SORT 33 (2) July-December 2009, 171-192

Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm

N. Balakrishnan[1], Víctor Leiva[2], Antonio Sanhueza[3] and Filidor Vilca[4]

Abstract

Scale mixtures of normal (SMN) distributions are used for modeling symmetric data. Members of this family have appealing properties such as robust estimates, easy number generation, and efficient computation of the ML estimates via the EM-algorithm. The Birnbaum-Saunders (BS) distribution is a positively skewed model that is related to the normal distribution and has received considerable attention. We introduce a type of BS distributions based on SMN models, produce a lifetime analysis, develop the EM-algorithm for ML estimation of parameters, and illustrate the obtained results with real data showing the robustness of the estimation procedure.

MSC: Primary 65C10; Secondary 60E05.

Keywords: Birnbaum-Saunders distribution, EM-algorithm, kurtosis, maximum likelihood methods, robust estimation, scale mixtures of normal distributions.

1 Introduction

The family of scale mixtures of normal (SMN) distributions has attracted considerable attention; see, for example, Kelker (1970), Efron and Olshen (1978), Lange and Sinheimer (1993), Gneiting (1997), Taylor and Verbyla (2004), Walker and Gutiérrez-Peña (2007), and Lachos and Vilca (2007). This family provides flexible thick-tailed

Corresponding author: Víctor Leiva, Departamento de Estadística, Universidad de Valparaíso, Casilla 5030, Valparaíso, Chile. Email: victor.leiva@uv.cl; victor.leiva@yahoo.com

[1] Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

[2] Department of Statistics, CIMFAV, Universidad de Valparaíso, Valparaíso, Chile

[3] Department of Mathematics and Statistics, Universidad de La Frontera, Temuco, Chile

[4] Department of Statistics, Universidade Estadual de Campinas, São Paulo, Brazil

Received: January 2009

Accepted: September 2009

distributions that are often used for robust estimation of parameters of a symmetric distribution; see Lange *et al.* (1989) and Lucas (1997). However, in many practical data involving variates such as lifetimes, pollutant concentrations, and family incomes, it is quite common to find skewed, heavy-tailed data. For this reason, it is necessary to have flexible distributions with good properties for fitting such kind of data. Distributions available with these characteristics are not abundant in the literature.

The Birnbaum-Saunders (BS) distribution is a positively skewed model with non-negative support that has also received considerable attention in the last two decades. This is primarily due to its derivation that is based on physical consideration, its attractive properties, and its close relationship to the normal distribution. These aspects of the BS model render it as an alternative to the normal model for data with non-negative support and positive skewness. For more details about various developments on the BS distribution, one may refer to Birnbaum and Saunders (1969a), Johnson *et al.* (1995, pp. 651-663), and Sanhueza *et al.* (2008).

Exploiting the relationship between the BS and normal distributions, it is possible to obtain a general class of BS distributions based on SMN models, which we call scale-mixture Birnbaum-Saunders (SBS) distributions. The three main reasons for developing this class of distributions are the following: (i) the use of the SBS specification to model observable data enables us to make robust estimation of parameters in a similar way to that of the SMS specification, which is not possible with the BS distribution or any other well-known compatible model such as the lognormal distribution, (ii) the theoretical arguments established in the genesis of the BS distribution can be transferred to the SBS one and thus it is an appropriate model for describing different phenomena that present accumulation of some type under stress, and (iii) SBS distributions allow us to efficiently compute the maximum likelihood (ML) estimates of the model parameters by using the EM-algorithm, which is not possible with the classical BS distribution; moreover, the estimation process proposed in this paper generalizes the one developed earlier by Birnbaum and Saunders (1969b). For more details about the EM-algorithm, see Dempster *et al.* (1977).

The rest of this paper is organized as follows. In Section 2, we introduce the SBS distributions and find their probability density function (pdf). In Section 3, we provide some properties, moments, conditional distributions, and some transformations of SBS models. In Section 4, we analyze some particular cases of these distributions. In Section 5, we produce a lifetime analysis mainly based on the failure rate function of SBS distributions. In Section 6, we describe the ML method for estimating the parameters of SBS models by means of the EM-algorithm. In Section 7, we provide an illustrative example that shows the usefulness of the SBS distributions for fitting three real data sets that are frequently utilized in the literature of this topic. Diagnostic and relative change procedures are used in this example, which show the inherent robustness of the estimation method based on SBS distributions. In addition, we discuss some aspects related to a computational implementation in `R` code for the results obtained in this paper. Finally, in Section 8, we draw some conclusions.

2 Scale-mixture Birnbaum-Saunders distributions

SMN models are related to the normal distribution through the stochastic representation

$$Y = \mu + \sqrt{g(U)}X, \tag{1}$$

where $X \sim \mathrm{N}(0, \sigma^2)$, U is a positive random variable (r.v.) independent of X with cumulative distribution function (cdf) $H(\cdot)$ indexed by a scalar or vector parameter $\boldsymbol{\nu}$ and $g(\cdot)$ is a positive function. Note that when $g(U) = 1/U$ in equation (1), the distribution of Y reduces to the normal/independent distribution discussed by Lange and Sinheimer (1993). Similarly, when $g(U) = U$ in equation (1), the distribution of Y reduces to the SMN distribution studied by Fernandez and Steel (1999).

An r.v. Y has a SMN distribution with location and scale parameters, $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, respectively, iff its pdf is of the form

$$\phi_{\mathrm{SMN}}(y) = \int_0^\infty \phi\left(y|\mu, g(u)\,\sigma^2\right) dH(u), \tag{2}$$

where $\phi\left(\,\cdot\,|\mu,\, g(\cdot)\,\sigma^2\right)$ is the pdf of the normal distribution with mean μ and variance $g(\cdot)\,\sigma^2$ and $H(\cdot)$ is the cdf of U introduced in equation (1). For an r.v. Y with pdf given as in equation (2), the notation $Y \sim \mathrm{SMN}(\mu, \sigma^2; H)$ is used. Now, when $\mu = 0$ and $\sigma^2 = 1$, we use the simpler notation $Y \sim \mathrm{SMN}(H)$.

The BS distribution is related to the normal model through the stochastic representation

$$T = \frac{\beta}{4}\left[\alpha Z + \sqrt{\{\alpha Z\}^2 + 4}\right]^2, \tag{3}$$

where $Z \sim \mathrm{N}(0,1)$, $\alpha > 0$ and $\beta > 0$. Thus, if an r.v. T has the BS distribution with shape and scale parameters, α and β, respectively, then the notation $T \sim \mathrm{BS}(\alpha, \beta)$ is used in this case. From equation (3), the r.v. Z can be stochastically represented in terms of T as

$$Z = \frac{1}{\alpha}\left[\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}}\right]. \tag{4}$$

In an analogous way, if the stochastic representation

$$T = \frac{\beta}{4}\left[\alpha\sqrt{g(U)}Z + \sqrt{\left\{\alpha\sqrt{g(U)}Z\right\}^2 + 4}\right]^2 \tag{5}$$

is considered, where $Y = \sqrt{g(U)}Z \sim \mathrm{SMN}(H)$, with $Z \sim \mathrm{N}(0,1)$, then the r.v. T follows a SBS distribution, which is denoted by $T \sim \mathrm{SBS}(\alpha, \beta; H)$. The stochastic representation given in equation (5) is useful for generating random numbers, deriving moments and implementing the EM-algorithm for ML estimation in SBS models, which is shown in the following sections.

