Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v12n1p26

A generalized Birnbaum-Saunders distribution with application to the air pollution data

By Tamandi, Jamalizadeh, Mahdizadeh

Published: 26 April 2019

Electronic Journal of Applied Statistical Analysis
Vol. 12, Issue 01, April 2019, 26-43
DOI: 10.1285/i20705948v12n1p26

A generalized Birnbaum-Saunders distribution with application to the air pollution data

Mostafa Tamandi*[a], Ahad Jamalizadeh[b], and M. Mahdizadeh[c]

[a,b]*Department of Statistics, Faculty of Mathematics and Computers, Shahid Bahonar University of Kerman, Kerman, Iran*
[c]*Department of Statistics, Hakim Sabzevari University, P.O. Box 397, Sabzevar, Iran*

Published: 26 April 2019

Birnbaum-Saunders (BS) distribution is a model with positive domain that is used in many fields including reliability and environmental studies. This article introduces a generalized version of the BS distribution which arises from the shape mixture of skew-normal distribution. A feasible EM type algorithm is developed to obtain maximum likelihood (ML) estimates of parameters of the new model. The asymptotic standard errors of ML estimates are obtained via the information-based approximation. The robustness and application of the proposed methodology are illustrated through simulation studies and air pollution analysis.

Keywords: ECM algorithm, Observed information matrix, Robustness, Shape mixtures.

1 Introduction

Birnbaum and Saunders (1969) introduced a two-parameter positive distribution, and used it in analyzing fatigue failure time data. Their distribution, here after called Birnbaum-Saunders (BS), is better suited for modeling data with extreme observations, as compared with lifetime distributions, such as gamma, Weibull and inverse Gaussian.

*Corresponding author: tamandi.m@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

Suppose Z is a standard normal random variable (denoted by $Z \sim N(0,1)$), and α and β are positive constants. Then, the random variable

$$T = \frac{\beta}{4}\left[\alpha Z + \sqrt{(\alpha Z)^2 + 4}\right]^2 \quad (1)$$

is said to have the BS distribution with shape parameter α and scale parameter β. The probability density function (PDF) of T is given by

$$g(t;\alpha,\beta) = \phi\left(a(t;\alpha,\beta)\right) A(t;\alpha,\beta), \quad t>0; \alpha,\beta>0$$

where $\phi(\cdot)$ is the PDF of Z,

$$a(t;\alpha,\beta) = \frac{1}{\alpha}\left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right] \quad (2)$$

and

$$A(t;\alpha,\beta) = \frac{t+\beta}{2\alpha\sqrt{\beta}\sqrt{t^3}}. \quad (3)$$

The BS distribution and its extensions have received considerable attention with regard to theoretical properties, inference and applications. Desmond (1985) presented a more general derivation of the distribution using a biological model. He motivated this model with several examples in the context of engineering. Louis Floyd (1998) provided an application in the area of risk management. Cordeiro and Lemonte (2011) proposed beta BS distribution, and used it in studying the number of successive failures for the air conditioning system. Cordeiro et al. (2016) introduced an extension of the BS distribution, and employed it in investigating shocks before failure and accelerated life testing in a system. Also the BS distribution and its generalizations have been utilized in environmental studies. See, among others Vilca et al. (2010), Ferreira et al. (2012), and Marchant et al. (2013).

Azzalini (1985) developed skew-normal (SN) distribution as an extension of the normal distribution. Let $\phi(\cdot)$ and $\Phi(\cdot)$ be the PDF and cumulative distribution function of $Z \sim N(0,1)$, respectively. The random variable Y is said to have SN model with shape parameter λ, denoted by $Y \sim SN(\lambda)$, if its PDF is given by

$$f_{SN}(y;\lambda) = 2\phi(y)\Phi(\lambda y), \quad y \in \mathbb{R}; \lambda \in \mathbb{R}.$$

Arellano-Valle et al. (2004) introduced shape mixture of skew-normal (SMSN) distribution. It has the following hierarchical formulation. If $\tau \sim N(\gamma,\delta)$ and $V|\tau \sim SN(\tau)$, then the marginal distribution of V is the SMSN, denoted by $V \sim SMSN(\gamma,\delta)$. A further representation of SMSN model is given by

$$V \stackrel{d}{=} \frac{1}{\sqrt{1+\tau^2}} U_1 + \frac{\tau}{\sqrt{1+\tau^2}} \left|U_0\right|, \quad (4)$$

where U_0 and U_1 are independently distributed as $N(0,1)$.

In this work, we propose a new generalization of the BS distribution called shape mixture of skew normal BS (SMSN-BS). It is constructed by replacing Z in (1) by a SMSN random variable. Owing to the additional flexibility of SN model over the standard normal distribution, the SMSN-BS distribution is expected to be a viable alternative to the BS distribution. The basic motivation of this study is to construct a model with adequate fit to data that are highly concentrated on the right-tail of the distribution. As compared with current modifications of the BS distribution, our model enjoys the advantage of better fit to data in the presence of outliers. It also covers a variety of functional forms including bimodal distributions, which happen in many practical situations.

In Section 2, the new model is introduced and some of its properties are investigated. In Section 3, an expectation-maximization (EM) algorithm is developed to find maximum likelihood (ML) estimates of the parameters. Also, a general information-based method for obtaining the asymptotic standard errors of the ML estimates is presented. In Section 4, properties of the ML estimators, and flexibility of the proposed distribution are studied using numerical experiments. Section 5 illustrates application of the SMSN-BS distribution for modeling air pollution data. Final conclusions are given in Section 6.

2 The proposed distribution

If $V \sim SMSN(\gamma, \delta)$, and α and β are positive constants, then the random variable

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

follows a SMSN-BS distribution, denoted by $T \sim SMSN - BS(\alpha, \beta, \gamma, \delta)$. This representation can be used to draw random samples from the SMSN-BS distribution.

To present a feasible EM-type algorithm for maximum likelihood estimation of the parameters, we introduce a hierarchical representation of SMSN-BS distribution. Before that, it is nessesary to address two distributions which are used in this representation.

