CHILEAN JOURNAL OF STATISTICS

Edited by Víctor Leiva

Volume 10 Number 1
April 2019
ISSN: 0718-7912 (print)
ISSN: 0718-7920 (online)

Published by the
Chilean Statistical Society

SOCHE
SOCIEDAD CHILENA DE ESTADÍSTICA

Chilean Journal of Statistics
Vol. 10, No. 1, April 2019, 55–76

Distribution Theory
Research Paper

A new Birnbaum-Saunders type distribution based on the skew-normal model under a centered parameterization

Nathalia L. Chaves[1], Caio L. N. Azevedo[1,*], Filidor Vilca-Labra[1], and Juvêncio S. Nobre[2]

[1]Department of Statistics, Universidade Estadual de Campinas, Brazil
[2]Department of Statistics and Applied Mathematics, Universidade Federal de Ceará, Brazil

(Received: 22 February 2019 · Accepted in final form: 23 April 2019)

Abstract

In this paper, we introduce a new distribution for positively skewed data by combining the Birnbaum-Saunders and centered skew-normal distributions. Several of its properties are developed. Our model accommodates both positively and negatively skewed data. Also, we show that our proposal circumvents some problems related to another Birnbaum-Saunders distribution based on the usual skew-normal model, previously presented in the literature. We derive both maximum likelihood and Bayesian inference, comparing them through a suitable simulation study. The convergence of the expectation conditional maximization (for maximum likelihood inference) and MCMC algorithms (for Bayesian inference) are verified and several factors of interest are compared. In general, as the sample size increases, the results indicate that the Bayesian approach provided the most accurate estimates. Our model accommodates the asymmetry of the data more properly than the usual Birnbaum-Saunders distribution, which is illustrated through real data analysis.

Keywords: Bayesian inference · Birnbaum-Saunders distribution · ECM algorithm · Frequentist inference · MCMC algorithms · R software

Mathematics Subject Classification: Primary 60E05 · Secondary 62F15.

1. Introduction

The Birnbaum-Saunders (BS) distribution is characterized by two parameters and defined in terms of the standard normal distribution. The BS distribution has been received considerable attention over the past few years, since it has been used quite effectively to model positively skewed data, especially lifetime and crack growth data. Since the pioneering work of Birnbaum and Saunders (1969a) was published, several extensions of the BS distribution have been proposed in the literature and its parameters estimated under both frequentist and Bayesian paradigms.

*Corresponding author. Email: cnaber@ime.unicamp.br

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

From a frequentist viewpoint, Birnbaum and Saunders (1969b) presented a discussion on the maximum likelihood (ML) parameter estimation. Mann et al. (1974) showed that the BS distribution is unimodal. Engelhardt et al. (1981) developed confidence intervals and hypothesis tests for both parameters. Desmond (1985) developed a BS-type distribution based on a biological model. Desmond (1986) investigated the relationship between the BS distribution and the inverse Gaussian distribution. Lu and Chang (1997) used bootstrap methods to construct prediction intervals for future observations. In the linear regression context, Rieck and Nedelman (1991) developed a related log-linear model and showed that it can be used for modeling accelerated life tests and to compare average lifetime of different populations.

From a Bayesian perspective, there are few works on the BS distribution. The first one is due to Achcar (1993) who developed Bayesian estimation using numerical approximations for the marginal posterior distributions of interest based on the Laplace approximation. Also, Xu and Tang (2011) presented a Bayesian study with partial information, while Wang et al. (2016) assumed that the two parameters follow mutually independently inverse gamma distributions. All these results were studied under a normal distribution for generating the BS distribution.

In terms of modeling, most of the generalizations of the BS distribution are based on elliptical and skew-elliptical laws, in order to obtain more robust and flexible models. Some works developed extensions based on symmetric distributions as Diaz-Garcia and Leiva (2005) who generalized the BS model using elliptical distributions that includes the Cauchy, Laplace, logistic, normal and Student-t distributions as particular cases. Other works are: the generalized BS distribution (Leiva et al., 2007), the Student-t BS distribution (Barros et al., 2008), and the scale-mixture of normal BS distribution (Balakrishnan et al., 2009), among others. More information can be found in Leiva (2016), who presented a review on the BS distribution. Other generalizations have been obtained in different ways to those aforementioned, as Owen and Padgett (1999), who developed a three-parameter BS distribution and the β-BS distribution presented in Cordeiro and Lemonte (2011). Also, Ferreira (2013) proposed a based BS distribution useful for modeling tail events and Mazucheli et al. (2018) presented a distribution on the unit interval based on the BS model. In addition, Balakrishnan et al. (2017) and Maehara (2018) provided new families of BS distribution based on the skew scale mixture of normal models. Also, extensions of the BS distribution based on the skew-elliptical models can be found in Vilca and Leiva (2006), Leiva et al. (2007, 2008) and Vilca et al. (2011). In these works, theoretical results were obtained, extending the properties of the BS and log-BS distributions.

A Bayesian perspective for the BS distributions based on skew-normal (SN) distribution did not receive much attention in the literature. Indeed, Vilca et al. (2011) considered, under a frequentist perspective, the BS distribution based on the SN model. However, even though the SN distribution has been applied with success in several fields, when the related asymmetry parameter is equals to zero, the associated Fisher information matrix is singular. Recently, to overcome this problem, Arellano and Azzalini (2008) and Azzalini (2013) explored a SN distribution under a convenient parameterization (proposed by Azzalini (1985) and deeper explored by Pewsey (2000)), the so-called centered parametrization (CP), which leads to a non-singular Fisher information matrix. Moreover, the relative profile log-likelihood function (RPLL) for the Pearson index of skewness exhibits a more regular behavior, closer to a quadratic function, and without a stationary point under null asymmetry . The resulting empirical distributions of the estimators under the CP, named CP estimators, are much closer to the normality than those obtained under the usual SN distribution, which is named direct parametrization estimators. All these desirable properties, related to the the CP, may be transferred to the respective BS distribution based on the centered SN (CSN) model. It is worthwhile to mention that all the aforementioned

BS models (that consider the SN model) used the direct parametrization that is, likely, they inherit the above problems.

The main objective of this work is to propose an alternative to the skew-normal BS (SNBS) distribution proposed by Vilca et al. (2011), considering the CSN distribution, as the generator variable. The resulting BS-type distribution has advantages, in inference terms, over the SNBS distributions (including those obtained as particular cases of the more general families as those of Balakrishnan et al. (2009) and Maehara (2018)), similarly to those related to the CSN distribution, compared with the SN distribution. The specific objectives of this work are: to develop a BS distribution based on the CSN model, named centered skew-normal BS (CSNBS) distribution, highlighting its advantages over the SNBS distribution proposed by Vilca et al. (2011), and its main properties. Also, estimation procedures under both frequentist and Bayesian approaches are developed and compared, considering different scenarios. In addition, some model comparison statistics are studied. Finally, two real data sets are analyzed showing some advantages of the CSNBS model compared to the usual BS distribution.

The paper is outlined as follows. In Section 2, we present our distribution and some motivation for its development. In Section 3, the estimation methods are proposed and some statistics of model comparison are presented. In Section 4, some simulation studies are presented and two real data sets are analyzed. Finally, in Section 5, some additional comments and conclusions are provided.

2. The Centered Skew-normal BS Distribution

2.1 The centered skew-normal distribution

A random variable Y is said to have a CSN distribution, denoted by $Y \sim \mathrm{CSN}(\mu, \sigma, \gamma)$, where μ, σ and γ are the mean, the standard deviation and the Pearson coefficient of skewness, respectively, if its density is given by

$$f_Y(y) = 2\frac{\sigma_z}{\sigma}\phi\Big(\mu_z + \frac{\sigma_z}{\sigma}(y-\mu)\Big)\Phi\Big[\lambda\Big(\mu_z + \frac{\sigma_z}{\sigma}(y-\mu)\Big)\Big]\frac{2}{\omega}\phi\Big(\frac{y-\xi}{\omega}\Big)\Phi\Big[\lambda\Big(\frac{y-\xi}{\omega}\Big)\Big], y \in \mathrm{R}, \tag{1}$$

where $\mu_z = r\delta$, $\sigma_z^2 = 1-\mu_z^2$, $\lambda = \gamma^{1/3}s/\sqrt{r^2 + s^2\gamma^{2/3}(r^2-1)}$, $r = \sqrt{2/\pi}$, $\gamma = r\delta^3(4/\pi - 1)(1-\mu_z^2)^{-3/2}$, $\delta = \lambda/\sqrt{1+\lambda^2}$, $\xi = \mu - \sigma\gamma^{1/3}s$, $\omega = \sigma\sqrt{1+\gamma^{2/3}s^2}$, and $s = [2/(4-\pi)]^{1/3}$. The quantity λ is the asymmetry parameter, see Azzalini (1985). For $\mu = 0$ and $\sigma = 1$, we have the standard CSN distribution, denoted by $Y \sim \mathrm{CSN}(0, 1, \gamma)$, whose density is given by

$$f_Y(y) = \frac{2}{\omega}\phi\Big(\frac{y-\xi}{\omega}\Big)\Phi\Big[\lambda\Big(\frac{y-\xi}{\omega}\Big)\Big], y \in \mathrm{R},$$

where $\xi = -\gamma^{1/3}s$ and $\omega = \sqrt{1+\gamma^{2/3}s^2}$. For inferential purposes, a useful stochastic representation of Y is given by

$$Y = \frac{1}{\sigma_z}\left\{\delta\,|X_0| + (1-\delta^2)X_1 - \mu_z\right\}, \tag{2}$$

where $X_i \sim \mathrm{N}(0,1)$, for $i = 0, 1$, are independent and so $H = |X_0|$ follows a half-normal (HN) distribution, denoted by $\mathrm{HN}(0,1)$.

2.2 The proposed distribution

Here, we present the CSNBS distribution, which is defined similarly to the usual BS and the SNBS distributions by

$$T = \beta\left[\frac{\alpha Y}{2} + \sqrt{\left(\frac{\alpha Y}{2}\right)^2 + 1}\right]^2, \tag{3}$$

where $Y \sim \mathrm{CSN}(0, 1, \gamma)$, α is the shape parameter, β is the location parameter, and γ is the asymmetry parameter. We use the following notation $T \sim \mathrm{CSNBS}(\alpha, \beta, \gamma)$. The vector $(\alpha, \beta, \gamma)^\top$ is called centered parameter and based on the SN distribution, that is, $(\alpha, \beta, \lambda)^\top$ is named direct parameters. Following the same steps as in the usual BS distribution, we have that its density is given by

$$f_T(t) = 2\phi\left[a_{t;\mu,\sigma}(\alpha, \beta)\right]\Phi\left[\lambda\, a_{t;\mu,\sigma}(\alpha, \beta)\right]A_{t;\sigma}(\alpha, \beta), t > 0, \tag{4}$$

where $a_{t;\mu,\sigma}(\alpha, \beta) = \mu_z + \sigma_z\, a_t(\alpha, \beta)$, $A_{t;\sigma}(\alpha, \beta) = \sigma_z A_t(\alpha, \beta)$, $a_t(\alpha, \beta) = \left(\sqrt{t/\beta} - \sqrt{\beta/t}\right)/\alpha$, $A_t(\alpha, \beta) = \mathrm{d}a_t(\alpha, \beta)/\mathrm{dt} = t^{-3/2}(t + \beta)/(2\alpha\beta^{1/2})$, and the other quantities are previously defined. Note that for $\gamma = 0$, we have the usual BS distribution. The mean and variance of T (see Appendix A for more details) are given, respectively, by

$$\mathrm{E}(T) = \beta\left(1 + \frac{\alpha^2}{2}\right) \quad \text{and} \quad \mathrm{Var}(T) = (\alpha\beta)^2\left\{1 + \frac{\alpha^2}{4}\Big[2\Delta - 1\Big]\right\},$$

where $\Delta = 2(\pi - 3)(4/\pi^2)\delta^4[1 - (2\delta^2/\pi)]^{-2} + 3$.

