Brazilian Journal of Probability and Statistics
2013, Vol. 27, No. 2, 133–149
DOI: 10.1214/11-BJPS160

A new extension of the Birnbaum–Saunders distribution

Artur J. Lemonte

Universidade de São Paulo

Abstract. In this paper, a new extension for the Birnbaum–Saunders distribution, which has been applied to the modeling of fatigue failure times and reliability studies, is introduced. The proposed model, called the Marshall–Olkin extended Birnbaum–Saunders distribution, arises based on the scheme introduced by Marshall and Olkin [*Biometrika* **84** (1997) 641–652]. The maximum likelihood estimators and statistical inference for the new distribution parameters and influence diagnostic for the new distribution are presented. Finally, the proposed new distribution is applied to model three real data sets.

1 Introduction

The crack growth caused by vibrations in commercial aircrafts motivated Birnbaum and Saunders (1969a, 1969b) to develop a new family of two-parameter distributions for modeling the failure time due to fatigue under cyclic loading. Relaxing assumptions made by Birnbaum and Saunders (1969a), Desmond (1985) presented a more general derivation of the Birnbaum–Saunders (BS) distribution based on a biological model. The relationship between the BS distribution and the inverse Gaussian distribution was investigated by Desmond (1986). The author also demonstrated that the BS distribution is an equal-weight mixture of an inverse Gaussian distribution and its complementary reciprocal. The two-parameter BS distribution is also known as the fatigue life distribution. This distribution is an attractive alternative to the Weibull, gamma and log-normal models, since its derivation considers the basic characteristics of the fatigue process.

The random variable T is said to have a BS distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, say, $\mathrm{BS}(\alpha, \beta)$, if its cumulative distribution function (c.d.f.) is given by $F(t) = \Phi(v)$, $t > 0$, where $\Phi(\cdot)$ is the standard normal distribution function, $v = \rho(t/\beta)/\alpha$ and $\rho(z) = z^{1/2} - z^{-1/2}$. Since $F(\beta) = \Phi(0) = 1/2$, β is the median of the distribution. For any constant $k > 0$, it follows that $kT \sim \mathrm{BS}(\alpha, k\beta)$. The reciprocal property holds for the BS distribution, that is, $T^{-1} \sim \mathrm{BS}(\alpha, \beta^{-1})$. The probability density function (p.d.f.) and hazard ratio function (h.r.f.) are given by (for $t > 0$)

$$f(t) = \kappa(\alpha, \beta)t^{-3/2}(t + \beta)\exp\left\{-\frac{\tau(t/\beta)}{2\alpha^2}\right\} \tag{1.1}$$

Key words and phrases. Birnbaum–Saunders distribution, fatigue life distribution, lifetime data, Marshall–Olkin extended distribution, maximum likelihood estimation.

Received October 2010; accepted June 2011.

and

$$h_{\mathrm{BS}}(t) = \frac{\kappa(\alpha,\beta)t^{-3/2}(t+\beta)\exp\{-\tau(t/\beta)/(2\alpha^2)\}}{1-\Phi(v)}, \tag{1.2}$$

respectively, where $\kappa(\alpha,\beta) = \exp(\alpha^{-2})/(2\alpha\sqrt{2\pi\beta})$ and $\tau(z) = z + z^{-1}$. For the applications of the BS distributions, read, for example, Balakrishnan et al. (2007) in reliability and Leiva et al. (2008, 2009) in other fields. The BS distribution has received significant attention over the last few years by many researchers such as Wu and Wong (2004), Kundu et al. (2008), Lemonte et al. (2007, 2008), Xu and Tang (2010) and Bhatti (2010), among others. Some generalizations and extensions of the BS distribution are proposed in Díaz-García and Leiva (2005), Owen (2006), Vilca and Leiva (2006), Gómes et al. (2009), Guiraud et al. (2009), Leiva et al. (2009), Leiva et al. (2010) and Cordeiro and Lemonte (2011). In this study, a new three-parameter extension for the BS distribution is proposed. An advantage of this new model over all the other generalizations of the BS distribution mentioned above is that this new extended model is geometrically extremely stable. This stable behavior makes that a bivariate lifetime distribution can be easily built (see Section 2).

Marshall and Olkin (1997) introduced an interesting method of adding a new parameter to an existing F distribution. The resulting distribution, known as the Marshall–Olkin (MO) extended-F distribution, includes the original distribution as a special case and gives more flexibility to model various types of data. Let $\bar{F}(t) = 1 - F(t)$ denote the survival function of a continuous random variable T. Then, the associated MO extended-F distribution has survival function given by

$$\bar{G}(t) = \frac{\eta\bar{F}(t)}{1-\bar{\eta}\bar{F}(t)} = \frac{\eta\bar{F}(t)}{F(t)+\eta\bar{F}(t)}, \qquad -\infty < t < \infty, \eta > 0, \tag{1.3}$$

where $\bar{\eta} = 1 - \eta$. Clearly, equation (1.3) provides a tool to obtain new parametric distributions from existing ones. For $\eta = 1$, $\bar{G}(t) = \bar{F}(t)$ and, therefore, $\bar{F}(t)$ is a basic exemplar of (1.3). The p.d.f. corresponding to (1.3), say, $g(t)$, is given by

$$g(t) = \frac{\eta f(t)}{\{1-\bar{\eta}\bar{F}(t)\}^2}, \qquad -\infty < t < \infty, \eta > 0, \tag{1.4}$$

where $f(t)$ is the density function corresponding to the baseline c.d.f. $F(t)$. Recently, Gómez-Déniz (2010) and García at al. (2010) used the MO scheme for generalizing the geometric and normal distributions, respectively.

In this note, we introduce a new distribution, called the MO extended BS (denoted with the prefix "MOEBS" for short) distribution of which density function can be obtained from (1.4) by taking $f(t)$ to be the p.d.f. of the BS(α,β) distribution. We discuss maximum likelihood estimation of the model parameters, derive the observed information matrix and present some properties of the new class of distributions. The article is outlined as follows. In Section 2, we introduce the MOEBS distribution and provide plots of the density function and hazard ratio function. Maximum likelihood estimation is addressed in Section 3. Influence diagnostic is presented in Section 4. Empirical applications are presented and discussed in Section 5. Finally, concluding remarks are given in Section 6.

