

Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution

ARTUR J. LEMONTE, ALEXANDRE B. SIMAS and FRANCISCO CRIBARI-NETO*

Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, Recife/PE, 50740-540, Brazil

(Received 00 Month 200x; In final form 00 Month 200x)

In this paper, we consider the two-parameter Birnbaum-Saunders distribution proposed by Birnbaum and Saunders [Birnbaum, Z.W. and Saunders, S.C., 1969a, A new family of life distributions. *Journal of Applied Probability*, **6**, 319–327], which is commonly used for modeling the lifetime of materials and equipments. We consider different strategies of bias correction of the maximum likelihood estimators for the parameters that index the distribution via bootstrap (parametric and nonparametric). The numerical evidence favors a particular bootstrap estimator based on parametric resampling. Finally, an example with real data is presented and discussed.

Keywords: Bootstrap; Bias correction; Birnbaum-Saunders distribution; Maximum likelihood estimation.

1 Introduction

Birnbaum and Saunders [1] proposed a new family of life distributions for modeling the lifetime of materials and equipments under the influence of dynamic loads. This two-parameter Birnbaum-Saunders distribution was derived from a model in which failures happen due to the development and growth of a dominant crack. Later, this distribution was derived in a more general way by Desmond [3], based on a biological model. Desmond also extended the physical justification for its use by relaxing some of the assumptions made in Birnbaum and Saunders [1]. Moreover, Desmond [4] investigated the relationship between the Birnbaum-Saunders distribution and the inverse Gaussian distribution.

A random variable T is said to follow the two-parameter Birnbaum-Saunders distribution with parameters $\alpha, \beta > 0$, denoted by $\mathcal{B}\text{-}\mathcal{S}(\alpha, \beta)$, if its cumulative

*Corresponding author. Email: cribari@de.ufpe.br

distribution function is given by

$$F_T(t) = P(T \leq t) = \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right], \quad t > 0, \quad (1)$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function; α is a shape parameter and β is a scale parameter, i.e., $T/\beta \sim \mathcal{B}\text{-}\mathcal{S}(\alpha, 1)$. β is also the median of the distribution, since $F_T(\beta) = \Phi(0) = 1/2$. Additionally, for any $k > 0$ it follows that $kT \sim \mathcal{B}\text{-}\mathcal{S}(\alpha, k\beta)$. The distribution $\mathcal{B}\text{-}\mathcal{S}(\alpha, \beta)$ has the well-known *reciprocal property*, that is, $T^{-1} \sim \mathcal{B}\text{-}\mathcal{S}(\alpha, \beta^{-1})$, which belongs to the same family of distributions (for more details, see [5]). Furthermore, Mann et al. [6, p. 155] showed that the distribution (1) is unimodal. Some recent papers on the Birnbaum-Saunders distribution are Chang and Tang [7,8], Dupuis and Mills [9], and Rieck [10,11]. For a detailed account of this distribution and its main properties, see Johnson, Kotz and Balakrishnan [12].

Maximum likelihood estimation of the parameters that index the Birnbaum-Saunders distribution was initially considered by Birnbaum and Saunders [13]; the asymptotic distribution of the maximum likelihood estimators (MLEs) was derived by Engelhardt et al. [14]. The authors showed that the MLEs of α and β are asymptotically independent. Therefore, confidence intervals for these parameters can be constructed using their asymptotic joint distribution.

Recently, Monte Carlo simulation results presented by Ng, Kundu and Balakrishnan [15] revealed that the MLEs of α and β are considerably biased. They used a simple method of bias correction to reduce the bias of the MLEs, and recommended the use of the adjusted estimators.

In this paper, we consider different schemes of bias correction in finite samples using parametric and nonparametric bootstrap resampling. Bootstrap is a computationally intensive method, first introduced by Bradley Efron [16], which can be used to achieve asymptotically refined inference without the need for tedious algebra. Ferrari and Cribari-Neto [17] explored the relationship between analytical bias correction, based on Edgeworth expansions, and bootstrap bias adjustment.

The paper unfolds as follows: in Section 2, we present the two-parameter Birnbaum-Saunders distribution; in Section 3, we present bootstrap methods for point estimation; Section 4 discusses the numerical results; Section 5 contains an empirical application; and, finally, in Section 6, we outline some concluding remarks.

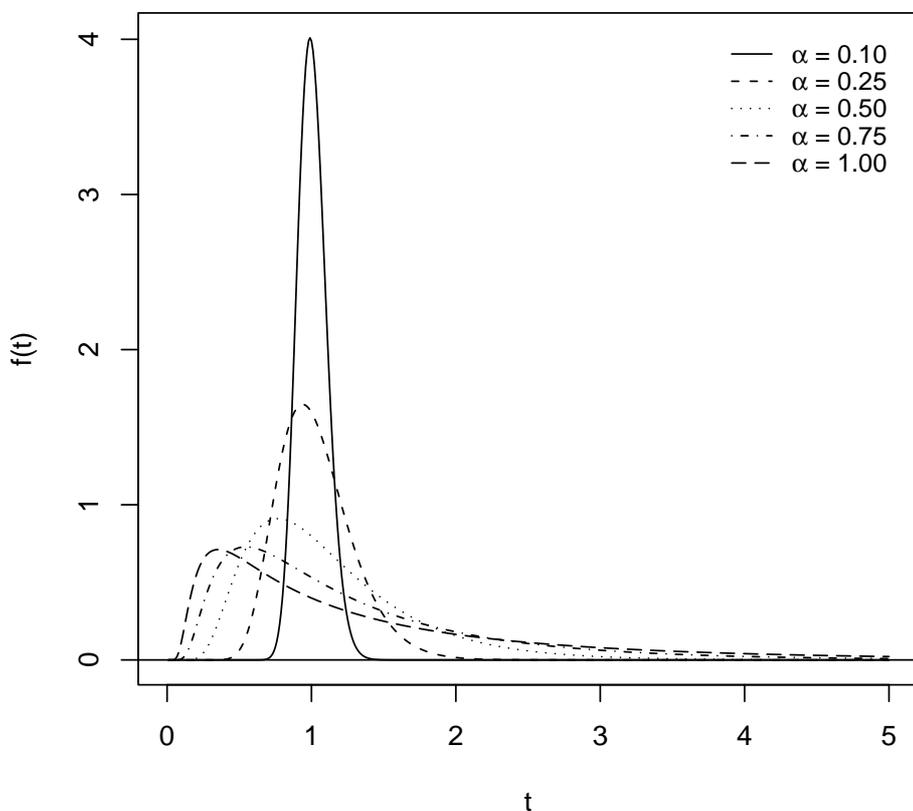


Figure 1. Two-parameter Birnbaum-Saunders density plots ($\beta = 1$ and different values of α).

