Inference in a bimodal Birnbaum-Saunders model

Rodney V. Fonseca\(^a\), Francisco Cribari-Neto\(^a\)

\(^a\)Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, Recife/PE, 50740-540, Brazil

Abstract

We address the issue of performing inference on the parameters that index a bimodal extension of the Birnbaum-Saunders distribution (BS). We show that maximum likelihood point estimation can be problematic since the standard nonlinear optimization algorithms may fail to converge. To deal with this problem, we penalize the log-likelihood function. The numerical evidence we present shows that maximum likelihood estimation based on such penalized function is made considerably more reliable. We also consider hypothesis testing inference based on the penalized log-likelihood function. In particular, we consider likelihood ratio, signed likelihood ratio, score and Wald tests. Bootstrap-based testing inference is also considered. We use a nonnested hypothesis test to distinguish between two bimodal BS laws. We derive analytical corrections to some tests. Monte Carlo simulation results and empirical applications are presented and discussed.

Keywords: Bimodal Birnbaum-Saunders distribution, Birnbaum-Saunders distribution, monotone likelihood, nonnested hypothesis test, penalized likelihood

1. Introduction

The Birnbaum-Saunders distribution was proposed by \(^7\) to model failure time due to fatigue under cyclic loading. In such a model, failure follows from
the development and growth of a dominant crack. Based on that setup, the authors obtained the following distribution function:

$$F(x) = \Phi\left[\frac{1}{\alpha} \left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)\right], \quad x > 0,$$

(1)

where $\alpha > 0$ and $\beta > 0$ are shape and scale parameters, respectively, and $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF). We write $X \sim BS(\alpha, \beta)$.

Maximum likelihood estimation of the parameters that index the BS distribution was first investigated by [6]. Bias-corrected estimators were obtained by [32] and [34]. Improved maximum likelihood estimation of the BS parameters was developed by [13]. [38] compared the finite-sample performance of maximum likelihood estimators (MLEs) to that of estimators obtained using the modified method of moments. For details on the BS distribution, its main properties and applications, readers are referred to [28].

Several extensions of the BS distribution have been proposed in the literature aiming at making the model more flexible. For instance, [18] and [45] used non-Gaussian kernels to extend the BS model. The BS distribution was also extended through the inclusion of additional parameters; see, e.g., [17], [40], and [41]. More recently, extensions of the BS model were proposed by [8], [11], [10] and [53]. Alternative approaches are the use of scale-mixture of normals, as discussed by [3] and [42], for example, and the use of mixtures of BS distributions, as in [2]. Again, details can be found in [28].

A bimodal BS distribution, which we denote by BBS distribution, was proposed by [39]. The authors used the approach described in [27] to obtain a variation of the BS model that can assume bimodal shapes. Another variant of the BS distribution that exhibits bimodality was discussed by [17] and [41], which the latter authors denoted by $\mathcal{GBS}_2$. In their model, bimodality takes place when two parameter values exceed certain thresholds. In what follows we shall work with the BBS model instead of the $\mathcal{GBS}_2$ distribution because in the former bimodality is controlled by a single parameter. Even though we shall focus on the BBS distribution, in some parts of the paper we shall consider the
Figure 1: Contour curves of the profile log-likelihood of $\alpha$ and $\gamma$ with $\beta = 0.69$ (fixed) for the runoff amounts data. Panel (a) corresponds to no penalization and panel (b) follows from penalizing the log-likelihood function.

$GBS_2$ law as an alternative model; see Section 6 for further details.

A problem with the $BBS$ distribution we detected is that log-likelihood maximizations based on Newton or quasi-Newton methods oftentimes fail to converge. In this paper we analyze some possible solutions to such a problem, such as the use of resampling methods and the inclusion of a penalization term in the log-likelihood function.

As a motivation, consider the data provided by [25] that consist of 25 observations on runoff amounts at Jug Bridge, in Maryland. Figure 1a shows log-likelihood contour curves obtained by varying the values of $\alpha$ and $\gamma$ while keeping the value of $\beta$ fixed. Notice that there is a region apparently flat of the profile log-likelihood function, which cause the optimization process to fail to converge. In Figure 1b, we present similar contour curves for a penalized version of the log-likelihood function. It can be seen that plausible estimates are obtained. We shall return to this application in Section 7.

The chief goal of our paper is to provide a solution to the convergence fail-
ure and implausible parameter estimates associated with log-likelihood maximization in the BBS model. We compare different estimation procedures and propose to include a penalization term in the log-likelihood function. In particular, regions of the parameter space where the likelihood is flat or nearly flat are heavily penalized. That approach considerably improves maximum likelihood parameter estimation. We also focus on hypothesis testing inference based on the penalized log-likelihood function. For instance, a one-sided hypothesis test is used to test whether the variate follows the BBS law with two modes. Analytical and bootstrap corrections are proposed to improve the finite sample performances of such test. Moreover, we present nonnested hypothesis tests that can be used to distinguish between two bimodal extensions of the BS distribution, the BBS and GBS2 models. The finite sample performances of all tests are numerically evaluated using Monte Carlo simulations.

The paper unfolds as follows. Section 2 presents the BBS distribution and its main properties. Simulation results are presented in Section 3 where we outline some possible solutions to the numerical difficulties associated with BBS log-likelihood maximization. Two-sided hypothesis tests in the BBS model are discussed in Section 4. In Section 5 we focus on one-sided tests where the main interest lies in detecting bimodality. Section 6 describes nonnested hypothesis testing inference. Empirical applications are presented and discussed in Section 7. Finally, some concluding remarks are offered in Section 8.

2. The bimodal Birnbaum-Saunders distribution

The Birnbaum-Saunders distribution proposed by [39] can be used to model positive data and is more flexible than the original BS distribution since it can accommodate bimodality. A random variable $X$ is $BBS(\alpha, \beta, \gamma)$ distributed if its probability density function (PDF) is given by

$$f(x) = \frac{x^{-3/2}(x + \beta)}{4\alpha\beta^{1/2}\Phi(-\gamma)}\phi(|t| + \gamma), \quad x > 0,$$

where $\alpha, \beta > 0, \gamma \in \mathbb{R}, t = \alpha^{-1}(\sqrt{x/\beta} - \sqrt{\beta/x})$ and $\phi(\cdot)$ is the standard normal PDF. Figure 2 shows plots of the density in (2) for some parameter values. We
Figure 2: \( \text{BBS}(\alpha, \beta, \gamma) \) densities for some parameter values.

Note that when \( \gamma < 0 \) the density is bimodal.

The CDF of \( X \) is

\[
F(x) = \left[ \frac{\Phi(t - \gamma)}{2\Phi(-\gamma)} \right] I(x, \beta) \left[ \frac{1}{2} + \frac{\Phi(t - \gamma) - \Phi(\gamma)}{2\Phi(-\gamma)} \right]^{1 - I(x, \beta)}, \quad x > 0, 
\]

where

\[
I(x, \beta) = \begin{cases} 
1 & \text{if } x < \beta, \\
0 & \text{if } x \geq \beta.
\end{cases}
\]

Some key properties of the \( \mathcal{BS} \) distribution also hold for the \( \mathcal{BBS} \) model, such as proportionality and reciprocity closure, i.e., \( aX \sim \text{BBS}(a\alpha, a\beta, \gamma) \) and \( X^{-1} \sim \text{BBS}(\alpha, \beta^{-1}, \gamma) \), respectively, where \( a \) is a positive scalar.

An expression for the \( r \)th ordinary moment of \( X \) is

\[
\mathbb{E}(X^r) = \frac{\beta^r}{\Phi(-\gamma)} \sum_{k=0}^{r} \sum_{j=0}^{k} \sum_{s=0}^{m} \binom{2r}{2k} \binom{m}{j} \binom{m}{s} \left( \frac{\alpha}{2} \right)^m (-\gamma)^{m-s} d_s(\gamma), 
\]

where \( r \in \mathbb{N} \) and \( d_a(r) \) is the \( r \)th standard normal incomplete moment:

\[
d_a(r) = \int_a^\infty t^r \phi(t) dt.
\]

A useful stochastic representation is \( Y = |T| + \gamma \). Here, \( Y \) follows the truncated standard normal distribution with support in \((\gamma, \infty)\), \( T = (\sqrt{X/\beta} - \sqrt{\beta/X})/\alpha \) and \( X \sim \text{BBS}(\alpha, \beta, \gamma) \). This relationship can be used to compute moments of the \( \text{BBS} \) distribution.
3. Log-likelihood functions

Consider a row vector \( x = (x_1, \ldots, x_n) \) of independent and identically distributed (IID) observations from the \( \text{BBS}(\alpha, \beta, \gamma) \) distribution. Let \( \theta = (\alpha, \beta, \gamma) \) be the vector of unknown parameters to be estimated. The log-likelihood function is

\[
\ell(\theta) = -n \log \left\{ 4\alpha\beta^{1/2}\Phi(-\gamma)(2\pi)^{1/2} \right\} - \frac{3}{2} \sum_{i=1}^{n} \log(x_i) + \sum_{i=1}^{n} \log(x_i + \beta)
- \frac{1}{2} \sum_{i=1}^{n} (|t_i| + \gamma)^2.
\] (6)

Differentiating the log-likelihood function with respect to each parameter we obtain the score function \( U_{\theta} = (U_{\alpha}, U_{\beta}, U_{\gamma}) \), where

\[
U_{\alpha} = \frac{\partial \ell(\theta)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{2} \sum_{i=1}^{n} t_i^2 + \frac{\gamma}{\alpha} \sum_{i=1}^{n} |t_i|,
\] (7)

\[
U_{\beta} = \frac{\partial \ell(\theta)}{\partial \beta} = -\frac{n}{2\beta} + \sum_{i=1}^{n} \frac{1}{x_i + \beta} + \sum_{i=1}^{n} \frac{\text{sign}(t_i)(|t_i| + \gamma)}{2\alpha\beta^{3/2}} \left( x_i^{1/2} + \frac{\beta}{x_i^{1/2}} \right),
\] (8)

\[
U_{\gamma} = \frac{\partial \ell(\theta)}{\partial \gamma} = n \frac{\phi(\gamma)}{\Phi(-\gamma)} - n\gamma - \sum_{i=1}^{n} |t_i|,
\] (9)

and \( \text{sign}(\cdot) \) represents the sign function.

The parameters MLEs, namely \( \hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}) \), can be obtained by solving \( U_{\theta} = 0 \). They cannot be expressed in closed-form and parameter estimates are obtained by numerically maximizing the log-likelihood function using a Newton or quasi-Newton algorithm. To that end, one must specify an initial point for the iterative scheme. We propose using as starting values for \( \alpha \) and \( \beta \) their modified method of moments estimates \[38\], and also using \( \gamma = 0 \) as a starting value; the latter means that the algorithm starts at the BS law. We used such starting values in the numerical evaluations, and they proved to work well.

Based on several numerical experiments we noted a serious shortcoming: iterative numerical maximization of the \( \text{BBS} \) log-likelihood function may fail to converge and may yield implausible parameter estimates. Indeed, that is very likely to happen, especially when \( \gamma > 0 \). It is not uncommon for one to obtain very large (thus implausible) \( \text{BBS} \) parameter estimates, which is indicative that
the likelihood function may be monotone; see [43]. We shall address this problem in the subsections that follow.

3.1. Log-likelihood function penalized by the Jeffreys prior

An interesting estimation procedure was proposed by [24], where the score function is modified in order to reduce the bias of the MLE. An advantage of this method is that maximum likelihood estimates need not be finite since the correction is applied in a preventive fashion. For models in the canonical exponential family, the correction can be applied directly to likelihood function:

\[
L^*(\theta|x) = L(\theta|x)|K|^{1/2},
\]

where \(|K|\) is the determinant of the expected information matrix. Thus, penalization of the likelihood function entails multiplying the likelihood function by the Jeffreys invariant prior.

Even though the BBS distribution is not a member of the canonical exponential family, we shall consider the above penalization scheme. In doing so, we follow [43] who used the same approach in speckled imagery analysis. We seek to prevent cases of monotone likelihood function that might lead to frequent optimization nonconvergences and implausible estimates. The BBS(\(\alpha, \beta, \gamma\)) expected information matrix was obtained by [39]. Its determinant is

\[
|K| = \left[ L_{\beta\beta} + \frac{1}{\alpha^2\beta^2} + \frac{\gamma(\gamma - \omega)}{4\beta^2} \right] \left[ \frac{(\gamma - \omega)\omega(3 - \gamma\omega - \gamma^2) + 2}{\alpha^2} \right],
\]

where \(\omega = \phi(\gamma)/\Phi(-\gamma)\) and \(L_{\beta\beta} = E[(X + \beta)^{-2}]\). Thus, the log-likelihood function penalized by the Jeffreys prior can be written as

\[
\ell^*(\theta) = -n \log \left\{ 4\alpha^{1/2}\beta \Phi(-\gamma)(2\pi)^{1/2} \right\} - \frac{3}{2} \sum_{i=1}^{n} \log(x_i) + \sum_{i=1}^{n} \log(x_i + \beta)
- \frac{1}{2} \sum_{i=1}^{n} (t_i^2 + 2t_i(\gamma + \gamma^2)) + \frac{1}{2} \log \left[ L_{\beta\beta} + \frac{1}{\alpha^2\beta^2} + \frac{\gamma(\gamma - \omega)}{4\beta^2} \right]
+ \frac{1}{2} \log \left[ \frac{(\gamma - \omega)\omega(3 + \gamma(\gamma - \omega)) + 2}{\alpha^2} \right].
\]

If the likelihood function is monotone, the function becomes very flat for large parameter values and the Jeffreys penalization described above essentially
eliminates such parameter range from the estimation. The likelihood of non-convergences taking place and implausible estimates being obtained should be greatly reduced.

