Brazilian Journal of Probability and Statistics
2020, Vol. 34, No. 3, 495–518
https://doi.org/10.1214/19-BJPS448

On a bimodal Birnbaum–Saunders distribution with applications to lifetime data

Roberto Vila[a], Jeremias Leão[b], Helton Saulo[a], Mirza Naveed Shahzad[c] and Manoel Santos-Neto[d,e]

[a]*University of Brasília*
[b]*Federal University of Amazonas*
[c]*University of Gujrat*
[d]*Federal University of São Carlos*
[e]*Federal University of Campina Grande*

Abstract. The Birnbaum–Saunders distribution is a flexible and useful model which has been used in several fields. In this paper, a new bimodal version of this distribution based on the alpha-skew-normal distribution is established. We discuss some of its mathematical and inferential properties. We consider likelihood-based methods to estimate the model parameters. We carry out a Monte Carlo simulation study to evaluate the performance of the maximum likelihood estimators. For illustrative purposes, three real data sets are analyzed. The results indicated that the proposed model outperformed some existing models in the literature, in special, a recent bimodal extension of the Birnbaum–Saunders distribution.

1 Introduction

Despite its broad applicability in many fields, see, for example, Balakrishnan, Leiva and López (2007), Bhatti (2010), Vilca et al. (2010), Paula et al. (2012), Saulo et al. (2010), Leiva et al. (2014a, 2014b), Leiva (2016) and Leao et al. (2017), the Birnbaum–Saunders (BS) distribution (Birnbaum and Saunders, 1969) is not suitable to model bimodal data. This distribution is positively skewed with positive support and is related to the normal distribution through the stochastic representation

$$T = \frac{\beta}{4}\Big[\alpha Z + \sqrt{(\alpha Z)^2 + 4}\Big]^2, \tag{1.1}$$

where $T \sim \mathrm{BS}(\alpha, \beta)$, $Z \sim \mathrm{N}(0, 1)$ and $\alpha > 0$, $\beta > 0$ are shape and scale parameters, respectively. The $\mathrm{BS}(\alpha, \beta)$ probability density function (PDF) and cumulative distribution function (CDF) are respectively, given by

$$f(t; \alpha, \beta) = \phi\big(a(t)\big)\frac{t^{-3/2}(t + \beta)}{2\alpha\beta^{1/2}} \quad \text{and} \quad F(t; \alpha, \beta) = \Phi\big(a(t)\big), \quad t > 0, \tag{1.2}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are standard normal PDF and CDF, respectively, and

$$a(t) = \frac{1}{\alpha}\left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right]. \tag{1.3}$$

Key words and phrases. Birnbaum–Saunders distribution, alpha-skew-normal distribution, bimodality, maximum likelihood estimation, Monte Carlo simulation.
Received February 2017; accepted April 2019.

Note that the k-th derivative of $a(t)$, denoted by $a^{(k)}(t)$, satisfies $a^{(k)}(t) > 0$ (or < 0) for k odd (or k even), where $k \geq 1$. Some special cases of these derivatives are

$$a'(t) = \frac{1}{2\alpha t}\left[\sqrt{\frac{t}{\beta}} + \sqrt{\frac{\beta}{t}}\right],$$
$$a''(t) = -\frac{1}{4\alpha t^2}\left[\sqrt{\frac{t}{\beta}} + 3\sqrt{\frac{\beta}{t}}\right] \quad \text{and} \tag{1.4}$$
$$a'''(t) = \frac{3}{8\alpha t^3}\left[\sqrt{\frac{t}{\beta}} + 5\sqrt{\frac{\beta}{t}}\right].$$

Note also that the function $a(\cdot)$ has inverse specified by (1.1). In order not to cause confusion, hereafter we will write $a^{-\perp}(\cdot)$ to denote the inverse of function $a(\cdot)$.

The stochastic representation in (1.1) allows us to obtain several generalizations of the BS model. For example, Díaz-García and Leiva (2005) assumed that Z follows a standard symmetric distribution in the real line and obtained the class of generalized BS distributions. On the same line, Balakrishnan et al. (2009) proposed scale-mixture BS distributions by assuming that Z belongs to the family of scale mixture of normal distributions. Many other generalizations can be obtained in order to obtain a new distribution with domain on the positive numbers; see Leiva (2016).

In general, one uses mixtures of distributions for describing bimodal data. However, it may be troublesome as identifiability problems may arise in the parameter estimation of the model; see Lin, Lee and Hsieh (2007), Lin, Lee and Yen (2007) and Gómez et al. (2011). In this sense, new mixture-free models which have the capacity to accommodate unimodal and bimodal data are very important. Some asymmetric bimodal models in the real line have been discussed by Azzalini and Capitanio (2003), Kim (2005), and Ma and Genton (2004), among others. In the context of bimodal BS models, Balakrishnan et al. (2011) introduced a mixture distribution of two different BS models (MXBS) and studied its characteristics. On the other hand, Olmos, Martínez-Flórez and Bolfarine (2017) introduced a bimodal extension of the BS distribution, denoted by BBSO, based on the approach described in Gómez et al. (2011). In addition, the authors also studied the probabilistic properties and moments of the BBSO distribution, and showed that this model can fit well both unimodal and bimodal data in comparison with the BS, log-normal and skew-normal BS models. A thorough inference study on the parameters that index the BBSO distribution was addressed by Fonseca and Cribari-Neto (2018).

In this paper, we introduce a new bimodal version of the BS distribution, denoted by BBS, by assuming that Z in (1.1) follows an alpha-skew-normal (ANS) distribution discussed by Elal-Olivero (2010). We present a statistical methodology based on the proposed BBS distribution including model formulation, mathematical properties, estimation and inference based on the maximum likelihood (ML) method. We evaluate the performance of the ML estimators by Monte Carlo (MC) simulations. Three real data illustrations indicated that the proposed BBS model provides better adjustment compared to the BBSO model proposed by Olmos, Martínez-Flórez and Bolfarine (2017). The proposed BBS distribution has some advantages over existing bimodal BS models: (i) unlike the BBSO model, the proposed BBS distribution does not suffer from convergence problems in the optimization process of the profile log-likelihood function as pointed out by Fonseca and Cribari-Neto (2018); (ii) the proposed BBS distribution does not present identifiability problems commonly encountered in mixture models, such as the MXBS distribution; and (iii) the proposed model does not present label switching problems (Celeux et al., 2006), that is, in a bimodal context with two groups, during the estimation an individual who was in group B can incorrectly stay in A and vice versa.

The rest of the paper proceeds as follows. In Section 2, we introduce the BBS distribution and discuss some related results. In Section 3, we consider likelihood-based methods to estimate the model parameters and to perform inference. In Section 4, we carry out a MC simulation study to evaluate the performance of the ML estimators. In Section 5, we illustrate the proposed methodology with three real data sets. Finally, in Section 6, we make some concluding remarks and discuss future research.

2 The BBS distribution

If a random variable (RV) X has an ASN distribution with parameter δ, denoted by $X \sim \text{ASN}(\delta)$, then its PDF and CDF are given by

$$g(x) = \frac{(1-\delta x)^2 + 1}{2+\delta^2}\phi(x) \quad \text{and} \quad G(x) = \Phi(x) + \delta\left(\frac{2-\delta x}{2+\delta^2}\right)\phi(x), \tag{2.1}$$

where $x, \delta \in \mathbb{R}$ and δ is an asymmetric parameter that controls the uni-bimodality effect; see Elal-Olivero (2010). The PDF of the BS distribution, based on the alpha-skew-normal model, is given by

$$f(t;\alpha,\beta,\delta) = \frac{(1-\delta a(t))^2 + 1}{2+\delta^2}\phi\big(a(t)\big)\frac{t^{-3/2}(t+\beta)}{2\alpha\beta^{1/2}}, \quad t > 0, \tag{2.2}$$

where $a(\cdot)$ is as in (1.3) and the notation $T \sim \text{BBS}(\alpha,\beta,\delta)$ is used. If $\delta = 0$, then the classical $\text{BS}(\alpha,\beta)$ distribution is obtained. The corresponding $\text{BBS}(\alpha,\beta,\delta)$ CDF is given by

$$F(t;\alpha,\beta,\delta) = \Phi\big(a(t)\big) + \delta\left(\frac{2-\delta a(t)}{2+\delta^2}\right)\phi\big(a(t)\big), \quad t > 0. \tag{2.3}$$

Note that $f(t;\alpha,\beta,\delta) = g(a(t))a'(t) = (G \circ a)'(t)$, $F(t;\alpha,\beta,\delta) = (G \circ a)(t)$ and $\lim_{\delta\to\pm\infty}\{F(t;\alpha,\beta,\delta) + \phi(a(t))a(t)\} = F(t;\alpha,\beta)$.

Differentiating the PDF of the BBS distribution (2.2), we obtain

$$f'(t;\alpha,\beta,\delta) = g'\big(a(t)\big)\big[a'(t)\big]^2 + g\big(a(t)\big)a''(t) \quad \text{and} \tag{2.4}$$

$$f''(t;\alpha,\beta,\delta) = g''\big(a(t)\big)\big[a'(t)\big]^3 + 3g'\big(a(t)\big)a'(t)a''(t) + g\big(a(t)\big)a'''(t), \tag{2.5}$$

where $g''(x) = -xg'(x) + \phi(x)(-3\delta^2x^2 + 4\delta x - 2(1-\delta^2))/(2+\delta^2)$ and

$$g'(x) = \frac{\phi(x)}{2+\delta^2}\big(-\delta^2x^3 + 2\delta x^2 - 2\big(1-\delta^2\big)x - 2\delta\big).$$

The survival and hazard functions, denoted by SF and HR, respectively, of the BBS distribution are given by $S(t;\alpha,\beta,\delta) = 1 - (G \circ a)(t)$ and

$$h(t;\alpha,\beta,\delta) = \frac{f(t;\alpha,\beta,\delta)}{1-F(t;\alpha,\beta,\delta)} = \frac{(G \circ a)'(t)}{S(t;\alpha,\beta,\delta)}, \quad t > 0,$$

respectively. From Figure 1, we note some different shapes of the BBS PDF for different combinations of parameters. These figures reveal clearly the bimodality effect caused by the parameter δ. Also, Figure 2 shows unimodal and bimodal shapes for the BBS HR.

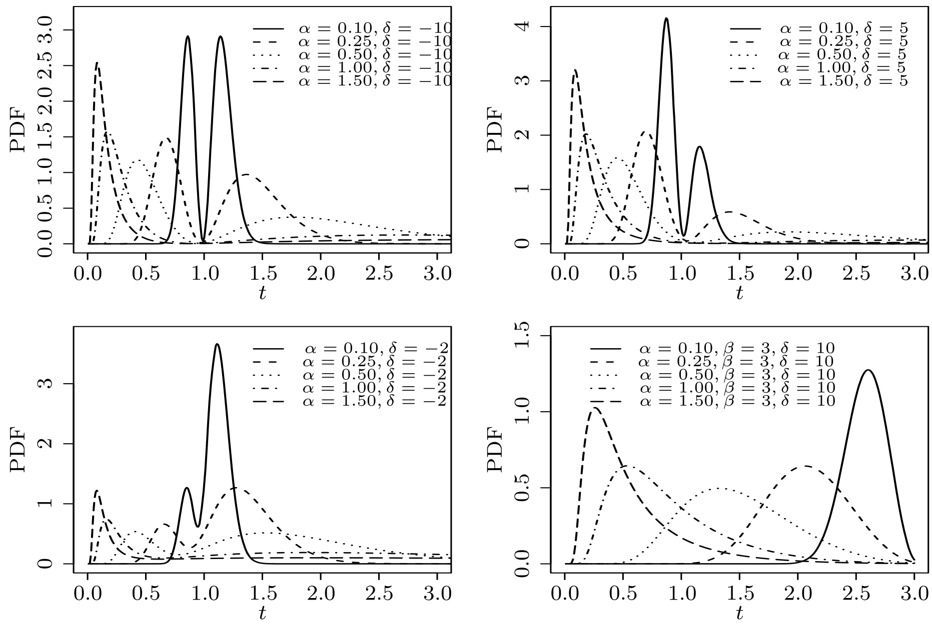

Figure 1 *BBS PDFs for some parameter values* ($\beta = 1.0$).

