Birnbaum-Saunders Distribution: A Review of Models, Analysis and Applications

N. Balakrishnan*and Debasis Kundu†

Abstract

Birnbaum and Saunders [43, 44] introduced a two-parameter lifetime distribution to model fatigue life of a metal, subject to cyclic stress. Since then, extensive work has been done on this model providing different interpretations, constructions, generalizations, inferential methods, and extensions to bivariate, multivariate and matrix-variate cases. More than two hundred papers and one research monograph have already appeared describing all these aspects and developments. In this paper, we provide a detailed review of all these developments and at the same time indicate several open problems that could be considered for further research.

Key Words and Phrases: Bayes estimators; EM algorithm; Fisher information matrix; hazard function; length-biased distribution; moments and inverse moments; probability density function; stochastic orderings; TP_2 property.

1 Introduction

Among all statistical distributions, the normal distribution is undoubtedly the most used one in practice. Several new distributions have been developed by employing some transformations on the normal distribution. Two-parameter Birnbaum-Saunders (BS) distribution is one such distribution which has been developed by making a monotone transformation

*Department of Mathematics and Statistic, McMaster University, Canada. Corresponding author.
†Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India.

on the standard normal random variable. The BS distribution has appeared in several different contexts, with varying derivations. It was given by Fletcher [83], and according to Schröedinger [245], it was originally obtained by Konstantinowsky [114]. Subsequently, it was obtained by Freudenthal and Shinozuka [86] as a model that will be useful in life-testing. But, it was the derivation of Birnbaum and Saunders [43] that brought the usefulness of this distribution into a clear focus. Birnbaum and Saunders [43] introduced the distribution, that has come to bear their names, specifically for the purpose of modeling fatigue life of metals subject to periodic stress; consequently, the distribution is also sometimes referred to as the fatigue-life distribution.

Since the re-introduction of the model in 1969, extensive work has been done on this specific distribution. Even though Birnbaum and Saunders [43] provided a natural physical justification of this model through fatigue failure caused under cyclic loading, Desmond [62] presented a more general derivation based on a biological model. He also strengthened the physical justification for the use of this distribution by relaxing some of the original assumptions made by Birnbaum and Saunders. Bhattacharyya and Fries [41] proved that a BS distribution can be obtained as an approximation of an inverse Gaussian (IG) distribution. Desmond [63] observed an interesting feature that a BS distribution can be viewed as an equal mixture of an IG distribution and its reciprocal. This mixing property becomes useful in deriving several properties of the BS distribution using well-known properties of the IG distribution, and it also becomes useful for estimation purposes as well. Balakrishnan et al. [18] used this mixture property effectively to define mixtures of BS distributions, and then establish various properties and inferential methods for such a mixture BS distribution.

Many different properties of BS distribution have been discussed by a number of authors. It has been observed that the probability density function (PDF) of the BS distribution is unimodal. The shape of the hazard function (HF) plays an important role in lifetime data analysis. In this regard, Mann et al. ([189], page 155) first conjectured that the HF is not an increasing function, but the average HF is nearly a non-decreasing function. Three decades later, Kundu et al. [122] and Bebbington et al. [38] proved formally that the HF of BS distribution is an unimodal function; see also Gupta and Akman [100] in this respect. In many real life situations, the HF may not be monotone and that it increases up to a point and then decreases. For example, in the study of recovery from the breast cancer, it has been observed by Langlands et al. [124] that the maximum mortality occurs about three years after the disease and then it decreases slowly over a fixed period of time. In a situation like this, BS distribution can be used quite effectively for modeling. Moreover, in this aspect, log-normal and BS distributions behave very similarly. Although the HF of the log-normal distribution tends to zero, as $t \to \infty$, the hazard function of the BS distribution tends to a constant. Hence, in the long run, the BS behaves like an exponential distribution. Cheng and Tang [50, 51] developed several reliability bounds and tolerance limits for BS distribution. Cheng and Tang [52] also developed a random number generator for the BS distribution, and a comparison of the performance of different BS generators can be found in Rieck [224].

The maximum likelihood (ML) estimators of the shape and scale parameters based on a complete sample were discussed originally by Birnbaum and Saunders [44], and their asymptotic distributions were obtained by Engelhardt et al. [75]. The existence and uniqueness of the ML estimators have been formally established by Balakrishnan and Zhu [27] who have shown that the ML estimators of the unknown parameters cannot be obtained in an explicit form and need to be obtained by solving a non-linear equation. Extensive work has been done on developing the point and interval estimation of the parameters in the case of complete as well as censored samples. Rieck [222] first developed the estimators of the unknown parameters in the case of a censored sample. Ng et al. [195] provided modified moment (MM) estimators of the parameters in the case of a complete sample which are in explicit simple form, and they later extended to the case of a Type-II censored sample; see Ng et al. [196]. From and Li [87] suggested four different estimation techniques for both complete and censored samples. Several other estimation procedures for complete or censored samples, mainly from the frequentist point of view, have been proposed by Dupuis and Mills [73], Wang et al. [263], Lemonte et al. [171, 173], Desmond et al. [65], Ahmed et al. [3], Balakrishnan et al. [20], and Balakrishnan and Zhu [28]. Padgett [205] first considered Bayesian inference for the scale parameter β of the model assuming the shape parameter α to be known, with the use of a non-informative prior. Achcar [1] considered the same problem when both parameters are unknown and Achcar and Moala [2] addressed the problem in the presence of censored data and covariates. Recently, Wang et al. [267] provided Bayes estimates and associated credible intervals for the parameters of BS distribution under a general set of priors.

Several other models associated with BS distribution, and their properties and statistical inferential methods, have been discussed in the literature. For example, Rieck and Nedelman [226] considered the log-linear model for the BS distribution. Desmond et al. [65] considered the BS regression model. The logarithmic version of BS (Log-BS) regression model has been discussed by Galea et al. [88] and Leiva et al. [131]. Lemonte and Cordeiro [161] considered the BS non-linear regression model. A length-biased version of the BS distribution and its different applications have been provided by Leiva et al. [142]. Balakrishnan et al. [18] considered three different mixture models associated with BS distribution and discussed their fitting methods and applications. Santos-Neto et al. [236] proposed a reparameterized BS regression model with varying precision and discussed associated inferential issues. Recently, Bourguingon et al. [45] and Desousa et al. [67] proposed transmuted BS and tobit-BS models, respectively, and discussed several inferential issues for these models. For a review of the BS model, we refer the readers to the monograph by Leiva [129]; see also Johnson et al. [109], Leiva and Saunders [150] and Leiva and Vivanco [154] in this regard.

Díaz-García and Leiva [69] introduced the generalized BS distribution by replacing the normal kernel in (1) with elliptically symmetric kernels. Since then, quite a bit of work has been done on this generalized BS distribution. Leiva et al. [144] developed a procedure for generating random samples from the generalized BS distribution, while Leiva et al. [139] discussed lifetime analysis based on generalized BS distribution. A comprehensive discussion on generalized BS distributions has been provided by Sanhueza et al. [231]. Kundu et al. [119] introduced the bivariate BS distribution and studied some of its properties and characteristics. The multivariate generalized BS distribution has been introduced, by replacing the normal kernel by an elliptically symmetric kernel, by Kundu et al. [119]. Caro-Lopera et al. [48] developed subsequently the generalized matrix-variate BS distribution and discussed its properties.

The main aim of this paper is to provide a review of all these developments on BS distribution and also to suggest some open problems along the way. The rest of this paper is organized as follows. In Section 2, we define the BS distribution and provide physical interpretations and some basic properties of the model. Point and interval estimation methods based on complete samples are detailed in Sections 3 and 4, respectively. Bayesian inference is discussed next in Section 5. In Section 6, we discuss the point and interval estimation of the model parameters based on censored samples. Some of the other univariate issues are discussed in Section 7. Bivariate and multivariate BS distributions are described in Sections 8 and 9., respectively. In Sections 10 and 11, we describe several related models and the BS regression model, respectively. Different generalizations are described in Section 12, and several illustrative examples are presented in Section 13. Finally, some concluding remarks are made in Section 14.

2 Definition & Some Basic Properties

In this section, we provide the definition, give some physical interpretations and discuss several properties of the two-parameter BS distribution.

2.1 Cumulative Distribution Function, Probability Density Function & Hazard Function

The cumulative distribution function (CDF) of a two-parameter BS random variable T can be written as

$$F_T(t;\alpha,\beta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{1/2} - \left(\frac{\beta}{t}\right)^{1/2}\right\}\right], \quad 0 < t < \infty, \quad \alpha, \beta > 0, \tag{1}$$

where $\Phi(\cdot)$ is the standard normal CDF. The parameters α and β in (1) are the shape and scale parameters, respectively.

If the random variable T has the BS distribution function in (1), then the corresponding probability density function (PDF) is

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

From here on, a random variable with PDF in (2) will be simply denoted by BS(α, β). It is clear that β is the scale parameter while α is the shape parameter. It can be seen that the PDF of BS(α, β) goes to 0 when $t \to 0$ as well as when $t \to \infty$. For all values of α and β, the PDF is unimodal. The plots of the PDF of BS(α, β) in (2), for different values of α, are presented in Figure 1.

The mode of BS(α, β) cannot be obtained in an explicit form. But, the mode of the PDF of BS$(\alpha, 1)$, say m_α, can be obtained as the solution of the following cubic equation [see

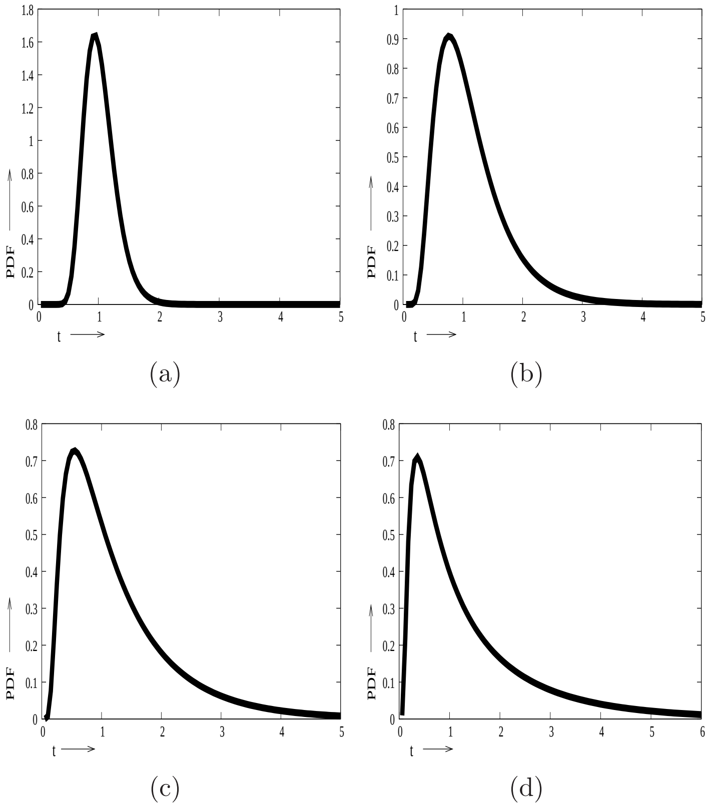

Figure 1: Plots of the PDF of BS(α, β) for different values of α and when $\beta = 1$: (a) $\alpha = 0.25$, (b) $\alpha = 0.5$ (c) $\alpha = 0.75$, (d) $\alpha = 1.0$.