Leiva et al. (2010) introduced an extended BS (EBS) distribution using a skewed sinh-normal model. A random variable T follows EBS distribution, denoted by $T \sim EBS(\alpha, \beta, \sigma, \nu, \lambda)$, if its PDF is given by

$$f(t; \alpha, \beta, \sigma, \nu, \lambda) = 2\phi(c(t))\Phi(\lambda c(t))C(t), \qquad t > 0,$$

where $c(t) = \nu + \frac{1}{\alpha}\left[(\frac{t}{\beta})^{1/\sigma} - (\frac{\beta}{t})^{1/\sigma}\right]$ and $C(t) = \frac{t^{2/\sigma}+\beta^{2/\sigma}}{\sigma\alpha\beta^{1/\sigma}t^{1+1/\sigma}}$.

The second distribution which will be used in hierarchical representation is truncated normal (TN). A random variable X has a TN distribution, denoted by $X \sim TN(\mu, \sigma^2; (a, b))$, if it has the following PDF

$$f(x; \mu, \sigma, a, b) = \frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\sigma\left(\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right)}, \qquad a \le x \le b.$$

Now, let $\psi = \sqrt{1+\tau^2}\,|U_0|$ be a reparameterized latent variable, where $\tau \sim N(\gamma,\delta)$ and $U_0 \sim N(0,1)$. The SMSN-BS distribution has the following hierarchical formulation. If

$$\begin{aligned} \tau &\sim N(\gamma,\delta) \\ \psi|\tau &\sim TN\left(0,1+\tau^2;(0,\infty)\right) \\ (T|\psi,\tau) &\sim EBS\left(\frac{\alpha}{\sqrt{1+\tau^2}},\beta,2,-\frac{\tau\psi}{\sqrt{1+\tau^2}},0\right), \end{aligned} \tag{6}$$

then the marginal distribution of T is the $SMSN-BS(\alpha,\beta,\gamma,\delta)$.

The joint PDF of T, ψ and τ is given by

$$\begin{aligned} f(t,\psi,\tau) &= f(t|\psi,\tau)f(\psi|\tau)f(\tau) \\ &= A(t;\alpha,\beta)\phi\left(\sqrt{1+\tau^2}\left[a(t;\alpha,\beta)-\frac{\psi\tau}{1+\tau^2}\right]\right)\phi\left(\frac{\psi}{\sqrt{1+\tau^2}}\right)\frac{1}{\sqrt{\delta}}\phi\left(\frac{\tau-\gamma}{\sqrt{\delta}}\right) \\ &= \frac{1}{\pi\sqrt{\delta}}A(t;\alpha,\beta)\phi\left(\frac{\tau-\gamma}{\sqrt{\delta}}\right)\exp\left\{-\frac{1}{2}\left[a^2(t;\alpha,\beta)+(\psi-a(t;\alpha,\beta)\tau)^2\right]\right\}, \end{aligned} \tag{7}$$

where $a(t;\alpha,\beta)$ and $A(t;\alpha,\beta)$ are defined in (2) and (3), respectively. Integrating out ψ in (7), we get

$$f(t,\tau) = 2A(t;\alpha,\beta)\phi\left(a(t;\alpha,\beta)\right)\Phi\left(\tau a(t;\alpha,\beta)\right)\frac{1}{\sqrt{\delta}}\phi\left(\frac{\tau-\gamma}{\sqrt{\delta}}\right). \tag{8}$$

Finally, integrating out τ in (8), the marginal PDF of the SMSN-BS distribution is derived as

$$f_{SMSN-BS}(t) = 2\phi\left(a(t;\alpha,\beta)\right)\Phi\left(\frac{\gamma a(t;\alpha,\beta)}{\sqrt{1+\delta\left(a^2(t;\alpha,\beta)\right)}}\right)A(t;\alpha,\beta), \quad t>0. \tag{9}$$

It should be mentioned that the $SMSN-BS(\alpha,\beta,\gamma,\delta)$ distribution includes the BS when $\gamma = 0$ or $\delta \to \infty$. Also, if $\delta = 0$, it reduces to an extension of the BS distribution constructed using the SN model, which is called SN-BS disribution (see Vilca et al. (2011) for details). Figure 1 displays the PDF of SMSN-BS model for several choices of the parameters. Assuming $\alpha > 1$, it can be seen that the SMSN-BS distribution has fatter tail to the right, as γ grows. Also, for $\alpha > 1$, the PDF of SMSN-BS tends to a bimodal PDF when δ increases. So this distribution is a good model for investigating the positive mixture data.

The moments of the SMSN-BS distribution can be expressed in terms of the moments of the BS distribution. In the following lemma, relationships between the mean, variance, skewness and kurtosis of the two distributions are presented.

Lemma 1 *Let $T_{SMBS} \sim SMSN-BS(\alpha,\beta,\gamma,\delta)$ and $T_{BS} \sim BS(\alpha,\beta)$. Also, assume that $\omega_k = E\left(V^k\sqrt{(\alpha V)^2+4}\right)$, with $V \sim SMSN(\gamma,\delta)$. Then the mean, variance,*