The following theorem is very useful to develop both classical and Bayesian approaches since they lead to conditional distributions that allow us to implement, more easily, the EM algorithm, and simplify the Bayesian developments. For the use of standard MCMC software, such as WinBUGS, OpenBUGS, JAGS or Stan, see Lunn et al. (2000), Lunn et al. (2009), Depaoli et al. (2016) and Carpenter et al. (2016).

Theorem 2.1 Let $T \sim \mathrm{CSNBS}(\alpha, \beta, \gamma)$ as in Equation (3), and Y and H as defined in Equation (2). Then,

(i) The conditional density of T, given $H = h$, can be expressed as

$$f_{T|H}(t|h) = \phi(\nu_h + a_t(\alpha_\delta, \beta))A_t(\alpha_\delta, \beta),$$

where $\alpha_\delta = \alpha\sqrt{(1 - \delta^2)/(1 - r^2\delta^2)}$ and $\nu_h = -(\delta(h - r))/\sqrt{1 - \delta^2}$.

(ii) $f_{H|T}(h|t) = \dfrac{\phi\left(h\,|\,\delta\sqrt{1-r^2\delta^2}\left(a_t(\alpha,\beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\right), 1-\delta^2\right)}{\Phi\left(\lambda\sigma_z\left(a_t(\alpha,\beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\right)\right)}$, $h > 0$. Moreover,

$$\mathrm{E}(H|T = t) = \eta_t + W_\Phi\left(\frac{\eta_t}{\tau}\right)\tau \quad \text{and} \quad \mathrm{E}(H^2|T = t) = \eta_t^2 + \tau^2 + W_\Phi\left(\frac{\eta_t}{\tau}\right)\eta_t\tau,$$

where $\eta_t = \delta\sqrt{1 - r^2\delta^2}\left(a_{t_i}(\alpha, \beta) + (r\delta)/\sqrt{1 - r^2\delta^2}\right)$.

The density in Theorem 2.1 corresponds to the extended Birnbaum-Saunders (EBS) discussed in Leiva et al. (2008) and denoted by $\mathrm{EBS}(\alpha_\delta, \beta, \sigma = 2, \nu_h)$. The proof of Theorem 2.1 is in the Appendix B.

Figures 1-3 present the density of the CSNBS distribution for different values of α, β and γ. From Figure 1, we have that for $\alpha = 0.2$ the density is concentrated around β ($\beta = 1$), and for $\alpha = 0.8$ the density is more asymmetric, with a higher variability. As α increases, the density becomes more flat, positively skewed and more dispersed, as it can be seen in Figure 2, for different values of α, fixing the other parameters. In addition, Figure 3 shows densities more concentrated around β for different values of α and β, with $\gamma = 0.9$. It is also possible to see that for large values of β, the density is more negatively skewed. Note that the distribution tends to be symmetric around β, for $\gamma = 0$ (the usual BS distribution) and/or for small values of α. Positive asymmetry is observed as α increases, β decreases and/or γ assumes positive values. In addition, negative asymmetry is observed as α decreases, β increases and/or γ assumes negative values. Another interesting point is that the CSNBS distribution may be negatively skewed, which is an unusual behavior for positive random variables. This feature makes our distribution a very useful alternative for modeling positive skewed data.

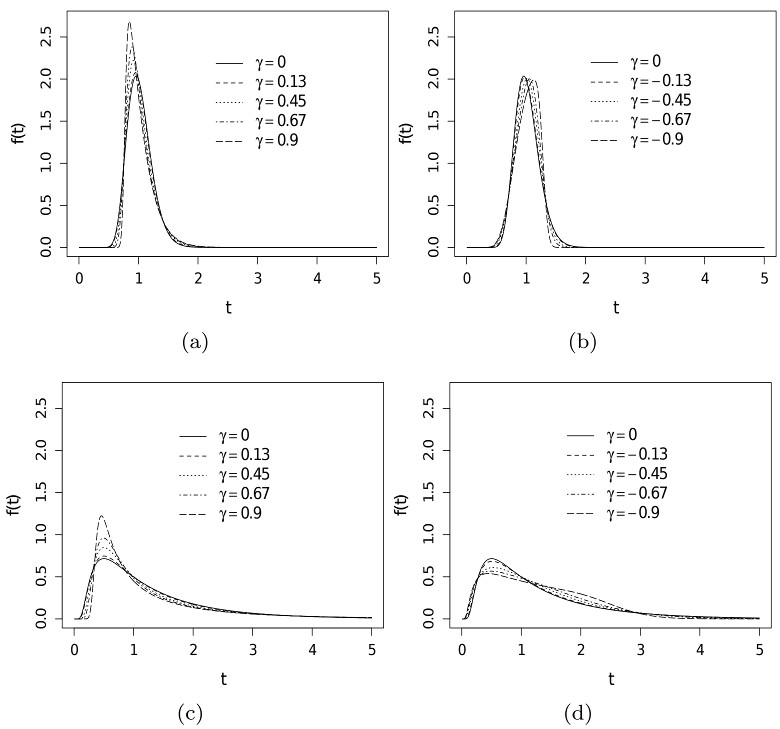

Figure 1. CSNBS density for different values of γ, with $\beta = 1$, $\alpha = 0.2$ (a)-(b) and $\alpha = 0.8$ (c)-(d).

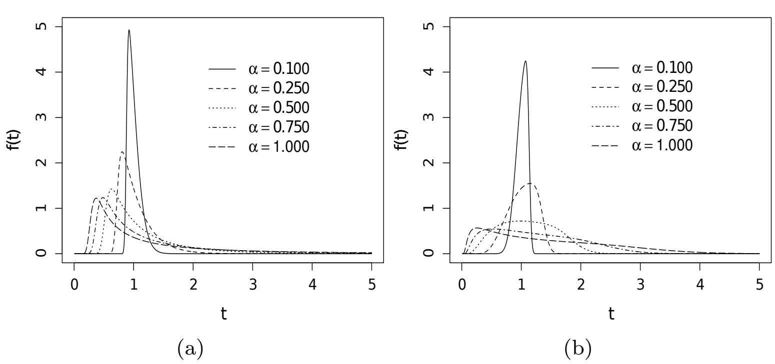

Figure 2. CSNBS density for different values of α, with $\beta = 1$, $\gamma = 0.9$ (a) and $\gamma = -0.9$ (b).

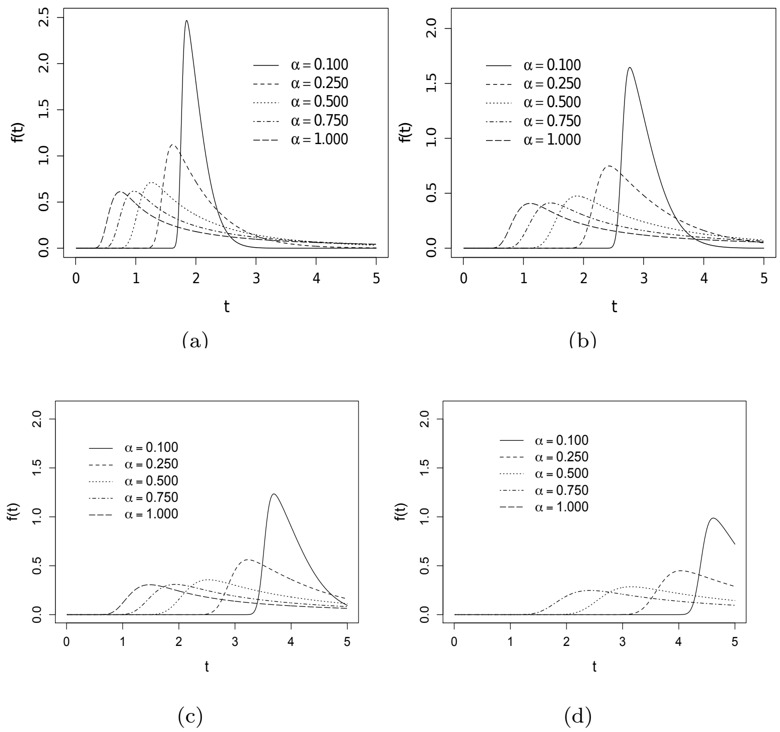

Figure 3. CSNBS density with $\beta = 2$ (a), $\beta = 3$ (b), $\beta = 4$ (c), and $\beta = 5$ (d) for indicated α and $\gamma = 0.9$.

2.3 Some motivational remarks on the proposal

(i) It is well known that there is some difficulty in estimating the parameters of the SN distribution by the ML approach, when the asymmetry parameter is close to zero. Some problems seem to persist even if one switched to the Bayesian inference, unless a strongly informative prior is considered, as pointed out by Arellano and Azzalini (2008). The SNBS distribution seems to inherit such problems. Thus, the proposed CSNBS distribution can circumvents these problems, since it is based on the CSN model.

(ii) When the asymmetry parameter is equals to zero, the Fisher information matrix is singular, even if all parameters are identifiable. This affects the behavior of the empirical distributions of the ML estimators and the Bayesian estimators. To get a direct perception of the problem, we generated 100 samples of size $n = 200$, from the SNBS distribution and for each sample, the ML and Bayesian estimates $(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})$ have been computed. In this case, we fix $\alpha = 0.5$, $\beta = 1$ and $\lambda = 1$, which induces a strong positively skewed behavior of the SNBS distribution. Figures 4 and 5 display the corresponding empirical distribution of $\widehat{\alpha}$ and $(\widehat{\alpha}, \widehat{\beta})$, through a histogram and scatter plot, respectively. Moreover, 100 samples of size $n = 200$ are generated from the CSNBS distribution, and the respective ML and Bayesian estimates $(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma})$ have been computed. In this case, we fix $\alpha = 0.5$, $\beta = 1$ and $\gamma = 0.137$, which induces a strong positively skewed behavior of the CSNBS model. The empirical distributions of $\widehat{\alpha}$ and $(\widehat{\alpha}, \widehat{\beta})$ are shown in Figures 6 and 7, respectively. Clearly these empirical distributions are much closer to normality than those in Figures 4 and 5. In fact, it can be shown that the singularity of the expected Fisher information matrix, when the asymmetry parameter is null, no longer occurs.