2 A new three-parameter BS distribution

In this section, we develop the MOEBS distribution. For this new distribution, we present the c.d.f., p.d.f., h.r.f., moments and discuss some properties. If a random variable T follows a MOEBS distribution, then the notation $T \sim \mathrm{MOEBS}(\eta, \alpha, \beta)$ is used. The c.d.f. of T can be written as

$$G(t) = \frac{\Phi(v)}{1 - \bar{\eta}\Phi(-v)}, \qquad t > 0. \tag{2.1}$$

The survival function of T is $\bar{G}(t) = 1 - G(t) = \eta\Phi(-v)/\{1 - \bar{\eta}\Phi(-v)\}$. The p.d.f. of T is given by

$$g(t) = \frac{\eta\kappa(\alpha, \beta)t^{-3/2}(t + \beta)}{[1 - \bar{\eta}\Phi(-v)]^2} \exp\left\{-\frac{\tau(t/\beta)}{2\alpha^2}\right\}, \qquad t > 0. \tag{2.2}$$

If $T \sim \mathrm{MOEBS}(\eta, \alpha, \beta)$, then $kT \sim \mathrm{MOEBS}(\eta, \alpha, k\beta)$, for $k > 0$, that is, the class of MOEBS distributions is closed under scale transformations, as in the case of the BS distribution. Additionally, this new distribution has an interesting property. If T_i $(i = 1, 2, \ldots)$ is a sequence of independent and identically distributed random variables with c.d.f. as in (2.1) and if N has a geometric distribution taken values $\{1, 2, \ldots\}$, then the random variables $U = \min\{T_1, \ldots, T_N\}$ and $V = \max\{T_1, \ldots, T_N\}$ are distributed as in (2.1). It implies that the MOEBS distribution is geometrically extremely stable (Marshall and Olkin, 1997). Thus, bivariate distributions can be built. Bivariate distributions generated using the above procedure are useful in lifetime distributions (see, for instance, Sarhan and Balakrishnan, 2007).

The h.r.f. of T has the form

$$h(t) = \frac{h_{\mathrm{BS}}(t)}{1 - \bar{\eta}\Phi(-v)}, \qquad t > 0, \tag{2.3}$$

where $h_{\mathrm{BS}}(t)$ is given in (1.2). From (2.3), note that $h(t)/h_{\mathrm{BS}}(t)$ is increasing in t for $\eta \geq 1$ and decreasing in t for $0 \leq \eta \leq 1$. Additionally, we have (for $\eta \geq 1$)

$$\frac{h_{\mathrm{BS}}(t)}{\eta} \leq h(t) \leq h_{\mathrm{BS}}(t), \qquad \Phi(-v) \leq \bar{G}(t) \leq \Phi(-v)^{1/\eta}$$

and

$$h_{\mathrm{BS}}(t) \leq h(t) \leq \frac{h_{\mathrm{BS}}(t)}{\eta}, \qquad \Phi(-v)^{1/\eta} \leq \bar{G}(t) \leq \Phi(-v)$$

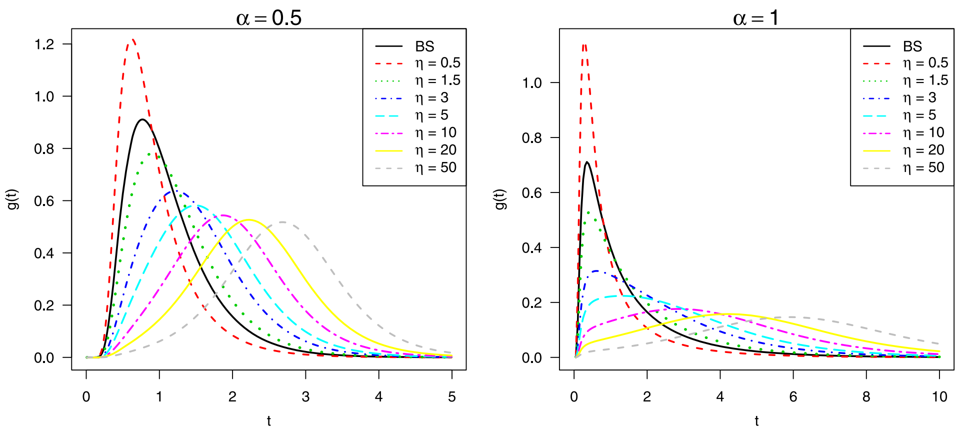

Figure 1 *Plots of the density function* (2.2) *for some parameter values*; $\beta = 1$.

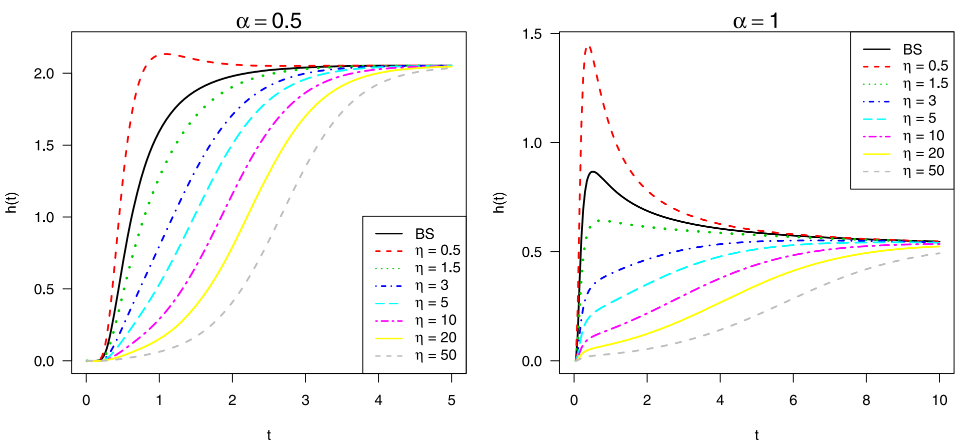

Figure 2 *Plots of the hazard ratio function* (2.3) *for some parameter values*; $\beta = 1$.

for $0 < \eta \leq 1$, where $\Phi(-v)$ is the survival function of the BS distribution. It can also be shown that

$$\lim_{t \to \infty} h(t) = \frac{1}{2\alpha^2 \beta}$$

and, hence, we have that the limit behavior of the h.r.f. of the MOEBS distribution is the same as that of the BS distribution.

Figures 1 and 2 illustrate some of the possible shapes of (2.2) and (2.3) for selected values of the parameters. The plots in these figures show that the MOEBS distribution is very versatile and that the value of η has a substantial effect on its skewness and kurtosis. It is evident that the MOEBS distribution is much more flexible than the BS distribution.

Table 1 *Numerical values of the function* $c_s(\eta, \alpha, \beta)$; $\alpha = 0.5$ *and* $\beta = 1$

	η							
s	0.5	1	1.5	3	5	10	20	50
1	1.864	1.125	0.836	0.499	0.339	0.198	0.114	0.054
2	2.231	1.594	1.300	0.902	0.680	0.453	0.295	0.161
3	3.447	2.805	2.458	1.920	1.570	1.163	0.836	0.517
4	6.772	6.018	5.550	4.725	4.115	3.316	2.588	1.778

The sth moment of the MOEBS(η, α, β) distribution is given by

$$\mu'_s = \mathrm{E}(T^s) = \eta c_s(\eta, \alpha, \beta), \qquad s > 0, \tag{2.4}$$

where

$$c_s(\eta, \alpha, \beta) = \int_0^\infty \frac{s t^{s-1} \Phi(-v)}{1 - \bar{\eta}\Phi(-v)}\, dt.$$

It is not known how $c_s(\eta, \alpha, \beta)$ can be reduced to a closed-form expression. However, this integral can be easily computed numerically in software such as MAPLE (Garvan, 2002), MATLAB (Sigmon and Davis, 2002), MATHEMATICA (Wolfram, 2003), Ox (Doornik, 2006) and R (R Development Core Team, 2010). Numerical values of the function $c_s(\eta, \alpha, \beta)$ for some values of the parameters are presented in Table 1. Note that for $\eta = 1$, the moments agree with the respective moments of the BS distribution. The skewness and kurtosis measures can be calculated from the ordinary moments given in (2.4) using well-known relationships.