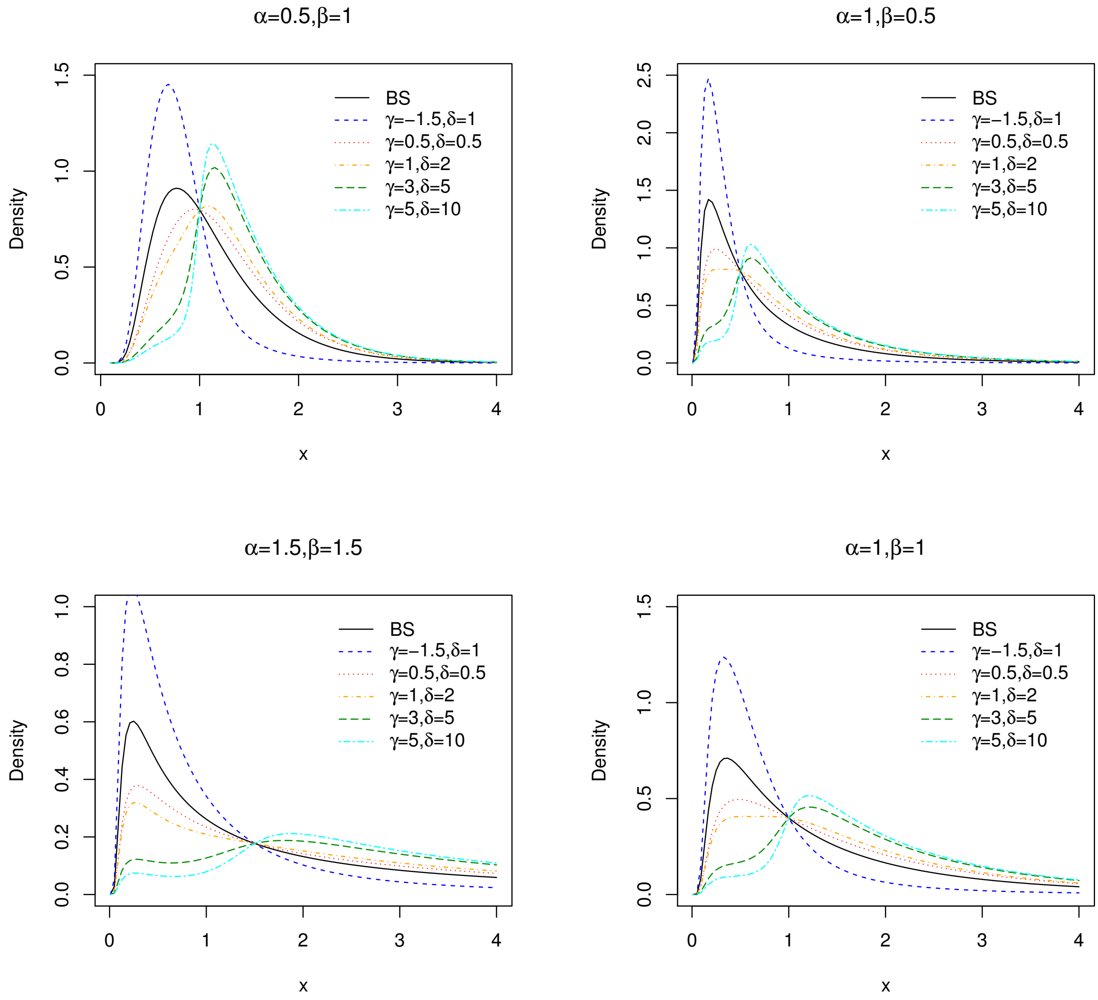

Figure 1: The PDF of the SMSN-BS distribution for several choices of the parameters

skewness (γ_3) and kurtosis (γ_4) of these two random variables satisfy the following relationships:

$$E(T_{SMBS}) = E(T_{BS}) + \frac{\alpha\beta}{2}\omega_1,$$

$$Var(T_{SMBS}) = Var(T_{BS}) + \left(\frac{\alpha\beta}{2}\right)^2 \alpha_\omega,$$

$$\gamma_3(T_{SMBS}) = \gamma_3(T_{BS}) \left(\frac{4+5\alpha^2}{4+5\alpha^2+\alpha_\omega}\right)^{3/2} + 2\frac{a_0+a_1\alpha+a_2\alpha^2}{(4+5\alpha^2+\alpha_\omega)^{3/2}},$$

and

$$\gamma_4(T_{SMBS}) = \left(\gamma_4(T_{BS}) + \frac{b_0+b_1\alpha+b_2\alpha^2+b_3\alpha^3}{(4+5\alpha^2)^2}\right)\frac{(4+5\alpha^2)^2}{(4+5\alpha^2+\alpha_\omega)^2},$$

where

$$a_0 = -6\omega_1 + \omega_1^3 + 2\omega_3, \quad a_1 = 3\omega_1^2 - 3\omega_1\omega_3,$$

$$a_2 = -6\omega_1 - 3\omega_3 + 2\omega_5, \quad \alpha_\omega = 2\alpha(\omega_3 - \omega_1) - \omega_1^2,$$

$$b_0 = 24\omega_1^2 - 3\omega_1^4 - 16\omega_1\omega_3, \quad b_1 = -96\omega_1 - 12\omega_1^3 - 16\omega_3 + 12\omega_1^2\omega_3 + 16\omega_5,$$

and

$$b_2 = 18\omega_1^2 + 24\omega_1\omega_3 - 16\omega_1\omega_5, \quad b_3 = 8\omega_7 - 168\omega_1 - 16\omega_5.$$

Proof: A stochastic representation of T can be obtained by combining (4) and (5). Then the moments of SMSN-BS random variable can be calculated with some algebra. It should be noted that ω_k's are calculated numerically. □

Table 1 presents values of some central moments (up to order five) for the SMSN-BS distribution, along with its skewness and kurtosis for different configurations of the parameters. It is observed that when $\gamma < 0$, the skewness and kurtosis of the SMSN-BS may be bigger than these quantities in the BS and SN-BS models. The next lemma plays a key role in the ML estimation for the SMSN-BS distribution.

Table 1: Values of central moments of the SMSN-BS distribution, along with its skewness and kurtosis for different configurations of the parameters

	$\alpha = 0.5$, $\beta = 1$				$\alpha = 1$, $\beta = 0.5$			
Moments	$\gamma = 0$, $\delta = 5$	$\gamma = -1$, $\delta = 0$	$\gamma = -1$, $\delta = 5$	$\gamma = 5$, $\delta = 5$	$\gamma = 0$, $\delta = 5$	$\gamma = -1$, $\delta = 0$	$\gamma = -1$, $\delta = 5$	$\gamma = 5$, $\delta = 5$
μ_1	1.12	0.82	0.99	1.52	0.74	0.39	0.59	1.20
μ_2	1.59	0.78	1.25	2.60	1.12	0.26	0.78	2.11
μ_3	2.80	0.88	2.03	5.08	2.75	0.28	1.84	5.37
μ_4	6.01	1.13	4.16	11.46	9.46	0.43	6.27	18.64
μ_5	15.42	1.66	10.38	30.04	42.14	0.88	27.80	83.10
Variance	0.32	0.11	0.26	0.28	0.56	0.11	0.42	0.66
Skewness	1.45	1.15	1.78	1.62	2.51	2.46	3.04	2.23
Kurtosis	6.44	5.18	8.02	7.37	12.85	12.89	17.12	10.94

Lemma 2 *If the hierarchical representation in (6) holds, then we have*

$$E\left(\tau|t\right) = \gamma + \frac{\delta u(t)}{\sqrt{1+\delta u^2(t)}} R\left(\frac{\gamma u(t)}{\sqrt{1+\delta u^2(t)}}\right),$$

$$E\left(\tau^2|t\right) = \gamma^2 + \delta + \frac{\gamma\delta u(t)}{\sqrt{1+\delta u^2(t)}}\left(1 + \frac{1}{1+\delta u^2(t)}\right) R\left(\frac{\gamma u(t)}{\sqrt{1+\delta u^2(t)}}\right),$$

and

$$E\left(\psi\tau|t\right) = u(t)\left(\gamma^2+\delta\right) + \gamma\sqrt{1+\delta u^2(t)} R\left(\frac{\gamma u(t)}{\sqrt{1+\delta u^2(t)}}\right),$$

where $u(t) = a(t;\alpha,\beta)$ *and* $R(x) = \phi(x)/\Phi(x)$.