Kundu et al. [122]]:

$$t^3 + t^2(\alpha^2 + 1) + t(3\alpha^2 - 1) - 1 = 0. \tag{3}$$

It can be easily shown that the cubic equation in (10) has a unique solution. The mode of BS(α, β) can then be obtained simply as $\beta\, m_\alpha$. It has been observed that m_α is an increasing function of α.

If $T \sim$ BS(α, β), consider the following transformation

$$Z = \frac{1}{\alpha}\left[\left(\frac{T}{\beta}\right)^{1/2} - \left(\frac{T}{\beta}\right)^{-1/2}\right], \tag{4}$$

or equivalently

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

It then readily follows from (1) that Z is a normal random variable with mean zero and variance one. From (4), it is immediate that

$$W = Z^2 = \frac{1}{\alpha^2}\left[\frac{T}{\beta} + \frac{\beta}{T} - 2\right] \sim \chi^2_1. \tag{6}$$

Eq. (5) can be used effectively to generate BS random variables from standard normal random variables; see, for example, Rieck [224]. Alternatively, the relation in (6) can also be used to generate BS random variables; see Chang and Tang [52]. Rieck [224] showed the generator based on (5) to be more efficient than the one based on (6).

Evidently, the q-th quantile of BS(α, β) random variable is

$$\frac{\beta}{4}\left[\alpha z_q + \sqrt{(\alpha z_q)^2 + 4}\right]^2, \tag{7}$$

with $0 < q < 1$, where z_q is the q-th quantile of the standard normal random variable. From (4), it is also immediate that β is the median of T.

We now discuss the shape characteristics of the HF of a BS random variable. With $T \sim$ BS(α, β), the HF of T is given by

$$h_T(t; \alpha, \beta) = \frac{f_T(t; \alpha, \beta)}{1 - F_T(t; \alpha, \beta)}, \quad t > 0. \tag{8}$$

Since the shape of the HF does not depend on the scale parameter β, we may take $\beta = 1$ without loss of generality. In this case, the HF in (8) takes the form

$$h_T(t; \alpha, 1) = \frac{\frac{1}{\sqrt{2\pi}\alpha}\epsilon'(t)e^{-\frac{1}{2\alpha^2}\epsilon^2(t)}}{\Phi\left(-\frac{\epsilon(t)}{\alpha}\right)}, \tag{9}$$

where

$$\epsilon(t) = t^{1/2} - t^{-1/2}, \quad \epsilon'(t) = \frac{1}{2t}(t^{1/2} + t^{-1/2}) \text{ and } \epsilon''(t) = -\frac{1}{4t^2}(t^{1/2} + 3t^{-1/2}). \tag{10}$$

Kundu et al. [122] then showed that the HF in (9), and hence the one in (8), is always unimodal. The plots of the HF of BS(α, β) in (9) for different values of α, are presented in Figure 2. From (9), it can be shown that $\ln(h_T(t; \alpha, 1)) \to 1/(2\alpha^2)$ as $t \to \infty$.

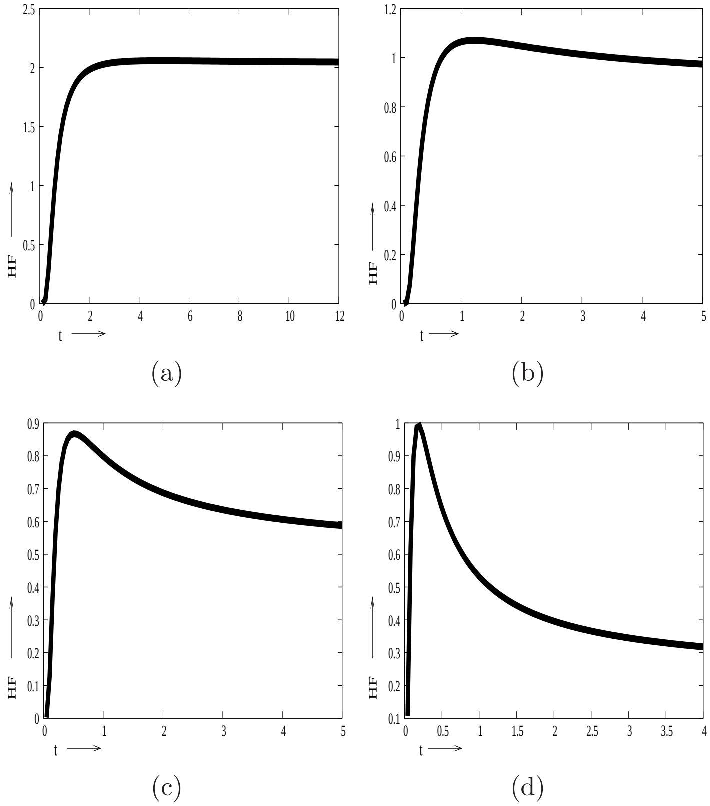

Figure 2: Plots of the HF of BS(α,β) for different values of α and when $\beta = 1$: (a) $\alpha = 0.50$, (b) $\alpha = 0.75$ (c) $\alpha = 1.0$, (d) $\alpha = 1.5$.

Kundu et al. [122] have shown that the change point, c_α, of the hazard function of BS$(\alpha, 1)$ in (9) can be obtained as a solution of the non-linear equation

$$\Phi\left(-\frac{1}{\alpha}\epsilon(t)\right)\left\{-(\epsilon'(t))^2\epsilon(t) + \alpha^2\epsilon''(t)\right\} + \alpha\phi\left(-\frac{1}{\alpha}\epsilon(t)\right)(\epsilon'(t))^2 = 0. \tag{11}$$

Since β is the scale parameter, the change point, $c_{\alpha,\beta}$, of the HF of BS(α,β) in (8) can be obtained immediately as $c_{\alpha,\beta} = \beta\ c_\alpha$. Values of c_α, for different choices of α, have been computed by Kundu et al. [122]. They have also observed that for $\alpha > 0.25$, the following approximation of c_α can be useful:

$$\widetilde{c_\alpha} = \frac{1}{(-0.4604 + 1.8417\alpha)^2}, \tag{12}$$

and that this approximation becomes quite accurate when $\alpha > 0.6$.

2.2 Physical Interpretations

It is well known that for the analysis of fatigue data, any of the two-parameter families of distributions, such as Weibull, log-normal or gamma, could be used to model effectively the region of central tendency. Moreover, since we usually have only small sample sizes in practice, none of them may get rejected by a goodness-of fit test, say Chi-square test or Kolmogorov-Smirnov test, for a given data. However, when the goal is to predict the safe life, say the one-thousandth percentile, we would expect to have a wide discrepancy in the results obtained from these models.

For this reason, Birnbaum and Saunders [43] proposed the fatigue failure life distribution based on a physical consideration of the fatigue process, rather than using an arbitrary parametric family of distributions. They obtained a two-parameter family of non-negative random variables as an idealization for the number of cycles necessary to force a fatigue crack to grow to a critical value, based on the following assumptions:

1. Fatigue failure is due to repeated application of a common cyclic stress pattern;

2. Due to the influence of this cyclic stress, a dominate crack in the material grows until a critical size w is reached at which point fatigue failure occurs;

3. The crack extension in each cycle is a random variable with the same mean and the same variance;

4. The crack extensions in cycles are statistically independent.

5. The total extension of the crack, after a large number of cycles, is approximately normally distributed, justified by the use of central limit theorem (CLT).

Specifically, at the j-th cycle, the crack extension X_j is assumed to be a random variable with mean μ_0 and variance σ_0^2. Then, due to CLT, $\sum_{i=1}^{n} X_i$ is approximately normally distributed with mean $n\mu_0$ and variance $n\sigma_0^2$, and consequently the probability that the crack does not exceed a critical length ω, say, is given by

$$\Phi\left(\frac{\omega - n\mu_0}{\sigma_0\sqrt{n}}\right) = \Phi\left(\frac{\omega}{\sigma_0\sqrt{n}} - \frac{\mu_0\sqrt{n}}{\sigma_0}\right). \tag{13}$$

It is further assumed that the failure occurs when the crack length exceeds ω. If T denotes the lifetime (in number of cycles) until failure, then the CDF of T is approximately

$$P(T \leq t) = F_T(t) \approx 1 - \Phi\left(\frac{\omega}{\sigma_0\sqrt{t}} - \frac{\mu_0\sqrt{t}}{\sigma_0}\right) = \Phi\left(\frac{\mu_0\sqrt{t}}{\sigma_0} - \frac{\omega}{\sigma_0\sqrt{t}}\right). \tag{14}$$

Note that in deriving (14), it is being assumed that the probability that X_j's take negative values is negligible. If (14) is assumed to be the exact lifetime model, then it is evident from (1) that T follows a BS distribution with CDF in (1), where

$$\beta = \frac{\omega}{\mu_0}, \qquad \text{and} \qquad \alpha = \frac{\sigma_0}{\sqrt{\omega\mu_0}}.$$

2.3 Moments and Inverse Moments

Using the relation in (5), and different moments of the standard normal random variable, moments of T, for integer r, can be obtained as follows:

$$\mathrm{E}(T^r) = \beta^r \sum_{j=0}^{r} \binom{2r}{2j} \sum_{i=0}^{j} \binom{i}{j} \frac{(2r-2j+2i)!}{2^{r-j+i}(r-j+i)!} \left(\frac{\alpha}{2}\right)^{2r-2j+2i}; \tag{15}$$

see Leiva et al. [142]. Rieck [223] also obtained $\mathrm{E}(T^r)$, for fractional values of r, in terms of Bessel function, from the moment generating function of $\mathrm{E}(\ln(T))$.

