = \frac{K_{p+1/2}(\alpha^{-2}) + K_{p-1/2}(\alpha^{-2})}{2K_{1/2}(\alpha^{-2})}. \tag{4}$$

The function $K_\nu(\cdot)$ is the modified Bessel function of the third kind and order ν.

The generalized beta distribution of the first kind (or beta type-I) may be characterized by the density function [8]

$$f_{\mathrm{Mc}}(t;a,b,c) = \frac{c}{B(ac^{-1},b)}t^{a-1}(1-t^c)^{b-1}, \quad 0 < t < 1, \tag{5}$$

where $a > 0$, $b > 0$ and $c > 0$ are shape parameters. Two important special models are the beta and Kumaraswamy [9] distributions for $c = 1$ and $a = c$, respectively.

For an arbitrary baseline distribution $G(t)$ with parameter vector $\boldsymbol{\tau}$ and density function $g(t)$, the McDonald generalized (denoted by the prefix 'McG' for short) cdf is defined by

$$F_{\mathrm{McG}}(t) = I_{G(t)^c}(ac^{-1},b) = \frac{1}{B(ac^{-1},b)}\int_0^{G(t)^c} \omega^{(a/c)-1}(1-\omega)^{b-1}\,\mathrm{d}\omega. \tag{6}$$

Here, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function, $\Gamma(\alpha) = \int_0^\infty w^{\alpha-1}\,\mathrm{e}^{-w}\,\mathrm{d}w$ is the gamma function, $I_y(a,b) = B_y(a,b)/B(a,b)$ is the incomplete beta function ratio, $B_y(a,b) = \int_0^y \omega^{a-1}(1-\omega)^{b-1}\,\mathrm{d}\omega$ is the incomplete beta function and $a > 0$, $b > 0$ and $c > 0$ are additional shape parameters to those in $\boldsymbol{\tau}$ to control skewness through the relative tail weights.

The density function corresponding to Equation (6) reduces to

$$f_{\mathrm{McG}}(t) = \frac{c}{B(ac^{-1},b)}g(t)G(t)^{a-1}[1-G(t)^c]^{b-1}. \tag{7}$$

Clearly, the McDonald density (5) is a basic exemplar of Equation (7) for $G(t) = t$, $t \in (0,1)$. Additionally, we obtain the classical beta and Kumaraswamy distributions for $c = 1$ and $a = c$, respectively. The distribution of Kumaraswamy [9] is commonly termed the 'minimax' distribution. Jones [10] advocates its tractability, especially in simulations because its quantile function takes a simple form, and its pedagogical appeal relative to the classical beta distribution. Equation (7) will be most tractable when both functions $G(t)$ and $g(t)$ have simple analytic expressions. Its major benefit is the ability to fit skewed data that cannot be properly fitted by existing distributions. Application of $T = G^{-1}(V^{1/c})$ to a beta random variable V with positive parameters a/c and b yields T with cumulative function (6).

The class of distributions (7) includes two important special sub-classes: the beta-generalized (BG) and Kumaraswamy-generalized (KwG) distributions defined by Eugene *et al.* [11] and Cordeiro and de Castro [12] when $c = 1$ and $a = c$, respectively. It follows immediately from Equation (7) that the McG distribution with parent $G(t)$ is the BG distribution with baseline $G(t)^c$. This simple transformation may facilitate the derivation of several of its properties. The BG distributions can be limited in one aspect. They have only two additional shape parameters and so they can add only a limited structure to the generated distribution. For instance, a BG distribution may have problems to capture the behaviour of random variables with symmetric but highly leptokurtic distributions. While the beta parameters offer explicit control over skewness when the parent is symmetric, they have less control over higher moments such as kurtosis. Further, a KwG distribution still introduces only two extra shape parameters, whereas three may be required to control both tail weights and the distribution of weight in the centre. Hence, the generated distribution (7) is a more flexible model since it has one more shape parameter than the classical beta or Kumaraswamy generators. This parameter can give additional control over both skewness and kurtosis.

In this note, we study some structural properties of a new five-parameter distribution, called the McDonald–BS (McBS) distribution, defined from Equation (7) by taking $G(t)$ and $g(t)$ to be the cdf and pdf of the BS(α, β) distribution, respectively. The McBS distribution can be widely applied in many areas of engineering and biology. Further, we propose a generalized regression model based on the logarithm of a random variable following the McBS distribution, i.e. the log-McBS (LMcBS) distribution. This regression model extends the log-BS (LBS) model, also referred to as the sinh-normal regression model [13]. The regression model introduced by Rieck and Nedelman [13] has been studied by several authors in the last few years. Some references are [14–21], among others. A sinh-normal nonlinear regression model is presented by Lemonte and Cordeiro [22].

The article is outlined as follows. In Section 2, we define the McBS distribution. Section 3 provides a useful expansion for its density function. Section 4 deals with non-standard measures for the skewness and kurtosis. A power series expansion for the McBS quantile function is derived in Section 5. In Section 6, we obtain two simple expansions for the moments. Section 7 provides two expansions for the moment-generating function (mgf). Mean deviations, Bonferroni and Lorenz curves and the reliability are investigated in Sections 8 and 9, respectively. Maximum-likelihood estimation is discussed in Section 10. Section 11 introduces the LMcBS distribution. The mgf of this distribution is derived in Section 12. We propose an extended LMcBS regression model in Section 13. Further, we provide applications to real data in Section 14. Finally, Section 15 offers some concluding remarks.

2. The McBS distribution

To avoid non-identifiability problems, we allow b to vary on the interval $[1, \infty)$ only. Let $\eta = b - 1$ which varies on $[0, \infty)$. From now on, we denote a random variable having the McBS$(a, \eta, c, \alpha, \beta)$ distribution by $T \sim$ McBS$(a, \eta, c, \alpha, \beta)$. The cdf of T reduces to $F(t) = I_{\Phi(v)^c}(ac^{-1}, \eta + 1)$. The density function of T (for $t > 0$) can be expressed from Equation (7) as

$$f(t) = \frac{c\kappa(\alpha,\beta)t^{-3/2}(t+\beta)}{B(ac^{-1},\eta+1)} \exp\left\{-\frac{\tau(t/\beta)}{2\alpha^2}\right\} \Phi(v)^{a-1}\{1-\Phi(v)^c\}^{\eta}, \tag{8}$$

where β is a scalar parameter and α, a, η and c are positive shape parameters. The hazard rate function associated with Equation (8) is given by

$$r(t) = \frac{c\kappa(\alpha,\beta)(t+\beta)t^{-3/2}}{B(ac^{-1},\eta+1)[1-I_{\Phi(v)^c}(ac^{-1},\eta+1)]} \exp\left\{-\frac{\tau(t/\beta)}{2\alpha^2}\right\} \Phi(v)^{a-1}\{1-\Phi(v)^c\}^{\eta}. \tag{9}$$

The study of the new distribution seems important since it extends some distributions previously considered in the literature. In fact, the BS model (with parameters α and β) is clearly a basic exemplar for $a = c = 1$ and $\eta = 0$, with a continuous crossover towards models with different shapes (e.g. a particular combination of skewness and kurtosis). The McBS model contains as sub-models the beta BS (BBS) [23] and Kumaraswamy–BS (KwBS) distributions when $c = 1$ and $a = c$, respectively. Plots of the McBS density and hazard rate functions for selected parameter

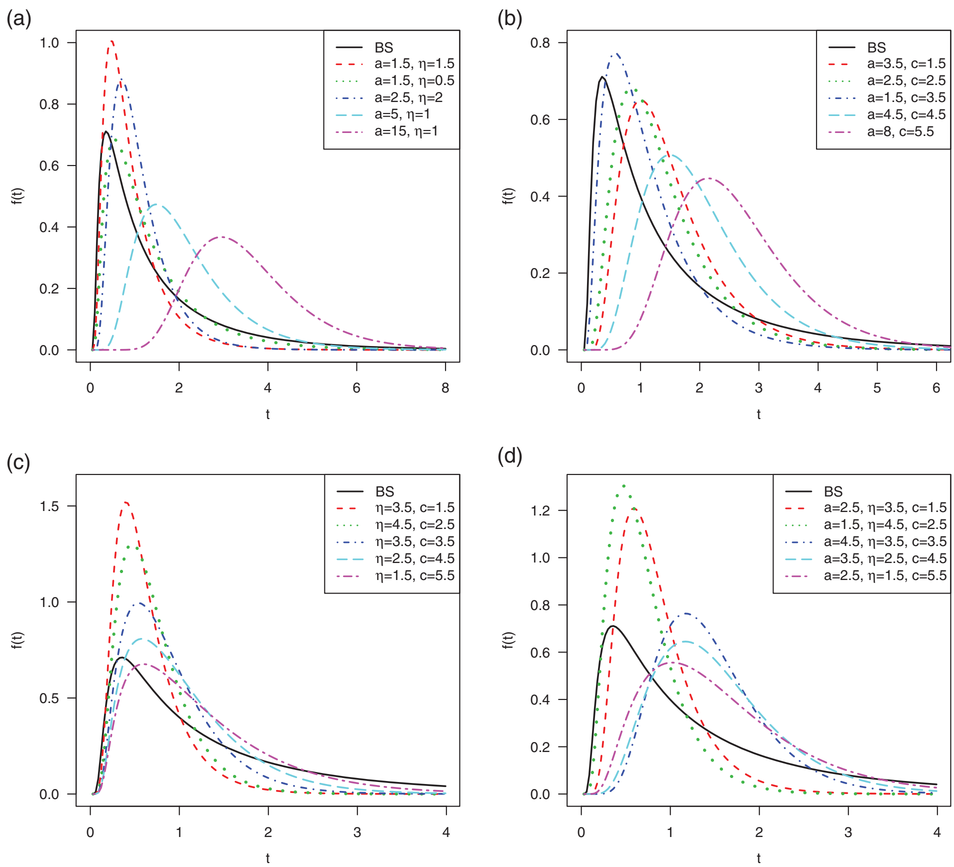

Figure 1. Plots of the density function (8) for some parameter values. (a) For values $\alpha = \beta = 1.0$ and $c = 1.5$; (b) for values $\alpha = \beta = 1.0$ and $\eta = 1.5$; (c) for values $\alpha = \beta = 1.0$ and $a = 1.5$ and (d) for values $\alpha = \beta = 1.0$.

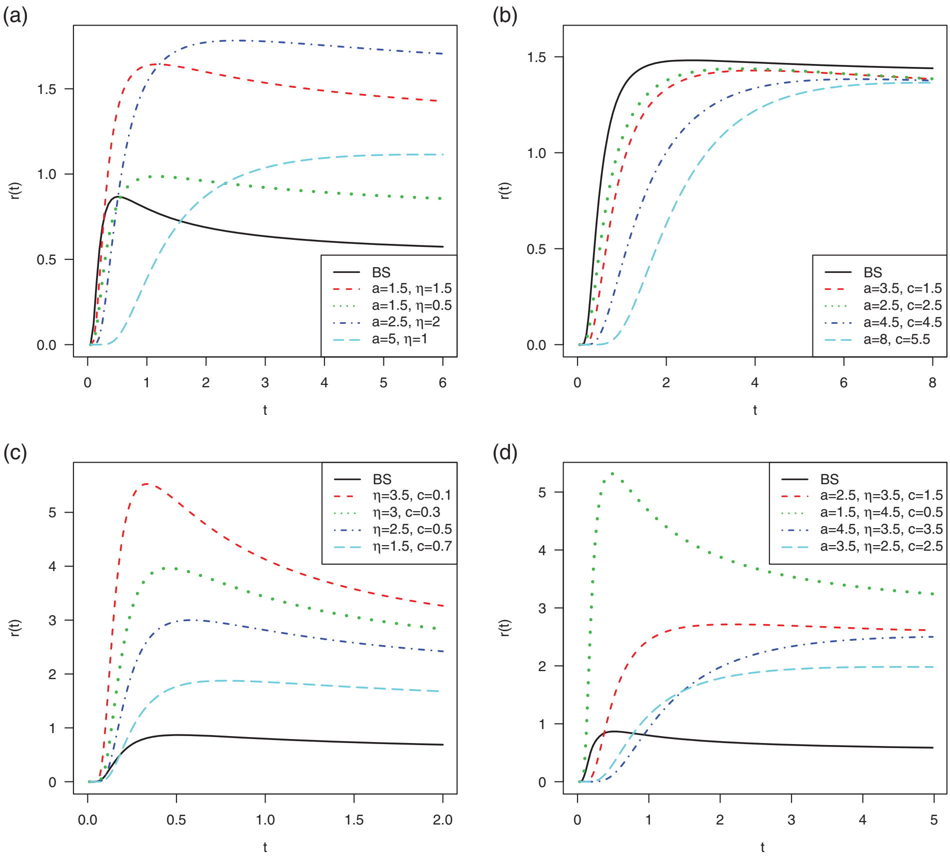

Figure 2. Plots of the hazard rate function (9) for some parameter values. (a) For values $\alpha = \beta = 1.0$ and $c = 1.5$; (b) for values $\alpha = \beta = 1.0$ and $\eta = 1.5$; (c) for values $\alpha = \beta = 1.0$ and $a = 1.5$ and (d) for values $\alpha = \beta = 1.0$.

values are given in Figures 1 and 2, respectively. The density function and hazard rate function can take various forms, depending on the parameter values. It is evident that the shapes of the McBS distribution are much more flexible than the BS distribution. Figure 3 provides some relationships among the models defined from the McBS distribution.