(iii) The CP circumvents the problem concerning the existence of an inflection point in the RPLL of this parameter. This can be seen in Figure 8, which refers to the plots of twice the RPLL function for λ, the asymmetry parameter of the SNBS distribution (left panel), and the for γ, the asymmetry parameter of the CSNBS distribution (right panel). The RPLL corresponds to $\ell(\widehat{\alpha}(\gamma), \widehat{\beta}(\gamma), \gamma) - \ell(\widehat{\alpha}(\gamma), \widehat{\beta}(\gamma), \widehat{\gamma})$, where $\ell(\cdot)$ represents the log-likelihood function. The respective plots are constructed by considering a random sample of both SNBS and CSBNS distributions, under suitable values of α, β and γ. We can notice a non-quadratic form of the log-likelihood function under the SNBS model, induced by the existence of an inflection point when the asymmetry parameter is very close to zero, making it difficult the obtaining of the ML estimates. However, the log-likelihood function of the CSNBS distribution presents a concave shape. Also, there is no inflection point when the asymmetry parameter is equals zero.

(iv) The posterior distribution of λ for the SNBS distribution has a non-quadratic form, as it can be seen in Figure 9 (a), and this occurs even if we consider an informative prior. However, the posterior distribution of γ for the CSNBS distribution is well-behaved, presenting a concave shape, as it can be seen in Figure 9 (b).

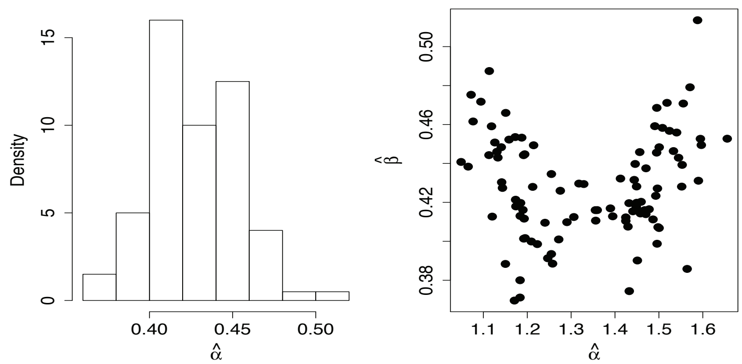

Figure 4. Estimated distributions of the ML estimates when samples of size $n = 200$ are drawn from SNBS; the left panel displays the histogram of $\widehat{\alpha}$, the right panel displays the scatter plot of $(\widehat{\alpha}, \widehat{\beta})$.

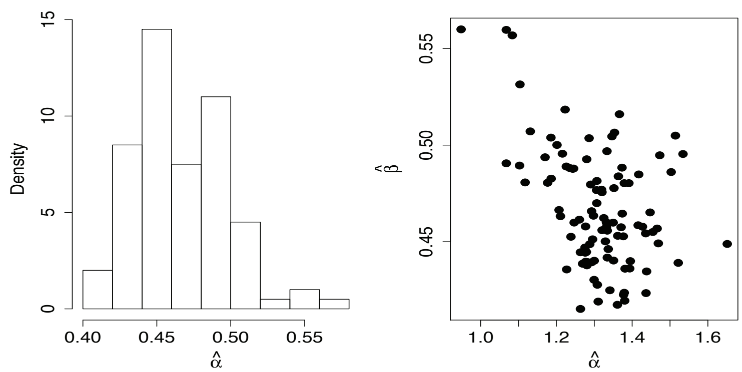

Figure 5. Estimated distributions of the Bayesian estimates when samples of size $n = 200$ are drawn from SNBS distribution; the left panel displays the histogram of $\widehat{\alpha}$, the right panel displays the scatter plot of $(\widehat{\alpha}, \widehat{\beta})$.

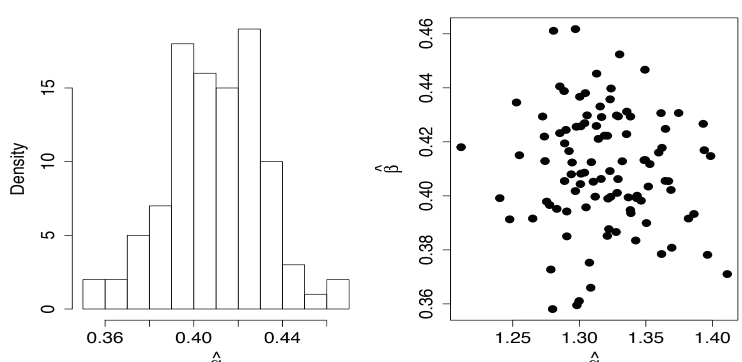

Figure 6. Estimated distributions of the ML estimates when samples of size $n = 200$ are drawn from CSNBS distribution; the left panel displays the histogram of $\widehat{\alpha}$, the right panel displays the scatter plot of $(\widehat{\alpha}, \widehat{\beta})$.

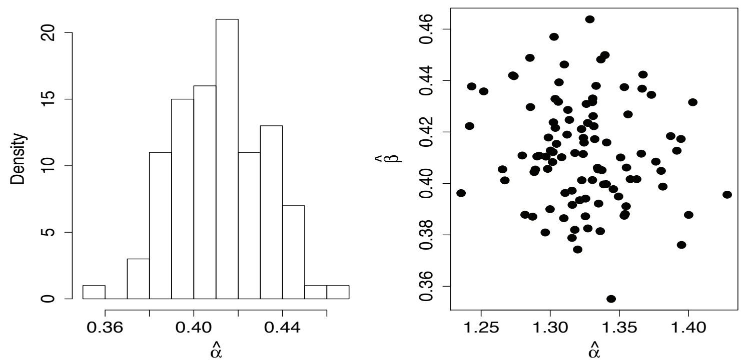

Figure 7. Estimated distributions of the Bayesian estimates when samples of size $n = 200$ are drawn from CSNBS distribution; histogram of $\widehat{\alpha}$ (left) and scatter plot of $(\widehat{\alpha}, \widehat{\beta})$ (right).

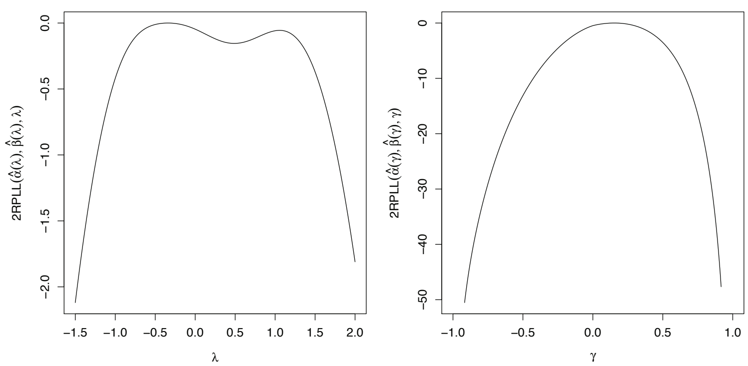

Figure 8. Twice the relative profiled log-likelihood function for the asymmetry parameter of the SNBS (left) and CSNBS (right) distributions.

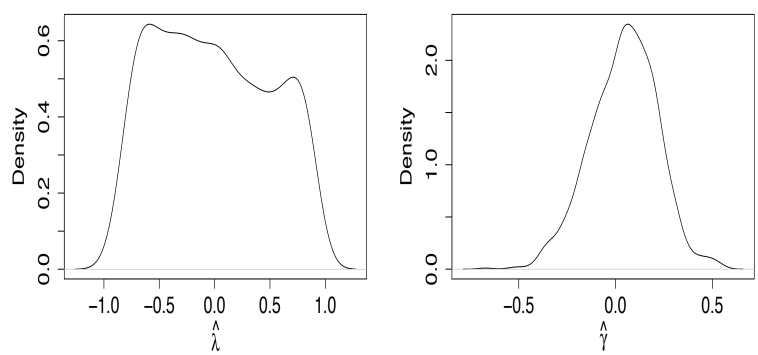

Figure 9. Posterior distribution of λ for the SNBS distribution (left) and of γ for the CSNBS distribution (right).

3. Estimation and Inference

3.1 General context

We present the ML estimation, based on the expectation conditional maximization (ECM) algorithm as in Meng and Rubin (1993), and the Bayesian approach, through MCMC algorithms. Let $T \sim \mathrm{CSNBS}(\alpha, \beta, \gamma)$ and then, recall that, from Theorem 2.1, we have $T|(H = h) \sim \mathrm{EBS}(\alpha_\delta, \beta, \sigma = 2, \nu_h)$, where $H = |X_0| \sim \mathrm{HN}(0, 1)$, $\alpha_\delta = \alpha\sqrt{(1-\delta^2)/(1-r^2\delta^2)}$ and $\nu_h = -\delta(h-r)/\sqrt{1-\delta^2}$. In Appendix B, we present some results that are useful for obtaining the ML estimators. For both methods, we consider a random sample $T_1, \ldots, T_n$ from $T \sim \mathrm{SNBS}(\alpha, \beta, \gamma)$, where $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$.

3.2 The ECM algorithm and ML estimation

Here, we discuss the ML estimation through the ECM algorithm. The log-likelihood function for $\boldsymbol{\theta}$ is given by $\ell(\boldsymbol{\theta}|\boldsymbol{t}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}|t_i)$, where

$$\ell_i(\boldsymbol{\theta}|t_i) = \log(2) + \log\left\{\phi\left[a_{t_i;\mu,\sigma}(\alpha,\beta)\right]\right\} + \log\left\{\Phi\left[\lambda\, a_{t_i;\mu,\sigma}(\alpha,\beta)\right]\right\} + \log\left[A_{t_i;\sigma}(\alpha,\beta)\right], \quad (5)$$

and $a_{t_i;\mu,\sigma}(\alpha,\beta)$ and $A_{t_i;\sigma}(\alpha,\beta)$ are given in Equation (4). Instead of considering the direct maximization of Equation (5), we obtain the ML estimates through the ECM algorithm, since it allows for a more tractable optimization process. In this case, we need to work with the so-called augmented likelihood function. Also, instead of working with $\boldsymbol{\theta}^* = (\alpha,\beta,\gamma)^\top$, we estimate $\boldsymbol{\theta} = (\alpha,\beta,\delta)^\top$, where δ is defined in Equation (1). Then, we recover γ through the invariance principle related to the ML estimators. This is performed since the related expressions (both analytically and computationally) are more tractable for $\boldsymbol{\theta}$.