2 The Birnbaum-Saunders distribution

The Birnbaum-Saunders density function is given by

$$f_T(t; \alpha, \beta) = \frac{1}{2\alpha\beta\sqrt{2\pi}} \left[\left(\frac{\beta}{t}\right)^{1/2} + \left(\frac{\beta}{t}\right)^{3/2} \right] \exp \left\{ -\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right\}, \quad (2)$$

where $t > 0$ and $\alpha, \beta > 0$. The expected value, variance, skewness and kurtosis are given, respectively, by

$$E(T) = \beta \left(1 + \frac{1}{2}\alpha^2 \right), \quad \text{Var}(T) = (\alpha\beta)^2 \left(1 + \frac{5}{4}\alpha^2 \right),$$

$$\mu_3 = \frac{16\alpha^2(11\alpha^2 + 6)}{(5\alpha^2 + 4)^3} \quad \text{and} \quad \mu_4 = 3 + \frac{6\alpha^2(93\alpha^2 + 41)}{(5\alpha^2 + 4)^2}.$$

As noted before, if $T \sim \mathcal{B}\text{-}\mathcal{S}(\alpha, \beta)$, then $T^{-1} \sim \mathcal{B}\text{-}\mathcal{S}(\alpha, \beta^{-1})$. Therefore,

$$E(T^{-1}) = \beta^{-1} \left(1 + \frac{1}{2} \alpha^2 \right) \quad \text{and} \quad \text{Var}(T^{-1}) = \alpha^2 \beta^{-2} \left(1 + \frac{5}{4} \alpha^2 \right).$$

The density (2) is asymmetric to the right for large values of α , the asymmetry decreasing with α . In Figure 1, we present the density (2) for different values of α , considering $\beta = 1$. Note that, as α decreases, the density becomes more symmetric around β , which is the median of the distribution. Note also that the variance also decreases with α .

3 Bootstrap methods

In what follows, we shall present different bootstrapping schemes that can be used to reduce the bias of the MLE. Let $\mathbf{y} = (y_1, \dots, y_n)^\top$ be a random sample of size n , where each element is a random draw from the random variable Y which has the distribution function $F = F_\theta(y)$. Here, θ is the parameter that indexes the distribution, and is viewed as a functional of F , i.e., $\theta = t(F)$. Finally, let $\hat{\theta}$ be an estimator of θ based on \mathbf{y} ; we write $\hat{\theta} = s(\mathbf{y})$.

The application of the bootstrap method proposed by Efron [16] consists in obtaining, from the original sample \mathbf{y} , a large number of pseudo-samples $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$, and then extracting information from these samples to improve inference. Bootstrap methods can be classified into two classes, depending on how the sampling is performed: parametric and nonparametric. In the parametric version, the bootstrap samples are obtained from $F(\hat{\theta})$, which we shall denote as $F_{\hat{\theta}}$, whereas in the nonparametric version they are obtained from the empirical distribution function \hat{F} , through sampling with replacement. Note that the nonparametric bootstrap does not entail parametric assumptions.

Let $B_F(\hat{\theta}, \theta)$ be the bias of the estimator $\hat{\theta} = s(\mathbf{y})$, that is,

$$B_F(\hat{\theta}, \theta) = E_F[\hat{\theta} - \theta] = E_F[s(\mathbf{y})] - t(F),$$

where the subscript F indicates that expectation is taken with respect to F . The bootstrap estimators of the bias in the parametric and nonparametric versions are obtained by replacing the true distribution F , which generated the original sample, with $F_{\hat{\theta}}$ and \hat{F} , respectively, in the above expression.

Therefore, the parametric and nonparametric estimates of the bias are given, respectively, by

$$B_{F_{\hat{\theta}}}(\hat{\theta}, \theta) = E_{F_{\hat{\theta}}}[s(\mathbf{y})] - t(F_{\hat{\theta}}) \quad \text{and} \quad B_{\hat{F}}(\hat{\theta}, \theta) = E_{\hat{F}}[s(\mathbf{y})] - t(\hat{F}).$$

If B bootstrap samples $(\mathbf{y}^{*1}, \mathbf{y}^{*2}, \dots, \mathbf{y}^{*B})$ are generated independently from the original sample \mathbf{y} , and the respective bootstrap replications $(\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*B})$ are calculated, where $\hat{\theta}^{*b} = s(\mathbf{y}^{*b})$, $b = 1, 2, \dots, B$, then it is possible to approximate the bootstrap expectations $E_{F_{\hat{\theta}}}[s(\mathbf{y})]$ and $E_{\hat{F}}[s(\mathbf{y})]$ by the mean $\hat{\theta}^{*(\cdot)} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}^{*b}$. Therefore, the bootstrap bias estimates based on B replications of $\hat{\theta}$ are

$$\hat{B}_{F_{\hat{\theta}}}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - s(\mathbf{y}) \quad \text{and} \quad \hat{B}_{\hat{F}}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - s(\mathbf{y}),$$

for the parametric and nonparametric versions, respectively.

An alternative bootstrap bias estimate was proposed by Efron [18]. His approach is nonparametric and uses an auxiliary vector known as the resampling vector, which records the proportions of the original observations $\mathbf{y} = (y_1, \dots, y_n)^\top$ included in the bootstrap sample. We denote the resampling vector by $\mathbf{P}^* = (P_1^*, P_2^*, \dots, P_n^*)$, where its components P_j^* , $j = 1, 2, \dots, n$, are defined with respect to a given bootstrap sample $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$ as $P_j^* = n^{-1}(\#\{y_k^* = y_j\})$, $j = 1, 2, \dots, n$. The vector $\mathbf{P}^0 = (1/n, 1/n, \dots, 1/n)$ corresponds to the original sample.

Note that a bootstrap replication $\hat{\theta}^*$ can be defined as a function of the resampling vector. For instance, if $\hat{\theta} = s(\mathbf{y}) = \bar{y}$, then

$$\begin{aligned} \hat{\theta}^* &= \frac{y_1^* + y_2^* + \dots + y_n^*}{n} = \frac{\#\{y_k^* = y_1\}y_1 + \dots + \#\{y_k^* = y_n\}y_n}{n} \\ &= \frac{(nP_1^*)y_1 + \dots + (nP_n^*)y_n}{n} = \mathbf{P}^* \mathbf{y}. \end{aligned}$$

Suppose we can write the estimate of interest, obtained from \mathbf{y} , as $G(\mathbf{P}^0)$. We can then obtain bootstrap estimates $\hat{\theta}^{*b}$ using the resampling vectors \mathbf{P}^{*b} , $b = 1, 2, \dots, B$, as $G(\mathbf{P}^{*b})$. The new bootstrap bias estimate, $\bar{B}_{\hat{F}}(\hat{\theta}, \theta)$, is defined as (Efron, [18])

$$\bar{B}_{\hat{F}}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - G(\mathbf{P}^{*(\cdot)}), \quad \text{where} \quad \mathbf{P}^{*(\cdot)} = \frac{1}{B} \sum_{b=1}^B \mathbf{P}^{*b},$$

which differs from $\hat{B}_{\hat{F}}(\hat{\theta}, \theta)$, since $\hat{B}_{\hat{F}}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - G(\mathbf{P}^0)$.

By using the three bootstrap bias estimates presented above, we arrive at the following three bias-corrected estimators:

$$\begin{aligned}\tilde{\theta}_1 &= s(\mathbf{y}) - \widehat{B}_{\widehat{F}}(\widehat{\theta}, \theta) = 2\widehat{\theta} - \widehat{\theta}^{*(\cdot)}, \\ \tilde{\theta}_2 &= s(\mathbf{y}) - \bar{B}_{\widehat{F}}(\widehat{\theta}, \theta) = \widehat{\theta} - \widehat{\theta}^{*(\cdot)} + G(\mathbf{P}^{*(\cdot)}), \\ \tilde{\theta}_3 &= s(\mathbf{y}) - \widehat{B}_{F_{\widehat{\theta}}}(\widehat{\theta}, \theta) = 2\widehat{\theta} - \widehat{\theta}^{*(\cdot)}.\end{aligned}$$

The corrected estimates $\tilde{\theta}_1$ and $\tilde{\theta}_3$ were called constant-bias-correcting (CBC) estimates by MacKinnon and Smith [19].

