symmetry
MDPI

Article

Flexible Birnbaum–Saunders Distribution

Guillermo Martínez-Flórez [1], Inmaculada Barranco-Chamorro [2], Heleno Bolfarine [3] and Héctor W. Gómez [4,*]

1 Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Córdoba 230027, Colombia; guillermomartinez@correo.unicordoba.edu.co
2 Departamento de Estadística e Investigación Operativa, Universidad de Sevilla, 41000 Sevilla, Spain; chamorro@us.es
3 Departamento de Estatítica, IME, Universidade de São Paulo, São Paulo 01000, Brazil; hbolfar@ime.usp.br
4 Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile
* Correspondence: hector.gomez@uantof.cl

Received: 5 September 2019; Accepted: 4 October 2019; Published: 16 October 2019

Abstract: In this paper, we propose a bimodal extension of the Birnbaum–Saunders model by including an extra parameter. This new model is termed flexible Birnbaum–Saunders (FBS) and includes the ordinary Birnbaum–Saunders (BS) and the skew Birnbaum–Saunders (SBS) model as special cases. Its properties are studied. Parameter estimation is considered via an iterative maximum likelihood approach. Two real applications, of interest in environmental sciences, are included, which reveal that our proposal can perform better than other competing models.

Keywords: flexible skew-normal distribution; skew Birnbaum–Saunders distribution; bimodality; maximum likelihood estimation; Fisher information matrix

---

1. Introduction

The BS distribution was originally introduced in [1] to model the fatigue in the lifetime of certain materials. During the last decades, mainly due to its good properties, the use of this model spread out to other fields, such as economics and environmental sciences. In these applied scenarios, quite often, departures of the BS model are found, and therefore it is necessary to introduce some improvements. In this paper, we focus on those situations in which extra asymmetry or bimodality are present in our data, and a generalization of the BS model should be considered to deal with these issues. To reach this end, a flexible BS model is introduced. Our proposal is based on the flexible skew-normal distribution introduced in [2], and includes, as particular cases, the BS and skew BS distribution. Next, we briefly describe the key aspects that properly combined result in the flexible BS model. These are asymmetry, bimodality and main features of the basic BS model.

1.1. Asymmetry

Earlier results on asymmetric models started with the pioneering works by [1,3,4]. This topic regained interest with the study in [5], which from a Bayesian point of view developed a new asymmetric model which was later studied in depth by Azzalini [6], from a classical point of view. Azzalini model was termed the skew-normal distribution. Following Azzalini's method, a general family of asymmetric models termed skew-symmetric models appeared in the literature. The following lemma, originally presented in [6], can be considered as the starting point for the development of these asymmetric models.

Lemma 1. *Let f_0 be a probability density function (pdf) which is symmetric around zero, and G a cumulative distribution function (cdf) such that G' exists and is a symmetric pdf around zero. Then*

$$f_Z(z;\lambda) = 2f_0(z)G(\lambda z), \qquad z \in \mathbb{R}, \tag{1}$$

is a pdf for $\lambda \in \mathbb{R}$.

Equation (1) provides the skew version of $f_0(\cdot)$ with skewing function $G(\cdot)$ and λ the skewness parameter. If $f_0(\cdot) = \phi(\cdot)$ and $G(\cdot) = \Phi(\cdot)$, the pdf and cdf, respectively, of the $N(0,1)$ distribution, then the skew-normal is obtained, whose pdf is

$$f_Z(z) = 2\phi(z)\Phi(\lambda z), \qquad z \in \mathbb{R},\ \lambda \in \mathbb{R}\,. \tag{2}$$

Other examples of skew models are: skew-t, skew-Cauchy, skew-elliptical, and generalized skew-elliptical. We highlight that all of them are unimodal distributions.

1.2. Bimodality

Another fundamental result in our proposal will be the following lemma, which was given in Gómez et al. [2]. These authors extended (1), by introducing a parameter δ in f_0, in such a way that for certain values of δ the resulting distribution is bimodal.

Lemma 2. *Let f be a symmetric pdf around zero, F the corresponding cdf and G an absolutely continuous cdf such that G' exists and is symmetric around zero. Then*

$$g(z;\delta,\lambda) = c_\delta f(|z| + \delta)G(\lambda z), \qquad z \in \mathbb{R}, \quad \lambda,\delta \in \mathbb{R} \tag{3}$$

is a pdf and $c_\delta^{-1} = 1 - F(\delta)$.

Taking $f(\cdot) = \phi(\cdot)$ and $G(\cdot) = \Phi(\cdot)$, in (3), the flexible skew-normal (FSN) model was obtained and studied in detail in [2]. There, it was proved that the FSN model can be bimodal for certain values of δ. Notice that the FSN model is obtained by adding an extra parameter, δ, to the skew-normal distribution proposed in [6]. That is a random variable (rv) Z follows a FSN distribution, $Z \sim FSN(\delta,\lambda)$, if its pdf is given by

$$f(z;\delta,\lambda) = c_\delta \phi(|z| + \delta)\Phi(\lambda z), \qquad z \in \mathbb{R}, \ \lambda,\delta \in \mathbb{R} \tag{4}$$

where ϕ and Φ are the pdf and cdf of the $N(0,1)$ distribution, respectively, and $c_\delta^{-1} = 1 - \Phi(\delta)$.

Other recent proposals in the contemporary literature dealing with bimodality are the extended two-pieces skew-normal model (ETN), introduced in [7] and the uni-bi-modal asymmetric power normal model given in [8] whose properties are based on results given in [9,10]. Applications of interest in Economics are given in [11]. All these references show the interest in the latest years for modelling bimodality.

1.3. BS Model

The BS or fatigue life distributions was proposed for modelling survival time data and material lifetime subject to stress in [12,13]. This model is asymmetric and only fits positive data. The pdf of a BS distribution is given by

$$f_T(t) = \phi(a_t)\frac{t^{-3/2}(t+\beta)}{2\alpha\sqrt{\beta}}, \qquad t > 0, \tag{5}$$

where

$$a_t = a_t(\alpha, \beta) = \frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right), \tag{6}$$

$\alpha > 0$ is a shape parameter, $\beta > 0$ is a scale parameter and the median of this distribution. (5) is denoted as $T \sim BS(\alpha, \beta)$. It is well known that α is the parameter that controls asymmetry. Specifically, (5) becomes more asymmetric as α increases and symmetric around β as α gets close to zero. It can be seen in [13] that (5) can be obtained as the distribution of the random variable

$$T = \beta\left[\frac{\alpha}{2}Z + \sqrt{\left(\frac{\alpha}{2}Z\right)^2 + 1}\right]^2, \tag{7}$$

where $Z \sim N(0,1)$.

The BS model has been applied to a variety of practical situations. However, quite often, although the data suggest a BS distribution, some deficiencies are observed in the fitted BS model. This problem has motivated an increasing interest in its generalizations. We highlight that, recently, this model was extended by [14] to the family of elliptical distributions, this is known in the literature as the generalized Birnbaum–Saunders (GBS) distribution. Later, [15] proposed an extension based on the elliptical asymmetric distributions, known as the doubly generalized Birnbaum–Saunders model. On the other hand, [16] presents the asymmetric BS distribution with five parameters called the extended Birnbaum–Saunders (EBS) distribution. Other types of extensions are the asymmetric epsilon-Birnbaum–Saunders model given in [17], models in [18] based on the slash-elliptical family of distributions, and the generalized modified slash Birnbaum–Saunders (GMSBS) proposed in [19], which is based on [20].

In these extensions, we find that the asymmetric BS models previously cited, such as [15,21], are designed to fit data with greater or smaller asymmetry (or kurtosis) than that of the ordinary BS model, but they are not appropriate for fitting bimodal data. On the other hand, we highlight that the extension given in [21], which can become bimodal for certain combination of parameters is unable to capture bimodality unless it is accentuated enough.

Therefore there exists a real need for an asymmetric model, based on the BS distribution, and able to fit data presenting bimodal features, which is not uncommon in the literature. So the present paper presents a flexible BS distribution able to model skewness and to fit data with and without bimodality.

The paper is organized as follows. Section 2 is devoted to the development of an asymmetric uni-bimodal BS model. Its properties are studied in depth. Specifically, a closed expression for the cumulative distribution function (cdf) is given in terms of the cdf of a bivariate normal distribution. Some of the models proposed in [15,22] are obtained as particular cases. The shape and bimodality of the distribution are studied. It is shown that this model is closed under a change of scale and reciprocity. Survival and hazard functions are also obtained. Section 3 deals with moments derivation and iterative maximum likelihood estimation methods for the new model. Section 4 is devoted to real data applications of interest in environmental sciences. The first one deals with a bimodal situation in which our proposal performs better than other BS models and a mixture of normal distributions. The second one is taken from [16], where the extended BS model was proposed as the best for this dataset. It is shown that the FBS outperforms the extended BS model.