Theorem 1 *Let $T \sim \mathrm{SBS}(\alpha,\beta;H)$. Then, the pdf of T is*

$$f_T(t) = \phi_{\mathrm{SMN}}(a(t))A(t), \quad t>0, \alpha>0, \beta>0, \tag{6}$$

where $\phi_{\mathrm{SMN}}(\cdot)$ is the pdf given in equation (2) with $\mu=0$ and $\sigma^2=1$, $a(t) = \left[\sqrt{t/\beta}-\sqrt{\beta/t}\right]/\alpha$, and $A(t) = t^{-3/2}[t+\beta]/[2\,\alpha\,\beta^{1/2}]$ is the derivative of $a(t)$ with respect to t.

Proof. The required result is directly obtained from the stochastic representation given in equation (5) and the change-of-variable method. ■

3 Properties, moments, conditional distributions, and transformations of SBS models

The following theorem provides some properties of SBS distributions.

Theorem 2 *Let $T \sim \mathrm{SBS}(\alpha,\beta;H)$. Then,*

(i) $cT \sim \mathrm{SBS}(\alpha, c\beta;H)$, *with $c>0$;*
(ii) $1/T \sim \mathrm{SBS}(\alpha,1/\beta;H)$.

Proof. Parts (i) and (ii) are directly obtained from the change-of-variable method. ■

Remark 1 Part (i) of Theorem 2 indicates that the SBS distributions belong to the scale family, while Part (ii) demonstrates that these distributions are closed under reciprocation; see Saunders (1974). In addition, Part (i) allows us to obtain a one-parameter SBS distribution by $\alpha T/\beta \sim \mathrm{SBS}(\alpha,\alpha;H)$.

The following theorem allows us to compute the moments of SBS distributions.

Theorem 3 *Let $T \sim \mathrm{SBS}(\alpha,\beta;H)$. If the r.v. $g(U)$ given in equation (1) has finite moments of all order, then the k-th moment of T is given by*

$$\mathbb{E}[T^k] = \beta^k \sum_{i=0}^{k} \binom{2k}{2i} \sum_{j=0}^{i} \binom{i}{j} \omega_{k+j-i} \left[\frac{\alpha}{2}\right]^{2[k+j-i]}, \quad k=1,2,\ldots,$$

where $\omega_r = \mathbb{E}[\{g(U)\}^r]$.

Proof. The required result is obtained from the stochastic representation given in equation (5) and by repeated application of the binomial theorem. ■

Remark 2 By using Theorem 2, the negative moments of T can be obtained by the fact that β/T and T/β have the same distribution. Consequently, we get $\mathbb{E}[T^{-k}] = \mathbb{E}[T^k]/\beta^{2k}$, for $k = 1,2,\ldots$.

Corollary 1 *Let* $T \sim \mathrm{SBS}(\alpha,\beta;H)$. *Then, the mean, variance, and the coefficients of variation (CV), skewness (CS) and kurtosis (CK) of* T *are given by*

$$\mathbb{E}[T] = \frac{\beta}{2}[2+\omega_1\alpha^2], \quad \mathrm{Var}[T] = \frac{\beta^2\alpha^2}{4}\left[\omega_1 + \{2\omega_2 - \omega_1^2\}\alpha^2\right],$$

$$\gamma[T] = \frac{\alpha\left[4\omega_1 + \{2\omega_2-\omega_1^2\}\alpha^2\right]^{1/2}}{2+\omega_1\alpha^2},$$

$$\alpha_3[T] = \frac{4\alpha\left[\{3\omega_2 - 3\omega_1^2\} + \frac{1}{2}\{2\omega_3 - 3\omega_1\omega_2 + \omega_1^3\}\alpha^2\right]}{\left[4\omega_1 + \{2\omega_2 - \omega_1^2\}\alpha^2\right]^{3/2}}, \quad \textit{and}$$

$$\alpha_4[T] = \frac{16\omega_2 + \{32\omega_3 - 48\omega_1\omega_2 + 24\omega_1^3\}\alpha^2 + \{8\omega_4 - 16\omega_1\omega_3 + 12\omega_1^2\omega_2 - \omega_1^4\}\alpha^4}{\left[4\omega_1 + \{2\omega_2 - \omega_1^2\}\alpha^2\right]^2},$$

respectively.

Remark 3 The dimensionless ratios $\gamma[T]$, $\alpha_3[T]$, and $\alpha_4[T]$ are functionally independent of the scale parameter β, with the skewness and kurtosis being basically controlled by the shape parameter α.

The following theorem and its corollary provide conditional distributions that are used in Section 5 to implement the EM-algorithm for the ML estimation of the parameters of SBS models.

Theorem 4 *Let* $T \sim \mathrm{SBS}(\alpha,\beta;H)$. *Then, the r.v.* T *given* $U = u$, *which is denoted by* $T|(U = u)$, *follows the classical BS distribution with parameters* $\sqrt{g(u)}\,\alpha$ *and* β, *i.e.,* $T|(U = u) \sim \mathrm{BS}\big(\sqrt{g(u)}\,\alpha,\beta\big)$.

Proof. By using equation (5) and given $U = u$, we have $T = \beta\left[\alpha_u Z + \sqrt{\{\alpha_u Z\}^2+4}\right]^2/4$, where $\alpha_u = \alpha\sqrt{g(u)}$, which establishes the required result. ∎

Corollary 2 *Let* $T \sim \mathrm{SBS}(\alpha,\beta;H)$. *Then:*

(i) *The pdf of the r.v.* $U|(T = t)$ *is given by*

$$h_{U|T}(u|t) = \frac{\phi\big(a(t)|0,g(u)\big)\,h_U(u)}{\phi_{\mathrm{SMN}}\big(a(t)\big)}, \quad u > 0;$$

(ii) *The moments of the r.v.* $g(U)|(T = t)$ *are given by*

$$\mathbb{E}\left[\{g(U)\}^s \big| (T=t)\right] = \frac{1}{\phi_{\rm SMN}\big(a(t)\big)} \int_0^\infty [g(u)]^{s-\frac{1}{2}}\, \phi\left(\frac{a(t)}{\sqrt{g(u)}}\right) dH(u), \quad s \in \mathbb{R}.$$

Next, we present some transformations related to SBS distributions.