As proposed by Efron [18], the bias corrected estimator $\tilde{\theta}_2$ requires the original estimator $\widehat{\theta}$ to have closed form. However, $\widehat{\theta}$ does not have closed form in general. To circumvent this problem, Cribari-Neto, Frery and Silva [20] proposed an adaptation of Efron's method, which can be applied to estimators that do not have closed form. Their proposal is to use the resampling vector to modify the log-likelihood function, and then maximize the modified log-likelihood. The idea is to write the log-likelihood function in terms of \mathbf{P}^0 , replace this quantity by $\mathbf{P}^{*(\cdot)}$, and then maximize the resulting modified log-likelihood function.

This adaptation of Efron's method is useful for performing inference in the two-parameter Birnbaum-Saunders distribution since the MLEs of the parameters do not have closed form. Let $\mathbf{t} = (t_1, \dots, t_n)^\top$ be a random sample of size n of the Birnbaum-Saunders distribution with parameters α and β . The log-likelihood function can be written in terms of \mathbf{P}^0 , apart from an unimportant constant, as

$$\ell(\alpha, \beta; \mathbf{t}) = -n \left(\log(\alpha\beta) - \frac{1}{\alpha^2} \right) + n\mathbf{P}^0 \mathbf{A}^\top - \frac{n}{2\alpha^2\beta} \mathbf{P}^0 \mathbf{B}^\top - \frac{n\beta}{2\alpha^2} \mathbf{P}^0 \mathbf{C}^\top,$$

where $\mathbf{A} = (\log[(\beta/t_1)^{1/2} + (\beta/t_1)^{3/2}], \dots, \log[(\beta/t_n)^{1/2} + (\beta/t_n)^{3/2}])$, $\mathbf{B} = (t_1, \dots, t_n)$, and $\mathbf{C} = (1/t_1, \dots, 1/t_n)$. One then replaces \mathbf{P}^0 by $\mathbf{P}^{*(\cdot)}$ after obtaining $\mathbf{P}^{*(\cdot)}$ from a nonparametric bootstrapping scheme based on B replications, and maximizes the modified log-likelihood function. That is, we maximize

$$\ell(\alpha, \beta; \mathbf{t}) = -n \left(\log(\alpha\beta) - \frac{1}{\alpha^2} \right) + n\mathbf{P}^{*(\cdot)} \mathbf{A}^\top - \frac{n}{2\alpha^2\beta} \mathbf{P}^{*(\cdot)} \mathbf{B}^\top - \frac{n\beta}{2\alpha^2} \mathbf{P}^{*(\cdot)} \mathbf{C}^\top,$$

instead of maximizing the original log-likelihood function.

MacKinnon and Smith [19] argue that the estimators $\tilde{\theta}_1$ and $\tilde{\theta}_3$, which they call CBC, are designed to work well when the bias function $B(\theta)$ is flat, that

is, when it does not depend on θ . They considered the situation when the bias function is linear with respect to θ , that is,

$$B(\theta) = a + c\theta.$$

The estimation of the bias function now involves the estimation of two parameters, namely a and c . We need to obtain estimates of two points of the bias line, and then use such points to obtain our estimates of a and c .

The procedure can be summarized as follows. Using the original sample \mathbf{y} , we compute the estimate $\hat{\theta} = s(\mathbf{y})$. In order to estimate the first point, we use a parametric bootstrapping scheme to obtain a bootstrap estimate of the bias of $\hat{\theta}$, denoted as \hat{B} and given by $\hat{\theta}^{*(\cdot)} - \hat{\theta}$. Next, in order to estimate the second point, we use another parametric bootstrapping scheme based on $\tilde{\theta}$, where $\tilde{\theta} = 2\hat{\theta} - \hat{\theta}^{*(\cdot)}$, and, for each bootstrap sample, we obtain the corresponding replications $\hat{\theta}_{F_{\tilde{\theta}}}^{*b}$, $b = 1, \dots, B$. Therefore, we estimate the bias of $\tilde{\theta}$ by the quantity \tilde{B} , where $\tilde{B} = \hat{\theta}_{F_{\tilde{\theta}}}^{*(\cdot)} - \tilde{\theta}$, $\hat{\theta}_{F_{\tilde{\theta}}}^{*(\cdot)}$ being the mean of the bootstrap replications of $\tilde{\theta}$. Note that here one needs to perform $2B$ bootstrap replications (instead of B). In other words, the number of replications required is the double of that used in the case of CBC estimators. Finally, using the point estimates $\hat{\theta}$ and $\tilde{\theta}$, and their respective estimated biases, \hat{B} and \tilde{B} , we arrive at the following system of two simultaneous equations:

$$\begin{cases} \hat{B} = \check{a} + \check{c}\hat{\theta}, \\ \tilde{B} = \check{a} + \check{c}\tilde{\theta}, \end{cases}$$

whose solution is

$$\check{a} = \hat{B} - \frac{\hat{B} - \tilde{B}}{\hat{\theta} - \tilde{\theta}} \quad \text{and} \quad \check{c} = \frac{\hat{B} - \tilde{B}}{\hat{\theta} - \tilde{\theta}}.$$

We now obtain the following bias-corrected estimator, known as the linear-bias-correcting (LBC) estimator, MacKinnon and Smith [19], and denoted by $\tilde{\theta}_4$:

$$\tilde{\theta}_4 = \frac{1}{1 + \check{c}}(\hat{\theta} - \check{a}).$$

Its variance is a function of the variance of $\hat{\theta}$:

$$\text{Var}(\tilde{\theta}_4) = \frac{1}{(1 + \check{c})^2} \text{Var}(\hat{\theta}).$$

If the estimated value of c , \check{c} , belongs to the set $\mathcal{A} = \{(-2, 0) \setminus \{-1\}\}$, then the variance of $\tilde{\theta}_4$ will exceed that of $\hat{\theta}$.

A detailed discussion of bootstrap methods and their applications can be found in Davison and Hinkley [21] and in Efron and Tibshirani [22].

4 Numerical results

The simulation of $\mathcal{B}\text{-}\mathcal{S}(\alpha, \beta)$ independent deviates can be performed using the monotone mapping:

$$X = \frac{1}{2} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right).$$

That is, from (1) we have that $X \sim \mathcal{N}(0, \frac{1}{4}\alpha^2)$. Therefore, from the equation above, T can be written as

$$T = \beta \{1 + 2X^2 + 2X(1 + X^2)^{1/2}\}. \quad (3)$$

In other words, pseudo-random numbers from the two-parameter Birnbaum-Saunders distribution can be obtained from pseudo-random normal numbers using (3). The Monte Carlo simulations were performed using the object-oriented matrix programming language `0x`¹ (Cribari-Neto and Zarkos [23]; Doornik [24]). We consider $R = 5000$ (number of Monte Carlo replications) and $B = 500$ (number of bootstrap replications). The MLEs of α and β are obtained by maximizing the log-likelihood function using the BFGS quasi-Newton non-linear optimization algorithm with analytical derivatives. This method is generally considered the best non-linear optimization method (Mittelhammer, Judge and Miller [25, p. 199]). Note that the simulations are computationally intensive since each experiment requires millions of non-linear maximizations.