The MOEBS quantile function is given by

$$t = Q(u) = \beta \left\{ \frac{\alpha}{2} \Phi^{-1}\left(\frac{\eta u}{1 - \bar{\eta} u}\right) + \left[1 + \frac{\alpha^2}{4} \Phi^{-1}\left(\frac{\eta u}{1 - \bar{\eta} u}\right)^2\right]^{1/2} \right\}^2,$$

where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal c.d.f., that is, $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal c.d.f. $\Phi(\cdot)$. Thus, the MOEBS distribution is easily simulated as follows: if $U \sim \mathcal{U}(0, 1)$, then $T = Q(U)$ has the MOEBS(η, α, β) distribution. This scheme is useful because of the existence of fast generators for uniform random variables. Additionally, it follows immediately that the median (M) of the distribution of T is given by $M = Q(1/2)$, and if $\eta = 1$, that is, when $T \sim \mathrm{BS}(\alpha, \beta)$, then $M = \beta$.

3 Maximum likelihood estimation

Let $\mathbf{t} = (t_1, \ldots, t_n)^\top$ denote a random sample of size n of the MOEBS distribution with unknown parameter vector $\boldsymbol{\theta} = (\eta, \alpha, \beta)^\top$. We shall consider estimation of the parameters of the MOEBS distribution by the method of maximum likelihood. The total log-likelihood function for $\boldsymbol{\theta}$ is

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}), \tag{3.1}$$

where

$$\begin{aligned}\ell_i(\boldsymbol{\theta}) = \log\{\eta\kappa(\alpha,\beta)\} - \frac{3}{2}\log(t_i) + \log(t_i+\beta)\\ - \frac{\tau(t_i/\beta)}{2\alpha^2} - 2\log\{1-\bar{\eta}\Phi(-v_i)\}.\end{aligned}$$

The components of the score vector $\mathbf{U}_{\boldsymbol{\theta}} = (U_\eta, U_\alpha, U_\beta)^\top$, which are obtained by taking the partial derivatives of the above log-likelihood function with respect to η, α and β, are given by

$$U_\eta = \frac{n}{\eta} - 2\sum_{i=1}^{n}\frac{\Phi(-v_i)}{1-\bar{\eta}\Phi(-v_i)},$$

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

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

where $\phi(\cdot)$ is the standard normal density 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 $i = 1,\dots,n$.

The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}} = (\widehat{\eta}, \widehat{\alpha}, \widehat{\beta})^\top$ of $\boldsymbol{\theta} = (\eta,\alpha,\beta)^\top$ is obtained by solving the likelihood equations, $U_\eta = 0$, $U_\alpha = 0$ and $U_\beta = 0$, simultaneously. These equations cannot be solved analytically. However, they can be solved by a numerical method through the implementation of a statistical software. For example, the Newton–Raphson iterative technique could be applied to solve the likelihood equations and obtain the estimate $\widehat{\boldsymbol{\theta}}$. The BFGS method (see, e.g., Nocedal and Wright, 1999; Press et al., 2007) with analytical derivatives has been used for maximizing the log-likelihood function $\ell(\boldsymbol{\theta})$. As starting values for the algorithm, we suggest for η, α and β, the values

$$\widetilde{\eta} = 1, \qquad \widetilde{\alpha} = \sqrt{\frac{\bar{s}}{\widetilde{\beta}} + \frac{\widetilde{\beta}}{\bar{r}} - 2} \quad\text{and}\quad \widetilde{\beta} = \sqrt{\bar{s}\bar{r}},$$

respectively, where

$$\bar{s} = \frac{1}{n}\sum_{i=1}^{n} t_i \quad\text{and}\quad \bar{r} = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{t_i}\right)^{-1}.$$

These starting values worked well in the applications considered in Section 5.

The asymptotic inference for the parameter vector $\boldsymbol{\theta} = (\eta, \alpha, \beta)^\top$ can be based on the normal approximation of the MLE of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\theta}} = (\widehat{\eta}, \widehat{\alpha}, \widehat{\beta})^\top$. Under some regular conditions stated in Cox and Hinkley (1974, Chapter 9) that are fulfilled for the parameters in the interior of the parameter space, we have $\widehat{\boldsymbol{\theta}} \overset{a}{\sim} \mathcal{N}_3(\theta, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$, for n large, where $\overset{a}{\sim}$ means approximately distributed and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is the asymptotic variance–covariance matrix of $\widehat{\boldsymbol{\theta}}$. The asymptotic behavior remains valid if $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is approximated by $-\ddot{\mathbf{L}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}^{-1}$, where $-\ddot{\mathbf{L}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}$ is the 3×3 observed information matrix evaluated at $\widehat{\boldsymbol{\theta}}$, obtained from

$$\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}} = \begin{bmatrix} \ddot{L}_{\eta\eta} & \ddot{L}_{\eta\alpha} & \ddot{L}_{\eta\beta} \\ \cdot & \ddot{L}_{\alpha\alpha} & \ddot{L}_{\alpha\beta} \\ \cdot & \cdot & \ddot{L}_{\beta\beta} \end{bmatrix},$$

whose elements are given in the Appendix. The multivariate normal $\mathcal{N}_3(\mathbf{0}, -\ddot{\mathbf{L}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}^{-1})$ distribution can be used to construct approximate confidence intervals for the parameters η, α and β, which are given, respectively, by $\widehat{\eta} \pm z_{\gamma/2} \times [\widehat{\mathrm{var}}(\widehat{\eta})]^{1/2}$, $\widehat{\alpha} \pm z_{\gamma/2} \times [\widehat{\mathrm{var}}(\widehat{\alpha})]^{1/2}$ and $\widehat{\beta} \pm z_{\gamma/2} \times [\widehat{\mathrm{var}}(\widehat{\beta})]^{1/2}$, where $\mathrm{var}(\cdot)$ is the diagonal element of $-\ddot{\mathbf{L}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}^{-1}$ corresponding to each parameter, and $z_{\gamma/2}$ is the quantile $100(1-\gamma/2)\%$ of the standard normal distribution.

We can easily check if the fit using the MOEBS model is statistically "superior" to a fit using the BS model by testing the null hypothesis $\mathcal{H}_0 : \eta = 1$ against $\mathcal{H}_1 : \eta \neq 1$. For testing $\mathcal{H}_0$, the likelihood ratio (LR) statistic is given by $w = 2\{\ell(\widehat{\eta}, \widehat{\alpha}, \widehat{\beta}) - \ell(1, \widetilde{\alpha}, \widetilde{\beta})\}$, where $\widehat{\eta}$, $\widehat{\alpha}$ and $\widehat{\beta}$ are the unrestricted MLEs obtained from the maximization of $\ell(\boldsymbol{\theta})$ under $\mathcal{H}_1$ and $\widetilde{\alpha}$ and $\widetilde{\beta}$ are the restricted MLEs obtained from the maximization of $\ell(\boldsymbol{\theta})$ under $\mathcal{H}_0$. The limiting distribution of this statistic is χ_1^2 under the null hypothesis. The null hypothesis is rejected if the test statistic exceeds the upper $100(1-\gamma)\%$ quantile of the χ_1^2 distribution.