Theorem 5 *Let* $T \sim \mathrm{SBS}(\alpha,\beta;H)$. *Then:*

(i) *The pdf of the r.v.* $V = T^\eta$, *with* $\eta > 0$, *is given by*

$$f_V(v) = \phi_{\rm SMN}\left(\frac{1}{\alpha}\left[\left\{\frac{v}{\delta}\right\}^{1/\sigma} - \left\{\frac{\delta}{v}\right\}^{1/\sigma}\right]\right) \frac{1}{\alpha\sigma v}\left[\left\{\frac{v}{\delta}\right\}^{1/\sigma} + \left\{\frac{\delta}{v}\right\}^{1/\sigma}\right], \quad v > 0,$$

where $\delta = \beta^\eta$ *and* $\sigma = 2\eta$;

(ii) *The pdf of the r.v.* $V = \log(T)$ *is given by*

$$f_V(v) = \phi_{\rm SMN}\left(\frac{2}{\alpha}\sinh\left(\frac{v-\rho}{2}\right)\right)\frac{1}{\alpha}\cosh\left(\frac{v-\rho}{2}\right), \quad -\infty < v < \infty,$$

where $\rho = \log(\beta)$;

(iii) *The pdf of the r.v.*

$$V = \left[\frac{1}{\alpha}\left\{\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}}\right\}\right]^k$$

is given by

$$f_V(v) = \begin{cases} \frac{1}{k}\, v^{\frac{1}{k}-1}\, \phi_{\rm SMN}\big(v^{\frac{1}{k}}\big), \ -\infty < v < \infty, & \text{if } k \text{ is odd}, \\ \frac{2}{k}\, v^{\frac{1}{k}-1}\, \phi_{\rm SMN}\big(v^{\frac{1}{k}}\big), \ v > 0, & \text{if } k \text{ is even}. \end{cases}$$

Proof. Parts (i), (ii) and (iii) are proved by using the change-of-variable method. ■

Remark 4 The density given in Theorem 5(i) corresponds to the pdf of an extension of the SBS family, which we call the three-parameter SBS distributions, denoted by $T \sim \mathrm{SBS}(\alpha,\delta,\sigma;H)$. Note that $\sigma = 2$ produces the SBS family. Similarly, the density given in Theorem 5(ii) can be seen as the pdf of an extension of the sinh-normal distribution introduced by Rieck and Nedelman (1991).

Corollary 3 *Let* $T \sim \mathrm{SBS}(\alpha,\beta;H)$ *and* $V = \left[\sqrt{T/\beta} - \sqrt{\beta/T}\right]/\alpha$. *Then:*

(i) *The pdf of* $V_1 = |V|$ *is* $f_{V_1}(v_1) = 2\phi_{\rm SMN}(v_1)$, *for* $v_1 > 0$;

(ii) *The pdf of* $V_2 = V^2$ *is* $f_{V_2}(v_2) = \phi_{\rm SMN}(v_2)/\sqrt{v_2}$, *for* $v_2 > 0$;

(iii) *The pdf of* $V_3 = \exp(V)$ *is* $f_{V_3}(v_3) = \phi_{\rm SMN}\big(\log(v_3)\big)/v_3$, *for* $v_3 > 0$.

Remark 5 From Corollary 3, we see that the random variables V_1, V_2, and V_3 follow the half-symmetric, generalized chi-square with one degree of freedom (d.f.) and log-symmetric distributions, respectively. For more details on these distributions, one may refer to Fang *et al.* (1990).

4 Special cases of the SBS family

In this section, some special cases of the SBS family are considered, which are based on the contaminated normal, slash and t models. These are obtained from the stochastic representation given in equation (5), with $g(U) = 1/U$ and U having a known pdf. In addition, from Corollary 2, the conditional distribution of $U|(T = t)$ is also considered for all these special cases.

4.1 The contaminated normal BS distribution

As is well-known, contaminated normal models can be used for describing symmetric data with outlying observations, where one of the parameters represents the percentage of outliers, while the other one can be interpreted as a scale factor; see Little (1988). The contaminated normal distribution can be utilized for generating a BS distribution, which we call contaminated normal Birnbaum-Saunders (CN-BS) distribution. This model can be used for describing positively skewed non-negative data in the presence of atypical observations.

Consider the case when $T \sim \mathrm{SBS}(\alpha, \beta; H)$, with H being the cdf of the r.v. U, which has a pdf of the form

$$h_U(u) = \nu\,\mathbb{I}_{\{\gamma\}}(u) + [1-\nu]\,\mathbb{I}_{\{1\}}(u), \quad 0 < \nu < 1, 0 < \gamma < 1, \tag{7}$$

where $\mathbb{I}_A(\cdot)$ denotes the indicator function of the set A. Then, from equations (2), (6) and (7), we have the pdf of the r.v. T to be

$$f_T(t) = \left[\nu\,\sqrt{\gamma}\,\phi\left(\sqrt{\gamma}\,a(t)\right) + [1-\nu]\phi\left(a(t)\right)\right]\frac{t^{-3/2}[t+\beta]}{2\,\alpha\,\sqrt{\beta}}, \quad t > 0, \tag{8}$$

with $\alpha > 0, \beta > 0, 0 < \nu < 1$, and $0 < \gamma < 1$, where $\phi(\cdot)$ is the standard normal pdf and $a(t)$ is given as in equation (6). The model with pdf given as in equation (8) is the CN-BS distribution. In this case, the pdf of $U|(T = t)$ is given by

$$h_{U|T}(u|t) = \nu\,p(t,u)\,\mathbb{I}_{\{\gamma\}}(u) + [1-\nu]\,p(t,u)\,\mathbb{I}_{\{1\}}(u), \tag{9}$$

where

$$p(t,u) = \frac{\sqrt{u}\exp\left(-\frac{u\,a(t)^2}{2}\right)}{\nu\sqrt{\gamma}\exp\left(-\frac{\gamma\,a(t)^2}{2}\right) + [1-\nu]\exp\left(-\frac{a(t)^2}{2}\right)}.$$

Thus,

$$\mathbb{E}\big[U|(T=t)\big] = \frac{1-\nu+\nu\gamma^{3/2}\exp\left(\frac{[1-\gamma]a(t)^2}{2}\right)}{1-\nu+\nu\sqrt{\gamma}\exp\left(\frac{[1-\gamma]a(t)^2}{2}\right)}. \tag{10}$$

4.2 The slash Birnbaum-Saunders distribution

The slash distribution presents heavier tails than the normal one. In addition, when its shape parameter converges to infinity this distribution approaches the normal one. As in the case of the CN-BS distribution, the slash model can be utilized for generating a BS distribution, which we call slash Birnbaum-Saunders (SL-BS) distribution. A study that relates the BS and slash distributions has been done by Gómez *et al.* (2009).

Consider the case when $T \sim \text{SBS}(\alpha,\beta;H)$, with H being the cdf of the r.v. $U \sim \text{Beta}(\nu,1)$, which has a pdf of the form

$$h_U(u) = \nu\,u^{\nu-1}, \quad 0<u<1, \nu>0. \tag{11}$$

Then, from equations (2), (6) and (11), we have the pdf of the r.v. T to be

$$f_T(t) = \left[\nu\int_0^1 u^{\nu-1}\phi\left(a(t)\big|0,\tfrac{1}{u}\right)du\right]\frac{t^{-3/2}[t+\beta]}{2\alpha\sqrt{\beta}}, \quad t>0,\alpha>0,\beta>0,\nu>0. \tag{12}$$

The model with pdf given as in equation (12) is the SL-BS distribution. In this case, $U|(T=t) \sim \text{Gamma}\left(1/2+\nu, a(t)^2/2\right)$ truncated at $[0,1]$. Thus,

$$\mathbb{E}\big[U|(T=t)\big] = \left[\frac{1+2\nu}{a(t)^2}\right]\frac{P_1\left(\frac{3}{2}+\nu,\frac{a(t)^2}{2}\right)}{P_1\left(\frac{1}{2}+\nu,\frac{a(t)^2}{2}\right)}, \tag{13}$$

where $P_x(a,b)$ denotes the cdf of the Gamma distribution of parameters a and b evaluated at x according to the parameterization established in the pdf given in equation (14).