Recall that, From Theorem 2.1, we have $T_i|H_i = h_i \overset{\text{IND}}{\sim} \text{EBS}(\alpha_\delta,\beta,\sigma=2,\nu_{h_i})$ and $H_i \overset{\text{IND}}{\sim} \text{HN}(0,1); i=1,\ldots,n$, where "IND" denotes "independent", $\alpha_\delta = \alpha\sqrt{(1-\delta^2)/(1-r^2\delta^2)}$ and $\nu_{h_i} = -(\delta(h_i - r))\sqrt{1-\delta^2}$. Then, defining $\boldsymbol{t}_c = (\boldsymbol{t}^\top,\boldsymbol{h}^\top)^\top$, with $\boldsymbol{t} = (t_1,\ldots,t_n)^\top$ and $\boldsymbol{h} = (h_1,\ldots,h_n)^\top$, the augmented log-likelihood function can be written as

$$\begin{aligned}
\ell(\boldsymbol{\theta}|\boldsymbol{t}_c) &= \sum_{i=1}^{n}\log[f_{T|H}(t_i|h_i)] + \sum_{i=1}^{n} f_H(h_i) \\
&= c - \frac{\delta^2}{2(1-\delta^2)}\sum_{i=1}^{n} h_i^2 + \frac{r\delta^2}{(1-\delta^2)}\sum_{i=1}^{n} h_i - \frac{nr^2\delta^2}{2(1-\delta^2)} \\
&\quad + \frac{\delta\sqrt{1-r^2\delta^2}}{1-\delta^2}\sum_{i=1}^{n} h_i a_{t_i}(\alpha,\beta) - \frac{r\delta\sqrt{1-r^2\delta^2}}{1-\delta^2}\sum_{i=1}^{n} a_{t_i}(\alpha,\beta) - \frac{1-r^2\delta^2}{2(1-\delta^2)}\sum_{i=1}^{n} a_{t_i}^2(\alpha,\beta) \\
&\quad + \frac{n}{2}\log(1-r^2\delta^2) + \sum_{i=1}^{n}\log(t_i+\beta) - \frac{n}{2}\log(1-\delta^2) - n\log(\alpha) - \frac{n}{2}\log(\beta).
\end{aligned}$$

For a current value of $\boldsymbol{\theta}$, say $\widehat{\boldsymbol{\theta}}$, the E-step requires the evaluation of $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \text{E}[\ell(\boldsymbol{\theta}|\boldsymbol{t}_c)|\boldsymbol{t},\widehat{\boldsymbol{\theta}}]$, where the expectation is taken with respect to the conditional distribution $H|(T=t)$, evaluated at $\widehat{\boldsymbol{\theta}}$. For a estimate of $\boldsymbol{\theta}$ at r-th iteration, say $\boldsymbol{\theta}^{(r)} = (\alpha^{(r)},\beta^{(r)},\delta^{(r)})^\top$, consider $\widehat{h}_i = \text{E}[H_i|\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},t_i]$ and $\widehat{h}_i^2 = \text{E}[H_i^2|\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},t_i]$, given in Theorem 2.1, that is,

$$\widehat{h}_i = \widehat{\eta}_{t_i} + W_\Phi\left(\frac{\widehat{\eta}_{t_i}}{\widehat{\tau}}\right)\widehat{\tau} \quad \text{and} \quad \widehat{h}_i^2 = \widehat{\eta}_{t_i}^2 + \widehat{\tau}^2 + W_\Phi\left(\frac{\widehat{\eta}_{t_i}}{\widehat{\tau}}\right)(\widehat{\eta}_{t_i}\widehat{\tau}), \quad (6)$$

respectively, where $\widehat{\eta}_{t_i} = \widehat{\delta}\sqrt{1-r^2\widehat{\delta}^2}\left(a_{t_i}(\widehat{\alpha},\widehat{\beta}) + r\widehat{\delta}/\sqrt{1-r^2\widehat{\delta}^2}\right)$, $\widehat{\tau} = \sqrt{1-\widehat{\delta}^2}$ and $W_\Phi(z) = \phi(z)/\Phi(z)$, $z\in\text{R}$. Then, let $\boldsymbol{\theta}^{(r)} = (\alpha^{(r)},\beta^{(r)},\delta^{(r)})^\top$ be the estimate of $\boldsymbol{\theta}$ at the k-th iteration. By considering Equation (6), we have that the augmented log-likelihood function becomes $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = \text{E}[\ell(\boldsymbol{\theta}|\boldsymbol{t}_c)|\mathbf{t},\widehat{\boldsymbol{\theta}}^{(r)}]$, where

$$
\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = {} & c - \frac{\delta^{2(r)}}{2\left(1-\delta^{2(r)}\right)}\sum_{i=1}^{n}\widehat{h}_i^{2(r)} + \frac{r\delta^{2(r)}}{\left(1-\delta^{2(r)}\right)}\sum_{i=1}^{n}\widehat{h}_i^{(r)} - \frac{nr^2\delta^{2(r)}}{2\delta^{2(r)}} \\
& + \frac{\delta^{(r)}\sqrt{1-r^2\delta^{2(r)}}}{\alpha^{(r)}\left(\delta^{2(r)}\right)}\sum_{i=1}^{n}\widehat{h}_i^{(r)} a_{t_i}(1,\beta^{(r)}) - \frac{r\delta^{(r)}\sqrt{1-r^2\delta^{2(r)}}}{\alpha^{(r)}\left(1-\delta^{2(r)}\right)} \\
& \times \sum_{i=1}^{n} a_{t_i}(1,\beta^{(r)}) - \frac{1-r^2\delta^{2(r)}}{2\alpha^{2(r)}\left(1-\delta^{2(r)}\right)}\sum_{i=1}^{n}\left[a_{t_i}(1,\beta^{(r)})\right]^2 + \frac{n}{2}\log\left(1-\delta^{2(r)}\right) \\
& + \sum_{i=1}^{n}\log\left(t_i+\beta^{(r)}\right) - \frac{n}{2}\log\left(1-\delta^{2(r)}\right) - n\log\left(\alpha^{(r)}\right) - \frac{n}{2}\log\left(\beta^{(r)}\right).
\end{aligned}
$$

Hence, the ECM algorithm corresponds to iterate the following steps:

E-step: Given $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(r)}$, compute $\widehat{h}_i$ and $\widehat{h}_i^2$, for $i = 1,\ldots,n$ by using Equation (6);

CM-step 1: Fix $\beta = \widehat{\beta}^{(r)}$ and $\delta = \widehat{\delta}^{(r)}$ and update $\widehat{\alpha}^{(r)}$ through the positive root of $\widehat{\alpha}^2 + \hat{b}^{(r)}\widehat{\alpha} + \hat{c}^{(r)} = 0$, where

$$
\hat{b}^{(r)} = \frac{1}{(1-\widehat{\delta}^{2(r)})}\left[\widehat{\delta}^{(r)}\sqrt{1-r^2\widehat{\delta}^{2(r)}}\frac{1}{n}\sum_{i=1}^{n}\widehat{h}_i a_{t_i}(1,\widehat{\beta}^{(r)}) - r\widehat{\delta}^{(r)}\sqrt{1-r^2\widehat{\delta}^{2(r)}}\frac{1}{n}\sum_{i=1}^{n}a_{t_i}(1,\widehat{\beta}^{(r)})\right].
$$

$$
\hat{c}^{(r)} = -\frac{(1-r^2\widehat{\delta}^{2(r)})}{(1-\widehat{\delta}^{2(r)})}\frac{1}{n}\sum_{i=1}^{n}\widehat{h}_i\left[a_{t_i}(1,\widehat{\beta}^{(r)})\right]^2,
$$

that is, $\widehat{\alpha}^{(r+1)} = (-b(r+1) + \sqrt{b^{2(r+1)} - 4c(r+1)})/2$;
CM-step 2: Fix $\alpha = \widehat{\alpha}^{(r+1)}$ and update $\widehat{\beta}^{(r)}$ and $\widehat{\delta}^{(r)}$ using

$$
\widehat{\beta}^{(r+1)} = \underset{\beta}{\operatorname{argmax}}\, Q\left(\widehat{\alpha}^{(r+1)},\beta,\widehat{\delta}^{(r)}\right) \quad \text{and } \widehat{\delta}^{(r+1)} = \underset{\delta}{\operatorname{argmax}}\, Q\left(\widehat{\alpha}^{(r+1)},\widehat{\beta}^{(r+1)},\delta\right).
$$

The updating of $\widehat{\beta}^{(r+1)}$ and $\widehat{\delta}^{(r+1)}$ needs to be done through some numerical optimization method. In this work we use the function `optim`, available on the `R` software (see R Development Core Team, 2017), considering the L-BFGS-B optimization algorithm (see Byrd et al., 1995)). Also, we start the ECM algorithm with initial values, say, $\widehat{\alpha}^{(0)}$, $\widehat{\beta}^{(0)}$ and $\widehat{\delta}^{(0)}$, using: $\widehat{\alpha}^{(0)} = [2(s/v)-1]^{1/2}$ and $\widehat{\beta}^{(0)} = (sv)^{1/2}$, where $s = (1/n)\sum_{i=1}^{n} t_i$ and $v = [(1/n)\sum_{i=1}^{n} 1/t_i]^{-1}$, as in Vilca et al. (2011). After obtaining $\widehat{\alpha}^{(0)}$ and $\widehat{\beta}^{(0)}$, we can define $z_i = (1/\widehat{\alpha}^{(0)})[(t_i/\widehat{\beta}^{(0)})^{1/2} - (\widehat{\beta}^{(0)}/t_i)^{1/2}]$, for $i = 1,\ldots,n$, which are observations related to a CSN distribution. Thus, $\widehat{\delta}^{(0)}$ can be obtained by maximizing (numerically) the log-likelihood function of a SN distribution with respect to δ, which is given by

$$
\ell(\theta) = \sum_{i=1}^{n}\Big[\log(2) + \log(\sigma_z) + \log\left[\phi\left(\mu_z + \sigma_z y_i\right)\right] + \log\left\{\Phi\left[\lambda(\mu_z + \sigma_z y_i)\right]\right\}\Big].
$$

According to Vilca et al. (2011), for ensuring that the true ML estimates are obtained, it is recommended to run the ECM algorithm using a range of different starting values and checking whether all of them result in similar estimates. The steps of the

ECM algorithm are repeated until a suitable convergence is attained, for example, using $\|\boldsymbol{\theta}^{(r)} - \boldsymbol{\theta}^{(r-1)}\| < \varepsilon$, with $\varepsilon > 0$. It is worthwhile to mention, under certain regularity conditions, that $\widehat{\boldsymbol{\theta}}$ converges in distribution to $\mathrm{N}_3(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}})$, as $n \to \infty$. We approximate $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ by $I^{-1}(\boldsymbol{\theta})$, where $I(\boldsymbol{\theta}) = -\ddot{\ell}$, $\ddot{\ell} = [\ddot{\ell}_{\theta_1\theta_2}]$, $\theta_1, \theta_2 = \alpha, \beta$ or γ is the Hessian matrix, and $\ddot{\ell}_{\theta_1\theta_2} = \ddot{\ell}_{\theta_2\theta_1} = \partial^2\ell(\boldsymbol{\theta})/\partial\theta_1\theta_2 = \sum_{i=1}^{n} \partial^2\ell_i(\boldsymbol{\theta})/\partial\theta_1\theta_2$. The second derivatives of $\ell_i(\boldsymbol{\theta})$ are provided in Appendix C. The approximate standard errors (SE) of $\widehat{\boldsymbol{\theta}}$ can be estimated with the square roots of the diagonal elements of $I^{-1}(\boldsymbol{\theta})$, replacing $\boldsymbol{\theta}$ by $\widehat{\boldsymbol{\theta}}$.