From (15), the mean and variance are obtained as

$$\mathrm{E}(T) = \frac{\beta}{2}(\alpha^2 + 2) \qquad \text{and} \qquad \mathrm{Var}(T) = \frac{\beta^2}{4}(5\alpha^4 + 4\alpha^2), \tag{16}$$

while the coefficients of variation (γ), skewness (δ) and kurtosis (κ) are found to be

$$\gamma = \frac{\sqrt{5\alpha^4 + 4\alpha^2}}{\alpha^2 + 2}, \quad \delta = \frac{44\alpha^3 + 24\alpha}{(5\alpha^2 + 4)^{3/2}} \quad \text{and} \quad \kappa = 3 + \frac{558\alpha^4 + 240\alpha^2}{(5\alpha^2 + 4)^2}, \tag{17}$$

respectively. It is clear that as $\alpha \to 0$, the coefficient of kurtosis approaches 3, and the behavior of the BS distribution appears to be approximately normal with mean and variance being approximately β and $\beta^2\alpha^2$, respectively. Hence, as $\alpha \to 0$, the BS distribution becomes degenerate at β. On the other hand, when $\alpha \to \infty$, both mean and variance diverge, while the coefficients of variation, skewness and kurtosis all converge to some fixed constants.

If $T \sim \text{BS}(\alpha, \beta)$, it can be easily observed from (2) that $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$. Therefore, for integer r, we readily obtain from (15) that

$$\text{E}(T^{-r}) = \beta^{-r} \sum_{j=0}^{r} \binom{2r}{2j} \sum_{i=0}^{j} \binom{i}{j} \frac{(2r - 2j + 2i)!}{2^{r-j+i}(r - j + i)!} \left(\frac{\alpha}{2}\right)^{2r-2j+2i}. \tag{18}$$

Also, the mean and variance of T^{-1} are obtained, from (16), to be

$$\text{E}(T^{-1}) = \frac{1}{2\beta}(\alpha^2 + 2) \quad \text{and} \quad \text{Var}(T^{-1}) = \frac{1}{4\beta^2}(5\alpha^4 + 4\alpha^2), \tag{19}$$

while the corresponding γ, δ and κ remain the same as presented in (17).

2.4 Relationships with Inverse Gaussian Distribution

A random variable X is said to have an inverse Gaussian (IG) distribution, with parameters $\mu > 0$ and $\lambda > 0$, if X has the PDF

$$f_X(t; \mu, \lambda) = \left(\frac{\lambda}{2\pi t^3}\right)^{1/2} \exp\left\{-\frac{\lambda}{2\mu^2 t}(t - \mu)^2\right\}, \quad t > 0. \tag{20}$$

An IG random variable, with PDF in (20), will be denoted by $\text{IG}(\mu, \lambda)$. The IG distribution has several interesting statistical and probabilistic properties, of which one of the most intriguing one is in the presence of χ^2 and F distributions in associated inferential methods.

Book length accounts of IG distribution can be found in Chhikara and Folks [54] and Seshadri [241, 242].

Desmond [63] first observed a relationship between BS and IG distributions. Consider two random variables X_1 and X_2 with $X_1 \sim \mathrm{IG}(\mu, \lambda)$ and $X_2^{-1} \sim \mathrm{IG}(\mu^{-1}, \lambda\mu^2)$. Now, let us define a random variable T as

$$T = \begin{cases} X_1 & \text{with probability} \quad 1/2, \\ X_2 & \text{with probability} \quad 1/2. \end{cases} \tag{21}$$

Then, the PDF of T becomes

$$f_T(t; \mu, \lambda) = \frac{1}{2} f_{X_1}(t; \mu, \lambda) + \frac{1}{2} f_{X_2}(t; \mu, \lambda), \tag{22}$$

where $f_{X_1}(t; \mu, \lambda)$ has the same form as in (20) and $f_{X_2}(t; \mu, \lambda) = t f_{X_1}(t; \mu, \lambda)/\mu$. A simple algebraic calculation shows that the random variable T in (21) has $\mathrm{BS}(\alpha, \beta)$, where $\alpha = \sqrt{\mu/\lambda}$ and $\beta = \mu$. Using the mixture representation in (22), many properties of the IG distribution can be readily transformed to those of the BS distribution. For example, by using the moment generating functions (MGF) of the IG distribution and the length-biased IG distribution, the MGFs of the BS distribution can be obtained easily. Specifically, if $T \sim \mathrm{BS}(\alpha, \beta)$, then the MGF of T is

$$M_T(t) = \frac{1}{2} \left[M_{X_1}(t) + M_{X_2}(t) \right],$$

where $M_{X_1}(t)$ and $M_{X_2}(t)$ are the MGFs of X_1 and X_2, respectively. Since it is known that [see Jorgensen et al. [111]]

$$\begin{aligned} M_{X_1}(t) &= \exp\left\{ (\lambda/\mu - ((\lambda/\mu)^2 - 2\lambda t)^{1/2}) \right\}, \\ M_{X_2}(t) &= \exp\left\{ (\lambda/\mu - ((\lambda/\mu)^2 - 2\lambda t)^{1/2}) \right\}, (1 - 2\mu^2 t/\lambda)^{-1/2} \end{aligned}$$

we readily find

$$M_T(t) = \frac{1}{2} \exp\left\{ (\lambda/\mu - ((\lambda/\mu)^2 - 2\lambda t)^{1/2}) \right\} (1 + (1 - 2\mu^2 t/\lambda)^{-1/2}).$$

This mixture representation becomes useful while estimating the parameters of BS(α, β) distribution through an EM algorithm, and also in estimating the parameters of a mixture of BS distributions. These details are presented in later sections.

3 Point Estimation: Complete Sample

In this section, we discuss different methods of estimation of the parameters α and β based on a random sample $\{T_1, \cdots, T_n\}$ with observations (data) $\{t_1, \cdots, t_n\}$ of size n from BS(α, β).

3.1 ML Estimators

Birnbaum and Saunders [44] considered the ML estimators of parameters α and β based on $\{t_1, \cdots, t_n\}$. The log-likelihood function, without the additive constant, is given by

$$l(\alpha, \beta|\text{data}) = -n\ln(\alpha) - n\ln(\beta) + \sum_{i=1}^{n} \ln\left[\left(\frac{\beta}{t_i}\right)^{1/2} + \left(\frac{\beta}{t_i}\right)^{3/2}\right] - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right). \tag{23}$$

Based on the observed sample, let us define the sample arithmetic and harmonic means as

$$s = \frac{1}{n}\sum_{i=1}^{n} t_i \quad \text{and} \quad r = \left[\frac{1}{n}\sum_{i=1}^{n} t_i^{-1}\right]^{-1},$$

respectively. Differentiating (23) with respect to α and equating it to zero, we obtain

$$\alpha^2 = \left[\frac{s}{\beta} + \frac{\beta}{r} - 2\right]. \tag{24}$$

Next, differentiating (23) with respect to β and equating it to zero and after substituting α^2 from (24), the following non-linear equation is obtained:

$$\beta^2 - \beta(2r + K(\beta)) + r(s + K(\beta)) = 0, \tag{25}$$

where

$$K(x) = \left[\frac{1}{n}\sum_{i=1}^{n}(x + t_i)^{-1}\right]^{-1} \quad \text{for } x \geq 0.$$

The ML estimator of β, say $\widehat{\beta}$, is then the positive root of (25). Birnbaum and Saunders [44] showed that $\widehat{\beta}$ is the unique positive root of (25) and furthermore, $r < \widehat{\beta} < s$. A numerical iterative procedure is needed to solve the non-linear equation in (25). In their work, Birnbaum and Saunders [44] proposed two different iterative methods to compute $\widehat{\beta}$ and showed that with any arbitrary initial guess value between r and s, both methods do converge to $\widehat{\beta}$. Once the ML estimator of β is obtained, the ML estimator of α can then be obtained as

$$\widehat{\alpha} = \left[\frac{s}{\widehat{\beta}} + \frac{\widehat{\beta}}{r} - 2 \right]^{1/2}. \tag{26}$$

Birnbaum and Saunders [44] also proved that the ML estimators are consistent estimates of the parameters. Recently, Balakrishnan and Zhu [27] showed that if $n = 1$, the ML estimators of α and β do not exist, but when $n > 1$, the ML estimators always exist and are unique.

Engelhardt et al. [75] showed that it is a regular family of distributions and that the Fisher information matrix is given by

$$\boldsymbol{I}(\alpha, \beta) = - \begin{bmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & n\left[\frac{1}{\alpha^2\beta^2} + E\left(\frac{1}{(T+\beta)^2}\right)\right] \end{bmatrix} = - \begin{bmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & \frac{n}{\alpha^2\beta^2}(1 + \alpha(2\pi)^{-1/2}h(\alpha)), \end{bmatrix}, \tag{27}$$

where

$$h(\alpha) = \alpha\sqrt{\pi/2} - \pi e^{2/\alpha^2}[1 - \Phi(2/\alpha)]. \tag{28}$$

By simple calculations, it has been shown by Engelhardt et al. [75] that the joint distribution of $\widehat{\alpha}$ and $\widehat{\beta}$ is bivariate normal, i.e.,

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} \sim \mathrm{N}\left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \frac{\alpha^2}{2n} & 0 \\ 0 & \frac{\beta^2}{n(0.25+\alpha^{-2}+I(\alpha))} \end{pmatrix} \right], \tag{29}$$

where

$$I(\alpha) = 2\int_0^\infty \left\{(1 + g(\alpha x))^{-1} - 1/2\right\}^2 \mathrm{d}\Phi(x) \tag{30}$$

with

$$g(y) = 1 + \frac{y^2}{2} + y\left(1 + \frac{y^2}{4}\right)^{1/2}.$$

It is of interest to note that $\widehat{\alpha}$ and $\widehat{\beta}$ are asymptotically independent. An asymptotic confidence interval for α can be easily obtained from (29), and an asymptotic confidence interval for β, for a given α, can also be obtained from (29).