We have evaluated, through Monte Carlo simulations, the performance of the MLE $\hat{\theta} = (\hat{\alpha}, \hat{\beta})^\top$ of the vector of parameters $\theta = (\alpha, \beta)^\top$ of the Birnbaum-Saunders distribution and its corrected versions: $\tilde{\theta}_1 = (\tilde{\alpha}_1, \tilde{\beta}_1)^\top$ (nonparametric CBC), $\tilde{\theta}_2 = (\tilde{\alpha}_2, \tilde{\beta}_2)^\top$ (based on the resampling vector $\mathbf{P}^{*(\cdot)}$), $\tilde{\theta}_3 = (\tilde{\alpha}_3, \tilde{\beta}_3)^\top$ (parametric CBC), and $\tilde{\theta}_4 = (\tilde{\alpha}_4, \tilde{\beta}_4)^\top$ (parametric LBC), in finite samples. We also computed the corrected MLE proposed by Ng, Kundu and Balakrishnan [15], denoted in this paper as $\bar{\theta} = (\bar{\alpha}, \bar{\beta})^\top$. The sample sizes considered

¹`0x` is freely distributed for academic purposes and available at <http://www.doornik.com>.

were $n = 10, 20, 40,$ and $60,$ and the values for the shape parameter α were $\alpha = 0.10, 0.25, 0.50, 0.75,$ and $1.00.$ Without loss of generality, the scale parameter β was set at 1, that is, $\beta = 1$ in all experiments.

For each sample size, we computed the following quantities: relative bias estimates (the relative bias of an estimator $\hat{\theta}$ of a scalar parameter θ is defined as $\{E(\hat{\theta}) - \theta\}/\theta,$ and the estimated relative bias is obtained by estimating $E(\hat{\theta})$ via Monte Carlo) and the square root of the mean squared error ($\sqrt{\text{MSE}}$) based on the 5000 estimates. In Tables 1 to 4, 'NP' denotes nonparametric bootstrap, and 'P' denotes parametric bootstrap.

In Tables 1 and 2 we present the estimated relative biases. Note in Table 1 that the estimators $\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4$ and $\bar{\alpha}$ displayed relative biases (in absolute values) that are smaller than that of the usual MLE $\hat{\alpha}$ in all sample sizes considered. Note also that the bias correction proposed by Ng, Kundu and Balakrishnan [15] (estimator $\bar{\alpha}$) was the least effective amongst all bias corrections considered. For instance, for $n = 20$ and $\alpha = 0.10,$ the estimated relative bias of the estimator $\bar{\alpha}$ was 0.01238, whereas the estimated relative bias of the estimators $\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3,$ and $\tilde{\alpha}_4$ were, in absolute values, 0.00286, 0.00361, 0.00167, and 0.00430, respectively. Furthermore, note that, amongst the bootstrap estimators, the estimator $\tilde{\alpha}_3$ (parametric CBC) displayed the best finite-sample performance, followed by the estimator $\tilde{\alpha}_1$ (non-parametric CBC). For instance, for $n = 40,$ and $\alpha = 0.25,$ the relative biases of $\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3,$ and $\tilde{\alpha}_4$ were, in absolute values, 0.00029, 0.00105, 0.00003, and 0.00483, respectively.

The figures in Table 2 show that the estimators $\tilde{\beta}_1$ and $\tilde{\beta}_3$ were again the most accurate for all sample sizes. For instance, for $n = 10$ and $\alpha = 0.25,$ the estimated relative biases of the estimators $\tilde{\beta}_1$ and $\tilde{\beta}_3$ were nearly 10 times smaller than that of the MLE $\hat{\beta}.$ Of these two bootstrap estimators, the one with best performance was $\tilde{\beta}_3$ (parametric CBC). In fact, this estimator was the most accurate in samples of small sizes.

In Tables 3 and 4 we present the square root of the estimated mean square error ($\sqrt{\text{MSE}}$) of the estimators of α and $\beta,$ respectively. Note in Table 3 that the estimates of $\sqrt{\text{MSE}}$ of $\tilde{\alpha}_1, \tilde{\alpha}_3$ and $\bar{\alpha}$ were similar to that of of the MLE $\hat{\alpha}.$ On the other hand, the estimates of $\sqrt{\text{MSE}}$ of $\tilde{\alpha}_2$ and $\tilde{\alpha}_4$ were significantly larger than that of $\hat{\alpha}.$ For instance, for $n = 60$ and $\alpha = 0.10,$ the estimates of $\sqrt{\text{MSE}}$ of $\hat{\alpha}, \tilde{\alpha}_2$ and $\tilde{\alpha}_4$ were 0.00902, 0.01268, and 0.02113, respectively.

Note in Table 4, similarly to what happens in Table 3, that the estimates of $\sqrt{\text{MSE}}$ of $\tilde{\beta}_1, \tilde{\beta}_3$ and $\hat{\beta}$ were close to that of $\hat{\beta}.$ On the other hand, the estimates of $\sqrt{\text{MSE}}$ of $\tilde{\beta}_2$ and $\tilde{\beta}_4$ were considerably larger than that of the MLE $\hat{\beta}.$ For instance, for $n = 60$ and $\alpha = 0.75,$ the estimates of $\sqrt{\text{MSE}}$ of $\hat{\beta}, \tilde{\beta}_2$ and $\tilde{\beta}_4$ were 0.08952, 0.12467, and 0.12026, respectively.

It is noteworthy that, as the shape parameter α increases, the estimates of the scale parameter β become less accurate (see Tables 2 and 4). For instance,

Table 1. Estimated relative bias ($\beta = 1$).

n	α	α estimates					
		$\hat{\alpha}$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\alpha}_4$	$\bar{\alpha}$
10	0.10	-0.07693	-0.00894	0.00979	-0.00546	0.00461	0.02563
	0.25	-0.07746	-0.00925	-0.01199	-0.00549	0.00455	0.02504
	0.50	-0.07933	-0.01041	-0.01428	-0.00582	0.00302	0.02297
	0.75	-0.08177	-0.01188	-0.01748	-0.00625	0.00637	0.02025
	1.00	-0.08444	-0.01348	-0.01982	-0.00686	0.00525	0.01729
20	0.10	-0.03824	-0.00286	-0.00361	-0.00167	0.00430	0.01238
	0.25	-0.03857	-0.00301	-0.00378	-0.00173	0.00066	0.01203
	0.50	-0.03964	-0.00349	-0.00439	-0.00191	0.01827	0.01091
	0.75	-0.04092	-0.00399	-0.00507	-0.00202	0.00636	0.00956
	1.00	-0.04225	-0.00452	-0.00579	-0.00217	0.00171	0.00816
40	0.10	-0.01848	-0.00023	-0.00093	0.00005	0.00314	0.00669
	0.25	-0.01866	-0.00029	-0.00105	0.00003	-0.00483	0.00651
	0.50	-0.01922	-0.00047	-0.00141	-0.00007	0.01841	0.00593
	0.75	-0.01988	-0.00062	-0.00173	-0.00010	-0.00059	0.00525
	1.00	-0.02055	-0.00079	-0.00210	-0.00014	0.00212	0.00456
60	0.10	-0.01233	-0.00014	-0.00053	0.00004	0.00158	0.00441
	0.25	-0.01246	-0.00019	-0.00061	0.00001	0.00248	0.00428
	0.50	-0.01281	-0.00026	-0.00074	-0.00002	-0.00178	0.00392
	0.75	-0.01323	-0.00031	-0.00083	-0.00001	-0.00130	0.00350
	1.00	-0.01367	-0.00037	-0.00095	-0.00001	0.00090	0.00305
Bootstrap		—	NP	NP	P	P	—

when $n = 20$, the estimated relative bias of the MLE $\hat{\beta}$ was 0.00039 for $\alpha = 0.10$, and 0.02150 for $\alpha = 1.00$, an increase of relative bias of almost 56 times.