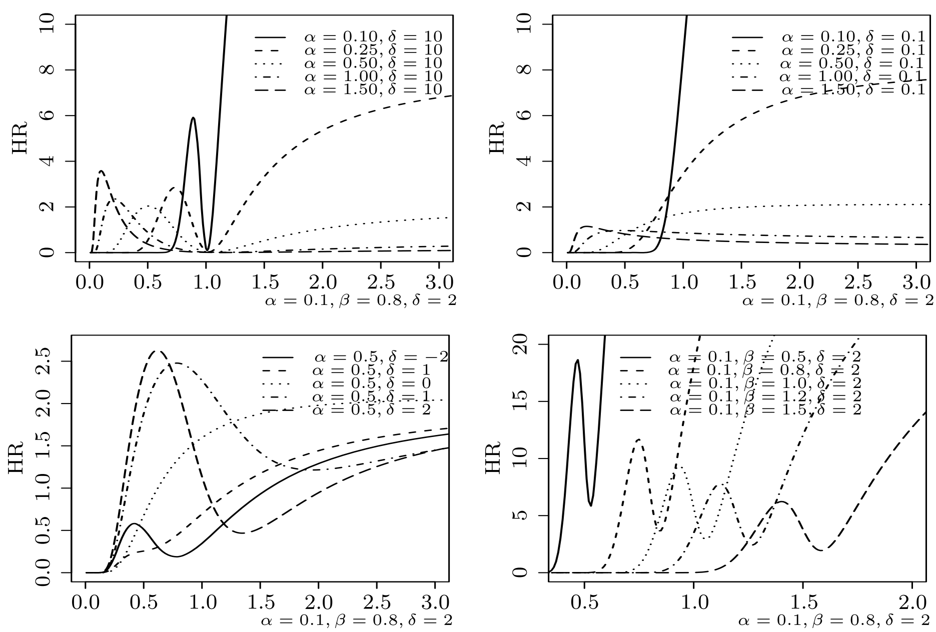

Figure 2 *BBS HRs for some parameter values* ($\beta = 1.0$).

2.1 Some properties of the BBS distribution

Lemma 1. *Let* $t_0 = \beta[(\alpha/\delta) + \sqrt{(\alpha/\delta)^2 + 4}]^2/4$. *The PDF of the BBS distribution* $t \mapsto f(t; \alpha, \beta, \delta)$ *defined in* (2.2) *is a decreasing function when*

1. $\delta = 0$ *and* $t > \beta$, *or*
2. $t < t_0$ ($t > t_0$), *for each* $\delta > 0$ ($\delta < 0$).

Proof. 1. If $\delta = 0$, we have $f(t;\alpha,\beta,\delta) = f(t;\alpha,\beta) = \phi(a(t))a'(t)$, $t > 0$. Since $a'(t) > 0$ and $a''(t) < 0$ (see (1.4)) we have that $t \mapsto \phi(a(t))$ and $t \mapsto a'(t)$ are decreasing functions whenever $t > \beta$. Therefore, since the PDF of the BBS distribution is a product of nonnegative decreasing functions, it is decreasing for each $t > \beta$.

2. Let $r(x) = ((1-\delta x)^2 + 1)/(2+\delta^2)$, $x \in \mathbb{R}$. Note that $(r \circ a)'(t) = -2\delta(1 - \delta a(t))a'(t)/(2+\delta^2)$, $t > 0$. If $\delta > 0$ ($\delta < 0$), then $(r \circ a)'(t) < 0$ whenever $t < t_0$ ($t > t_0$). Since, by Item 1, the function $t \mapsto f(t;\alpha,\beta)$ is decreasing and $f(t;\alpha,\beta,\delta) = r(a(t)) f(t;\alpha,\beta)$, $t > 0$, the proof follows. □

Proposition 2.1. *Let $T \sim \mathrm{BBS}(\alpha,\beta,\delta)$ as defined in (2.1). Then,*

1. $a(T) \sim ASN(\delta)$;
2. $cT \sim \mathrm{BBS}(\alpha, c\beta, \delta)$, *with* $c > 0$;
3. $T^{-1} \sim \mathrm{BBS}(\alpha, \beta^{-1}, -\delta)$.

Proof. Since $\mathbb{P}(a(T) \le t) = F(a^{-\perp}(t);\alpha,\beta,\delta)$, we have that the PDF of $a(T)$ is equal to $f(a^{-\perp}(t);\alpha,\beta,\delta)/a'(a^{-\perp}(t)) = g(t)$. Then $a(T) \sim \mathrm{ASN}(\delta)$. The proof of the Items 2 and 3 are immediate, after making convenient variables transformations. □

Proposition 2.2. *Let $T \sim \mathrm{BBS}(\alpha,\beta,\delta)$ and $X \sim ASN(\delta)$ and suppose $\mathbb{E}[T^n]$ exists, $n \ge 1$. Then, we have*

1. $\mathbb{E}[T] = \frac{\beta}{2}(-\alpha\frac{2\delta}{2+\delta^2} + \omega_{0,1})$;
2. $\mathbb{E}[T^2] = (\frac{\beta}{2})^2(4 + 2\alpha^2\frac{2+3\delta^2}{2+\delta^2} + 2\alpha\omega_{1,1})$;
3. $\mathbb{E}[T^3] = (\frac{\beta}{2})^3(-24\alpha(\alpha^2+1)\frac{\delta}{2+\delta^2} + 3\alpha^2\omega_{2,1} + \alpha\omega_{1,1} + \omega_{0,3})$;
4. $\mathbb{E}[T^4] = (\frac{\beta}{2})^4(16 + 8\alpha^2\frac{8+12\delta^2+6\alpha^2+15\delta^2\alpha^2}{2+\delta^2} + 4\alpha(\alpha^2\omega_{3,1} + \omega_{1,3}))$ *and*
5. $\mathrm{Var}[T] = (\frac{\beta}{2})^2(2\alpha^2\frac{4+6\delta^2+3\delta^4}{(2+\delta^2)^2} + \omega_{0,1}^2 + 2\alpha(\omega_{1,1} + \omega_{0,1}\frac{2\delta}{2+\delta^2}))$,

where $\omega_{r,k} = \mathbb{E}[X^r(\sqrt{\alpha^2X^2+4})^k]$.

Proof. By Proposition 2.1 Item 1, we have $a(T) \sim \mathrm{ASN}(\delta)$ which implies that $\mathbb{E}[T^n] = \mathbb{E}[\{a^{-\perp}(X)\}^n]$. Then the proof is immediate since

$$\mathbb{E}[\{a^{-\perp}(X)\}^n] = \left(\frac{\beta}{2}\right)^n \mathbb{E}[(\alpha X + \sqrt{(\alpha X)^2+4})^n], \quad n \ge 1 \tag{2.6}$$

and $\mathbb{E}[X] = -2\delta/(2+\delta^2)$, $\mathbb{E}[X^2] = 1 - \delta\mathbb{E}[X]$, $\mathbb{E}[X^3] = 3\mathbb{E}[X]$, and $\mathbb{E}[X^4] = 3(1 - 2\delta\mathbb{E}[X])$. □

Remark 1. By using the Binomial Theorem and (2.6) note that $\mathbb{E}[T^n]$ exists iff $\omega_{r,n-r}$ (defined in Proposition 2.2) exists, with $r = 0,\dots,n$. By Jensen's inequality (see, e.g., Chung (2001)), we obtain

$$|\omega_{0,1}| \le \sqrt{\alpha^2\mathbb{E}[X^2]+4} = \sqrt{\alpha^2\left(\frac{2+3\delta^2}{2+\delta^2}\right)+4} < +\infty,$$

and by Minkowski inequality (see, e.g., Natanson (1955)), we have

$$|\omega_{1,1}| \le \sqrt{\mathbb{E}[X^2]}\sqrt{\alpha^2\mathbb{E}[X^2]+4} = \sqrt{\frac{2+3\delta^2}{2+\delta^2}}\sqrt{\alpha^2\left(\frac{2+3\delta^2}{2+\delta^2}\right)+4} < +\infty.$$

Then, the expected value $\mathbb{E}[T]$ and variance $\mathrm{Var}[T]$ always exist. Note also that higher order moments can also be easily obtained from the expression of $\mathbb{E}[T^n]$.

Proposition 2.3. *Let $X \sim ASN(\delta)$ and $T = a^{-\perp}(X)$. Then, $T \sim \mathrm{BBS}(\alpha, \beta, \delta)$.*

Proof. Since $\mathbb{P}(a^{-\perp}(X) \le t) = G(a(t))$, we have that the PDF of the RV $a^{-\perp}(X)$ is equal to $g(a(t))a'(t) = f(t; \alpha, \beta, \delta)$. □

Proposition 2.4. *Let $T \sim \chi_3^2$, where χ_3^2 denotes the chi-squared distribution with 3 degrees of freedom. We have the following relation*

$$f_{a^{-\perp}(\sqrt{T})}(t)\big[(1/a(t) - \delta)^2 + (1/a(t))^2\big] = 2(2 + \delta^2) f(t; \alpha, \beta, \delta), \quad t > 0,$$

where $f_{a^{-\perp}(\sqrt{T})}(\cdot)$ denotes the PDF of the RV $a^{-\perp}(\sqrt{T})$.

Proof. Since $\mathbb{P}(a^{-\perp}(\sqrt{T}) \le t) = \mathbb{P}(T \le a^2(t))$, we have that $f_{a^{-\perp}(\sqrt{T})}(t) = 2a^2(t) \times \phi(a(t))a'(t)$, from where the proof follows. □

2.2 Some properties of the HR of the BBS distribution

Let $s(t) = -f'(t; \alpha, \beta, \delta)/f(t; \alpha, \beta, \delta)$, where $f(\cdot; \alpha, \beta, \delta)$ denotes the PDF of the BBS distribution (2.2). It is straightforward to show that

$$s(t) = \frac{a'(t)m(t)}{(1 - \delta a(t))^2 + 1},$$

where

$$m(t) = \delta^2 a^3(t) - 2\delta a^2(t) - 2a(t)(\delta^2 - 1) + 2\delta - \big((1 - \delta a(t))^2 + 1\big)\frac{a''(t)}{[a'(t)]^2}.$$

Remark 2. Using the identities in (1.4), note that $m(t) = 0$ iff

$$\begin{aligned}
&\frac{1}{\alpha^2\beta} t\{\delta^2 a^3(t) - 2\delta a^2(t) - 2(\delta^2 - 1)a(t) + 2\delta\} \\
&\quad + \frac{1}{4} a(t)\{\delta^2 a^4(t) - 2\delta a^3(t) - (5\delta^2 - 2)a^2(t) + 8\delta a(t) - 6\} \\
&\quad + \frac{1}{\alpha\sqrt{\beta}}\sqrt{t}\{-\delta^2 a^4(t) + 2\delta a^3(t) + (3\delta^2 - 2)a^2(t) - 4\delta a(t) + 2\} = 0.
\end{aligned}$$

Consider also the function spaces

$$B = \left\{\ell : \mathbb{R}^+ \to \mathbb{R}^+ \text{ differentiable} : \begin{array}{c} \ell'(t) < 0 \text{ for } t \in (0, t_0),\ \ell'(t_0) = 0, \\ \ell'(t) > 0 \text{ for } t > t_0 \end{array}\right\},$$

$$U = \left\{\ell : \mathbb{R}^+ \to \mathbb{R}^+ \text{ differentiable} : \begin{array}{c} \ell'(t) > 0 \text{ for } t \in (0, t_0),\ \ell'(t_0) = 0, \\ \ell'(t) < 0 \text{ for } t > t_0 \end{array}\right\}.$$

Each function $\ell \in B$ or $\ell \in U$ is said bathtub shaped or upside down bathtub shaped, respectively.

The following results due to Glaser (1980) helps us to characterize the shape of the failure rates, through the function $s(\cdot)$.

1. If $t \mapsto s(t)$ is increasing, then the HR is increasing in t.
2. If $t \mapsto s(t)$ is decreasing, then the HR is decreasing in t.
3. If $t \mapsto s(t) \in B$ and if there exist a t^* such that $h'(t^*; \alpha, \beta, \delta) = 0$, then the HR belongs to B, otherwise the HR is increasing in t.

4. If $t \mapsto s(t) \in U$ and if there exist a t^* such that $h'(t^*; \alpha, \beta, \delta) = 0$, then the HR belongs to U, otherwise the HR is decreasing in t.

Using the expressions of the derivatives of $a(\cdot)$ in (1.4), we can get the monotonicity of the HR of the BBS distribution from the following equation

$$s'(t) = \frac{m'(t)a'(t) + m(t)a''(t) + 2\delta(1 - \delta a(t))a'(t)s(t)}{(1 - \delta a(t))^2 + 1},$$

where

$$\frac{m'(t)}{a'(t)} = 3\delta^2 a^2(t) - 4\delta a(t) - 2(\delta^2 - 1) + 2\delta(1 - \delta a(t))\frac{a''(t)}{[a'(t)]^2} - \big((1 - \delta a(t))^2 + 1\big)\frac{a'''(t)a'(t) - 2[a''(t)]^2}{[a'(t)]^4}.$$

For example, if $\delta = 0$ and $\alpha > 2$, we have that $m(t) > 0$ iff $t > \beta$. Defining the set $\mathscr{L}_\beta^\alpha = \{t : t^4 + (4 - \alpha^2)\beta t^3 + 6(1 - \alpha^2)\beta^2 t^2 + (4 + 3\alpha^2)\beta^3 t + \beta^4 < 0\}$ note that $m'(t) = a'(t)(2 - 2(a'''(t)a'(t) - 2[a''(t)]^2)/[a'(t)]^4) < 0$ on $\mathscr{L}_\beta^\alpha$. That is, $s'(t) = (m'(t)a'(t) + m(t)a''(t))/2 < 0$ on $\{t \in \mathscr{L}_\beta^\alpha : t > \beta\}$. Therefore, by Item 2 above, the HR $t \mapsto h(t; \alpha, \beta, \delta = 0)$ is decreasing on $\{t \in \mathscr{L}_\beta^\alpha : t > \beta\}$. On the other hand, if $\delta = 0$ and $\alpha < 1$, $m(t) < 0$ iff $t < \beta$. In this case, note that $m'(t) > 0$ on $[\mathscr{L}_\beta^\alpha]^c = \mathbb{R}^+$, hence $s'(t) > 0$ for each $t < \beta$. Then, using the Item 1 above, the HR is increasing for each $t < \beta$.