2. Results in Flexible Birnbaum-Saunders

Based on the flexible skew-normal model proposed in [2], we extend the Birnbaum–Saunders. The main idea is to apply (7) with $Z \sim FSN(\delta, \lambda)$ introduced in (4). This new model is called the flexible Birnbaum–Saunders (FBS) distribution whose pdf is given by

$$f(t; \alpha, \beta, \delta, \lambda) = \frac{t^{-3/2}(t+\beta)}{2\alpha\beta^{1/2}(1-\Phi(\delta))}\phi(|a_t| + \delta)\Phi(\lambda a_t), \tag{8}$$

with a_t defined in (6), $t > 0$, $\alpha > 0$, $\beta > 0$, $\delta \in \mathbb{R}$, $\lambda \in \mathbb{R}$, $\phi(\cdot)$ and $\Phi(\cdot)$ the pdf and cdf of a $N(0,1)$, respectively. We use the notation $T \sim FBS(\alpha, \beta, \delta, \lambda)$. The inclusion of parameters δ and λ makes our approach more flexible than the extensions previously discussed. λ is a parameter that controls asymmetry (skewness) and δ is a shape parameter related to bimodality of our proposal.

If $\lambda = 0$ then we obtain, as a particular case, the model introduced by [22].

Figures 1 and 2 depict the behaviour of (8) for some values of parameters, illustrating that it can be bimodal for some combinations of them.

2.1. *Interpretation of Parameters.*

In both figures the values of parameters α and β are fixed. We study the effects of

(i) λ positive versus λ negative.
(ii) Increasing $\delta > 0$ in Figure 1. Decreasing $\delta < 0$ in Figure 2.

Figure 1 suggests that, for α and β fixed, if a positive value of δ is considered then we have a unimodal distribution and the peak of the distribution increases when δ increases: $\delta = 0.75$ (red solid line), $\delta = 1.5$ (green dashed line), $\ldots$, $\delta = 3$ (blue dashed dotted line). This happens for positive and negative values of λ.

On the other hand, in Figure 2, we have different situations. This plot suggests that, for α and β fixed, if a negative value of δ is considered then a bimodal distribution can be obtained. For positive λ, if δ decreases: $\delta = -0.75$ (red solid line), $\delta = -1.5$ (green dashed line), $\ldots$, $\delta = -3$ (blue dashed dotted line), then the peaks decrease and bimodality becomes more accentuated. For negative λ, if δ decreases, then main peak increases and bimodality becomes less accentuated.

Also, note in Figures 1 and 2, that in the FBS model the pdf for negative λ is no longer the specular image of plot for positive λ.

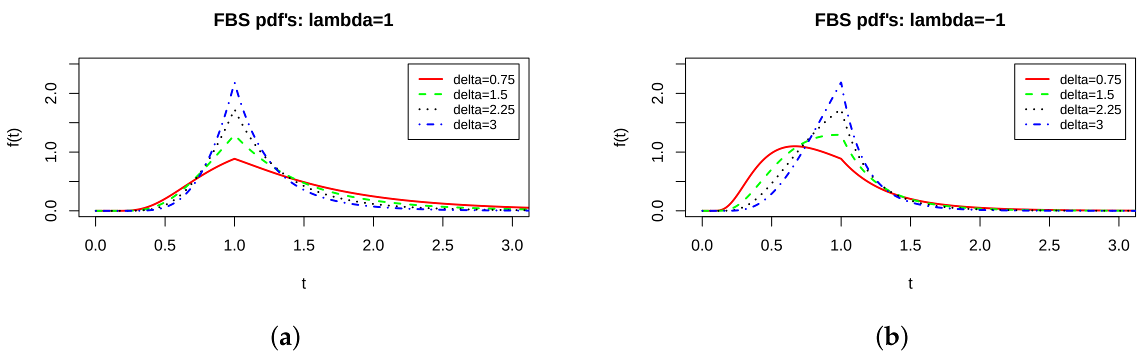
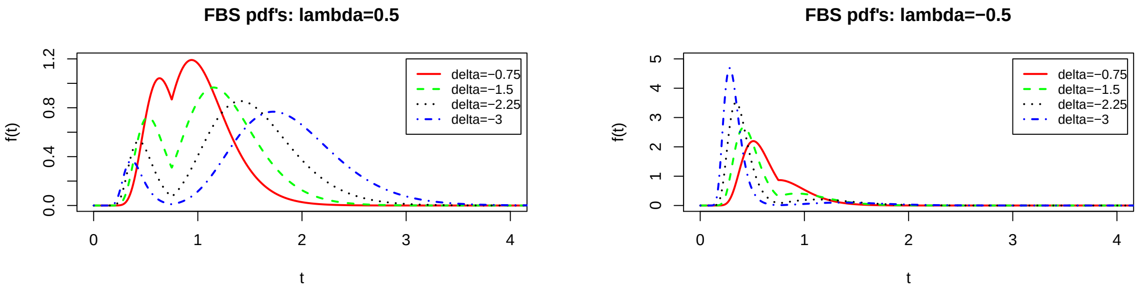


Figure 1. FBS distributions for $\alpha = 0.75$, $\beta = 1$ (both fixed). In (**a**) $\lambda = 1$ versus (**b**) $\lambda = -1$. Increasing values of $\delta > 0$: $\delta = 0.75$ (red solid line), 1.5 (green dashed line), 2.25 (black dotted line) and 3.0 (blue dashed and dotted line).



(**a**) (**b**)

Figure 2. Flexible Birnbaum–Saunders (FBS) distributions for $\alpha = 0.30,\ \beta = 0.75$ (both fixed). In (**a**) $\lambda = 0.5$ versus (**b**) $\lambda = -0.5$. Decreasing values of $\delta < 0$: $\delta = -0.75$ (red solid line), -1.5 (green dashed line), -2.25 (black dotted line) and -3.0 (blue dashed and dotted line).

2.2. Properties

Next, important properties of the FBS model are presented. First an explicit expression for the cdf is given in terms of the cdf of a bivariante normal distribution.

Proposition 1. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$. Then the cdf of T is*

$$F_T(t) = \begin{cases} c_\delta \Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, a_t - \delta\right), & \text{if } 0 < t < \beta \\ c_\delta \left[\Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, -\delta\right) + \Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, a_t + \delta\right) - \Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, \delta\right)\right], & \text{if } t \geq \beta, \end{cases} \tag{9}$$

where $\Phi_{BN_\lambda}(x, y)$ is the cdf of a bivariate normal distribution, with mean vector $\boldsymbol{\mu}' = (0, 0)$ and covariance matrix

$$\Omega_\lambda = \begin{pmatrix} 1 & \rho_\lambda \\ \rho_\lambda & 1 \end{pmatrix} \quad \text{where } \rho_\lambda = -\frac{\lambda}{\sqrt{1+\lambda^2}}\,. \tag{10}$$

Proof. It can be seen in Appendix A. □

Next some particular cases of interest for λ and δ parameters are discussed. Results about the shape of $f_T(\cdot)$ are included.

2.2.1. Effect of λ.

Corollary 1. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$. If $\lambda = 0$ then the cdf of T is*

$$F_T(t) = \begin{cases} \frac{c_\delta}{2}\Phi(a_t - \delta), & \text{if } 0 < t < \beta \\ \frac{c_\delta}{2}\left\{\Phi(a_t + \delta) + 1 - 2\Phi(\delta)\right\}, & \text{if } t \geq \beta\,. \end{cases} \tag{11}$$

Proof. If $\lambda = 0$ then ρ_λ, defined in (10), is equal to zero, and since in the bivariate normal distribution uncorrelation implies independence, we have that

$$\Phi_{BN_{\lambda=0}}(x, y) = \Phi(x)\,\Phi(y), \qquad \forall (x, y)\,.$$

Taking into account that $\Phi(0) = 1/2$ and $\Phi(-\delta) = 1 - \Phi(\delta)$, we have that (9) reduces to (11). □

Result in Corollary 1 corresponds to the model studied in [22].

2.2.2. Effect of δ.

Corollary 2. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$. If $\delta = 0$ then F_T reduces to $F_T(t) = 2\Phi_{BN_\lambda}(0, a_t)$, for $t > 0$.*

Corollary 2 is a particular case of models studied in [15].

2.2.3. Shape of $f_T(\cdot)$.

Proposition 2. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$. Then the pdf given in (8) is nondifferentiable at $t = \beta$.*

Proof. It follows from (8), by noting that if $t = \beta$ then $a_t = 0$ and the absolute value function is not differentiable at zero. □

Proposition 3. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$. The pdf given in (8) can be bimodal. The modes are the solution of the following non-linear equations.*

1. $0 < t_1^* < \beta$ *solution of*
$$a_{t_1} = \delta + \lambda \frac{\phi(\lambda a_{t_1})}{\Phi(\lambda a_{t_1})} + \frac{a''_{t_1}}{\left\{a'_{t_1}\right\}^2}. \tag{12}$$
2. $t_2^* > \beta$ *solution of*
$$a_{t_2} = -\delta + \lambda \frac{\phi(\lambda a_{t_2})}{\Phi(\lambda a_{t_2})} + \frac{a''_{t_2}}{\left\{a'_{t_2}\right\}^2}, \tag{13}$$

With a_t given in (6), a'_t and a''_t the first and second derivatives of a_t with respect to t, respectively.