Table 5 contains the mean of the 5000 estimates of c , denoted as \bar{c} , as well as the proportion of the 5000 estimates \check{c} which belonged to the set \mathcal{A} , previously defined. Note that all \bar{c} values for the shape parameter α belonged to \mathcal{A} .

Table 2. Estimated relative bias ($\beta = 1$).

n	α	β estimates					
		$\hat{\beta}$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$	$\tilde{\beta}_4$	$\bar{\beta}$
10	0.10	0.00046	0.00002	0.00053	0.00002	0.00025	0.00018
	0.25	0.00304	0.00032	0.00189	0.00030	0.00141	0.00130
	0.50	0.01185	0.00150	0.00503	0.00124	-0.17898	0.00491
	0.75	0.02510	0.00355	0.00931	0.00260	0.01349	0.00956
	1.00	0.04116	0.00619	0.01333	0.00409	0.01295	0.01368
20	0.10	0.00039	0.00018	0.00004	0.00014	-0.00038	0.00026
	0.25	0.00195	0.00054	0.00025	0.00045	0.00012	0.00112
	0.50	0.00669	0.00126	0.00088	0.00104	0.00314	0.00341
	0.75	0.01357	0.00226	0.00190	0.00174	-0.00091	0.00618
	1.00	0.02150	0.00333	0.00297	0.00234	0.02775	0.00837
40	0.10	0.00024	0.00011	0.00022	0.00011	0.00015	0.00017
	0.25	0.00107	0.00029	0.00057	0.00030	0.00073	0.00066
	0.50	0.00348	0.00058	0.00126	0.00059	0.00023	0.00188
	0.75	0.00691	0.00093	0.00211	0.00088	0.00527	0.00330
	1.00	0.01081	0.00131	0.00309	0.00114	0.01127	0.00440
60	0.10	0.00004	-0.00003	-0.00023	-0.00004	0.00008	0.00000
	0.25	0.00042	-0.00007	-0.00057	-0.00010	0.00031	0.00015
	0.50	0.00171	-0.00016	-0.00121	-0.00023	0.00056	0.00065
	0.75	0.00367	-0.00024	-0.00174	-0.00037	0.00008	0.00129
	1.00	0.00595	-0.00031	-0.00217	-0.00052	0.00097	0.00172
Bootstrap		—	NP	NP	P	P	—

Furthermore, note that these values are close to zero, especially in samples of small size (less than 40 observations). This suggests that the parameter c may be approximated by zero when estimating α , which corresponds to a constant bias function. Note also that the \bar{c} values for α were similar in all sample sizes. For instance, for $n = 20$, the values of \bar{c} were -0.0779 , -0.0781 , -0.0786 , -0.0790 , and -0.0793 , respectively, for $\alpha = 0.10, 0.25, 0.50, 0.75$,

Table 3. Square root of the estimated mean square error ($\beta = 1$).

n	α	α estimates					
		$\hat{\alpha}$	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\alpha}_4$	$\bar{\alpha}$
10	0.10	0.02333	0.02371	0.03080	0.02378	0.02420	0.02460
	0.25	0.05833	0.05929	0.07519	0.05949	0.06061	0.06145
	0.50	0.11684	0.11875	0.15117	0.11936	0.13545	0.12265
	0.75	0.17576	0.17858	0.22726	0.17979	0.18978	0.18365
	1.00	0.23552	0.23918	0.30435	0.24103	0.24527	0.24490
20	0.10	0.01594	0.01607	0.02161	0.01608	0.02053	0.01633
	0.25	0.03984	0.04019	0.05400	0.04022	0.08043	0.04080
	0.50	0.07972	0.08043	0.10788	0.08054	0.47666	0.08147
	0.75	0.11976	0.12081	0.16176	0.12107	0.17284	0.12206
	1.00	0.16010	0.16145	0.21583	0.16187	0.39681	0.16275
40	0.10	0.01116	0.01123	0.01530	0.01124	0.01487	0.01131
	0.25	0.02790	0.02807	0.03825	0.02809	0.12112	0.02826
	0.50	0.05579	0.05614	0.07635	0.05619	0.59965	0.05644
	0.75	0.08375	0.08429	0.11441	0.08439	0.18762	0.08461
	1.00	0.11182	0.11253	0.15251	0.11269	0.27302	0.11283
60	0.10	0.00902	0.00904	0.01268	0.00906	0.02113	0.00909
	0.25	0.02254	0.02261	0.03168	0.02265	0.04589	0.02273
	0.50	0.04506	0.04521	0.06330	0.04530	0.07546	0.04540
	0.75	0.06761	0.06784	0.09488	0.06799	0.13039	0.06806
	1.00	0.09020	0.09051	0.12647	0.09072	0.12946	0.09073
Bootstrap		—	NP	NP	P	P	—

and 1.00. Yet, most of the estimates $\check{\alpha}$, corresponding to the shape parameter α , belonged to \mathcal{A} (more than 50% in all considered cases), thus explaining the large variability of the estimator $\tilde{\alpha}_4$ relative to the MLE $\hat{\alpha}$ (see Table 3).

Note in Table 5 that the \bar{c} values for the scale parameter β vary considerably. For instance, for $n = 20$, the \bar{c} values were -2.0242 , 3.6810 , 0.7426 , 0.0686 and -0.3013 , for $\alpha = 0.10, 0.25, 0.50, 0.75$, and 1.00 , respectively. As a result,

Table 4. Square root of the estimated mean square error ($\beta = 1$).

n	α	β estimates					
		$\hat{\beta}$	$\tilde{\beta}_1$	$\tilde{\beta}_2$	$\tilde{\beta}_3$	$\tilde{\beta}_4$	$\bar{\beta}$
10	0.10	0.03176	0.03175	0.04413	0.03177	0.03287	0.03175
	0.25	0.07925	0.07898	0.11042	0.07906	0.10503	0.07907
	0.50	0.15720	0.15526	0.22005	0.15539	11.68536	0.15584
	0.75	0.23250	0.22701	0.32574	0.22685	0.72599	0.22823
	1.00	0.30394	0.29346	0.42732	0.29222	0.74609	0.29474
20	0.10	0.02251	0.02253	0.03131	0.02255	0.05065	0.02251
	0.25	0.05602	0.05598	0.07791	0.05601	0.10928	0.05595
	0.50	0.11024	0.10958	0.15300	0.10964	0.13444	0.10973
	0.75	0.16082	0.15873	0.22254	0.15875	0.23888	0.15923
	1.00	0.20629	0.20207	0.28485	0.20191	1.37750	0.20284
40	0.10	0.01577	0.01578	0.02229	0.01580	0.01766	0.01576
	0.25	0.03921	0.03920	0.05548	0.03926	0.04261	0.03919
	0.50	0.07693	0.07670	0.10896	0.07683	0.11120	0.07675
	0.75	0.11165	0.11090	0.15834	0.11108	0.20490	0.11109
	1.00	0.14222	0.14068	0.20203	0.14089	0.89815	0.14101
60	0.10	0.01275	0.01276	0.01774	0.01278	0.01309	0.01275
	0.25	0.03169	0.03169	0.04407	0.03173	0.03341	0.03168
	0.50	0.06196	0.06186	0.08622	0.06194	0.06757	0.06187
	0.75	0.08952	0.08916	0.12467	0.08927	0.12026	0.08924
	1.00	0.11348	0.11271	0.15815	0.11285	0.24111	0.11286
Bootstrap		—	NP	NP	P	P	—

we do not recommend the use of the estimator $\tilde{\beta}_4$, for it is possible that such oscillations lead to poor estimates of the slope of the bias function, thus yielding poor estimates of β . For instance, note in Table 2 that, for $n = 10$ and $\alpha = 0.50$, the estimated relative bias of the estimators $\hat{\beta}$ and $\tilde{\beta}_4$ were, in absolute values, 0.01185, and 0.17898, respectively. Still, as the value of the