3.2 Moment and Modified Moment Estimators

The moment estimators of α and β can be obtained by equating the sample mean and sample variance, respectively, with $\mathrm{E}(T)$ and $\mathrm{Var}(T)$, provided in (16). Thus, if s and v denote the sample mean and sample variance, respectively, then the moment estimators of α and β are obtained by solving

$$s = \frac{\beta}{2}(\alpha^2 + 2), \quad v = \frac{\beta^2}{4}(5\alpha^4 + 4\alpha^2). \tag{31}$$

It can be easily seen that the moment estimator of α can be obtained as the root of the non-linear equation

$$\alpha^4(5s^2 - v) + 4\alpha^2(s^2 - 4) - 4v = 0. \tag{32}$$

It is clear that if the sample coefficient of variation is less than $\sqrt{5}$, then the moment estimator of α is

$$\widehat{\widehat{\alpha}} = \left[\frac{-2(s^2-4) + 2\sqrt{(s^2-4)^2 + v(5s^2 - v)}}{(5s^2 - v)}\right]^{1/2} \tag{33}$$

and the moment estimator of β is then

$$\widehat{\widehat{\beta}} = \frac{2s}{\widehat{\widehat{\alpha}}^2 + 2}. \tag{34}$$

However, if the sample coefficient of variation is greater than $\sqrt{5}$, then the moment estimator of α may not exist.

Since the ML estimators cannot be obtained in explicit form and that the moment estimators may not always exist, Ng et al. [195] suggested MM estimators of the parameters α and β by utilizing the fact that if $T \sim \mathrm{BS}(\alpha, \beta)$, then $T^{-1} \sim \mathrm{BS}(\alpha, \beta^{-1})$. Therefore, equating the sample arithmetic and harmonic means, s and r, respectively, with the corresponding population versions, we obtain

$$s = \beta\left(1 + \frac{1}{2}\alpha^2\right) \quad \text{and} \quad r^{-1} = \beta^{-1}\left(1 + \frac{1}{2}\alpha^2\right). \tag{35}$$

Thus, the MM estimators of α and β are obtained as

$$\widetilde{\alpha} = \left\{2\left[\left(\frac{s}{r}\right)^{1/2} - 1\right]\right\}^{1/2} \quad \text{and} \quad \widetilde{\beta} = (sr)^{1/2}. \tag{36}$$

The interesting point about the MM estimators is that they always exist unlike the moment estimators, and moreover they have simple explicit forms. Using the CLT on $(1/n)\sum_{i=1}^{n} T_i$ and $(1/n)\sum_{i=1}^{n} T_i^{-1}$, when T_i's are independent and identically distributed (*i.i.d.*) BS(α, β), Ng et al. [195] proved that the joint asymptotic distribution of $\widetilde{\alpha}$ and $\widetilde{\beta}$ is bivariate normal,i.e.,

$$\begin{pmatrix} \widetilde{\alpha} \\ \widetilde{\beta} \end{pmatrix} \sim \text{N}\left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \frac{\alpha^2}{2n} & 0 \\ 0 & \frac{\alpha\beta^2}{n}\left(\frac{1+\frac{3}{4}\alpha^2}{(1+\frac{1}{2}\alpha^2)^2}\right) \end{pmatrix}\right]. \tag{37}$$

Interestingly, in this case also, we observe that the MM estimators are asymptotically independent.

Ng et al. [195] performed extensive Monte Carlo simulations to compare the performances of the ML estimators and MM estimators for different sample sizes and for different choices of parameter values. They then observed that the performances of ML estimators and MM estimators are almost identical for any sample size when the shape parameter α is not large (say, < 0.5). For small sample sizes, both estimators are highly biased if α is large. It has been observed that

$$\text{Bias}(\widehat{\alpha}) \approx \text{Bias}(\widetilde{\alpha}) \approx -\frac{\alpha}{n} \quad \text{and} \quad \text{Bias}(\widehat{\beta}) \approx \text{Bias}(\widetilde{\beta}) \approx -\frac{\alpha^2}{4n}. \tag{38}$$

So, the bias-corrected ML estimators and MM estimators perform quite well. Alternatively, jackknifing [see Efron [74]] can also be performed to determine bias-corrected ML estimators and MM estimators, and these estimators have also been observed to perform quite well.

3.3 From and Li Estimators

From and Li [87] proposed four other estimators of α and β, which are as follows.

From and Li Estimator-1: If $T \sim \mathrm{BS}(\alpha,\beta)$, then Z defined in (4) follows the standard normal distribution and based on this fact, an estimator of (α,β) can be obtained by solving the following two equations:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\alpha}\left[\left(\frac{t_i}{\beta}\right)^{1/2}-\left(\frac{t_i}{\beta}\right)^{-1/2}\right]=0 \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\alpha^2}\left[\left(\frac{t_i}{\beta}\right)+\left(\frac{t_i}{\beta}\right)^{-1}-2\right]=1, \qquad (39)$$

giving rise to the estimators

$$\widehat{\beta}_{FL,1}=\frac{\sum_{i=1}^{n} t_i^{1/2}}{\sum_{i=1}^{n} t_i^{-1/2}} \quad \text{and} \quad \widehat{\alpha}_{FL,1}=\left(\sum_{i=1}^{n}\frac{1}{n}\left[\left(\frac{t_i}{\widehat{\beta}_{FL,1}}\right)+\left(\frac{t_i}{\widehat{\beta}_{FL,1}}\right)^{-1}-2\right]\right)^{1/2}.$$

These can be seen to be variations of the moment estimators.

From and Li Estimator-2: Since β is the median of T irrespective of α, From and Li [87] obtained estimators by equating the sample median and sample variance to their corresponding population versions. Thus, the estimator of β in this case is simply

$$\widehat{\beta}_{FL,2}=\mathrm{median}\{t_1,\cdots,t_n\},$$

while an estimator of α can be obtained from the equation

$$v=\frac{\widehat{\beta}_{FL,2}}{4}\left(5\alpha^4+4\alpha^2\right),$$

for which the solution is

$$\widehat{\alpha}_{FL,2}=\sqrt{\frac{-2+2\sqrt{1+5v/\widehat{\beta}_{FL,2}}}{5}}.$$

Here, v is the sample variance as before.

From and Li Estimator-3: If we denote $t_{1:n}<\cdots<t_{n:n}$ as the ordered sample of $t_1,\cdots,t_n$, then take $\widehat{\beta}_{FL,3}$ same as $\widehat{\beta}_{FL,2}$, and solve for α from

$$F_T(t_{i:n};\alpha,\widehat{\beta}_{FL,3})=\frac{i}{n+1}, \quad i=1,\cdots,n,$$

i.e.,

$$\widehat{\alpha}(i) = \frac{\epsilon(t_{i:n}/\widehat{\beta}_{FL,3})}{\Phi^{-1}(i/(n+1))}, \quad i = 1, \cdots, n.$$

Here, $\epsilon(\cdot)$ is same as defined earlier in (10). Finally, an estimator of α is proposed as

$$\widehat{\alpha}_{FL,3} = \text{median}\{\widehat{\alpha}(1), \cdots, \widehat{\alpha}(n)\}.$$

From & Li Estimator-4: Instead of using the whole data, From and Li [87] proposed the following estimators of α and β using only the middle portion of the data, which produces more robust estimators. For $1 \leq n_1 < n_2 \leq n$, these estimators are as follows:

$$\widehat{\beta}_{FL,4} = \frac{\sum_{i=n_1}^{n_2} t_{i:n}}{\sum_{i=n_1}^{n_2} (1/\sqrt{t_{i:n}})} \quad \text{and} \quad \widehat{\alpha}_{FL,4} = \sqrt{\frac{\sum_{i=n_1}^{n_2} \epsilon^2(t_{i:n}/\widehat{\beta}_{FL,4})}{\sum_{i=n_1}^{n_2} \left(\Phi^{-1}(i/(n+1))\right)^2}}; \tag{40}$$

here, n_1 and n_2 are chosen such that $n_1/n < 0.5$ and $n_2/n > 0.5$. Since the estimators in (40) use only the middle portion of the data, it is expected that the estimators will not be affected by the presence of outliers, but a loss in efficiency is to be expected in the case when there are no outliers in the data.

From and Li [87] performed extensive Monte Carlo simulation experiments to compare the performance of ML estimators with those of their four estimators. They observed that among their estimators, From and Li Method-1 performed nearly the same as the ML estimators if the model is true and that there are no outliers in the data. If some outliers are present, then the ML estimators do not perform well, but From and Li Method-4 performed the best in this case.

3.4 Balakrishnan and Zhu Estimator

Balakrishnan and Zhu [28] developed their estimators of α and β based on some key properties of the BS distribution. Suppose $T_1, \ldots, T_n$ are *i.i.d.* BS(α, β) random variables. Then, consider the following $\binom{n}{2}$ pairs of random variables (Z_{ij}, Z_{ji}), where

$$Z_{ij} = \frac{T_i}{T_j} \quad \text{for} \quad 1 \leq i \neq j \leq n.$$

Now, observe that

$$\mathrm{E}(Z_{ij}) = \mathrm{E}\left(\frac{T_i}{T_j}\right) = \mathrm{E}(T_i)\mathrm{E}\left(\frac{1}{T_j}\right) = \left(1 + \frac{\alpha^2}{2}\right)^2.$$

Upon equating the sample mean of Z_{ij} with its population mean, we obtain the equation

$$\bar{z} = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} z_{ij} = \left(1 + \frac{\alpha^2}{2}\right)^2. \tag{41}$$

Here, z_{ij} denotes the sample value of Z_{ij}. Upon solving (41), an estimator of α is obtained as

$$\widehat{\alpha}_{BZ} = \left(2(\sqrt{\bar{z}} - 1)\right)^{1/2}. \tag{42}$$

Balakrishnan and Zhu [28] then proposed two different estimators of β based on $\widehat{\alpha}_{BZ}$ in (42), as follows. Since

$$\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} T_i\right) = \beta\left(1 + \frac{1}{2}\alpha^2\right),$$

an estimator β can be obtained immediately as

$$\widehat{\beta}_{BZ,1} = \frac{2s}{(\widehat{\alpha}_{BZ})^2 + 2} = \frac{s}{\sqrt{\bar{z}}};$$

here, s is the sample mean of the observed $t_1, \ldots, t_n$. Moreover, based on the fact that

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(\frac{1}{T_i}\right) = \frac{1}{\beta}\left(1 + \frac{\alpha^2}{2}\right),$$

Balakrishnan and Zhu [28] proposed another estimator of β as

$$\widehat{\beta}_{BZ,2} = r\left(1 + \frac{1}{2}(\widehat{\alpha}_{BZ})^2\right) = r\sqrt{\bar{z}};$$

here, r is the harmonic mean of $t_1, \ldots, t_n$, as defined before.