Another easy case to study is when $\delta = 1$. In this case, $m(t) > 0$ iff $t > \beta$. Note also that $m'(t) < 0$ on the set $\mathscr{L}_{\alpha,\beta} = \{t : 3a^2(t) - 4a(t) + 2(1 - a(t))a''(t)/[a'(t)]^2 < 0\}$. Then $s'(t) < 0$ on $\{t \in \mathscr{L}_{\alpha,\beta} : t > t_1\}$ where $t_1 = \beta[\alpha + \sqrt{\alpha^2 + 4}]^2/4$. Therefore, by Item 2 above, the HR $t \mapsto h(t; \alpha, \beta, \delta = 1)$ is decreasing on $\{t \in \mathscr{L}_{\alpha,\beta} : t > t_1\}$. Similar analyzes can be done for the other possible cases.

We emphasize that $h'(t; \alpha, \beta, \delta) = 0$ iff the PDF of the BBS distribution is a decreasing function. But, by Lemma 1 this happens when $\delta = 0$ and $t > \beta$ or $t < t_0$ $(t > t_0)$, for each $\delta > 0$ $(\delta < 0)$ with $t_0 = \beta[(\alpha/\delta) + \sqrt{(\alpha/\delta)^2 + 4}]^2/4$. So, to see if the HR belongs (or not) to B or to U it would be sufficient to verify that $t \mapsto s(t)$ belongs (or not) to B or to U.

2.3 Bimodality properties

In this subsection, some results on the bimodality properties of BBS distribution are obtained.

Proposition 2.5. *A mode of the* BBS(α, β, δ) *is any point* $t_0 = t_0(\alpha, \beta, \delta)$ *that satisfies*

$$t_0 = -\frac{\alpha^2\beta}{p_3(t)}\left[\frac{1}{4}a(t)p_4(t) + \frac{1}{\alpha\sqrt{\beta}}\sqrt{t}\,\widetilde{p}_4(t)\right],$$

where $p_3(t) = \delta^2 a^3(t) - 2\delta a^2(t) - 2(\delta^2 - 1)a(t) + 2\delta$, $p_4(t) = \delta^2 a^4(t) - 2\delta a^3(t) - (5\delta^2 - 2)a^2(t) + 8\delta a(t) - 6$ *and* $\widetilde{p}_4(t) = -\delta^2 a^4(t) + 2\delta a^3(t) + (3\delta^2 - 2)a^2(t) - 4\delta a(t) + 2$.

Proof. A mode of the BBS(α, β, δ) is any point t that satisfies $f'(t; \alpha, \beta, \delta) = 0$. But this happens iff $s(t) = 0$ which is equivalent to $m(t) = 0$, where $s(t)$ and $m(t)$ were defined in Section 2.2. Then, using Remark 2 and solving for t gives the result. □

Proposition 2.6. *The function* $t \mapsto (g \circ a)(t)$ *and the PDF of the BBS distribution* (2.2) *have different modes.*

Proof. We will do the proof by contradiction. Let's suppose that t_0 is a mode for both $(g \circ a)(\cdot)$ (which always exists, since g is bimodal) and $f(\cdot;\alpha,\beta,\delta)$. Then $g'(a(t_0))a'(t_0)=0$ and $g''(a(t_0))[a'(t_0)]^2<0$.

Since $f'(t_0;\alpha,\beta,\delta)=0$ and $f''(t_0;\alpha,\beta,\delta)<0$, using (2.4) and (2.5) we obtain that $g(a(t_0))=0$, which is impossible. Then, the proof follows. □

Remark 3. As a consequence of the proof of the Proposition 2.6, we have that, if t_0 is a maximum point of $t \mapsto (g \circ a)(t)$ then the maximum points of the BBS distribution must be to the left side of t_0. On the other hand, if t_1 is a minimum point of $t \mapsto (g \circ a)(t)$ then, the minimum points of the BBS distribution must be to the right side of t_1.

Proposition 2.7. *The PDF of the BBS distribution* (2.2) *has at most one mode when* $\delta = 0$.

Proof. If $\delta=0$, then the classical BS(α,β) distribution is obtained, that is, $f(t;\alpha,\beta,\delta)=f(t;\alpha,\beta)=\phi(a(t))a'(t)$, $t>0$. Differentiating $f(t;\alpha,\beta)$, we obtain

$$f'(t;\alpha,\beta)=\phi\big(a(t)\big)\big(a''(t)-a(t)\big[a'(t)\big]^2\big).$$

Using (1.4), it is straightforward to show that $f'(t;\alpha,\beta)=0$ iff

$$t^3+\beta(1+\alpha^2)t^2-\beta^2 t-\beta^3=0. \tag{2.7}$$

The discriminant of a cubic polynomial ax^3+bx^2+cx+d is given by $\Delta_3=b^2c^2-4ac^3-4b^3d-27a^2d^2+18abcd$. In our case, we have

$$\Delta_3=\beta^6\big(4(1+\alpha^2)^3+(1+\alpha^2)^2+18(1+\alpha^2)-23\big).$$

Note that $\Delta_3>0$ for each $\alpha>0$, then the equation (2.7) has three distinct real roots.

Let t_1, t_2 and t_3 be the three distinct real roots of (2.7), by Vieta's formula (see, e.g., Vinberg (2003)), it is valid that

$$t_1+t_2+t_3=-\beta(1+\alpha^2),$$
$$t_1t_2+t_1t_3+t_2t_3=-\beta^2,$$
$$t_1t_2t_3=\beta^3.$$

From the first and third equations above we conclude that there must be two negative and one positive roots, hence $f(t;\alpha,\beta)$ has at most one mode. □

Proposition 2.8. *If* $\delta=-\alpha$, *then one of the modes of the BBS distribution* (2.2) *occurs at* $t=\beta$.

Proof. Since $a(\beta)=0$, $a'(\beta)=1/\alpha\beta$, $a''(\beta)=-1/\alpha\beta^2$, $g(0)=2/\sqrt{2\pi}(2+\delta^2)$ and $g'(0)=-g(0)\delta$, by (2.4) we have that

$$f'(\beta;\alpha,\beta,\delta)=-\frac{2}{\alpha\beta^2(2+\delta^2)\sqrt{2\pi}}\left(\frac{\delta}{\alpha}+1\right)=0 \quad \text{since } \delta=-\alpha. \tag{2.8}$$

That is, $t=\beta$ is one of the critical points of f when $\delta=-\alpha$.

Since $a'''(\beta)=9/4\alpha\beta^3$ and $g''(0)=-g'(0)(1+\delta^2)$, using (2.5), note that

$$f''(\beta;\alpha,\beta,\delta)=\frac{2}{\alpha\beta^3(2+\delta^2\sqrt{2\pi})}\left(\delta(1+\delta^2)\frac{1}{\alpha^2}+3\frac{\delta}{\alpha}-\frac{9}{4}\right).$$

As $\delta=-\alpha$, we obtain

$$f''(\beta;\alpha,\beta,\delta)=-\frac{2}{\alpha\beta^3(2+\alpha^2)\sqrt{2\pi}}\left(\frac{1+\alpha^2}{\alpha}+\frac{21}{4}\right)<0.$$

Therefore, the PDF of the BBS distribution is concave downward when $\delta=-\alpha$. □

Example 2.1 (Bimodality). Consider $\alpha = \beta = 1$ and $\delta = -\alpha$. By Proposition 2.8, the point $t = 1$ is one of the modes of $f(\cdot; \alpha, \beta, \delta)$. Using (2.4) note that $f'(t; \alpha, \beta, \delta) = 0$ iff

$$p(y) = y^{10} + 2y^9 + y^6 - 4y^5 + 3y^4 - 8y^3 + 4y^2 + 2y - 1 = 0,$$

where $y = t^{1/2}$.

We have that $p(0) = -1 < 0$,

$$p(1/2) = \frac{85}{1024} > 0,$$
$$p(3/4) = -\frac{252{,}223}{1{,}048{,}576} < 0 \quad \text{and}$$
$$p(5/4) \approx 15.27127 > 0.$$

Therefore, $p(y)$ has roots in the intervals (0, 1/2), (1/2, 3/4) and (3/4, 5/4). It is not hard to show that $p(y) > 0$ for $y > 1$. Thus, all real roots of the polynomial $p(y)$ lie in the interval $(0, 5/4)$. Computationally it can be verified that $y_0 \approx 0.419703$, $y_1 \approx 0.646914$ and $y_2 = 1$ are the only roots of $p(y)$ on $\{y : y > 0\}$. Hence, $t_0 = y_0^2 \approx 0.1761$, $t_1 = y_1^2 \approx 0.4184$ and $t_2 = y_2^2 = 1$ are the only roots of $f'(t; \alpha, \beta, \delta) = 0$. It can be verified that

$a^{(k)}(t)$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$t = t_0$	-1.9633	7.963	-61.0996	848.6550
$t = t_1$	-0.8991	2.6204	-7.5471	42.8876

and

$g^{(k)}(t)$	$k = 0$	$k = 1$	$k = 2$
$t = t_0$	$1.9219\phi(t_0)/3$	$1.8585\phi(t_0)/3$	$-0.0616\phi(t_0)/3$
$t = t_1$	$1.0101\phi(t_1)/3$	$1.1100\phi(t_1)/3$	$2.1692\phi(t_1)/3$

where $a^{(0)} \equiv a$ and $g^{(0)} \equiv g$. Using (2.5) and the quantities above, we obtain

$$\begin{aligned} f''(t_0; \alpha, \beta, \delta) &= g''(a(t_0))[a'(t_0)]^3 + 3g'(a(t_0))a'(t_0)a''(t_0) + g(a(t_0))a'''(t_0) \\ &\approx -1107.6637\frac{\phi(t_0)}{3} < 0, \end{aligned}$$

and similarly $f''(t_1; \alpha, \beta, \delta) \approx 60.3992\phi(t_1)/3 > 0$.

Therefore, the PDF of the BBS distribution, with parameters $\alpha = \beta = 1$ and $\delta = -\alpha$, has exactly two modes at $t = t_0$ and $t = t_2$.

Remark 4. Let $\alpha = \beta = 1$ and $\delta = -\alpha$. It can be verified that the point $t_{\max} \approx a^{-\perp}(0.83929) = 2.26240$ is the only maximum point of the function $(g \circ a)(\cdot)$. The Remark 3 assures us that the maximum points of the PDF $f(\cdot; \alpha, \beta, \delta)$ must be to the left-hand side of $t_{\max}$. This statement was verified in the previous example.

2.4 Shannon entropy

For a continuous PDF $f(t)$ on an interval I, its entropy is defined as

$$H(f) = -\int_I f(t) \log f(t)\, \mathrm{d}t.$$

This definition of entropy, introduced by Shannon and Weaver (1949), resembles a formula for a thermodynamic notion of entropy. In our probabilistic context, if X is an absolutely continuous RV with PDF $f_X(t)$, the quantity $H(X) = H(f_X) = -\mathbb{E}[\log f_X(X)]$ is viewed as a measure of uncertainty associated with a RV. Note that $H(X)$ is not necessarily well-defined, since the integral does not always exist.

Consider $T \sim \text{BBS}(\alpha, \beta, \delta)$. The Shannon entropy of T satisfies the following identity.