Table 5. Linear parametric bootstrap ($\beta = 1$).

n	α	\bar{c}		% in \mathcal{A}	
		α	β	α	β
10	0.10	-0.0925	1.3077	67.50	28.54
	0.25	-0.0939	-0.2799	67.76	37.56
	0.50	-0.0979	-1.1699	69.08	50.78
	0.75	-0.1026	-1.0142	70.46	56.44
	1.00	-0.1064	-0.3078	71.62	57.60
20	0.10	-0.0779	-2.0242	56.62	26.92
	0.25	-0.0781	3.6810	56.80	31.08
	0.50	-0.0786	0.7426	57.22	41.32
	0.75	-0.0790	0.0686	58.10	47.64
	1.00	-0.0793	-0.3013	58.78	50.54
40	0.10	-0.1236	2.0441	52.74	25.12
	0.25	-0.1197	2.5700	52.84	27.62
	0.50	-0.1127	-2.8554	52.86	33.98
	0.75	-0.1056	1.1646	53.02	39.50
	1.00	-0.0998	0.3012	53.36	44.12
60	0.10	-0.1359	-1.1559	52.10	25.42
	0.25	-0.2307	1.1046	52.16	27.00
	0.50	-0.1562	1.5288	52.36	32.28
	0.75	-0.1927	-3.0283	52.56	37.74
	1.00	-0.1596	1.1644	52.74	41.48

shape parameter α increases, the proportions of the estimates \check{c} which belonged to \mathcal{A} , corresponding to the scale parameter β , also increased significantly. For instance, for $n = 10$, the percentage was 28.54% with $\alpha = 0.10$, and 57.60% with $\alpha = 1.00$.

In order to further understand some of the results presented above, we have produced two figures (Figures 2 and 3). The figures include estimated density plots and box plots. Density estimation was performed using kernel methods. (A Gaussian kernel was used.) Figure 2 corresponds to the estimation of α

Table 6. Real data from application.

152.7	172.0	172.5	173.3	193.0	204.7	216.5	234.9	262.6	422.6
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

whereas Figure 3 corresponds to the estimation of β . Two values of α were used, namely: $\alpha = 0.25, 0.75$. We note from these figures that the distributions of $\tilde{\alpha}_2$ and $\tilde{\beta}_2$ (weighted bootstrap estimators) are clearly less peaked than the remaining ones, thus resulting in larger variance. The box plots also show the existence of atypical values, especially in the upper tail. The larger variability of $\tilde{\alpha}_4$ and $\tilde{\beta}_4$ (linear parametric bootstrap estimators) stems from a few very atypical realizations in the upper tail of the distribution; their densities are not markedly less peaked than that of the parameteric and nonparametric bootstrap estimators.

It is noteworthy that the weighted nonparametric bootstrap estimator, i.e., $\tilde{\theta}_2 = (\tilde{\alpha}_2, \tilde{\beta}_2)^\top$ (based on the resampling vector $\mathbf{P}^{*(\cdot)}$), was clearly outperformed by other bootstrap-based estimators. This contrasts with the numerical results in Cribari-Neto, Frery and Silva [2], in the context of processing radar image data, where this estimator was the best performing amongst several bootstrap-based estimators.

The MLEs $\hat{\alpha}$ and $\hat{\beta}$ of α and β are considerably biased (see Tables 1 and 2), especially in small samples. We thus strongly recommend their bias correction. To that end, we recommend the use of the bias-adjusted parametric bootstrap estimator (CBC estimator), which displayed the best finite-sample performance.

5 Application to real data

The source of the data is McCool [26]. The data describe the lifetime, in hours, of 10 sustainers of a certain type. They were used as an illustration of the three-parameter Weibull distribution in Cohen, Whitten and Ding [27] and are given in Table 6.

The maximum likelihood estimates of the parameters α and β , and their corrected versions, together with their respective standard errors (in parentheses) are presented in Table 7, where ‘NP’ stands for nonparametric bootstrap and ‘P’ stands for parametric bootstrap. Note that the corrected estimates $\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4$, and $\tilde{\alpha}$ are larger than the maximum likelihood estimate $\hat{\alpha}$ and that the corrected estimates $\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3, \tilde{\beta}_4$, and $\tilde{\beta}$ are smaller than the maximum likelihood estimate $\hat{\beta}$. Note also that the estimates $\tilde{\alpha}_2$ and $\tilde{\beta}_2$ have large standard errors relative to $\hat{\alpha}$ and $\hat{\beta}$, respectively.

In Figure 4 we present the density (2) with α and β replaced by their corresponding estimates, which are given in Table 7. Note that the densities