Proof. It is given in Appendix A. □

Comments on the use of (12) and (13) are included in Appendix A, Remark A1.

Remark 1. *Equations obtained in (12) and (13) are similar to those we have in the skew normal and BS model.*

1. *Let $Z \sim SN(\lambda)$, $\lambda \in \mathbb{R}$. Then Z is unimodal and the mode, z^*, is given by the solution of the non-linear equation*
$$z = \lambda \frac{\phi(\lambda z)}{\Phi(\lambda z)}.$$
2. *Let $T \sim BS(\alpha, \beta)$, $\alpha, \beta > 0$. Then T is unimodal and the mode, t^*, is given by the solution of the non-linear equation*
$$-a_t \left\{a'_t\right\}^2 + a''_t = 0.$$

Next it is shown that the p-th quantile of T can be given in terms of the pth quantile of the $FSN(\delta, \lambda)$. Also it is proved that the FBS model is closed under change of scale and reciprocity.

Theorem 1. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$, with $\alpha,\ \beta \in \mathbb{R}^+$ and $\delta,\ \lambda \in \mathbb{R}$. Then*

(i) *Let t_p be the pth quantile of T, $0 < p < 1$.*
$$t_p = \beta \left(\frac{\alpha}{2} z_p + \sqrt{\left(\frac{\alpha}{2} z_p\right)^2 + 1} \right)^2 \tag{14}$$
where z_p denotes the pth quantile of $Z \sim FSN(\delta, \lambda)$.

(ii) $kT \sim FBS(\alpha, k\beta, \delta, \lambda)$ *for* $k > 0$.

(iii) $T^{-1} \sim FBS(\alpha, \beta^{-1}, \delta, -\lambda)$.

Proof. It can be seen in Appendix A. □

2.2.4. Lifetime Analysis

The BS model is commonly used to explain survival and material resistance data. The survival and risk (or hazard) functions are important indicators in such fields. For the FBS model these functions are given next.

Proposition 4. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$ with $\alpha,\ \beta \in \mathbb{R}^+$ and $\delta,\ \lambda \in \mathbb{R}$. Then*

(i) The survival function is $S(t) = P[T > t] = 1 - F_T(t)$ with $F_T(\cdot)$ given in (9).
(ii) The hazard function, $r(t) = f(t)/S(t)$, is

$$r(t) = \begin{cases} \dfrac{c_\delta a_t' \phi(|a_t|+\delta)\Phi(\lambda a_t)}{1-c_\delta \Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, a_t-\delta\right)}, & \text{if } \ 0 < t < \beta \\ \dfrac{c_\delta a_t' \phi(|a_t|+\delta)\Phi(\lambda a_t)}{1-c_\delta\left[\Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, -\delta\right)+\Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, a_t+\delta\right)-\Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, \delta\right)\right]}, & \text{if } \ t \geq \beta \end{cases}$$

with $\Phi_{BN_\lambda}(\cdot)$ the cdf of the bivariate normal given in Proposition 1.

In Figure 3, the hazard function for those pdf's considered in Figures 1 and 2 are plotted. These graphs show that, for the FBS distribution, the hazard function admits a variety of shapes, which is interesting from an applied point of view.

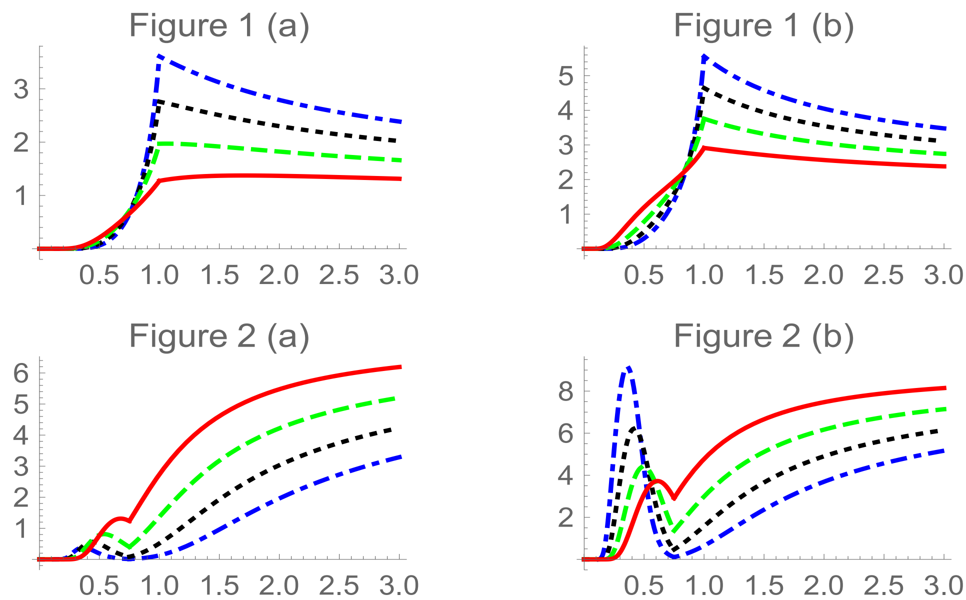


Figure 3. Hazard function of the FBS distribution for plots corresponding to Figure 1 (a), (b): $\alpha = 0.75$, $\beta = 1$ (both fixed), $\delta = 0.75$ (red solid line), 1.5 (green dashed line), 2.25 (black dotted line) and 3.0 (blue dashed dotted line), in (a) $\lambda = 1$ versus (b) $\lambda = -1$. For plots corresponding to Figure 2 (a), (b): $\alpha = 0.30$, $\beta = 0.75$ (both fixed), $\delta = -0.75$ (red solid line), –1.5 (green dashed line), –2.25 (black dotted line) and –3.0 (blue dashed dotted line), in (a) $\lambda = 0.5$ versus (b) $\lambda = -0.5$.

Remark 2. *More complicated hazard functions than the traditional ones are obtained when we are dealing with models with complex structure, as it happens with the FBS. For instance, in Figure 3, we have two situations:*

1. *$r(t)$ corresponding to Figure 1a,b. These are, first, quickly increasing, later decreasing more slowly or even in a flat way. It can be applied in practical situations in which the risk of failure increases quickly until certain point in which its behaviour becomes flatter. As [23] points out, the flat area is very interesting in survival analysis and reliability contexts.*

2. *$r(t)$ corresponding to Figure 2a,b are increasing-decreasing-increasing. This kind of hazard functions has been recently introduced and discussed in literature, due to its interest in reliability of systems, see for instance [23] or [24] (and references therein). In plot for Figure 2b, $r(t)$ is (quickly) increasing—or (quickly) decreasing. On the other hand, for Figure 2a the initial effect increasing—decreasing is less accentuated.*

3. Moments and Maximum Likelihood Estimation

Moments of the FBS model can be obtained from the moments of the flexible skew-normal model given in [2]. The following results present important properties relating those distributions, and the expressions for the first moment and variance in the FBS model.

Theorem 2. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$ and $Z \sim FSN(\delta, \lambda)$. Then $\mathbb{E}(T^r)$, $r = 0, 1, \ldots$, always exists. Moreover*

$$\mathbb{E}(T^r) = \frac{\beta^r}{4^r} \sum_{k=0}^{2r} \binom{2r}{k} \mathbb{E}\left[(\alpha Z)^k \left(\alpha^2 Z^2 + 4 \right)^{\frac{2r-k}{2}} \right], \qquad r = 0, 1, \ldots. \tag{15}$$

Proof. From (7), we can write

$$T = \frac{\beta}{4} \left\{ \alpha Z + \left(\alpha^2 Z^2 + 4 \right)^{1/2} \right\}^2 .$$

Taking expectation of the rth-power of T

$$\mathbb{E}(T^r) = \frac{\beta^r}{4^r} \mathbb{E}\left[\left\{ \alpha Z + \left(\alpha^2 Z^2 + 4 \right)^{1/2} \right\}^{2r} \right] . \tag{16}$$

From (16), note that for $r = 0, 1, \ldots$, $\mathbb{E}(T^r)$ exists if and only if $\mathbb{E}(Z^{2r})$ exists. On the other hand, it can be seen in [2] that $\mathbb{E}(Z^{2r})$ always exists, and therefore $\mathbb{E}(T^r)$ too.