Proposition 2.9. *If $T \sim \text{BBS}(\alpha, \beta, \delta)$, there exists a constant $C(\alpha, \beta, \delta)$ such that*

$$H(T) = C(\alpha, \beta, \delta) + \mathbb{E}\left[\log\left\{\frac{T^{3/2}/(T+\beta)}{(1-\delta a(T))^2 + 1}\right\}\right]. \tag{2.9}$$

Proof. It is straightforward to verify that

$$\begin{aligned} H(T) = {} & \log(2+\delta^2) + \log(2\alpha\beta^{1/2}) + \log(\sqrt{2\pi}) + \frac{1}{2}\mathbb{E}[a^2(T)] \\ & + \int_0^\infty \log\left\{\frac{t^{3/2}/(t+\beta)}{(1-\delta a(t))^2+1}\right\} f(t; \alpha, \beta, \delta)\,\mathrm{d}t. \end{aligned}$$

Since $a(T) \sim \text{ASN}(\delta)$, by Proposition 2.1 we have $\mathbb{E}[a^2(T)] = 1 + 2\delta^2/(2+\delta^2)$. Therefore, the identity (2.9) is verified considering $C(\alpha, \beta, \delta) = \log(2+\delta^2) + \log(2\alpha\beta^{1/2}) + \log(\sqrt{2\pi}) + (1 + 2\delta^2/(2+\delta^2))/2$. □

Remark 5. If $T \sim \text{BBS}(\alpha, \beta, \delta)$ and $T \geq 1$, then the Shannon entropy always exists. In fact, by Jensen's inequality (see, e.g., Chung (2001)), Minkosky inequality (see, e.g., Natanson (1955)) and Remark 1 we obtain

$$\begin{aligned} &|\mathbb{E}[\log(T^{3/2})]| \leq \log \mathbb{E}[T^{3/2}] \leq \log(\mathbb{E}[T^2])^{1/2} + \log(\mathbb{E}[T])^{1/2} < +\infty, \\ &|\mathbb{E}[\log(T+\beta)]| \leq \log(\mathbb{E}[T] + \beta) < +\infty \quad \text{and} \\ &|\mathbb{E}[\log((1-\delta a(T)^2 + 1)]| \leq \log(2 + \delta^2\mathbb{E}[a^2(T)] - 2\delta\mathbb{E}[a(T)]) < +\infty, \end{aligned}$$

because $a(T) \sim \text{ASN}(\delta)$. Then, using (2.9) and the above inequalities, the proof follows.

3 Estimation and inference

3.1 Maximum likelihood estimation

Let $(t_1, \ldots, t_n)$ be a random sample of size n from the BBS distribution with PDF in (2.2). Considering δ known, it follows that the log-likelihood function, without the constant, is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}) = {} & -n\log(\alpha) - \frac{n}{2}\log(\beta) + \sum_{i=1}^n \log\big(1 + \big(1 - \delta a(t_i)\big)^2\big) \\ & - \frac{1}{2}\sum_{i=1}^n a^2(t_i) + \sum_{i=1}^n \log(t_i + \beta), \end{aligned}$$

where $\boldsymbol{\theta} = (\alpha, \beta)$. Since

$$\frac{\partial}{\partial\alpha}a(t) = -\frac{t^{-1/2}}{\alpha^2\beta^{1/2}}(t-\beta) \quad \text{and} \quad \frac{\partial}{\partial\beta}a(t) = -\frac{t^{-1/2}}{2\alpha}\big(\beta^{-3/2}t + 2\big), \tag{3.1}$$

taking the first derivatives with respect to α and β and equating them to zero, we have

$$\frac{\partial}{\partial\alpha}\ell(\boldsymbol{\theta}) = -\frac{n}{\alpha} - \sum_{i=1}^{n}\left(\frac{2\delta(1-\delta a(t_i))}{1+(1-\delta a(t_i))^2} + a(t_i)\right)\frac{\partial}{\partial\alpha}a(t_i) = 0 \quad \text{and}$$

$$\frac{\partial}{\partial\beta}\ell(\boldsymbol{\theta}) = -\frac{n}{2\beta} - \sum_{i=1}^{n}\left(\frac{2\delta(1-\delta a(t_i))}{1+(1-\delta a(t_i))^2} + a(t_i)\right)\frac{\partial}{\partial\beta}a(t_i) + \sum_{i=1}^{n}\frac{1}{t_i+\beta} = 0. \tag{3.2}$$

The ML estimates $\widehat{\alpha}$ and $\widehat{\beta}$ of α and β, respectively, are obtained by solving an iterative procedure for non-linear optimization of the system of equations in (3.2), such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton method; see Mittelhammer, Judge and Miller (2000). We implement the BFGS algorithm in the `R` software, available at http://cran.r-project.org, by the function `optim`.

The ML estimator $\widehat{\boldsymbol{\theta}}$, under some standard regularity conditions (see Section 3.2), is consistent and follows a normal joint asymptotic distribution with mean $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$. Furthermore, $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$ can be obtained from the corresponding expected Fisher information matrix, $\mathcal{I}(\boldsymbol{\theta})$ say. Thus, we have

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} \mathrm{N}_2\big(\mathbf{0}_{(2)\times 1}, \boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}) = \mathcal{J}(\boldsymbol{\theta})^{-1}\big), \quad \text{as } n \to \infty,$$

where $\xrightarrow{D}$ denotes convergence in distribution, $\mathbf{0}_{(2)\times 1}$ is a $(2) \times 1$ vector of zeros and $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}\frac{1}{n}\mathcal{I}(\boldsymbol{\theta})$. Here, we approximate the expected Fisher information matrix by its observed version, and the square root of each diagonal element of its inverse matrix is used to approximate the associated standard error (SE); see Efron and Hinkley (1978).

We can use the profile log-likelihood for finding the value of δ. In fact, this parameter is assumed to be fixed in the log-likelihood function, because some difficulties in calculating it by the ML method were reported. Generally, two steps are required to estimate δ:

(i) Let $\delta_i = i$ and for each $i = -20, \ldots, 0, \ldots, 20$ compute the ML estimates of α and β by solving the system of equations in (3.2);

(ii) Select the final estimate of δ as the one which maximizes the log-likelihood function and also select the associated estimates of α and β as final ones.

Case of random censoring. Suppose that the time to the event of interest is not completely observed and it may be subject to right censoring. Let c_i denote the censoring time and t_i the time to the event of interest. We observe $y_i = \min\{t_i, c_i\}$, whereas $\tau_i = I(t_i \leq c_i)$ is such that $\tau_i = 1$ if y_i is the time to the event of interest and $\tau_i = 0$ if it is right censored, for $i = 1, \ldots, n$. Let $\boldsymbol{\theta} = (\alpha, \beta)$ denote the parameter vector of the BBS model given in (2.2) with δ known. From n pairs of times and censoring indicators $(t_1, \tau_1), \ldots, (t_n, \tau_n)$, the corresponding likelihood function obtained under uninformative censoring can be expressed as

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} f(t_i; \alpha, \beta, \delta)^{\tau_i}\big(1 - F(t_i; \alpha, \beta, \delta)\big)^{1-\tau_i} \\ &= \prod_{i=1}^{n}\left\{\frac{1+(1-\delta a(t_i))^2}{2+\delta^2}\phi\big(a(t_i)\big)\frac{t_i^{-3/2}(t_i+\beta)}{2\alpha\beta^{1/2}}\right\}^{\tau_i} \\ &\quad \times \left[1 - \Phi\big(a(t_i)\big) + \delta\left(\frac{2-\delta t_i}{2+\alpha^2}\right)\phi\big(a(t_i)\big)\right]^{1-\tau_i}. \end{aligned} \tag{3.3}$$

Therefore, the log-likelihood function for the BBS model obtained from (3.3) is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}) = &-\omega_i \eta(\boldsymbol{\theta}) + \sum_{i=1}^{n} \tau_i \log\big(1 + \big(1 - \delta a(t_i)\big)^2\big) + \sum_{i=1}^{n} \tau_i \log\big(\phi\big(a(t_i)\big)\big) \\ &- \frac{3}{2}\sum_{i=1}^{n} \tau_i \log(t_i) + \sum_{i=1}^{n} \tau_i \log(t_i + \beta) \\ &+ \sum_{i=1}^{n}(1 - \tau_i) \log\bigg[1 - \Phi\big(a(t_i)\big) + \delta\Big(\frac{2 - \delta t_i}{2 + \alpha^2}\Big)\phi\big(a(t_i)\big)\bigg], \end{aligned} \tag{3.4}$$

where $\omega_i = \sum_{i=1}^{n} \tau_i$ and $\eta(\boldsymbol{\theta}) = \log(2\alpha\beta^{1/2}(2 + \delta^2))$. The parameter vector $\boldsymbol{\theta}$ may be estimated using an iterative procedure for non-linear optimization (BFGS method) of the log-likelihood function (3.4). The estimation of δ can be performed using the profile log-likelihood as mentioned earlier in Section 3.1.

3.2 Confidence intervals

In this subsection, we present confidence intervals (CIs) for $S(t; \alpha, \beta, \delta)$, $\mathbb{E}[T]$ and $\mathrm{Var}[T]$, where $T \sim \mathrm{BBS}(\alpha, \beta, \delta)$ and δ is known.

Let $\{T_n, n \geq 1\}$ be a sequence of RVs. We will say that $\{T_n\}$ is asymptotically normal (AN) with mean μ_n and variance σ_n^2, and write $T_n \sim \mathrm{AN}(\mu_n, \sigma_n^2)$, if $\sigma_n > 0$ and as $n \to \infty$,

$$\frac{T_n - \mu_n}{\sigma_n} \longrightarrow N(0, 1).$$

Here μ_n is not necessarily the mean of T_n and σ_n^2, not necessarily its variance. This is, for sufficiently large n, for each $t \in \mathbb{R}$ we can approximate the probability $\mathbb{P}(T_n \leq t)$ by $\mathbb{P}(Z \leq ((t - \mu_n)/\sigma_n))$ where Z is $N(0, 1)$.

Let $\boldsymbol{\theta} = (\alpha, \beta)^\top$ in $\boldsymbol{\Theta}$ and $\rho \in (0, 1)$. The random interval $(\underline{\theta}(T_1, \ldots, T_n), \overline{\theta}(T_1, \ldots, T_n))$ will be called a CI at confidence level $1 - \rho$ for the parameter $\boldsymbol{\theta}$, provided that

$$\mathbb{P}\big(\underline{\theta}(T_1, \ldots, T_n) < \boldsymbol{\theta} < \overline{\theta}(T_1, \ldots, T_n)\big) \geq 1 - \rho.$$

In what follows, we assume $\ell(\boldsymbol{\theta})$ holds the following standard regularity conditions:

1. The parameter space, defined by $\boldsymbol{\Theta}$, is open and $\ell(\boldsymbol{\theta})$ has a global maximum at $\boldsymbol{\Theta}$;
2. For almost all t, the fourth-order log-likelihood derivatives with respect to the model parameters exist and are continuous in an open subset of $\boldsymbol{\Theta}$ that contains the true parameter $\boldsymbol{\theta}$;
3. The support set of $t \mapsto f(t; \boldsymbol{\theta}, \delta)$, for $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}$, does not depend on $\boldsymbol{\theta}$;
4. The expected information matrix $\mathcal{I}(\boldsymbol{\theta})$ is positive definite and finite. We remember that the information matrix $\mathcal{I}(\boldsymbol{\theta})$ is a 2×2 matrix with elements $\mathcal{I}_{j,k}(\boldsymbol{\theta})$ $j, k = 1, 2$, defined by

$$\mathcal{I}_{j,k}(\boldsymbol{\theta}) = \mathrm{Cov}\bigg(\frac{\partial}{\partial\theta_j} \log f(T; \boldsymbol{\theta}, \delta), \frac{\partial}{\partial\theta_k} \log f(T; \boldsymbol{\theta}, \delta)\bigg), \quad \theta_j, \theta_k \in \{\alpha, \beta\}.$$

These regularity conditions are not restrictive and hold for the models cited in this work. Let

$$\mathcal{V}(t; \boldsymbol{\theta}, \delta) = \frac{2\delta[1 + (1 - \delta a(t))^2] - 4\delta^2(1 - \delta a(t))^2}{[1 + (1 - \delta a(t))^2]^2} - 1,$$

$$\mathcal{W}(t; \boldsymbol{\theta}, \delta) = \frac{2\delta(1 - \delta a(t))}{1 + (1 - \delta a(t))^2} + a(t).$$

The Fisher information matrix may also be written as

$$\mathcal{I}_{j,k}(\boldsymbol{\theta}) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta_j\partial\theta_k}\log f(T;\boldsymbol{\theta},\delta)\right), \quad \theta_j,\theta_k \in \{\alpha,\beta\},$$

where

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha^2}\log f(t;\boldsymbol{\theta},\delta) &= \frac{1}{\alpha^2} + \mathcal{V}(t;\boldsymbol{\theta},\delta)\left(\frac{\partial}{\partial\alpha}a(t)\right)^2 - \mathcal{W}(t;\boldsymbol{\theta},\delta)\frac{\partial^2}{\partial\alpha^2}a(t),\\
\frac{\partial^2}{\partial\beta^2}\log f(t;\boldsymbol{\theta},\delta) &= \frac{1}{4\beta^2} + \mathcal{V}(t;\boldsymbol{\theta},\delta)\left(\frac{\partial}{\partial\beta}a(t)\right)^2\\
&\qquad - \mathcal{W}(t;\boldsymbol{\theta},\delta)\frac{\partial^2}{\partial\beta^2}a(t) - \frac{1}{(t+\beta)^2},\\
\frac{\partial^2}{\partial\alpha\partial\beta}\log f(t;\boldsymbol{\theta},\delta) &= \frac{\partial^2}{\partial\beta\partial\alpha}\log f(t;\boldsymbol{\theta},\delta)\\
&= \mathcal{V}(t;\boldsymbol{\theta},\delta)\frac{\partial}{\partial\alpha}a(t)\frac{\partial}{\partial\beta}a(t) - \mathcal{W}(t;\boldsymbol{\theta},\delta)\frac{\partial^2}{\partial\alpha\partial\beta}a(t).
\end{aligned}$$

The above first-order partial derivatives of $a(\cdot)$ with respect to α and β were calculated in (3.1) and the respective second-order partial derivatives are given by

$$\begin{aligned}
&\frac{\partial^2}{\partial\alpha^2}a(t) = \frac{2t^{-1/2}}{\alpha^3\beta^{1/2}}(t-\beta), \qquad \frac{\partial^2}{\partial\beta^2}a(t) = \frac{3t^{-1/2}\beta^{-5/2}}{4\alpha},\\
&\frac{\partial^2}{\partial\alpha\partial\beta}a(t) = \frac{\partial^2}{\partial\beta\partial\alpha}a(t) = \frac{t^{-1/2}}{2\alpha^2}(\beta^{-3/2}+2).
\end{aligned}$$

Here, the mixed partial differentiations are commutative at a given point $\boldsymbol{\theta}$ in $\mathbb{R}^2$ because the corresponding functions have continuous second partial derivatives at that point (Schwarz's theorem).