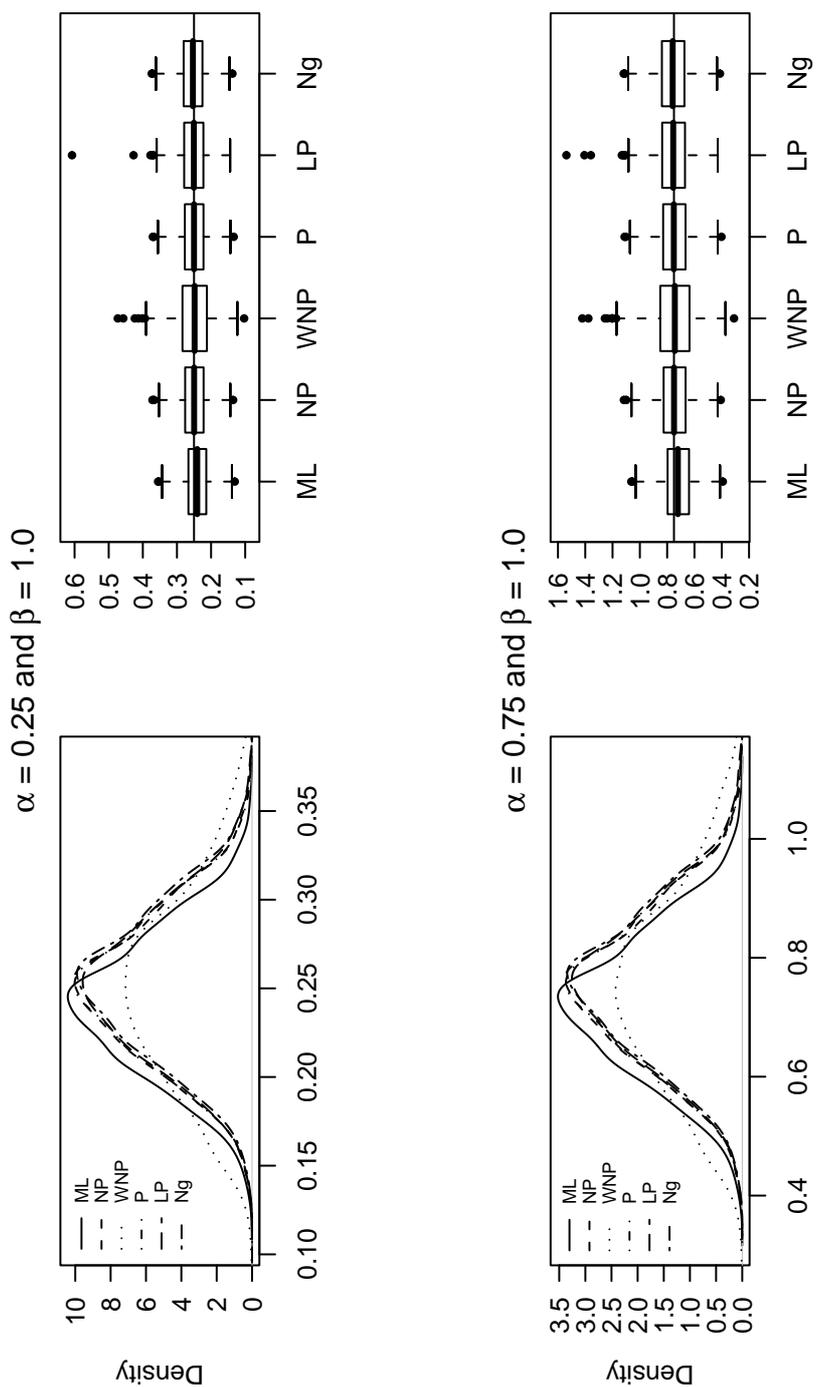


Figure 2. Estimated densities and box plots (estimators of α). (ML: maximum likelihood, NP: nonparametric bootstrap, WNP: weighted nonparametric bootstrap, P: parametric bootstrap, LP: linear parametric bootstrap, Ng: estimator proposed by Ng et al. (2003).)

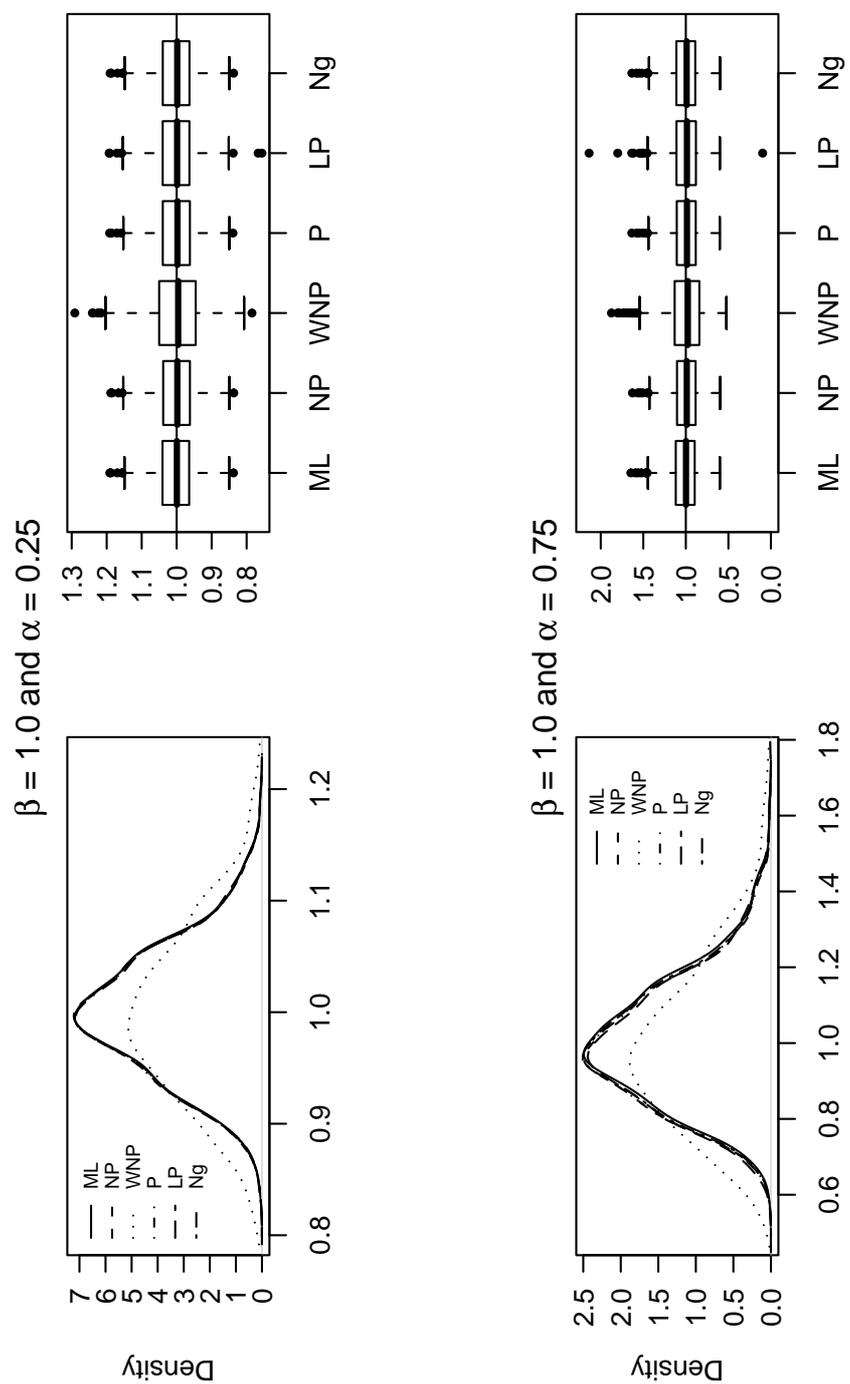


Figure 3. Estimated densities and box plots (estimators of β). (ML: maximum likelihood, NP: nonparametric bootstrap, WNP: weighted nonparametric bootstrap, P: parametric bootstrap, LP: linear parametric bootstrap, Ng: estimator proposed by Ng et al. (2003).)

Table 7. Point estimates of the parameters.

Estimator	Parameters		Bootstrap
	α	β	
$(\hat{\alpha}, \hat{\beta})$	0.28249 (0.06317)	212.04910 (18.75291)	—
$(\tilde{\alpha}_1, \tilde{\beta}_1)$	0.30942 (0.08319)	211.83562 (18.47368)	NP
$(\tilde{\alpha}_2, \tilde{\beta}_2)$	0.32455 (0.10493)	209.74397 (25.96065)	NP
$(\tilde{\alpha}_3, \tilde{\beta}_3)$	0.30578 (0.06373)	210.84103 (19.55948)	P
$(\tilde{\alpha}_4, \tilde{\beta}_4)$	0.31112 (0.06708)	210.89152 (19.75174)	P
$(\bar{\alpha}, \bar{\beta})$	0.31388 (0.07018)	211.52811 (20.73594)	—

obtained using the corrected estimates $(\tilde{\alpha}_1, \tilde{\beta}_1)$, $(\tilde{\alpha}_2, \tilde{\beta}_2)$, $(\tilde{\alpha}_3, \tilde{\beta}_3)$, $(\tilde{\alpha}_4, \tilde{\beta}_4)$ and $(\bar{\alpha}, \bar{\beta})$ are close to each other. It is also noteworthy that the density obtained using the maximum likelihood estimates $(\hat{\alpha}, \hat{\beta})$ is the ‘most peaked’; this happens due to underestimation of the shape parameter α .

6 Concluding remarks

This paper considered different bootstrapping schemes (parametric and non-parametric) for correcting the biases of the maximum likelihood estimators of the parameters that index the two-parameter Birnbaum-Saunders distribution, as well as the corrected estimators proposed by Ng, Kundu and Balakrishnan [15]. The numerical evidence showed that the bias correcting schemes are generally effective, even when the sample size is small.