3.2. Log-likelihood function modified by the better bootstrap

An alternative approach uses the method proposed by [14], where bootstrap samples are used to improve maximum likelihood estimation similarly to the approach introduced by [21] and known as ‘the better bootstrap’. The former, however, does not require the estimators to have closed-form expressions. Based on the sample \( x = (x_1, \ldots, x_n) \) of \( n \) observations, we obtain pseudo-samples \( x^* \) of the same size by sampling from \( x \) with replacement. Let \( P^*_i \) denote the proportion of times that observation \( x_i \) is selected, \( i = 1, \ldots, n \). We obtain the row vector \( P^*_b = (P^*_1, \ldots, P^*_n) \) for the \( b \)th pseudo-sample, \( b = 1, \ldots, B \). Now compute a row vector \( P^*(\cdot) \) as

\[
P^*(\cdot) = \frac{1}{B} \sum_{b=1}^{B} P^*_b,
\]
i.e., compute the vector of mean selection frequencies using the \( B \) bootstrap samples. The vector \( P^*(\cdot) \) is then used to modify the log-likelihood function in the following manner:

\[
\ell(\theta) = -n \log \left\{ 4\alpha\beta^{1/2} \Phi(-\gamma)(2\pi)^{1/2} \right\} - \frac{3n}{2} P^*(\cdot) \log(x)^\top + n P^*(\cdot) \log(x + \beta)^\top
- \frac{n}{2} P^*(\cdot)t_\gamma^\top,
\]

where \( \log(x) = (\log(x_1), \ldots, \log(x_n)) \), \( \log(x + \beta) = (\log(x_1 + \beta), \ldots, \log(x_n + \beta)) \) and \( t_\gamma = ((|t_1| + \gamma)^2, \ldots, (|t_n| + \gamma)^2) \) are row vectors. Hence, \( P^*(\cdot) \) is used to obtain a weighted average of the log-likelihood function terms that involve the data. The motivation behind the method is to approximate the ideal bootstrap estimates (which corresponds to \( B = \infty \)) faster than with the usual nonparametric bootstrap approach. In this paper we shall investigate whether this method is able to attenuate the numerical difficulties associated with BBS log-likelihood function maximization.
3.3. Log-likelihood function with a modified Jeffreys prior penalization

Monotone likelihood cases can arise with considerable frequency in models based on the asymmetric normal distribution, with some samples leading to situations where maximum likelihood estimates of the asymmetry parameter may not be finite, as noted by [36]. A solution to such a problem was proposed by [46], who used the score function transformation proposed by [24] in the asymmetric normal and Student-t models. A more general solution was proposed by [1], who penalized the log-likelihood function as follows:

$$\ell^*(\theta) = \ell(\theta) - Q,$$

where $\ell(\theta)$ and $\ell^*(\theta)$ denote the log-likelihood function and its modified version, respectively. The authors imposed some restrictions on $Q$, namely: (i) $Q \geq 0$; (ii) $Q = 0$ when the asymmetry parameter equals zero (values close to zero can lead to monotone likelihood cases in the asymmetric normal model); (iii) $Q \to \infty$ when the asymmetry parameter in absolute value tends to infinity. Additionally, $Q$ should not depend on the data or, at least, be $O_p(1)$. According to [1], when these conditions are satisfied, the estimators obtained using $\ell^*(\theta)$ are finite and have the same asymptotic properties as standard MLEs, such as consistency and asymptotic normality.

We shall now use a similar approach for the BBS model. In particular, we propose modifying the Jeffreys penalization term so that the new penalization satisfies the conditions listed by [1]. Since the numerical problems are mainly associated with $\alpha$ and $\gamma$, only terms involving these parameters were used. We then arrive at the following penalization term:

$$Q = Q_\gamma + Q_\alpha = -\frac{1}{2} \log \left\{ \frac{1}{2} \log \left( \frac{(\gamma - \omega)\omega + 3 + 2(\gamma - \omega)}{2} + 1 \right) + \frac{1}{2} \log(1 + \alpha^2) \right\}, \quad (12)$$

where, as before, $\omega = \phi(\gamma)/\Phi(-\gamma)$.

We note that $Q_\alpha \geq 0$. Additionally, $Q_\alpha \to 0$ when $\alpha \to 0$, and $Q_\alpha \to \infty$ when $\alpha \to \infty$. It can be shown that $Q_\gamma \to \infty$ when $\gamma \to \infty$ and that $Q_\gamma \to 0$ when $\gamma \to -\infty$, such that $Q_\gamma \geq 0$. Figure 3 shows the penalization terms as a function of the corresponding parameters. The quantities $Q_\gamma$ and $Q_\alpha$ penalize
large positive values of $\gamma$ and $\alpha$, helping avoid estimates that are unexpectedly large. Therefore, the proposed penalization satisfies the conditions indicated by \[1\]. An advantage of the penalization scheme we propose is that, unlike the Jeffreys penalization, it does not require the computation of $L_{\beta\beta}$. In what follows we shall numerically evaluate the effectiveness of the proposed correction when performing point estimation and we shall also consider the issue of carrying out testing inference on the parameters that index the BBS model.

3.4. Numerical evaluation

A numerical evaluation of the methods described in this section was performed. We considered different BBS estimation strategies. In what follows we shall focus on the estimation of the bimodality parameter $\gamma$.

The Monte Carlo simulations were carried out using the Ox matrix programming language [21]. Numerical maximizations were performed using the BFGS quasi-Newton method. We considered alternative nonlinear optimization algorithms such as Newton-Raphson and Fisher’s scoring, but they did not out-
perform the BFGS algorithm. We then decided to employ the BFGS method, which is typically regarded as the best performing method [37, Section 8.13]. The results are based on 5,000 Monte Carlo replications for values of $\gamma$ ranging from $-2$ to 2 and samples of size $n = 50$. In each replication, maximum likelihood estimates were computed and it was verified whether the nonlinear optimization algorithm converged. At the end of the experiment, the frequency of nonconvergences (proportion of samples for which there was no convergence) was computed for each method (denoted by $\text{pnf}$). Figure 4 shows the proportion of nonconvergences corresponding to the standard MLEs, the MLEs obtained using the better bootstrap (MLE$_{bboot}$) and the MLEs obtained from the log-likelihood function penalized using the Jeffreys prior (MLE$_{jp}$) and its modified version (MLE$_{p}$) as a function of $\gamma$. Notice that MLE and MLE$_{bboot}$ are the worst performers when $\gamma > 0$; they display the largest rates of nonconvergence. The methods based on penalized log-likelihood function display the smallest values of $\text{pnf}$, with slight advantage for MLE$_{jp}$.

In order to evaluate the impact of the sample size on nonconvergence rates, a numerical study similar to the previous one was performed, but now with the value of the bimodality parameter fixed at $\gamma = 1$. The samples sizes are $n \in \{30, 45, 60, 75, \ldots, 300\}$. The number of Monte Carlo replications was 5,000 for each value of $n$. The results are displayed in Figure 5. We note that the sample size does not seem to influence the MLE and MLE$_{bboot}$ nonconvergence rates. The corresponding optimizations failed in approximately 40% of the samples regardless of the sample size. In contrast, the MLE$_{jp}$ and MLE$_{p}$ failure rates display a slight increase and then stabilize as $n$ increases. Recall that one of the conditions imposed by [1] on the penalization term is that it should remain $O_p(1)$ as $n \to \infty$, i.e., the penalization influence seems to decrease as larger sample sizes are used, which leads to slightly larger nonconvergence frequencies in larger samples.

A second set of Monte Carlo simulations was carried out, this time only considering the estimator that uses the better bootstrap resampling scheme and also estimators based on the two penalized likelihood functions, i.e., we now only
consider MLE_{bboot}, MLE_{jp} and MLE_{p}. Again, 5,000 Monte Carlo replications were performed. We estimated the bias (denoted by B) and mean squared errors (denoted by MSE) of the three estimators. The number of nonconvergences is denoted by nf. Tables 1, 2 and 3 contain the results for MLE_{jp}, MLE_{p} and MLE_{bboot}, respectively. Overall, MLE_{p} outperforms MLE_{jp}. For instance, when n = 30 in the last combination of parameter values, the MSEs of \( \hat{\alpha}_{jp} \), \( \hat{\beta}_{jp} \) and \( \hat{\gamma}_{jp} \) are, respectively, 0.0211, 0.0022 and 2.6671, whereas the corresponding values for \( \hat{\alpha}_{p} \), \( \hat{\beta}_{p} \) and \( \hat{\gamma}_{p} \) are 0.0167, 0.0020 and 2.0471. MLE_{bboot} is typically less biased when it comes to the estimation of \( \alpha \) and \( \gamma \), but there are more convergence failures when computing better bootstrap estimates. Overall, the estimator based on the log-likelihood function that uses the penalization term we proposed typically yields more accurate estimates than MLE_{jp} and outperforms MLE_{bboot} in terms of convergence rates.

Next, we shall evaluate how changes in the penalization term impact the frequency of nonconvergences when computing MLE_{p}. In particular, we consider
Table 1: Bias and mean squared error of MLE for some combinations of parameter values.

| n  | \( \hat{B}(\hat{\alpha}_{jp}) \) | \( \hat{B}(\hat{\beta}_{jp}) \) | \( \hat{B}(\hat{\gamma}_{jp}) \) | MSE(\(\hat{\alpha}_{jp}\)) | MSE(\(\hat{\beta}_{jp}\)) | MSE(\(\hat{\gamma}_{jp}\)) | nf |
|----|----------------------------------|----------------------------------|----------------------------------|-----------------------------|-----------------------------|-----------------------------|----|
| 30 | -0.0851                          | -0.0048                          | -0.4369                          | 0.0134                      | 0.0118                      | 0.3770                      | 1  |
| 50 | -0.0567                          | -0.0032                          | -0.2844                          | 0.0079                      | 0.0068                      | 0.2022                      | 7  |
| 100| -0.0303                          | -0.0025                          | -0.1498                          | 0.0038                      | 0.0032                      | 0.0876                      | 2  |
| 150| -0.0208                          | -0.0016                          | -0.1018                          | 0.0025                      | 0.0021                      | 0.0574                      | 0  |
|    | \( \alpha = 0.5, \beta = 1 \) and \( \gamma = -1 \) |                                  |                                  |                             |                             |                             |    |
| 30 | -0.1532                          | -0.0037                          | -0.8487                          | 0.0303                      | 0.0103                      | 0.9668                      | 67 |
| 50 | -0.1103                          | -0.0017                          | -0.5798                          | 0.0184                      | 0.0057                      | 0.5201                      | 124|
| 100| -0.0692                          | -0.0002                          | -0.3523                          | 0.0095                      | 0.0026                      | 0.2476                      | 185|
| 150| -0.0484                          | -0.0006                          | -0.2412                          | 0.0061                      | 0.0017                      | 0.1510                      | 181|
|    | \( \alpha = 0.5, \beta = 1 \) and \( \gamma = 0 \) |                                  |                                  |                             |                             |                             |    |
| 30 | -0.2295                          | 0.0001                           | -1.5358                          | 0.0585                      | 0.0064                      | 2.6802                      | 373|
| 50 | -0.1865                          | -0.0018                          | -1.1795                          | 0.0403                      | 0.0031                      | 1.6360                      | 729|
| 100| -0.1332                          | -0.0009                          | -0.8061                          | 0.0229                      | 0.0012                      | 0.8366                      | 1520|
| 150| -0.1070                          | -0.0005                          | -0.6337                          | 0.0164                      | 0.0008                      | 0.5793                      | 2115|
|    | \( \alpha = 0.5, \beta = 1 \) and \( \gamma = 1 \) |                                  |                                  |                             |                             |                             |    |
| 30 | -0.0520                          | 0.0004                           | -0.4429                          | 0.0048                      | 0.0050                      | 0.3812                      | 5  |
| 50 | -0.0353                          | -0.0005                          | -0.2874                          | 0.0029                      | 0.0028                      | 0.2068                      | 2  |
| 100| -0.0174                          | -0.0013                          | -0.1400                          | 0.0013                      | 0.0013                      | 0.0875                      | 2  |
| 150| -0.0119                          | -0.0003                          | -0.0970                          | 0.0009                      | 0.0009                      | 0.0578                      | 1  |
|    | \( \alpha = 0.3, \beta = 1 \) and \( \gamma = -1 \) |                                  |                                  |                             |                             |                             |    |
| 30 | -0.0914                          | -0.0023                          | -0.8477                          | 0.0109                      | 0.0039                      | 0.9736                      | 86 |
| 50 | -0.0663                          | -0.0018                          | -0.5873                          | 0.0067                      | 0.0022                      | 0.5309                      | 127|
| 100| -0.0401                          | -0.0003                          | -0.3406                          | 0.0033                      | 0.0010                      | 0.2366                      | 205|
| 150| -0.0298                          | 0.0000                           | -0.2472                          | 0.0022                      | 0.0006                      | 0.1512                      | 208|
|    | \( \alpha = 0.3, \beta = 1 \) and \( \gamma = 0 \) |                                  |                                  |                             |                             |                             |    |
| 30 | -0.1380                          | 0.0002                           | -1.5331                          | 0.0211                      | 0.0022                      | 2.6671                      | 419|
| 50 | -0.1114                          | 0.0001                           | -1.1799                          | 0.0144                      | 0.0011                      | 1.6382                      | 745|
| 100| -0.0808                          | -0.0009                          | -0.8149                          | 0.0084                      | 0.0004                      | 0.8538                      | 1652|
| 150| -0.0649                          | 0.0000                           | -0.6450                          | 0.0059                      | 0.0003                      | 0.5863                      | 2211|
Table 2: Bias and mean squared error of MLE for some combinations of parameter values.