3.3 Bayesian inference

Next, we present the developments related to the Bayesian inference through MCMC algorithms. We present the prior and the respective posterior distributions, along with suitable MCMC algorithms to sample from the respective marginal posterior distributions of interest. Consider both original and augmented likelihood functions (in order to compare them). The first of them is given by

$$L(\boldsymbol{\theta}|\boldsymbol{t}) = \prod_{i=1}^{n} 2\phi\left[a_{t_i;\mu,\sigma}(\alpha,\beta)\right] \Phi\left[\lambda\, a_{t_i;\mu,\sigma}(\alpha,\beta)\right] A_{t_i;\sigma}(\alpha,\beta).$$

We assume the following prior distributions: $\alpha \sim \text{gamma}(\mathrm{r}_\alpha; \lambda_\alpha)$, $\beta \sim \text{gamma}(r_\beta; \lambda_\beta)$ and $\gamma \sim \mathrm{U}(a; b)$, mutually independent, where $\text{gamma}(r, \lambda)$ stands for a gamma distribution such that $\mathrm{E}(\alpha) = r/\lambda$ and $\mathrm{Var}(\alpha) = r/\lambda^2$ and $U(a; b)$ stands for a continuous uniform distribution over the interval $[a, b]$. Combining the likelihood function with the prior distribution, we have that the joint posterior distribution is given by

$$\pi(\boldsymbol{\theta}|\boldsymbol{t}) \propto \alpha^{\mathrm{r}_\alpha - 1}\beta^{\mathrm{r}_\beta - 1} \exp\left[-(\alpha\lambda_\alpha + \beta\lambda_\beta)\right] \prod_{i=1}^{n} \phi\left[a_{t_i;\mu,\sigma}(\alpha,\beta)\right] \Phi\left[\lambda\, a_{t_i;\mu,\sigma}(\alpha,\beta)\right] A_{t_i;\sigma}(\alpha,\beta),$$

and the respective full conditional distributions, given by

$$\pi(\alpha|\beta,\gamma,\boldsymbol{t}) \propto \alpha^{\mathrm{r}_\alpha - 1} \exp(-\alpha\lambda_\alpha) \prod_{i=1}^{n} \phi\left[a_{t_i;\mu,\sigma}(\alpha,\beta)\right] \Phi\left[\lambda\, a_{t_i;\mu,\sigma}(\alpha,\beta)\right] A_{t_i;\sigma}(\alpha,\beta),$$

$$\pi(\beta|\alpha,\gamma,\boldsymbol{t}) \propto \beta^{\mathrm{r}_\beta - 1} \exp(-\beta\lambda_\beta) \prod_{i=1}^{n} \phi\left[a_{t_i;\mu,\sigma}(\alpha,\beta)\right] \Phi\left[\lambda\, a_{t_i;\mu,\sigma}(\alpha,\beta)\right] A_{t_i;\sigma}(\alpha,\beta),$$

$$\pi(\gamma|\alpha,\beta,\boldsymbol{t}) \propto \prod_{i=1}^{n} \phi\left[a_{t_i;\mu,\sigma}(\alpha,\beta)\right] \Phi\left[\lambda\, a_{t_i;\mu,\sigma}(\alpha,\beta)\right] A_{t_i;\sigma}(\alpha,\beta).$$

In addition, the augmented likelihood function is given by

$$L(\boldsymbol{\theta}|\boldsymbol{t}_c) = \prod_{i=1}^{n} \sqrt{2/\pi}\,\phi\left[\nu_{h_i} + a_{t_i}(\alpha,\beta)\right] A_{t_i}(\alpha,\beta) \exp\left(-\frac{h_i^2}{2}\right).$$

Similarly, combining the augmented likelihood function with the above prior distribution, we have that the posterior distribution is expressed as

$$\pi(\boldsymbol{\theta},\boldsymbol{h}|\boldsymbol{t}) \propto \alpha^{\mathrm{r}_\alpha - 1}\beta^{\mathrm{r}_\beta - 1} \prod_{i=1}^{n} \phi\left[a_{t_i,h_i}(\alpha,\beta)\right] A_{t_i}(\alpha,\beta) \exp\left[-\frac{1}{2}\left(h_i^2 + 2\alpha\lambda_\alpha + 2\beta\lambda_\beta\right)\right]$$

and the respective full conditional distributions are given by

$$\pi(\boldsymbol{h}|\alpha,\beta,\gamma,\boldsymbol{t}_c) \propto \prod_{i=1}^{n} \phi\left[a_{t_i,h_i}(\alpha,\beta)\right] A_{t_i}(\alpha,\beta) \exp\left(-\frac{h_i^2}{2}\right),$$

$$\pi(\alpha|\beta,\gamma,\boldsymbol{t}_c) \propto \alpha^{\mathrm{r}_\alpha-1} \prod_{i=1}^{n} \phi\left[a_{t_i,h_i}(\alpha,\beta)\right] A_{t_i}(\alpha,\beta) \exp\left[-\frac{1}{2}\left(h_i^2+2\alpha\lambda_\alpha\right)\right],$$

$$\pi(\beta|\alpha,\gamma,\boldsymbol{t}_c) \propto \beta^{\mathrm{r}_\beta-1} \prod_{i=1}^{n} \phi\left[a_{t_i,h_i}(\alpha,\beta)\right] A_{t_i}(\alpha,\beta) \exp\left[-\frac{1}{2}\left(h_i^2+2\beta\lambda_\beta\right)\right],$$

$$\pi(\gamma|\alpha,\beta,\boldsymbol{t}_c) \propto \prod_{i=1}^{n} \phi\left[a_{t_i,h_i}(\alpha,\beta)\right] A_{t_i}(\alpha,\beta) \exp\left(-\frac{h_i^2}{2}\right),$$

where $a_{t_i,h_i}(\alpha,\beta) = \nu_{h_i} + a_{t_i}(\alpha,\beta)$. We can see that both posterior distributions are not analytically tractable. Therefore, some numerical method must be employed to obtain suitable numerical approximations for the respective marginal posterior distributions. The above full conditional distributions do not correspond to known distributions, but they can be simulated through some auxiliary algorithm such as the Metropolis-Hastings, slice sampling or adaptive rejection. All these algorithms can be easily implemented in the `R` program. In addition, which is the approach pursued here, we can use a general MCMC computational framework, such `OpenBUGS`, see Lunn et al. (2009). In this case, it is necessary to provide the original or the augmented likelihood function, along with the prior distributions, such that the full conditional distributions are simulated through suitable algorithms, following a pre-defined hierarchy available on the `OpenBUGS`. We made all simulations using the `R` package `R2OpenBUGS`.

4. Numerical Aspects

4.1 Simulation study I

A simulation study is conducted to assess the behavior of the ECM algorithm, in terms of parameter recovery, and the accuracy of the corresponding SEs, calculated through the observed Fisher information matrix. For that, $N = 1,000$ replications are generated considering $n = 500$ and $\boldsymbol{\theta}^\top = (\alpha,\beta,\gamma) = (0.5, 1.0, 0.67)$, which induces a strong positively skewed behavior of the SNBS distribution. In Table 1 we can see the mean of the estimates $(\bar{\widehat{\boldsymbol{\theta}}})$, the mean of the theoretical (asymptotic) SE $(\mathrm{SE}(\widehat{\boldsymbol{\theta}}))$ and the empirical SE $(\mathrm{SE}_{\mathrm{emp}})$. We can notice that the parameters are well recovered and that the empirical SE are close to the theoretical ones, which indicates that the use of the observed Fisher information matrix, to obtain the corresponding SE, is appropriate.

Table 1. Results of the simulation study I.

	$\bar{\widehat{\boldsymbol{\theta}}}$	$\mathrm{SE}(\widehat{\boldsymbol{\theta}})$	$\mathrm{SE}_{\mathrm{emp}}$
$\widehat{\alpha}$	0.495	0.019	0.021
$\widehat{\beta}$	1.003	0.032	0.028
$\widehat{\gamma}$	0.667	0.015	0.012

4.2 Simulation study II

We consider a total of 30 scenarios, resulting from the combination of the levels of: three different sample sizes (n) (10, 50, 200), under $\alpha \in (0.5; 1.5)$, $\beta = 1$ and $\gamma \in (-0.67; -0.45; 0; 0.45; 0.67)$. The sample sizes are chosen in order to verify the properties of the estimators, as consistency, and to compare their behavior, in terms of accuracy. The values of α and β are chosen in order to induce different shapes and small variability, whereas the values of γ induce from null to high positive/negative asymmetry. We calculated the usual statistics to measure the accuracy of the estimates: bias, variance (Var), root mean squared error (RMSE) and absolute value of the relative bias (AVRB). Let θ be the parameter of interest, $\widehat{\theta}_r$ be some estimate related to the replica r and $\overline{\widehat{\theta}} = (1/R)\sum_{r=1}^{R}\widehat{\theta}_r$. The adopted statistics are: Bias $= \overline{\widehat{\theta}} - \theta$, Variance $= (1/R)\sum_{r=1}^{R}(\widehat{\theta}_r - \overline{\widehat{\theta}})^2$, RMSE $= ((1/R)\sum_{r=1}^{R}(\theta - \widehat{\theta}_r)^2)^{1/2}$, AVRB $= |\overline{\widehat{\theta}} - \theta|/|\theta|$.

The usual tools for monitoring the convergence of the MCMC algorithms, see Gamerman and Lopes (2006), indicate that a burn-in of 4,000, a thin of 100, simulating a total of 100,000 values, are enough to produce valid MCMC samples of size 1,000 for each parameter. Since the other results are similar (they are omitted here but they are available under request from the authors), we present only those related to the scenario where $\alpha = 0.5$, $\beta = 1$, $\gamma = -0.67$, varying the sample size. We used (< 0.001) to represent positive values (statistics and/or estimates) and (> -0.001) to denote negative values, when they are close to zero. In addition, we refer the Bayesian estimates as "augmented", when the augmented likelihood function is used, and "original", whenever the original likelihood function is considered. The selected results can be seen in Table 2. In general, we can see that, as the sample size increases, the estimates obtained by the three approaches tend to the correspondent the respective true values. When $\alpha = 0.5$, the ML estimates are more accurate than the Bayesian ones, especially considering the bias and AVRB metrics. In other scenarios (not shown), when $\alpha = 1.5$, the opposite occurs for all sample sizes. Concerning β and γ, it is possible to notice that, under the smallest sample size ($n = 10$), the ML approach presents more accurate estimates than the Bayesian ones. In addition, for $n = 50$ and $n = 200$, Bayesian estimates, for both parameters, are closer to the respective true values. In conclusion, we can say that all estimators, mainly the Bayesian ones, are consistent, since both bias and RMSE tend to decrease, as the sample size increases. Furthermore, the results indicate (including those not shown here) that the Bayesian approach provided the most accurate estimates. Moreover, we can notice that the original and augmented approaches, performed quite similarly. Therefore, we decide to use the original likelihood function) approach, since it is easier to implement and faster.