4.3 The Student-t BS distribution

The Student-*t* distribution with ν d.f., denoted by t_ν, has been used as an alternative model to the normal one for obtaining qualitatively robust parameter estimates; see Lange *et al.* (1989) and Lucas (1997). Special cases of the t_ν distribution are the Cauchy model, when $\nu = 1$, and the normal model, when $\nu \to \infty$. As in the case of the CN-BS and SL-BS distributions, the t_ν model can be utilized for generating a BS distribution, which we call Student-*t* Birnbaum-Saunders (t_ν-BS) distribution. The t_ν-BS distribution can be used for obtaining qualitatively robust parameter estimates with respect to the BS distribution; see Balakrishnan *et al.* (2007), Leiva *et al.* (2008), and Barros *et al.* (2008).

Consider the case when $T \sim \mathrm{SBS}(\alpha, \beta; H)$, with H being the cdf of the r.v. $U \sim \mathrm{Gamma}(\nu/2, \nu/2)$, which has a pdf of the form

$$h_U(u) = \frac{\left[\frac{\nu}{2}\right]^{\frac{\nu}{2}} u^{\frac{\nu}{2}-1}}{\Gamma(\frac{\nu}{2})} \exp\left(-\frac{\nu u}{2}\right), \quad u > 0, \nu > 0. \tag{14}$$

Then, from equations (2), (6) and (14), we have the pdf of the r.v. T to be

$$f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\sqrt{\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{1}{\nu\alpha^2}\left\{\frac{t}{\beta} + \frac{\beta}{t} - 2\right\}\right]^{-\frac{\nu+1}{2}} \frac{t^{-3/2}[t+\beta]}{2\,\alpha\sqrt{\beta}}, \quad t > 0, \tag{15}$$

with $\alpha > 0$, $\beta > 0$, and $\nu > 0$. The model with pdf given as in equation (15) is the t_ν-BS distribution. In this case, we have $U|(T = t) \sim \mathrm{Gamma}\big([\nu+1]/2, [\nu + a(t)^2]/2\big)$. Thus,

$$\mathbb{E}\big[U|(T = t)\big] = \frac{\nu+1}{\nu + a(t)^2}.$$

5 Lifetime analysis in the SBS family

A useful indicator in lifetime analysis is the failure rate, which, for a non-negative r.v. T with pdf f_T and cdf F_T, is defined as $r_T(t) = f_T(t)/[1 - F_T(t)]$, for $t > 0$, and $0 < F_T(t) < 1$. Although the distribution of T may be characterized equally in terms of the pdf or of the failure rate, according to Cox and Oakes (1984, pp. 24-28) and Balakrishnan *et al.* (2007), it is convenient to check the behaviour of the failure rate because distributions with densities whose shapes are similar could have failure rates with different shapes. If $r_T(t)$ is an increasing or decreasing function in t, then the distribution T belongs to the class of increasing failure rate (IFR) or decreasing failure rate (DFR) distributions, respectively. If $r_T(t) = \lambda > 0$, for $t > 0$, we have $F_T(t) = 1 - \exp(-\lambda t)$, and F_T is the exponential cdf with parameter λ. However, there are distributional families that have a non-monotone failure rate. In this case, an important value for lifetime analysis is the change point of the failure rate of T (denoted by t_c), which is the value where the hazard changes its behaviour. Within the class of distributions with a non-monotone failure rate, we can identify $\cap$-or-$\cup$ shapes. Particularly, for the $\cap$-shaped case, we also have two cases, when the failure rate is initially increasing until its change point and then: (i) decreases to zero, as in the case of the lognormal distribution or (ii) decreases until it becomes stabilized at a positive constant, as in the case of the classical BS distribution. For this reason, for distributional families with a non-monotone failure rate, their change point and their limiting behaviour are aspects that should be studied. For more details about life distributions and lifetime analysis, see Johnson *et al.* (1995, pp. 651-663), Marshall and Olkin (2007), and Saunders (2007).

As mentioned, the BS model belongs to the upside-down (or $\cap$-shaped) class and its failure rate approaches $1/[2\alpha^2\beta]$ as $t \to \infty$; see Chang and Tang (1993). A complete study of the change point of the BS failure rate can be found in Kundu *et al.* (2008) and Bebbington *et al.* (2008). Next, we give some results related to the failure rate of SBS distributions.

Theorem 6 *Let $T \sim \mathrm{SBS}(\alpha, \beta; H)$. Then, the failure rate of T is*

$$r_T(t) = \frac{\phi_{\mathrm{SMN}}\big(a(t)\big)\,A(t)}{\Phi_{\mathrm{SMN}}\big(-a(t)\big)}, \quad t > 0,\ 0 < \Phi_{\mathrm{SMN}}(\cdot) < 1,$$

where $a(t)$ and $A(t)$ are given as in equation (6) and $\Phi_{\mathrm{SMN}}(\cdot)$ is the cdf of the SMN family.

Proof. It follows immediately direct from the definition of the failure rate and the SMN symmetry. ∎

Theorem 7 *Let $T \sim \mathrm{SBS}(\alpha, \beta; H)$ and $r_T(\cdot)$ be its failure rate. Then,*

$$\lim_{t\to\infty} r_T(t) = \frac{1}{2\alpha^2\beta} \lim_{t\to\infty} W_{g,H}\left(a(t)^2\right), \tag{16}$$

where

$$W_{g,H}\big(a(t)^2\big) = \frac{\int_0^\infty g^{-3/2}(u)\exp\left(-\frac{a(t)^2}{2g(u)}\right) dH(u)}{\int_0^\infty g^{-1/2}(u)\exp\left(-\frac{a(t)^2}{2g(u)}\right) dH(u)}.$$

Proof. For $T \sim \mathrm{SBS}(\alpha, \beta; H)$, we have $f_T(t) = \phi_{\mathrm{SMN}}(a(t))A(t)$ and a function $f(\cdot)$ such that $\phi_{\mathrm{SMN}}(y) = f(y^2)$, for all $y \in \mathbb{R}$. In this case,

$$f(w) = \int_0^\infty \frac{1}{\sqrt{2\pi\, g(u)}} \exp\left(-\frac{w}{2g(u)}\right) dH(u), \quad w \geq 0. \tag{17}$$

Thus, $W_f(w) = f'(w)/f(w) = -W_{g,H}(w)/2$. As $g(\cdot)$ is a positive function, we have $W_{g,H}(w) \geq 0$. The proof of this theorem is similar to the one given in Theorem 4 of Leiva *et al.* (2008). ■

Theorem 8 *Let $T \sim \text{SBS}(\alpha, \beta; H)$ and that the distribution of U is unimodal. Then, the pdf of the T is unimodal and the mode, denoted by t_m, is obtained as solution of*

$$W_{g,H}\big(a(t_m)^2\big) = -\frac{\alpha^2 \beta t_m\,[t_m + 3\beta]}{[t_m - \beta][t_m + \beta]},$$

where $0 < t_m < \beta$.