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.0575                     | 0.0078                      | -0.3129                     | 0.0114                  | 0.0121                  | 0.3151                  | 12 |
| 50 | -0.0364                     | 0.0006                      | -0.1918                     | 0.0072                  | 0.0068                  | 0.1730                  | 4  |
| 100| -0.0190                     | 0.0018                      | -0.0999                     | 0.0035                  | 0.0032                  | 0.0795                  | 0  |
| 150| -0.0128                     | 0.0004                      | -0.0670                     | 0.0025                  | 0.0021                  | 0.0546                  | 0  |

$\alpha = 0.5$, $\beta = 1$ and $\gamma = -1$

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.0477                     | 0.0017                      | -0.2505                     | 0.0082                  | 0.0026                  | 0.2069                  | 287|
| 50 | -0.0323                     | 0.0009                      | -0.1847                     | 0.0057                  | 0.0017                  | 0.1369                  | 318|

$\alpha = 0.5$, $\beta = 1$ and $\gamma = 0$

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.0165                     | 0.0035                      | -1.3012                     | 0.0474                  | 0.0057                  | 2.0986                  | 857|
| 50 | -0.1535                     | 0.0002                      | -0.9703                     | 0.0321                  | 0.0029                  | 1.2726                  | 1398|

$\alpha = 0.5$, $\beta = 1$ and $\gamma = 1$

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.0326                     | 0.0012                      | -0.2940                     | 0.0042                  | 0.0049                  | 0.3119                  | 14 |
| 50 | -0.0197                     | 0.0005                      | -0.1725                     | 0.0025                  | 0.0029                  | 0.1672                  | 7  |

$\alpha = 0.3$, $\beta = 1$ and $\gamma = -1$

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.0682                     | 0.0012                      | -0.6436                     | 0.0085                  | 0.0036                  | 0.7375                  | 252|
| 50 | -0.0474                     | 0.0020                      | -0.4288                     | 0.0053                  | 0.0021                  | 0.4085                  | 298|

$\alpha = 0.3$, $\beta = 1$ and $\gamma = 0$

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.0682                     | 0.0012                      | -0.6436                     | 0.0085                  | 0.0036                  | 0.7375                  | 252|
| 50 | -0.0474                     | 0.0020                      | -0.4288                     | 0.0053                  | 0.0021                  | 0.4085                  | 298|

$\alpha = 0.3$, $\beta = 1$ and $\gamma = 1$

| n  | $\hat{B}(\hat{\alpha}_p)$ | $\hat{B}(\hat{\beta}_p)$ | $\hat{B}(\hat{\gamma}_p)$ | MSE$(\hat{\alpha}_p)$ | MSE$(\hat{\beta}_p)$ | MSE$(\hat{\gamma}_p)$ | nf |
|----|-----------------------------|-----------------------------|-----------------------------|-------------------------|-------------------------|-------------------------|----|
| 30 | -0.1161                     | 0.0005                      | -1.2810                     | 0.0167                  | 0.0020                  | 2.0471                  | 979|
| 50 | -0.0920                     | -0.0005                     | -0.9769                     | 0.0116                  | 0.0010                  | 1.2924                  | 1475|

$\alpha = 0.3$, $\beta = 1$ and $\gamma = 0$
Table 3: Bias and mean squared error of MLE_{\text{bboot}} for some combinations of parameter values.

| n   | $\hat{B}(\hat{\alpha}_{\text{bboot}})$ | $\hat{B}(\hat{\beta}_{\text{bboot}})$ | $\hat{B}(\hat{\gamma}_{\text{bboot}})$ | $\text{MSE}(\hat{\alpha}_{\text{bboot}})$ | $\text{MSE}(\hat{\beta}_{\text{bboot}})$ | $\text{MSE}(\hat{\gamma}_{\text{bboot}})$ | nf |
|-----|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-----|
| 30  | -0.0283                         | 0.0091                          | -0.1575                         | 0.0141                          | 0.0121                          | 0.3221                          | 126 |
| 50  | -0.0130                         | 0.0035                          | -0.0761                         | 0.0086                          | 0.0070                          | 0.1898                          | 38  |
| 100 | -0.0075                         | 0.0017                          | -0.0430                         | 0.0040                          | 0.0033                          | 0.0861                          | 2   |
| 150 | -0.0041                         | 0.0018                          | -0.0230                         | 0.0027                          | 0.0021                          | 0.0568                          | 0   |

$\alpha = 0.5$, $\beta = 1$ and $\gamma = -1$

| n   | $\hat{B}(\hat{\alpha}_{\text{bboot}})$ | $\hat{B}(\hat{\beta}_{\text{bboot}})$ | $\hat{B}(\hat{\gamma}_{\text{bboot}})$ | $\text{MSE}(\hat{\alpha}_{\text{bboot}})$ | $\text{MSE}(\hat{\beta}_{\text{bboot}})$ | $\text{MSE}(\hat{\gamma}_{\text{bboot}})$ | nf |
|-----|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-----|
| 30  | -0.0811                         | 0.0056                          | -0.4596                         | 0.0281                          | 0.0093                          | 0.7186                          | 1494 |
| 50  | -0.0476                         | 0.0038                          | -0.2658                         | 0.0177                          | 0.0054                          | 0.4071                          | 1224 |
| 100 | -0.0215                         | 0.0001                          | -0.1192                         | 0.0093                          | 0.0025                          | 0.2052                          | 860  |
| 150 | -0.0138                         | 0.0001                          | -0.0780                         | 0.0061                          | 0.0016                          | 0.1365                          | 657  |

$\alpha = 0.5$, $\beta = 1$ and $\gamma = 0$

| n   | $\hat{B}(\hat{\alpha}_{\text{bboot}})$ | $\hat{B}(\hat{\beta}_{\text{bboot}})$ | $\hat{B}(\hat{\gamma}_{\text{bboot}})$ | $\text{MSE}(\hat{\alpha}_{\text{bboot}})$ | $\text{MSE}(\hat{\beta}_{\text{bboot}})$ | $\text{MSE}(\hat{\gamma}_{\text{bboot}})$ | nf |
|-----|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-----|
| 30  | -0.1609                         | 0.0030                          | -1.0607                         | 0.0498                          | 0.0053                          | 1.9644                          | 4640 |
| 50  | -0.1203                         | 0.0026                          | -0.7657                         | 0.0380                          | 0.0028                          | 1.2973                          | 5096 |
| 100 | -0.0735                         | 0.0003                          | -0.4606                         | 0.0228                          | 0.0012                          | 0.7213                          | 5877 |
| 150 | -0.0562                         | 0.0001                          | -0.3442                         | 0.0154                          | 0.0007                          | 0.4782                          | 5926 |

$\alpha = 0.5$, $\beta = 1$ and $\gamma = 1$

| n   | $\hat{B}(\hat{\alpha}_{\text{bboot}})$ | $\hat{B}(\hat{\beta}_{\text{bboot}})$ | $\hat{B}(\hat{\gamma}_{\text{bboot}})$ | $\text{MSE}(\hat{\alpha}_{\text{bboot}})$ | $\text{MSE}(\hat{\beta}_{\text{bboot}})$ | $\text{MSE}(\hat{\gamma}_{\text{bboot}})$ | nf |
|-----|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-----|
| 30  | -0.0159                         | 0.0016                          | -0.1554                         | 0.0050                          | 0.0049                          | 0.3204                          | 150  |
| 50  | -0.0090                         | 0.0028                          | -0.0820                         | 0.0031                          | 0.0029                          | 0.1927                          | 43   |
| 100 | -0.0040                         | 0.0008                          | -0.0384                         | 0.0015                          | 0.0013                          | 0.0883                          | 2    |
| 150 | -0.0032                         | 0.0008                          | -0.0292                         | 0.0009                          | 0.0008                          | 0.0551                          | 0    |

$\alpha = 0.3$, $\beta = 1$ and $\gamma = -1$

| n   | $\hat{B}(\hat{\alpha}_{\text{bboot}})$ | $\hat{B}(\hat{\beta}_{\text{bboot}})$ | $\hat{B}(\hat{\gamma}_{\text{bboot}})$ | $\text{MSE}(\hat{\alpha}_{\text{bboot}})$ | $\text{MSE}(\hat{\beta}_{\text{bboot}})$ | $\text{MSE}(\hat{\gamma}_{\text{bboot}})$ | nf |
|-----|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-----|
| 30  | -0.0472                         | 0.0024                          | -0.4434                         | 0.0101                          | 0.0034                          | 0.7141                          | 1543 |
| 50  | -0.0280                         | -0.0000                         | -0.2633                         | 0.0070                          | 0.0020                          | 0.4315                          | 1252 |
| 100 | -0.0127                         | 0.0003                          | -0.1150                         | 0.0034                          | 0.0009                          | 0.2104                          | 813  |
| 150 | -0.0090                         | 0.0007                          | -0.0823                         | 0.0023                          | 0.0006                          | 0.1416                          | 698  |

$\alpha = 0.3$, $\beta = 1$ and $\gamma = 0$

| n   | $\hat{B}(\hat{\alpha}_{\text{bboot}})$ | $\hat{B}(\hat{\beta}_{\text{bboot}})$ | $\hat{B}(\hat{\gamma}_{\text{bboot}})$ | $\text{MSE}(\hat{\alpha}_{\text{bboot}})$ | $\text{MSE}(\hat{\beta}_{\text{bboot}})$ | $\text{MSE}(\hat{\gamma}_{\text{bboot}})$ | nf |
|-----|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|-----|
| 30  | -0.0983                         | 0.0007                          | -1.0801                         | 0.0179                          | 0.0019                          | 1.9779                          | 4528 |
| 50  | -0.0716                         | 0.0002                          | -0.7586                         | 0.0131                          | 0.0010                          | 1.2703                          | 5128 |
| 100 | -0.0428                         | 0.0004                          | -0.4492                         | 0.0094                          | 0.0004                          | 0.7788                          | 5844 |
| 150 | -0.0328                         | -0.0000                         | -0.3410                         | 0.0059                          | 0.0003                          | 0.5082                          | 6142 |

$\alpha = 0.3$, $\beta = 1$ and $\gamma = 1$
the following penalized log-likelihood function:

\[ \ell^*(\theta) = \ell(\theta) - Q^\phi, \]

with \( \phi > 0 \) fixed. This additional quantity controls for the penalization strength, with \( \phi = 1 \) resulting in MLE\(_p\) and different values of \( \phi \) leading to stronger or weaker penalizations. A Monte Carlo study was performed to evaluate the accuracy of the parameter estimates for \( \gamma \in \{0, 0.1, \ldots, 2.0\} \) and \( \phi \in \{0.1, 0.2, \ldots, 2.1\} \). The parameter values are \( \alpha = 0.5 \) and \( \beta = 1 \), and the sample size is \( n = 50 \). Again, 5,000 replications were performed for each combination of values of \( \phi \) and \( \gamma \). Samples for which there was convergence failure were discarded. Figure 6 shows the estimated MSEs and the number of non-convergences for each combination of \( \gamma \) and \( \phi \). Figures 6a and 6b show that estimates of \( \alpha \) are less accurate than those of \( \beta \), both being considerably more accurate than the estimates of \( \gamma \) (Figure 6c). The MSE of the estimator of \( \gamma \) tends to be smaller when the value of \( \phi \) is between 0.4 and 1, especially for larger values of \( \gamma \). Visual inspection of Figure 6d shows that larger values of
\(\gamma\) lead to more nonconvergences, which was expected in light of our previous results. Furthermore, \(n_f\) tends to decrease when larger values of \(\phi\) are used.