Finally note that (15) is the result of applying the binomial formula to (16). □

Next, explicit expressions for the expected value and variance of $T \sim FBS(\alpha, \beta, \delta, \lambda)$ are given. In these expressions, $\kappa_j = \mathbb{E}_{SFN}\left(\frac{Z^j}{2} \sqrt{\alpha^2 Z^2 + 4} \right)$ with $Z \sim FSN(\delta, \lambda)$.

Theorem 3. *Let $T \sim FBS(\alpha, \beta, \delta, \lambda)$. Then*

$$\mathbb{E}(T) = \beta \left[1 + \alpha \kappa_1 + \frac{\alpha^2}{2} c_\delta \left\{ (1 + \delta^2)(1 - \Phi(\delta)) - \delta \phi(\delta) \right\} \right] , \tag{17}$$

$$\begin{aligned} \mathbb{E}(T^2) = \beta^2 & \left[\frac{7\alpha^4 c_\delta}{16} \left((3 + 6\delta^2 + \delta^4)(1 - \Phi(\delta)) - \delta(5 + \delta^2)\phi(\delta) \right) + \alpha^3 \kappa_3 + 2\alpha\kappa_1 + 1 \right] \\ & + 2\alpha^2 \beta^2 c_\delta \left((1 + \delta^2)(1 - \Phi(\delta)) - \delta \phi(\delta) \right) \end{aligned}$$

and

$$\begin{aligned} Var(T) = \beta^2 & \left[\frac{7\alpha^4 c_\delta}{16} \left((3 + 6\delta^2 + \delta^4)(1 - \Phi(\delta)) - \delta(5 + \delta^2)\phi(\delta) \right) + \alpha^3 \kappa_3 - \alpha^2 \kappa_1 + 1 \right] \\ - \frac{\alpha^2 \beta^2 c_\delta}{4} & \left((1 + \delta^2)(1 - \Phi(\delta)) - \delta \phi(\delta) \right) \left(\alpha^2 c_\delta \left((1 + \delta^2)(1 - \Phi(\delta)) - \delta \phi(\delta) \right) + 4(1 - \alpha \kappa_1) \right) \end{aligned}$$

where $\kappa_j = \mathbb{E}_{FSN}\left(\frac{Z^j}{2} \sqrt{\alpha^2 Z^2 + 4} \right)$ and $Z \sim FSN(\delta, \lambda)$.

Proof. These results follows from Theorem 2 and the expressions of κ_j, which have been computed by using the results for moments of $Z \sim FSN(\delta, \lambda)$ obtained in [2].

As illustration, note that for the case $r = 1$, (15) reduces to

$$\begin{aligned}\mathbb{E}(T) &= \frac{\beta}{4}\left[\mathbb{E}(\alpha^2 Z^2 + 4) + 2\mathbb{E}\left((\alpha Z)\sqrt{\alpha^2 Z^2 + 4)}\right) + \mathbb{E}(\alpha^2 Z^2)\right] \\ &= \beta\left[1 + \alpha\kappa_1 + \frac{\alpha^2}{2}\mathbb{E}(Z^2)\right],\end{aligned}$$

it can be seen in [2] that $\mathbb{E}(Z^2) = c_\delta(1+\delta^2)\,(1-\Phi(\delta)) - \delta\phi(\delta))$, and so (17) is obtained. □

3.1. Maximum Likelihood Estimators

Parameter estimation in the BS model has been the topic of interest in many papers. Among others, we mentioned [25–27]. To estimate the parameters in the usual BS model, the modified moment method (MME) and maximum likelihood (MLE) are commonly used. To start the maximum likelihood approach moment estimators are used which are given by

$$\hat{\beta}_M = \sqrt{sr}, \qquad \hat{\alpha}_M = \sqrt{2\left(\sqrt{\frac{s}{r}} - 1\right)}.$$

where $s = \frac{1}{n}\sum_{i=1}^n t_i$ and $r = \left(\frac{1}{n}\sum_{i=1}^n \frac{1}{t_i}\right)^{-1}$ are the sample (arithmetic) and harmonic mean, respectively. Relevant aspects of this distribution such as its robustness with respect to parameter estimation and $O(n^{-1})$ bias corrections for MLEs, can be seen in [25–27].

In the following, we discuss MLE estimation for the FBS model in depth. Thus, given n observations independent and identically distributed, $T_1, T_2, ..., T_n$, with $T_i \sim FBS(\alpha, \beta, \delta, \lambda)$, the log-likelihood function for the parameter vector $\boldsymbol{\theta} = (\alpha, \beta, \delta, \lambda)'$ is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}) = -n\left(\log(\alpha) + \frac{1}{2}\log(\beta) + \log\left(1 - \Phi(\delta)\right)\right) - \frac{3}{2}\sum_{i=1}^n \log(t_i) + \sum_{i=1}^n \log(t_i + \beta) \\ - \frac{1}{2}\sum_{i=1}^n (a_{t_i}^2 + 2\delta|a_{t_i}| + \delta^2) + \sum_{i=1}^n \log\left(\Phi(\lambda a_{t_i})\right).\end{aligned} \tag{18}$$

To maximize $l(\boldsymbol{\theta})$ in $\boldsymbol{\theta}$, consider the first derivatives of $l(\boldsymbol{\theta})$ with respect to α, β, δ and λ, denoted as $\dot{l}_\alpha$, $\dot{l}_\beta$, $\dot{l}_\delta$ and $\dot{l}_\lambda$, respectively. From $\dot{l}_\alpha = 0$, $\dot{l}_\beta = 0$, $\dot{l}_\delta = 0$ and $\dot{l}_\lambda = 0$, the likelihood equations are given by

$$-n + \sum_{i=1}^n a_{t_i}^2 - \delta\sum_{i=1}^n |a_{t_i}| - \lambda\sum_{i=1}^n a_{t_i}\frac{\phi(\lambda a_{t_i})}{\Phi(\lambda a_{t_i})} = 0 \tag{19}$$

$$-\frac{n}{2\beta} + \sum_{i=1}^n \frac{1}{t_i + \beta} + \frac{1}{2\alpha\beta^{3/2}}\sum_{i=1}^n \left(sgn(a_{t_i})\,(|a_{t_i}| + \delta) - \lambda\frac{\phi(\lambda a_{t_i})}{\Phi(\lambda a_{t_i})}\right)\frac{t_i + \beta}{\sqrt{t_i}} = 0 \tag{20}$$

$$\delta - \frac{\phi(\delta)}{1 - \Phi(\delta)} = -\frac{1}{n}\sum_{i=1}^n |a_{t_i}| \tag{21}$$

$$\sum_{i=1}^n a_{t_i}\frac{\phi(\lambda a_{t_i})}{\Phi(\lambda a_{t_i})} = 0 \tag{22}$$

in which $sgn(\cdot)$ denotes the *sign* function.

The solution to the previous system of equations must be obtained by iterative methods such as the Newton-Raphson or quasi-Newton procedures, which can be implemented using the statistical software R, [28].

As initial estimates of α and β can be proposed the estimates of these parameters obtained in the basic BS model, denoted as $\widehat{\alpha}_0$ and $\widehat{\beta}_0$. These estimates can be plugged into (21) and (22) to obtain preliminar estimates of δ and λ, $\widehat{\delta}_0$ and $\widehat{\lambda}_0$, and so, start the recursion.

3.2. Expected and Observed Information Matrices

Recall that, the Fisher information matrix is given by

$$I(\boldsymbol{\theta}) \;=\; (j_{i,j})_{i,j=\alpha,\beta,\delta,\lambda}\,,$$

which entries are equal to minus the second partial derivatives of the log-likelihood function given in (18) with respect to the parameters in the model. They are denoted as $j_{\alpha\alpha} = -\frac{\partial^2}{\partial\alpha^2} l(\boldsymbol{\theta})$, and so on. So we have

$$j_{\alpha\alpha} = -\frac{n}{\alpha^2} + \frac{1}{\alpha^2}\sum_{i=1}^{n}\left(3a_{t_i}^2 + 2\delta|a_{t_i}|\right) + \frac{\lambda}{\alpha^2}\sum_{i=1}^{n} a_{t_i} w_i \left(2 + \lambda a_{t_i} B_i\right),$$

$$j_{\beta\alpha} = -\frac{1}{\alpha^3\beta^2}\sum_{i=1}^{n}\left(\frac{\beta^2 - t_i^2}{t_i}\right) + \frac{1}{2\alpha^2\beta^{3/2}}\sum_{i=1}^{n}\left(\delta sgn(a_{t_i}) + \lambda w_i(-1 + \lambda a_{t_i} B_i)\right)\left(\frac{t_i + \beta}{\sqrt{t_i}}\right),$$