Proof: Dividing (8) by (9) gives

$$f(\tau|t) = \frac{\Phi\left(\tau a(t;\alpha,\beta)\right)}{\Phi\left(\frac{\gamma a(t;\alpha,\beta)}{\sqrt{1+\delta a^2(t;\alpha,\beta)}}\right)} \frac{1}{\sqrt{\delta}}\phi\left(\frac{\tau-\gamma}{\sqrt{\delta}}\right). \tag{10}$$

So $\tau|t$ follows the extended SN (ESN) distribution (Azzalini and Capitanio, 2014), denoted by $ESN(\gamma, \delta, \sqrt{\delta}a(t,\alpha,\beta), \gamma a(t,\alpha,\beta))$. The first two expectations in the lemma are obtained from the properties of ESN model, see Azzalini and Capitanio (2014).

Dividing (7) by (8) yields

$$f(\psi|t,\tau) = \frac{\phi\left(\psi - \tau a(t;\alpha,\beta)\right)}{\Phi\left(\tau a(t;\alpha,\beta)\right)}. \tag{11}$$

It can be seen that $(\psi|t,\tau) \sim TN\left(\tau a(t;\alpha,\beta), 1; (0,\infty)\right)$. The last expectation in the lemma is concluded by applying (11). □

3 Parameter estimation

The ML estimation is one of the most widely used methods for estimating the parameters of a model. It selects the set of values of the model parameters that maximizes the likelihood function. In doing so, the "agreement" of the selected model with the observed data is maximized. Generally, the ML estimation for the SMSN-BS distribution is not an easy job due to complexity of the associated likelihood function. To sidestep this problem, an EM-type algorithm is proposed in the following. A method of obtaining the asymptotic standard errors of the ML estimates is also provided.

3.1 The ECM algorithm

The EM type algorithm, introduced by Dempster (1977), is a widely applicable method for iterative computation of the ML estimates. The power of the EM procedure lies in its ability in preserving implementation simplicity and monotonic convergence. However, it is not directly practicable to estimating the SMSN-BS model because the the M-step involves intractable computations. To solve this critical limitation, expectation-conditional maximization (ECM) algorithm, proposed by Meng and Rubin (1993), can be utilized. It substitutes the M-step of the EM algorithm by simpler conditional maximization steps.

Let $T_1, \ldots, T_n$ be a random sample from $SMSN - BS(\alpha, \beta, \gamma, \delta)$. From the representation (6), it follows that

$$\tau_i \overset{iid}{\sim} N(\gamma, \delta),$$

$$\psi_i|\tau_i \overset{iid}{\sim} TN\left(0, 1+\tau_i^2; (0,\infty)\right),$$

and

$$(T_i|\psi_i,\tau_i) \overset{iid}{\sim} EBS\left(\frac{\alpha}{\sqrt{1+\tau_i^2}}, \beta, 2, -\frac{\tau_i\psi_i}{\sqrt{1+\tau_i^2}}, 0\right).$$

Assume that $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \delta)$, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)$ and $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_n)$. Also, the vector of observed T_i's is denoted by $\mathbf{t} = (t_1, \ldots, t_n)$. Then, complete log-likelihood function of the data is given by

$$
\begin{aligned}
\ell_c(\boldsymbol{\theta}|\boldsymbol{t}, \boldsymbol{\psi}, \boldsymbol{\tau}) &= \sum_{i=1}^{n} \log f(t_i, \psi_i, \tau_i) \\
&= -\frac{n}{2}(3\log\pi + \log 2 + \log\delta) + \sum_{i=1}^{n} \log A(t_i; \alpha, \beta) \\
&- \frac{1}{2}\sum_{i=1}^{n} \left\{ a^2(t_i; \alpha, \beta) + (\psi_i - a(t_i; \alpha, \beta)\tau_i)^2 + \frac{(\tau_i - \gamma)^2}{\delta} \right\}
\end{aligned}
$$

Given the estimate of $\boldsymbol{\theta}$ at the k-th iteration, say $\hat{\boldsymbol{\theta}}^{(k)} = (\hat{\alpha}^{(k)}, \hat{\beta}^{(k)}, \hat{\gamma}^{(k)}, \hat{\delta}^{(k)})$, the expected value of the log-likelihood function becomes

$$
\begin{aligned}
Q\left(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}\right) &= E\left\{ \ell_c(\boldsymbol{\theta}|\boldsymbol{t}, \boldsymbol{\psi}, \boldsymbol{\tau}) \Big| \boldsymbol{t}, \hat{\boldsymbol{\theta}}^{(k)} \right\} \\
&\propto \frac{\gamma}{\delta}\sum_{i=1}^{n} \hat{s}_{1i}^{(k)} - \frac{1}{2}\sum_{i=1}^{n}(a^2(t_i; \alpha, \beta) + \frac{1}{\delta})\hat{s}_{2i}^{(k)} + \sum_{i=1}^{n} a(t_i; \alpha, \beta)\hat{s}_{3i}^{(k)} \\
&- \frac{n\gamma^2}{2\delta} - \frac{n}{2}\log\delta - \frac{1}{2}\sum_{i=1}^{n} a^2(t_i; \alpha, \beta) + \sum_{i=1}^{n} \log A(t_i; \alpha, \beta), \qquad (12)
\end{aligned}
$$

where

$$
\hat{s}_{1i}^{(k)} = E(\tau_i|t_i, \hat{\boldsymbol{\theta}}^{(k)}),
$$

$$
\hat{s}_{2i}^{(k)} = E(\tau_i^2|t_i, \hat{\boldsymbol{\theta}}^{(k)}),
$$

and

$$
\hat{s}_{3i}^{(k)} = E(\psi_i\tau_i|t_i, \hat{\boldsymbol{\theta}}^{(k)}). \qquad (13)
$$

The above quantities are easily obtained using lemma 2.