3.2.1 *Confidence interval for* $S(t;\boldsymbol{\theta},\delta)$. Let $\widehat{\alpha}$ and $\widehat{\beta}$ be ML estimates of α and β, respectively. It is known that the ML estimate of $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha},\widehat{\beta})^\top$ has normal asymptotic distribution, with null mean vector and asymptotic covariance matrix given by the inverse of the information matrix $I(\boldsymbol{\theta})$. That is,

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim \mathrm{AN}\big(\mathbf{0}, [I(\boldsymbol{\theta})]^{-1}\big).$$

Since the function $\boldsymbol{\theta} \mapsto S(t;\boldsymbol{\theta},\delta)$, $\forall t > 0$, is continuously differentiable, by the Delta method we have

$$\sqrt{n}\big(S(t;\widehat{\boldsymbol{\theta}},\delta) - S(t;\boldsymbol{\theta},\delta)\big) \sim \mathrm{AN}\big(\mathbf{0}, J_S(\boldsymbol{\theta})[I(\boldsymbol{\theta})]^{-1}J_S(\boldsymbol{\theta})^\top\big),$$

where $J_S(\boldsymbol{\theta}) = J_S(\boldsymbol{\theta};t) = [\frac{\partial}{\partial\alpha}S(t;\boldsymbol{\theta},\delta) \quad \frac{\partial}{\partial\beta}S(t;\boldsymbol{\theta},\delta)]_{1\times 2}$ is the Jacobian of the function $\boldsymbol{\theta} \mapsto S(t;\boldsymbol{\theta},\delta)$.

As $\widehat{\boldsymbol{\theta}}$ is a ML estimate of $\boldsymbol{\theta}$, the asymptotic variance of $S(t;\widehat{\boldsymbol{\theta}},\delta)$ can be estimated by

$$\mathrm{Var}\big[S(t;\widehat{\boldsymbol{\theta}},\delta)\big] \approx J_S(\widehat{\boldsymbol{\theta}})\big[I(\widehat{\boldsymbol{\theta}})\big]^{-1}J_S(\widehat{\boldsymbol{\theta}})^\top.$$

As $\widehat{\boldsymbol{\theta}}$ is consistent (because it is a ML estimate), by Slutsky's theorem we have

$$\sqrt{n}\big(S(t;\widehat{\boldsymbol{\theta}},\delta) - S(t;\boldsymbol{\theta},\delta)\big) \sim \mathrm{AN}\big(\mathbf{0}, \mathrm{Var}\big[S(t;\widehat{\boldsymbol{\theta}},\delta)\big]\big). \tag{3.5}$$

If $0 < \rho < 1$, using (3.5), the CI at confidence level $1-\rho$ for $S(t;\boldsymbol{\theta},\delta)$ is obtained from the following identity:

$$\lim_{n\to\infty}\mathbb{P}\bigg(\frac{|S(t;\widehat{\boldsymbol{\theta}},\delta)-S(t;\boldsymbol{\theta},\delta)|}{\widehat{\sigma}(t)}<\frac{z_{\rho/2}}{\sqrt{n}}\bigg)=\mathbb{P}\big(|Z|<z_{\rho/2}\big)\geq 1-\rho,$$

where $z_{\rho/2}$ is the $\rho/2$-quantile of the normal distribution and $\widehat{\sigma}^2(t)=\mathrm{Var}[S(t;\widehat{\boldsymbol{\theta}},\delta)]$. Then, the random interval

$$\bigg(S(t;\widehat{\boldsymbol{\theta}},\delta)-\frac{z_{\rho/2}}{\sqrt{n}}\widehat{\sigma}(t),\,S(t;\widehat{\boldsymbol{\theta}},\delta)+\frac{z_{\rho/2}}{\sqrt{n}}\widehat{\sigma}(t)\bigg) \tag{3.6}$$

is a CI at confidence level $1-\rho$ for $S(t;\boldsymbol{\theta},\delta)$, $\forall t>0$.

3.2.2 *Confidence interval for* $\mathbb{E}[T|\boldsymbol{\theta}]=\mathbb{E}[T]$. Since $T\sim\mathrm{BBS}(\boldsymbol{\theta},\delta)$ is a positive RV, we have the identity $\mathbb{E}[T|\boldsymbol{\theta}]=\int_0^\infty S(t;\boldsymbol{\theta},\delta)\,\mathrm{d}t$. Using this identity and denoting $\widehat{\sigma}^2(t)=\mathrm{Var}[S(t;\widehat{\boldsymbol{\theta}},\delta)]$ note that (3.6) implies that the set

$$\bigg\{\mathbb{E}[T|\widehat{\boldsymbol{\theta}}]-\frac{z_{\rho/2}}{\sqrt{n}}\int_0^\infty\widehat{\sigma}(t)\,\mathrm{d}t<\mathbb{E}[T|\boldsymbol{\theta}]<\mathbb{E}[T|\widehat{\boldsymbol{\theta}}]+\frac{z_{\rho/2}}{\sqrt{n}}\int_0^\infty\widehat{\sigma}(t)\,\mathrm{d}t\bigg\}$$

contains the set

$$\bigg\{S(t;\widehat{\boldsymbol{\theta}},\delta)-\frac{z_{\rho/2}}{\sqrt{n}}\widehat{\sigma}(t)<S(t;\boldsymbol{\theta},\delta)<S(t;\widehat{\boldsymbol{\theta}},\delta)+\frac{z_{\rho/2}}{\sqrt{n}}\widehat{\sigma}(t)\bigg\}.$$

Therefore, the random interval

$$\bigg(\mathbb{E}[T|\widehat{\boldsymbol{\theta}}]-\frac{z_{\rho/2}}{\sqrt{n}}\int_0^\infty\widehat{\sigma}(t)\,\mathrm{d}t,\,\mathbb{E}[T|\widehat{\boldsymbol{\theta}}]+\frac{z_{\rho/2}}{\sqrt{n}}\int_0^\infty\widehat{\sigma}(t)\,\mathrm{d}t\bigg)$$

provides us a CI at confidence level $1-\rho$ for $\mathbb{E}[T|\boldsymbol{\theta}]$. If the lower limit of the CI is negative, we will replace it with zero.

3.2.3 *Confidence interval for* $\mathrm{Var}[T|\boldsymbol{\theta}]=\mathrm{Var}[T]$. Let $\widehat{L}_\pm(t)=S(t;\widehat{\boldsymbol{\theta}},\delta)\pm z_{\rho/2}\widehat{\sigma}(t)/\sqrt{n}$ where $\widehat{\sigma}^2(t)=\mathrm{Var}[S(t;\widehat{\boldsymbol{\theta}},\delta)]$, $t>0$. Assume that $\widehat{L}_-(t)>0$, otherwise we replace this lower limit with zero.

Let $A=\{\widehat{L}_-(t)<S(t;\boldsymbol{\theta},\delta)<\widehat{L}_+(t)\}$ and

$$B=\bigg\{\int_0^\infty\widehat{L}_-(t)\,\mathrm{d}t<\mathbb{E}[T|\boldsymbol{\theta}]<\int_0^\infty\widehat{L}_+(t)\,\mathrm{d}t\bigg\}.$$

Using the identity $\mathbb{E}[T^2|\boldsymbol{\theta}]=2\int_0^\infty tS(t;\boldsymbol{\theta},\delta)\,\mathrm{d}t$, let's denote also $C=\{2\int_0^\infty t\times\widehat{L}_-(t)\,\mathrm{d}t<\mathbb{E}[T^2|\boldsymbol{\theta}]<2\int_0^\infty t\widehat{L}_+(t)\,\mathrm{d}t\}$ and

$$D=\bigg\{-\bigg(\int_0^\infty\widehat{L}_+(t)\,\mathrm{d}t\bigg)^2<-\big(\mathbb{E}[T|\boldsymbol{\theta}]\big)^2<-\bigg(\int_0^\infty\widehat{L}_-(t)\,\mathrm{d}t\bigg)^2\bigg\}.$$

Note that $A\subseteq B,C,D$ and $B\cap D=B$. Hence, if $(\widehat{L}_-(t),\widehat{L}_+(t))$ is a random CI for $S(t;\boldsymbol{\theta},\delta)$ with confidence coefficient $1-\rho$ (by Section 3.2.1), for each $t>0$, then $(\int_0^\infty\widehat{L}_-(t)\,\mathrm{d}t,\int_0^\infty\widehat{L}_+(t)\,\mathrm{d}t)$ and $(2\int_0^\infty t\widehat{L}_-(t)\,\mathrm{d}t,2\int_0^\infty t\widehat{L}_+(t)\,\mathrm{d}t)$ are also (random) CIs for $\mathbb{E}[T|\boldsymbol{\theta}]$ and $\mathbb{E}[T^2|\boldsymbol{\theta}]$ respectively, with confidence coefficient $1-\rho$ each.

Since

$$\mathrm{I}_{\mathrm{Var}}=\big\{2J(\widehat{L}_-,\widehat{L}_+)<\mathrm{Var}(T|\boldsymbol{\theta})<2J(\widehat{L}_+,\widehat{L}_-)\big\}\supseteq B\cap C\cap D=B\cap C,$$

where J denotes the operator $J(f,g)=\int_0^\infty tf(t)\,\mathrm{d}t-(\int_0^\infty g(t)\,\mathrm{d}t)^2$, we have

$$\mathbb{P}(\mathrm{I}_{\mathrm{Var}})\geq\mathbb{P}(B\cap C)\geq\mathbb{P}(B)+\mathbb{P}(C)-1\geq 1-2\rho.$$

Therefore, $(2J(\widehat{L}_-,\widehat{L}_+),2J(\widehat{L}_+,\widehat{L}_-))$ is a (random) CI for $\mathrm{Var}(T|\boldsymbol{\theta})$ with confidence coefficient $1-2\rho$. Again, if the lower limit of the CI is negative, we will replace it with zero.

Remark 6. Analogously to that done in Section 3.2.1, we can construct a CI for the function $\log(-\log(S(t; \alpha, \beta, \delta)))$.

4 Monte Carlo simulation

Two MC simulation studies were carried out to evaluate the performance of the ML estimators of the proposed BBS model. The first study considers simulated data generated from the BBS distribution, whereas the second one has as its data generating process the BS, log-normal (LN) and MXBS distributions. All numerical evaluations were done in the `R` software; see R-Team (2018). The used `R` codes are available upon request.

4.1 Simulation study 1

In this first study, we evaluate the performance of the ML estimators for the proposed BBS model, considering the simulated data generated from the same model. The simulation scenario assumes the sample sizes $n \in \{10, 50\}$, the values of the shape parameter as $\alpha \in \{0.10, 0.50, 1.00, 1.50\}$, the values of the asymmetric parameter as $\delta \in \{-10, -5, -1, 1, 5, 10\}$, and 10,000 MC replications. The censoring proportion is $p \in \{0.0, 0.1, 0.3\}$; see Section 3.1. Note that the values of the shape parameter α have been chosen in order to study the performance under low, moderate and high skewness.

For each value of the parameter δ, sample size and censoring proportion, the empirical values for the bias (Bias) and mean squared error (MSE) of the ML estimators are reported in Tables 1–2. From these tables, note that, as the sample size increases, the ML estimators become more efficient, as expected. We can also note that, as the censoring proportion increases, the performances of the estimators of α and β, deteriorate. It is interesting to note two points on the increasing of the bias of $\widehat{\beta}$: (i) when the skewness increases, the bias of $\widehat{\beta}$ increases, which is expected as the original distribution occurs in the BS, see, for example, Lemonte, Simas and Cribari-Neto (2008); and (ii) note that there seems to be an increase in the bias of $\widehat{\beta}$ when we decrease the values of the parameter δ, see the cases $\delta = \{-1, 1\}$. In general, all of these results show the good performance of the proposed model.