We note from Figure [6] that the estimates of \(\gamma\) are the ones most sensitive to changes in the values of \(\gamma\) and \(\phi\). Figure [7] presents the number of nonconvergences (right vertical axis) and the MSE of \(\hat{\gamma}\) (left vertical axis) as a function of \(\gamma\) for three different values of \(\phi\). Figure [7a] shows that although \(\phi = 0.5\) yields more accurate estimates it also leads to more nonconvergences. For \(\phi = 1.5\), the number of nonconvergences did not exceed 1400, but \(\text{MSE}(\hat{\gamma})\) was larger relative to other values of \(\phi\). Overall, \(\phi = 1\) seems to balance well accuracy and the likelihood of convergence. In what follows we shall use \(\phi = 1\).

4. Two-sided hypothesis tests

In this section we consider two-sided hypothesis tests in the BBS model. Our interest lies in investigating the finite-sample performances of tests based on MLE\(_p\). The first test we consider is the penalized likelihood ratio test, denoted by LR. Consider a model parametrized by \(\theta = (\psi, \lambda)\), where \(\psi\) is the parameter of interest and \(\lambda\) is a nuisance parameter vector. Our interest lies in testing \(H_0: \psi = \psi_0\) against a two-sided alternative hypothesis. The LR test statistic is

\[ W = 2\{\ell^*(\hat{\theta}) - \ell^*(\tilde{\theta})\}, \]

where \(\hat{\theta}\) is the unrestricted MLE\(_p\) of \(\theta\), i.e., \(\hat{\theta}\) is obtained by maximizing \(\ell^*(\theta)\) without imposing restrictions on the parameters, and \(\tilde{\theta}\) is the restricted MLE\(_p\), which follows from the maximization of \(\ell^*(\theta)\) subject to the restrictions in \(H_0\). Critical values at the \(\epsilon \times 100\%\) significance level are obtained from the null distribution of \(W\) which, based on the results in [1], can be approximated by \(\chi^2_1\) when \(\psi\) is scalar. When \(\psi\) is a vector of dimension \(q (\leq 3)\), the test is performed in similar fashion with the single difference that the critical value is obtained from \(\chi^2_q\).

It is also possible to test \(H_0: \psi = \psi_0\) against \(H_1: \psi \neq \psi_0\) using the score and Wald tests. To that end, we use the score function and the expected
Figure 6: Mean squared errors of the estimators of $\alpha$ (a), $\beta$ (b) and $\gamma$ (c), and the number of nonconvergences $nf$ (d), for different values of $\gamma$ and $\phi$. 
Figure 7: Number of nonconvergences (solid line) and MSE of $\hat{\gamma}$ (dashed line) for different values of $\gamma$ with $\phi \in \{0.5, 1.0, 1.5\}$. The $nf$ values are shown in the left vertical axis and the values of $MSE(\hat{\gamma})$ are shown in the right vertical axis.

Information matrix obtained using the penalized log-likelihood function. The score and Wald test statistics are given, respectively, by

\[
W_S = U^\ast(\hat{\theta})^T K^\ast(\hat{\theta})^{-1} U^\ast(\hat{\theta}),
\]

\[
W_W = (\hat{\psi} - \psi_0)^2 / K^\ast(\hat{\theta})^{psi},
\]

where $U^\ast(\theta)$ and $K^\ast(\theta)$ denote the score function and the expected information, respectively, obtained using the penalized log-likelihood function and $K^\ast(\theta)^{psi}$ is the diagonal element of the inverse of $K^\ast(\theta)$ corresponding to $\psi$. Both test statistics are asymptotically distributed as $\chi^2_1$ under the null hypothesis. If instead of a scalar $\psi$ we considered a vector of dimension $q \leq 3$, the test statistic asymptotic null distribution would be $\chi^2_q$. An alternative test is the gradient test. We shall not consider it in the Monte Carlo simulations. For details on the gradient test, see [31].

A Monte Carlo simulation study was performed to evaluate the finite sample performances of the LR, score (denoted by S) and Wald tests in the BBS model. Log-likelihood maximizations were carried out using the BFGS quasi-Newton method. The number of Monte Carlo replications is 5,000 replica-
Table 4: Null rejection rates of the LR, score and Wald tests for testing of $H_0 : \alpha = 0.5$ against $H_1 : \alpha \neq 0.5$ in the $BBS(0.5, 1, 0)$ model.

| $n$ | LR   | S     | Wald  |
|-----|------|-------|-------|
|     | $\epsilon = 0.1$ |       |       |
| 30  | 0.2660 | 0.2166 | 0.4460 |
| 50  | 0.2008 | 0.1656 | 0.3374 |
| 100 | 0.1472 | 0.1288 | 0.2392 |
| 150 | 0.1346 | 0.1202 | 0.2078 |
|     | $\epsilon = 0.05$ |       |       |
| 30  | 0.1632 | 0.1116 | 0.3850 |
| 50  | 0.1156 | 0.0820 | 0.2836 |
| 100 | 0.0844 | 0.0664 | 0.1914 |
| 150 | 0.0716 | 0.0574 | 0.1524 |
|     | $\epsilon = 0.01$ |       |       |
| 30  | 0.0508 | 0.0104 | 0.2706 |
| 50  | 0.0352 | 0.0100 | 0.1822 |
| 100 | 0.0230 | 0.0094 | 0.1094 |
| 150 | 0.0166 | 0.0084 | 0.0766 |
Table 5: Null rejection rates of the LR, score and Wald tests for testing of $H_0: \beta = 1$ against $H_1: \beta \neq 1$ in the $BSS(0.5, 1, 0)$ model.

| n  | LR   | S   | Wald |
|----|------|-----|------|
|    |      |     |      |
| $\epsilon = 0.1$ |      |     |      |
| 30 | 0.1692 | 0.0832 | 0.2728 |
| 50 | 0.1416 | 0.0826 | 0.2116 |
| 100| 0.1190 | 0.0894 | 0.1650 |
| 150| 0.1170 | 0.0934 | 0.1494 |
| $\epsilon = 0.05$ |      |     |      |
| 30 | 0.1040 | 0.0340 | 0.2162 |
| 50 | 0.0802 | 0.0370 | 0.1480 |
| 100| 0.0620 | 0.0442 | 0.1090 |
| 150| 0.0558 | 0.0438 | 0.0808 |
| $\epsilon = 0.01$ |      |     |      |
| 30 | 0.0298 | 0.0048 | 0.1094 |
| 50 | 0.0218 | 0.0080 | 0.0734 |
| 100| 0.0112 | 0.0076 | 0.0354 |
| 150| 0.0112 | 0.0090 | 0.0268 |
Table 6: Null rejection rates of the LR, score and Wald tests for testing $H_0 : \gamma = 0$ against $H_1 : \gamma \neq 0$ in the $BSS(0.5, 1, 0)$ model.

| n  | LR   | S     | Wald  |
|----|------|-------|-------|
|    | \(\epsilon = 0.1\) |       |       |
| 30 | 0.2570 | 0.2434 | 0.3704 |
| 50 | 0.1974 | 0.1920 | 0.2896 |
| 100| 0.1368 | 0.1320 | 0.2014 |
| 150| 0.1364 | 0.1326 | 0.1774 |
|    | \(\epsilon = 0.05\) |       |       |
| 30 | 0.1678 | 0.1522 | 0.2992 |
| 50 | 0.1208 | 0.1076 | 0.2134 |
| 100| 0.0860 | 0.0762 | 0.1454 |
| 150| 0.0734 | 0.0660 | 0.1140 |
|    | \(\epsilon = 0.01\) |       |       |
| 30 | 0.0520 | 0.0334 | 0.1632 |
| 50 | 0.0336 | 0.0234 | 0.1100 |
| 100| 0.0218 | 0.0150 | 0.0658 |
| 150| 0.0154 | 0.0114 | 0.0474 |
tions, the sample sizes are \( n \in \{30, 50, 100, 150\} \) and the significance levels are \( \epsilon \in \{0.1, 0.05, 0.01\} \). The tests were performed for each parameter of the model \( BBS(0.5, 1, 0) \). It is noteworthy that by testing \( H_0 : \gamma = 0 \) against \( H_1 : \gamma \neq 0 \) we test whether the data follows the \( BS \) law, i.e., the original version of the Birnbaum-Saunders distribution. The data were generated according to the model implied by the null hypothesis and samples for which convergence did not take place were discarded.

Tables 4 to 6 contain the null rejection rates of the tests of \( H_0 : \alpha = 0.5 \), \( H_0 : \beta = 1 \) and \( H_0 : \gamma = 0 \), respectively, against two-sided alternative hypotheses. We note that all tests are considerably liberal when the sample size is small (30 or 50). We also note that the score test outperforms the competition. The Wald test was the worst performer.

The tests null rejection rates converge to the corresponding nominal levels as \( n \to \infty \). Such convergence, however, is rather slow. More accurate testing inference can be achieved by using bootstrap resampling; see [16]. The tests employ critical values that are estimated in the bootstrapping scheme instead of asymptotic (approximate) critical values. \( B \) bootstrap samples are generated imposing the null hypothesis and the test statistic is computed for each artificial sample. The critical value of level \( \epsilon \times 100\% \) is obtained as the \( 1 - \epsilon \) upper quantile of the \( B \) test statistics, i.e., of the test statistics computed using the bootstrap samples. The bootstrap tests are indicated by the subscript ‘pb’. We also use bootstrap resampling to estimate the Bartlett correction factor to the likelihood ratio test as proposed by [44]. The bootstrap Bartlett corrected test is indicated by the subscript ‘bbc’. For details on bootstrap tests, Bartlett-corrected tests and Bartlett corrections based on the bootstrap, the reader is referred to [9]. Since the Wald test proved to be considerably unreliable we shall not consider it.

Next, we shall numerically evaluate the finite sample performances of the \( LR_{pb} \), \( LR_{bbc} \) and \( S_{pb} \) tests under the same scenarios considered for the results presented in Tables 4 to 6. The number of Monte Carlo replications is as before. Samples for which the optimization methods failed to reach convergence were
Table 7: Null rejection rates of the bootstrap versions of the LR and score tests of \( H_0 : \alpha = 0.5 \) against \( H_1 : \alpha \neq 0.5 \) in the BBS(0.5, 1, 0) model.

| \( \epsilon \) | \( \text{LR}_{pb} \) | \( \text{LR}_{bbc} \) | \( S_{pb} \) |
|-------------|-------------|-------------|-------------|
|             |             |             |             |
| \( n = 30 \) |             |             |             |
| 0.10        | 0.1010      | 0.0934      | 0.0994      |
| 0.05        | 0.0542      | 0.0388      | 0.0500      |
| 0.01        | 0.0100      | 0.0038      | 0.0082      |
| \( n = 50 \) |             |             |             |
| 0.10        | 0.0990      | 0.0970      | 0.0964      |
| 0.05        | 0.0484      | 0.0420      | 0.0466      |
| 0.01        | 0.0096      | 0.0058      | 0.0108      |
| \( n = 100 \) |             |             |             |
| 0.10        | 0.1006      | 0.1026      | 0.0988      |
| 0.05        | 0.0510      | 0.0520      | 0.0488      |
| 0.01        | 0.0106      | 0.0102      | 0.0106      |
| \( n = 150 \) |             |             |             |
| 0.10        | 0.0950      | 0.0970      | 0.0940      |
| 0.05        | 0.0474      | 0.0484      | 0.0486      |
| 0.01        | 0.0096      | 0.0096      | 0.0110      |
Table 8: Null rejection rates of the bootstrap versions of the LR and score tests of $H_0: \beta = 1$ against $H_1: \beta \neq 1$ in the BBS(0.5, 1, 0) model.

| $\epsilon$ | $\text{LR}_{pb}$ | $\text{LR}_{bbc}$ | $S_{pb}$ |
|------------|------------------|------------------|--------|
| $n = 30$   |                  |                  |        |
| 0.10       | 0.1088           | 0.1088           | 0.1104 |
| 0.05       | 0.0496           | 0.0490           | 0.0514 |
| 0.01       | 0.0108           | 0.0100           | 0.0114 |
| $n = 50$   |                  |                  |        |
| 0.10       | 0.1064           | 0.1054           | 0.1096 |
| 0.05       | 0.0532           | 0.0528           | 0.0538 |
| 0.01       | 0.0092           | 0.0094           | 0.0112 |
| $n = 100$  |                  |                  |        |
| 0.10       | 0.1016           | 0.1002           | 0.1164 |
| 0.05       | 0.0538           | 0.0524           | 0.0574 |
| 0.01       | 0.0102           | 0.0112           | 0.0102 |
| $n = 150$  |                  |                  |        |
| 0.10       | 0.1048           | 0.1058           | 0.1140 |
| 0.05       | 0.0510           | 0.0510           | 0.0554 |
| 0.01       | 0.0098           | 0.0106           | 0.0104 |
Table 9: Null rejection rates of the bootstrap versions of the LR and score tests of $H_0 : \gamma = 0$ against $H_1 : \gamma \neq 0$ in the BBS(0.5, 1, 0) model.