We thus propose the following ECM algorithm to perform the ML estimation for the SMSN-BS distribution:

- **E-step:** Given $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(k)}$, compute $\hat{s}_{1i}^{(k)}, \hat{s}_{2i}^{(k)}$ and $\hat{s}_{3i}^{(k)}$ in (13), for $i = 1, \ldots, n$.
- **CM-step 1:** Maximize (12) with respect to γ, δ and α, and obtain the following estimates

$$
\hat{\gamma}^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n} \hat{s}_{1i}^{(k)},
$$

$$
\hat{\delta}^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n} \hat{s}_{2i}^{(k)} - \hat{\gamma}^{2(k+1)},
$$

and

$$\hat{\alpha}^{(k+1)} = \frac{\sqrt{b_k^2 + 4c_k} - b_k}{2},$$

where

$$b_k = \frac{1}{n} \sum_{i=1}^{n} \left(\sqrt{\frac{t_i}{\hat{\beta}^{(k)}}} - \sqrt{\frac{\hat{\beta}^{(k)}}{t_i}} \right) \hat{s}_{3i}^{(k)},$$

and

$$c_k = \frac{1}{n} \sum_{i=1}^{n} \left(\sqrt{\frac{t_i}{\hat{\beta}^{(k)}}} - \sqrt{\frac{\hat{\beta}^{(k)}}{t_i}} \right)^2 (1 + \hat{s}_{2i}^{(k)}).$$

- **CM-step 2:** Use $\hat{\alpha}^{(k+1)}$, $\hat{\gamma}^{(k+1)}$ and $\hat{\delta}^{(k+1)}$ in the previous step, and obtain $\hat{\beta}^{(k+1)}$ from

$$\hat{\beta}^{(k+1)} = \arg\max_{\beta} Q(\hat{\alpha}^{(k+1)}, \beta, \hat{\gamma}^{(k+1)}, \hat{\delta}^{(k+1)} | \hat{\boldsymbol{\theta}}^{(k+1)}).$$

The above procedure is repeated until a suitable stopping criterion is met. To facilitate the process of determining the actual convergence, we adopt an acceleration method due to Aitken (1926). Given the sequence of observed log-likelihood $\{\ell^{(k)}\}_{k=0}^{\infty}$, the Aitken's acceleration criterion is calculated as $a^{(k)} = (\ell^{(k+1)} - \ell^{(k)})/(\ell^{(k)} - \ell^{(k-1)})$. This yields the asymptotic estimate of the log-likelihood $\ell_{\infty}^{(k+1)} = \ell^{(k)} + (\ell^{(k+1)} - \ell^{(k)})/(1 - a^{(k)})$, which can be computed in advance at $(k+1)$-th iteration. Then, the algorithm is considered to be convergent if $\ell_{\infty}^{(k+1)} - \ell^{(k)} < \epsilon$, where $\epsilon = 10^{-6}$ is the default tolerance employed in our experimental study.

It is well-known that the EM-type algorithm is likely to get trapped in one of the many local maxima of the likelihood function. To overcome this problem, we generate a variety of reasonable initial values, and the set of parameters associated with the highest converged log-likelihood is selected. We use modified moment estimators of α and β as initial values. This quantities are obtained from Vilca et al. (2011). Thus we have $\hat{\alpha}_{(0)} = \sqrt{2(s/r)^{1/2} - 1}$ and $\hat{\beta}_{(0)} = \sqrt{sr}$, where s and r are arithmetic and harmonic means of samples. Each of these values are multiplied by a fixed number that is randomly drawn from the uniform distribution on the interval (0.5,2). Finally, initial values of the shape parameters γ and δ are randomly chosen from the uniform distribution on the interval (1,10).

3.2 The observed information matrix

The Fisher information is a measure of information contained in the sample about parameters of the parent distribution. It also has connections to the ML estimation method. Under regularity conditions (see Cramer, 1946), the asymptotic covariance matrix of the ML estimator $\hat{\boldsymbol{\theta}}$ can be approximated by the inverse of this matrix. Specifically, the Fisher information matrix is defined as

$$I(\boldsymbol{\theta}) = -E\left\{ \sum_{i=1}^{n} H_i(\boldsymbol{\theta}) \right\}.$$

where $H_i(\boldsymbol{\theta}) = \left[\frac{\partial^2}{\partial\theta_k\partial\theta_j}\ell(\theta \mid t_i)\right]$ denotes the Hessian matrix associated with t_i.

Unfortunately, the above expectation cannot be obtained analytically for the SMSN-BS model. As remarked by Efron and Hinkley (1978), the Fisher information matrix can be consistently estimated by $-\sum_{i=1}^n H_i(\hat{\boldsymbol{\theta}})$, which is called the observed information matrix.

From (9), the log-likelihood function of $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \delta)$ given the observed data $\mathbf{t} = (t_1, ..., t_n)$ is $\ell(\boldsymbol{\theta}|\boldsymbol{t}) = \sum_{i=1}^n \ell(\boldsymbol{\theta}|t_i)$, where

$$\ell(\boldsymbol{\theta}|t_i) \propto \log A(t_i;\alpha,\beta) + \log\phi(a(t_i;\alpha,\beta)) + \log\Phi(x_i),$$

with $x_i = \frac{\gamma a(t_i;\alpha,\beta)}{\sqrt{1+\delta a^2(t_i;\alpha,\beta)}}$.

After some algebra, the second derivatives of $\ell(\boldsymbol{\theta}|t_i)$ are derived as

$$\begin{aligned}
\frac{\partial^2\ell(\boldsymbol{\theta}|t_i)}{\partial\theta_1\partial\theta_2} = & - \frac{1}{A^2(t_i;\alpha,\beta)}\frac{\partial A(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial A(t_i;\alpha,\beta)}{\partial\theta_2} + \frac{1}{A(t_i;\alpha,\beta)}\frac{\partial^2 A(t_i;\alpha,\beta)}{\partial\theta_1\partial\theta_2} \\
& - a(t_i;\alpha,\beta)\frac{\partial^2 a(t_i;\alpha,\beta)}{\partial\theta_1\partial\theta_2} - \frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_2} \\
& + w(\theta_1,\theta_2)R(x_i) + w(\theta_1)w(\theta_2)R'(x_i), \qquad \theta_1,\theta_2 \in \{\alpha,\beta\}.
\end{aligned}$$

with

$$\begin{aligned}
w(\theta_j) &= \frac{\gamma}{(1+\delta a^2(t_i;\alpha,\beta))^{3/2}}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_j}, \quad j = 1,2, \\
w(\theta_1,\theta_2) &= \frac{\gamma}{(1+\delta a^2(t_i;\alpha,\beta))^{3/2}} \\
& \times \left[\frac{\partial^2 a(t_i;\alpha,\beta)}{\partial\theta_1\partial\theta_2} - 3\frac{\delta a(t_i;\alpha,\beta)}{(1+\delta a^2(t_i;\alpha,\beta))}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_2}\right],
\end{aligned}$$