4.2 Simulation study 2

In this second simulation study, we consider the BS, LN and MXBS distributions as data generating processes, and the BBS and BBSO distributions are fitted to the simulated data. Note that the BBS and BBSO models are the closest competitors, since both models do not require mixture of distributions to produce bimodality. The purpose is to evaluate how the estimators behave when the data generating process is wrong (the assumed model is different from the data generating model). In addition, we also compare the adjustments of the BBS and BBSO models by means of the fitted log-likelihood (log-lik) values. The $\mathrm{BS}(\alpha, \beta)$, $\mathrm{LN}(\mu, \sigma)$ and $\mathrm{MXBS}(\alpha_1, \beta_1, \alpha_2, \beta_2, p)$ samples were generated by considering the following PDFs $f_{\mathrm{BS}}(t; \alpha, \beta) = \phi(a(t))[t^{-3/2}(t + \beta)]/[2\alpha\beta^{1/2}]$, $f_{\mathrm{LN}}(t; \mu, \sigma) = 1/[t\sigma\sqrt{2\pi}]\exp([\log(t) - \mu]^2/[2\sigma^2])$ and $f_{\mathrm{MXBS}}(t; \alpha_1, \beta_1, \alpha_2, \beta_2, p) = p f_{\mathrm{BS}}(t; \alpha_1, \beta_1) + [1 - p] f_{\mathrm{BS}}(t; \alpha_2, \beta_2)$, $t > 0$, where $\phi(\cdot)$ and $a(\cdot)$ are as in (1.2) and (1.3). Moreover, the BBSO PDF is given by $f_{\mathrm{BBSO}}(t; \alpha, \beta, \gamma) = [t^{-3/2}(t + \beta)]/[4\alpha\beta^{1/2}\Phi(-\gamma)]\phi(|a(t)| + \gamma)$, where $\phi(\cdot)$, $\Phi(\cdot)$ and $a(\cdot)$ are as in (1.2) and (1.3).

The simulation scenario considers: sample sizes $n \in \{10, 50\}$, the values of the shape parameters as $\alpha, \sigma \in \{0.10, 1.00, 1.50, 2.50, 4.00\}$, the values of the mixing parameter as $p \in \{0.25, 0.50, 0.75\}$, and 1,000 MC replications. In this case, we do not consider censoring as in Simulation 1. The values of the shape parameters α, σ cover different levels of skewness. Note that the $\mathrm{BS}(\alpha, \beta)$ and $\mathrm{LN}(\mu, \sigma)$ PDFs are unimodal, whereas the

Table 1 *Simulated values of biases (MSEs within parentheses) of the estimators of the BBS model*

Censoring %	n	δ	BBS($\alpha=0.1, \beta=1.0, \delta$) Bias($\widehat{\alpha}$)	BBS($\alpha=0.1, \beta=1.0, \delta$) Bias($\widehat{\beta}$)	BBS($\alpha=0.5, \beta=1.0, \delta$) Bias($\widehat{\alpha}$)	BBS($\alpha=0.5, \beta=1.0, \delta$) Bias($\widehat{\beta}$)
0%	10	−10	−0.0011 (0.0002)	−0.0010 (0.0009)	−0.0066 (0.0062)	0.0028 (0.0184)
		−5	−0.0021 (0.0003)	−0.0034 (0.0013)	−0.0117 (0.0075)	−0.0042 (0.0224)
		−1	−0.0240 (0.0012)	−0.0418 (0.0054)	−0.1224 (0.0308)	−0.1581 (0.0824)
		1	−0.0235 (0.0012)	0.0480 (0.0066)	−0.1192 (0.0303)	0.2824 (0.2153)
		5	−0.0022 (0.0003)	0.0050 (0.0015)	−0.0111 (0.0078)	0.0293 (0.0376)
		10	−0.0012 (0.0002)	0.0020 (0.0010)	−0.0056 (0.0063)	0.0169 (0.0244)
	50	−10	−0.0001 (<0.0001)	−0.0001 (0.0001)	−0.0012 (0.0004)	0.0004 (0.0026)
		−5	−0.0003 (<0.0001)	0.0002 (0.0001)	−0.0025 (0.0010)	0.0027 (0.0028)
		−1	−0.0024 (0.0002)	−0.0182 (0.0018)	−0.0139 (0.0040)	−0.0597 (0.0254)
		1	−0.0024 (0.0002)	0.0205 (0.0022)	−0.0133 (0.0040)	0.0927 (0.0475)
		5	−0.0003 (<0.0001)	0.0001 (0.0001)	−0.0019 (0.0010)	0.0019 (0.0028)
		10	−0.0001 (<0.0001)	0.0003 (0.0001)	−0.0010 (0.0009)	0.0034 (0.0026)
10%	10	−10	0.0039 (0.0009)	−0.0241 (0.0081)	−0.0084 (0.0083)	0.0393 (0.0392)
		−5	0.0070 (0.0013)	−0.0466 (0.0133)	−0.0184 (0.0112)	0.0276 (0.0575)
		−1	−0.0006 (0.0011)	−0.1308 (0.0183)	0.0023 (0.0301)	−0.4827 (0.2416)
		1	0.0012 (0.0011)	0.1533 (0.0258)	0.0431 (0.0381)	1.0269 (1.1883)
		5	−0.0029 (0.0004)	0.0002 (0.0043)	−0.0233 (0.0095)	0.0376 (0.1120)
		10	−0.0007 (0.0004)	−0.0047 (0.0033)	−0.0059 (0.0079)	0.0281 (0.1520)
	50	−10	0.0019 (0.0002)	0.0109 (0.0022)	0.0053 (0.0011)	0.0659 (0.0074)
		−5	0.0070 (0.0007)	−0.0027 (0.0062)	0.0095 (0.0015)	0.1094 (0.0170)
		−1	0.0091 (0.0003)	−0.1387 (0.0195)	0.0391 (0.0083)	−0.5053 (0.2571)
		1	0.0116 (0.0004)	0.1632 (0.0271)	0.0904 (0.0171)	1.0871 (1.2133)
		5	0.0015 (0.0003)	−0.0118 (0.0031)	0.0001 (0.0011)	−0.0680 (0.0078)
		10	0.0006 (0.0001)	−0.0094 (0.0008)	0.0003 (0.0011)	−0.0423 (0.0040)
30%	10	−10	0.0875 (0.0123)	−0.2085 (0.0677)	0.0195 (0.0232)	0.0854 (0.1273)
		−5	0.0811 (0.0106)	−0.2206 (0.0638)	0.0035 (0.0255)	0.0741 (0.1459)
		−1	0.0031 (0.0013)	−0.1311 (0.0187)	−0.0680 (0.0286)	−0.4564 (0.2194)
		1	0.0004 (0.0012)	0.1497 (0.0250)	0.0521 (0.0550)	1.0430 (1.4937)
		5	0.0410 (0.0062)	−0.0509 (0.0390)	0.0283 (0.0343)	0.1593 (0.4103)
		10	0.0679 (0.0102)	−0.1303 (0.0538)	0.0307 (0.0290)	0.0776 (0.2059)
	50	−10	0.1319 (0.0185)	−0.2877 (0.0945)	0.0324 (0.0037)	0.1056 (0.0234)
		−5	0.1144 (0.0139)	−0.2838 (0.0828)	0.0483 (0.0082)	0.1750 (0.0597)
		−1	0.0126 (0.0005)	−0.1400 (0.0199)	−0.0200 (0.0056)	−0.4831 (0.2352)
		1	0.0124 (0.0005)	0.1635 (0.0273)	0.1065 (0.0219)	1.0849 (1.2121)
		5	0.0657 (0.0097)	−0.0825 (0.0538)	0.0101 (0.0039)	−0.0336 (0.0613)
		10	0.0913 (0.0145)	−0.1506 (0.0695)	0.0189 (0.0066)	−0.0005 (0.0656)

MXBS($\alpha_1, \beta_1, \alpha_2, \beta_2, p$) PDF is either unimodal or bimodal. In special, the parameters of the latter distribution have been chosen to provide bimodal shapes.

The ML estimation results are presented in Tables 3 and 4. The empirical means for the ML estimates and fitted log-likelihood values are reported. A look at the results in Tables 3 and 4 allows us to conclude that the proposed BBS model provides better adjustment compared to the BBSO model based on the log-likelihood values.

5 Real data analysis

The proposed BBS model is now used to analyze three lifetime data sets. For comparison, the results of the bimodal BBSO model (bimodal BS distribution proposed by Olmos,

Table 2 *Simulated values of biases (MSEs within parentheses) of the estimators of the BBS model*

Censoring %	n	δ	BBS($\alpha = 1.0, \beta = 1.0, \delta$) Bias($\widehat{\alpha}$)	BBS($\alpha = 1.0, \beta = 1.0, \delta$) Bias($\widehat{\beta}$)	BBS($\alpha = 1.5, \beta = 1.0, \delta$) Bias($\widehat{\alpha}$)	BBS($\alpha = 1.5, \beta = 1.0, \delta$) Bias($\widehat{\beta}$)
0%	10	−10	−0.0162 (0.0263)	0.0202 (0.0507)	−0.0244 (0.0608)	0.0248 (0.0728)
		−5	−0.0258 (0.0326)	0.0070 (0.0575)	−0.0477 (0.0779)	0.0166 (0.0859)
		−1	−0.2455 (0.1298)	−0.2136 (0.1767)	−0.3764 (0.3075)	−0.2438 (0.2449)
		1	−0.2441 (0.1295)	0.5896 (0.9632)	−0.3801 (0.3087)	0.8931 (2.3962)
		5	−0.0269 (0.0314)	0.0728 (0.1911)	−0.0411 (0.0714)	0.0980 (0.3867)
		10	−0.0160 (0.0253)	0.0367 (0.0831)	−0.0258 (0.0588)	0.0537 (0.1826)
	50	−10	−0.0023 (0.0037)	0.0014 (0.0073)	−0.0033 (0.0083)	0.0012 (0.0105)
		−5	−0.0042 (0.0042)	0.0046 (0.0081)	−0.0049 (0.0093)	0.0058 (0.0114)
		−1	−0.0391 (0.0179)	−0.0630 (0.0486)	−0.0703 (0.0456)	−0.0616 (0.0618)
		1	−0.0385 (0.0175)	0.1410 (0.1272)	−0.0728 (0.0448)	0.1612 (0.1832)
		5	−0.0035 (0.0042)	0.0029 (0.0078)	−0.0050 (0.1832)	0.0021 (0.0115)
		10	−0.0033 (0.0037)	0.0057 (0.0072)	−0.0030 (0.0115)	0.0068 (0.0107)
10%	10	−10	0.0088 (0.0575)	0.1620 (0.4960)	0.0456 (0.1813)	0.2316 (0.6773)
		−5	−0.0074 (0.0713)	0.2010 (0.2840)	0.0573 (0.2973)	0.4123 (1.6875)
		−1	0.0005 (0.1111)	−0.6838 (0.4792)	−0.0498 (0.1986)	−0.7573 (0.5941)
		1	0.1053 (0.1851)	2.4846 (7.1727)	0.1370 (0.3345)	3.7838 (18.1242)
		5	−0.0250 (0.0393)	0.0338 (0.5171)	−0.0236 (0.1127)	0.0605 (1.5485)
		10	−0.0142 (0.0276)	0.0270 (0.1993)	−0.0380 (0.0587)	0.0121 (0.2071)
	50	−10	0.0187 (0.0053)	0.1260 (0.0256)	0.0272 (0.0115)	0.1458 (0.0361)
		−5	0.0219 (0.0060)	0.2016 (0.0569)	0.0314 (0.0134)	0.2368 (0.0820)
		−1	0.0746 (0.0268)	−0.7159 (0.5141)	0.0963 (0.0492)	−0.7910 (0.6272)
		1	0.2470 (0.0974)	2.5919 (6.9146)	0.3733 (0.2098)	3.4797 (12.8450)
		5	0.0028 (0.0043)	−0.1110 (0.0205)	0.0004 (0.0092)	−0.1286 (0.0275)
		10	0.0057 (0.0043)	−0.0619 (0.0103)	0.0077 (0.0093)	−0.0763 (0.0146)
30%	10	−10	0.2393 (0.3985)	0.6891 (4.0713)	0.4426 (1.2441)	1.0896 (7.0831)
		−5	0.1985 (0.2985)	0.6169 (1.9729)	0.7262 (2.1753)	1.9417 (12.9445)
		−1	−0.1458 (0.1069)	−0.6624 (0.4552)	−0.3442 (0.2392)	−0.7627 (0.5970)
		1	0.1667 (0.3212)	2.4961 (7.9861)	0.2472 (0.5322)	3.5858 (15.6473)
		5	0.0959 (0.3339)	0.4870 (5.5723)	0.2082 (1.0045)	0.7895 (11.7672)
		10	0.0774 (0.1579)	0.1944 (0.9839)	0.1840 (0.6911)	0.4070 (4.2938)
	50	−10	0.1884 (0.3646)	0.6260 (8.5780)	0.1899 (0.1168)	0.3775 (0.4992)
		−5	0.2240 (0.1207)	0.6080 (0.7797)	0.4197 (0.4973)	0.9224 (2.5520)
		−1	−0.1144 (0.0284)	−0.6953 (0.4855)	−0.2762 (0.1004)	−0.7957 (0.6349)
		1	0.3004 (0.1388)	2.5498 (6.7053)	0.4939 (0.3467)	3.5056 (12.8719)
		5	0.0436 (0.0771)	0.0425 (1.5990)	0.0406 (0.0558)	−0.0484 (0.9883)
		10	0.0347 (0.0223)	−0.0014 (0.4995)	0.0410 (0.0156)	−0.0309 (0.0156)

Martínez-Flórez and Bolfarine (2017)) and MXBS distribution introduced by Balakrishnan et al. (2011), in addition to classical BS and LN models, are given as well.