Next, a small Monte Carlo simulation is conducted to evaluate the estimations of the MOEBS distribution parameters. The simulation was performed using the Ox matrix programming language (Doornik, 2006). The Ox program is freely distributed for academic purposes and available at http://www.doornik.com. The number of Monte Carlo replications was $R = 1000$. For maximizing the log-likelihood function, we use the subroutine `MaxBFGS` with analytical derivatives.

The evaluation of point estimation was performed based on the following quantities for each sample size: the empirical mean and the root mean squared error, $\sqrt{\mathrm{MSE}}$, where MSE is the mean squared error estimated from R Monte Carlo replications. We set the sample size at $n = 100$ and 200, the parameter α at $\alpha = 0.5$, 1.0 and 1.5, and the parameter η at $\eta = 0.4$, 0.8, 1.2, 1.8, 2.5 and 5. Without loss of generality, the scale parameter β was fixed at 1.0. It can be seen from Table 2 that the estimates are quite stable and, more important, are close to the true values for the sample sizes considered.

Table 2 *Empirical means and root mean squared error in parentheses*

		$n = 100$			$n = 200$		
α	η	$\widehat{\eta}$	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\eta}$	$\widehat{\alpha}$	$\widehat{\beta}$
0.5	0.4	0.522 (0.396)	0.502 (0.044)	1.013 (0.210)	0.457 (0.235)	0.501 (0.031)	1.007 (0.144)
	0.8	0.982 (0.719)	0.502 (0.038)	1.018 (0.189)	0.885 (0.431)	0.501 (0.026)	1.010 (0.131)
	1.2	1.440 (1.072)	0.502 (0.038)	1.022 (0.186)	1.312 (0.639)	0.501 (0.026)	1.013 (0.130)
	1.8	2.125 (1.644)	0.501 (0.040)	1.029 (0.190)	1.952 (0.976)	0.500 (0.028)	1.016 (0.133)
	2.5	2.923 (2.371)	0.500 (0.044)	1.035 (0.199)	2.700 (1.404)	0.500 (0.030)	1.020 (0.140)
	5.0	5.764 (5.374)	0.497 (0.056)	1.058 (0.234)	5.392 (3.225)	0.498 (0.040)	1.031 (0.166)
1.0	0.4	0.459 (0.223)	0.998 (0.086)	1.012 (0.248)	0.428 (0.141)	0.998 (0.061)	1.007 (0.173)
	0.8	0.883 (0.402)	1.000 (0.073)	1.023 (0.228)	0.839 (0.264)	0.999 (0.052)	1.013 (0.159)
	1.2	1.302 (0.584)	0.999 (0.073)	1.032 (0.228)	1.247 (0.391)	0.999 (0.051)	1.018 (0.159)
	1.8	1.922 (0.866)	0.998 (0.078)	1.043 (0.238)	1.857 (0.588)	0.998 (0.055)	1.024 (0.165)
	2.5	2.636 (1.205)	0.996 (0.086)	1.054 (0.253)	2.563 (0.828)	0.997 (0.060)	1.030 (0.174)
	5.0	5.121 (2.502)	0.987 (0.112)	1.092 (0.313)	5.056 (1.764)	0.992 (0.079)	1.049 (0.211)
1.5	0.4	0.436 (0.157)	1.492 (0.129)	1.008 (0.245)	0.417 (0.102)	1.495 (0.092)	1.005 (0.171)
	0.8	0.851 (0.290)	1.496 (0.109)	1.022 (0.226)	0.823 (0.195)	1.497 (0.077)	1.013 (0.158)
	1.2	1.261 (0.425)	1.495 (0.109)	1.032 (0.227)	1.228 (0.290)	1.496 (0.077)	1.018 (0.158)
	1.8	1.870 (0.631)	1.493 (0.117)	1.045 (0.239)	1.832 (0.437)	1.495 (0.082)	1.025 (0.165)
	2.5	2.573 (0.879)	1.489 (0.130)	1.058 (0.258)	2.533 (0.614)	1.493 (0.090)	1.032 (0.176)
	5.0	5.029 (1.818)	1.475 (0.170)	1.104 (0.333)	5.011 (1.292)	1.485 (0.119)	1.054 (0.217)

4 Influence diagnostic

The detection of atypical cases is an important step in the estimation procedure. The first technique developed to assess the individual impact of cases on the estimation process is based on case-deletion (see, e.g., Cook and Weisberg, 1982) in which the effects are studied of completely removing cases from the analysis. This is a global influence analysis, since the effect of the case is evaluated by dropping it from the data. The local influence method is recommended when the concern is related to investigate the model sensibility under some minor perturbations in the model (or data).

Note that in equation (3.1) the contributions $\ell_i(\boldsymbol{\theta})$ are equally weighted. A perturbed log-likelihood function, allowing different weight for the different observations, can be defined in the form $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^{n} \omega_i \ell_i(\boldsymbol{\theta})$, where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$ is a n-dimensional vector of perturbations, which corresponds to a vector of weights of the contributions from each case to the log-likelihood function. Also, let $\boldsymbol{\omega}_0 = (1, \ldots, 1)^\top$ be the vector of no perturbation such that $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta})$. This perturbation is intended to evaluate whether the contribution of the cases with different weights affects the MLE of $\boldsymbol{\theta}$. The influence of minor perturbations on $\widehat{\boldsymbol{\theta}}$ can be assessed by using the log-likelihood displacement $LD_{\boldsymbol{\omega}} = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}$, where $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denotes the maximum likelihood estimate under $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$. This is the most commonly used method to evaluate the influence of a small modification of the model. This method is briefly described as follows.

The idea for assessing local influence as advocated by Cook (1986) is essentially the analysis of the local behavior of $LD_{\boldsymbol{\omega}}$ around $\boldsymbol{\omega}_0$ by evaluating the curvature of the plot of $LD_{\boldsymbol{\omega}_0+a\mathbf{d}}$ against a, where $a \in \Re$ and $\mathbf{d}$ is a unit direction. One of the measures of particular interest is the direction $\mathbf{d}_{\max}$ corresponding to the largest curvature $C_{\mathbf{d}_{\max}}$. The index plot of $\mathbf{d}_{\max}$ may evidence those observations that have considerable influence on $LD_{\boldsymbol{\omega}}$ under minor perturbations. Cook (1986) showed that the normal curvature at the direction $\mathbf{d}$ is given by $C_{\mathbf{d}}(\boldsymbol{\theta}) = 2|\mathbf{d}^\top \boldsymbol{\Delta}^\top \ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} \boldsymbol{\Delta}\mathbf{d}|$, where $\boldsymbol{\Delta} = \partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\boldsymbol{\theta}\, \partial\boldsymbol{\omega}^\top$ and $-\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ is the observed information matrix, both $\boldsymbol{\Delta}$ and $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ are evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$. Hence, $C_{\mathbf{d}_{\max}}/2$ is the largest eigenvalue of $\mathbf{B} = -\boldsymbol{\Delta}^\top \ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} \boldsymbol{\Delta}$ and $\mathbf{d}_{\max}$ is the corresponding unit norm eigenvector. The index plot of $\mathbf{d}_{\max}$ for the matrix $\mathbf{B}$ may show how to obtain large changes in the estimate of $\boldsymbol{\theta}$.