Theoretically, it has been shown by Balakrishnan and Zhu [28] that $\widehat{\alpha}_{BZ}$, $\widehat{\beta}_{BZ,1}$ and $\widehat{\beta}_{BZ,2}$ always exist and that $\widehat{\alpha}_{BZ}$ is negatively biased. Based on extensive simulations, they have observed that the performances of their estimators are very similar to those of the ML estimators and the MM estimators in terms of bias and MSE.

3.5 New Estimators

We may propose the following estimators which are in explicit form. First, we may take $\widehat{\beta}_{\text{new}} = \text{median}\{t_1, \cdots, t_n\}$. Then, we consider the new transformed variables

$$u_i = \left(\frac{t_i}{\widehat{\beta}_{\text{new}}}\right)^{1/2} - \left(\frac{\widehat{\beta}_{\text{new}}}{t_i}\right)^{1/2}, \quad i = 1, \cdots, n.$$

It is known that if $T \sim \text{BS}(\alpha, \beta)$, then $\left\{(T/\beta)^{1/2} - (\beta/T)^{1/2}\right\} \sim \text{N}(0, \alpha^2)$, and using this fact, we may propose an estimate of α as

$$\widehat{\alpha}_{\text{new}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} u_i^2}.$$

Its properties and relative performance may need further investigation.

3.6 Robust Estimators

Dupuis and Mills [73] proposed a robust estimation procedure for α and β by employing the influence function (IF) approach of Hampel [102]. It has been observed by Hampel [102] that the robustness of an estimator can be measured by means of its IF. It is well known that [see Hampel [102]] the IF of an ML estimator is proportional to its score function, which for the BS model is given by

$$\underline{s}(t; \alpha, \beta) = \begin{bmatrix} \frac{\partial \ln(f_T(t;\alpha,\beta))}{\partial \alpha} \\ \frac{\partial \ln(f_T(t;\alpha,\beta))}{\partial \beta} \end{bmatrix}, \tag{43}$$

where

$$\frac{\partial \ln(f_T(t;\alpha,\beta))}{\partial \alpha} = \frac{-\alpha^2\beta t + t^2 - 2\beta t + \beta^2}{\alpha^3\beta t}, \tag{44}$$

$$\frac{\partial \ln(f_T(t;\alpha,\beta))}{\partial \beta} = \frac{-\beta\alpha^2 t^2 + \beta^2\alpha^2 t + t^3 - \beta^2 t + \beta t^2 - \beta^3}{2\beta^2\alpha^2 t(t+\beta)}. \tag{45}$$

Clearly, the score functions are unbounded in t, which implies that the corresponding IF is also unbounded in t. For this reason, the ML estimators become biased and inefficient when the model is incorrect.

Dupuis and Mills [73], therefore, proposed the following optimal biased robust estimators (OBREs) of α and β. The OBREs of α and β, say $\widehat{\alpha}_R$ and $\widehat{\beta}_R$, respectively, are the solutions of

$$\sum_{i=1}^{n} \underline{\Psi}(t_i;\alpha,\beta) = 0 \tag{46}$$

for some function $\underline{\Psi} : \mathbb{R}_+^3 \to \mathbb{R}_+^2$. Moreover, the IF function of an M-estimator $\underline{\Psi}$, at $F_T(t;\alpha,\beta)$, is given by

$$\begin{aligned} IF(t;\underline{\Psi},F_T) &= [\boldsymbol{M}(\underline{\Psi},F_T)]^{-1}\underline{\Psi}(t;\alpha,\beta), \\ \boldsymbol{M}(\underline{\Psi},F_T) &= -\int_0^\infty \underline{\Psi}(t;\alpha,\beta)[\underline{\Psi}(t;\alpha,\beta)]^\top \mathrm{d}F_T(t;\alpha,\beta). \end{aligned}$$

In any robust M-estimation procedure, $\underline{\Psi}(\cdot)$ needs to be chosen properly. There are several versions of OBREs that are available in the literature. Dupuis and Mills [73] proposed the standardized OBREs. They can be defined as follows for a given bound c on the IF. The function $\underline{\Psi}(\cdot)$ is defined implicitly by

$$\sum_{i=1}^{n} \underline{\Psi}(t_i;\alpha,\beta) = \sum_{i=1}^{n} \left\{ \underline{s}(t_i;\alpha,\beta) - \underline{a}(\alpha,\beta) W_c(t_i;\alpha,\beta) \right\},$$

where

$$W_c(t;\theta) = \min\left[1, \frac{c}{||\boldsymbol{A}(\alpha,\beta)[\underline{s}(t;\alpha,\beta) - \underline{a}(\alpha,\beta)]||}\right]$$

with $||\cdot||$ being the Euclidean norm. The 2×2 matrix $\boldsymbol{A}(\alpha,\beta)$ and the 2×1 vector $\underline{a}(\alpha,\beta)$ are defined implicitly as

$$\begin{aligned}\mathrm{E}\left[\underline{\Psi}(t;\alpha,\beta)[\underline{\Psi}(t;\alpha,\beta)]^\top\right] &= \{\boldsymbol{A}(\alpha,\beta)^\top\boldsymbol{A}(\alpha,\beta)\}^{-1},\\ E\left[\underline{\Psi}(t;\alpha,\beta)\right] &= \underline{0}.\end{aligned}$$

The following algorithm can be used to compute the OBREs, as suggested by Dupuis and Mills [73]:

Algorithm

Step 1: Fix an initial value of (α,β);

Step 2: Fix η, c, $\underline{a}=\underline{0}$, and $\boldsymbol{A}=(\boldsymbol{I}(\alpha,\beta))^{-1/2}$. Here, the matrix $\boldsymbol{I}(\alpha,\beta)$ is the Fisher information matrix given in (27);

Step 3: Solve the following two equations for $\underline{a}$ and $\boldsymbol{A}$:

$$\begin{aligned}\boldsymbol{A}^\top\boldsymbol{A} &= \boldsymbol{M}_2^{-1},\\ \underline{a} &= \frac{\int_0^\infty \underline{s}(t;\alpha,\beta)W_c(t;\alpha,\beta)\mathrm{d}F_T(t;\alpha,\beta))}{\int_0^\infty W_c(t;\alpha,\beta)\mathrm{d}F_T(t;\alpha,\beta)},\end{aligned}$$

where for $k=1$ and 2,

$$\boldsymbol{M}_k=\int_0^\infty\{(s(t;\alpha,\beta)-\underline{a})(\underline{s}(t;\alpha,\beta)-\underline{a})^\top\}(W_c(t;\alpha,\beta))^k\mathrm{d}F_T(t;\alpha,\beta).$$

The current values of α, β, $\underline{a}$ and $\boldsymbol{A}$ are used as starting values for solving the equations in this step;

Step 4: Compute $\boldsymbol{M}_1$, using $\underline{a}$ and $\boldsymbol{A}$ from Step 3, and

$$\Delta(\alpha,\beta)=\boldsymbol{M}_1^{-1}\left[\frac{1}{n}\sum_{i=1}^n(\underline{s}(t_i;\alpha,\beta)-\underline{a})(W_c(t_i;\alpha,\beta))\right];$$

Step 5: If

$$\max\left\{\left|\frac{\Delta\alpha}{\alpha}\right|,\left|\frac{\Delta\beta}{\beta}\right|\right\}>\eta,$$

then a new (α, β) can be obtained as previous $(\alpha, \beta) + \Delta(\alpha, \beta)$, and then return to Step 3. Otherwise, we stop the iteration.

4 Interval Estimation

In the last section, we have discussed several point estimators of the parameters α and β. In this section, we discuss different interval estimators of α and β based on a random sample $\{t_1, \cdots, t_n\}$ of size n from BS(α, β).

4.1 Ng-Kundu-Balakrishnan Estimators

Ng et al. [195] proposed interval estimation of the parameters based on the asymptotic distribution of the ML and MM estimators, as provided in (29) and (37), respectively. Simulation results revealed that the ML and MM estimators are highly biased, and that they are of the order given in (38). Almost unbiased bias-corrected ML (UML) and MM (UMM) estimators have been proposed by these authors as follows:

$$\begin{aligned} \widehat{\alpha}^* &= \left(\frac{n}{n-1}\right)\widehat{\alpha}, \qquad \widehat{\beta}^* = \left(1 + \frac{\widehat{\alpha}^{*2}}{4n}\right)^{-1}\widehat{\beta}, \\ \widetilde{\alpha}^* &= \left(\frac{n}{n-1}\right)\widetilde{\alpha}, \qquad \widetilde{\beta}^* = \left(1 + \frac{\widetilde{\alpha}^{*2}}{4n}\right)^{-1}\widetilde{\beta}. \end{aligned}$$

Here, $\widehat{\alpha}$ and $\widehat{\beta}$ are the ML estimators, while $\widetilde{\alpha}$ and $\widetilde{\beta}$ are the MM estimators of α and β, respectively. Based on the UML and UMM estimators, 100(1-γ)% confidence intervals of α and β can be obtained as follows:

$$\left[\widehat{\alpha}^*\left(\sqrt{\frac{n}{2}}\frac{z_{\gamma/2}}{(n-1)} + 1\right)^{-1}, \widehat{\alpha}^*\left(\sqrt{\frac{n}{2}}\frac{z_{1-\gamma/2}}{(n-1)} + 1\right)^{-1}\right],$$

$$\left[\widehat{\beta}^*\left(\frac{n}{h_1(\widehat{\alpha}^*)}\frac{4z_{\gamma/2}}{(4n + \widehat{\alpha}^{*2})} + 1\right)^{-1}, \widehat{\beta}^*\left(\frac{n}{h_1(\widehat{\alpha}^*)}\frac{4z_{1-\gamma/2}}{(4n + \widehat{\alpha}^{*2})} + 1\right)^{-1}\right],$$

and

$$\left[\widetilde{\alpha}^*\left(\sqrt{\frac{n}{2}}\frac{z_{\gamma/2}}{(n-1)}+1\right)^{-1}, \widetilde{\alpha}^*\left(\sqrt{\frac{n}{2}}\frac{z_{1-\gamma/2}}{(n-1)}+1\right)^{-1}\right],$$

$$\left[\widetilde{\beta}^*\left(\frac{n}{h_2(\widetilde{\alpha}^*)}\frac{4z_{\gamma/2}}{(4n+\widetilde{\alpha}^{*2})}+1\right)^{-1}, \widetilde{\beta}^*\left(\frac{n}{h_2(\widetilde{\alpha}^*)}\frac{4z_{1-\gamma/2}}{(4n+\widetilde{\alpha}^{*2})}+1\right)^{-1}\right],$$

respectively; see also Wu and Wong [272]. Here, $h_1(x) = 0.25 + x^{-2} + I(x)$, $h_2(x) = (1+3/4x^2)/(1+x^2/2)^2$, with $I(\cdot)$ being as defined in (30), and z_p is the p-th percentile point of a standard normal distribution. Extensive simulation results suggest that the performance of the bias-corrected technique is quite good if the sample size is at least 20. If the sample size is less than 20, the coverage percentages are slightly lower than the corresponding nominal levels.