Example 5.1. The first data set corresponds to the duration of the eruption for the Old Faithful geyser in Yellowstone National Park, Wyoming, USA; see Azzalini and Bowman (1990). Descriptive statistics for the Old Faithful data set are the following: 272 (sample size), 43 (minimum), 96 (maximum), 76 (median), 70.897 (mean), 13.595 (standard deviation), 19.176 (coefficient of variation), −0.414 (coefficient of skewness) and −1.156 (coefficient of kurtosis). Table 5 reports the ML estimates, computed by the BFGS method, SEs and log-likelihood (log-lik) values for the BBS, BBSO, MXBS, BS and LN models. Furthermore, we report the Akaike (AIC) and Bayesian information (BIC) criteria. From this table, we note that the BBS and MXBS models provide better adjustments compared

Table 3 *Empirical mean from simulated BS and LN data for the indicated model, estimator, generator, α and n*

			BBS				BBSO			
Generator	n	α	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	log-lik	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\gamma}$	log-lik
$\mathrm{BS}(\alpha, \beta = 1.0)$	10	0.50	0.3442	1.0235	0.2160	−4.4525	0.2616	1.0101	−1.6996	−5.1787
		1.00	0.6927	1.0639	0.0270	−10.6575	0.5264	1.0486	−1.6618	−11.3504
		1.50	1.0527	1.0904	−0.0320	−13.8437	0.7942	1.0798	−1.6259	−14.5051
		2.50	1.7765	1.1310	−0.0160	−17.1733	1.3337	1.1228	−1.5805	−17.7920
		4.00	2.8464	1.1624	−0.0280	−19.4770	2.1465	1.1526	−1.5497	−20.0728
	50	0.50	0.4750	1.0187	−0.0020	−33.6220	0.3385	1.0027	−1.2025	−36.7761
		1.00	0.9652	1.0376	0.0180	−64.9110	0.6777	1.0110	−1.1975	−67.9465
		1.50	1.4547	1.0266	0.0030	−81.0343	1.0176	1.0170	−1.1928	−83.9877
		2.50	2.4309	1.0236	−0.0130	−98.0011	1.6989	1.0212	−1.1869	−100.8794
		4.00	3.8981	1.0188	−0.0240	−109.9385	2.7212	1.0229	−1.1836	−112.7911
$\mathrm{LN}(\mu = 1.0, \sigma)$	10	0.50	0.3570	1.0256	0.1610	−4.6916	0.2720	1.0140	−1.6841	−5.4663
		1.00	0.7970	1.1156	0.0710	−11.4631	0.6112	1.0936	−1.5989	−12.3236
		1.50	1.3940	1.2790	0.0310	−15.3910	1.1025	1.2483	−1.4743	−16.3613
		2.50	3.6272	1.9682	−0.1580	−20.5330	3.0648	1.9584	−1.2293	−21.7278
		4.00	14.4744	5.8215	0.0030	−26.1805	13.3911	6.1023	−0.9624	−27.6768
	50	0.50	0.4888	1.0238	0.0000	−34.9625	0.3564	1.0034	−1.1701	−38.7044
		1.00	1.0946	1.0938	0.0070	−69.8828	0.8280	1.0206	−1.0800	−74.9645
		1.50	1.9469	1.1917	0.0060	−91.6809	1.5722	1.0720	−0.9653	−98.3310
		2.50	5.5552	1.5043	0.0030	−125.9847	5.1288	1.3933	−0.7391	−134.8456
		4.00	30.5304	3.0715	0.0040	−175.5051	30.9898	2.9942	−0.5108	−185.1701

to the other models based on the values of AIC and BIC. The null hypothesis of a BS distribution ($\delta = 0$) against an alternative BBS distribution ($\delta \neq 0$) can be tested by using the likelihood ratio (LR) test $\mathrm{LR} = -2(\ell_{\mathrm{BS}}(\widehat{\alpha}, \widehat{\beta}) - \ell_{\mathrm{BBS}}(\widehat{\alpha}, \widehat{\beta}, \widehat{\delta}))$. In this case, we obtain $\mathrm{LR} = -2(-1107.849 + 1050.592) = 114.514$ and comparing it to the 5% critical value from the chi-square distribution with one degree of freedom ($\chi_1^2 = 3.84$), it supports rejection of the null hypothesis, thus the BBS model outperforms, in terms of fitting, the BS one for the data under study.

Figure 3 shows the histogram of the data set superimposed with the fitted curves of the BBS, BBSO, MXBS, BS and LN distributions. From this figure, we clearly note that the BBS captures quite well the inherent bimodality of the data.

Example 5.2. The data used here, which are given by Andrews and Herzberg (1985) who attribute them to a study by Barlow, Toland and Freeman (1984), present the stress rapture life in hours of Kevlar-49/epoxy strands when subjected to a constant sustained pressure until failure. A descriptive summary for the Kevlar-49/epoxy data set provides the following values: 49 (sample size), 1051 (minimum), 17,568 (maximum), 8831 (median), 8805.694 (mean), 4553.915 (standard deviation), 51.176 (coefficient of variation), 0.094 (coefficient of skewness) and −0.915 (coefficient of kurtosis).

Table 6 reports the ML estimates, SEs and log-lik values associated with the BBS, BBSO, MXBS, BS and LN models. Furthermore, we report the values of AIC and BIC. From this table, observe that the proposed BBS model has the lowest values for the AIC and BIC, suggesting that this model provides the best fit to Kevlar-49/epoxy data. To test the null hypothesis of a BS distribution ($\delta = 0$) against an alternative BBS distribution ($\delta \neq 0$), we use the LR test. The result $\mathrm{LR} = -2(-488.4345 + 480.049) = 16.771$ supports the BBS model assumption and rejects the BS model for this data set. This result suggest that the BBS distribution is indeed a good model for the Kevlar-49/epoxy data. A graphical comparison of the fitted BBS, BBSO, MXBS, BS and LN distributions is given in Figure 4.

Table 4 *Empirical mean from simulated MXBS data for the indicated model, estimator, generator, p and n*

			BBS				BBSO			
Generator	n	p	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\delta}$	log-lik	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\gamma}$	log-lik
MXBS($\alpha_1 = 0.1, \beta_1 = 0.5, \alpha_2 = 1.0, \beta_2 = 2.0, p$)	10	0.25	0.6078	1.8486	0.2890	−14.8013	0.4622	1.7921	−1.6841	−15.4262
		0.50	0.5013	1.5308	0.7390	−10.8898	0.3790	1.4707	−1.7584	−11.5081
		0.75	0.3556	1.1169	1.1480	−4.3414	0.2639	1.0798	−1.8864	−5.0188
	50	0.25	0.8350	1.9243	0.3120	−85.1100	0.5803	1.7684	−1.2486	−87.6034
		0.50	0.6721	1.6953	0.7320	−65.5651	0.4772	1.4628	−1.3057	−68.2009
		0.75	0.4502	1.2270	1.2940	−32.7972	0.3395	1.0801	−1.3767	−36.8795
MXBS($\alpha_1 = 1.0, \beta_1 = 0.5, \alpha_2 = 0.5, \beta_2 = 5.0, p$)	10	0.25	0.3295	4.0274	−0.0520	−17.7737	0.2517	3.9848	−1.6814	−18.4817
		0.50	0.3136	2.9319	−0.2210	−14.1957	0.2419	2.8967	−1.6687	−14.9392
		0.75	0.3161	1.8058	−0.2920	−9.4978	0.2434	1.7908	−1.6703	−10.2457
	50	0.25	0.4509	4.0453	0.0570	−99.7776	0.3215	3.9677	−1.2011	−102.9411
		0.50	0.4309	2.9188	−0.0040	−81.8596	0.3080	2.8846	−1.1967	−85.1153
		0.75	0.4322	1.7842	−0.0130	−57.9865	0.3085	1.7790	−1.2038	−61.1831
MXBS($\alpha_1 = 2.5, \beta_1 = 1.0, \alpha_2 = 0.5, \beta_2 = 1.0, p$)	10	0.25	0.5487	1.7084	1.2980	−11.7933	0.4194	1.6472	−1.7132	−12.6414
		0.50	0.7658	2.0200	1.9640	−15.2106	0.5700	1.9589	−1.8182	−15.9973
		0.75	1.0252	1.9457	1.6240	−17.0309	0.7539	1.8879	−1.8284	−17.6494
	50	0.25	0.7016	1.9251	0.9980	−68.7779	0.5432	1.6049	−1.2276	−74.1506
		0.50	0.9403	2.2861	1.4130	−85.5513	0.7108	1.9387	−1.4005	−89.7587
		0.75	1.2874	2.1852	1.2490	−94.6959	0.9241	1.8571	−1.4378	−96.8759

Table 5 *ML estimates and model selection measures for fit to the Old Faithful data*

Model	Parameter	ML estimate	SE	log-lik	AIC	BIC
BBS	α	0.1255	0.0034	−1050.592	2107.184	2118.001
	β	66.8612	0.4739			
	δ	−4				
BBSO	α	0.0893	0.0047	−1054.396	2114.792	2125.609
	β	65.7730	0.4128			
	γ	−2.1803	0.1432			
MXBS	α_1	0.1150	0.0046	−1032.681	2075.362	2093.391
	α_2	0.0697	0.0108			
	β_1	54.8174	0.4824			
	β_2	80.1850	0.7607			
	p	0.3762	0.0317			
BS	α	0.2055	0.0088	−1107.849	2221.698	2226.91
	β	69.4289	0.8608			
LN	μ	4.2411	0.0124	−1108.300	2222.6	2227.812
	σ	0.2048	0.0087			

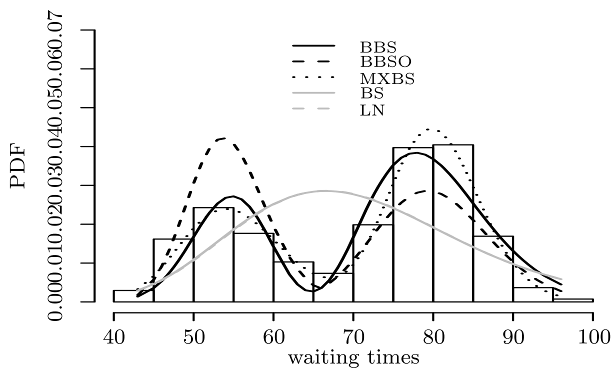

Figure 3 *Histogram of waiting times until the next eruption (from the Old Faithful data) overlaid with the fitted densities.*

Example 5.3. The third data set corresponds to the lifetimes of adult flies in days after exposure to a pest control technique, which consists of using small portions of food laced with an insecticide that kills the flies. The experiment was carried out at the Department of Entomology of the Luiz de Queiroz School of Agriculture, University of São Paulo, Brazil. In this technique, the period was set at 51 days such that larvae that survived beyond this period are considered as censored cases; see Silvia et al. (2013) for more details about this experiment. Descriptive statistics for the Entomology data are the following: sample size = 172 (four cases are lost), minimum = 1.000, maximum = 51.000, median = 21.000, mean = 21.878, standard deviation = 11.674, coefficient of variation = 53.30, coefficient of skewness = 0.818 and coefficient of kurtosis = 0.569.