For the MOEBS distribution, after some algebra, the matrix $\boldsymbol{\Delta}$ is given by

$$\boldsymbol{\Delta} = [\,\boldsymbol{\Delta}_\eta^\top \quad \boldsymbol{\Delta}_\alpha^\top \quad \boldsymbol{\Delta}_\beta^\top\,]^\top,$$

where

$$\boldsymbol{\Delta}_\eta = (\widehat{k}_{11}, \dots, \widehat{k}_{1n}), \qquad \boldsymbol{\Delta}_\alpha = (\widehat{k}_{21}, \dots, \widehat{k}_{2n}), \qquad \boldsymbol{\Delta}_\beta = (\widehat{k}_{31}, \dots, \widehat{k}_{3n}),$$

with

$$k_{1i} = \frac{1}{\eta} - \frac{2\Phi(-v_i)}{1-\bar{\eta}\Phi(-v_i)},$$

$$k_{2i} = -\frac{1}{\alpha}\left(1+\frac{2}{\alpha^2}\right) + \frac{1}{\alpha^3}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right) + \frac{2\bar{\eta}}{\alpha}\,\frac{v_i\phi(v_i)}{[1-\bar{\eta}\Phi(-v_i)]},$$

$$k_{3i} = -\frac{1}{2\beta} + \frac{1}{t_i+\beta} + \frac{1}{2\alpha^2\beta}\left(\frac{t_i}{\beta}-\frac{\beta}{t_i}\right) + \frac{\bar{\eta}}{\alpha\beta}\,\frac{\tau(\sqrt{t_i/\beta})\phi(v_i)}{[1-\bar{\eta}\Phi(-v_i)]}$$

for $i = 1, \dots, n$. Here, the hat indicates evaluation at $\widehat{\boldsymbol{\theta}} = (\widehat{\eta}, \widehat{\alpha}, \widehat{\beta})^\top$.

5 Applications

Here, we present three empirical applications to demonstrate the flexibility and applicability of the MOEBS distribution. We compare the results of the fits of the MOEBS and BS distributions. All the computations were done using the Ox matrix programming language (Doornik, 2006).

First, we shall consider an uncensored data set from Nichols and Padgett (2006) on breaking stress of carbon fibres (in Gba). Table 3 lists the MLEs (and the corresponding standard errors in parentheses) of the model parameters and the following statistics: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and HQIC (Hannan–Quinn Information Criterion). These results show

Table 3 *MLEs (standard errors in parentheses) and AIC, BIC and HQIC; first data set*

	Estimates			Statistic		
Distribution	η	α	β	AIC	BIC	HQIC
MOEBS	22.3238 (24.0611)	0.5792 (0.1295)	0.9879 (0.4275)	288.58	296.40	291.75
BS		0.4622 (0.0327)	2.3660 (0.1064)	304.12	309.33	306.23

Table 4 *MLEs (standard errors in parentheses) and AIC, BIC and HQIC; second data set*

	Estimates			Statistic		
Distribution	η	α	β	AIC	BIC	HQIC
MOEBS	2.7338 (0.3522)	1.2750 (0.0391)	3.5815 (0.2977)	5391.10	5405.15	5396.50
BS		1.2347 (0.0309)	6.0625 (0.2195)	5451.86	5461.22	5455.45

that the MOEBS distribution has the lowest AIC, BIC and HQIC values, and so it could be chosen as the best model. The LR statistic for testing the null hypothesis $\mathcal{H}_0 : \eta = 1$ (BS) against $\mathcal{H}_1 : \eta \neq 1$ (MOEBS) is 17.5379 (p-value < 0.001). Thus, we reject the null hypothesis in favor of the MOEBS distribution at any usual significance level, that is, the MOEBS model is significantly better than the BS model based on the LR statistic.

Next, as a second application, we consider the data set corresponding to a record of 799 intervals between pulses along a nerve fibre presented in Cox and Lewis (1966) and reported in Jørgensen (1982). Table 4 lists the MLEs (standard errors in parentheses) of the parameters and the statistics AIC, BIC and HQIC. Based on these statistics, the MOEBS model should be preferred to the BS model. The LR statistic for testing the hypothesis $\mathcal{H}_0$: BS against $\mathcal{H}_1$: MOEBS is 62.757 (p-value < 0.001) and, hence, we strongly reject the null hypothesis in favor of the MOEBS distribution at any usual significance level. Therefore, according to the LR statistic, the MOEBS model is significantly better than the BS model.

The third real data set corresponds to daily ozone concentrations in New York during May–September, 1973. They were taken from Nadarajah (2008), provided by the New York State Department of Conservation. The MLEs (standard errors in parentheses) of the parameters and the statistics AIC, BIC and HQIC are given in Table 5. Based on the AIC, BIC and HQIC, the MOEBS model should be preferred to the BS model. The null hypothesis $\mathcal{H}_0$: BS against $\mathcal{H}_1$: MOEBS is rejected, since the value of the LR statistic is 10.7089 (p-value < 0.001). Thus, the MOEBS model is significantly better than the MOEBS model according to this statistic.

Table 5 *MLEs (standard errors in parentheses) and AIC, BIC and HQIC; third data set*

	Estimates			Statistic		
Distribution	η	α	β	AIC	BIC	HQIC
MOEBS	3.8289 (1.6021)	1.0646 (0.1031)	14.5053 (3.6853)	1093.49	1101.75	1096.84
BS		0.9823 (0.0645)	28.0234 (2.2644)	1102.19	1107.70	1104.43

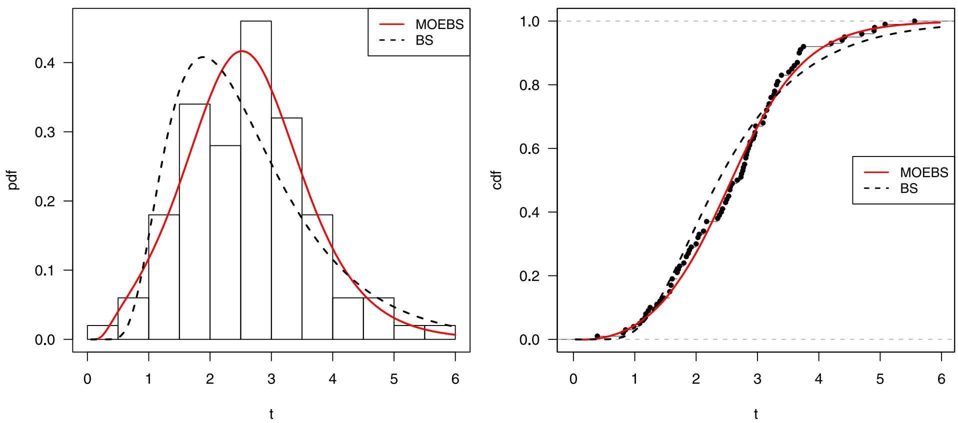

Figure 3 *Estimated p.d.f. and c.d.f. of the MOEBS and BS distributions; first data set.*

Plots of the estimated p.d.f. and c.d.f. of the MOEBS and BS models are given in Figures 3, 4 and 5 for the first, second and third data set, respectively. From these figures, the MOEBS model provides a better fit than the BS model in all the cases. QQ-plots are presented in Figures 6, 7 and 8 for the first, second and third real data set, respectively. Note that the MOEBS model outperforms the BS model in all the cases.