$$\begin{aligned} j_{\beta\beta} = {}& -\frac{n}{2\beta^2} + \sum_{i=1}^{n}\frac{1}{t_i + \beta} + \frac{1}{\alpha^2\beta^3}\sum_{i=1}^{n} t_i + \frac{1}{4\alpha\beta^{5/2}}\sum_{i=1}^{n}\left(\delta sgn(a_{t_i}) - \lambda w_i\right)\left(\frac{3t_i + \beta}{\sqrt{t_i}}\right) \\ & + \frac{\lambda^2}{4\alpha^2\beta^{3/2}}\sum_{i=1}^{n} w_i B_i \left(\frac{t_i + \beta}{\sqrt{t_i}}\right)^2, \end{aligned}$$

$$j_{\delta\alpha} = -\frac{1}{\alpha}\sum_{i=1}^{n}|a_{t_i}|, \qquad j_{\delta\beta} = -\frac{1}{2\alpha^2\beta^{3/2}}\sum_{i=1}^{n} sgn(a_{t_i})\left(\frac{t_i + \beta}{\sqrt{t_i}}\right),$$

$$j_{\lambda\alpha} = \frac{1}{\alpha}\sum_{i=1}^{n} a_{t_i} w_i \left(1 - \lambda a_{t_i} B_i\right), \qquad j_{\lambda\beta} = \frac{1}{2\alpha\beta^{3/2}}\sum_{i=1}^{n} w_i\left(1 + \lambda a_{t_i} B_i\right)\left(\frac{t_i + \beta}{\sqrt{t_i}}\right),$$

$$j_{\delta\delta} = n\left(w_\delta(\delta - w_\delta) + 1\right), \qquad j_{\delta\alpha} = j_{\delta\beta} = j_{\lambda\delta} = 0, \qquad j_{\lambda\lambda} = \sum_{i=1}^{n} a_{t_i}^2 w_i B_i$$

where $w = \frac{\phi(\lambda a_t)}{\Phi(\lambda a_t)}$, $w_\delta = \phi(\delta)/(1 - \Phi(\delta))$ and $B = \lambda a_t + w$.

The Fisher (expected) information matrix would be obtained by computing the expected values of the above second derivatives. Taking in this matrix $\delta = \lambda = 0$, that is, $T \sim BS(\alpha, \beta)$, and, using results in [21], we have

$$I(\boldsymbol{\theta}) = \begin{pmatrix} \frac{2}{\alpha^2} & 0 & -\frac{1}{\alpha}\sqrt{\frac{2}{\pi}} & 0 \\ 0 & \alpha^{-2}\beta^{-2}\left(1 + \frac{\alpha q(\alpha)}{\sqrt{2\pi}}\right) & 0 & \frac{1}{\sqrt{2\pi}}\frac{1}{\alpha\beta^{3/2}}A_1(t) \\ -\frac{1}{\alpha}\sqrt{\frac{2}{\pi}} & 0 & 1 - \frac{2}{\pi} & 0 \\ 0 & \frac{1}{\sqrt{2\pi}}\frac{1}{\alpha\beta^{3/2}}A_1(t) & 0 & \frac{2}{\pi} \end{pmatrix},$$

where $A_1(t) = \mathbb{E}\left(\frac{t_i + \beta}{\sqrt{t_i}}\right)$, $q(\alpha) = \alpha\sqrt{\frac{2}{\pi}} - \frac{\pi\exp(\frac{2}{\alpha^2})}{2} erfc(\frac{2}{\alpha})$, with $erfc(\cdot)$ the error function, i.e., $erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty \exp(-t^2)dt$, see [29].

It can be shown that $|I(\boldsymbol{\theta})| \neq 0$, so the Fisher information matrix is not singular at $\delta = \lambda = 0$.

Hence, for large samples, the MLE, $\widehat{\boldsymbol{\theta}}$, of $\boldsymbol{\theta}$ is asymptotically normal, that is,

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \xrightarrow{\mathcal{L}} N_4(\mathbf{0},\ I(\boldsymbol{\theta})^{-1}),$$

resulting that the asymptotic variance of the MLE, $\hat{\boldsymbol{\theta}}$, is the inverse of Fisher information matrix $I(\boldsymbol{\theta})$. Since the parameters are unknown, usually the observed information matrix is considered where the unknown parameters are estimated by ML.

Asymptotic confidence intervals for the parameters in the FBS model can be obtained from these results.

4. Numerical Illustrations

The numerical illustrations introduced next are aimed to show that the FBS model can be an alternative to modelling unimodal or bimodal data from different areas. First illustration is related to nickel content in soil samples analyzed at the Mining Department (Departamento de Minas) of Universidad de Atacama, Chile. We start by showing that both, BS and skew-BS (SBS) models are not able to capture bimodality present in this data set. Thus, the FBS model turned out to be a good option, to fit the data even better than a mixture of two normal distributions, which is another competing alternative to fit bimodal data. Second illustration is related to air pollution in New York city in USA, which was previously analyzed in [16,30]. In this case, it is shown that FBS model again provides a better fit than BS and SBS. As competing model the extended Birnbaum-Saunders (EBS) is also considered. Recall that the $EBS(\alpha, \beta, \sigma, \nu, \lambda)$ is a five-parameter model proposed in [16] where the parameter σ affects the kurtosis; ν and λ affect the skewness; and α and β the shape and scale as in the usual BS model. We highlight that, for this dataset, the $FBS(\alpha, \beta, \delta, \lambda)$ model provides a better fit than that given by the $EBS(\alpha, \beta, \sigma, \nu, \lambda)$ in [16] with the merit of using less parameters.

4.1. Nickel Concentration

For illustrative purposes, we apply the FBS model to a data set related to nickel content in soil samples. This data set encompasses 85 observations of the variable concentration of nickel with sample mean = 21.588, sample standard deviation = 16.573, sample asymmetry = 2.392 and sample kurtosis = 8.325, much higher than expected with the ordinary BS distribution.

4.1.1. FBS versus the BS and SBS distributions

To fit the nickel concentration variable, we use the BS, skew BS (SBS) and FBS models. Using function optim from the R-package, [28], the following point estimates (and their standard errors) are obtained for each of the three models under consideration
BS model: $\hat{\alpha} = 0.789\ (0.060)$ and $\hat{\beta} = 16.382\ (1.296)$.
SBS model: $\hat{\alpha} = 1.073\ (0.201)$, $\hat{\beta} = 8.841\ (1.998)$ and $\hat{\lambda} = 1.252\ (0.590)$.
FBS model: $\hat{\alpha} = 0.870\ (0.104)$, $\hat{\beta} = 5.072\ (0.763)$, $\hat{\delta} = -1.520\ (0.282)$ and $\hat{\lambda} = 1.405\ (0.341)$.

The bimodal hypothesis can be formally tested as follows

$$H_0 : \delta = 0 \quad \text{versus} \quad H_1 : \delta \neq 0,$$

which is equivalent to compare models SBS versus FBS. Given the nonsingularity of the Fisher information matrix, and since these models are nested, we can consider the likelihood ratio statistics, namely

$$\Lambda_1 = L_{SBS}(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda}) / L_{FBS}(\widehat{\alpha}, \widehat{\beta}, \widehat{\delta}, \widehat{\lambda}).$$

It is obtained $-2\log(\Lambda_1) = 5.618$, which is greater than the 5% chi-square critical value with one degree of freedom (df), which is equal to 3.84. Therefore, the null hypothesis of no-bimodality is rejected at the 5% critical level, leading to the conclusion that FBS model fits better than the unimodal SBS model to the nickel concentration data.

To compare the FBS model with the BS model, consider to test the null hypothesis of a BS distribution versus a FBS distribution, that is

$$H_0 : (\delta, \lambda) = (0, 0) \quad \text{vs} \quad H_1 : (\delta, \lambda) \neq (0, 0)$$

using the likelihood ratio statistics based on the ratio $\Lambda_2 = L_{BS}(\widehat{\alpha}, \widehat{\beta}) / L_{FBS}(\widehat{\alpha}, \widehat{\beta}, \widehat{\delta}, \widehat{\lambda})$. After substituting the estimated values, we obtain $-2\log(\Lambda_2) = 7.628$, which is greater than the 5% chi-square critical value with 2 df, which is 5.99. Therefore the FBS is preferred to BS model for this data set.

4.1.2. FBS versus a Mixture of Normal Distributions

Another model widely applied in such situations of bimodality is the mixture of two normal distributions. The normal mixture model is given by:

$$f(x; \mu_1, \sigma_1, \mu_2, \sigma_2, p) = p f_1(x, \mu_1, \sigma_1) + (1 - p) f_2(x; \mu_2, \sigma_2) \tag{23}$$

where f_j is a normal distribution with parameters (μ_j, σ_j), $j = 1, 2$ and $0 < p < 1$. (23) is denoted by $\text{MN}(\mu_1, \sigma_1, \mu_2, \sigma_2, p)$.