4.3 Real data analysis I

We analyze a data set corresponding to self-efficacy, which is available in the `R` software and can be accessed from the `EstCRM` package through the command `data(SelfEff)`. A group of 307 pre-service teachers, graduated from various departments in the college of education, are asked to check on a 11 cm line segment with two end points (can not do at all, highly certain can do) using their own judgment for the 10 items that measure teacher self-efficacy on different activities. We take, as response variable, the teacher self-efficacy in the creation of learning environments in which students can effectively express themselves. Table 3 presents some descriptive statistics, including location measures, standard deviation (SD), coefficient of skewness (CS), and kurtosis (CK). We can notice that the distribution is strongly negatively skewed. We fit the CSNBS and BS distributions, using the Bayesian augmented and the ML method, to the data. The results obtained considering the frequentist approach are omitted here but they are available under request from

Table 2. Results of simulation study II with $\gamma = -0.67$.

Parameter	n	Method	Mean	Variance	Bias	RMSE	AVRB
		Augmented	0.577	< 0.001	0.077	0.081	0.154
	10	Original	0.578	0.001	0.078	0.082	0.156
		ML	0.520	0.071	0.020	0.267	0.040
α		Augmented	0.511	< 0.001	0.011	0.016	0.022
	50	Original	0.511	< 0.001	0.011	0.015	0.021
		ML	0.498	0.001	-0.002	0.033	0.004
		Augmented	0.502	< 0.001	0.002	0.005	0.004
	200	Original	0.502	< 0.001	0.002	0.005	0.004
		ML	0.490	< 0.001	-0.010	0.012	0.019
		Augmented	1.006	< 0.001	0.006	0.023	0.006
	10	Original	1.004	< 0.001	0.004	0.021	0.004
		ML	1.105	0.214	0.105	0.474	0.105
β		Augmented	0.996	< 0.001	-0.004	0.009	0.004
	50	Original	0.997	< 0.001	-0.003	0.009	0.003
		ML	1.039	0.018	0.039	0.140	0.039
		Augmented	0.999	< 0.001	-0.001	0.005	0.001
	200	Original	0.999	< 0.001	-0.001	0.005	0.001
		ML	0.997	< 0.001	-0.003	0.004	0.003
		Augmented	-0.157	0.067	0.513	0.575	0.766
	10	Original	-0.182	0.054	0.488	0.540	0.728
		ML	-0.603	0.028	0.067	0.179	0.100
γ		Augmented	-0.493	0.059	0.177	0.301	0.264
	50	Original	-0.505	0.049	0.165	0.276	0.247
		ML	-0.569	0.012	0.101	0.148	0.150
		Augmented	-0.614	0.017	0.056	0.142	0.083
	200	Original	-0.601	0.015	0.069	0.141	0.103
		ML	-0.523	0.002	0.147	0.153	0.220

the authors. The prior distributions are the same used in Section 3. In Table 4, in addition to the posterior expectations (PE), the posterior standard deviations (PSD) and the 95% equi-tailed credibility intervals (CI), we also present the model selection criteria. We consider the usual statistics of model comparison for both frequentist (AIC, BIC) and Bayesian (DIC, EAIC, EBIC and LPLM) see, respectively (Akaike, 1974; Schwarz, 1978; Spiegelhalter et al., 2014). The smaller values of AIC and BIC indicates the model that fits the data better. In addition, the smaller the values of DIC, EAIC, EBIC, the better the model fit, occurring the opposite with the LPML. We can notice that the estimates of α and γ (under the CSNBS model) indicate that the distribution is strongly negatively skewed. Notice also that we have indications that the asymmetry parameter is different from zero, since this value does not belong to the CI. Moreover, the criteria indicated the CSNBS model is the best. Figure 10 (left) presents the histogram of the observations and estimated densities. We can notice that the CSNBS distribution presents an advantage over the BS model. From Figure 10, we can notice that the CSNBS distribution predicts better the observations than the BS distribution. In conclusion, we can say that the CSNBS model is preferable to the BS model.

Table 3. Descriptive statistics for the teacher self-efficacy data.

Mean	Median	Minimum	Maximum	SD	Asymmetry	Kurtosis
9.205	9.700	1.650	10.900	1.365	-1.752	7.781

4.4 Real data analysis II

We analyze now a data set corresponding to prices of bottles of Barolo wine and discussed in Azzalini (2013). It concerns the price (in euros) of bottles (75 cl) of Barolo wine. The data have been obtained in July 2010 from the websites of four Italian wine resellers, selecting only quotations of Barolo wine, which is produced in the Piedmont region of

Table 4. Posterior expectations (PE), posterior standard deviations (PSD), equi-tailed 95% CI and model selection criteria.

Parameter	PE	PSD	$\mathrm{CI}_{95\%}$
		CSNBS	
α	0.154	0.002	[0.151; 0.157]
β	8.871	0.016	[8.836; 8.903]
γ	-0.971	0.003	[-0.978; -0.966]
EAIC		1,021.912	
EBIC		1,033.093	
DIC		3,047.154	
LPML		-508.531	
		BS	
α	0.205	0.008	[0.190; 0.222]
β	9.016	0.105	[8.815; 9.229]
EAIC		1,252.772	
EBIC		1,260.226	
DIC		3,744.335	
LPML		-632.9564	

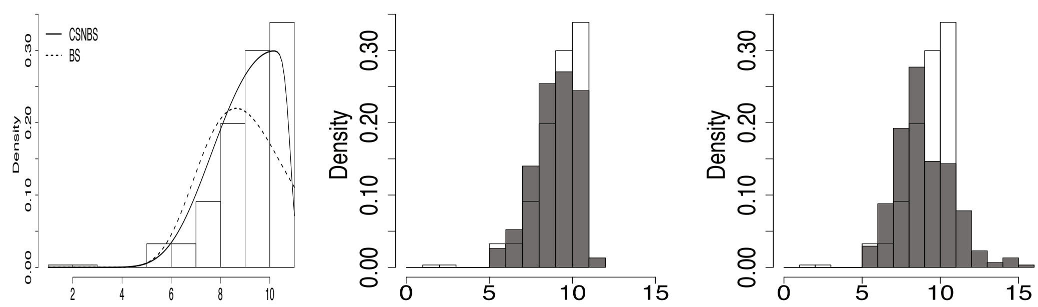

Figure 10. Histogram of the observations and estimated densities (left), histogram of the predicted and observed distributions for the CSNBS (center) and BS (right) models.

Italy. The price does not include the delivery charge. In Table 5 and Figure 11 (left), we present a descriptive analysis. It is possible to see that the distribution is positively skewed and more concentrated in the first class [0,100]. We fit the CSNBS and BS distributions, using the Bayesian augmented and the ML method, to the data. The results obtained considering the frequentist approach are omitted here but they are available under request from the authors. The prior distributions are the same used in Section 3. In Table 6, in addition to the posterior expectations (PE), the posterior standard deviations (PSD) and the 95% equi-tailed CI, we also present the Bayesian criteria. Table 6 shows that the estimates of α and γ (under the CSNBS model) indicate that the distribution of the prices is strongly positively skewed. Notice also that we have indications that the asymmetry parameter is different from zero, since this value does not belong to the CI. Moreover, the criteria indicated the CSNBS model is the best. Also, we construct QQ plots with simulated envelopes. Similar to Vilca et al. (2011), we considered the Bayesian estimates of α and β in $d(\alpha,\beta) = (1/\alpha^2)(T/\beta + \beta/T - 2)$. When $T \sim \mathrm{BS}(\alpha,\beta)$, it is know that $d(\alpha,\beta) \sim N(0,1)$. Since the observations $d(\widehat{\alpha},\widehat{\beta})$ are expected to follow a standard normal distribution, under the well fit of the model, the envelopes are simulated from the standard normal distribution, as described in Atkinson (1985). Similarly, if $T \sim \text{CSNBS}(\alpha, \beta, \gamma)$, thus $d(\alpha, \beta) \sim \text{CSN}(0, 1, \gamma)$. Since the observations $d(\widehat{\alpha}, \widehat{\beta})$ are expected to follow a CSN distribution, under the well fit of the model, the envelopes are simulated from the CSN distribution. These plots are presented in Figure 11 (lines represent the 5th percentile, the mean, and the 95th percentile of 100 simulated points). From those figures, we conclude that the CSNBS distribution provides a better fit than the BS model. Specifically, from the QQ plot shown in Figure 11 (a), we notice that the observations appear to form a slight upward-facing concave. However, the QQ plot shown in Figures 11 (b) indicate that the CSNBS distribution offers an excellent fit, provided that the majority of observations are inside of the envelope.

Table 5. Descriptive statistics for the prices of bottles of Barolo wine.

Mean	Median	Minimum	Maximum	SD	Asymmetry	Kurtosis
124.617	72	14	1000	37.041	2.903	12.982

Table 6. Posterior expectations (PE), posterior standard deviations (PSD), equi-tailed 95% CI and model selection criteria.

Parameter	PE	PSD	$\text{CI}_{95\%}$
		CSNBS	
α	0.844	0.037	[0.775; 0.917]
β	89.576	3.911	[82.260; 97.871]
γ	0.690	0.070	[0.541; 0.809]
EAIC		3,437.879	
EBIC		3,449.060	
DIC		10,292.690	
LPML		-1,718.110	
		BS	
α	0.858	0.035	[0.794; 0.929]
β	92.444	4.264	[84.778; 101.302]
EAIC		3,474.893	
EBIC		3,482.346	
DIC		10,410.620	
LPML		-1,736.669	

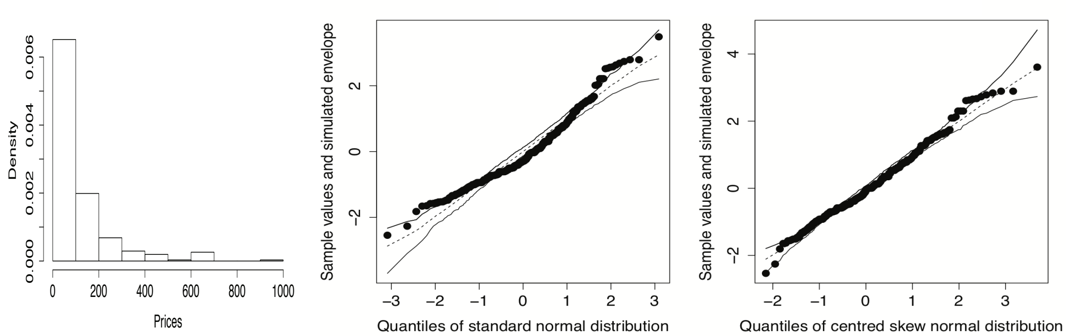

Figure 11. Histogram of the prices of bottles of Barolo wine (left), QQ plots with envelopes for BS (center) and CSNBS (right) distributions for the data of Barolo wine bottle prices.