Figure 9 gives the influence index plot for the MOEBS model for the first, second and third real data sets. An inspection of these plots [Figures 9(a) and 9(b)] reveal that no observation appears with outstanding influence on the MLEs of the model parameters in the first and second real data sets. However, Figure 9(c) indicates that the case #18 has more pronounced influence on the MLEs than the other observations. It corresponds to the smallest observation. The relative change (RC), in percentage, of each parameter estimate is used to evaluate the effect of the potentially influential case. The RC is defined by $\mathrm{RC}_\theta(i) = |(\widehat{\theta} - \widehat{\theta}_{(i)}/\widehat{\theta}| \times 100\%$, where $\widehat{\theta}_{(i)}$ denotes the MLE of θ after removing the ith observation. We have $\mathrm{RC}_\eta(18) = 64.36\%$, $\mathrm{RC}_\alpha(18) = 19.34\%$ and $\mathrm{RC}_\beta(18) = 83.39\%$. Note the potential influence of this observation, mainly on the MLE of the parameters η and β.

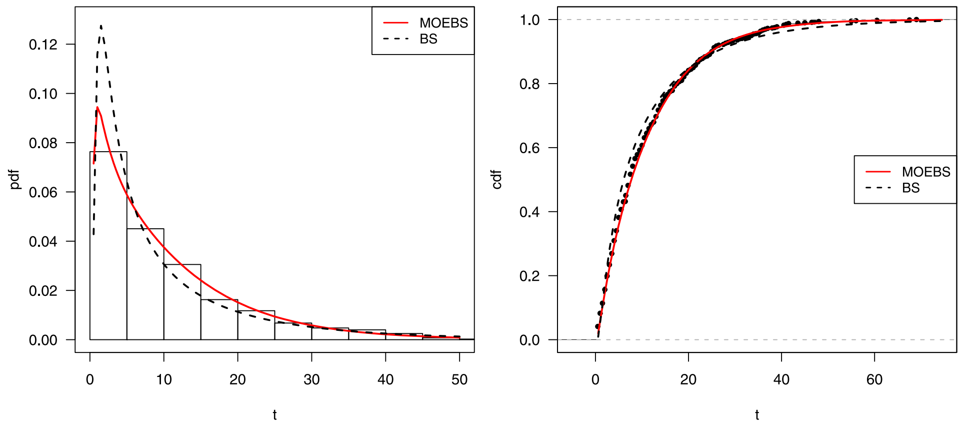

Figure 4 *Estimated p.d.f. and c.d.f. of the MOEBS and BS distributions*; *second data set.*

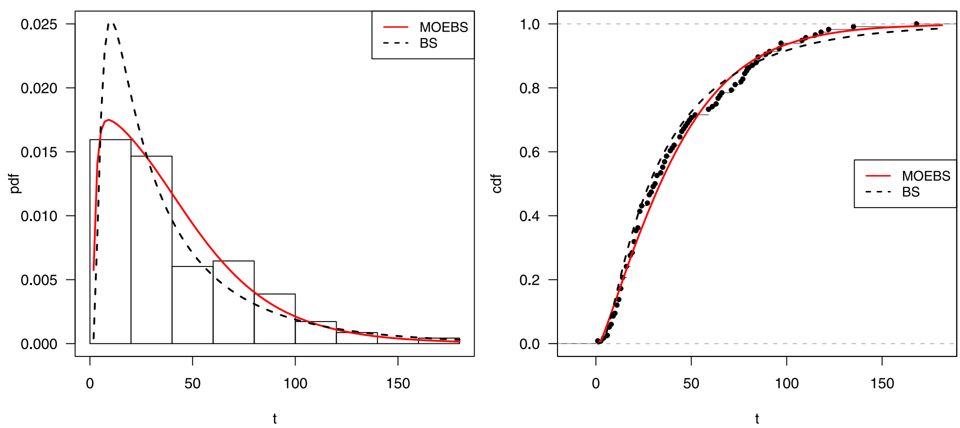

Figure 5 *Estimated p.d.f. and c.d.f. of the MOEBS and BS distributions*; *third data set.*

We fit the MOEBS and BS distributions when the potentially influential case #18 is not considered in the analysis. In this case, the classical BS distribution presents a better fit than the MOEBS model. For example, we have AIC = 1071.44, BIC = 1079.67 and HQIC = 1074.78 for the MOEBS model, whereas AIC = 1069.96, BIC = 1075.45 and HQIC = 1072.19 for the BS model.

6 Concluding remarks

Marshall and Olkin (1997) proposed a simple transformation of a distribution function by inserting an additional parameter $\eta > 0$ in order to obtain a larger class of distribution functions which contains the original one, in case of $\eta = 1$. Based on this approach, a new three parameter Birnbaum–Saunders distribution