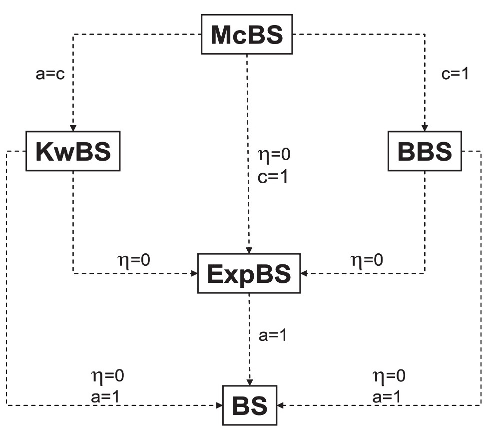

Figure 3. Relationships of the McBS sub-models.

The new model is easily simulated as follows: if V is a beta random variable with parameters a/c and $\eta+1$, then

$$T=\beta\left\{\frac{\alpha\Phi^{-1}(V)}{2}+\left[1+\frac{\alpha^2\Phi^{-1}(V)^2}{4}\right]^{1/2}\right\}^2$$

has the McBS(a,η,c,α,β) distribution. This scheme is useful because of the existence of fast generators for beta random variables and the standard normal quantile function is available in most statistical packages.

3. Density function expansion

An expansion for Equation (8) can be derived using the concept of exponentiated distributions. We define a random variable Y having the exponentiated BS (EBS) distribution with parameters α, β and $a>0$, say $Y\sim \text{EBS}(\alpha,\beta,a)$, if its cdf and pdf are given by $H(y;\alpha,\beta,a)=\Phi(v)^a$ and $h(y;\alpha,\beta,a)=ag_{\alpha,\beta}(y)\Phi(v)^{a-1}$, respectively, where $v=\alpha^{-1}\rho(y/\beta)$. The properties of some exponentiated distributions have been studied by several authors. In particular, the reader is referred to [24–27].

By expanding the binomial in Equation (8), we obtain the linear combination representation (with $b=\eta+1$)

$$f(t)=\sum_{k=0}^{\infty}w_k h(t;\alpha,\beta,kc+a), \tag{10}$$

where $h(t;\alpha,\beta,kc+a)$ denotes the EBS$(\alpha,\beta,kc+a)$ density function and the weights w_k are given by

$$w_k=\frac{(-1)^k c\binom{\eta}{k}}{B(ac^{-1},\eta+1)(kc+a)}.$$

The McBS density function is then a linear combination of EBS density functions. The EBS$(\alpha,\beta,kc+a)$ density function follows directly from Equation (8) by setting $\eta=0$ and $c=1$ and replacing a by $kc+a$. So, some McBS properties can be obtained by knowing those of the EBS distribution.

By integrating Equation (10), we obtain

$$F(t)=\sum_{k=0}^{\infty}w_k\Phi(v)^{kc+a}. \tag{11}$$

If a is a positive real non-integer, we can expand $\Phi(v)^a$ as

$$\Phi(v)^a=\sum_{r=0}^{\infty}s_r(a)\Phi(v)^r, \tag{12}$$

where

$$s_r(m)=\sum_{j=r}^{\infty}(-1)^{r+j}\binom{m}{j}\binom{j}{r}.$$

Thus, from Equations (2), (10) and (12), we can write

$$f(t) = g_{\alpha,\beta}(t) \sum_{r=0}^{\infty} e_r \Phi(v)^r, \tag{13}$$

where $e_r = \sum_{k=0}^{\infty} d_k s_r(kc + a - 1)$.

4. Quantile measures

The McBS quantile function, say $Q(u) = F^{-1}(u)$, can be expressed in terms of the BS quantile function ($Q_{\rm BS}(\cdot)$) and beta quantile function ($Q_\beta(\cdot)$). The BS quantile function is straightforward to be computed from the standard normal quantile function $x = Q_N(u) = \Phi^{-1}(u)$ by [23]

$$Q_{\rm BS}(u) = \frac{\beta}{2}\{2 + \alpha^2 Q_N(u)^2 + \alpha Q_N(u)[4 + \alpha^2 Q_N(u)^2]^{1/2}\}. \tag{14}$$

Let $w = Q_\beta(u)$ be the quantile function of a beta random variable with parameters ac^{-1} and $\eta + 1$. By inverting $F(t) = I_{\Phi(v)^c}(ac^{-1}, \eta + 1) = u$, the McBS quantile function can be determined by $t = Q(u) = Q_{\rm BS}(Q_\beta(u)^{1/c})$.

The effect of the shape parameters a, b and c on skewness and kurtosis can be considered based on quantile measures. The shortcomings of the classical kurtosis measure are well known. One of the earliest skewness measures to be suggested is the Bowley skewness [28] 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 [29] is based on octiles

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

These measures 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 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, $Q(3/8) - Q(1/8) \to 0$ and $Q(7/8) - Q(5/8) \to 0$, $M \to 0$ and half of the total probability mass is concentrated in the neighbourhoods of the octiles $Q(2/8)$ and $Q(6/8)$. Clearly, $M > 0$ and there is a good concordance with the usual kurtosis measures for some distributions. For the normal distribution, $B = M = 0$.

Figures 4–6 provide plots of the measures B and M for some parameter values, respectively. These plots show that both measures depend on all shape parameters. Figure 6 also shows that they can be very sensitive to the extra third parameter c even in the case when $a = b = \eta + 1$. The BBS distribution has only two extra shape parameters and so it can add only a limited structure to the skewness and kurtosis of the generated distribution.

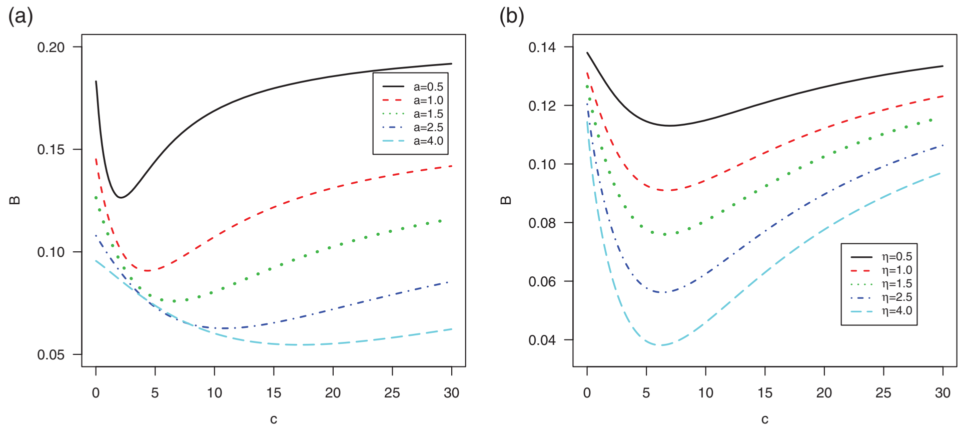

Figure 4. Plots of the measure B for some parameter values. (a) For values $\alpha = 0.5$, $\beta = 1.0$ and $\eta = 1.5$ and (b) for values $\alpha = 0.5$, $\beta = 1.0$ and $a = 1.5$.

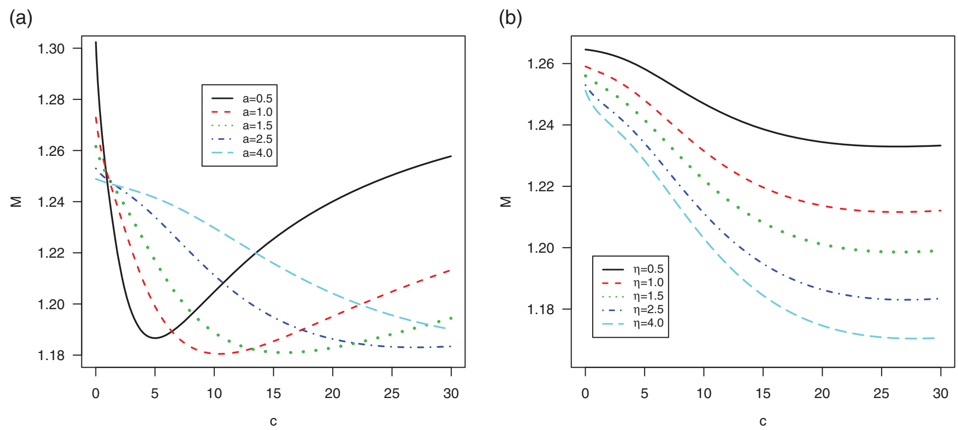

Figure 5. Plots of the measure M for some parameter values. (a) For values $\alpha = 0.5$, $\beta = 1.0$ and $\eta = 1.5$ and (b) for values $\alpha = 0.5$, $\beta = 1.0$ and $a = 1.5$.

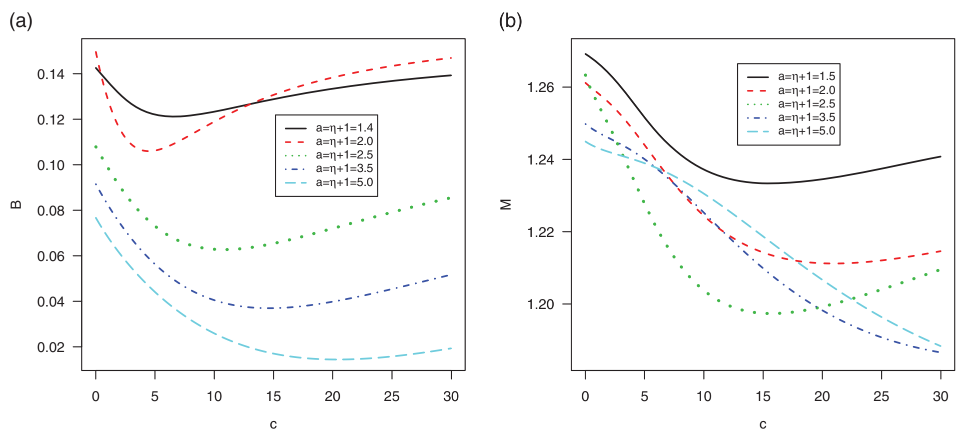

Figure 6. Plots of the measures (a) B and (b) M for some parameter values with $\alpha = 0.5$ and $\beta = 1.0$.

5. Quantile expansion function

Power series methods are at the heart of many aspects of applied mathematics and statistics. In this section, we derive a power series expansion for $Q(u)$ that can be useful to determine some mathematical measures of the new distribution. Let $F(t) = I_w(ac^{-1}, \eta + 1) = u$, where $w = \Phi(v)^c$. We can express w as a power series expansion of u given by [23]

$$w = \sum_{r=0}^{\infty} f_r u^{rc/a}, \tag{15}$$

where $f_0 = 0$ and $f_i = q_i\{ac^{-1}B(ac^{-1}, \eta + 1)\}^{i/a}$ for $i \geq 1$, and the quantities q_i's for $i \geq 2$ can be obtained from the cubic recursive equation

$$q_i = \frac{1}{[i^2 + (ac^{-1} - 2)i + (1 - ac^{-1})]} \Bigg\{ (1 - \delta_{i,2}) \sum_{r=2}^{i-1} q_r q_{i+1-r}[r(1 - ac^{-1})(i - r) - r(r - 1)]$$
$$+ \sum_{r=1}^{i-1} \sum_{s=1}^{i-r} q_r q_s q_{i+1-r-s}[r(r - ac^{-1}) + s(ac^{-1} + \eta - 1)(i + 1 - r - s)] \Bigg\},$$

where $\delta_{i,2} = 1$ if $i = 2$ and $\delta_{i,2} = 0$ if $i \neq 2$. The quadratic term in the last expression contributes only for $i \geq 3$. We have $q_0 = 0$, $q_1 = 1$, $q_2 = \eta/(ac^{-1} + 1)$, $q_3 = \eta[a^2c^{-2} + 3ac^{-1}(\eta + 1) - ac^{-1} + 5\eta + 1]/[2(ac^{-1} + 1)^2(ac^{-1} + 2)]$ and

$$q_4 = \frac{\eta[a^4c^{-4} + (6\eta + 5)a^3c^{-3} + (\eta + 3)(8\eta + 3)a^2c^{-2} + (33\eta^2 + 36\eta + 7)ac^{-1} + (\eta + 1)(31\eta - 16) + 18]}{[3(ac^{-1} + 1)^3(ac^{-1} + 2)(ac^{-1} + 3)]}.$$

Equation (15) yields w in $(0, 1)$, since it is an expansion for the beta quantile function.