4.2 Wu and Wong Estimators

Since the ML, UML, MM and UMM estimators do not work satisfactorily in case of small sample sizes, Wu and Wong [272] used the higher-order likelihood-based method, as proposed by Barndorff-Nielsen [32, 33], for constructing confidence intervals for the parameters. This is known as the modified signed log-likelihood ratio statistic, and is known to have a higher-order accuracy, and it performs quite well even when the sample is quite small. To be specific, let

$$r^*(\beta) = r(\beta) + r(\beta)^{-1}\ln\left\{\frac{q(\beta)}{r(\beta)}\right\},$$

where $r(\beta)$ is the signed log-likelihood ratio statistic defined by

$$r(\beta) = \text{sgn}(\widehat{\beta}-\beta)\{2[l(\widehat{\underline{\theta}}) - l(\widehat{\underline{\theta}}_\beta)]\}^{1/2},$$

with $\underline{\theta} = (\alpha, \beta)$ being the unknown set of parameters, $l(\cdot)$ being the log-likelihood function as defined in (23), $\widehat{\underline{\theta}} = (\widehat{\alpha}, \widehat{\beta})$ being the overall ML estimator of $\underline{\theta}$, and $\widehat{\underline{\theta}}_\beta = (\widehat{\alpha}_\beta, \beta)$ being the constrained ML estimator of $\underline{\theta}$ for a given β, *i.e.*,

$$\widehat{\alpha}_\beta = \left[\frac{s}{\beta} + \frac{\beta}{r} - 2\right]^{1/2}.$$

Moreover, a general form of $q(\beta)$ takes on the following form [see Fraser et al. [85]]:

$$q(\beta) = \frac{|l_{;V}(\widehat{\underline{\theta}}) - l_{;V}(\widehat{\underline{\theta}}_\beta) l_{\alpha;V}(\widehat{\underline{\theta}}_\beta)|}{|l_{\theta;V}(\widehat{\underline{\theta}})|} \left\{ \frac{|j_{\underline{\theta}\,\underline{\theta}}(\widehat{\underline{\theta}})|}{|j_{\alpha\alpha}(\widehat{\underline{\theta}}_\beta)|} \right\}^{1/2},$$

where $j_{\alpha\alpha}(\underline{\theta})$ is the observed Fisher information matrix and $j_{\alpha\alpha}(\underline{\theta})$ is the observed nuisance information matrix. The quantity $l_{;V}(\underline{\theta})$ is known as the likelihood gradient, and is defined as

$$\underline{V} = (v_1, v_2) = -\left(\frac{\partial z(t;\underline{\theta})}{\partial t}\right)^{-1} \left(\frac{\partial z(t;\underline{\theta})}{\partial \underline{\theta}}\right)\Bigg|_{\widehat{\underline{\theta}}};$$

it is a vector array with a vector pivotal quantity $\underline{z}(t;\underline{\theta}) = (z_1(t;\underline{\theta}), \cdots, z_n(t;\underline{\theta}))$. The likelihood gradient becomes

$$l_{;V} = \left\{ \frac{d}{dv_1} l(\underline{\theta};t), \frac{d}{dv_2} l(\underline{\theta};t) \right\}^\top,$$

where

$$\frac{d}{dv_k} l(\underline{\theta};t) = \sum_{i=1}^{n} l_{t_i}(\underline{\theta};t) v_{ki}; \quad k = 1, 2,$$

is the directional derivative of the log-likelihood function taken in the direction $\underline{v}_k = (v_{k1}, \cdots, v_{kn})$ on the data space with gradient $l_{t_i}(\underline{\theta};t) = (\partial/\partial t_i) l(\underline{\theta};t)$, $i = 1, \cdots, n$. Moreover,

$$l_{\underline{\theta};V}(\underline{\theta}) = \frac{\partial l_{;V}(\underline{\theta})}{\partial \underline{\theta}}.$$

It has been shown by Fraser et al. [85] that $r^*(\beta)$ is asymptotically distributed as $N(0,1)$, and with order of accuracy $O(n^{-3/2})$, and a 100(1-γ)% confidence interval for β based on $r^*(\beta)$ is then

$$\{\beta : |r^*(\beta)| \leq z_{\gamma/2}\}.$$

Based on a complete data from the BS distribution, Wu and Wong [272] constructed confidence intervals of α and β by considering $z = (z_1, \cdots, z_n)$ as the pivotal quantity, since

$$z_i = z_i(t_i;\underline{\theta}) = \frac{1}{\alpha}\left[\left(\frac{t_i}{\beta}\right)^{1/2} - \left(\frac{t_i}{\beta}\right)^{-1/2}\right], \quad i = 1, \cdots, n,$$

are distributed as $\mathrm{N}(0,1)$. These authors then performed an extensive Monte Carlo simulation study mainly for small sample sizes and from it, they observed that even for a sample of size 5, the coverage probabilities are quite close to the corresponding nominal values.

5 Bayes Estimations

In the last two sections, we have discussed point and interval estimation of the parameters based on the frequentist approach. In this section, we take on the Bayesian approach.

5.1 Achcar Estimators

Achcar [1] first considered the Bayesian inference of the parameters of a BS distribution. He considered different non-informative priors on α and β for developing the Bayesian inference. The Jeffreys prior density for α and β is given by

$$\pi(\alpha,\beta) \propto \{\det \boldsymbol{I}(\alpha,\beta)\}^{1/2},$$

where $\boldsymbol{I}(\alpha,\beta)$ is the Fisher information matrix given in (27). Considering Laplace approximation $E(T+\beta)^{-2} \approx 1/(4\beta)^2$, Jeffreys' non-informative prior takes on the form

$$\pi(\alpha,\beta) \propto \frac{1}{\alpha\beta}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{1/2}, \quad \alpha>0, \beta>0. \tag{47}$$

Another non-informative prior considered by Achcar [1] is as follows:

$$\pi(\alpha,\beta) \propto \frac{1}{\alpha\beta}, \quad \alpha>0, \beta>0. \tag{48}$$

Note that the prior in (48) can be obtained from the prior in (47) by assuming $H(\alpha^2)=1$, where

$$H(\alpha^2) = \left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{-1/2}. \tag{49}$$

Based on the prior in (47), the posterior density of α and β becomes

$$\pi(\alpha, \beta|\text{data}) \propto \frac{\prod_{i=1}^{n}(\beta + t_i)exp\{-A(\beta)/\alpha^2\}}{\alpha^{n+1}\beta^{(n/2)+1}H(\alpha^2)}, \tag{50}$$

where $H(\alpha^2)$ is as defined in (49) and

$$A(\beta) = \frac{ns}{2\beta} + \frac{n\beta}{2r} - n.$$

Based on the prior in (48), the posterior density α and β becomes the same as in (50), with $H(\alpha^2) = 1$. Using Laplace approximation, Achcar [1] then provided approximate posterior density functions of α, based on the priors in (47) and (48), to be

$$\pi(\alpha|\text{data}) \propto \alpha^{-(n+1)}(4 + \alpha^2)^{1/2} \exp\left\{-\frac{n}{\alpha^2}(\sqrt{s/r} - 1)\right\}$$

and

$$\pi(\alpha|\text{data}) \propto \alpha^{-n} \exp\left\{-\frac{n}{\alpha^2}(\sqrt{s/r} - 1)\right\},$$

respectively. Similarly, the approximate posterior density functions of β, based on the priors in (47) and (48), are obtained as

$$\pi(\beta|\text{data}) \propto \frac{\prod_{i=1}^{n}(\beta + t_i)\left\{4 + [2n/(n+2)][s/(2\beta) + \beta/(2r) - 1]\right\}^{1/2}}{\beta^{n/2+1}\left\{s/(2\beta) + \beta/(2r) - 1\right\}^{(n+1)/2}}, \quad \beta > 0,$$

and

$$\pi(\beta|\text{data}) \propto \frac{\beta^{-(n/2+1)}\prod_{i=1}^{n}(\beta + t_i)}{(s/(2\beta) + \beta/(2r) - 1)^{n/2}}, \quad \beta > 0,$$

respectively. Evidently, the Bayes estimates of α and β cannot be obtained in closed-form. Achcar [1] proposed to use the mode of the corresponding posterior density functions as Bayes estimates of the unknown parameters. He noted that the posterior modes are quite close to the corresponding ML estimates of the parameters.