5. Concluding Remarks

In this paper, we introduced a new distribution for modeling positive data which can present both positive and negative asymmetry, by combining the Birnbaum-Saunders and the centered skew normal distributions. We developed both maximum likelihood and Bayesian estimation procedures, comparing them through a suitable simulation study. The convergence of the conditional expectation maximization and MCMC algorithms were verified and several factors of interest were compared in the parameter recovery study. In general, as the sample size increases, the results indicated that the Bayesian approach provided the most accurate estimates. In future works we can consider the development of predictive posterior checking to detect the goodness of fit. Furthermore, we suggest the use of Jeffreys-rule prior and independence Jeffreys prior. Other auxiliary algorithms as the Hamiltonian Monte Carlo (see Homand and Gelman, 2014; Carpenter et al., 2016)), adaptive reject sampling and slice sampling (see Gamerman and Lopes, 2006) can be used and compared. Other family of distributions could be used instead of the centered skew normal distribution, as the scale mixture of the SN distributions, to generate new family of Birnbaum-Saunders-type distributions. Finally, other numerical methods to obtain approximation for the marginal posterior distributions, such as the INLA algorithm, can be considered (see Rue and Martino, 2009).

6. Appendix

Appendix A. Moments of the CSNBS Distribution

Theorem A.1 Let $T \sim \text{CSNBS}(\alpha, \beta, \gamma)$ and $Y \sim \text{CSN}(0, 1, \gamma)$. If $\mathrm{E}\left[Y^{2(r-j+i)}\right] < \infty$, then the moments of T are given by

$$\mathrm{E}(T^r) = \beta^r \sum_{j=0}^{r} \binom{2r}{2j} \sum_{i=0}^{j} \mathrm{E}\left[Y^{2(r-j+i)}\right] (\alpha/2)^{2(r-j+i)}.$$

Proof of Theorem A.1 From Equation (3), we have that

$$\mathrm{E}\left[\left(\frac{T}{\beta}\right)^r\right] = \mathrm{E}\left\{\left[\left(\frac{\alpha Y}{2} + \sqrt{\left(\frac{\alpha Y}{2}\right)^2 + 1}\right)^2\right]^r\right\}.$$

From the binomial theorem, that is, $(a+b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} b^k$, we have that

$$\mathrm{E}\left[\left(\frac{T}{\beta}\right)^r\right] = \sum_{k=0}^{2r} \binom{2r}{k} \mathrm{E}\left\{\left[\left(\frac{\alpha Y}{2}\right)^2 + 1\right]^{k/2} \left(\frac{\alpha Y}{2}\right)^{2r-k}\right\}.$$

Considering $k = 2j$, that is, $j = k/2$, it comes that

$$\mathrm{E}\left[\left(\frac{T}{\beta}\right)^r\right] = \sum_{j=0}^{r} \binom{2r}{2j} \mathrm{E}\left\{\left[\left(\frac{\alpha Y}{2}\right)^2 + 1\right]^{j} \left(\frac{\alpha Y}{2}\right)^{2(r-j)}\right\}.$$

From the binomial theorem again, we have

$$\begin{aligned}
\mathrm{E}\left[\left(\frac{T}{\beta}\right)^r\right] &= \sum_{j=0}^{r}\binom{2r}{2j}\mathrm{E}\left\{\sum_{i=0}^{j}\binom{j}{i}\left(\frac{\alpha Y}{2}\right)^{2i}\left(\frac{\alpha Y}{2}\right)^{2(r-j)}\right\}\\
&= \sum_{j=0}^{r}\binom{2r}{2j}\sum_{i=0}^{j}\binom{j}{i}\mathrm{E}\left[\left(\frac{\alpha Y}{2}\right)^{2(r-j+i)}\right]\\
&= \sum_{j=0}^{r}\binom{2r}{2j}\sum_{i=0}^{j}\binom{j}{i}\mathrm{E}\left[Y^{2(r-j+i)}\right]\left(\frac{\alpha}{2}\right)^{2(r-j+i)}.
\end{aligned}$$

Therefore,

$$\mathrm{E}(T^r) = \beta^r\sum_{j=0}^{r}\binom{2r}{2j}\sum_{i=0}^{j}\binom{j}{i}\mathrm{E}\left[Y^{2(r-j+i)}\right](\alpha/2)^{2(r-j+i)}. \tag{A1}$$

From Equation (A1), we get

$$\mathrm{E}(T) = \beta\sum_{j=0}^{1}\binom{2}{2j}\sum_{i=0}^{j}\binom{j}{i}\mathrm{E}\left[Y^{2(1-j+i)}\right](\alpha/2)^{2(1-j+i)}.$$

For $j=0$, the first term of the sum in Equation (A1) is equal to $\beta\mathrm{E}(Y^2)(\alpha/2)^2$. For $j=1$, the second term of the sum in Equation (A1) is equal to $\beta\left[1+\mathrm{E}(Y^2)(\alpha/2)^2\right]$. Hence, by adding these two terms, we have

$$\mathrm{E}(T) = \beta\left[1+(\alpha^2/2)\right].$$

Furthermore, from Equation (A1), we have

$$\mathrm{E}(T^2) = \beta^2\sum_{j=0}^{2}\binom{4}{2j}\sum_{i=0}^{j}\binom{j}{i}\mathrm{E}\left[Y^{2(2-j+i)}\right](\alpha/2)^{2(2-j+i)}.$$

Developing the above sum in j, we obtain

$$\mathrm{E}(T^2) = \beta^2\left[1+\frac{\alpha^4}{2}\Delta+2\alpha^2\right].$$

Thus,

$$\begin{aligned}
\mathrm{Var}(T) &= \mathrm{E}(T^2)-[\mathrm{E}(T)]^2\\
&= (\alpha\beta)^2\left\{1+\frac{\alpha^2}{4}\left[2\Delta-1\right]\right\}.
\end{aligned}$$

where $\Delta=\mathrm{E}(Y^4)=2(\pi-3)(4/\pi^2)\delta^4[1-(2\delta^2/\pi)]^{-2}+3$.

Appendix B. The ECM Algorithm

The following result is used in the proof of Theorem 2.1.

Lemma 1. Let $X \sim N(\eta, \tau^2)$, thus $\forall a \in \mathrm{R}$

$$\mathrm{E}(X|X > a) = \eta + \frac{\phi\left(\frac{a-\eta}{\tau}\right)}{1-\Phi\left(\frac{a-\eta}{\tau}\right)}\tau;\ \mathrm{E}(X^2|X > a) = \eta^2 + \tau^2 + \frac{\phi\left(\frac{a-\eta}{\tau}\right)}{1-\Phi\left(\frac{a-\eta}{\tau}\right)}(\eta + a)\tau.$$

Proof of Theorem 2.1

(i) Since $Y \sim \mathrm{CSN}(0, 1, \gamma)$, using the stochastic representation given by Equation (2), we can define

$$Y = \frac{1}{\sigma_z}\left[\delta H + \sqrt{1-\delta^2}X_1 - \mu_z\right] = \frac{1}{\alpha}\left[\sqrt{T/\beta} - \sqrt{\beta/T}\right].$$

Therefore,

$$Y|(H = h) = \frac{1}{\alpha}\left(\sqrt{T/\beta} - \sqrt{\beta/T}\right)\Big|(H = h) \sim \mathrm{N}(\mu_h, \sigma^2),$$

where $\mu_h = \delta(h - r)/(1 - r^2\delta^2)^{1/2}$ and $\sigma^2 = (1 - \delta^2)/(1 - r^2\delta^2)$. Then,

$$W|(H = h) = -\frac{\mu_h}{\sigma} + \frac{1}{\sigma\alpha}\left(\sqrt{T/\beta} - \sqrt{\beta/T}\right)\Big|(H = h) \sim \mathrm{N}(0, 1)$$

and

$$T = \beta\left[\frac{\alpha}{2}(\sigma W + \mu_h) + \sqrt{\left[\frac{\alpha}{2}(\sigma W + \mu_h)\right]^2 + 1}\right].$$

From the above result, the proof is completed.

(ii) As $f_H(h) = 2\phi(h|0, 1), h > 0$ and

$$\phi(\nu_h + a_t(\alpha_{\delta,\beta})) = \frac{\sqrt{1-\delta^2}}{\sqrt{1-r^2\delta^2}}\phi\left(a_t(\alpha, \beta)\,\Bigg|\,\frac{\delta(h-r)}{\sqrt{1-r^2\delta^2}}, \frac{1-\delta^2}{1-r^2\delta^2}\right).$$

Then, we have

$$\phi\left(a_t(\alpha,\beta)\Big|\frac{\delta(h-r)}{\sqrt{1-r^2\delta^2}}, \frac{1-\delta^2}{1-r^2\delta^2}.\right)\phi(h|0,1) = \phi\Big(a_t(\alpha,\beta)\Big| - \frac{r\delta}{\sqrt{1-r^2\delta^2}}; \frac{1}{1-r^2\delta^2}\Big)$$
$$\times\, \phi\Big(h\Big|\delta\sqrt{1-r^2\delta^2}(a_t(\alpha,\beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}), 1-\delta^2\Big).$$

Therefore, the proof of (i) follows directly from that $f_{H|T}(h|t) = f_{T|H}(t|h)f_H(h)/f_T(t)$. To demonstrate (ii)-(iii), notice, for $k = 1, 2$, we have that

$$\mathrm{E}\left[H^k|T\right] = \frac{1}{\Phi\Big(\lambda\sigma_z\big(a_t(\alpha,\beta) + \frac{r\delta}{\sqrt{1-r^2\delta^2}}\big)\Big)}\int_0^\infty h^k\phi\left\{h|\eta_t, 1-\delta^2\right\}\mathrm{d}h = \mathrm{E}(X^k|X > 0).$$

Then, using some properties of the HN distribution from Lemma 1, the proof is completed.

Appendix C. The Observed Fisher Information Matrix

The necessary expressions are given below. For the sake of simplicity, we consider the following notation to obtain the necessary expressions, $a_{t_i;\mu,\sigma} = a_{t_i;\mu,\sigma}(\alpha,\beta)$ and $A_{t_i;\sigma} = A_{t_i;\sigma}(\alpha,\beta)$.