We use throughout an equation in Section 0.314 of Gradshteyn and Ryzhik [30] for a power series raised to a positive integer j given by

$$\left(\sum_{i=0}^{\infty} a_i x^i \right)^j = \sum_{i=0}^{\infty} c_{j,i} x^i, \tag{16}$$

whose coefficients $c_{j,i}$ (for $i = 1, 2, \ldots$) are determined from the recurrence equation

$$c_{j,i} = (ia_0)^{-1} \sum_{m=1}^{i} (jm - i + m) a_m c_{j,i-m} \tag{17}$$

and $c_{j,0} = a_0^j$. Hence, the coefficients $c_{j,i}$ can be computed directly from $c_{j,0}, \ldots, c_{j,i-1}$ and, therefore, from $a_0, \ldots, a_i$. They can be given explicitly in terms of the a_i's, although it is not necessary for programming numerically our expansions in any algebraic or numerical software.

Following again Cordeiro and Lemonte [23], we can invert $\Phi(v) = w^{1/c}$ if the condition $-2 < (t/\beta)^{1/2} - (\beta/t)^{1/2} < 2$ holds, to express t as a power series expansion of w

$$t = Q(u) = \sum_{k=0}^{\infty} m_k \left(w^{1/c} - \frac{1}{2} \right)^k. \tag{18}$$

The coefficients m_k in Equation (18) can be expressed in terms of known constants

$$m_k = (2\pi)^{k/2} \sum_{j=0}^{\infty} p_j e_{j,k},$$

where $p_0 = \beta$, $p_{2j+1} = \beta\alpha^{2j+1} \binom{1/2}{j} 2^{-2j}$ for $j \geq 0$, $p_2 = \beta\alpha^2/2$ and $p_{2j} = 0$ for $j \geq 2$ and the quantities $e_{j,k}$ follow recursively from Equations (16) and (17) by $e_{j,0} = d_0^j$ and

$$e_{j,k} = (kd_0)^{-1} \sum_{m=1}^{k} (jm - k + m) d_m e_{j,k-m}.$$

Here, the quantities d_m are defined by $d_m = 0$ (for $m = 0, 2, 4, \ldots$) and $d_m = b_{(m-1)/2}$ (for $k = 1, 3, 5, \ldots$), where the b_m's are calculated recursively from

$$b_{m+1} = \frac{1}{2(2m+3)} \sum_{r=0}^{m} \frac{(2r+1)(2m-2r+1) b_r b_{m-r}}{(r+1)(2r+1)}.$$

We have $b_0 = 1$, $b_1 = 1/6$, $b_2 = 7/120$, $b_3 = 127/7560, \ldots$.

Combining Equations (15) and (18), we obtain

$$t = \sum_{k=0}^{\infty} m_k \left[\left(\sum_{r=0}^{\infty} f_r u^{rc/a} \right)^{1/c} - \frac{1}{2} \right]^k.$$

Hence,

$$t = \sum_{k=0}^{\infty} m_k \sum_{i=0}^{k} \binom{k}{i} \left(-\frac{1}{2}\right)^{k-i} \left(\sum_{r=0}^{\infty} f_r u^{rc/a} \right)^{i/c}.$$

Since the last sum belongs to the interval $(0, 1)$, we can use Equation (12) to obtain

$$\left[\sum_{r=0}^{\infty} f_r u^{rc/a} \right]^{i/c} = \sum_{m=0}^{\infty} s_m(i/c) \left(\sum_{r=0}^{\infty} f_r u^{rc/a} \right)^m,$$

where

$$s_m(i/c) = \sum_{j=m}^{\infty} (-1)^{m+j} \binom{i/c}{j} \binom{j}{m}.$$

Further, we can write

$$\left(\sum_{r=0}^{\infty} f_r u^{rc/a} \right)^m = \sum_{r=0}^{\infty} c_{m,r} u^{rc/a}, \tag{19}$$

whose coefficients $c_{m,r}$ (for $r = 1, 2, \ldots$) can be obtained from $c_{m,0} = f_0^m$ and

$$c_{m,r} = (rf_0)^{-1} \sum_{l=1}^{r} (ml - r + l) f_l c_{m,r-l}.$$

Hence,

$$t = Q(u) = \sum_{r=0}^{\infty} q_r u^{rc/a}, \tag{20}$$

where

$$q_r = \sum_{k,m=0}^{\infty} m_k c_{m,r} \sum_{i=0}^{k} \binom{k}{i} \left(-\frac{1}{2}\right)^{k-i} s_m(i/c). \tag{21}$$

Equation (20) provides a power series expansion for the McBS quantile function. Some of its measures of interest (such as moments and mgf) can be derived from Equation (20).

6. Moments

The ordinary moments of the McBS random variable T can be derived from the probability weighted moments [31] of the BS distribution formally defined for p and r non-negative integers by

$$\tau_{p,r} = \int_0^{\infty} t^p g_{\alpha,\beta}(t) \Phi(v)^r \, \mathrm{d}t. \tag{22}$$

The integral (22) can be easily computed numerically in software such as `MAPLE`, `MATLAB`, `MATHEMATICA`, `Ox` and `R`. Cordeiro and Lemonte [23] proposed an alternative representation to compute $\tau_{p,r}$ as

$$\tau_{p,r} = \frac{\beta^p}{2^r} \sum_{j=0}^{r} \binom{r}{j} \sum_{k_1,\ldots,k_j=0}^{\infty} A(k_1,\ldots,k_j) \sum_{m=0}^{2s_j+j} (-1)^m \binom{2s_j+j}{m} I\left(p + \frac{(2s_j+j-2m)}{2}, \alpha\right),$$

where $s_j = k_1 + \cdots + k_j$, $A(k_1,\ldots,k_j) = \alpha^{-2s_j-j} a_{k_1},\ldots,a_{k_j}$, $a_k = (-1)^k 2^{(1-2k)/2}[\sqrt{\pi}(2k+1)k!]^{-1}$ and $I(p + (2s_j + j - 2m)/2, \alpha)$ is calculated from Equation (4).

The sth moment of T can be written from Equation (13) as

$$\mu'_s = \mathrm{E}(T^s) = \sum_{r=0}^{\infty} e_r \tau_{s,r}, \tag{23}$$

where $\tau_{s,r}$ is obtained from Equation (22) and e_r is given after Equation (13). Equation (23) can be computed numerically in any symbolic software (e.g. `MAPLE`, `MATLAB` and `MATHEMATICA`) by taking in the sum a large number of summands. These algebraic software have currently the ability to deal with analytic expressions of formidable size and complexity.

As a simple application of Equation (20), we obtain an alternative expression for μ'_s (for a real non-integer) using the quantile function $Q(u)$. We can write from Equation (13)

$$\mu'_s = \sum_{p=0}^{\infty} e_r \int_0^1 Q(u)^s u^p \, \mathrm{d}u.$$

Further, we can easily evaluate the integral in (0, 1) from Equation (20) as

$$\int_0^1 \left(\sum_{r=0}^{\infty} q_r u^{rc/a}\right)^s u^p \, \mathrm{d}u = \sum_{r=0}^{\infty} d_{s,r} \int_0^1 u^{p+rc/a} \, \mathrm{d}u = \sum_{r=0}^{\infty} \frac{d_{s,r}}{(p+1+rc/a)},$$

where $d_{s,r}$ follows the recurrence equation (with $d_{s,0} = q_0^s$)

$$d_{s,r} = (rq_0)^{-1} \sum_{m=1}^{t} (sm - r + m) q_m d_{s,t-m}.$$

Finally, we obtain

$$\mu_s' = \sum_{p,r=0}^{\infty} \frac{e_r d_{s,r}}{(p + 1 + rc/a)},$$

which is an alternative formula to Equation (23).

Plots of the skewness and kurtosis of the McBS distribution as a function of c for selected values of a and η for $\alpha = 0.5$ and $\beta = 1.0$ are shown in Figures 7 and 8, respectively.

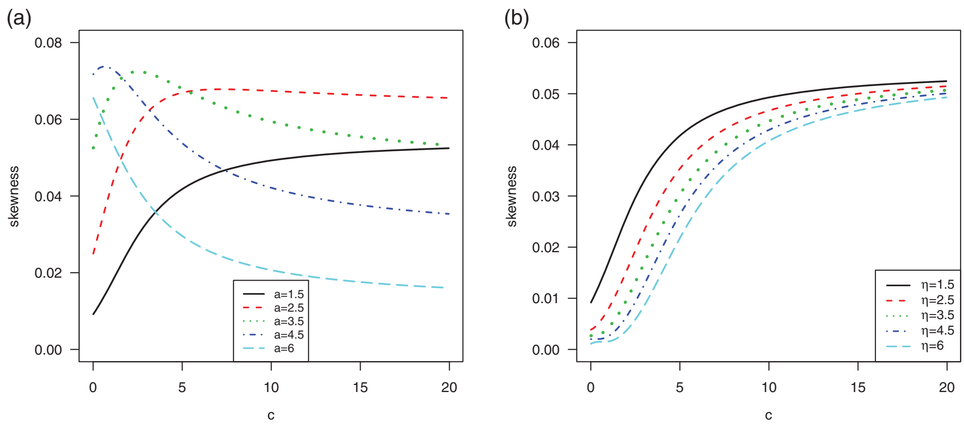

Figure 7. Plots of the skewness of the McBS distribution as a function of c for some values of a and η for $\alpha = 0.5$ and $\beta = 1.0$. (a) $\eta = 1.5$ and (b) $a = 1.5$.

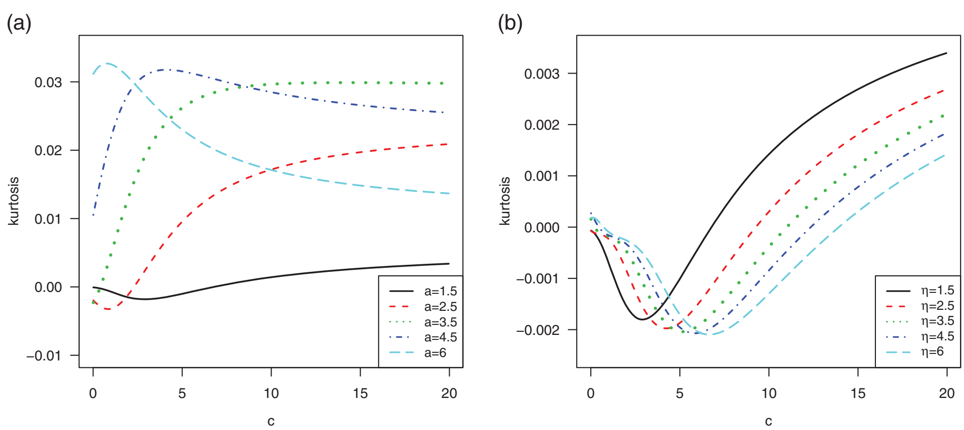

Figure 8. Plots of the kurtosis of the McBS distribution as a function of c for some values of a and η for $\alpha = 0.5$ and $\beta = 1.0$. (a) $\eta = 1.5$ and (b) $a = 1.5$.