The ML estimates and log-lik values for the BBS, BBSO, MXBS, BS and LN models are reported in Table 7. Furthermore, the AIC and BIC values are also reported in this table. From Table 7, we note that the proposed BBS model has the lowest AIC and BIC values, and therefore it could be chosen as the best model. Using the LR statistic to compare the fits of the BS and BBS models, that is, the null hypothesis of a BS distribution ($\delta = 0$) against an alternative BBS distribution ($\delta \neq 0$), we obtain $\text{LR} = -2(-676.913 + 610.523) = 132.780$

Table 6 *ML estimates and model selection measures for fit to the Kevlar-49/epoxy data*

Model	Parameter	ML estimate	SE	log-lik	AIC	BIC
BBS	α	0.5933	0.0504	−480.049	966.098	971.773
	β	4507.365	376.7557			
	δ	−2				
BBSO	α	0.4679	0.08655	−490.208	986.417	992.092
	β	5036.113	437.0908			
	γ	−1.5345	0.4708			
MXBS	α_1	0.3412	0.0166	−593.987	1197.974	1207.433
	α_2	0.0697	0.0123			
	β_1	10,278.38	979.1201			
	β_2	4577.872	50.2859			
	p	0.6708	0.0991			
BS	α	0.7520	0.0759	−488.434	982.869	984.652
	β	6800.546	679.8509			
LN	μ	8.8925	0.1001	−487.873	981.746	983.530
	σ	0.7012	0.0708			

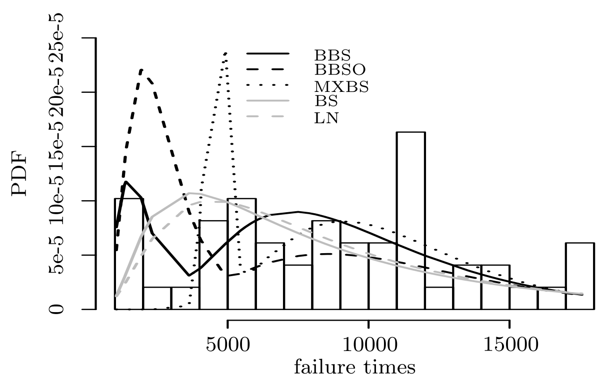

Figure 4 *Histogram of Kevlar 49/epoxy strands failure times (70% pressure) overlaid with the fitted densities.*

and then we could accept the BBS model. Figure 5 shows the fitted PDFs and SFs (by Kaplan–Meier (KM) estimator) of the BBS, BBSO, BS and LN distributions.

6 Concluding remarks

In this work, we have introduced a bimodal generalization of the Birnbaum–Saunders distribution, based on the alpha-skew-normal distribution. We have discussed some of its properties. We have considered estimation and inference based on likelihood methods. We have carried out a Monte Carlo simulation study to evaluate the behavior of the maximum likelihood estimators of the corresponding parameters. Three real data sets were considered to illustrate the potentiality of the proposed model. In general, the results have shown that the proposed bimodal Birnbaum–Saunders distribution outperforms some existing models in the literature. As part of future research, it is of interest to study univariate and multivariate bimodal Birnbaum–Saunders regression models; see Rieck and Nedelman (1991), Balakrishnan and Zhu (2015) and Marchant, Leiva and Cysneiros (2016). Moreover, time series models based on the bimodal Birnbaum–Saunders distribution with corresponding influ-

Table 7 *ML estimates and model selection measures for fit to the Entomology data*

Model	Parameter	ML estimate	SE	log-lik	AIC	BIC
BBS	α	0.6922	0.0385	−610.523	1227.052	1236.494
	β	8.7636	0.5311			
	δ	−2				
BBSO	α	0.4975	0.0437	−663.144	1332.287	1341.730
	β	7.5959	0.5250			
	γ	−2.2434	0.2859			
MXBS	α_1	1.2349	0.0115	−631.137	1266.273	1272.568
	α_2	0.2104	0.0009			
	β_1	14.6760	2.7919			
	β_2	19.9960	0.4717			
	p	0.60				
BS	α	0.8912	0.0500	−676.913	1357.862	1364.157
	β	16.1512	1.0078			
LN	μ	2.9139	0.0583	−660.230	1324.461	1330.756
	σ	0.7613	0.0708			

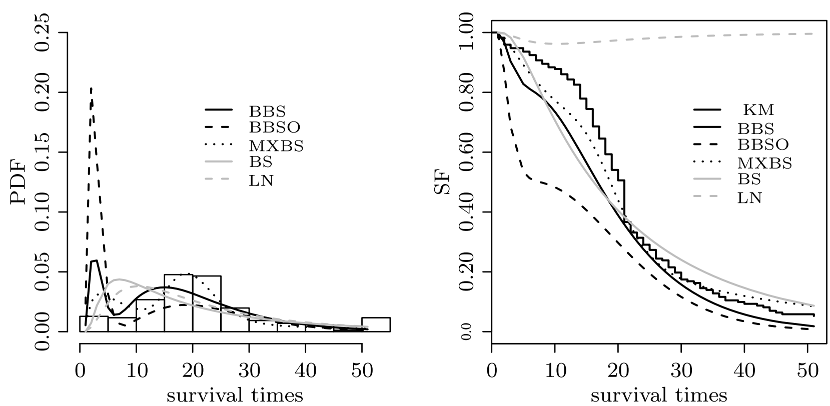

Figure 5 *Histogram and SF fitted by KM with the Entomology data.*

ence diagnostic tools can also be considered; see Saulo et al. (2019). Work on these problems is currently under progress and we hope to report these findings in a future paper.

Acknowledgments

The authors thank the Editors and reviewers for their constructive comments on an earlier version of this manuscript. The research was partially supported by CNPq and CAPES grants from the Brazilian federal government and by FAPEAM grants from the government of the State of Amazonas, Brazil.

References

Andrews, D. F. and Herzberg, A. M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. *Springer Series in Statistics*. New York: Springer. https://doi.org/10.1007/978-1-4612-5098-2. https://doi.org/10.1007/978-1-4612-5098-2

Azzalini, A. and Bowman, A. W. (1990). A look at some data on the old faithful geyser. *Applied Statistics* **39**, 357–365. https://doi.org/10.2307/2347385

Azzalini, A. and Capitanio, A. (2003). Distributions generate by perturbation of symmetry with emphasis on a multivariate skew-*t* distribution. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **65**, 367–389. MR1983753

Balakrishnan, N., Gupta, R. C., Kundu, D., Leiva, V. and Sanhueza, A. (2011). On some mixture models based on the Birnbaum–Saunders distribution and associated inference. *Journal of Statistical Planning and Inference* **141**, 2175–2190. MR2775197 https://doi.org/10.1016/j.jspi.2010.12.005

Balakrishnan, N., Leiva, V. and López, J. (2007). Acceptance sampling plans from truncated life tests based on the generalized Birnbaum–Saunders distribution. *Communications in Statistics Simulation and Computation* **36**, 643–656. https://doi.org/10.1080/03610910701207819. MR2370927 https://doi.org/10.1080/03610910701207819

Balakrishnan, N., Leiva, V., Sanhueza, A. and Vilca, F. (2009). Estimation in the Birnbaum–Saunders distribution based on scale-mixture of normals. *Statistics and Operations Research Transactions* **33**, 171–192. MR2643505

Balakrishnan, N. and Zhu, X. (2015). Inference for the Birnbaum–Saunders lifetime regression model with applications. *Communications in Statistics Simulation and Computation* **44**, 2073–2100. https://doi.org/10.1080/03610918.2013.844838. MR3345743 https://doi.org/10.1080/03610918.2013.844838

Barlow, R. E., Toland, R. H. and Freeman, T. (1984). A Bayesian analysis of stress-rupture life of Kevlar 49/epoxy spherical pressure vessels. In *Proceedings of the Canadian Conference in Applied Statistics* (T. D. Dwivedi, ed.) New York: Marcel Dekker.

Bhatti, C. R. (2010). The Birnbaum–Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation* **80**, 2062–2078. MR2665320

Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability* **6**, 319–327. MR0253493

Celeux, G., Forbes, F., Robert, C. P. and Titterington, D. M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis* **4**, 651–673. MR2282197 https://doi.org/10.1214/06-BA122

Chung, K. L. (2001). *A Course in Probability Theory*, 3rd ed. San Diego, CA: Academic Press, Inc. MR1796326

Díaz-García, J. A. and Leiva, V. (2005). A new family of life distributions based on elliptically contoured distributions. *Journal of Statistical Planning and Inference* **128**, 445–457. MR2102769 https://doi.org/10.1016/j.jspi.2003.11.007

Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed vs. expected Fisher information. *Biometrika* **65**, 457–487. MR0521817 https://doi.org/10.1093/biomet/65.3.457

Elal-Olivero, D. (2010). Alpha-skew-normal distribution. *Proyecciones* **29**, 224–240. MR2749567

Fonseca, R. and Cribari-Neto, F. (2018). Inference in a bimodal Birnbaum–Saunders model. *Mathematics and Computers in Simulation* **146**, 134–159. MR3739543 https://doi.org/10.1016/j.matcom.2017.11.004

Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association* **75**, 667–672. MR0590699

Gómez, H. W., Elal-Olivero, D., Salinas, H. S. and Bolfarine, H. (2011). Bimodal extension based on the skew-normal distribution with application to pollen data. *EnvironMetrics* **22**, 50–62. MR2843336

Kim, H. J. (2005). On a class of two-piece skew-normal distributions. *Statistics* **39**, 537–553. MR2211732

Leao, J., Leiva, V., Saulo, H. and Tomazella, V. (2017). Birnbaum–Saunders frailty regression models: Diagnostics and application to medical data. *Biometrical Journal* **59**, 291–314. MR3623342

Leiva, V. (2016). *The Birnbaum–Saunders Distribution*. Amsterdam: Elsevier/Academic Press. MR3430824

Leiva, V., Marchant, C., Saulo, H., Aslam, M. and Rojas, F. (2014a). Capability indices for Birnbaum–Saunders processes applied to electronic and food industries. *Journal of Applied Statistics* **41**, 1881–1902. MR3292616

Leiva, V., Saulo, H., Leão, J. and Marchant, C. (2014b). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics & Data Analysis* **79**, 175–191. MR3227995

Lemonte, A. J., Simas, A. B. and Cribari-Neto, F. (2008). Bootstrap-based improved estimators for the two-parameter Birnbaum–Saunders distribution. *Journal of Statistical Computation and Simulation* **78**, 37–49. MR2412760

Lin, T. I., Lee, J. C. and Hsieh, W. J. (2007). Robust mixture models using the skew-*t* distribution. *Statistics and Computing* **17**, 81–92. MR2380638

Lin, T. I., Lee, J. C. and Yen, S. Y. (2007). Finite mixture modeling using the skew-normal distribution. *Statistica Sinica* **17**, 81–92. MR2408641

Ma, Y. and Genton, M. G. (2004). Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics* **31**, 459–468. MR2087837

Marchant, C., Leiva, V. and Cysneiros, F. J. A. (2016). A multivariate log-linear model for Birnbaum–Saunders distributions. *IEEE Transactions on Reliability* **65**, 816–827. https://doi.org/10.1109/TR.2015.2499964

Mittelhammer, R. C., Judge, G. G. and Miller, D. J. (2000). *Econometric Foundations*. Cambridge, UK: Cambridge University Press. MR1789434

Natanson, I. P. (1955). *Theory of Functions of a Real Variable (Translated by Leo F. Boron with the Collaboration of Edwin Hewitt)*. New York: Frederick Ungar Publishing Co. MR0067952

Olmos, N. O., Martínez-Flórez, M. and Bolfarine, H. (2017). Bimodal Birnbaum–Saunders distribution with applications to non-negative measurements. *Communications in Statistics Theory and Methods* **46**, 6240–6257. MR3631510

Paula, G. A., Leiva, V., Barros, M. and Liu, S. (2012). Robust statistical modeling using the Birnbaum–Saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry* **28**, 16–34. MR2898899

R-Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum–Saunders distribution. *Technometrics* **33**, 51–60. https://doi.org/10.2307/1269007

Saulo, H., Leao, J., Leiva, V. and Aykroyd, R. G. (2019). Birnbaum–Saunders autoregressive conditional duration models applied to high-frequency financial data. *Statistical Papers*. **60**, 1605–1629. https://doi.org/10.1007/s00362-017-0888-6

Saulo, H., Leiva, V., Ziegelmann, F. A. and Marchant, C. (2010). A nonparametric method for estimating asymmetric densities based on skewed Birnbaum–Saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment* **27**, 1479–1491. https://doi.org/10.1007/s00477-012-0684-8.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: Univ. of Illinois Press. MR0032134

Silvia, M. A., Bezerra-Silva, G. C. D., Vendramim, J. D. and Mastrangelo, T. (2013). Sublethal effect of neem extract on Mediterranean fruit fly adults. *Revista Brasileira de Fruticultura* **35**, 93–101.

Vilca, F., Sanhueza, A., Leiva, V. and Christakos, G. (2010). An extended Birnbaum–Saunders model and its application in environmental quality in Santiago. *Stochastic Environmental Research and Risk Assessment* **24**, 771–782. https://doi.org/10.1007/s00477-009-0363-6.

Vinberg, E. B. (2003). *A Course in Algebra*. Providence, R.I.: American Mathematical Society. ISBN 0-8218-3413-4. MR1974508 https://doi.org/10.1090/gsm/056

R. Vila
H. Saulo
Department of Statistics
University of Brasília
Brasília, DF 70910-900
Brazil
E-mail: rovig161@gmail.com; heltonsaulo@gmail.com

J. Leao
Department of Statistics
Federal University of Amazonas
Manaus, AM 69077-000
Brazil
E-mail: leaojeremiass@gmail.com

M. N. Shahzad
Department of Statistics
University of Gujrat
Gujrat 50700
Pakistan
E-mail: nvd.shzd@uog.edu.pk

M. Santos-Neto
Department of Statistics
Federal University of Campina Grande
Campina Grande, PB 58429-900
Brazil
E-mail: mn.neco@gmail.com
URL: www.santosnetoce.com.br