| $\epsilon$ | LR$_{pb}$ | LR$_{bbc}$ | S$_{pb}$ |
|------------|-----------|------------|---------|
| $n = 30$   |           |            |         |
| 0.10       | 0.1034    | 0.0948     | 0.1022  |
| 0.05       | 0.0522    | 0.0392     | 0.0492  |
| 0.01       | 0.0094    | 0.0030     | 0.0106  |
| $n = 50$   |           |            |         |
| 0.10       | 0.1040    | 0.1028     | 0.1014  |
| 0.05       | 0.0532    | 0.0488     | 0.0514  |
| 0.01       | 0.0098    | 0.0062     | 0.0106  |
| $n = 100$  |           |            |         |
| 0.10       | 0.0980    | 0.1012     | 0.0976  |
| 0.05       | 0.0464    | 0.0488     | 0.0458  |
| 0.01       | 0.0082    | 0.0078     | 0.0088  |
| $n = 150$  |           |            |         |
| 0.10       | 0.0986    | 0.0992     | 0.1014  |
| 0.05       | 0.0468    | 0.0488     | 0.0456  |
| 0.01       | 0.0082    | 0.0080     | 0.0082  |
discard, even for bootstrap samples. The same 1,000 bootstrap samples \((B = 1,000)\) were used in all tests. The null rejection rates of the tests used for making inferences on \(\alpha\), \(\beta\) and \(\gamma\) are presented in Tables 7, 8 and 9, respectively. It is noteworthy that the tests size distortions are now considerably smaller. For instance, when making inference on \(\alpha\) based on a sample of size \(n = 30\) and \(\epsilon = 0.05\), the LR and score null rejection rates are 16.32\% and 11.16\% (Table 4), whereas the corresponding figures for their bootstrap versions LRbp, LRbbc and Sbp are 5.42\%, 3.88\% and 5\%, respectively, which are much closer to 5\%. When testing restrictions on \(\beta\) with \(n = 30\) and \(\epsilon = 0.10\), the LR and score null rejection rates are, respectively, 16.92\% and 8.32\% (Table 5) whereas their bootstrap versions, LRbp, LRbbc and Sbp, display null rejection rates of 10.88\%, 10.88\% and 11.04\%, respectively. Finally, when the interest lies in making inferences on \(\gamma\) with \(n = 30\) and \(\epsilon = 0.01\), the LR and score null rejection rates are 5.20\% and 3.34\% (Table 6); for the bootstrap-based tests LRbp, LRbbc and Sbp we obtain 0.94\%, 0.3\% and 1.06\%, respectively. Figure 8 shows the quantile-quantile (QQ) plots of the LR and LRbbc test statistics for Figure 8: Quantile-quantile plots for the LR and LRbbc test statistics with \(n = 50\), for the tests on \(\alpha\) (a), on \(\beta\) (b) and on \(\gamma\) (c).
samples of size \( n = 50 \). It is noteworthy that the empirical quantiles of the LR_{cb} test statistic are much more closer of the corresponding asymptotic quantiles than those of \( W \). Hence, we note that testing inference in small samples can be made considerably more accurate by using bootstrap resampling.

5. One-sided hypothesis tests

One-sided tests on a scalar parameter can be performed using the signed likelihood ratio (SLR) statistic, which is particularly useful in the BBS model since it allows practitioners to make inferences on \( \gamma \) in a way that makes it possible to detect bimodality. The signed penalized likelihood ratio test statistic is

\[
R = \text{sign}(\hat{\psi} - \psi_0) \sqrt{W} = \text{sign}(\hat{\psi} - \psi_0) \sqrt{2(\ell^*({\hat{\theta}}) - \ell^*(\tilde{\theta}))}.
\]  

(13)

The statistic \( R \) is asymptotically distributed as standard normal under the null hypothesis. An advantage of the SLR test over the tests described in Section 4 is that it can be used to perform two-sided and one-sided tests. In this section, we shall focus on one-sided hypothesis testing inference. Our interest lies in detecting bimodality. The null hypothesis is \( H_0 : \gamma \geq 0 \) which is tested against \( H_1 : \gamma < 0 \). Rejection of \( H_0 \) yields evidence that the data came from a bimodal distribution. On the other hand, when \( H_0 \) is not rejected, there is evidence that the data follow a distribution with a single mode.

Consider the sample \( x = (x_1, \ldots, x_n) \) for a model with parameter vector \( \theta = (\psi, \lambda) \) of dimension \( 1 \times p \), where the parameter of interest \( \psi \) is a scalar and the vector of nuisance parameters \( \lambda \) has dimension \( 1 \times (p - 1) \). The test statistic \( R \), given in Equation (13), is asymptotically distributed as standard normal with error of order \( O(n^{-1/2}) \) when the null hypothesis is true. Such an approximation may not be accurate when the sample size is small. Some analytical corrections for \( R \) were proposed in the literature. They can be used to improve the test finite sample behavior.

An important contribution was made by [4, 5]. The author proposed a
correction term $U$ of the form

$$R^* = R + \log(U/R)/R,$$

where $R$ represents the SLR statistic and $R^*$ is its corrected version. Let $\ell(\theta)$ be the log-likelihood function of the parameters. Its derivatives shall be denoted by

$$\ell_\theta(\theta) = \frac{\partial \ell(\theta)}{\partial \theta} \quad \text{and} \quad \ell_{\theta\theta}(\theta) = \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top}.$$

The observed information matrix is given by $J_{\theta\theta}(\theta) = -\ell_{\theta\theta}(\theta)$. To obtain the correction proposed by [4], the sufficient statistic has to be of the form $(\hat{\theta}, a)$, where $\hat{\theta}$ is the MLE of $\theta$ and $a$ is an ancillary statistic. Additionally, it is necessary to compute sample space derivatives of the log-likelihood, such as

$$\ell_{\hat{\theta}}(\theta) = \frac{\partial \ell(\theta)}{\partial \hat{\theta}} \quad \text{and} \quad \ell_{\theta\hat{\theta}}(\theta) = \frac{\partial \ell_{\hat{\theta}}(\theta)}{\partial \theta}.$$

where derivatives are taken with respect to some functions of the sample while keeping other terms fixed, as explained in [48]. The quantity $U$ is given by

$$U = \begin{vmatrix} \ell_{\hat{\theta}}(\theta) - \ell_{\hat{\theta}}(\hat{\theta}) & \ell_{\hat{\theta}\hat{\theta}}(\theta) \\ J_{\theta\theta}(\hat{\theta})^{1/2} & J_{\theta\theta}(\theta)^{1/2} \end{vmatrix}^{1/2},$$

where $\hat{\theta}$ is the restricted MLE of $\theta$ and the indices indicate which components are being used in each vector or matrix.

The null distribution of $R^*$ is standard normal with error of order $O(n^{-3/2})$. Although the null distribution of $R^*$ is better approximated by the limiting distribution than that of $R$, the computation of $U$ is restricted to some specific classes of models, such as exponential family and transformation models [48].

Some alternatives to $R^*$ were proposed in the literature. They approximate the sample space derivatives used in $U$. For instance, approximations were obtained by [19], [26] and [47]. They were computed by [52] for the BS model and by [33] for a Birnbaum-Saunders regression model. Other recent contributions are [22] and [49].
In this paper, we apply the approximations proposed by [47] and [26] for log-likelihood functions without penalization. Our interest is to evaluate the effectiveness of the corrections when applied to the statistic $R$, computed using the penalized log-likelihood function. We also compare the performances of the corrected tests to those of the SLR test and its bootstrap version.

Using the same notation as [33], the approximation proposed by [26] (denoted by $\text{SLR}_{c1}$) for $U$ can be written as

\[
U_{c1} = \frac{\Gamma_{\theta}}{J_{\lambda \lambda}(\hat{\theta})^{1/2}} \frac{\Psi_{\lambda \theta}}{J_{\theta \theta}(\hat{\theta})^{1/2}},
\]

where

\[
\Gamma_{\theta} = [\ell_{x}(\theta) - \ell_{x}(\hat{\theta})]V(\hat{\theta})[\ell_{x}(\theta)\lambda V(\hat{\theta})]^{-1}J_{\theta \theta}(\hat{\theta}),
\]

with

\[
\Psi_{\theta \theta} = \begin{bmatrix} \Psi_{\varphi \theta} \\
\Psi_{\lambda \theta} \end{bmatrix} = [\ell_{\theta \varphi}(\hat{\theta})V(\hat{\theta})]J_{\theta \theta}(\hat{\theta})^{-1}J_{\theta \theta}(\hat{\theta}),
\]

where $\ell_{x}(\theta) = \partial l(\theta)/\partial x$ is a $1 \times n$ vector, $\ell_{\theta \varphi}(\hat{\theta}) = \partial^2 l(\theta)/\partial \theta \partial x$ is a $p \times n$ matrix and

\[
V(\theta) = \left[ \frac{\partial z(x; \theta)}{\partial x} \right]^{-1} \left[ \frac{\partial z(x; \theta)}{\partial \theta} \right],
\]

is an $n \times p$ matrix, $z(x; \theta)$ being a vector of pivotal quantities.

The corrected SLR statistic obtained using the approximation given by [26] is $R_{c1} = R + \log(U_{c1}/R)$, which has asymptotic standard normal distribution with error of order $O(n^{-3/2})$ under the null hypothesis. We derived the quantities needed to obtain $U_{c1}$ in the BBS model, which are presented below.

Consider the random variable $Y = |T| + \gamma$, where $T = \alpha^{-1}(\sqrt{X/\beta} - \sqrt{\beta/X})$ and $X \sim \text{BBS}(\alpha, \beta, \gamma)$. The distribution of $Y$ is truncated standard normal with support $(\gamma, \infty)$, its distribution function being given by

\[
F_Y(y) = \begin{cases} 0 & \text{if } y < \gamma, \\ \frac{\Phi(y) - \Phi(\gamma)}{1 - \Phi(\gamma)} & \text{if } y \geq \gamma. \end{cases}
\]
Therefore, $Z = F_Y(Y)$ is uniformly distributed in the standard interval, $(0, 1)$. Hence, it is a pivotal quantity that can be used for obtaining the approximations to sample space derivatives proposed by \[26\]. Let $x = (x_1, \ldots, x_n)$ be a random $BBS(\alpha, \beta, \gamma)$ sample. It follows that $\partial z_i/\partial x_j = 0$ when $i \neq j$ and $\partial z_i/\partial x_i = \phi(y_i)\text{sign}(t_i)(x_i + \beta)/[\Phi(-\gamma)2\alpha\beta^{1/2}x_i^{3/2}]$, with $t_i = \alpha^{-1}(\sqrt{x_i/\beta} - \sqrt{\beta/x_i})$.