7. Generating function

Here, we provide two representations for the mgf of the McBS(a, b, c, α, β) distribution, say $M(s) = \mathrm{E}\{\exp(sX)\}$. From expansion (13), we obtain a first expansion

$$M(s) = \sum_{r=0}^{\infty} e_r \int_0^{\infty} \exp(st) g_{\alpha,\beta}(t) \Phi(v)^r \,\mathrm{d}t = \sum_{r,p=0}^{\infty} \frac{e_r \tau_{p,r}}{p!} s^p.$$

A second representation for the mgf is based on the quantile expansion (20). We have

$$M(s) = \int_0^{\infty} \exp(st) f(t) \,\mathrm{d}t = \int_0^1 \exp\left\{ s \left(\sum_{r=0}^{\infty} q_r u^{rc/a} \right) \right\} \mathrm{d}u,$$

where q_r is given by Equation (21). The polynomial expansion

$$M(s) = \sum_{i,r=0}^{\infty} \frac{p_{i,r} s^i}{i!} \int_0^1 u^{rc/a} \,\mathrm{d}u = \sum_{i=0}^{\infty} t_i s^i,$$

where

$$t_i = \sum_{r=0}^{\infty} \frac{a p_{i,r}}{(rc + a) i!}$$

follows and the quantities $p_{i,r}$ can be obtained from the recurrence equation (for $i = 1, 2, \ldots$) (with $p_{i,0} = q_0^i$)

$$p_{i,r} = (r q_0)^{-1} \sum_{m=1}^{r} (im - r + m) q_m p_{i,r-m}.$$

8. Mean deviations

The deviations from the mean and from the median can be used as a measure of spread in a population. Let T be a random variable having the McBS$(a, \eta, c, \alpha, \beta)$ distribution. We can derive the mean deviations about the mean and about the median from the relations

$$\delta_1(T) = \mathrm{E}(|T - \mu_1'|) = \int_0^{\infty} |t - \mu_1'| f(t) \,\mathrm{d}t \quad \text{and} \quad \delta_2(T) = \mathrm{E}(|T - m|) = \int_0^{\infty} |t - m| f(t) \,\mathrm{d}t,$$

respectively, where the mean $\mu_1' = \mathrm{E}(T)$ is calculated from Equation (23) and the median m is given by $m = Q_{\mathrm{BS}}(Q_\beta(0.5)^{1/c})$. The measures $\delta_1(T)$ and $\delta_2(T)$ can be expressed as

$$\delta_1(T) = 2\mu_1' F(\mu_1') - 2J(\mu_1') \quad \text{and} \quad \delta_2(T) = \mu_1' - 2J(m),$$

where $F(\mu_1') = I_{\Phi(w)^c}(ac^{-1}, \eta + 1)$, $J(q) = \int_0^q t f(t) \,\mathrm{d}t$ and $w = \alpha^{-1} \rho(\mu_1'/\beta)$. In what follows, we obtain an expression for the integral $J(q)$. From Equation (13), $J(q)$ can be written as

$$J(q) = \sum_{r=0}^{\infty} e_r \rho(q, r), \tag{24}$$

where $\rho(q,r) = \int_0^q t g_{\alpha,\beta}(t)\Phi(v)^r \, \mathrm{d}t$. From Cordeiro and Lemonte [23], we have

$$\Phi(v)^r = \frac{1}{2^r}\sum_{j=0}^{r}\binom{r}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\beta^{-(2s_j+j)/2}A(k_1,\ldots,k_j)\sum_{m=0}^{2s_j+j}(-\beta)^m\binom{2s_j+j}{m}t^{(2s_j+j-2m)/2},$$

where s_j and $A(k_1,\ldots,k_j)$ are defined in Section 6. Thus,

$$\rho(q,r) = \frac{\kappa(\alpha,\beta)}{2^r}\sum_{j=0}^{r}\binom{r}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\beta^{-(2s_j+j)/2}A(k_1,\ldots,k_j)\sum_{m=0}^{2s_j+j}(-\beta)^m\binom{2s_j+j}{m}$$
$$\times\int_0^q t^{(2s_j+j-2m-1)/2}(t+\beta)\exp\left\{-\frac{\tau(t/\beta)}{2\alpha^2}\right\}\mathrm{d}t.$$

Let

$$D(p,q) = \int_0^q t^p\exp\left\{-\frac{(t/\beta+\beta/t)}{2\alpha^2}\right\}\mathrm{d}t = \beta^{p+1}\int_0^{q/\beta}u^p\exp\left\{-\frac{(u+u^{-1})}{2\alpha^2}\right\}\mathrm{d}u.$$

From Terras [32], we can write

$$D(p,q) = 2\beta^{p+1}K_{p+1}(\alpha^{-2}) - q^{p+1}K_{p+1}\left(\frac{q}{2\alpha^2\beta},\frac{\beta}{2\alpha^2 q}\right),$$

where $K_\nu(x_1,x_2)$ denotes the incomplete Bessel function with arguments x_1 and x_2 and order ν (for further details, see [33–35]). Hence, we obtain

$$\rho(q,r) = \frac{\kappa(\alpha,\beta)}{2^r}\sum_{j=0}^{r}\binom{r}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\beta^{-(2s_j+j)/2}A(k_1,\ldots,k_j)\sum_{m=0}^{2s_j+j}(-\beta)^m$$
$$\times\binom{2s_j+j}{m}\left\{D\left(\frac{2s_j+j-2m+1}{2},q\right)+\beta D\left(\frac{2s_j+j-2m-1}{2},q\right)\right\},$$

which can be calculated from the function $D(p,q)$. Hence, we can use this expression for $\rho(q,r)$ to compute $J(q)$ from Equation (24).

From Equation (24), we obtain the Bonferroni and Lorenz curves defined by $B(\pi) = J(q)/(\pi\mu_1')$ and $L(\pi) = J(q)/(\mu_1')$, respectively, where $q = Q(\pi) = Q_{\mathrm{BS}}(Q_\beta(\pi)^{1/c})$ is calculated for a given probability π. These curves have applications in economics, reliability, demography, insurance and medicine.

9. Reliability

In the context of reliability, the stress–strength model describes the life of a component which has a random strength T_1 that is subjected to a random stress T_2. The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $T_1 > T_2$. Hence, $R = \Pr(T_2 < T_1)$ is a measure of component reliability which has many applications in engineering. Here, we derive the reliability R when T_1 and T_2 have independent McBS$(a_1,\eta_1,c_1,\alpha,\beta)$ and McBS$(a_2,\eta_2,c_2,\alpha,\beta)$ distributions with the same shape parameters α and β.

The pdf of T_1 and the cdf of T_2 can be written from Equations (10) and (11) as

$$f_1(t) = \sum_{k=0}^{\infty} w_{1k}(kc_1 + a_1)g_{\alpha,\beta}(t)\Phi(v)^{kc_1+a_1-1} \quad \text{and} \quad F_2(t) = \sum_{j=0}^{\infty} w_{2j}\Phi(v)^{jc_2+a_2},$$

respectively, where

$$w_{1k} = \frac{(-1)^k c_1 \binom{\eta_1}{k}}{B(a_1 c_1^{-1}, \eta_1 + 1)(kc_1 + a_1)} \quad \text{and} \quad w_{2j} = \frac{(-1)^j c_2 \binom{\eta_2}{j}}{B(a_2 c_2^{-1}, \eta_2 + 1)(jc_2 + a_2)}.$$

We have

$$R = \int_0^{\infty} f_1(t)F_2(t)\,\mathrm{d}t$$

and then

$$R = \sum_{k,j=0}^{\infty} w_{1k}w_{2j}(kc_1 + a_1)\int_0^{\infty} g_{\alpha,\beta}(t)\Phi(v)^{kc_1+a_1+jc_2+a_2-1}.$$

From Equation (12), we can write

$$\Phi(v)^{kc_1+a_1+jc_2+a_2-1} = \sum_{r=0}^{\infty} s_r(kc_1 + a_1 + jc_2 + a_2 - 1)\Phi(v)^r$$

and then R reduces to

$$R = \sum_{k,j=0}^{\infty} w_{1k}w_{2j}(kc_1 + a_1)\sum_{r=0}^{\infty} s_r(kc_1 + a_1 + jc_2 + a_2 - 1)\tau_{0,r}.$$

10. Estimation

The estimation of the model parameters of the McBS distribution will be investigated by maximum likelihood. Let $\boldsymbol{t} = (t_1, \ldots, t_n)^\top$ denote a random sample of size n obtained from the McBS model and let $\boldsymbol{\theta} = (a, \eta, c, \alpha, \beta)^\top$ be the parameter vector. The total log-likelihood function for $\boldsymbol{\theta}$ can be reduced to

$$\begin{aligned}\ell(\boldsymbol{\theta}) = n\log\{c\kappa(\alpha,\beta)\} - n\log\{B(ac^{-1}, \eta + 1)\} - \frac{3}{2}\sum_{i=1}^{n}\log(t_i) + \sum_{i=1}^{n}\log(t_i + \beta)\\ - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\tau\left(\frac{t_i}{\beta}\right) + (a-1)\sum_{i=1}^{n}\log\{\Phi(v_i)\} + \eta\sum_{i=1}^{n}\log\{1 - \Phi(v_i)^c\}.\end{aligned}$$

By taking the derivatives of the log-likelihood function in relation to the parameters, the components of the score vector $\boldsymbol{U_\theta} = (U_a, U_\eta, U_c, U_\alpha, U_\beta)^\top$ are given by

$$U_a = -\frac{n\psi(ac^{-1})}{c} + \frac{n\psi(ac^{-1} + \eta + 1)}{c} + \sum_{i=1}^{n}\log\{\Phi(v_i)\},$$

$$U_\eta = -n\psi(\eta+1) + n\psi(ac^{-1}+\eta+1) + \sum_{i=1}^{n} \log\{1-\Phi(v_i)^c\},$$

$$U_c = \frac{n}{c} + \frac{n\psi(ac^{-1})a}{c^2} - \frac{n\psi(a+\eta+1)a}{c^2} - \eta \sum_{i=1}^{n} \frac{\Phi(v_i)^c \log\{\Phi(v_i)\}}{1-\Phi(v_i)^c},$$

$$U_\alpha = -\frac{n}{\alpha}\left(1+\frac{2}{\alpha^2}\right) + \frac{1}{\alpha^3}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right) - \frac{1}{\alpha}\sum_{i=1}^{n}\frac{v_i\phi(v_i)}{\Phi(v_i)}\left\{(a-1) - \frac{\eta c\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c}\right\},$$

$$U_\beta = -\frac{n}{2\beta} + \sum_{i=1}^{n}\frac{1}{t_i+\beta} + \frac{1}{2\alpha^2\beta}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}-\frac{\beta}{t_i}\right)$$

$$- \frac{1}{2\alpha\beta}\sum_{i=1}^{n}\frac{\tau(\sqrt{t_i/\beta})\phi(v_i)}{\Phi(v_i)}\left\{(a-1) - \frac{\eta c\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c}\right\},$$

where $\phi(\cdot)$ is the standard normal density function, $\psi(\cdot)$ is the digamma function, $v_i = \alpha^{-1}\rho(t_i/\beta) = \alpha^{-1}\{(t_i/\beta)^{1/2} - (\beta/t_i)^{1/2}\}$ and $\tau(\sqrt{t_i/\beta}) = (t_i/\beta)^{1/2} + (\beta/t_i)^{1/2}$ for $i = 1, \ldots, n$. The maximum-likelihood estimate (MLE) $\hat{\boldsymbol{\theta}} = (\hat{a}, \hat{\eta}, \hat{c}, \hat{\alpha}, \hat{\beta})^\top$ of $\boldsymbol{\theta} = (a, \eta, c, \alpha, \beta)^\top$ is obtained by setting $U_a = 0$, $U_\eta = 0$, $U_c = 0$, $U_\alpha = 0$ and $U_\beta = 0$ and solving the nonlinear equations simultaneously. These equations can be solved by iterative techniques such as a Newton–Raphson-type algorithm to obtain the estimate $\hat{\boldsymbol{\theta}}$. The Broyden–Fletcher–Goldfarb–Shanno method [36,37] with analytical derivatives has been used to maximize $\ell(\boldsymbol{\theta})$. We considered the `Ox` matrix programming language [38] to obtain the MLEs of the model parameters by using the subroutine MaxBFGS.

The normal approximation for the MLE of $\boldsymbol{\theta}$ can be used to construct approximate confidence intervals and for testing hypotheses on the parameters a, η, c, α and β. Under conditions that are fulfilled for the parameters in the interior of the parameter space, we obtain $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{A}{\sim} \mathcal{N}_5(\mathbf{0}, \boldsymbol{K}_{\boldsymbol{\theta}}^{-1})$, where $\overset{A}{\sim}$ means approximately distributed and $\boldsymbol{K}_{\boldsymbol{\theta}}$ is the unit expected information matrix. The asymptotic result $\boldsymbol{K}_{\boldsymbol{\theta}} = \lim_{n\to\infty} n^{-1}\boldsymbol{J}_n(\boldsymbol{\theta})$ holds, where $\boldsymbol{J}_n(\boldsymbol{\theta})$ is the observed information matrix. The average matrix evaluated at $\widehat{\boldsymbol{\theta}}$, say $n^{-1}\boldsymbol{J}_n(\hat{\boldsymbol{\theta}})$, can estimate $\boldsymbol{K}_{\boldsymbol{\theta}}$. The observed information matrix $\boldsymbol{J}_n(\boldsymbol{\theta}) = -\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$ is given in Appendix 1.