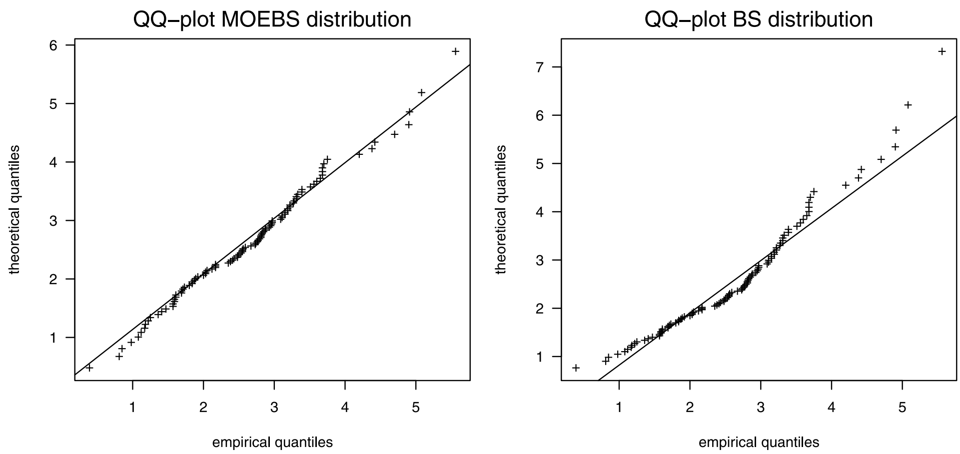

Figure 6 *QQ-plots: first data set.*

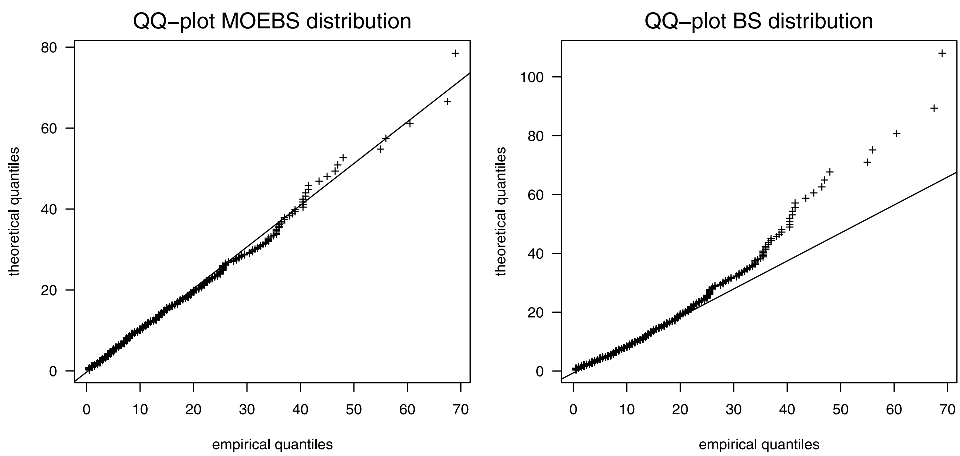

Figure 7 *QQ-plots: second data set.*

was developed. The probability density function, cumulative distribution function, hazard function, moments and some plots were provided. Some properties of this new model were obtained. The estimation of the parameters is approached by the method of maximum likelihood and the observed information matrix is derived. Moreover, it can be depicted from the Monte Carlo simulation study that the maximum likelihood estimators of the model parameters of the MOEBS distribution present a quite nice behavior in terms of empirical means and mean squared error. We also discussed influence diagnostic for the new model. Applications of the MOEBS distribution to three real data sets are given to show that the new distribution provides consistently better fits than the BS distribution. It also illustrates the fact that there is still room for improving the BS model. We hope that this

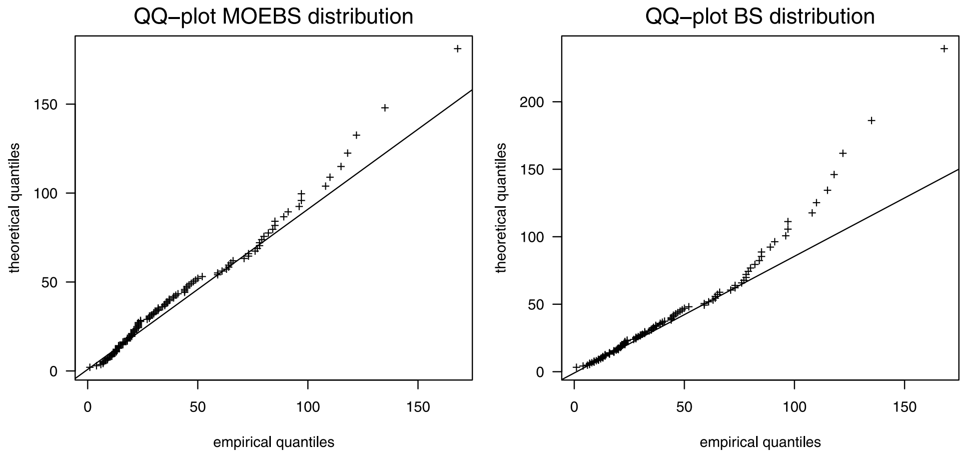

Figure 8 *QQ-plots: third data set.*

generalization may attract wider applications in the literature of the fatigue life distributions and survival analysis.

Appendix

After algebraic manipulations, the elements of the matrix $\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ are given by

$$\ddot{L}_{\eta\eta} = -\frac{n}{\eta^2} + 2\sum_{i=1}^{n}\left(\frac{\Phi(-v_i)}{1-\bar{\eta}\Phi(-v_i)}\right)^2, \qquad \ddot{L}_{\eta\alpha} = -\frac{2}{\alpha}\sum_{i=1}^{n}\frac{v_i\phi(v_i)}{[1-\bar{\eta}\Phi(-v_i)]^2},$$

$$\ddot{L}_{\eta\beta} = -\frac{1}{\alpha\beta}\sum_{i=1}^{n}\frac{\tau(\sqrt{t_i/\beta})\phi(v_i)}{[1-\bar{\eta}\Phi(-v_i)]^2},$$

$$\ddot{L}_{\alpha\alpha} = \frac{n}{\alpha^2} + \frac{6n}{\alpha^4} - \frac{3}{\alpha^4}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i}\right) - \frac{2\bar{\eta}}{\alpha^2}\sum_{i=1}^{n}\frac{v_i\phi(v_i)}{1-\bar{\eta}\Phi(-v_i)}\left\{2 - v_i^2 - \frac{\bar{\eta}v_i\phi(v_i)}{1-\bar{\eta}\Phi(-v_i)}\right\},$$

$$\ddot{L}_{\alpha\beta} = -\frac{1}{\alpha^3\beta}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} - \frac{\beta}{t_i}\right) - \frac{\bar{\eta}}{\alpha^2\beta}\sum_{i=1}^{n}\frac{\tau(\sqrt{t_i/\beta})\phi(v_i)}{1-\bar{\eta}\Phi(-v_i)}\left\{1 - v_i^2 - \frac{\bar{\eta}v_i\phi(v_i)}{1-\bar{\eta}\Phi(-v_i)}\right\},$$