Moreover, $\partial z_i/\partial \alpha = -\phi(y_i)\text{sign}(t_i)t_i/\Phi(-\gamma)\alpha$, $\partial z_i/\partial \beta = -\phi(y_i)\text{sign}(t_i)\sqrt{x_i/\beta + \sqrt{\beta/x_i}}/[\Phi(-\gamma)2\alpha\beta]$ and $\partial z_i/\partial \gamma = [\Phi(y_i) - \Phi(\gamma)]\phi(\gamma)/\Phi^2(-\gamma) + [\phi(y_i) - \phi(\gamma)]/\Phi(-\gamma)$, where $y_i = |t_i| + \gamma$ and $z_i = F_Y(y_i)$. Therefore, $v_{\alpha i} = 2\beta^{1/2}x_i^{3/2}t_i/(x_i + \beta)$, $v_{\beta i} = x_i/\beta$ and $v_{\gamma i} = -2\alpha\beta^{1/2}x_i^{3/2}(\Phi(y_i) - \Phi(\gamma))\phi(\gamma)/\Phi(-\gamma) + \phi(y_i) - \phi(\gamma))]/[\phi(y_i)\text{sign}(t_i)(x_i + \beta)]$. The vectors $v_\alpha$, $v_\beta$ and $v_\gamma$ are used to form the matrix $V(\theta)$. For instance, $v_\alpha = (v_{\alpha 1}, \ldots, v_{\alpha n})^\top$ is a $n \times 1$ vector. Here, $V(\theta) = [v_\alpha \ v_\beta \ v_\gamma]$. Furthermore, we have that

\[
\ell_{\alpha i}(\theta) = -\frac{3}{2x_i} + \frac{1}{x_i + \beta} - (|t_i| + \gamma)\text{sign}(t_i)\frac{(x_i + \beta)}{2\beta^{1/2}x_i^{3/2}},
\]

\[
\ell_{\alpha \alpha i}(\theta) = \frac{\text{sign}(t_i)(x_i + \beta)}{2\beta^{1/2}x_i^{3/2}}\alpha^{-1}(2|t_i| + \gamma),
\]

\[
\ell_{\beta \beta i}(\theta) = \frac{(-1)}{(x_i + \beta)^2} + \frac{(x_i + \beta)}{4\alpha^2\beta^{3/2}x_i^{3/2}}\left(x_i^{1/2}/\beta^{1/2} + \beta^{1/2}/x_i^{1/2}\right) + \frac{\text{sign}(t_i)(|t_i| + \gamma)(x_i - \beta)}{4\alpha\beta^{3/2}x_i^{3/2}},
\]

\[
\ell_{\gamma \gamma i}(\theta) = -\frac{\text{sign}(t_i)(x_i + \beta)}{2\beta^{1/2}x_i^{3/2}}.
\]

The method proposed by \[47\] (denoted by $\text{SLR}_{c2}$) approximates the sample space derivatives by covariances of the log-likelihood function. The main idea is to use the sample to obtain the covariance values empirically. Using again the notation of \[33\], the approximation of $U$ proposed by \[47\] is given by

\[
U_2 = \frac{\Delta_\theta}{\Sigma_{\lambda\theta}} \frac{1}{\sqrt{J_{\lambda\lambda}(\hat{\theta})}} \frac{1}{\sqrt{J_{\theta\theta}(\hat{\theta})}},
\]

with

\[
\Delta_\theta = [Q(\hat{\theta}; \hat{\theta}) - Q(\hat{\theta}; \hat{\theta})]I(\hat{\theta}; \hat{\theta})^{-1}J_{\theta\theta}(\hat{\theta})
\]
and
\[
\Sigma_{\theta\theta} = \begin{bmatrix}
\Sigma_{\psi\theta} \\
\Sigma_{\lambda\theta}
\end{bmatrix} = I(\hat{\theta}; \hat{\theta})I(\hat{\theta}; \hat{\theta})^{-1} J_{\theta\theta}(\hat{\theta}),
\]
where \( Q(\theta; \theta_0) = \sum_{i=1}^n \ell^{(i)}(\theta)\ell^{(i)}(\theta_0)\) is an \(1 \times p\) vector and \( I(\theta; \theta_0) = \sum_{i=1}^n \ell^{(i)}(\theta)\ell^{(i)}(\theta_0)\) is a \(p \times p\) matrix, the index \((i)\) indicating that the quantity corresponds to the \(i\)th sample observation. The corrected statistic proposed by \[47\] is \( R_{c2} = R + \log(U_2/R)/R \). Its null distribution is standard normal with error of order \(O(n^{-1})\). The score function and the observed information matrix, which can be found in \[39\], are used to obtain \( U_2 \) in the \textit{BBS} model.

Alternatively, bootstrap resampling can be used to obtain critical values for the SLR test. Since we test \( H_0 : \gamma \geq 0 \) against \( H_1 : \gamma < 0 \), the critical value of level \( \epsilon \times 100\% \) is obtained as the \( \epsilon \) quantile of the \( B \) test statistics computed using the bootstrap samples.

A simulation study was performed to evaluate the sizes and powers of the SLR, SLR\(_{c1}\), SLR\(_{c2}\) and SLR\(_{bp}\) tests. We tested \( H_0 : \gamma \geq 0 \) against \( H_1 : \gamma < 0 \). The true parameter values are \( \gamma \in \{-1, -0.5, 0, 0.5, 1\} \). The most reliable tests are those with large power (i.e., higher probability of rejecting \( H_0 \) when \( \gamma < 0 \)) and small size distortions. Again, 5,000 Monte Carlo replication were performed. The SLR\(_{bp}\) test is based on 1,000 bootstrap samples. The simulation results are presented in Table \[10\] The most powerful tests are SLR, SLR\(_{c2}\) and SLR\(_{c1}\), in that order, whereas the tests with the smallest size distortions are SLR\(_{bp}\), SLR\(_{c1}\) and SLR\(_{c2}\). We recommend that testing inference be based on either SLR\(_{c1}\) or SLR\(_{c2}\), since these tests display a good balance between size and power.

6. Nonnested hypothesis tests for the bimodal Birnbaum-Saunders model

In the previous section we presented a test that is useful for detecting whether the data came from a bimodal \textit{BBS} law. That was done by testing a restriction on \( \gamma \). In this section we shall present tests that are useful for distinguishing
Table 10: Null rejection rates of the SLR, SLR\(_{c1}\), SLR\(_{c2}\) and SLR\(_{bp}\) tests of \(H_0 : \gamma \geq 0\) against \(H_1 : \gamma < 0\) in a sample of size 30 of the model BBS(0.5, 1, \(\gamma\)).

| \(\epsilon\) | SLR | SLR\(_{c1}\) | SLR\(_{c2}\) | SLR\(_{bp}\) |
|---------------|-----|-------------|-------------|------------|
| \(\gamma = -1\) |     |             |             |            |
| 0.10          | 0.7488 | 0.5768     | 0.6310    | 0.4960 |
| 0.05          | 0.6300 | 0.4276     | 0.4728    | 0.3488 |
| 0.01          | 0.3560 | 0.1910     | 0.2100    | 0.1376 |
| \(\gamma = -0.5\) |     |             |             |            |
| 0.10          | 0.4634 | 0.2766     | 0.3300    | 0.2248 |
| 0.05          | 0.3242 | 0.1762     | 0.2042    | 0.1326 |
| 0.01          | 0.1328 | 0.0588     | 0.0652    | 0.0376 |
| \(\gamma = 0\) |     |             |             |            |
| 0.10          | 0.2614 | 0.1334     | 0.1678    | 0.1042 |
| 0.05          | 0.1658 | 0.0746     | 0.0892    | 0.0498 |
| 0.01          | 0.0486 | 0.0202     | 0.0210    | 0.0106 |
| \(\gamma = 0.5\) |     |             |             |            |
| 0.10          | 0.1546 | 0.0722     | 0.0928    | 0.0526 |
| 0.05          | 0.0890 | 0.0378     | 0.0442    | 0.0222 |
| 0.01          | 0.0214 | 0.0098     | 0.0102    | 0.0046 |
| \(\gamma = 1\) |     |             |             |            |
| 0.10          | 0.1144 | 0.0488     | 0.0640    | 0.0340 |
| 0.05          | 0.0606 | 0.0246     | 0.0292    | 0.0150 |
| 0.01          | 0.0144 | 0.0050     | 0.0046    | 0.0022 |
between the BBS model and another extension of the BS distribution that can display bimodality.

As noted in the Introduction, another variant of the BS distribution that can exhibit bimodality is the model recently discussed by [41], which the authors denoted by $\mathcal{GBS}_2$. Let $X \sim \mathcal{GBS}_2(\alpha, \beta, \nu)$. Its PDF is given by

$$g(x) = \frac{\nu}{\alpha x} \left[ \left( \frac{x}{\beta} \right)^{\nu} + \left( \frac{\beta}{x} \right)^{\nu} \right] \phi \left( \frac{1}{\alpha} \left[ \left( \frac{x}{\beta} \right)^{\nu} - \left( \frac{\beta}{x} \right)^{\nu} \right] \right), \quad x > 0,$$

where $\alpha > 0$, $\beta > 0$ and $\nu > 0$. According to [41], the $\mathcal{GBS}_2$ density is bimodal when $\alpha > 2$ and $\nu > 2$ (simultaneously).

Therefore, when bimodality is detected the subsequent data analysis may be carried out with either the BBS distribution or the $\mathcal{GBS}_2$ model. It would then be useful to have a hypothesis test that could be used to distinguish between the two models. Obviously, the tests discussed so far cannot be used to that end.

BS model selection criteria were considered by [28] and [30]. Model selection is usually based on the Bayes factor and also on the Schwarz and Akaike information criteria. We shall use a different approach: we shall develop tests for nonnested hypotheses. Notice that the $\mathcal{GBS}_2$ distribution cannot be obtained from the BBS distribution by imposing restrictions on the model parameters, and vice-versa. Hence, the two models are not nested.

The literature of nonnested models began with [12, 13]. The author introduced likelihood ratio tests for some nonnested models. His main results were generalized by [50], who considered nested, nonnested and overlapping models and derived the required asymptotics. For nonnested models, [50] established the relationship between the likelihood ratio statistic and the Kullback-Leibler information. Let $F$ and $G$ be competing nonnested models. The author presented a test of the null hypothesis $H_0$ that both models are equivalent, the alternative hypotheses being: $H_f$: model $F$ is better and $H_g$: model $G$ is better. An alternative approach for testing nonnested models was considered by [51] and [35]. The authors only considered tests of the hypothesis $H_f$ and $H_g$. They proposed to consider $H_f$ and $H_g$ sequentially.

We shall consider the hypothesis involving the BBS and $\mathcal{GBS}_2$ models as:
The test statistic we consider is the following likelihood ratio statistic:
\[
W_{ne} = \log \left( \frac{\hat{f}}{\hat{g}} \right) = \hat{\ell}_f - \hat{\ell}_g,
\]
where \( \hat{f} \) and \( \hat{g} \) denote the likelihood functions of the \textit{BBS} and \textit{GBS}_2 models, respectively, evaluated at the respective maximum likelihood estimates, \( \ell \) representing the log-likelihood function of the model indicated by its index. Then, for a given sample \( x \), a large positive value of \( W_{ne} \) yields evidence in favor \( H_f \) and against \( H_g \); on the other hand, a large negative value of \( W_{ne} \) favors \( H_g \).

The \textit{BBS} parameters are estimated using the penalized log-likelihood function and those of \textit{GBS}_2 are estimated using the standard log-likelihood function.

In the test introduced by [50] for nonnested models, the test statistic asymptotic null distribution is standard normal. In some Monte Carlo simulations not reported here, this test, based on asymptotic critical values, indicated the models equivalence too frequently. Since the test is based on a large sample approximation, superior finite sample performance can be achieved by using bootstrap resampling. Application of the bootstrap method is not, however, straightforward for the test at hand because one would need to define an model equivalent to \textit{BBS} and \textit{GBS}_2 in order to generate pseudo-samples under the null hypothesis. Thus, an approach similar to the one employed by [35], which only considers the hypotheses \( H_f \) and \( H_g \) in the test, will be used in this paper.

The null hypothesis \( H_f \) can be tested, using bootstrap resampling, as follows:

1. Compute \( W_{ne} \) using sample \( x \);
2. With the MLEs of the parameters from the \textit{BBS} model, generate a bootstrap sample \( x^* \), and then compute \( W_{ne}^* \) using that sample;
3. Execute step 2 \( B \) times and obtain the bootstrap p-value: \( p_b = \frac{\#\{W_{ne}^* < W_{ne}\} + 1}{B+1} \).

Hence, at the \( \epsilon \times 100\% \) significance level, \( H_f \) is rejected if \( p_b < \epsilon \), i.e., we reject the hypothesis that the data originated from the \textit{BBS} law and conclude that the
GBS$_2$ distribution is more adequate. Similar testing inference can be performed by taking $H_g$ as the null hypothesis. The test is carried out as follows:

1. Compute $W_{ne}$ using sample $x$;
2. With the MLEs of the parameters from the GBS$_2$ model, generate a bootstrap sample $x^*$, and then compute $W_{ne}^*$ using that sample;
3. Execute step 2 $B$ times and obtain the bootstrap $p$-value: $p_b = \frac{\#\{W_{ne}^* > W_{ne}\} + 1}{B+1}$.

It is noteworthy that step 3 is different from the corresponding step in the first procedure, since the rejection region changes when we consider $H_g$ as the null hypothesis. Again, at the $\epsilon \times 100\%$ significance level, the null hypothesis is rejected if $p_b < \epsilon$, but now that means that we reject the hypothesis $H_g$ according to which the data came from the GBS$_2$ distribution and conclude that the BBS model is more adequate.

The problem with this approach is that four inference results can happen, as noted by [51] and [35]:

- **R1** The two null hypotheses, $H_f$ and $H_g$, are not rejected, and we conclude that both models are adequate;
- **R2** We do not reject $H_f$, but $H_g$ is rejected, thus indicating that the BBS model is more adequate;
- **R3** We do not reject $H_g$, but $H_f$ is rejected, thus indicating that the GBS$_2$ model is more adequate;
- **R4** We reject both null hypotheses, $H_f$ and $H_g$, and conclude that neither model is adequate;

Under some regularity conditions, [50] has shown that, in nonnested models, an adjusted likelihood ratio test statistic tends to infinity under $H_f$ when $n \to \infty$ and that under $H_g$ it tends to minus infinity when $n \to \infty$. That way, the test statistic tends to indicate the correct model as the sample size increases. Therefore, when result **R1** is reached, model selection can be based on $W_{ne}$: the
Figure 9: Densities \( BBS(0.2, 1, -1) \) (solid line) and \( GBS_2(5, 1, 5) \) (dashed line).