We can compute the maximum values of the unrestricted and restricted log-likelihood functions to obtain likelihood ratio (LR) statistics for testing some McBS sub-models. We consider the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ for the vector of parameters of the McBS distribution, where $\boldsymbol{\theta}_1$ is a subset of parameters of interest and $\boldsymbol{\theta}_2$ is a subset of the remaining parameters. The LR statistic for testing the null hypothesis $\mathcal{H}_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ against the alternative hypothesis $\mathcal{H}_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$ is given by $w = 2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\tilde{\boldsymbol{\theta}})\}$, where $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ are the MLEs under the alternative and null hypotheses, respectively, and $\boldsymbol{\theta}_1^{(0)}$ is a specified parameter vector. The statistic w is asymptotically (as $n \to \infty$) distributed as χ_k^2, where k is the dimension of the subset $\boldsymbol{\theta}_1$ of interest. For example, the McBS and BBS models are compared by testing $\mathcal{H}_0 : c = 1$ versus $\mathcal{H}_1 : c \neq 1$ and the LR statistic becomes $w = 2\{\ell(\hat{a}, \hat{\eta}, \hat{c}, \hat{\alpha}, \hat{\beta}) - \ell(\tilde{a}, \tilde{\eta}, 1, \tilde{\alpha}, \tilde{\beta})\}$, where $\hat{a}$, $\hat{\eta}$, $\hat{c}$, $\hat{\alpha}$ and $\hat{\beta}$ are the MLEs under $\mathcal{H}_1$ and $\tilde{a}$, $\tilde{\eta}$, $\tilde{\alpha}$ and $\tilde{\beta}$ are the estimates under $\mathcal{H}_0$.

As pointed out by an anonymous referee, when the expected information matrix is not available, the Wald statistic by considering the observed information matrix is used frequently in practice for hypothesis testing, especially in conjunction with higher-order asymptotic methods. So, this statistic may be used instead of the LR statistic since it avoids the potential problems in finding the restricted MLEs. Based on the referee's argument, a future research could be conducted to compare the LR and Wald statistics for testing its model parameters.

11. Log-McBS distribution

In this section, we extend the log-BS model [13] by replacing the BS distribution by the McBS distribution. First, let T be a random variable having the McBS density function (8). The random variable $W = \log(T)$ has a LMcBS distribution (also referred to as the beta sinh-normal distribution). After some algebra, the density function of W, parametrized in terms of $\mu = \log(\beta)$, can be expressed as

$$f_W(w) = \frac{c\xi_{01}\exp(-\xi_{02}^2/2)\Phi(\xi_{02})^{a-1}[1-\Phi(\xi_{02})^c]^\eta}{2\sqrt{2\pi}B(ac^{-1},\eta+1)}, \quad w \in \mathbb{R},$$

where $\xi_{01} = 2\alpha^{-1}\cosh((w-\mu)/2)$ and $\xi_{02} = 2\alpha^{-1}\sinh((w-\mu)/2)$. The parameter $\mu \in \mathbb{R}$ is a location parameter and a, η, c and α are positive shape parameters. Now, we define the standardized random variable $Z = (W-\mu)/2$ having density function

$$\pi(z) = \frac{2c\cosh(z)\exp\{-2\sinh^2(z)/\alpha^2\}}{B(ac^{-1},\eta+1)\sqrt{2\pi}\alpha}\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{a-1}\left[1-\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^c\right]^\eta, \quad z \in \mathbb{R}. \tag{25}$$

Some sub-models can be immediately obtained from Equation (25). The special cases $c = 1$ and $a = c$ correspond to the log-BBS (LBBS) and log-KwBS (LKwBS) distributions, respectively. Further, $c = 1$ and $\eta = 0$ yields the log-EBS model and $a = c = 1$, $\eta = 0$ and $\sigma = 2$ gives the sinh-normal distribution or LBS distribution [13].

Let $Y = \mu + \sigma Z$, whose density function takes the form

$$f(y) = \frac{c\xi_1\exp(-\xi_2^2/2)\Phi(\xi_2)^{a-1}[1-\Phi(\xi_2)^c]^\eta}{\sqrt{2\pi}\sigma B(ac^{-1},\eta+1)}, \quad -\infty < y < \infty, \tag{26}$$

where

$$\xi_1 = \frac{2}{\alpha}\cosh\left(\frac{y-\mu}{\sigma}\right) \quad \text{and} \quad \xi_2 = \frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right).$$

Here, $\sigma > 0$ is a kind of scale parameter. If Y is a random variable having density function (26), we write $Y \sim \text{LMcBS}(a,\eta,c,\alpha,\mu,\sigma)$. Thus, if $T \sim \text{McBS}(a,\eta,c,\alpha,\beta)$, then $Y = \mu + \sigma[\log(T)-\mu]/2 \sim \text{LMcBS}(a,\eta,c,\alpha,\mu,\sigma)$. The survival function corresponding to Equation (26) is

$$S(y) = 1 - \frac{1}{B(ac^{-1},\eta+1)}\int_0^{\Phi(\xi_2)^c} w^{ac^{-1}-1}(1-w)^\eta = 1 - I_{\Phi(\xi_2)^c}(ac^{-1},\eta+1).$$

Plots of the density function (26) for selected parameter values are given in Figure 9. These plots show great flexibility for different values of the shape parameters a, η and c. They indicate that the density function (26) is very flexible and hence can be used in many practical situations.

12. LMcBS generating function

Here, we obtain the mgf of the standardized LMcBS distribution (25) parametrized in terms of a, η, c and α. First, we can write from the binomial expansion

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{a-1}\left[1-\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^c\right]^\eta = \sum_{i=0}^{\infty}(-1)^i\binom{\eta}{i}\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{ci+a-1}$$

and then using Equation (12)

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{a-1}\left[1-\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{c}\right]^{\eta}=\sum_{i,r=0}^{\infty}(-1)^{i}\binom{\eta}{i}s_r(ci+a-1)\,\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{r},$$

where as before

$$s_r(ci+a-1)=\sum_{j=r}^{\infty}(-1)^{r+j}\binom{ci+a-1}{j}\binom{j}{r}.$$

Hence,

$$M_Z(s)=\sum_{i,r=0}^{\infty}p_{i,r}\int_{-\infty}^{\infty}\exp(sz)\cosh(z)\exp\left\{-\frac{2\sinh^2(z)}{\alpha^2}\right\}\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^{r}\,\mathrm{d}z,$$

where $p_{i,r}=p_{i,r}(a,\eta,c,\alpha)=(-1)^{i}2c(\frac{\eta}{i})s_r(ci+a-1)[B(ac^{-1},\eta+1)\sqrt{2\pi}\alpha]^{-1}$.

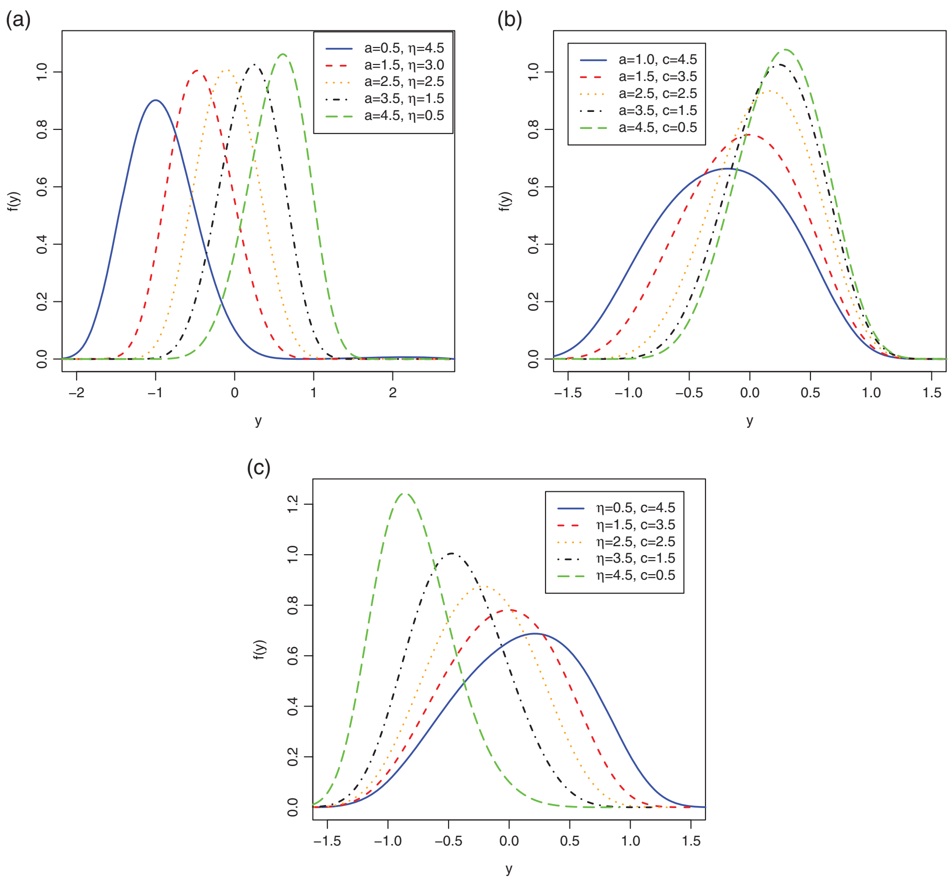

Figure 9. Plots of the LMcBS density for some parameter values. (a) For values $c=1.5$, $\alpha=1.5$, $\mu=0$ and $\sigma=1$; (b) for values $\eta=1.5$, $\alpha=1.5$, $\mu=0$ and $\sigma=1$ and (c) for values $a=1.5$, $\alpha=1.5$, $\mu=0$ and $\sigma=1$.

The last integral, say $I_r(s,\alpha)$, follows from the error function $\mathrm{erf}(\cdot)$

$$\Phi(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right] \quad \text{and} \quad \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-y^2)\,\mathrm{d}y$$

and its power series expansion $\mathrm{erf}(x/\sqrt{2}) = \sum_{m=0}^{\infty} b_m x^{2m+1}$, where

$$b_m = (-1)^m[(2m+1)2^{m/2}m!\sqrt{\pi}]^{-1}.$$

We have

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^r = \frac{1}{2^r}\left\{1 + \sum_{m=0}^{\infty} d_m \sinh(z)^{2m+1}\right\}^r,$$

where $d_m = 2^{2m+1} b_m \alpha^{-(2m+1)}$. Thus, using Equation (17), we obtain

$$\Phi\left(\frac{2}{\alpha}\sinh(z)\right)^r = \frac{1}{2^r}\sum_{k=0}^{r}\binom{r}{k}\left(\sum_{m=0}^{\infty} d_m \sinh(z)^{2m+1}\right)^k = \sum_{m=0}^{\infty} e_{m,r}\sinh(z)^{2m+1},$$

where $e_{m,r} = 2^{-r}\sum_{k=0}^{r}\binom{r}{k} g_{k,m}$, $g_{k,0} = d_0^j$ and

$$g_{k,m} = (id_0)^{-1}\sum_{\ell=1}^{m}(k\ell - m + \ell)d_\ell g_{k,m-\ell}.$$

Further,

$$I_r(s,\alpha) = \sum_{m=0}^{\infty} e_{m,r}\int_{-\infty}^{\infty}\exp(sz)\cosh(z)\sinh(z)^{2m+1}\exp\left\{-2\frac{\sinh^2(z)}{\alpha^2}\right\}\mathrm{d}z.$$

Now, using the identity $\cosh(2z) = 2\sinh^2(z) + 1$, the definition of $\sinh(z)$ and $\cosh(z)$, and expanding the binomial, we obtain after some algebra

$$\begin{aligned} I_r(s,\alpha) &= \sum_{m=0}^{\infty} e_{m,r}\int_{-\infty}^{\infty}\exp(sz)\cosh(z)\sinh(z)^{2m+1}\exp\left\{-\frac{2\sinh^2(z)}{\alpha^2}\right\}\mathrm{d}z \\ &= \exp\left(\frac{1}{\alpha^2}\right)\sum_{m=0}^{\infty}\frac{e_{m,r}}{2^{m+3}}\sum_{j=0}^{2m+1}(-1)^j\binom{2m+1}{j} \\ &\quad\times\int_{-\infty}^{\infty}\left\{\exp\left[\left(m+1-j+\frac{s}{2}\right)x\right] + \exp\left[\left(m-j+\frac{s}{2}\right)x\right]\right\}\exp\left\{-\frac{\cosh(x)}{\alpha^2}\right\}\mathrm{d}x. \end{aligned}$$

Using the integral representation $K_\nu(\beta) = 0.5\int_{-\infty}^{\infty}\exp\{-\beta\cosh(x) - \nu x\}\,\mathrm{d}x$, it follows that

$$\begin{aligned} I_r(s,\alpha) = \exp\left(\frac{1}{\alpha^2}\right)&\sum_{m=0}^{\infty}\frac{e_{m,r}}{2^{m+2}}\sum_{j=0}^{2m+1}(-1)^j\binom{2m+1}{j} \\ &\times\left[K_{-(m+1-j+s/2)}\left(\frac{1}{\alpha^2}\right) + K_{-(m-j+s/2)}\left(\frac{1}{\alpha^2}\right)\right]. \end{aligned} \tag{27}$$

Finally, the LMcBS generating function can be expressed as

$$M_Z(s) = \sum_{i,r=0}^{\infty} p_{i,r} I_r(s,\alpha),$$

where $I_r(s,\alpha)$ is calculated from Equation (27).