5.2 Xu and Tang Estimators

More recently, Xu and Tang [275] considered the Bayesian inference of the parameters based on the prior in (48) and showed that the prior in (48) is the reference prior, as introduced by Bernardo [40] and further developed by Berger and Bernardo [39]. Based on Lindley's [177] method, the approximate Bayes estimates of α and β, with respect to squared error loss function, are obtained as

$$\widehat{\alpha}_B = \widehat{\alpha} + \frac{1}{2}\left[\left(-\frac{\widehat{\alpha}}{2n} + \frac{3}{\widehat{\alpha}n^2}\sum_{i=1}^{n}\left(\frac{t_i}{\widehat{\beta}} + \frac{\widehat{\beta}}{t_i} - 2\right)\right) + \frac{\sum_{i=1}^{n}\widehat{\alpha}t_i}{2n^2\widehat{\beta}(1+\widehat{\alpha}(2\pi)^{-1/2}H(\widehat{\alpha}))}\right] - \frac{\widehat{\alpha}}{2n}$$

and

$$\begin{aligned}\widehat{\beta}_B &= \widehat{\beta} + \frac{1}{2}\left(-\frac{n}{\widehat{\beta}^3} + \sum_{i=1}^{n}\frac{2}{(t_i+\widehat{\beta})^3} + \sum_{i=1}^{n}\frac{3t_i}{\widehat{\alpha}^2\widehat{\beta}^4}\right)\left(\frac{1+\widehat{\alpha}(2\pi)^{-1/2}H(\widehat{\alpha})}{\widehat{\alpha}^2\widehat{\beta}^2}\right)^2 \\ &\quad + \frac{3\sum_{i=1}^{n}\widehat{\beta}^2/t_i - \widehat{\beta}t_i}{4n^2\widehat{\beta}(1+\widehat{\alpha}(2\pi)^{-1/2}H(\widehat{\alpha}))} - \frac{\widehat{\alpha}^3\widehat{\beta}^2}{2n(1+\widehat{\alpha}(2\pi)^{-1/2}H(\widehat{\alpha}))},\end{aligned}$$

respectively. Here, $\widehat{\alpha}$ and $\widehat{\beta}$ are the ML estimators of α and β, respectively, and $h(\cdot)$ is as defined earlier in (49). Although Bayes estimates with respect to squared error loss function can be obtained using Lindley's approximation, the associated credible intervals cannot be obtained by the same method. So, Xu and Tang [275] used Gibbs sampling procedure to generate posterior samples of α and β and based on those posterior samples, Bayes estimates and the associated credible intervals were obtained. Simulation results showed that the Bayes estimates based on reference priors performed quite well even for small sample sizes. Based on large scale simulations, it was also observed that the Bayes estimates based on Gibbs sampling technique performed better than the ML estimates as well as the approximate Bayes estimates based on Lindley's approximation, and that the performance was satisfactory even for small sample sizes.

6 Point and Interval Estimation Based on Censored Samples

In this section, we discuss point and interval estimation of the parameters of a BS distribution when the data are Type-II and progressively Type-II censored.

6.1 Type-II Censoring

Ng et al. [196] considered the point and interval estimation of the two-parameter BS distribution based on Type-II censored samples. They mainly considered the ML estimators of the parameters and also used the observed Fisher information matrix to construct confidence intervals for the parameters. Let $\{t_{1:n}, \ldots, t_{r:n}\}$ be an ordered Type-II right censored sample from n units placed on a life-testing experiment. Suppose the lifetime distribution of each unit follows the two-parameter BS distribution with PDF as in (2), and that the largest $(n-r)$ lifetimes have been censored. Based on the observed Type-II right censored sample, the log-likelihood can be written as [see Balakrishnan and Cohen [16]]

$$\begin{aligned} \ln(L) &= \text{const.} + (n-r)\ln\left\{1 - \Phi\left[\frac{1}{\alpha}\epsilon\left(\frac{t_{r:n}}{\beta}\right)\right]\right\} - r\ln(\alpha) - r\ln(\beta) \\ &\quad + \sum_{i=1}^{r}\epsilon'\left(\frac{t_{i:n}}{\beta}\right) - \frac{1}{2\alpha^2}\sum_{i=1}^{r}\epsilon^2\left(\frac{t_{i:n}}{\beta}\right), \end{aligned} \tag{51}$$

where $\epsilon(\cdot)$ and $\epsilon'(\cdot)$ are as defined earlier in (10).

The ML estimators of the parameters can then be obtained by differentiating the log-likelihood function in (51) with respect to the parameters α and β and equating them to zero. They cannot be obtained in explicit forms. If we use the notation

$$t^*_{i:n} = \frac{t_{i:n}}{\beta}, \quad H(x) = \frac{\phi(x)}{\Phi(x)}, \quad g(\alpha,\beta) = \frac{\alpha(n-r)}{r}H\left[\frac{1}{\alpha}\epsilon(t^*_{r:n})\right], \quad h_1(\beta) = \epsilon(t^*_{r:n}),$$

$$h_2(\beta) = -\frac{1}{r}\sum_{i=1}^{r}\epsilon^2(t^*_{i:n}), \quad h_3(\beta) = \left[1 + \sum_{i=1}^{r}\frac{t^*_{i:n}\epsilon''(t^*_{i:n})}{\epsilon'(t^*_{i:n})}\right]^{-1} t^*_{r:n}\epsilon'(t^*_{r:n}),$$

$$h_4(\beta) = \left[1 + \sum_{i=1}^{r}\frac{t^*_{i:n}\epsilon''(t^*_{i:n})}{\epsilon'(t^*_{i:n})}\right]^{-1}\left[-\frac{1}{r}\sum_{i=1}^{r}t^*_{i:n}\epsilon(t^*_{i:n})t^*_{r:n}\epsilon'(t^*_{i:n})\right],$$

$$\psi^2(\beta) = \frac{h_2(\beta)h_3(\beta) - h_1(\beta)h_4(\beta)}{h_1(\beta) - h_3(\beta)}, \quad u^* = \frac{1}{r}\sum_{i=1}^{r}t^*_{i:n}, \quad v^* = \left[\frac{1}{r}\sum_{i=1}^{r}(t^*_{i:n})^{-1}\right]^{-1},$$

$$K^*(\beta) = \left[\frac{1}{r}\sum_{i=1}^{r}(1+t^*_{i:n})^{-1}\right]^{-1}, \quad K^{'*}(\beta) = [K^*(\beta)]^2\left[\frac{1}{r}\sum_{i=1}^{r}(1+t^*_{i:n})^{-2}\right],$$

then the ML estimator of β can be obtained as the unique solution of the non-linear equation

$$Q(\beta) = \psi^2(\beta)\left[\frac{1}{2} - \frac{1}{K^*(\beta)}\right] - \frac{u^*}{2} + \frac{1}{2v^*} - \frac{\psi(\beta)(n-r)}{r} H\left[\frac{\epsilon(t^*_{r:n})}{\psi(\beta)}\right] t^*_{r:n}\epsilon'(t^*_{r:n}) = 0,$$

and once $\widehat{\beta}$, the ML estimator of β, is obtained, the ML estimator of α can be obtained as

$$\widehat{\alpha} = \psi(\widehat{\beta});$$

see Ng et al. [196] for pertinent details. Balakrishnan and Zhu [26] showed that for $r = 1$, the ML estimators of α and β do not exist. For $n > r > 1$, the ML estimators of α and β may not always exist, and that if they do exist, they are unique. They provided some sufficient conditions for the existence of ML estimators of the parameters in this case.

For the construction of confidence intervals, Ng et al. [196] used the observed Fisher information matrix and the standard asymptotic properties of the ML estimators. They also performed extensive Monte Carlo simulations to check the performance of their method. They observed that the ML estimator of α is especially biased, and if the sample size is small, it is of the form

$$\text{Bias}(\widehat{\alpha}) \approx -\frac{\alpha}{n}\left[1 + 2.5\left(1 - \frac{r}{n}\right)\right]. \tag{52}$$

Bias-correction method, as proposed in Section 4, has been suggested, and it has been observed that the biased-corrected ML estimators, as well as the confidence intervals based on biased-corrected ML estimators, both perform quite well even for small sample sizes. Barreto et al. [34] provided improved Birnbaum-Saunders estimators under Type-II censoring. Similar methods may be adopted for Type-I and hybrid censoring schemes as well. Recently, Balakrishnan and Zhu [28] discussed the existence and uniqueness of the ML estimators of the parameters α and β of BS distribution in the case of Type-I, Type-II and hybrid censored samples.

6.2 Progressive Censoring

Although Type-I and Type-II censoring schemes are the most popular censoring schemes, progressive censoring scheme has received considerable attention in the past two decades. Progressive censoring scheme was first introduced by Cohen [55], but it became very popular since the publication of the book by Balakrishnan and Aggarwala [15]. A progressive censoring scheme can be briefly described as follows. For a given n and m, choose m non-negative integers, $R_1, \ldots, R_m$, such that

$$R_1 + \ldots + R_m = n - m.$$

Now, consider the following experiment. Suppose n identical units are placed on a life-test. At the time of the first failure, say $t_{1:m:n}$, R_1 units from the $(n-1)$ surviving units are chosen at random and removed. Similarly, at the time of the second failure, say $t_{2:m:n}$, R_2 units from the $n - R_1 - 1$ surviving units are removed, and so on. Finally, at the time of the m-th failure, say $t_{m:m:n}$, all remaining R_m surviving units are removed, and the experiment terminates at $t_{m:m:n}$. For more elaborate details on progressive censoring, the readers are referred to the discussion article by Balakrishnan [14] and the recent book by Balakrishnan and Cramer [17].

Pradhan and Kundu [214] considered the inference for the parameters α and β of the BS distribution when the data are Type-II progressively censored. In a Type-II progressive censoring experiment, the data are of the form $\{(t_{1:m:n}, R_1), \ldots, (t_{m:m:n}, R_m)\}$. Based on such a progressively censored data, without the additive constant, the log-likelihood function of the observed data can be written as

$$l(\alpha, \beta|\text{data}) = \sum_{i=1}^{m} \{\ln(f_T(t_{i:m:n}; \alpha, \beta)) + R_i \ln(\Phi(-g(t_{i:m:n}; \alpha, \beta)))\},$$

where $f_T(\cdot)$ is the PDF of the BS distribution given in (2) and

$$g(t;\alpha,\beta) = \frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{1/2} - \left(\frac{\beta}{t}\right)^{1/2}\right\}.$$

As expected, the ML estimators of the parameters cannot be obtained in closed-form. Using the property that a BS distribution can be written as a mixture of an inverse Gaussian distribution and its reciprocal as mentioned earlier in Section 3.3, Pradhan and Kundu [214] provided an EM algorithm which can be used quite effectively to compute the ML estimates of the parameters. It is observed that in each 'E-step' of the algorithm, the corresponding 'M-step' can be obtained in an explicit form. Moreover, by using the method of Louis [180], the confidence intervals of the parameters can also be obtained conveniently.

7 Other Univariate Issues

In this section we will discuss some of the other issues related to the univariate BS distribution. First, we will consider the stochastic orderings of BS classes of distribution functions under different conditions. Then, we will discuss mixture of two BS distributions, its properties and different inferential aspects.