and

$$R'(v) = -R(v)(v + R(v)).$$

Furthermore, we have

$$\begin{aligned}
\frac{\partial^2\ell(\boldsymbol{\theta}|t_i)}{\partial\delta\partial\theta_j} &= -\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_j}\frac{\gamma a^2(t_i;\alpha,\beta)}{2\left(1+\delta a^2(t_i;\alpha,\beta)\right)^{5/2}}\left[3R(x_i) + x_iR'(x_i)\right], \\
\frac{\partial^2\ell(\boldsymbol{\theta}|t_i)}{\partial\gamma\partial\theta_j} &= \frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_j}\frac{1}{\left(1+\delta a^2(t_i;\alpha,\beta)\right)^{3/2}}\left[R(x_i) + x_iR'(x_i)\right], \\
\frac{\partial^2\ell(\boldsymbol{\theta}|t_i)}{\partial\gamma\partial\delta} &= -\frac{a^3(t_i;\alpha,\beta)}{2(1+\delta a^2(t_i;\alpha,\beta))^{3/2}}\left[R(x_i) + x_iR'(x_i)\right], \\
\frac{\partial^2\ell(\boldsymbol{\theta}|t_i)}{\partial\delta^2} &= \frac{\gamma a^5(t_i;\alpha,\beta)}{4\left(1+\delta a^2(t_i;\alpha,\beta)\right)^{5/2}}\left[3R(x_i) + x_iR'(x_i)\right],
\end{aligned}$$

and

$$\frac{\partial^2\ell(\boldsymbol{\theta}|t_i)}{\partial\gamma^2} = \frac{a^2(t_i;\alpha,\beta)}{1+\delta a^2(t_i;\alpha,\beta)}R'(x_i).$$

The standard errors of estimates can be obtained by the square root of the diagonal elements of the inverse of $-\sum_{i=1}^{n} H_i(\hat{\boldsymbol{\theta}})$. The calculation standard errors is particularly useful for inferential purposes, e.g. construction of confidence intervals, or hypothesis testing.

4 Numerical results

In this section, two simulation studies are conducted to assess finite-sample performance of our proposed model and the ECM algorithm. The objective of the first simulation is to investigate properties of the ML estimates. The second simulation aims to compare the SMSN-BS distribution with the BS, EBS and SN-BS models in terms of robustness and flexibility. The EBS distribution, introduced by Vilca et al. (2010), is more flexible than the classical BS and some its generalizations. On the other hand, as mentioned before, the SN-BS distribution is a special case of SMSN-BS model when $\gamma = \lambda$ and $\delta = 0$. For properties and further details about SN-BS model, see Vilca et al. (2011).

We first study properties of the ML estimates obtained from the ECM algorithm described in the previous section. To this end, 500 samples of sizes $n = 250, 500, 1000, 2000$ are generated from the SMSN-BS distribution with parameters $\alpha = 0.5$, $\beta = 1$, $\gamma = 1.5$ and $\delta = 2$. The parameters are estimated from each sample via the ECM algorithm. The resulting values are then used to determine absolute bias (AB) and mean squared error (MSE) of any estimator as

$$AB = \frac{1}{500}\sum_{i=1}^{500}\left|\widehat{\theta}_i - \theta_{true}\right| \quad \text{and} \quad MSE = \frac{1}{500}\sum_{i=1}^{500}\left(\widehat{\theta}_i - \theta_{true}\right)^2.$$

where $\widehat{\theta}_i$ denotes estimate of a specific parameter at the ith replication.

Furthermore, we compute the standard deviations (SDs) of the ML estimates across 500 simulated samples and compare them with the average values of the approximate standard errors (ASEs) obtained through Hessian matrix. The SD is computed as

$$SD = \left(\frac{1}{499}\left[\sum_{i=1}^{500}\widehat{\theta}_i^2 - \frac{1}{500}\left(\sum_{i=1}^{500}\widehat{\theta}_i\right)^2\right]\right)^{1/2}.$$

Numerical results in Table 2 support consistency of the ML estimates because the AB and MSE values shrink toward zero as n increases. We note that the SD values are in good agreement with the corresponding ASE values for all sample sizes.

Moreover, it should be noted that the likelihood function for the SMSN-BS distribution tends to be relatively flat near the ML estimates of γ and δ. Therefore, as evidenced in Table 2, the parameters γ and δ have larger MSE values as well as estimated variances than the corresponding values for the other parameters.

In the second simulation design, we evaluate robustness of the proposed model and some of its competitors. To do so, 300 samples of sizes $n = 500, 1000$ are generated from the BS distribution with parameters $\alpha = 0.5$ and $\beta = 1$. To create heavy tails, the

Table 2: Simulation results for assessing the consistency of parameter estimates and accuracy of approximate standard errors for various sample sizes.

n	Measure	α	β	γ	δ
250	AB	0.0037	0.0025	0.2986	2.0537
	MSE	0.0015	0.0039	0.6901	38.040
	SD	0.0380	0.0685	0.7650	4.7655
	ASE	0.0357	0.0600	0.7713	4.2012
500	AB	0.0004	0.0010	0.0767	0.4806
	MSE	0.0009	0.0022	0.1881	3.6532
	SD	0.0263	0.0477	0.4417	2.5023
	ASE	0.0254	0.0429	0.5105	2.4036
1000	AB	0.0011	0.0032	0.0338	0.3098
	MSE	0.0003	0.0011	0.0919	1.5251
	SD	0.0187	0.0327	0.2878	1.1428
	ASE	0.0180	0.0302	0.3395	1.3999
2000	AB	0.00008	0.0001	0.0347	0.1659
	MSE	0.0001	0.0005	0.0341	0.4378
	SD	0.0141	0.0236	0.2018	0.6940
	ASE	0.0127	0.0209	0.2334	0.8895

size of each sample is increased by 2% through adding data generated from a uniform distribution on the interval (10,20). Finally, the EBS, SN-BS and SMSN-BS distributions are fitted to each simulated data set. For any model and sample size, averages and SDs of the log-likelihood maxima, Akaike information criterion (AIC) and Bayesian information criterion (BIC) are computed based on 300 replications. The AIC and BIC criteria are defined as

$$\text{AIC} = 2m - 2\ell_{max} \quad \text{and} \quad \text{BIC} = m \log n - 2\ell_{max}$$

where n is the sample size, m is the number of parameters, and $\ell_{\max}$ is the maximized log-likelihood (see Akaike, 1973 and Schwarz, 1978).