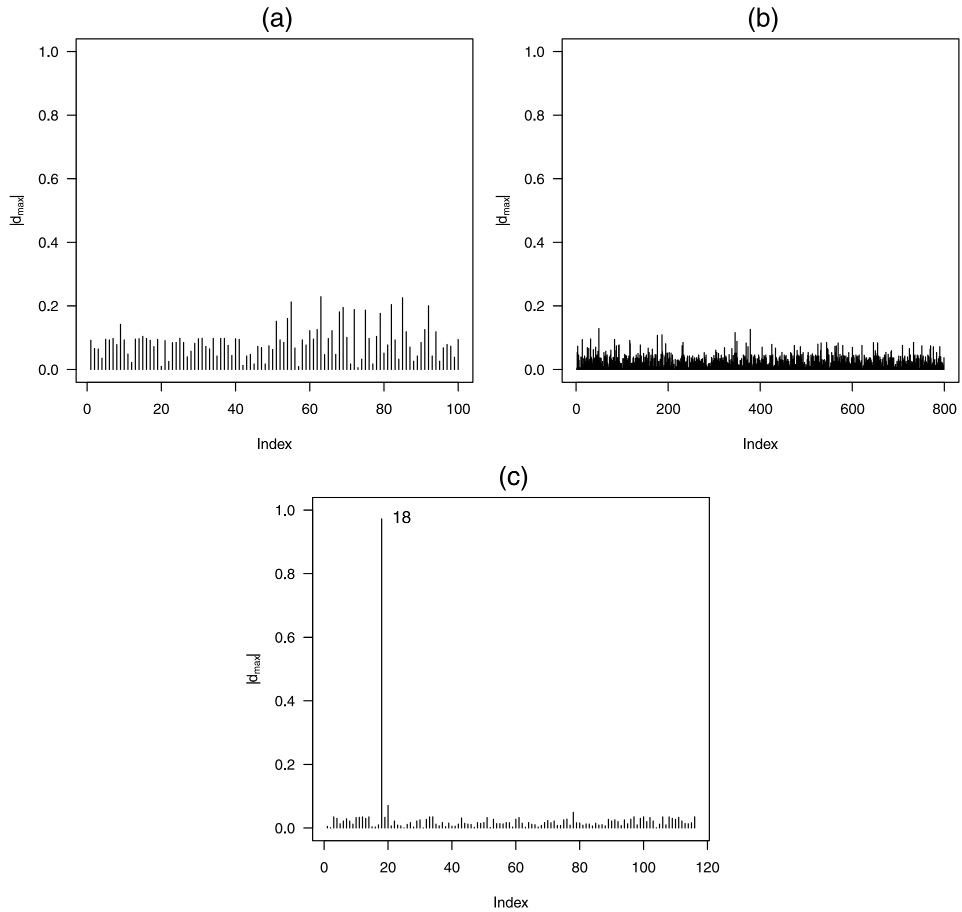

Figure 9 *Influence index plots for the MOEBS model for the first* (a), *second* (b) *and third* (c) *real data sets*.

$$\ddot{L}_{\beta\beta} = \frac{n}{2\beta^2} - \sum_{i=1}^{n} \frac{1}{(t_i+\beta)^2} - \frac{1}{\alpha^2\beta^3}\sum_{i=1}^{n} t_i - \frac{\bar{\eta}}{2\alpha\beta^2}\sum_{i=1}^{n}\frac{\phi(v_i)}{1-\bar{\eta}\Phi(-v_i)}\left\{3\sqrt{\frac{t}{\beta}} + \sqrt{\frac{\beta}{t}} - \frac{\tau(\sqrt{t_i/\beta})^2}{\alpha}\left[v_i + \frac{\bar{\eta}\phi(v_i)}{1-\bar{\eta}\Phi(-v_i)}\right]\right\}.$$

Acknowledgments

The author thanks two anonymous referees for helpful comments. The financial support from FAPESP (Brazil) is gratefully acknowledged.

References

Balakrishnan, N., Leiva, V. and López, J. (2007). Acceptance sampling plans from truncated life tests from generalized Birnbaum–Saunders distribution. *Communications in Statistics—Simulation and Computation* **36**, 643–656. MR2370927

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. (1969a). A new family of life distributions. *Journal of Applied Probability* **6**, 319–327. MR0253493

Birnbaum, Z. W. and Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* **6**, 328–347. MR0251807

Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society, Ser. B* **48**, 133–169. MR0867994

Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression*. London: Chapman & Hall. MR0675263

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

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*. London: Chapman & Hall. MR0370837

Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events*. London: Methuem. MR0199942

Desmond, A. F. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics* **13**, 171–183. MR0818323

Desmond, A. F. (1986). On the relationship between two fatigue-life models. *IEEE Transactions on Reliability* **35**, 167–169.

Díaz-García, J. A. and Leiva, V. (2005). A new family of life distributions based on the elliptically contoured distributions. *Journal of Statistical Planning and Inference* **128**, 445–457. MR2102769

Doornik, J. A. (2006). *An Object-Oriented Matrix Language—Ox 4*, 5th ed. London: Timberlake Consultants Press.

García, V. J., Gómez-Déniz, E. and Vázquez-Polo, F. J. (2010). A new skew generalization of the normal distribution: Properties and applications. *Computational Statistics and Data Analysis* **54**, 2021–2034. MR2640305

Garvan, F. (2002). *The Maple Book*. London: Chapman & Hall/CRC.

Gómes, H. W., Olivares-Pacheco, J. F. and Bolfarine, H. (2009). An extension of the generalized Birnbaum–Saunders distribution. *Statistics and Probability Letters* **79**, 331–338. MR2493016

Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. *Test* **19**, 399–415. MR2677735

Guiraud, P., Leiva, V. and Fierro, R. (2009). A non-central version of the Birnbaum–Saunders distribution for reliability analysis. *IEEE Transactions on Reliability* **58**, 152–160.

Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution*. New York: Springer-Verlag. MR0648107

Kundu, D., Kannan, N. and Balakrishnan, N. (2008). On the hazard function of Birnbaum–Saunders distribution and associated inference. *Computational Statistics and Data Analysis* **52**, 2692–2702. MR2419535

Leiva, V., Barros, M., Paula, G. A. and Sanhueza, A. (2008). Generalized Birnbaum–Saunders distributions applied to air pollutant concentration. *Environmetrics* **19**, 235–249. MR2420468

Leiva, V., Sanhueza, A. and Angulo, J. M. (2009). A length-biased version of the Birnbaum–Saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment* **23**, 299–307. MR2476075

Leiva, V., Vilca, F., Balakrishnan, N. and Sanhueza, A. (2010). A skewed sinh-normal distribution and its properties and application to air pollution. *Communications in Statistics—Theory and Methods* **39**, 426–443. MR2745286

Lemonte, A. J., Cribari-Neto, F. and Vasconcellos, K. L. P. (2007). Improved statistical inference for the two-parameter Birnbaum–Saunders distribution. *Computational Statistics and Data Analysis* **51**, 4656–4681. MR2364472

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

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **84**, 641–652. MR1603936

Nadarajah, S. (2008). A truncated inverted beta distribution with application to air pollution data. *Stochastic Environmental Research and Risk Assessment* **22**, 285–289. MR2412572

Nichols, M. D. and Padgett, W. J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International* **22**, 141–151.

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. New York: Springer. MR1713114

Owen, W. J. (2006). A new three-parameter extension to the Birnbaum–Saunders distribution. *IEEE Transactions on Reliability* **55**, 475–479.

Press, W. H., Teulosky, S. A., Vetterling, W. T. and Flannery, B. P. (2007). *Numerical Recipes in C: The Art of Scientific Computing*, 3rd ed. Cambridge: Cambridge Univ. Press. MR2371990

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sarhan, A. M. and Balakrishnan, N. (2007). A new class of bivariate distributions and its mixture. *Journal of Multivariate Analysis* **98**, 1508–1527. MR2364132

Sigmon, K. and Davis, T. A. (2002). *MATLAB Primer*, 6th ed. London: Chapman & Hall/CRC.

Vilca, F. and Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics—Theory and Methods* **35**, 229–244. MR2274046

Wolfram, S. (2003). *The Mathematica Book*, 5th ed. New York: Cambridge Univ. Press.

Wu, J. and Wong, A. C. M. (2004). Improved interval estimation for the two-parameter Birnbaum–Saunders distribution. *Computational Statistics and Data Analysis* **47**, 809–821. MR2101553

Xu, A. and Tang, Y. (2010). Reference analysis for Birnbaum–Saunders distribution. *Computational Statistics and Data Analysis* **54**, 185–192. MR2558469

Departamento de Estatística
Universidade de São Paulo
São Paulo/SP, 05508-090
Brazil
E-mail: arturlemonte@gmail.com