To compare FBS model with the MN model, we propose the Akaike information criterion (AIC), see [31], namely $AIC = -2\hat{\ell}(\cdot) + 2k$, the modified AIC criterion (CAIC), typically called the consistent AIC, namely $CAIC = -2\hat{\ell}(\cdot) + (1 + \log(n))k$ and the Bayesian Information Criterion, BIC, $BIC = -2\hat{\ell}(\cdot) + \log(n)k$, where k is the number of parameters and $\hat{\ell}(\cdot)$ is the log-likelihood function evaluated at the MLEs of parameters. The best model is the one with the smallest AIC or CAIC or BIC.

Now we compare the FBS with $\text{MN}(\mu_1, \sigma_1, \mu_2, \sigma_2, p)$. The estimated mixture model is

$$\text{MN}(15.348,\ 6.622,\ 40.908,\ 21.960,\ 0.755)$$

with $AIC = 674.849$, $CAIC = 692.061$ and $BIC = 687.062$. On the other hand, for the FBS model, we have $AIC = 671.859$, $CAIC = 685.628$ and $BIC = 681.630$. According to these criteria, the FBS model provides a better fit to the data of nickel concentration.

4.1.3. FBS versus a Mixture of Log-Normal Distributions

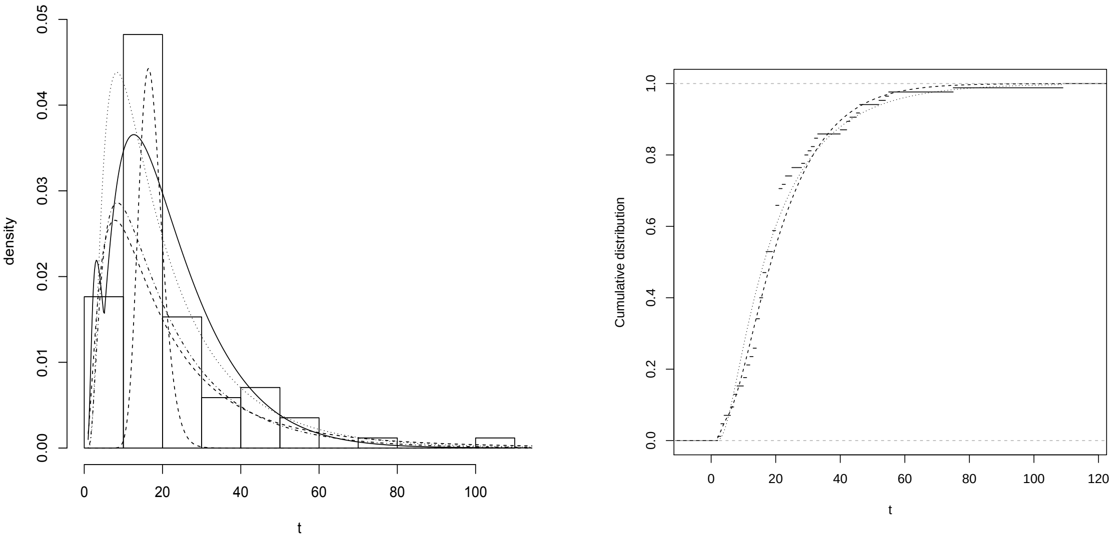
Following reviewer's recommendations, a mixture of two log-normal distributions is also considered. The log-normal mixture model will be given by (23) with f_j the pdf of a log-normal distribution with parameters (μ_j, σ_j), $j = 1, 2$ and $0 < p < 1$, and it is denoted by $\text{MLN}(\mu_1, \sigma_1, \mu_2, \sigma_2, p)$. The estimated mixture model is

$$\text{MLN}(2.829,\ 0.177,\ 2.8275,\ 0.877,\ 0.327)$$

with $AIC = 663.571$, $CAIC = 680.784$ and $BIC = 675.784$. All of them are less than those corresponding to FBS. So, according to these criteria, the mixture ot two log-normal distributions provides a better fit to this dataset than the FBS model.

This discussion illustrates that, quite often, the final selection of a model is a matter of choice. FBS model can be considered as appropriate if we want to use a more parsimonious model, and this is better than other BS models and a mixture of two normal distributions. On the other hand, based on AIC, CAIC and BIC, the mixture of two log-normal would be preferred but this model has more parameters than the FBS distribution and may present problems of identifiability. Anyway, the final choice must be properly justified.

Figure 4 depicts maximized likelihoods and empirical cdf for variable nickel concentration revealing that FBS model fitting is quite good.



(**a**) (**b**)

Figure 4. (**a**) Plots for FBS, (solid line), MLN (dashed line), BS (dotted line) and SBS (dotted and dashed line) models . (**b**) Empirical cdf with estimated FBS cdf (dashed line) and estimated BS cdf (dotted line).

Remark 3. *Going through the origin of this data set, the bimodal behavior of the nickel concentration statistical model seems to be due to the fact that the samples were taken according to different lithologies. Lithology classifies according to the physical and chemical elements in rock formation. Mining operations found different lithologies in these samples, as it is depicted in Figure 4.*

4.2. Air Pollution

The concentration of average air pollutants has been used in epidemiological surveillance as an indicator of the level of atmospheric contamination. Among its associated adverse effects in humans, diseases such as bronchitis are found. The distribution of this concentration has a bias to the right, and is always positive. It is typically assumed that the data on air pollutant concentrations are uncorrelated and independent and thus they do not require the diurnal or cyclic trend analysis, see [32]. The data set studied in this section corresponds to daily measures of ozone levels (in $ppb = ppm \times 1000$) in the city of New York, USA, from May to September, 1973, collected by the New York State Conservation Department. The sample mean, standard deviation, asymmetry and kurtosis coefficients are given, respectively, by 42.129, 32.987, 1.209 and 1.112.

4.2.1. FBS versus the BS and SBS Distributions

Maximum likelihood estimators, their estimated standard errors (in parenthesis), for the BS, SBS and FBS models, were computed, the results are:
BS model: $\hat{\alpha} = 0.982\ (0.064)$ and $\hat{\beta} = 28.031\ (2.265)$.
SBS model: $\hat{\alpha} = 1.270\ (0.235)$, $\hat{\beta} = 14.831\ (4.019)$ and $\hat{\lambda} = 1.066\ (0.533)$.
FBS: $\hat{\alpha} = 5.160\ (0.481)$, $\hat{\beta} = 78.000\ (0.008)$, $\hat{\delta} = 3.991\ (0.050)$ and $\hat{\lambda} = -9.135\ (2.417)$.
For this data set, the log-likelihood ratio statistics to test BS vs FBS and SBS vs FBS are given by

$$-2\log(\Lambda_1) = 12.812 \ \text{and} \ -2\log(\Lambda_2) = 5.828,$$

which are greater than the corresponding 5% critical values from the chisquare distribution, which are 5.99 (with 2 df) and 3.84 (with one df), respectively. So, we can conclude that the unimodal FBS model provides a better fit to this dataset than BS and FBS models.

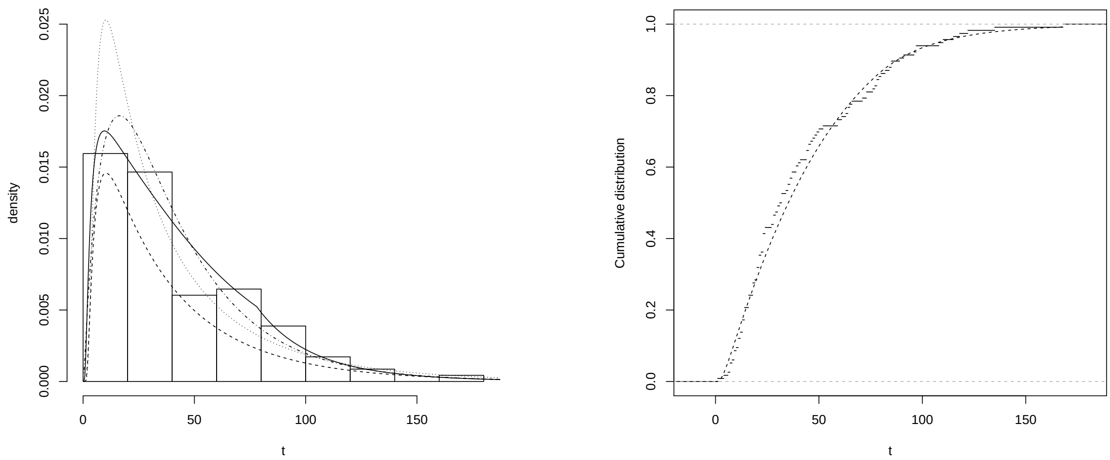
4.2.2. FBS versus the Extended BS (EBS) Model

Fitting the five-parameter EBS model, $EBS(\alpha, \beta, \sigma, \nu, \lambda)$, proposed in [16] as the best for this dataset, and whose point estimates for the parameters and summaries for comparison are equal to

$$EBS(0.780, 0.596, 3.618, -3.539, -0.096),$$

$CAIC = 1111.154$ and $BIC = 1106.154$. On the other hand, for the FBS model, it was obtained $CAIC = 1108.396$ and $BIC = 1104.396$. Then, according to the CAIC and BIC criteria, FBS model presents the best fit to this data set dealing with the daily ozone level concentration in the atmosphere.