Proof. From equation (17), we have that $f(w)$ is a monotonic non-increasing function for all $w > 0$, and so $\phi_{\text{SMN}}(\cdot)$ is a unimodal function. The rest of the proof follows by using Equation (8) of Leiva *et al.* (2008), replacing $-W_{g,H}(u)/2$ by $w_g(u)$, for $u > 0$. ■

Theorem 9 *The failure rate of SBS distributions is an upside-down function for all values of α and β.*

Proof. Following the same procedure as in Kundu *et al.* (2008), we can write the SBS failure rate as

$$r_T(t) = \frac{\phi_{\text{SMN}}(a(t))\,A(t)}{\Phi_{\text{SMN}}\big(-a(t)\big)}, \quad t > 0.$$

Thus, it is enough to prove that $\lim_{t\to 0} r_T(t) = 0$. As $f_T(t) = \phi_{\text{SMN}}(a(t))\,A(t)$, this can be expressed as

$$f_T(t) = \frac{1}{2\alpha\beta^{1/2}}\left[\int_0^\infty \frac{1}{\sqrt{2\pi g(u)}}\,\Delta_1(t,u)\,dH(u) + \beta\int_0^\infty \frac{1}{\sqrt{2\pi g(u)}}\,\Delta_2(t,u)\,dH(u)\right],$$

where $\Delta_1(t,u) = t^{-1/2}\exp\big(-a(t)^2/[2g(u)]\big)$ and $\Delta_2(t,u) = t^{-3/2}\exp\big(-a(t)^2/[2g(u)]\big)$. Then, following Kundu *et al.* (2008), we have that $\lim_{t\to 0}\Delta_1(t,u) = \lim_{t\to 0}\Delta_2(t,u) = 0$. Thus, since $\lim_{t\to 0} f_T(t) = 0$ and $\lim_{t\to 0}\Phi_{\text{SMN}}(-a(t)) = 1$, we have $\lim_{t\to 0} r_T(t) = 0$. ■

Remark 6 Note that Theorems 6 and 7 contain the expression of $W_{g,H}(\cdot)$. Next, we specify this expression for some particular cases (indicated in brackets) and obtain the limit of $r_T(t)$ as $t \to \infty$.

(i) [**CN-BS distribution**] Since

$$W_{g,H}\big(a(t)^2\big) = \frac{1 - \nu + \nu\gamma^{3/2}\exp\left(\frac{[1-\gamma]a(t)^2}{2}\right)}{1 - \nu + \nu\sqrt{\gamma}\exp\left(\frac{[1-\gamma]a(t)^2}{2}\right)},$$

then $\lim_{t\to\infty} r_T(t) = \gamma/[2\alpha^2\beta]$; note that if $\gamma = 1$, we have the case of the classical BS distribution;

(ii) [**SL-BS distribution**] Since

$$W_{g,H}\left(a(t)^2\right) = \left[\frac{1+2\nu}{a(t)^2}\right] \frac{P_1\left(\frac{3}{2}+\nu, \frac{a(t)^2}{2}\right)}{P_1\left(\frac{1}{2}+\nu, \frac{a(t)^2}{2}\right)},$$

where $P_x(a,b)$ denotes the cdf of the Gamma distribution of parameters a and b evaluated at x according to the parameterization established in the pdf given in equation (14), then $\lim_{t\to\infty} r_T(t) = 0$; note that, in this case, the BS class has a failure rate similar to that of the lognormal distribution;

(iii) [t_ν**-BS distribution**] Since

$$W_{g,H}\left(a(t)^2\right) = \frac{\nu+1}{\nu+a(t)^2},$$

then $\lim_{t\to\infty} r_T(t) = 1/[2\alpha^2\beta]$, if $\nu \to \infty$, which corresponds to the case of the classical BS distribution; however, if $\nu \to 0$, then $\lim_{t\to\infty} r_T(t) = 0$, which is also the case when $\nu = 1$ corresponding to the Cauchy-BS distribution, such as occurs with the failure rate of the lognormal and SL-BS distributions.

Figure 1 shows different shapes of the failure rate of SBS distributions through which is possible compare their shapes to those of the classical BS model. This graphical analysis is coherent with the results given in this section.

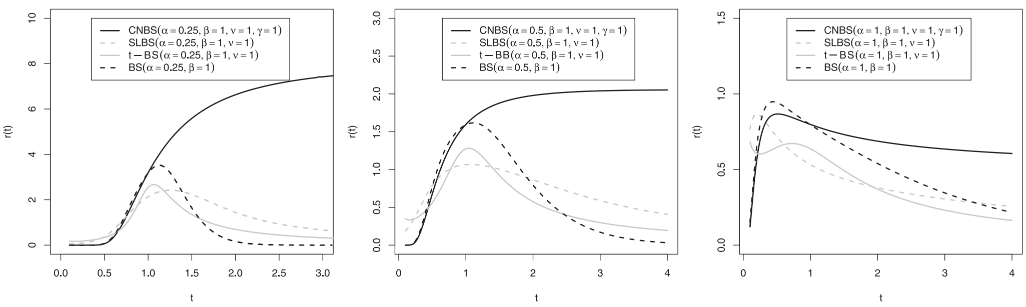


Figure 1: *Failure rate plots for the indicated distributions for some choices of the parameters.*

6 ML estimation via EM-algorithm in the SBS class

The EM-algorithm is a well-known tool for ML estimation when unobserved (or missing) data or latent variables are present while modeling. This algorithm enables the computationally efficient determination of the ML estimates when iterative procedures are required. Specifically, let $\mathbf{t} = [t_1,\ldots,t_n]^\top$ and $\mathbf{u} = [u_1,\ldots,u_n]^\top$ denote observed and unobserved data, respectively. The complete data $\mathbf{t}_c = [\mathbf{t}^\top,\mathbf{u}^\top]^\top$ corresponds to the original data $\mathbf{t}$ augmented with $\mathbf{u}$. We now detail the implementation of the ML estimation of parameters of SBS distributions by using the EM-algorithm.

Let $T_1,\ldots,T_n$ be a random sample of size n, where $T_i \sim \mathrm{SBS}(\alpha,\beta;H)$, for $i = 1,\ldots,n$. Here, the parameter vector is $\boldsymbol{\theta} = [\alpha,\beta]^\top$, with $\boldsymbol{\theta} \in \boldsymbol{\Theta} \equiv \mathbb{R}^+ \times \mathbb{R}^+$. Let $\ell_c(\boldsymbol{\theta}\,|\mathbf{t}_c)$ and $Q(\boldsymbol{\theta}\,|\widehat{\boldsymbol{\theta}}) = \mathbb{E}[\ell_c(\boldsymbol{\theta}\,|\mathbf{t}_c)|\mathbf{t},\widehat{\boldsymbol{\theta}}]$ denote the complete-data log-likelihood function and its expected value conditioned to the observed-data, respectively. Each iteration of the EM algorithm involves two steps, i.e., the expectation step (E-step) and the maximization step (M-step), which are defined by:

E-step. Compute $Q\big(\boldsymbol{\theta}\,|\widehat{\boldsymbol{\theta}}^{(r)}\big)$, for $r = 1,2,\ldots$.