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial\theta_1\partial\theta_2} = -\frac{1}{A^2_{t_i;\sigma}}\frac{\partial A_{t_i;\sigma}}{\partial\theta_1}\frac{\partial A_{t_i;\sigma}}{\partial\theta_2} + \frac{1}{A_{t_i;\sigma}}\frac{\partial^2 A_{t_i;\sigma}}{\partial\theta_1\partial\theta_2} - \frac{1}{2}\frac{\partial^2 a^2_{t_i;\mu,\sigma}}{\partial\theta_1\partial\theta_2} + \lambda^2 W'_\Phi[\lambda a_{t_i;\mu,\sigma}]\frac{\partial a_{t_i;\mu,\sigma}}{\partial\theta_1}\frac{\partial a_{t_i;\mu,\sigma}}{\partial\theta_2}$$
$$+\lambda W_\Phi[\lambda a_{t_i;\mu,\sigma}]\frac{\partial^2 a_{t_i;\mu,\sigma}}{\partial\theta_1\partial\theta_2}, \quad \theta_1,\theta_2 = \alpha,\beta$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial\theta_3\partial\gamma} = -\frac{1}{A^2_{t_i;\sigma}}\frac{\partial A_{t_i;\sigma}}{\partial\theta_3}\frac{\partial A_{t_i;\sigma}}{\partial\gamma} + \frac{1}{A_{t_i;\sigma}}\frac{\partial^2 A_{t_i;\sigma}}{\partial\theta_3\partial\gamma} - \frac{1}{2}\frac{\partial^2 a^2_{t_i;\mu,\sigma}}{\partial\theta_3\partial\gamma} + \lambda W_\Phi[\lambda a_{t_i;\mu,\sigma}]\frac{\partial^2 a_{t_i;\mu,\sigma}}{\partial\theta_3\partial\gamma}$$
$$+\left\{\lambda W'_\Phi[\lambda a_{t_i;\mu,\sigma}]\frac{\partial\lambda a_{t_i;\mu,\sigma}}{\partial\gamma} + W_\Phi[\lambda a_{t_i;\mu,\sigma}]\frac{\partial\lambda}{\partial\gamma}\right\}\frac{\partial a_{t_i;\mu,\sigma}}{\partial\theta_3}, \quad \theta_3 = \alpha,\beta$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial\gamma^2} = -\frac{1}{A^2_{t_i;\sigma}(\alpha,\beta)}\frac{\partial^2 A_{t_i;\sigma}}{\partial\gamma^2} + \frac{1}{A_{t_i;\sigma}(\alpha,\beta)}\frac{\partial^2 A_{t_i;\sigma}}{\partial\gamma^2} - \frac{1}{2}\frac{\partial^2 a^2_{t_i;\mu,\sigma}}{\partial\gamma^2} + W'_\Phi[\lambda a_{t_i;\mu,\sigma}]\left(\frac{\partial\lambda a_{t_i;\mu,\sigma}}{\partial\gamma}\right)^2$$
$$+W_\Phi[\lambda a_{t_i;\mu,\sigma}]\frac{\partial^2\lambda a_{t_i;\mu,\sigma}}{\partial\gamma^2},$$

where $W'_\Phi(x) = -W_\Phi(x)[x+W_\Phi(x)]$ is the derivative of $W_\Phi(x)$ with respect to x, see Vilca et al. (2011), and the other quantities are as before defined.

Acknowledgements

The authors would like to thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the financial support through a Master Scholarship granted to the first author under the guidance of the second and also for Conselho Nacional de Desenvolvimento Científico e Tecnológico, grants numbers 308339/2015-0 and 305336/2017-7, for a research scholarship granted to the second and fourth authors, respectively.

References

Achcar, J.A., 1993. Inferences for the Birnbaum-Saunders fatigue life model using Bayesian methods. Computational Statistics and Data Analysis, 15, 367-380.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.

Atkinson, A.C., 1985. Plots, Transformations and Regressions. Oxford University Press, Oxford, UK.

Arellano-Valle, R. and Azzalini, A., 2008. The centered parametrization for the multivariate skew-normal distribution. Journal of Multivariate Analysis, 99, 1362-1382.

Azzalini, A., 1985. A class of distribution which includes the normal ones. Scandinavian Journal of Statistics, 12, 171-178.

Azzalini, A., 2013. The skew-normal and related families. Cambridge University Press, Cambridge, UK..

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

Balakrishnan, N.;,Saulo, H., and Leao, J., 2017. On a new class of skewed BirnbaumSaunders models. Journal of Statistical Theory and Practice, 11, 573-593.

Barros, M., Paula, G.A., and Leiva, V., 2008. A New Class of Survival Regression Models with Heavy-Tailed Errors: Robustness and Diagnostics. Lifetime Data Analysis, 14, 1-17.

Birnbaum, Z.W. and Saunders, SC., 1969a. A new family of life distributions. Journal of Applied Probability, 6, 637-652.

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

Byrd, R.H., Lu, P., Nocedal, J., and Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM Journal on Scientific Computing, 15, 1190-1208.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M.A., Guo, J., Li, P., and Riddell, A., 2016. Stan: A probabilistic programming language. Journal of Statistical Software, 20, 1-37.

Cordeiro, G.M. and Lemonte, A.J., 2011. The β-Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. Computational Statistics and Data Analysis, 55, 116-124.

Depaoli, S., Clifton, J.P., and Cobb, P. R., 2016. Just another Gibbs sampler (JAGS) flexible software for MCMC implementation. Journal of Educational and Behavioral Statistics, 41, 628-649.

Desmond, A.F., 1985. Stochastic models of failure in random environments. Canadian Journal of Statistics, 13, 171-183.

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

Diaz-Garcia, J.A. and Leiva, V., 2005. New family of life distributions based on the elliptically contoured distributions. Journal of Statistical Planning and Inference, 128, 445-457.

Engelhardt, M., Bain, L.J., and Wright, F. ., 1981. Inferences on the parameters of the Birnbaum-Saunders fatigue life distribution based on maximum likelihood estimation. Technometrics, 23, 251-256.

Ferreira, M.S., 2013. A study of exponential-type tails applied to Birnbaum-Saunders models. Chilean Journal of Statistics, 4, 87-97.

Gamerman, D. and Lopes, H., 2006. MCMC-Stochastic Simulation for Bayesian Inference. Chapman and Hall/CRC, Boca Raton, FL, US.

Homan, M.D. and Gelman, A., 2014. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. The Journal of Machine Learning Research, 15, 1593-1623.

Leiva, V., Barros, M., Paula, G.A., and Sanhueza, A., 2007. Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. Environmetrics, 19, 235-249.

Leiva, V., Vilca, F., Balakrishnan, N., and Sanhueza, A., 2008. A skewed sinh-normal distribution and its properties and application to air pollution. Communications in Statistics Theory and Methods, 39, 426-443.

Leiva, V., 2016. The Birnbaum-Saunders Distribution. Academic Press, New York.

Lu, M.C. and Chang, D.S., 1997. Bootstrap prediction intervals for the Birnbaum-Saunders distribution. Microelectronics Reliability, 37, 1213-1216.

Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing, 10, 325-337.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N., 2009. The BUGS project: Evolution, critique and future directions (with discussion). Statistics in Medicine, 28, 3049-3082.

Maehara, R.P. (2018). An extension of Birnbaum-Saunders distributions based on scale mixtures of skew-normal distributions with applications to regression models. PhD thesis, University of Sao Paulo, available at http://www.teses.usp.br/teses/disponiveis/45/45133/tde-14052018-202935/es.php.

Mann, N.R., Singpurwalla, N.D., and Schafer, R.E., 1974. Methods for Statistical Analysis of Reliability and Life Data. Wiley, New York.

Mazucheli, J.M., Menezes, A.F.B., and Dey, S., 2018.The unit-Birnbaum-Saunders distribution with applications. Chilean Journal of Statistics, 9, 47-57.

Meng, X.L. and Rubin, D.B., 1993. Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika, 80, 267-278.

Owen, W.J. and Padgett, W.J., 1999. Accelerated test models for system strength based on Birnbaum-Saunders distributions. Lifetime Data Analysis, 5, 133-147.

Pewsey, A., 2000. Problems of inference for Azzalini's skew-normal distribution. Journal of Applied Statistics, 27, 859-870.

R Development Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rieck, J.R. and Nedelman, J.R., 1991. A log-linear model for the Birnbaum-Saunders distribution. Technometrics, 33, 51-60.

Rue, H. and Martino, S., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society B, 71, 319-392.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics, 6, 461-464.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Linde, A., 2014. The deviance information criterion: 12 years on. Journal of the Royal Statistical Society B, 76, 485-493.

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

Vilca, F., Santana, L., Leiva, V., and Balakrishnan, N., 2011. Estimation of extreme percentiles in Birnbaum-Saunders distributions. Computational Statistics and Data Analysis, 55, 1665-1678.

Wang, M., Sun, X., and Park, C., 2016. Bayesian analysis of BirnbaumSaunders distribution via the generalized ratio-of-uniforms method. Computational Statistics, 31, 207-225.

Xu, A. and Tang, Y., 2011. Bayesian analysis of Birnbaum-Saunders distribution with partial information. Computational Statistics and Data Analysis, 55, 2324-2333.

Information for authors

The editorial board of the Chilean Journal of Statistics (ChJS) is seeking papers, which will be refereed. We encourage the authors to submit a PDF electronic version of the manuscript in a free format to Víctor Leiva, Editor-in-Chief of the ChJS (E-mail: chilean.journal.of.statistics@gmail.com). Submitted manuscripts must be written in English and contain the name and affiliation of each author followed by a leading abstract and keywords. The authors must include a"cover letter" presenting their manuscript and mentioning: "We confirm that this manuscript has been read and approved by all named authors. In addition, we declare that the manuscript is original and it is not being published or submitted for publication elsewhere".

Preparation of accepted manuscripts

Manuscripts accepted in the ChJS must be prepared in Latex using the ChJS format. The Latex template and ChJS class files for preparation of accepted manuscripts are available at http://chjs.mat.utfsm.cl/files/ChJS.zip. Such as its submitted version, manuscripts accepted in the ChJS must be written in English and contain the name and affiliation of each author, followed by a leading abstract and keywords, but now mathematics subject classification (primary and secondary) are required. AMS classification is available at http://www.ams.org/mathscinet/msc/. Sections must be numbered 1, 2, etc., where Section 1 is the introduction part. References must be collected at the end of the manuscript in alphabetical order as in the following examples:

Arellano-Valle, R., (1994). Elliptical Distributions: Properties, Inference and Applications in Regression Models. Unpublished Ph.D. Thesis. Department of Statistics, University of São Paulo, Brazil.

Cook, R.D., (1997). Local influence. In Kotz, S., Read, C.B. and Banks, D.L. (Eds.), Encyclopedia of Statistical Sciences, Update, Vol. 1, Wiley, New York, pp. 380-385.

Rukhin, A.L., (2009). Identities for negative moments of quadratic forms in normal variables. Statistics and Probability Letters 79, 1004-1007.

Stein, M.L., (1999). Statistical Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York.

Tsay, R.S., Pena, D., and Pankratz, A.E., (2000). Outliers in multivariate time series. Biometrika 87, 789-804.

References in the text must be given by the author's name and year of publication, e.g., Gelfand and Smith (1990). In the case of more than two authors, the citation must be written as Tsay et al. (2000).

Copyright