Figure 5 depicts the histograms and the fitted density curves for the data set studied and empirical cdf for variable daily ozone level concentration in the atmosphere, while the dashed line corresponds to the cfd for FBS model.



(**a**) (**b**)

Figure 5. (**a**) Plots for FBS, (solid line), BS (dotted line), SBS (dashed line) and EBS (dotted and dashed line) models. (**b**) Empirical cdf with estimated FBS cdf (dashed line).

5. Conclusions

We have introduced a new family of distributions able to model skewness, unimodality and bimodality in the BS distribution. We have discussed several of its properties. Explicit expressions for the cdf are given in terms of the cdf of a bivariate normal variable. Non-linear equations to obtain the modes of this distribution are provided. The estimation of parameters is carried out via maximum likelihood. We highlight that the ML equations must be solved by using iterative methods. The information matrix is non-singular and therefore likelihood ratio tests to compare this model with other nested models can be implemented. The interest and flexibility of our proposal is supported with two illustrations to real data in which we show that:

(i) the FBS model provides consistently better fits than the BS and SBS models (they can be considered relevant precedents of our proposal)
(ii) the FBS distribution can improve the fit provided by other competing models designed to deal with bimodality (such as a mixture of normal distributions). It can also perform better for unimodal situations in which a generalized BS model with skewness parameters must be applied, such as the EBS model proposed in [16]. We highlight that in both situations FBS provides a better fit with a more parsimonious model (less number of parameters), and the problem of identifiability of mixtures can be circumvented.

Therefore the outcome of these practical demonstrations show that the new family is very flexible and widely applicable.

Author Contributions: All the authors contributed significantly to the present paper.

Funding: The research of G. Martínez-Flórez was supported by Grant Proyecto Universidad de Córdoba: Distribuciones de probabilidad asimétrica bimodal con soporte positivo (Colombia). The research of I. Barranco-Chamorro was supported by Grant CTM2015-68276-R (Spain). H.W. Gómez was supported by Grant SEMILLERO UA-2019 (Chile).

Acknowledgments: The authors would like to thank the editor and the anonymous referees for their comments and suggestions, which significantly improved our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Proposition 1 (cdf in the FBS model). From Equation (7), T is a monotonically increasing function of $Z \sim FSN(\delta, \lambda)$. Therefore the cdf of T is given by

$$F_T(t) = F_Z(a_t) \tag{A1}$$

where $F_Z(\cdot)$ denotes the cdf of Z and a_t was given in (6).

(i) First, we obtain the cdf of $Z \sim FSN(\delta, \lambda)$

It can be seen in Gómez et al. [2], Proposition 4, that the pdf of $Z \sim FSN(\delta, \lambda)$ is

$$f_Z(z) = \begin{cases} c_\delta \phi(z-\delta)\Phi(\lambda z), & \text{if } z < 0 \\ c_\delta \phi(z+\delta)\Phi(\lambda z), & \text{if } z \geq 0. \end{cases}$$

Let us consider the case for $z < 0$

$$F_Z(z) = \int_{-\infty}^{z} f_Z(t)dt \ = \int_{-\infty}^{z} c_\delta \phi(t-\delta)\Phi(\lambda t)dt$$

By making the change of variable $v = t - \delta$, and later, taking into account that $\Phi(\cdot)$ is the cdf of a $N(0,1)$ distribution, we have that

$$F_Z(z) = c_\delta \int_{-\infty}^{z-\delta} \phi(v)\Phi(\lambda(v+\delta))dv = c_\delta \int_{-\infty}^{z-\delta} \int_{-\infty}^{\lambda(v+\delta)} \phi(v)\phi(s)dsdv \tag{A2}$$

The integrand in (A2) is the joint pdf of two independent $N(0,1)$ rv's, (S, V), i.e.,

$$\begin{pmatrix} S \\ V \end{pmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

Note that (A2) can be rewritten as

$$\begin{aligned} F_Z(z) &= c_\delta \, Pr\left[S - \lambda V \leq \lambda\delta, \ V \leq z - \delta\right] \\ &= c_\delta \, Pr\left[\frac{S - \lambda V}{\sqrt{1+\lambda^2}} \leq \frac{\lambda\delta}{\sqrt{1+\lambda^2}}, \ V \leq z - \delta\right] \\ &= c_\delta \Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, z - \delta\right) \end{aligned} \tag{A3}$$

where $\Phi_{BN_\lambda}(x, y)$ denotes the cdf of a bivariate normal distribution, with mean vector $\boldsymbol{\mu}' = (0, 0)$ and covariance matrix Ω_λ given in (10).

For $z > 0$, we have that

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{0} f_Z(t)dt \ + \int_{0}^{z} f_Z(t)dt \\ &= F_Z(0) + c_\delta \int_0^z \phi(t+\delta)\Phi(\lambda t)dt \, . \end{aligned} \tag{A4}$$

From (A3), it follows that

$$F_Z(0) = \lim_{z \to 0^-} F_Z(z) \;=\; c_\delta \Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, -\delta\right) \tag{A5}$$

On the other hand, proceeding similarly to the previous case (change of variable $v = t + \delta$), it can be proved that

$$\begin{aligned} \int_0^z \phi(t+\delta)\Phi(\lambda t)dt &= Pr\left[\frac{S-\lambda V}{\sqrt{1+\lambda^2}} \le -\frac{\lambda\delta}{\sqrt{1+\lambda^2}},\ \delta < V \le z+\delta\right] \\ &= \Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, z+\delta\right) - \Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, \delta\right) \end{aligned} \tag{A6}$$

Therefore, from (A3)–(A6), we have just proved that the cdf of Z is

$$F_Z(z) = \begin{cases} c_\delta \Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, z-\delta\right), & \text{if } z < 0 \\ c_\delta\left[\Phi_{BN_\lambda}\left(\frac{\lambda\delta}{\sqrt{1+\lambda^2}}, -\delta\right) + \Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}},\ z+\delta\right) - \Phi_{BN_\lambda}\left(-\frac{\lambda\delta}{\sqrt{1+\lambda^2}},\ \delta\right)\right], & \text{if } z \ge 0 \end{cases} \tag{A7}$$

(ii) Finally, the expression for the cdf of $T \sim FBS(\alpha, \beta, \delta, \lambda)$ given in (9) follows from (A1) and (A7). □

Proof of Proposition 3 (Modes in the FBS model). Recall that, from (A1), the pdf of T is given by

$$f_T(t) = f_Z(a_t)a_t' = c_\delta a_t' \phi(|a_t| + \delta)\Phi(\lambda a_t)$$

where $f_Z(\cdot)$ denotes the pdf of $Z \sim FSN(\delta, \lambda)$, a_t was given in (6) and

$$a_t' = \frac{\partial}{\partial t}a_t = \frac{t^{-3/2}\beta^{-1/2}}{2\alpha}(t+\beta)\,. \tag{A8}$$

For $a_t < 0$, or equivalently $0 < t < \beta$, consider the first derivative with respect to t of $f_T(\cdot)$ and equating to zero, we have

$$f_T'(t) = c_\delta \frac{\partial}{\partial t}\left\{a_t'\phi(-a_t+\delta)\Phi(\lambda a_t)\right\} = 0 \tag{A9}$$

By using that $\phi'(z) = -z\phi(z)$, it can be proved that (A9) is equivalent to

$$\left\{a_t'\right\}^2\left[(\delta - a_t)\Phi(\lambda a_t) + \lambda\phi(\lambda a_t)\right] + a_t''\Phi(\lambda a_t) = 0 \tag{A10}$$

Since $a_t' > 0, \forall t > 0$ ($\beta > 0$), we have that (A10) is equivalent to (12).
Similarly, for $a_t > 0$, i.e., $t > \beta$, from $f_T'(t) = c_\delta \frac{\partial}{\partial t}\left\{a_t'\phi(a_t+\delta)\Phi(\lambda a_t)\right\} = 0$, (13) is obtained. □

Remark A1 (Comments to Proposition 3)**.** *In order to illustrate the use of Equations (12) and (13) next cases are considered.*

1. *Consider the pdf given in Figure 1a, case $\alpha = 0.75$, $\beta = 1$, $\lambda = 1$, $\delta = 0.75$. In this setting there do not exist $t_1^* \in (0, \beta)$ and $t_2^* > \beta$ satisfying (12) and (13), respectively. It can be checked than the distribution is unimodal and the mode is at β.*
2. *Figure 1b, case $\alpha = 0.75$, $\beta = 1$, $\lambda = -1$, $\delta = 0.75$. There exists $t_1^* \in (0, \beta)$ satisfying (12) and there does not exists $t_2^* > \beta$ satisfying (13). Then the distribution is unimodal and the mode is at t_1^*.*
3. *Figure 2a,b, in all cases under consideration, there exist $t_1^* \in (0, \beta)$ and $t_2^* > \beta$ satisfying (12) and (13). Then the distribution is bimodal and the modes are t_1^* and t_2^*.*

Proof of Theorem 1 (*pth* quantile, change of scale and reciprocity). (i) (14) follows from the fact that (7) is one-to-one function preserving the order from $\mathbb{R}$ to $\mathbb{R}^+$.