Table 3 reports the above mentioned quantities. As a general rule, smaller values of AIC or BIC indicate a better-fitting model. The number of times (out of 300 replications) that each model is selected by a given criterion comes with Freq label. It can be observed that all criteria tend to select the SMSN-BS model. Moreover, it is clear that the performance of SMSN-BS significantly improves with larger sample size.

5 Illustration

Recently, the air quality of many big cities has seriously deteriorated. Research studies show that some of air contaminant concentrations can be very harmful for human health. For example, tropospheric ozone (O_3) remains in the atmosphere for a long time and irritates the respiratory system. The O_3 can increase the severity of chronic respiratory

Table 3: Comparison of averages and SDs of the log-likelihood maxima, AIC and BIC for different models, and frequencies supported by Criteria

		$n = 500$			$n = 1000$		
Criterion		EBS	SN-BS	SMSN-BS	EBS	SN-BS	SMSN-BS
$\ell_{\max}$	Mean	-513.52	-511.07	-468.50	-1030.88	-1025.97	-932.33
	SD	15.97	16.07	27.29	21.11	21.29	40.12
	Freq	0	31	269	0	15	285
AIC	Mean	1033.04	1028.14	945.00	2067.76	2057.95	1872.66
	SD	31.94	32.15	54.58	42.23	42.59	80.25
	Freq	0	45	255	0	24	276
BIC	Mean	1045.74	1040.84	961.93	2082.54	2072.74	1892.37
	SD	31.94	32.15	54.58	42.23	42.59	80.25
	Freq	0	47	253	0	29	271

Table 4: Descriptive statistics for the ozone data

Sample size	Mean	Standard deviation	Skewness	Kurtosis	D'Agostino test (p-value)	Anscombe Glynn test (p–value)
116	42.12	32.98	1.22	4.18	4.65 ($< 1e-4$)	2.20 (0.027)

diseases including bronchitis, and emphysema (WHO, 2006). Therefore modelling and analysis of air pollution quantities can be helpful to decrease their effects.

Usually, air contaminant concentrations are considered as continuous positive random variables. These random variables often show asymmetric PDF's and present positive skewness and high kurtosis. So symmetric models such as normal distribution are not good choices for describing the environmental random variables. The BS distribution and its generalizations have been largely applied to environmental data. For example, see Leiva et al. (2008), Leiva et al. (2010), Balakrishnan et al. (2009), Vilca et al. (2010), Vilca et al. (2011), and Ferreira et al. (2012).

To illustrate applicability of the SMSN-BS model and computational methods proposed in this paper, we study the daily ozone concentrations in New York during May-September 1973, used by Nadarajah (2008), provided by the New York State Department of Conservation. Vilca et al. (2011) fitted the SN-BS distribution to these data and compare it with the BS model. Table 4 contains descriptive statistics of these data. Results of D'Agostino test (D'Agostino, 1970) for skewness, and Anscombe-Glynn test (Anscombe and Glynn, 1983) for kurtosis suggest that the observed data could be more adequately modelled by some specific skew distributions rather than the normal distribution.

We implemented the ECM algorithm described in Section 3 for fitting the SMSN-BS distribution to the data set. For the sake of comparison, the BS, EBS, SN-BS, BS-logistic (BS-LOG) and BS-t student (BS-T) distributions were are also fitted. For details of the last two models see Leiva et al. (2008). The models are compared based on the $\ell(\hat{\theta})$, AIC

Table 5: The ML estimates and information criteria for the ozone data. Here, ϵ refers to additional shape parameters in the EBS, BS-T and SN-BS distributions.

Model	α	β	ϵ	γ	δ	$\ell(\hat{\theta})$	AIC	SABIC
BS	0.9823	28.0234	–	–	–	–549.09	1102.19	1101.38
	(0.064)	(2.261)						
BS-LOG	0.5292	30.4946	–	–	–	–544.35	1092.70	1091.89
	(0.040)	(2.459)						
BS-T	0.8101	30.878	7.241	–	–	–543.39	1092.78	1091.56
	(0.073)	(2.48)	(3.38)					
EBS	1.0308	75.402	0.669	–	–	–545.11	1096.23	1091.89
	(0.073)	(11.83)	(0.098)					
SN-BS	1.2702	14.8351	1.0667	–	–	–545.60	1097.21	1095.98
	(0.236)	(4.028)	(0.534)					
SMSN-BS	1.5224	11.2994	–	3.4067	2.4188	–540.84	1089.46	1087.83
	(0.117)	(0.972)		(1.148)	(2.222)			

and sample size adjusted BIC (SABIC) criteria. When the sample size is not big enough or distinguishing between two models is difficult, the BIC criterion is not suggested for model selection. Dziak (2012) showed that in these cases, the SABIC is better than BIC criterion. The SABIC is defined as

$$\text{SABIC} = m \log\left(\frac{n+2}{24}\right) - 2\ell_{max},$$

where n is the sample size, m is the number of parameters, and $\ell_{\max}$ is the maximized log-likelihood, (Sclove, 1987). The above quantities along with the parameters' estimates are given in Table 5. The standard errors (SEs) of the ML estimates appear in parentheses.

The Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests have been widely used for goodness-of-fit tests. See Smirnov (1948) and Anderson and Darling (1954). The latter test assigns more weights to the tails of a distribution, and thus it is preferable for application in the case of the BS distribution and its generalizations. Values of these statistics along with the corresponding p-values are reported in Table 6. Based on the AIC and SABIC criteria, and the AD test, the SMSN-BS model provides the best fit for the ozone data. Figure 2 shows the histogram for data, and density of the fitted models. One can see that the SMSN-BS distribution has a good fit to data. Also in this figure, profile log-likelihood of the shape parameters for the SMSN-BS model is plotted for ozone data, where red triangle indicates the location of maximum of the log-likelihood function. As it can be seen in this figure, the ML estimates for γ and δ are numerically stable for these data, since there is a sharp peak in the perspective surface surrounding the stationary point.