M-step. Find $\boldsymbol{\theta}^{(r+1)}$ such that $Q\big(\boldsymbol{\theta}^{(r+1)}\,|\widehat{\boldsymbol{\theta}}^{(r)}\big) = \max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} Q\big(\boldsymbol{\theta}\,|\widehat{\boldsymbol{\theta}}^{(r)}\big)$, for $r = 1,2,\ldots$.

Note that, by using Theorem 4, the above setup can be written as

$$T_i|(U_i = u_i) \overset{\text{ind}}{\sim} \sim \mathrm{BS}\big(\sqrt{g(u_i)}\,\alpha,\,\beta\big), \tag{18}$$

$$U_i \overset{\text{ind}}{\sim} h_U(u_i), \quad i = 1,\ldots,n. \tag{19}$$

We assume that the parameter vector $\boldsymbol{\nu}$ that indexes the pdf $h_U(\cdot)$ is known. An optimal value of $\boldsymbol{\nu}$ can then be chosen by using the Schwarz information criterion; see Spiegelhalter *et al.* (2002). Thus, under the hierarchical representation given in equations (18) and (19), it follows that the complete log-likelihood function associated with $\mathbf{t}_c = [\mathbf{t}^\top,\mathbf{u}^\top]^\top$ is given by

$$\ell_c(\boldsymbol{\theta}\,|\mathbf{t}_c) \propto -n\log(\alpha) - \frac{n}{2}\log(\beta) - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\frac{1}{g(u_i)}\left[\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right] + \sum_{i=1}^{n}\log(t_i+\beta). \tag{20}$$

Letting $\widehat{u}_i = \mathbb{E}[1/g(U_i)|t_i,\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}]$, for $i = 1,\ldots,n$, it follows that the conditional expectation of the complete log-likelihood function has the form

$$Q(\boldsymbol{\theta}\,|\widehat{\boldsymbol{\theta}}) \propto -n\log(\alpha) - \frac{n}{2}\log(\beta) - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\widehat{u}_i\left[\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right] + \sum_{i=1}^{n}\log(t_i+\beta). \tag{21}$$

We then have the EM-algorithm for the ML estimation of the parameters of the SBS distributions as follows:

E-step. Given $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, compute $\widehat{u}_i$, for $i = 1,\ldots,n$;

M-step. Update $\widehat{\boldsymbol{\theta}}$ by maximizing $Q(\boldsymbol{\theta}\,|\widehat{\boldsymbol{\theta}})$ in equation (21) over $\boldsymbol{\theta}$, which leads to the following expressions:

$$\widehat{\alpha}^2 = \frac{S_u}{\widehat{\beta}} + \frac{\widehat{\beta}}{R_u} - 2\bar{u} \quad \text{and} \quad \widehat{\beta}^2 - \widehat{\beta}\left[k(\widehat{\beta}) + 2\bar{u}R_u\right] + R_u\left[\bar{u}k(\widehat{\beta}) + S_u\right] = 0, \tag{22}$$

where

$$\bar{u} = \frac{1}{n}\sum_{i=1}^{n}\widehat{u}_i, \quad S_u = \frac{1}{n}\sum_{i=1}^{n}\widehat{u}_i t_i, \quad R_u = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{u}_i}{t_i}}, \quad \text{and} \quad k(x) = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x+t_i}}. \tag{23}$$

Remark 7 Note that if $g(U) = 1$ in the EM-algorithm presented above (i.e., if the r.v. U is degenerate), then the M-step equations reduce to those when the BS distribution is used. Thus, the EM-algorithm here generalizes the results provided earlier by Birnbaum and Saunders (1969b). Moreover, the presented procedure provides an EM-algorithm for the t_ν-BS distribution, which has been studied recently by Balakrishnan *et al.* (2007), Leiva *et al.* (2008), and Barros *et al.* (2008). Useful starting values necessary to implement this algorithm can be the ML estimates of the parameters of the BS distribution.

7 Illustrative numerical example

In this section, for the purpose of illustration, we analyze the data of Birnbaum and Saunders (1969b). These data correspond to fatigue life represented by cycles ($\times 10^{-3}$) until failure of aluminum specimens of type 6061-T6. These specimens were cut parallel to the direction of rolling and oscillating at 18 cycles per seconds. They were exposed to a pressure with maximum stress of 21,000 (Psi21), 26,000 (Psi26) and 31,000 (Psi31) pounds per square inch (psi) for $n = 101, 102$, and 101 specimens, respectively. All specimens were tested until failure.

We first present an exploratory data analysis. Table 1 provides a descriptive summary while Figure 2 shows the histograms and boxplots for Psi21, Psi26, and Psi31.

A careful look at Table 1 and Figure 2 reveals slightly positively skewed distributions with moderate kurtosis and some atypical observations, which can be potentially influential on the ML estimates of the parameters of the BS distribution. SBS distributions should consider the degrees of skewness and kurtosis present in the data. In addition, they also enable the estimation of the parameters of the model in a robust manner when atypical observations are present.

We now find the ML estimates of the parameters α and β of SBS distributions. Several authors have suggested to fix the parameter $\boldsymbol{\nu}$ for the distribution of the r.v. U defined in equation (1) and assume it to be a known value or otherwise get information for it from the data. For instance, in the case of the Student-t_ν distribution, the reason for doing it is that only when the parameter ν is fixed, the influence function is bounded, which allows us to obtain qualitatively robust estimators of parameters.

Table 1: *Descriptive statistics for the indicated data sets*

Data set	Mean	Median	StDev	CV	CS	CK	Range	Min.	Max.	n
Psi21	1400.84	1416.00	391.01	27.91%	0.14	−0.28	2070	370	2440	101
Psi26	397.88	400.00	62.32	15.66%	0.01	−0.21	327	233	560	102
Psi31	133.73	133.00	22.36	16.70%	0.33	0.97	142	70	212	101

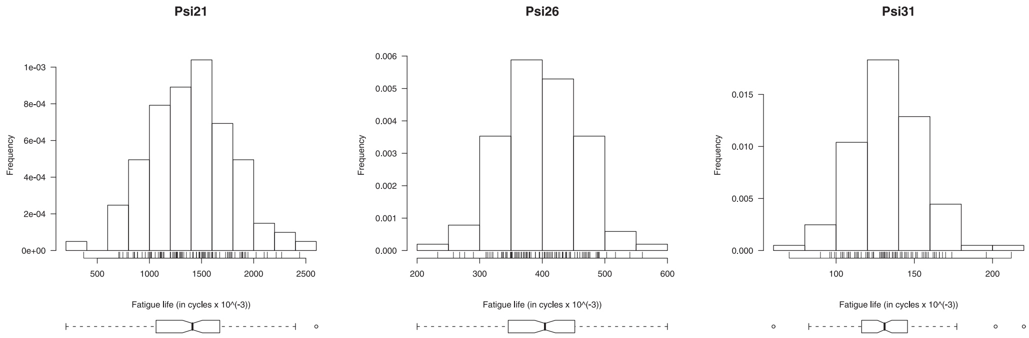


Figure 2: *Histograms and boxplots for the indicated data sets.*

For more details about these proposals and their justification, one can refer to Lucas (1997); see also Lange *et al.* (1989) and Leiva *et al.* (2008).