Table 11: Proportions of outcomes of the test of \( H_f \) against \( H_g \) when the data generating function is \( BBS(0.2, 1, -1) \) (first four columns), and proportions of \( BBS \) and \( GBS_2 \) model selection and proportion of no model selected. The test significance level is \( \epsilon \).

| \( \epsilon \) | \( R1 \) | \( R2 \) | \( R3 \) | \( R4 \) | \( BBS \) | \( GBS_2 \) | None |
|--------------|--------|--------|--------|--------|--------|--------|------|
| \( n = 30 \) |
| 0.10         | 0.6676 | 0.2378 | 0.0946 | 0.0000 | 0.4354 | 0.5646 | 0.0000 |
| 0.05         | 0.8284 | 0.1322 | 0.0394 | 0.0000 | 0.4354 | 0.5646 | 0.0000 |
| 0.01         | 0.9650 | 0.0292 | 0.0058 | 0.0000 | 0.4354 | 0.5646 | 0.0000 |
| \( n = 50 \) |
| 0.10         | 0.5454 | 0.3578 | 0.0968 | 0.0000 | 0.5574 | 0.4426 | 0.0000 |
| 0.05         | 0.7230 | 0.2326 | 0.0444 | 0.0000 | 0.5572 | 0.4428 | 0.0000 |
| 0.01         | 0.9174 | 0.0754 | 0.0072 | 0.0000 | 0.5572 | 0.4428 | 0.0000 |
| \( n = 100 \) |
| 0.10         | 0.3420 | 0.5538 | 0.1036 | 0.0006 | 0.6952 | 0.3042 | 0.0006 |
| 0.05         | 0.5344 | 0.4138 | 0.0518 | 0.0000 | 0.6940 | 0.3060 | 0.0000 |
| 0.01         | 0.8024 | 0.1884 | 0.0092 | 0.0000 | 0.6940 | 0.3060 | 0.0000 |
| \( n = 150 \) |
| 0.10         | 0.1920 | 0.6908 | 0.1138 | 0.0034 | 0.7682 | 0.2284 | 0.0034 |
| 0.05         | 0.3772 | 0.5646 | 0.0582 | 0.0000 | 0.7656 | 0.2344 | 0.0000 |
| 0.01         | 0.6720 | 0.3118 | 0.0162 | 0.0000 | 0.7654 | 0.2346 | 0.0000 |

37
Table 12: Proportions of outcomes of the test of $H_f$ against $H_g$ when the data generating function is $GBS_2(5,1,5)$ (first four columns), and proportions of $BBS$ and $GBS_2$ model selection and proportion of no model selected. The test significance level is $\epsilon$.

| $\epsilon$ | $R1$   | $R2$   | $R3$   | $R4$   | $BBS$  | $GBS_2$ | None |
|------------|--------|--------|--------|--------|--------|---------|------|
|            |        |        |        |        |        |         |      |
| $n = 30$   |        |        |        |        |        |         |      |
| 0.10       | 0.4438 | 0.1080 | 0.4482 | 0.0000 | 0.1784 | 0.8216  | 0.0000|
| 0.05       | 0.6658 | 0.0538 | 0.2804 | 0.0000 | 0.1784 | 0.8216  | 0.0000|
| 0.01       | 0.9040 | 0.0112 | 0.0848 | 0.0000 | 0.1784 | 0.8216  | 0.0000|
| $n = 50$   |        |        |        |        |        |         |      |
| 0.10       | 0.2204 | 0.1010 | 0.6786 | 0.0000 | 0.1270 | 0.8730  | 0.0000|
| 0.05       | 0.4502 | 0.0466 | 0.5032 | 0.0000 | 0.1238 | 0.8762  | 0.0000|
| 0.01       | 0.7624 | 0.0090 | 0.2286 | 0.0000 | 0.1238 | 0.8762  | 0.0000|
| $n = 100$  |        |        |        |        |        |         |      |
| 0.10       | 0.0146 | 0.0662 | 0.8804 | 0.0388 | 0.0672 | 0.8940  | 0.0388|
| 0.05       | 0.1068 | 0.0560 | 0.8350 | 0.0022 | 0.0712 | 0.9266  | 0.0022|
| 0.01       | 0.4000 | 0.0112 | 0.5888 | 0.0000 | 0.0674 | 0.9326  | 0.0000|
| $n = 150$  |        |        |        |        |        |         |      |
| 0.10       | 0.0002 | 0.0158 | 0.8968 | 0.0872 | 0.0158 | 0.8970  | 0.0872|
| 0.05       | 0.0136 | 0.0310 | 0.9344 | 0.0210 | 0.0320 | 0.9470  | 0.0210|
| 0.01       | 0.1610 | 0.0122 | 0.8268 | 0.0000 | 0.0322 | 0.9678  | 0.0000|

38
\( BBS \) distribution is selected if \( W_{ne} > 0 \) and the \( GBS_2 \) distribution is selected if \( W_{ne} < 0 \).

A simulation study was performed to evaluate the performances of the nonnested hypothesis tests involving the \( BBS \) and \( GBS_2 \) distributions. The models considered were \( BBS(0.2, 1, -1) \) and \( GBS_2(5, 1, 5) \). Figure 9 shows the two densities. The number of Monte Carlo replications used was 5,000. First, we considered the case in which the true distribution is \( BBS(0.2, 1, -1) \); in each replication, \( B = 1,000 \) bootstrap samples were generated under \( H_f \) and other \( B = 1,000 \) bootstrap samples were generated under \( H_g \), thus reaching one of the previously indicated results (\( R_1, R_2, R_3 \) or \( R_4 \)) in each replication. Table 11 contains the proportions of times that each inference was reached and also the proportions of times each distribution was chosen as the most suitable model, which is: the \( BBS \) model when we obtain result \( R_1 \) and \( W_{ne} > 0 \) or when we obtain result \( R_2 \); the \( GBS_2 \) model when we obtain result \( R_1 \) and \( W_{ne} < 0 \) or when we obtain result \( R_3 \); none of the considered distributions under result \( R_4 \). The same procedure was used when the true model was the \( GBS_2(5, 1, 5) \) distribution, the corresponding results being presented in Table 12.

Based on the figures in Table 11 we note that the null rejection rates of the true hypothesis (\( H_f \)) are close to the nominal levels. For instance, when \( n = 30 \) and \( \epsilon = 0.10 \), by adding the cells corresponding to \( R_3 \) and \( R_4 \), we see that the rejection rate of \( H_f \) is 9.46\%, which is close to the test nominal level. We also note that for small or moderate sample sizes the tests tend to indicate equivalence of both models, but as \( n \) increases, the tests tend to indicate the \( BBS \) model as the most suitable model with increasing frequency. For example, when \( \epsilon = 0.05 \), in the column that corresponds to \( R_2 \), that happens for 13.22\% of the samples when \( n = 30 \), whereas for \( n = 150 \), for 56.46\% of the samples the \( BBS \) distribution is considered the most adequate model. This can also be observed in the fifth column, where we can see that as \( n \) increases, the \( BBS \) model is selected more frequently.

Table 12 contains the results obtained by taking \( H_g \) as the true null hypothesis. Once again, the null rejection rates stayed close of the nominal levels.
When $n = 30$ and $\epsilon = 0.05$, the sum of the cells relative to $R2$ and $R4$ equals 5.38%, which is quite close to the test significance level. Moreover, we note that the results in Table 12 are superior to those obtained under $H_f$. In the column corresponding to $R3$ we find the proportions of samples for which the $GBS_2$ law was chosen as the most adequate model by disregarding the sign of $W_{ne}$. When $\epsilon = 0.05$, for $n = 30$ that happens for 28.04% of all samples and for $n = 150$ that happens for 93.44% of the samples, a performance which is superior to that observed when $H_f$ was the true model. This can also be seen in the sixth column of the table. The proportions of samples for which the $GBS_2$ distribution is correctly chosen are higher than the corresponding values in the fifth column of Table 11.

Therefore, we conclude that the bootstrap-based nonnested hypothesis tests used for distinguishing between the $BBS$ and $GBS_2$ models perform well. For both distributions the null rejection rates are close to the nominal levels. We also note that, as $n$ increases, the tests tend to single out the correct model with increasing frequency. We also note that the tests perform better when the true model is the $GBS_2$ distribution.

7. Empirical applications

7.1. Runoff amounts

We shall now return to the data briefly described in Section 1, which we used to illustrate the problem of nonconvergence of optimization processes during $BBS$ parameter estimation. The data, provided by [25], consist of 25 runoff amounts at Jug Bridge, in Maryland. Table 13 contains some descriptive statistics. We note that the data have large kurtosis, i.e., they come from a leptokurtic distribution, and have small variance, which might be indicative that the data are concentrated around the mean and median values. These characteristics may suggest that the data came from a unimodal population.

The models fitted to the data were $BS(\alpha, \beta)$ and $BBS(\alpha, \beta, \gamma)$. For the former, we obtained $\hat{\alpha} = 0.66$ (0.0936) and $\hat{\beta} = 0.69$ (0.0865); the numbers in
Table 13: Descriptive statistics for the runoff data.

|        | min | max | median | mean   | variance | asymmetry | kurtosis |
|--------|-----|-----|--------|--------|----------|-----------|----------|
|        | 0.17| 2.92| 0.7    | 0.84   | 0.3459   | 1.7953    | 6.7493   |

Parentheses are standard errors. For the second model, the maximum likelihood estimates could not be obtained because the optimization algorithm failed to converge. As shown in Figure 1b, the log-likelihood function has a region which is apparently flat for some values of the parameters $\alpha$ and $\gamma$, with the value of $\beta$ being fixed at 0.69. In contrast, the BBS penalized maximum likelihood estimates ($\text{MLE}_p$) were easily obtained: $\hat{\alpha} = 0.63$ (0.2287), $\hat{\beta} = 0.69$ (0.0817) and $\hat{\gamma} = -0.13$ (0.8449). Notice that the standard error of $\hat{\gamma}$ is large relative to the point estimate which indicates that the BS model is adequate.

Figure 10 contains the data histogram and the two fitted densities. It is noteworthy that the fitted densities are very similar. Since the BS distribution is simpler than the BBS distribution, it is to preferred. As a confirmation check, we tested the null hypothesis $H_0: \gamma = 0$ against a two-sided alternative. The $p$-values of the LR, score, Wald, LR$_{pb}$, LR$_{bbc}$ and $S_{pb}$ tests were equal to 0.85, 0.81, 0.87, 0.92, 0.89 and 0.91, respectively. Therefore, there is strong evidence that the BS fit is adequate.

7.2. Depressive condition data

The second empirical application uses data on the emotional condition of 134 children. In particular, the interest lies in modeling depression measures. The data were analyzed, for example, by [29] and [2]. In both papers, mixtures of distributions were used.

Table 14 presents some descriptive statistics. The data are right-skewed, leptokurtic and highly dispersed. The following models were fitted: BS, BBS and GBS$_2$. For each model, we computed the Akaike (AIC) and Schwarz (BIC) information criteria. The BS estimates (standard errors in parentheses) are $\hat{\alpha} = 0.603$ (0.0368) and $\hat{\beta} = 7.58$ (0.3773), with AIC and BIC values of 780.09 and 785.89, respectively. For the BBS model, $\hat{\alpha} = 0.42$ (0.0481), $\hat{\beta} = 7.54$.
Figure 10: Histogram of the runoff data with the fitted densities obtained with $BS(0.66, 0.69)$ (dashed line) and $BBS(0.63, 0.69, -0.13)$ (dotted line).

The point estimates are $\hat{\alpha} = 2.38 (0.6290)$, $\hat{\beta} = 7.74 (0.3045)$ and $\hat{\nu} = 1.53 (0.2525)$, with AIC and BIC values of 771.78 and 780.47, respectively. The data histogram and the fitted densities are presented in Figure 11.

The information criteria favor the $GBS_2$ model, followed by the $BBS$ law. We performed a nonnested hypothesis test to distinguish between the two models. For these data, $W_{ne} = -0.0167$ with $p$-value of 0.0189 under $H_f$ ($BBS$ is taken to be the true model) and $p$-value of 0.6453 under $H_g$ ($GBS_2$ is assumed to be the true model). Hence, we select $GBS_2$ model.