Moreover, for testing the null hypothesis $H_0 : \gamma = 0$ against the alternative hypothesis $H_1 : \gamma \neq 0$, we calculate the likelihood ratio test (LRT) statistic as $\Lambda = -2(\ell_0 - \ell_1)$, where ℓ_0 (ℓ_1) is the maximum value of the log-likelihood function under H_0 (H_1). For the

Table 6: Goodness of fit tests and corresponding p-values for the ozone data.

Test	BS	BS-LOG	BS-T	EBS	SN-BS	SMSN-BS
K-S test	0.08	0.065	0.063	0.068	0.062	0.064
	(0.40)	(0.70)	(0.73)	(0.63)	(0.715)	(0.70)
A-D test	1.34	0.51	0.49	0.68	0.819	0.47
	(0.22)	(0.73)	(0.77)	(0.57)	(0.443)	(0.78)

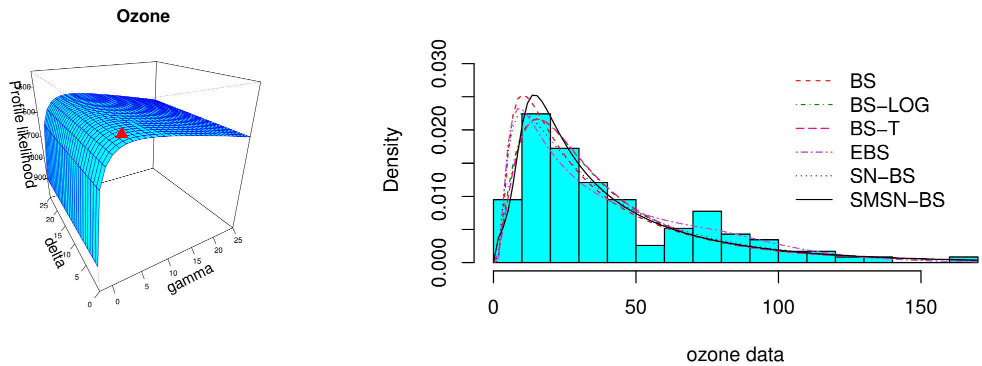

Figure 2: Profile log-likelihood of the shape parameters for the SMSN-BS model. (Left) Histograms of the data. The lines represent fitted densities using different models (Right).

enough large sample size n, H_0 is rejected at the significance level of 0.05 if $\Lambda > \chi^2_{1,0.05}$, where $\chi^2_{1,0.05} = 3.84$ is 95th percentile of chi-square distribution with one degree of freedom. Under H_0, the model is reduced to BS distribution for which $\ell_0 = -549.09$. Under H_1, we have the SMSN-BS distribution for which $\ell_1 = -540.84$. Thus $\Lambda = 16.5$, and the null hypothesis is rejected at the significance level of 0.05.

Testing $H_0 : \delta = 0$ against $H_1 : \delta \neq 0$ can be done similarly. Under H_0, the model is reduced to SN-BS distribution for which $\ell_0 = -545.60$. In this case, $\Lambda = 9.52$ and the null hypothesis is again rejected at the significance level of 0.05. These results indicate that both shape parameters γ and δ are significant for these data.

6 Conclusion

This article deals with a new extension of the BS distribution based on a shape mixture of skew normal model. The main strength of this new distribution lies in the fact that it offers additional flexibility in modelling data with varying degrees of peakedness and tail heaviness. The likelihood function associated with the new model turns out to be complicated. An EM-type algorithm is therefore presented to facilitate the ML estimation method. Practical issues regarding specification of the starting values, the stopping rule, and the provision of estimates' SEs are also addressed. Properties of the ML estimators, and flexibility of the proposed distribution are investigated using numerical results. Application of the new model is illustrated using an environmental data set. The proposed distribution can be generalized to multivariate setup. This will be considered in a subsequent work.

Acknowledgments

We are grateful to the Editor and two anonymous reviewers for their valuable comments and suggestions which have led to improvement in this paper.

References

Aitken, A. C. (1926). On Bernoulli's numerical solution of algebraic equations. *Proc R Soc Edinburgh* 46, 289–305.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *In 2nd Int. Symp. on Information Theory (Edited by B. N. Petrov and F. Csaki)*, 267–281.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49, 765–769.

Anscombe, F. J. and Glynn, W. J. (1983). Distribution of kurtosis statistic for normal statistics. *Biometrika*, 70, 227–234.

Arellano-Valle, R. B., Gómez, H. W., and Quintana, F. A. (2004). A new class of skew-normal distributions. *Communication in Statistics: Theory and Methods*, 33, 1465–1480.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171–178.

Azzalini, A., and Capitanio, A. (2014). *The Skew-Normal and Related Families*, Cambridge University Press, Cambridge, UK.

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

Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability*, 6, 319–327.

Cramér, H. (1946) *Mathematical Methods of Statistics*, Princeton University Press, Princeton, New Jersey.

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

Cordeiro, G. M., Lima, M.D.C.S., Cysneiros, A.H., Pascoa, M. A., Pescim R. R., and

Ortega, E. M. (2016). An extended Birnbaum-Saunders distribution: Theory, estimation, and applications. *Communications in Statistics: Theory and Methods*, 45, 2268-2297.

D'Agostino, R. B. (1970). Transformation to Normality of the Null Distribution of g_1. *Biometrika*, 57, 679–681.

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

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

Dziak, J. J., Coffman, D. L., Lanza S. T. and Li, R. (2012). Sensitivity and specificity of information criteria. *The Methodology Center and Department of Statistics, Penn State, The Pennsylvania State University*, 1–10.

Efron, B. and Hinkley, D. V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information, *Biometrika*, 65, 457–482.

Ferreira, M., Gomes M. I., and Leiva, V. (2012). On an extreme value version of the Birnbaum-Saunders distribution, *Revstat Statistical Journal*, 10, 181–210.

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

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

Louis Floyd, F. (1998). Reducing product development cycle times without increasing risk. *Journal of Coatings Technology*, 70, 71–81.

Marchant, C., Leiva, V., Cavieres, M. F., and Sanhueza, A. (2013) Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *In Reviews of Environmental Contamination and Toxicology*, 223, 1–31.

Meng, X. L., and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika*, 80, 267–278.

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

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52, 333–343.

Smirnov, N. V. (1948). Tables for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19, 279–281.

Vilca, F., Sanhueza, A., Leiva, V., and Christakos, G. (2010). An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile, *Stochastic Environmental Research and Risk Assessment*, 24, 771–782.

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.

World Health Organization (WHO). (2006). Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide. *World Health Organization.*