In order to select the best SBS model that fits the data, we have implemented the estimation procedure described in Section 6 in `R` code (`http://www.R-project.org`); see R Development Core Team (2008). As mentioned earlier, we use the ML estimates of α and β of the classical BS distribution as starting values in the numerical iterative procedure, which can be easily obtained from an `R` package called `bs` that is available from CRAN (`http://CRAN.R-project.org`); see Leiva *et al.* (2006). For ML estimation via EM-algorithm in SBS models, we have implemented the command `smnbsEstimation()`, which automatically chooses the distribution that best fits the data set among the CN-BS, SL-BS and t_ν-BS distributions by maximizing the likelihood function. This command also computes the ML estimates of the parameters of SBS models separately. For instance, in the case of the t_ν-BS distribution, the following algorithm can be used for estimating its parameters:

(A1) For $\nu = 1$ to $\nu = 100$ by 1:
- **(A1.1)** Determine the ML estimates of the parameters α and β of the t_ν-BS model via the EM-algorithm proposed in Section 6 by beginning with the ML estimates of α and β of the BS distribution as starting values for the numerical procedure;
- **(A1.2)** Compute the likelihood function;

(A2) Choose the value of ν that maximizes the likelihood function and then establish as ML estimates of α and β those associated with that maximum likelihood function.

The real data sets used in this example are implemented in the `bs` package, which are called `psi21`, `psi26` and `psi31` and obtained by using the instructions `data(psi21)`, `data(psi26)`, and `data(psi31)`, respectively. Now, if the following commands are used

```
> smnbsEstimation(psi21)
> smnbsEstimation(psi26)
> smnbsEstimation(psi31)
```

then the distributions that best fit the Psi21, Psi26 and Psi31 data set are chosen among the CN-BS, SL-BS and t_ν-BS models. In addition, the ML estimates of α and β of these models are computed. The results can be saved in `R` variables as follows:

```
> estimatespsi21 <- smnbsEstimation(psi21)
> estimatespsi26 <- smnbsEstimation(psi26)
> estimatespsi31 <- smnbsEstimation(psi31)
```

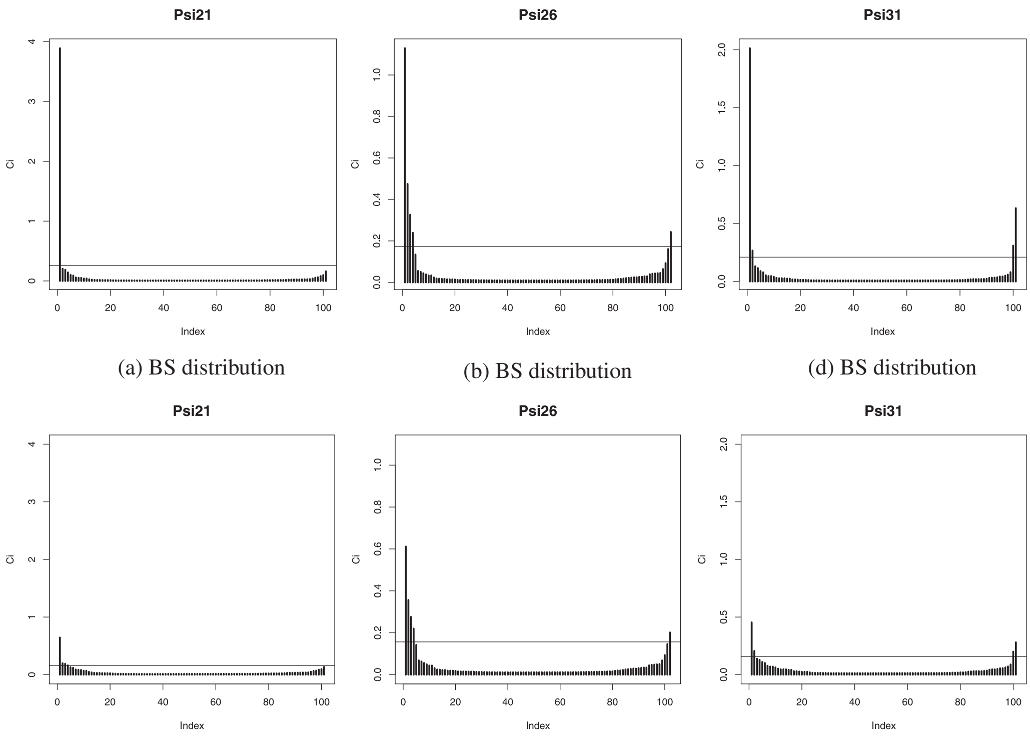
obtaining, respectively,

```
> smnbsEstimation(psi21)
$Best model
[1] "CN-BS"
$alpha
[1] 0.2737684
$beta
[1] 1356.624
$nu
[1] 0.02
$gamma
[1] 0.08
$logLikelihood
[1] -747.3013
> smnbsEstimation(psi26)
$Best model
[1] "BS-t"
$alpha
[1] 0.1534232
$beta
[1] 393.9034
$nu
[1] 18
$logLikelihood
[1] -567.4124
```

and

```
> smnbsEstimation(psi31)
$Best model
[1] "CN-BS"
$alpha
[1] 0.1465890
$beta
[1] 132.2940
$nu
[1] 0.06
$gamma
[1] 0.18
$logLikelihood
[1] -455.4714
```

From these results and within the three considered distributions, we can see that the CN-BS, t_{18}-BS, CN-BS distributions present the best fit to the Psi21, Psi26 and Psi31 data sets, respectively.



(a) BS distribution (b) BS distribution (d) BS distribution

(d) SBS distribution (e) SBS distribution (f) SBS distribution

Figure 3: *Influence index plots for the indicated data sets and models*

In order to show the inherent robustness of the estimation procedure based on Birnbaum-Saunders distributions from scale-mixture of normals, we carry out a brief diagnostic analysis based on local influence and relatives changes. For more details about local influence, see Cook (1986).

In Figure 3, we can observe the inherent robustness of the estimation procedure based on SBS distributions. Values of C_i, for $i = 1,\ldots,n$, –total local influence for the ith case– show a more pronounced potential influence of observations for the classical BS model. For more details about the local influence techniques in BS models, see Galea *et al.* (2004) and Leiva *et al.* (2007).

Table 2 presents the relative changes (RC), in percentage, of each parameter estimate, defined by $\text{RC}_{\theta_j} = |[\hat{\theta}_j - \hat{\theta}_{j(\text{I})}]/\hat{\theta}_j| \times 100\%$, for $j = 1,2$, with $\theta_1 = \alpha$ and $\theta_2 = \beta$, where $\hat{\theta}_{j(\text{I})}$ denotes the ML estimate of θ_j after the set I of cases has been removed. From this table, we note that the RC values are greater for the classical BS model than for the SBS models. Thus, all the specimens are retained in the analysis as they do not greatly affect the ML estimates under the SBS models.