13. The LMcBS regression model

In many practical applications, the lifetimes t_i are affected by explanatory variables such as the cholesterol level, blood pressure and many other factors. Let $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^\top$ be the explanatory variable vector associated with the ith response variable y_i for $i = 1, \ldots, n$. Consider a sample $(y_1, \boldsymbol{x}_1), \ldots, (y_n, \boldsymbol{x}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(t_i), \log(c_i)\}$, and $\log(t_i)$ and $\log(c_i)$ are the log-lifetime and log-censoring, respectively. We assume non-informative censoring and assume that the observed lifetimes and censoring times are independent.

For the first time, we propose a linear regression model for the response variable y_i based on the LMcBS distribution given by

$$y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \ldots, n, \tag{28}$$

where the random error z_i follows the density function (25), $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is a p-vector $(p < n)$ of regression parameters, $\sigma > 0$, $a > 0$, $\eta > 0$ and $c > 0$ are unknown scalar parameters and $\boldsymbol{x}_i$ is the vector of explanatory variables modelling the location parameter $\mu_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}$. Hence, the location parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ of the LMcBS model has a linear structure $\boldsymbol{\mu} = \boldsymbol{X}\boldsymbol{\beta}$, where $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^\top$ is a known model matrix of full rank, i.e. $\operatorname{rank}(\boldsymbol{X}) = p$. The LBS (or the sinh-normal) regression model comes from Equation (28) with $a = c = 1$, $\eta = 0$ and $\sigma = 2$.

Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. The total log-likelihood function for the model parameters $\boldsymbol{\theta} = (a, \eta, c, \alpha, \sigma, \boldsymbol{\beta}^\top)^\top$ can be expressed from Equations (25) and (28) as

$$\begin{aligned}\ell(\boldsymbol{\theta}) = {} & q \log\left[\frac{(2\pi)^{-1/2} c}{B(ac^{-1}, \eta + 1)\sigma}\right] + \sum_{i \in F} \log(\xi_{1i}) - \frac{1}{2} \sum_{i \in F} \xi_{2i}^2 + (a - 1) \sum_{i \in F} \log\{\Phi(\xi_{2i})\} \\ & + \eta \sum_{i \in F} \log\{1 - \Phi^c(\xi_{2i})\} + \sum_{i \in C} \log\{1 - I_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)\},\end{aligned} \tag{29}$$

where q is the observed number of failures and

$$\xi_{1i} = \xi_{1i}(\boldsymbol{\theta}) = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mu_i}{\sigma}\right), \quad \xi_{2i} = \xi_{2i}(\boldsymbol{\theta}) = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mu_i}{\sigma}\right), \quad i = 1, \ldots, n.$$

The score functions for the parameters $a, \eta, c, \alpha, \sigma$ and β_j $(j = 1, \ldots, p)$ are given by

$$U_a(\boldsymbol{\theta}) = -\frac{q}{c}[\psi(ac^{-1}) - \psi(ac^{-1} + \eta + 1)] + \sum_{i \in F} \log[\Phi(\xi_{2i})] - \sum_{i \in C} \frac{[\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)]_a}{1 - I_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)},$$

$$\begin{aligned}U_\eta(\boldsymbol{\theta}) = {} & -q[\psi(\eta + 1) - \psi(ac^{-1} + \eta + 1)] + \sum_{i \in F} \log[1 - \Phi^c(\xi_{2i})] \\ & - \sum_{i \in C} \frac{[\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)]_\eta}{1 - I_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)},\end{aligned}$$

$$\begin{aligned}U_c(\boldsymbol{\theta}) = {} & \frac{q}{c}\left[1 - \frac{\psi(ac^{-1} + \eta + 1)}{c} + \frac{\psi(ac^{-1})}{c}\right] - \eta \sum_{i \in F} \frac{\Phi^c(\xi_{2i}) \log[\Phi(\xi_{2i})]}{1 - \Phi(\xi_{2i})} \\ & - \sum_{i \in C} \frac{[\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)]_c}{1 - I_{\Phi^c(\xi_{2i})}(ac^{-1}, \eta + 1)},\end{aligned}$$

$$U_\alpha(\boldsymbol{\theta}) = -\frac{2}{\alpha^2}\sum_{i\in F}\xi_{1i}^{-1}\cosh(z_i) + \frac{2}{\alpha^2}\sum_{i\in F}\xi_{2i}\sinh(z_i) - \frac{2(a-1)}{\alpha^2}\sum_{i\in F}\phi(\xi_{2i})\Phi^{-1}(\xi_{2i})\sinh(z_i)$$
$$+\frac{2\eta c}{\alpha^2}\sum_{i\in F}\frac{\phi(\xi_{2i})\Phi^{c-1}(\xi_{2i})\sinh(z_i)}{1-\Phi^c(\xi_{2i})} - \sum_{i\in C}\frac{[\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)]_\alpha}{1-I_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)},$$

$$U_\sigma(\boldsymbol{\theta}) = -\frac{2}{\alpha\sigma}\sum_{i\in F}\xi_{1i}^{-1}\sinh(z_i)z_i + \frac{2}{\alpha\sigma}\sum_{i\in F}\xi_{2i}\cosh(z_i)z_i - \frac{2(a-1)}{\alpha\sigma}\sum_{i\in F}\phi(\xi_{2i})\Phi^{-1}(\xi_{2i})$$
$$\times\cosh(z_i)z_i + \frac{2\eta c}{\alpha\sigma}\sum_{i\in F}\times\frac{\phi(\xi_{2i})\Phi^{c-1}(\xi_{2i})\cosh(z_i)z_i}{1-\Phi^c(\xi_{2i})} - \sum_{i\in C}\frac{[\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)]_\sigma}{1-I_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)},$$

$$U_{\beta_j}(\boldsymbol{\theta}) = -\frac{2}{\alpha\sigma}\sum_{i\in F}\xi_{1i}^{-1}\sinh(z_i)x_{ij} + \frac{2}{\alpha\sigma}\sum_{i\in F}\xi_{2i}\cosh(z_i)x_{ij} - \frac{2(a-1)}{\alpha\sigma}\sum_{i\in F}\phi(\xi_{2i})\Phi^{-1}(\xi_{2i})$$
$$\times\cosh(z_i)x_{ij} + \frac{2\eta c}{\alpha\sigma}\sum_{i\in F}\frac{\phi(\xi_{2i})\Phi^{c-1}(\xi_{2i})\cosh(z_i)x_{ij}}{1-\Phi^c(\xi_{2i})} - \sum_{i\in C}\frac{[\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)]_{\beta_j}}{1-I_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)},$$

where

$$z_i = \frac{y_i-\mu_i}{\sigma} \quad \text{and} \quad [\dot{I}_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)]_k = \frac{\partial I_{\Phi^c(\xi_{2i})}(ac^{-1},\eta+1)}{\partial k}.$$

The MLE $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by solving the nonlinear-likelihood equations $U_a(\boldsymbol{\theta})=0$, $U_\eta(\boldsymbol{\theta})=0$, $U_c(\boldsymbol{\theta})=0$, $U_\alpha(\boldsymbol{\theta})=0$, $U_\sigma(\boldsymbol{\theta})=0$ and $U_{\beta_j}(\boldsymbol{\theta})=0$. They cannot be solved analytically and statistical software can be used to solve the equations numerically. As initial values, we suggest

$$\bar{\boldsymbol{\beta}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y}, \quad \bar{\alpha} = \sqrt{\frac{4}{n}\sum_{i=1}^n\sinh^2\left(\frac{y_i-\boldsymbol{x}_i^\top\bar{\boldsymbol{\beta}}}{2}\right)},$$

for $\boldsymbol{\beta}$ and α, respectively, where $\boldsymbol{y}=(y_1,\ldots,y_n)^\top$, and take the values 1, 0, 1 and 2 for a, η, c and σ, respectively.

For interval estimation and tests of hypotheses on the parameters is necessary to obtain the $(p+5)\times(p+5)$ observed information matrix corresponding to the parameters a,η,c,α,σ and $\boldsymbol{\beta}$. This matrix is too complicated to be presented here. The LR statistic, for example, can be used to discriminate between the LBBS and LMcBS regression models since they are nested models. In this case, for testing the null hypothesis $\mathcal{H}_0: c=1$ against $\mathcal{H}_1:\mathcal{H}_0$ is not true, it becomes $w=2\{\ell(\hat{\boldsymbol{\theta}})-\ell(\tilde{\boldsymbol{\theta}})\}$, where $\tilde{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ under $\mathcal{H}_0$. The null hypothesis is rejected if $w>\chi^2_{1-\gamma}(1)$, where $\chi^2_{1-\gamma}(1)$ is the upper γth quantile of the chi-square distribution with one degree of freedom.

From the fitted model (28), the survival function for y_i can be estimated by

$$\hat{S}(y_i) = 1 - I_{\Phi(\hat{\xi}_{2i})^{\hat{c}}}(\hat{a}\hat{c}^{-1},\hat{\eta}+1), \quad i=1,\ldots,n, \tag{30}$$

where $\hat{\xi}_{2i}=\xi_{2i}(\hat{\boldsymbol{\theta}})$.

14. Applications

14.1. *Data: breaking stress*

First, we provide an application of the McBS model and its BBS, KwBS, exponentiated BS (EBS) and BS sub-models. We compare the results of the fits of these models by considering an uncensored data set from Nichols and Padgett [39] on breaking stress of carbon fibres (in Gba). The data are: 3.70, 2.74, 2.73, 2.50, 3.60, 3.11, 3.27, 2.87, 1.47, 3.11, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.90, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22, 3.39, 2.81, 4.20, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56, 3.15, 2.35, 2.55, 2.59, 2.38, 2.81, 2.77, 2.17, 2.83, 1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59, 3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2.00, 1.22, 1.12, 1.71, 2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.70, 2.03, 1.80, 1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, 3.65. All the computations were done using the `Ox` matrix programming language [38]. `Ox` is freely distributed for academic purposes and available at http://www.doornik.com.

Table 1 lists the MLEs of the model parameters and the following statistics: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hannan–Quinn Information Criterion (HQIC). These results show that the KwBS distribution has the lowest AIC, BIC and HQIC values among all fitted models, and so it could be chosen as the best model. Additionally, it is evident that the BS distribution presents the worst fit to the current data and then the proposed models outperform this distribution. In order to assess if the model is appropriate, the estimated pdf and cdf of the fitted distributions are shown in Figure 10. From these plots, we conclude that the McBS and KwBS models yield the best fits and that they could be adequate for these data. Note that the McBS and KwBS models present similar fitted density curves.

Table 1. MLEs and the measures AIC, BIC and HQIC.