Table 14: Descriptive statistics for the depressive condition data.

| min | max | median | mean | variance | asymmetry | kurtosis |
|-----|-----|--------|------|----------|-----------|----------|
| 3   | 28  | 8      | 8.96 | 28.73    | 1.11      | 3.88     |
7.3. Adhesive strength

The third data set analyzed is provided by [22] and was also analyzed by [39], who used the $BBS$ distribution in the analysis. The data consist of 48 observations on the adhesive strength to concrete of bars reinforced with glass fiber. Some descriptive statistics are presented in Table 15. Notice the large kurtosis coefficient (in excess of 5), the positive asymmetry and also the fact that the variance is considerably larger than both the mean and the median.

| min  | max  | median | mean   | variance | asymmetry | kurtosis |
|------|------|--------|--------|----------|-----------|----------|
| 3.4  | 25.5 | 5.95   | 8.08   | 23.7017  | 1.448     | 5.0345   |

Once again, the $BS$, $BBS$ and $GBS_2$ models were fitted to the data. The $BS$ point estimates are $\hat{\alpha} = 0.54$ (0.0553) and $\hat{\beta} = 7.05$ (0.5316), the AIC and BIC values being 264.52 and 268.26, respectively. For the $BBS$ model, $\hat{\alpha} = 0.31$ (0.0460), $\hat{\beta} = 7.39$ (0.3162) and $\hat{\gamma} = -1.38$ (0.3525), with AIC and BIC values
of 260.06 and 265.67, respectively. Finally, for the $\mathcal{GBS}_2$ distribution, $\hat{\alpha} = 3.19 (1.5536)$, $\hat{\beta} = 8.05 (0.5371)$ and $\hat{\nu} = 1.99 (0.5203)$, with AIC and BIC values of 262.26 and 267.88, respectively. The data histogram and the fitted densities are shown in Figure 12.

For this data set, the best fit according to the information criteria is the $BBS$ fit, followed by $\mathcal{GBS}_2$. The nonnested hypothesis test statistic is $W_{ne} = 0.0229$, with $p$-value of 0.6543 under $H_f$ and with $p$-value of 0.0489 under $H_g$. Thus, there is substantial evidence that the $BBS$ distribution is the most adequate model for these data.

For the $BBS$ model, we tested unimodality versus bimodality. The hypotheses in the test were $H_0 : \gamma \geq 0$ and $H_1 : \gamma < 0$. The $p$-values of the SLR, SLR$_{c1}$, SLR$_{c2}$ and SLR$_{bp}$ tests were 0.0002, 0.0007, 0.0006 and 0.002, respectively. Hence, all tests reject $H_0$ in favor of $H_1$, which implies that there is strong evidence in favor of $\gamma < 0$, i.e., that the data came from a bimodal $BBS$ law.
8. Concluding remarks

Optimization processes may often fail to reach convergence when used to obtain maximum likelihood estimates of the parameters that index the BBS model, an interesting extension of the well known Birnbaum-Saunders model that may display bimodality. A penalization of the log-likelihood function that uses the Jeffreys prior was proposed. Alternative strategies to circumvent the problem were also explored. Overall, the best results were obtained using a penalized log-likelihood function based on a modified version of the Jeffreys prior.

We also considered hypothesis testing inference. For that, we used the BBS log-likelihood function penalized using the modified Jeffreys prior. The likelihood ratio, score and Wald tests were shown to be liberal in small samples, the Wald test being the worst performer. We have also shown that their bootstrap variants are typically quite accurate. One-sided tests based on the signed likelihood ratio statistic were also considered. We derived analytical corrections to the test statistic and also used bootstrap resampling. Overall, the analytically corrected tests displayed superior performance. We also developed tests for distinguishing between nonnested models. Our interest lied in distinguishing between the BBS model and an alternative version of the Birnbaum-Saunders distribution that also exhibits bimodality. Since in this case there are two distributions, the test was performed by considering two null hypotheses. It was shown that bootstrap-based nonnested testing inference can be quite accurate.

Three empirical applications were presented and discussed. In the first application, it was not possible to obtain the BBS maximum likelihood point estimates since it was not possible to maximize the log-likelihood function. Parameter estimates were easily obtained when the penalized log-likelihood function proposed in this paper was used. Other two applications were presented. In one of them, the BBS model was selected as the best model and it was shown that there was substantial evidence that the true data generating process is bimodal.
Acknowledgements

We gratefully acknowledge partial financial support from CAPES and CNPq. We also thank two anonymous referees for comments and suggestions.

References

[1] A. Azzalini, R.B. Arellano-Valle, Maximum penalized likelihood estimation for skew-normal and skew-t distributions, J. Stat. Plan. Infer. 143 (2013) 419–433.

[2] N. Balakrishnan, R.C. Gupta, D. Kundu, V. Leiva, A. Sanhueza, On some mixture models based on the Birnbaum-Saunders distribution and associated inference, J. Stat. Plann. Infer. 141 (2011) 2175–2190.

[3] N. Balakrishnan, V. Leiva, A. Sanhueza, F. Vilca, Estimation in the Birnbaum–Saunders distribution based on scale-mixture of normals and the EM-algorithm, SORT 33 (2009) 171–192.

[4] O.E. Barndorff-Nielsen, Inference on full or partial parameters based on the standardized signed log likelihood ratio, Biometrika 73 (1986) 307–322.

[5] O.E. Barndorff-Nielsen, Modified signed log likelihood ratio, Biometrika 78 (1991) 557–563.

[6] Z.W. Birnbaum, S.C. Saunders, Estimation for a family of life distributions with applications to fatigue, J. Appl. Probab. 6 (1969) 328–347.

[7] Z.W. Birnbaum, S.C. Saunders, A new family of life distributions, J. Appl. Probab. 6 (1969) 319–327.

[8] M. Bourguignon, R.B. Silva, G.M. Cordeiro, A new class of fatigue life distributions, J. Stat. Comput. Simul. 84 (2014) 2619–2635.
[9] G.M. Cordeiro, F. Cribari-Neto, An Introduction to Bartlett Correction and Bias Reduction, Springer, New York, 2014.

[10] G.M. Cordeiro, A.J. Lemonte, The exponentiated generalized Birnbaum–Saunders distribution, Appl. Math. Comput. 247 (2014) 762–779.

[11] G.M. Cordeiro, A.J. Lemonte, E.M. Ortega, An extended fatigue life distribution, Statistics 47 (2013) 626–653.

[12] D.R. Cox, Tests of separate families of hypotheses, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, volume 1, Birckbeck College, London, 1961, pp. 105–123.

[13] D.R. Cox, Further results on tests of separate families of hypotheses, J. Roy. Stat. Soc. B 24 (1962) 406–424.

[14] F. Cribari-Neto, A.C. Frery, M.F. Silva, Improved estimation of clutter properties in speckled imagery, Comput. Stat. Data Anal. 40 (2002) 801–824.

[15] A.H. Cysneiros, F. Cribari-Neto, C.A. Araújo, On Birnbaum-Saunders inference, Comput. Stat. Data Anal. 52 (2008) 4939–4950.

[16] A.C. Davison, D.V. Hinkley, Bootstrap Methods and Their Application, Cambridge University Press, New York, 1997.

[17] J.A. Díaz-García, J.R. Domnguez-Molina, Some generalisations of Birnbaum-Saunders and sinh-normal distributions, Int. Math. Forum 1 (2006) 1709–1727.

[18] J.A. Díaz-García, V. Leiva, A new family of life distributions based on the elliptically contoured distributions, J. Stat. Plan. Infer. 128 (2005) 445–457.

[19] T.J. DiCiccio, M.A. Martin, Simple modifications for signed roots of likelihood ratio statistics, J. Roy. Stat. Soc. B 55 (1993) 305–316.
[20] J.A. Doornik, An Object-Oriented Matrix Programming Language Ox 6., Timberlake Consultants Press, London, 2009.

[21] B. Efron, More efficient bootstrap computations, J. Am. Stat. Assoc. 85 (1990) 79–89.

[22] M.R. Ehsani, H. Saadatmanesh, S. Tao, Design recommendations for bond of GFRP rebars to concrete, J. Struct. Eng. 122 (1996) 247–254.

[23] S.L. Ferrari, E.C. Pinheiro, Small-sample one-sided testing in extreme value regression models, Adv. Stat. Anal. 100 (2016) 79–97.

[24] D. Firth, Bias reduction of maximum likelihood estimates, Biometrika 80 (1993) 27–38.

[25] J. Folks, R. Chhikara, The inverse Gaussian distribution and its statistical application—a review, J. Roy. Stat. Soc. B 40 (1978) 263–289.

[26] D.A.S. Fraser, N. Reid, J. Wu, A simple general formula for tail probabilities for frequentist and Bayesian inference, Biometrika 86 (1999) 249–264.

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

[28] V. Leiva, The Birnbaum-Saunders Distribution, Academic Press, London, 2015.

[29] V. Leiva, A. Sanhuez, S. Kotz, N. Araneda, A unified mixture model based on the inverse Gaussian distribution, Pak. J. Stat. 26 (2010) 445–460.

[30] V. Leiva, M. Tejo, P. Guiraud, O. Schmachtenberg, P. Orio, F. Marmolejo-Ramos, Modeling neural activity with cumulative damage distributions, Biol. Cybern. 109 (2015) 421–433.

[31] A.J. Lemonte, The Gradient Test: Another Likelihood-Based Test, Academic Press, New York, 2016.
[32] A.J. Lemonte, F. Cribari-Neto, K.L.P. Vasconcellos, Improved statistical inference for the two-parameter Birnbaum–Saunders distribution, Comput. Stat. Data Anal. 51 (2007) 4656–4681.

[33] A.J. Lemonte, S.L. Ferrari, Signed likelihood ratio tests in the Birnbaum–Saunders regression model, J. Stat. Plan. Infer. 141 (2011) 1031–1040.

[34] A.J. Lemonte, A.B. Simas, F. Cribari-Neto, Bootstrap-based improved estimators for the two-parameter Birnbaum–Saunders distribution, J. Stat. Comput. Simul. 78 (2008) 37–49.

[35] F. Lewis, A. Butler, L. Gilbert, A unified approach to model selection using the likelihood ratio test, Methods Ecol. Evol. 2 (2011) 155–162.

[36] B. Liseo, La classe delle densita normali sghembe: aspetti inferenziali da un punto di vista Bayesiano, Statistica 50 (1990) 71–79.

[37] R.C. Mittelhammer, G.G. Judge, D.J. Miller, Econometric Foundations, Cambridge University Press, New York, 2000.

[38] H. Ng, D. Kundu, N. Balakrishnan, Modified moment estimation for the two-parameter Birnbaum–Saunders distribution, Comput. Stat. Data Anal. 43 (2003) 283–298.

[39] N. Olmos, G. Martinez-Florez, H. Bolfarine, Bimodal Birnbaum-Saunders distribution with applications to non-negative measurements, Commun. Stat.-Theor. Methods 46 (2017) 6240–6257.

[40] W.J. Owen, A new three-parameter extension to the Birnbaum–Saunders distribution, IEEE Trans. Reliab. 55 (2006) 475–479.

[41] W.J. Owen, H.K.T. Ng, Revisit of relationships and models for the Birnbaum–Saunders and inverse-Gaussian distributions, J. Stat. Distrib. Appl. 2 (2015) 1–23.

[42] A.G. Patriota, On scale-mixture Birnbaum–Saunders distributions, J. Stat. Plan. Infer. 142 (2012) 2221–2226.
[43] D.M. Pianto, F. Cribari-Neto, Dealing with monotone likelihood in a model for speckled data, Comput. Stat. Data Anal. 55 (2011) 1394–1409.

[44] D.M. Rocke, Bootstrap Bartlett adjustment in seemingly unrelated regression, J. Am. Stat. Assoc. 84 (1989) 598–601.

[45] A. Sanhueza, V. Leiva, N. Balakrishnan, The generalized Birnbaum–Saunders distribution and its theory, methodology, and application, Commun. Stat.-Theor. Methods 37 (2008) 645–670.

[46] N. Sartori, Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions, J. Stat. Plan. Infer. 136 (2006) 4259–4275.

[47] T.A. Severini, An empirical adjustment to the likelihood ratio statistic, Biometrika 86 (1999) 235–247.

[48] T.A. Severini, Likelihood Methods in Statistics, Oxford University Press, New York, 2000.

[49] B. Smith, S. Wang, A. Wong, X. Zhou, A penalized likelihood approach to parameter estimation with integral reliability constraints, Entropy 17 (2015) 4040–4063.

[50] Q.H. Vuong, Likelihood ratio tests for model selection and non-nested hypotheses, Econometrica 57 (1989) 307–333.

[51] D.A. Williams, Discrimination between regression models to determine the pattern of enzyme synthesis in synchronous cell cultures, Biometrics 26 (1970) 23–32.

[52] J. Wu, A.C. Wong, Improved interval estimation for the two-parameter Birnbaum–Saunders distribution, Comput. Stat. Data Anal. 47 (2004) 809–821.

[53] X. Zhu, N. Balakrishnan, Birnbaum–Saunders distribution based on Laplace kernel and some properties and inferential issues, Stat. Probab. Lett. 101 (2015) 1–10.