Table 2: *RC (in %) for the indicated parameters, models and data sets*

		SBS		BS	
Data set	Dropped case(s)	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
S21	$\{1\}$	3.00	1.02	9.83	1.60
	$\{101\}$	1.41	0.56	1.48	0.61
	$\{1,101\}$	4.44	0.49	11.5	0.99
S26	$\{1\}$	2.35	0.20	4.96	0.53
	$\{102\}$	1.66	0.22	1.96	0.35
	$\{1,102\}$	4.00	0.02	7.02	0.18
S31	$\{1\}$	2.73	0.19	6.90	0.66
	$\{101\}$	2.25	0.18	3.56	0.48
	$\{1,101\}$	4.94	0.00	10.7	0.18

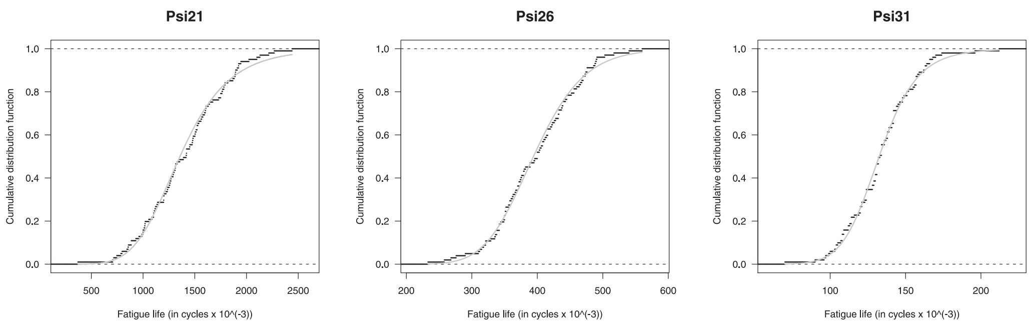


Figure 4: *Empirical cdf (in bold) against estimated SBS theoretical cdf (in gray) for the indicated data.*

In order to show how the SBS family fits the data, we use the invariance property of the ML estimators for obtaining the estimated SBS cdf, which is shown in Figure 4 on the empirical cdf of the data. In addition, the application of the Kolmogorov-Smirnov test provides the p-values 0.765, 0.974, and 0.799 for the Psi21, Psi26 and Psi31 data sets, respectively. These results suggest an excellent agreement between the SBS models and the data.

8 Concluding remarks

We have introduced a general class of Birnbaum-Saunders distributions based on scale mixtures of normal distributions. This class allows us to obtain qualitatively robust maximum likelihood estimates and efficiently compute these by using the EM-algorithm. Specifically, we have found the pdf, shown some properties, computed the moments, considered some transformations, and carried out a lifetime analysis based on the failure rate of scale-mixture Birnbaum-Saunders distributions. We have also presented some particular cases of these distributions based on the contaminated normal, slash and t models. In addition, we have implemented in `R` code different aspects pertaining to the considered distributions, including the mentioned EM-algorithm for determining the ML estimates of their parameters, which can make this model more attractive to users. Moreover, we have illustrated the results obtained for this class of distributions and discussed the computational implementation of them by using a numerical example with three different data sets, which display the flexibility, adequacy, and inherent robustness of the estimation procedure based on a Birnbaum-Saunders distribution from scale-mixture of normals.

8 Acknowledgments

The authors wish to thank the editors and referees for their helpful comments that aided in improving this article. This study was partially supported by a FAPESP grant from Brazil and FONDECYT 1080326, DIPUV 29-2006 and FONDECYT 1090265 grants from Chile.

References

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

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, 316-332.

Bebbington, M., Lai, C-D, and Zitikis, R. (2008). A proof of the shape of the Birnbaum-Saunders hazard rate function. *The Mathematical Scientist*, 33, 49-56.

Birnbaum, Z. W. and Saunders, S. C. (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.

Chang, D. S. and Tang, L. C. (1993). Reliability bounds and critical time for the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability*, 42, 464-469.

Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of The Royal Statistical Society Series B–Statistical Methodology*, 48, 133-169.

Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data.* Chapman & Hall, London.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of The Royal Statistical Society Series B–Statistical Methodology*, 39, 1-38.

Efron, B. and Olshen, R. A. (1978). How broad is the class of normal scale mixtures. *Annals of Statistics*, 6, 1159-1164.

Fang, K. T., Kotz, S. and Ng. K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman & Hall, London.

Fernandez, C. and Steel, M. (1999). Multivariate student *t* regression models: pitfalls and inference. *Biometrika*, 86, 153-167.

Kelker, D. (1970). Distribution theory of spherical distributions and a location scale parameter generalization. *Sankhyā A*, 32, 419-430.

Galea, M., Leiva, V. and Paula, G. A. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics*, 31, 1049-1064.

Gneiting, T. (1997). Normal scale mixtures and dual probability densities. *Journal of Statistical Computation and Simulation*, 59, 375-384.

Gómez, H. W., Olivares-Pacheco, J. F. and Bolfarine, H. (2008). An extension of the generalized Birnbaum-Saunders distribution. *Statistical and Probability Letters*, 79, 331-338.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions-Vol. 2*, Second edition. Wiley, New York.

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

Lachos, V. H. and Vilca, F. (2007). Skew-normal/independent distributions with applications. Technical report UNICAMP, Brazil. `http://www.unicamp.br/anuario/2007/IMECC/DE/DE--0040.html`

Lange, K. L., Little, J. A. and Taylor, M. G. J. (1989). Robust statistical modeling using the *t* distribution. *Journal of the American Statistical Association*, 84, 881-896.

Lange, K. and Sinsheimer, J. S. (1993). Normal/independent distributions and their applications in robust regression. *Journal of Computational and Graphical Statistics*, 2, 175-198.

Leiva, V., Hernández, H. and Riquelme, M. (2006). A new package for the Birnbaum-Saunders distribution. *The R Journal (R News)*, 6, 35-40.

Leiva, V., Barros, M., Paula, G. A. and Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis*, 51, 5694-5707.

Leiva, V., Riquelme, M., Balakrishnan, N. and Sanhueza, A. (2008). Lifetime analysis based on the generalized Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*, 52, 2079-2097.

Little, R. J. A. (1988). Robust estimation of the mean and covariance matrix from data with missing values. *Applied Statistics*, 37, 23-38.

Lucas, A. (1997). Robustness of the student t based M-estimator. *Communications in Statistics: Theory and Methods*, 26, 1165-1182.

Marshall, A. W. and Olkin, I. (2007). *Life Distributions*. Springer, New York.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

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

Sanhueza, A., Leiva, V. and Balakrishnan, N. (2008). The generalized Birnbaum-Saunders distribution and its theory, methodology and application. *Communications in Statistics: Theory and Methods*, 37, 645-670.

Saunders, S. C. (1974). A family of random variables closed under reciprocation. *Journal of the American Statistical Association*, 69, 533-539.

Saunders, S. C. (2007). *Reliability, Life Testing and Prediction of Services Lives*. Springer, New York.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of complexity and fit. *Journal of The Royal Statistical Society Series B–Statistical Methodology*, 64, 1-34.

Taylor, J. and Verbyla, A. (2004). Joint modeling of location and scale parameters of t distribution. *Statistical Modelling*, 4, 91-112.

Walker, S. G. and Gutiérrez-Peña, E. (2007). Bayesian parametric inference in a nonparametric framework. *Test*, 16, 188-197.