	Estimates					Statistic		
Model	α	β	a	η	c	AIC	BIC	HQIC
McBS	5.8339	0.0366	18.9355	5.7196	45.3233	292.50	305.52	297.77
BBS	1.0681	43.2830	0.2232	263.5054		295.61	306.03	299.83
KwBS	4.1227	0.1650	16.9030	20.9144		290.57	300.99	294.78
EBS	0.0938	5.2869	0.0171			298.28	306.09	301.44
BS	0.4622	2.3660				304.12	309.33	306.23

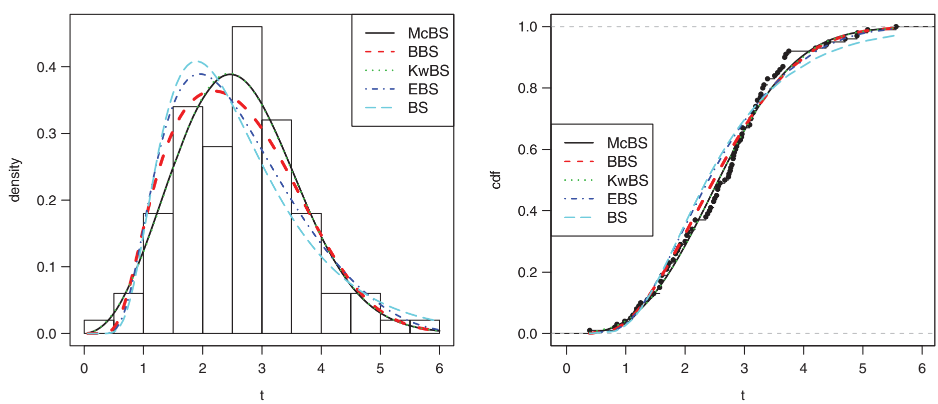

Figure 10. Estimated density and cumulative functions of the McBS, BBS, KwBS, EBS and BS distributions.

Table 2. LR tests.

Model	w	p-Value
McBS versus BBS	5.1108	0.0238
McBS versus KwBS	0.0674	0.7952
McBS versus EBS	9.7780	0.0075
McBS versus BS	17.6236	0.0005
BBS versus EBS	4.6672	0.0307
BBS versus BS	12.5128	0.0019
KwBS versus EBS	9.7106	0.0018
KwBS versus BS	17.5563	0.0002
EBS versus BS	7.8456	0.0051

In addition to comparing the models, we consider LR statistics and formal tests. First, the McBS model includes some sub-models, thus allowing their evaluation relative to each other and to a more general model. The values of the LR statistics for testing some sub-models of the McBS distribution are given in Table 2. The figures in this table indicate that there is no difference among the fits of the McBS and KwBS models to the current data. In addition, these two models provide a better representation for these data than the BS model based on the LR test at any usual significance level.

Secondly, we apply formal tests in order to verify which distribution better fits these data. We consider the Cramér–von Mises (W^*) and Anderson–Darling (A^*) statistics. In general, the smaller the values of the statistics W^* and A^*, the better the fit to the data. Let $H(x; \boldsymbol{\theta})$ be the cdf, where the form of H is known but $\boldsymbol{\theta}$ (a k-dimensional parameter vector, say) is unknown. To obtain the statistics W^* and A^*, we can proceed as follows: (i) compute $v_i = H(x_i; \hat{\boldsymbol{\theta}})$, where the x_i's are in ascending order, $y_i = \Phi^{-1}(v_i)$, where $\Phi^{-1}(\cdot)$ is the standard normal quantile function and $u_i = \Phi\{(y_i - \bar{y})/s_y\}$, where $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$ and $s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$; (ii) calculate $W^2 = \sum_{i=1}^{n}\{u_i - (2i-1)/(2n)\}^2 + 1/(12n)$ and $A^2 = -n - n^{-1}\sum_{i=1}^{n}\{(2i-1)\log(u_i) + (2n+1-2i)\log(1-u_i)\}$ and (iii) modify W^2 into $W^* = W^2(1 + 0.5/n)$ and A^2 into $A^* = A^2(1 + 0.75/n + 2.25/n^2)$. For further details, the reader is referred to Chen and Balakrishnan [40]. The values of the statistics W^* and A^* for all models are given in Table 3. Based on these statistics, it follows that the McBS model fits the current data better than the BBS, EBS and BS sub-models, and it is slightly better than the KwBS model.

14.2. *Data: components*

The data set consists of failure times (T) of eight components at three different temperatures. The data were obtained from Murthy *et al.* [41]. The original sample size was $n = 24$ components. The following variables are associated with each component: t_i, observed time (in years); x_{i1}, temperatures (temperature category: 100, 120 and 140), for $i = 1, 2, \ldots, 24$. We analyse these

Table 3. Formal tests.

	Statistic	
Distribution	W^*	A^*
McBS	0.06619	0.39066
BBS	0.11992	0.68449
KwBS	0.06826	0.40021
EBS	0.19431	1.04883
BS	0.29785	1.61816

data using the LMcBS regression model. First, we consider the regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \sigma z_i, \quad i = 1, \ldots, 24,$$

where the random errors z_i's ($i = 1, \ldots, n$) are independent random variables having density function (25).

Table 4 lists the MLEs of the parameters for the LMcBS, LBBS, LKwBS and LBS regression models fitted to these data using the NLMixed procedure in SAS. Iterative maximization of the logarithm of the likelihood function (29) starts with initial values for $\boldsymbol{\beta}$ and α taken from the fit of the LBS regression model with $a = c = 1$, $\eta = 0$ and $\sigma = 2$. We note from the fitted LMcBS regression model that x_1 is significant at 1% and that there is a significant difference among the levels of the temperature for the failure times.

The values of the AIC, Corrected AIC (CAIC) and BIC statistics to compare the LMcBS, LBBS, LKwBS and LBS regression models are given in Table 5. Note that the LMcBS and LKwBS regression models outperform the LBBS and LBS models irrespective of the criteria and then the proposed regression model can be used effectively in the analysis of these data. A comparison of the McBS regression model with some of its sub-models using LR statistics is performed in Table 6. The figures in this table, specially the p-values, indicate that the LMcBS and LKwBS regression models yield better fits to these data than the other sub-models.

A graphical comparison among the LMcBS, LBBS and LBS models is given in Figure 11. The curves displayed in this figure are the empirical survival function and the estimated survival

Table 4. MLEs (p-values between parentheses) for the LMcBS, LBBS, LKwBS and LBS regression models fitted to the component data.

Model	Estimates						
	a	η	c	α	σ	β_0	β_1
LMcBS	117.09	85.1230	0.3823	86.1542	0.5359	6.7077	−0.0306
						(<0.01)	(<0.01)
LBBS	9.1965	9.3392		0.7435	8.1351	7.8052	−0.0330
						(<0.01)	(<0.01)
LKwBS	2.0085	1E − 8		10.4550	0.5219	6.6927	−0.0307
						(<0.01)	(<0.01)
LBS				0.9317		7.6629	−0.0344
						(<0.01)	(<0.01)

Table 5. Statistics AIC, BIC and CAIC.

Model	Statistic		
	AIC	CAIC	BIC
LMcBS	46.7	67.7	68.9
LBBS	73.2	78.2	80.3
LKwBS	58.4	63.4	65.5
LBS	66.4	67.6	69.9

Table 6. LR statistics.

Model	Hypotheses	w	p-Value
LMcBS versus LBBS	$\mathcal{H}_0 : a = b = 1$ versus $\mathcal{H}_1 : \mathcal{H}_0$ is false	14.5	<0.01
LMcBS versus LKwBS	$\mathcal{H}_0 : a = c$ versus $\mathcal{H}_1 : \mathcal{H}_0$ is false	0.1	0.9512
LMcBS versus LBS	$\mathcal{H}_0 : a = b = 1, \eta = 0$ vs $\mathcal{H}_1 : \mathcal{H}_0$ is false	13.7	<0.01

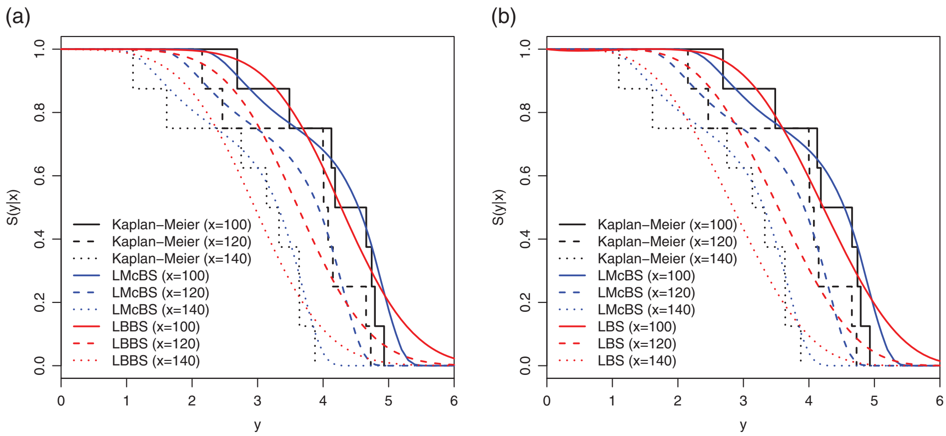

Figure 11. Estimated survival functions and the empirical survival: (a) LMcBS versus LBBS; (b) LMcBS versus LBS.

functions are given by Equation (30). Based on these plots, it is evident that the LMcBS model provides a superior fit.

15. Concluding remarks

We introduce a five-parameter continuous distribution, called the McBS distribution, which extends the BS and the BBS [23] distributions. We provide a mathematical treatment of the new distribution including expansions for the density function, moments, generating and quantile functions, mean deviations and reliability. The model parameters are estimated by the method of maximum likelihood and the observed information matrix is derived. An application of the new distribution to real data is given to show that it can provide consistently better fits than the other special models. Further, based on the logarithm of the McBS distribution, we propose an extended regression model which generalizes the well-known LBS regression model [13]. This extended regression model is very flexible and can be used in many practical situations. Its usefulness is also illustrated in an analysis of real data. Our formulas in connection with the new distribution and with the extended regression model are manageable, and with the use of modern computer resources and their analytic and numerical capabilities, the proposed models may prove to be an useful addition to the arsenal of applied statisticians.

Acknowledgements

We gratefully acknowledge grants from CNPq and FAPESP (Brazil). The authors thank two referees for comments and suggestions that led to a much improved paper.

References

[1] Z.W. Birnbaum and S.C. Saunders, *A new family of life distributions*, J. Appl. Probab. 6 (1969), pp. 319–327.
[2] Z.W. Birnbaum and S.C. Saunders, *Estimation for a family of life distributions with applications to fatigue*, J. Appl. Probab. 6 (1969), pp. 328–377.
[3] D. Kundu, N. Kannan, and N. Balakrishnan, *On the function of Birnbaum–Saunders distribution and associated inference*, Comput. Statist. Data Anal. 52 (2008), pp. 2692–2702.
[4] J. Wu and A.C.M. Wong, *Improved interval estimation for the two-parameter Birnbaum–Saunders distribution*, Comput. Statist. Data Anal. 47 (2004), pp. 809–821.
[5] A.J. Lemonte, F. Cribari-Neto, and K.L.P. Vasconcellos, *Improved statistical inference for the two-parameter Birnbaum–Saunders distribution*, Comput. Statist. Data Anal. 51 (2007), pp. 4656–4681.