The best performing estimator was the CBC parametric bootstrap estimator. We, therefore, strongly recommend that practitioners use this estimator when modeling data using the Birnbaum-Saunders distribution.

Acknowledgments

The authors gratefully acknowledge the financial support from CAPES and CNPq. We also thank an anonymous referee for comments and suggestions.

References

- [1] Birnbaum, Z. W. and Saunders, S. C., 1969a, A new family of life distributions. *Journal of Applied Probability*, **6**, 319–327.

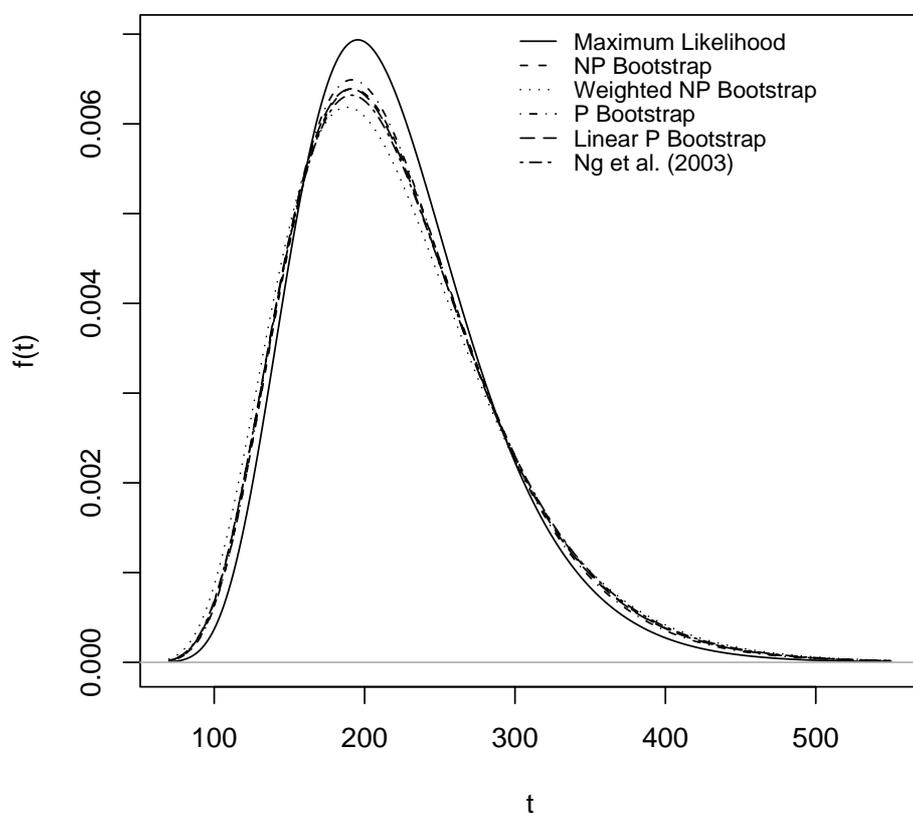


Figure 4. Estimated densities.

- [2] Cribari-Neto, F., Frery, A. C. and Silva, M. F., 2002, Improved estimation of clutter properties in speckled image. *Computational Statistics and Data Analysis*, **40**, 801–824.
- [3] Desmond, A. F., 1985, Stochastic models of failure in random environments. *Canadian Journal of Statistics*, **13**, 171–183.
- [4] Desmond, A. F., 1986, On the relationship between two fatigue-life models. *IEEE Transactions on Reliability*, **35**, 167–169.
- [5] Saunders, S. C., 1974, A family of random variables closed under reciprocation. *Journal of the American Statistical Association*, **69**, 533–539.
- [6] Mann, N. R., Schafer, R. E. and Singpurwalla, N., 1974, *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley and Sons, New York.
- [7] Chang, D. S. and Tang, L. C., 1993, Reliability bounds and critical time for the Birnbaum–Saunders distribution. *IEEE Transactions on Reliability*, **42**, 464–469.
- [8] Chang, D. S. and Tang, L. C., 1994, Percentile bounds and tolerance limits for the Birnbaum–Saunders distribution. *Communications in Statistics – Theory and Methods*, **23**, 2853–2863.
- [9] Dupuis, D. J. and Mills, J. E., 1998, Robust estimation of the Birnbaum–Saunders distribution. *IEEE Transactions on Reliability*, **47**, 88–95.
- [10] Rieck, J. R., 1995, Parametric estimation for the Birnbaum–Saunders distribution based on

- symmetrically censored samples. *Communications in Statistics – Theory and Methods*, **24**, 1721–1736.
- [11] Rieck, J. R., 1999, A moment-generating function with application to the Birnbaum-Saunders distribution. *Communications in Statistics – Theory and Methods*, **28**, 2213–2222.
- [12] Johnson, N., Kotz, S. and Balakrishnan, N., 1995, *Continuous Univariate Distributions – Vol. 2*, 2nd ed. John Wiley and Sons, New York.
- [13] Birnbaum, Z. W. and Saunders, S. C., 1969b, Estimation for a family of life distributions with applications to fatigue. *Journal Applied Probability*, **6**, 328–347.
- [14] Engelhardt, M., Bain, L. J. and Wright, F. T., 1981, Inferences on the parameters of the Birnbaum–Saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics*, **23**, 251–255.
- [15] Ng, H. K. T., Kundu, D. and Balakrishnan, N., 2003, Modified moment estimation for the two-parameter Birnbaum–Saunders distribution. *Computational Statistics and Data Analysis*, **43**, 283–298.
- [16] Efron, B., 1979, Bootstrap methods: another look at the jackknife. *Annals of Statistics*, **7**, 1–26.
- [17] Ferrari, S. L. P. and Cribari-Neto, F., 1998, On bootstrap and analytical bias corrections. *Economics Letters*, **58**, 7–15.
- [18] Efron, B., 1990, More efficient bootstrap computations. *Journal of the American Statistical Association*, **85**, 79–89.
- [19] MacKinnon, J. G. and Smith, J. A. A., 1998, Approximate bias correction in econometrics. *Journal of Econometrics*, **85**, 205–230.
- [20] Cribari-Neto, F., Frery, A. C. and Silva, M. F., 2002, Improved estimation of clutter properties in speckled imagery. *Computational Statistics and Data Analysis*, **40**, 801–824.
- [21] Davison, A. C. and Hinkley, D. V., 1997, *Bootstrap Methods and Their Application*. Cambridge University Press, New York.
- [22] Efron, B. and Tibshirani, R. J., 1993, *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- [23] Cribari-Neto, F. and Zarkos, S. G., 1999, Econometric and statistical computing using Ox. *Computational Economics*, **21**, 277–295.
- [24] Doornik, J. A., 2001, *Ox: An Object-Oriented Matrix Language*, 4th ed. Timberlake Consultants Press, London; Oxford, <http://www.doornik.com>.
- [25] Mittelhammer, R. C., Judge, G. G. and Miller, D. J., 2000, *Econometric Foundations*, New York: Cambridge University Press.
- [26] McCool, J. I., 1974, Inferential techniques for Weibull populations. Aerospace Research Laboratories Report ARL TR74-0180. Wright–atterson Air Force Base, Dayton, OH.
- [27] Cohen, A. C., Whitten, B. J. and Ding, Y., 1984, Modified moment estimation for the three-parameter Weibull distribution. *Journal of Quality Technology*, **16**, 159–167.