(ii) Note that the pdf of T can be rewritten as

$$f_T(t) = c_\delta a_t'(\alpha,\beta)\phi(|a_t(\alpha,\beta)| + \delta)\Phi(\lambda a_t(\alpha,\beta)) \tag{A11}$$

with $c_\delta = (1-\Phi(\delta))^{-1}$, $a_t = a_t(\alpha,\beta)$ given in (6) and

$$a_t' = a_t'(\alpha,\beta) = \frac{\partial}{\partial t} a_t(\alpha,\beta) = \frac{t^{-3/2}\beta^{-1/2}}{2\alpha}(t+\beta)\,. \tag{A12}$$

Let $Y = kT$ with $k > 0$. By applying the Jacobian technique $f_Y(y) = |J| f_T(\frac{y}{k};\alpha,\beta,\delta,\lambda)$ with $|J| = \frac{1}{k}$. From (6), $a_{y/k}(\alpha,\beta) = a_y(\alpha,k\beta)$, and from (A12)

$$|J|a_{y/k}'(\alpha,\beta) = \frac{y^{-3/2}(k\beta)^{-1/2}}{2\alpha}(y+k\beta) = a_y'(\alpha,k\beta)\,.$$

Therefore

$$f_Y(y) = c_\delta a_y'(\alpha,k\beta)\phi(|a_y(\alpha,k\beta)| + \delta)\Phi(\lambda a_y(\alpha,k\beta)),$$

i.e., $Y \sim FBS(\alpha,k\beta,\delta,\lambda)$.

(iii) Let be $Y = T^{-1}$. In this case $|J| = Y^{-2}$, $a_{y^{-1}}(\alpha,\beta) = -a_y(\alpha,\beta^{-1})$, and $|J|a_{y-1}'(\alpha,\beta) = a_y'(\alpha,\beta^{-1})$. Therefore

$$f_Y(y) = |J| f_T(y^{-1};\alpha,\beta,\delta,\lambda) = c_\delta a_y'(\alpha,\beta^{-1})\phi\left(|a_y(\alpha,\beta^{-1})| + \delta\right)\Phi\left(-\lambda a_y(\alpha,\beta^{-1})\right),$$

i.e., $Y = T^{-1} \sim FBS(\alpha,\beta^{-1},\delta,-\lambda)$.

□

References

1. Birnbaum, Z.W. Effect of linear truncation on a multinormal population. *Ann. Math. Statist.* **1950**, *21*, 272–279. [CrossRef]
2. Gómez, H.W.; Elal-Olivero, D.; Salinas, H.S.; Bolfarine, H. Bimodal Extension Based on the Skew-normal Distribution with Application to Pollen Data. *Environmetrics* **2009**, *22*, 50–62. [CrossRef]
3. Lehmann, E.L. The power of rank tests. *Ann. Mathematical Stat.* **1953**, *24*, 23–43. [CrossRef]
4. Roberts, C. A Correlation Model Useful in the Study of Twins. *J. Am. Stat. Assoc.* **1966**, *61*, 1184–1190. [CrossRef]
5. O'Hagan, A.; Leonard, T. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **1976**, *63*, 201–203. [CrossRef]
6. Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
7. Arnold, B.C.; Gómez, H.W.; Salinas, H.S. On multiple contraint skewed models'. *Statistics* **2009**, *43*, 279–293. [CrossRef]
8. Bolfarine, H.; Martínez–Flórez, G.; Salinas, H.S. Bimodal symmetric-asymmetric families. *Commun. Stat. Theory Methods* **2018**, *47(2)*, 259–276. [CrossRef]
9. Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [CrossRef]
10. Pewsey, A.; Gómez, H.W.; Bolfarine, H. Likelihood-based inference for power distributions. *Test* **2012**, *21*, 775–789. [CrossRef]
11. Chan, J.; Choy, S.; Makov, U.; Landsman, Z. Modelling insurance losses using contaminated generalised beta type-II distribution. *ASTIN Bull.* **2018**, *48*, 871–904.
12. Birnbaum, Z.W.; Saunders, S.C. Estimation for a family of life distributions with applications to fatigue. *J. Appl. Probab.* **1969**, *6*, 328–347. [CrossRef]

13. Birnbaum, Z.W.; Saunders, S.C. A New Family of Life Distributions. *J. Appl. Probab.* **1969**, *6*, 319–327. [CrossRef]
14. Díaz-García, J.A.; Leiva-Sánchez, V. A new family of life distributions based on the elliptically contoured distributions. *J. Statist. Plann. Inference* **2005**, *128*, 445–457. [CrossRef]
15. Vilca-Labra, F.; Leiva-Sanchez, V. A new fatigue life model based on the family of skew-elliptical distributions. *Commun. Stat. Theory Methods* **2006**, *35*, 229–244. [CrossRef]
16. Leiva, V.; Vilca-Labra, F.; Balakrishnan, N.; Sanhueza, A. A skewed sinh-normal distribution and its properties and application to air pollution. *Commun. Stat. Theory Methods* **2010**, *39*, 426–443. [CrossRef]
17. Castillo, N.O.; Gómez, H.W.; Bolfarine, H. Epsilon Birnbaum-Saunders distribution family: Properties and inference. *Stat. Pap.* **2011**, *52*, 871–883. [CrossRef]
18. Gómez, H.W.; Olivares, J.; Bolfarine, H. An extension of the generalized Birnbaum- Saunders distribution. *Stat. Probab. Lett.* **2009**, *79*, 331–338. [CrossRef]
19. Reyes, J.; Barranco-Chamorro, I.; Gallardo, D.I.; Gómez, H.W. Generalized Modified Slash Birnbaum-Saunders Distribution. *Symmetry* **2018**, *10*, 724. [CrossRef]
20. Reyes, J.; Barranco-Chamorro, I.; Gómez, H.W. Generalized modified slash distribution with applications. *Commun. Stat. Theory Methods* **2019**, doi:10.1080/03610926.2019.1568484. [CrossRef]
21. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. An alpha-power extension for the Birnbaum-Saunders distribution. *Statistics* **2014**, *48*, 896–912. [CrossRef]
22. Olmos, N.M.; Martínez-Flórez, G.; Bolfarine, H. Bimodal Birnbaum-Saunders Distribution with Application to Corrosion Data. *Commun. Stat. Theory Methods* **2017**, *46*, 6240-6257. [CrossRef]
23. Shakhatreh, M.; Lemonte, A.; Moreno-Arenas, G. The log-normal modified Weibull distribution and its reliability implications. *Reliab. Eng. Syst. Saf.* **2019**, *188* , 6–22. [CrossRef]
24. Prataviera, F.; Ortega, E.M.M.; Cordeiro, G.M.; Pescim, R.R. A new generalized odd log-logistic flexible Weibull regression model with applications in repairable systems. *Reliab. Eng. Syst. Saf.* **2018**, *176*, 13–26. [CrossRef]
25. Ng, H.K.T.; Kundu, D.; Balakrishnan, N. Modified moment estimation for the two-parameter Birnbaum-Saunders distribution. *Comput. Statist. Data Anal.* **2003**, *43*, 283–298. [CrossRef]
26. Barros, M.; Paula, G.A.; Leiva, V. A new class of survival regression models with heavy-tailed errors: Robustness and diagnostics. *Lifetime Data Anal.* **2008**, *14*, 316–332. [CrossRef]
27. Lemonte, A.; Cribari-Neto, F.; Vasconcellos, K. Improved statistical inference for the two-parameter Birnbaum-saunders distribution. *Comput. Stat. Data Anal.* **2007**, *51*, 4656–4681. [CrossRef]
28. R Development Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.
29. Gradshteyn, I.S.; Ryzhik, I.M. *Table of Integrals, Series, and Products*; Academic Press: New York, NY, USA, 2007.
30. Nadarajah, S. A truncated inverted beta distribution with application to air pollution data. *Stoch. Environ. Res. Risk Assess.* **2008**, *22*, 285–289. [CrossRef]
31. Akaike, H. A new look at statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
32. Gokhale, S.; Khare, M. Statistical behavior of carbon monoxide from vehicular exhausts in urban environments. *Environ. Model. Softw.* **2007**, *22*, 526–535. [CrossRef]