[6] A.J. Lemonte, A.B. Simas, and F. Cribari-Neto, *Bootstrap-based improved estimators for the two-parameter Birnbaum–Saunders distribution*, J. Stat. Comput. Simul. 78 (2008), pp. 37–49.
[7] J.R. Rieck, *A moment-generating function with application to the Birnbaum–Saunders distribution*, Comm. Statist. Theory Methods 28 (1999), pp. 2213–2222.
[8] J.B. McDonald, *Some generalized functions for the size distribution of income*, Econometrica 52 (1984), pp. 647–663.
[9] P. Kumaraswamy, *Generalized probability density function for double-bounded random processes*, J. Hydrol. 46 (1980), pp. 79–88.
[10] M.C. Jones, *Kumaraswamy's distribution: A beta-type distribution with some tractability advantages*, Stat. Methodol. 6 (2009), pp. 70–91.
[11] N. Eugene, C. Lee, and F. Famoye, *Beta-normal distribution and its applications*, Comm. Statist. Theory Methods 31 (2002), pp. 497–512.
[12] G.M. Cordeiro and M. de Castro, *A new family of generalized distributions*, J. Stat. Comput. Simul. 81 (2011), pp. 883–898.
[13] J.R. Rieck and J.R. Nedelman, *A log-linear model for the Birnbaum–Saunders distribution*, Technometrics 33 (1991), pp. 51–60.
[14] F.C. Xi and B.C. Wei, *Diagnostics analysis for log-Birnbaum–Saunders regression models*, Comput. Statist. Data Anal. 51 (2007), pp. 4692–4706.
[15] A.F. Desmond, G.A. Rodríguez-Yam, and X. Lu, *Estimation of parameters for a Birnbaum–Saunders regression model with censored data*, J. Stat. Comput. Simul. 78 (2008), pp. 983–997.
[16] A.J. Lemonte, S.L.P. Ferrari, and F. Cribari-Neto, *Improved likelihood inference in Birnbaum–Saunders regressions*, Comput. Statist. Data Anal. 54 (2010), pp. 1307–1316.
[17] Q. Xiao, Z. Liu, N. Balakrishnan, and X. Lu, *Estimation of the Birnbaum–Saunders regression model with current status data*, Comput. Statist. Data Anal. 54 (2010), pp. 326–332.
[18] V.G. Cancho, E.E.M. Ortega, and G.A. Paula, *On estimation and influence diagnostics for log-Birnbaum–Saunders Student-t regression models: Full Bayesian analysis*, J. Statist. Plann. Inference 140 (2010), pp. 2486–2496.
[19] A.J. Lemonte and S.L.P. Ferrari, *Size and power properties of some tests in the Birnbaum–Saunders regression model*, Comput. Statist. Data Anal. 55 (2011), pp. 1109–1117.
[20] A.J. Lemonte and S.L.P. Ferrari, *Small-sample corrections for score tests in Birnbaum–Saunders regressions*, Comm. Statist. Theory Methods 40 (2011), pp. 232–243.
[21] A.J. Lemonte and S.L.P. Ferrari, *Signed likelihood ratio tests in the Birnbaum–Saunders regression model*, J. Statist. Plann. Inference 141 (2011), pp. 1031–1040.
[22] A.J. Lemonte and G.M. Cordeiro, *Birnbaum–Saunders nonlinear regression models*, Comput. Statist. Data Anal. 53 (2009), pp. 4441–4452.
[23] G.M. Cordeiro and A.J. Lemonte, *The β-Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling*, Comput. Statist. Data Anal. 55 (2011), pp. 1445–1461.
[24] G.S. Mudholkar and A.D. Hutson, *The exponentiated Weibull family: Some properties and a flood data application*, Comm. Statist. Theory Methods 25 (1996), pp. 3059–3083.
[25] R.C. Gupta and D. Kundu, *Exponentiated exponential distribution: An alternative to gamma and Weibull distributions*, Biom. J. 43 (2001), pp. 117–130.
[26] S. Nadarajah and A.K. Gupta, *The exponentiated gamma distribution with application to drought data*, Calcutta Statist. Assoc. Bull. 59 (2007), pp. 29–54.
[27] A.J. Lemonte and G.M. Cordeiro, *The exponentiated generalized inverse Gaussian distribution*, Statist. Probab. Lett. 81 (2011), pp. 506–517.
[28] J.F. Kenney and E.S. Keeping, *Mathematics of Statistics, Part 1*, 3rd ed., Van Nostrand, Princeton, NJ, 1962, pp. 101–102.
[29] J.J.A. Moors, *A quantile alternative for kurtosis*, J. Roy. Statist. Soc. D 37 (1998), pp. 25–32.
[30] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 2007.
[31] J.A. Greenwood, J.M. Landwehr, N.C. Matalas, and J.R. Wallis, *Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form*, Water Resour. Res. 15 (1979), pp. 1049–1054.
[32] R. Terras, *A Miller algorithm for an incomplete Bessel function*, J. Comput. Phys. 39 (1981), pp. 233–240.
[33] D.S. Jones, *Incomplete Bessel functions*, Proc. Edinb. Math. Soc. 50 (2007), pp. 173–183.
[34] D.S. Jones, *Incomplete Bessel functions. Asymptotic expansions for large agument*, Proc. Edinb. Math. Soc. 50 (2007), pp. 711–723.
[35] F.E. Harris, *Incomplete Bessel, generalized incomplete gamma, or leaky aquifer functions*, J. Comput. Appl. Math. 215 (2008), pp. 260–269.
[36] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer, New York, 1999.
[37] W.H. Press, S.A. Teulosky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 3rd ed., Cambridge University Press, New York, 2007.
[38] J.A. Doornik, *An Object-Oriented Matrix Language – Ox 4*, 5th ed., Timberlake Consultants Press, London, 2006.
[39] M.D. Nichols and W.J. Padgett, *A Bootstrap control chart for Weibull percentiles*, Qual. Reliab. Eng. Int. 22 (2006), pp. 141–151.
[40] G. Chen and N. Balakrishnan, *A general purpose approximate goodness-of-fit test*, J. Qual. Technol. 27 (1995), pp. 154–161.
[41] D.N.P. Murthy, M. Xie, and R. Jiang, *Weibull Models*, John Wiley & Sons, New Jersey, 2004.

Appendix 1. Observed information matrix for the McBS model

The observed information matrix $\boldsymbol{J}_n(\boldsymbol{\theta})$ for the parameters a, η, c, α and β can be derived after extensive algebraic manipulations. It is given by

$$\boldsymbol{J}_n(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = -\begin{pmatrix} U_{aa} & U_{a\eta} & U_{ac} & U_{a\alpha} & U_{a\beta} \\ \cdot & U_{\eta\eta} & U_{\eta c} & U_{\eta\alpha} & U_{\eta\beta} \\ \cdot & \cdot & U_{cc} & U_{c\alpha} & U_{c\beta} \\ \cdot & \cdot & \cdot & U_{\alpha\alpha} & U_{\alpha\beta} \\ \cdot & \cdot & \cdot & \cdot & U_{\beta\beta} \end{pmatrix},$$

whose elements are

$$\begin{aligned}
U_{\alpha\alpha} &= \frac{n}{\alpha^2} + \frac{6n}{\alpha^4} - \frac{3}{\alpha^4}\sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i}\right) + \frac{2(a-1)}{\alpha^2}\sum_{i=1}^n \frac{v_i\phi(v_i)}{\Phi(v_i)} - \frac{2\eta c}{\alpha^2}\sum_{i=1}^n \frac{v_i\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} \\
&\quad - \frac{(a-1)}{\alpha^2}\sum_{i=1}^n \left\{\frac{v_i^3\phi(v_i)}{\Phi(v_i)} + \frac{v_i^2\phi(v_i)^2}{\Phi(v_i)^2}\right\} + \frac{\eta c(1-c)}{\alpha^2}\sum_{i=1}^n \frac{v_i^2\phi(v_i)^2\Phi(v_i)^{c-2}}{1-\Phi(v_i)^c} \\
&\quad + \frac{\eta c}{\alpha^2}\sum_{i=1}^n \frac{v_i^3\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} - \frac{\eta c^2}{\alpha^2}\sum_{i=1}^n \frac{v_i^2\phi(v_i)^2\Phi(v_i)^{2c-2}}{\{1-\Phi(v_i)^c\}^2}, \\
U_{\alpha\beta} &= -\frac{1}{\alpha^3\beta}\sum_{i=1}^n \left(\frac{t_i}{\beta} - \frac{\beta}{t_i}\right) + \frac{(a-1)}{2\alpha^2\beta}\sum_{i=1}^n \frac{\tau(\sqrt{t_i/\beta})\phi(v_i)}{\Phi(v_i)}\left\{1 - v_i^2 - \frac{v_i\phi(v_i)}{\Phi(v_i)}\right\} \\
&\quad + \frac{\eta c(1-c)}{2\alpha^2\beta}\sum_{i=1}^n \frac{v_i\tau(\sqrt{t_i/\beta})\phi(v_i)^2\Phi(v_i)^{c-2}}{1-\Phi(v_i)^c} - \frac{\eta c}{2\alpha^2\beta}\sum_{i=1}^n \frac{(1-v_i^2)\tau(\sqrt{t_i/\beta})\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} \\
&\quad - \frac{\eta c^2}{2\alpha^2\beta}\sum_{i=1}^n \frac{v_i\tau(\sqrt{t_i/\beta})\phi(v_i)^2\Phi(v_i)^{2c-2}}{\{1-\Phi(v_i)^c\}^2}, \\
U_{\beta\beta} &= \frac{n}{2\beta^2} - \sum_{i=1}^n \frac{1}{(t_i+\beta)^2} - \frac{1}{\alpha^2\beta^3}\sum_{i=1}^n t_i + \frac{(a-1)}{\alpha}\sum_{i=1}^n \frac{z_i\phi(v_i)}{\Phi(v_i)} \\
&\quad - \frac{(a-1)}{4\alpha^2\beta^2}\sum_{i=1}^n \frac{\tau(\sqrt{t_i/\beta})^2\phi(v_i)}{\Phi(v_i)}\left\{v_i + \frac{\phi(v_i)}{\Phi(v_i)}\right\} \\
&\quad + \frac{\eta c(1-c)}{4\alpha^2\beta^2}\sum_{i=1}^n \frac{\tau(\sqrt{t_i/\beta})^2\phi(v_i)^2\Phi(v_i)^{c-2}}{1-\Phi(v_i)^c} - \frac{\eta c}{\alpha}\sum_{i=1}^n \frac{z_i\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} \\
&\quad + \frac{\eta c}{4\alpha^2\beta^2}\sum_{i=1}^n \frac{v_i\tau(\sqrt{t_i/\beta})^2\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} - \frac{\eta c^2}{4\alpha^2\beta^2}\sum_{i=1}^n \frac{\tau(\sqrt{t_i/\beta})^2\phi(v_i)^2\Phi(v_i)^{2c-2}}{\{1-\Phi(v_i)^c\}^2}, \\
U_{cc} &= -\frac{n}{c^2} - \frac{2n\psi(ac^{-1})a}{c^3} - \frac{n\psi'(ac^{-1})a^2}{c^4} + \frac{2n\psi(ac^{-1}+\eta+1)a}{c^3} + \frac{n\psi'(ac^{-1}+\eta+1)a^2}{c^4} \\
&\quad - \eta\sum_{i=1}^n \frac{\Phi(v_i)^2[\log(\Phi(v_i))]^2}{1-\Phi(v_i)^c} - \eta\sum_{i=1}^n \frac{\Phi(v_i)^{2c}[\log(\Phi(v_i))]^2}{\{1-\Phi(v_i)^c\}^2}, \\
U_{c\alpha} &= \frac{\eta}{\alpha}\sum_{i=1}^n \frac{v_i\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} + \frac{\eta c}{\alpha}\sum_{i=1}^n \frac{v_i\phi(v_i)\Phi(v_i)^{c-1}\log(\Phi(v_i))}{\{1-\Phi(v_i)^c\}^2}, \\
U_{c\beta} &= \frac{\eta}{2\alpha\beta}\sum_{i=1}^n \frac{\tau(\sqrt{t_i/\beta})\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c} + \frac{\eta c}{2\alpha\beta}\sum_{i=1}^n \frac{\tau(\sqrt{t_i/\beta})\phi(v_i)\Phi(v_i)^{c-1}\log(\Phi(v_i))}{\{1-\Phi(v_i)^c\}^2}, \\
U_{\eta\eta} &= -n\psi'(\eta+1) + n\psi'(ac^{-1}+\eta+1), \\
U_{\eta c} &= -\frac{n\psi'(ac^{-1}+\eta+1)a}{c^2} - \sum_{i=1}^n \frac{\Phi(v_i)^c\log(\Phi(v_i))}{1-\Phi(v_i)^c},
\end{aligned}$$

$$U_{\eta\alpha} = \frac{c}{\alpha}\sum_{i=1}^{n}\frac{v_i\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c}, \quad U_{\eta\beta} = \frac{c}{2\alpha\beta}\sum_{i=1}^{n}\frac{\tau(\sqrt{t_i/\beta})\phi(v_i)\Phi(v_i)^{c-1}}{1-\Phi(v_i)^c},$$

$$U_{aa} = -\frac{n\psi'(ac^{-1})}{c^2} + \frac{n\psi'(ac^{-1}+\eta+1)}{c^2}, \quad U_{a\eta} = \frac{n\psi'(ac^{-1}+\eta+1)}{c},$$

$$U_{ac} = \frac{n\psi(ac^{-1})}{c^2} + \frac{n\psi'(ac^{-1})a}{c^3} - \frac{n\psi(ac^{-1}+\eta+1)}{c^2} - \frac{n\psi'(ac^{-1}+\eta+1)a}{c^3},$$

$$U_{a\alpha} = -\frac{1}{\alpha}\sum_{i=1}^{n}\frac{v_i\phi(v_i)}{\Phi(v_i)}, \quad U_{a\beta} = -\frac{1}{2\alpha\beta}\sum_{i=1}^{n}\frac{\tau(\sqrt{t_i/\beta})\phi(v_i)}{\Phi(v_i)},$$

where $\psi'(\cdot)$ is the trigamma function and $z_i = \{3(t_i/\beta)^{1/2} + (\beta/t_i)^{1/2}\}/(4\beta^2)$, for $i = 1, \ldots, n$.