7.1 Stochastic Orderings

Stochastic orderings have been studied rather extensively for various lifetime distributions; see, for example, the book by Shaked and Shantikumar [244]. However, very little work seems to have been done for the BS model in this regard with the first work possibly by Fang et al. [77]. By considering two independent and non-identically distributed BS random variables, they established stochastic ordering results for parallel and series systems. Specifically, let $X_i \sim \text{BS}(\alpha_i, \beta_i)$, independently, for $i = 1, \ldots, n$. Further, let $X_i^* \sim \text{BS}(\alpha_i^*, \beta_i^*)$, independently, for $i = 1, \ldots, n$. Then, Fang et al. [77] established that when

$\alpha_1 = \ldots = \alpha_n = \alpha_1^* = \ldots = \alpha_n^*$, $(\beta_1, \ldots, \beta_n) \geq_m (\beta_1^*, \ldots, \beta_n^*)$ implies $X_{n:n} \geq_{\text{st}} X_{n:n}^*$. Here, $X \geq_{\text{st}} Y$ denotes that Y is smaller than X in the usual stochastic order meaning $S_X(x) \geq S_Y(x)$ for all x, where $S_X(\cdot)$ and $S_Y(\cdot)$ denote the survival functions of X and Y, respectively. Also, if $(\lambda_1, \ldots, \lambda_n)$ and $(\lambda_1^*, \ldots, \lambda_n^*)$ are two real vectors and $\lambda_{[1]} \geq \ldots \geq \lambda_{[n]}$ and $\lambda_{[1]}^* \geq \ldots \geq \lambda_{[n]}^*$ denote their ordered components, then $(\lambda_1, \ldots, \lambda_n) \geq_m (\lambda_1^*, \ldots, \lambda_n^*)$ denotes that the second vector is majorized by the first vector meaning $\sum_{i=1}^{k} \lambda_{[i]} \geq \sum_{i=1}^{k} \lambda_{[i]}^*$ for $k = 1, \ldots, n-1$ and $\sum_{i=1}^{n} \lambda_{[i]} = \sum_{i=1}^{n} \lambda_{[i]}^*$. Similarly, Fang et al. [77] established that when $\alpha_1 = \ldots = \alpha_n = \alpha_1^* = \ldots = \alpha_n^*$, $(1/\beta_1, \ldots, 1/\beta_n) \geq_m (1/\beta_1^*, \ldots, 1/\beta_n^*)$ implies $X_{1:n}^* \geq_{\text{st}} X_{1:n}$. They also further proved that when $\beta_1 = \ldots = \beta_n = \beta_1^* = \ldots = \beta_n^*$, $(1/\alpha_1, \ldots, 1/\alpha_n) \geq_m (1/\alpha_1^*, \ldots, 1/\alpha_n^*)$ implies $X_{n:n} \geq_{\text{st}} X_{n:n}^*$ and $X_{1:n}^* \geq_{\text{st}} X_{1:n}$.

There are several open problems in this direction; for example, are there such results for other order statistics, and for other stochastic orderings like hazard rate ordering and likelihood-ratio ordering?

7.2 Mixture of Two BS Distributions

Balakrishnan et al. [18] studied the properties of a mixture of two BS distributions and then discussed inferential issues for the parameters of such a mixture model. The mixture of two BS distributions has its PDF as

$$f_Y(t; \alpha_1, \beta_1, \alpha_2, \beta_2) = p f_{T_1}(t; \alpha_1, \beta_1) + (1-p) f_{T_2}(t; \alpha_2, \beta_2), \qquad t > 0, \tag{53}$$

where $T_i \sim \text{BS}(\alpha_i, \beta_i)$, $f_{T_i}(\cdot)$ denotes the PDF of T_i, for $i = 1$, 2, and $0 \leq p \leq 1$. The distribution of the random variable Y, with PDF in (53), is known as the mixture distribution of two BS models (MTBS), and is denoted by $\text{MTBS}(\alpha_1, \beta_1, \alpha_2, \beta_1, p)$. Although it has not been formally proved, it has been observed graphically that the PDF and HF of Y can be either unimodal or bimodal depending on the parameter values. Note that the first two raw moments of Y can be easily obtained as

$$\begin{aligned}
E(Y) &= p\beta_1\left[1+\frac{1}{2}\alpha_1^2\right]+(1-p)\beta_2\left[1+\frac{1}{2}\alpha_2^2\right],\\
E(Y^2) &= \frac{3}{2}p\beta_1^2\left[\alpha_1^4+2\alpha_1^2+\frac{2}{3}\right]+\frac{3}{2}(1-p)\beta_2^2\left[\alpha_2^4+2\alpha_2^2+\frac{2}{3}\right].
\end{aligned}$$

The higher-order moments also can be obtained involving some infinite series expressions. Moreover, if $Y \sim \text{MTBS}(\alpha_1, \beta_1, \alpha_2, \beta_2, p)$, then $1/Y \sim \text{MTBS}(\alpha_1, 1/\beta_1, \alpha_2, 1/\beta_2, p)$. Therefore, the inverse moments of Y also can be obtained as done earlier.

We shall now discuss the estimation of the parameters of $\text{MTBS}(\alpha_1, \beta_1, \alpha_2, \beta_2, p)$ based on a random sample $\{y_1, \cdots, y_n\}$ from Y. The ML estimators of the parameters can be obtained by solving a five-dimensional optimization problem. To avoid that, Balakrishnan et al. [18] suggested the use of the EM algorithm, which can be implemented rather easily as follows. Let us rewrite the PDF of Y in (53) as

$$f_Y(y;\mu_1,\lambda_1,\mu_2,\lambda_2,p_1,p_2)=\frac{1}{2}\sum_{j=1}^{2}p_j f_{X_1}(y;\mu_j,\lambda_j)+\frac{1}{2}\sum_{j=1}^{2}p_j f_{X_2}(y;\mu_j,\lambda_j),$$

where $f_{X_1}(y;\mu_j,\lambda_j)$ is defined in (20), $f_{X_2}(y;\mu_j,\lambda_j) = y f_{X_1}(y;\mu_j,\lambda_j)/\mu_j$ as defined in Section 2.4, $\mu_j = \beta_j$, $\lambda_j = \beta_j/\alpha_j^2$, for $j = 1, 2$, $p_1 = p$ and $p_2 = 1-p$. To implement the EM algorithm, we may treat this as a missing value problem. For the random variable Y, define an associated random vector $W = (U_1, V_1, U_2, V_2)$ as follows. Here, each U_j and V_j can take on values 0 or 1, with $\sum_{j=1}^{2}(U_j + V_j) = 1$, where $P(U_j = 1) = P(V_j = 1) = p_j/2$, for $j = 1, 2$. Further, $Y|(U_j = 1)$ and $Y|(V_j = 1)$ have densities $f_{X_1}(\cdot;\mu_j,\lambda_j)$ and $f_{X_2}(\cdot;\mu_j\lambda_j)$, respectively, for $j = 1, 2$. From these specifications, it is possible to get uniquely the joint distribution of (Y, W). Now, suppose we have the complete observations (y_i, w_i), where $w_i = (u_{i1}, v_{i1}, u_{i2}, v_{i2})$, for $i = 1, \cdots, n$. Then, with $\underline{\theta} = (\mu_1, \lambda_1, p_1, \mu_2, \lambda_2, p_2)$ being the parameter vector, the complete data log-likelihood function is given by

$$l^{(c)}(\underline{\theta}|y,w) = c \;+\; \sum_{j=1}^{2}(u_{\cdot j}+v_{\cdot j})\log(p_j)+\sum_{i=1}^{n}\sum_{j=1}^{2}u_{ij}\log(f_{X_1}(y_i;\mu_j,\lambda_j))$$

$$+ \quad \sum_{i=1}^{n}\sum_{j=1}^{2} v_{ij}\log(f_{X_2}(y_i;\mu_j,\lambda_j)), \tag{54}$$

where $u_{\cdot j} = \sum_{i=1}^{n} u_{ij}$, $v_{\cdot j} = \sum_{i=1}^{n} v_{ij}$, and c is a constant independent of the unknown parameters. Based on the complete data log-likelihood function in (54), it can be shown that the ML estimators of the model parameters are as follows:

$$\widehat{p}_j = \frac{u_{\cdot j} + v_{\cdot j}}{n}, \qquad j = 1, 2,$$

and

$$\begin{aligned} \widehat{\mu}_j &= \frac{B_j(A_j - B_j) + \sqrt{(B_j(A_j - B_j))^2 + A_jC_jD_j(2B_j - A_j)}}{D_jA_j}, \\ \widehat{\lambda}_j &= \frac{[u_{\cdot j} + v_{\cdot j}]\widehat{\mu}_j^2}{\sum_{i=1}^{n}(u_{ij} + v_{ij})(y_i - \mu_j)^2/y_i}. \end{aligned}$$

Here, $A_j = v_{\cdot j}$, $B_j = (u_{\cdot j} + v_{\cdot j})/2$, $C_j = \sum_{i=1}^{n} y_i(u_{ij} + v_{ij})y_i/2$ and $D_j = \sum_{i=1}^{n} y_i(u_{ij} + v_{ij})/(2y_i)$. Thus, in the case of complete data, the ML estimators can be expressed in explicit forms. The existence of the explicit expressions of the ML estimators simplifies the implementation of the EM algorithm, and also makes it computationally efficient. Now, the EM algorithm can be described very easily. Suppose, at the m-th stage of the EM algorithm, the estimate of the unknown parameter $\underline{\theta}$ is $\underline{\theta}^{(m)} = (\mu_1^{(m)}, \lambda_1^{(m)}, p_1^{(m)}, \mu_2^{(m)}, \lambda_2^{(m)}, p_2^{(m)})$. Before proceeding further, we need to introduce the following notation:

$$a_{ij}^{(m)} = \mathrm{E}(U_{ij}|y, \underline{\theta}^{(m)}) \quad \text{and} \quad b_{ij}^{(m)} = \mathrm{E}(V_{ij}|y, \underline{\theta}^{(m)}).$$

Note here that $a_{ij}^{(m)}$ and $b_{ij}^{(m)}$ are the usual posterior probabilities associated with the i-th observation, which are given by

$$\begin{aligned} a_{ij}^{(m)} &= \frac{p_j^{(m)} f_{X_1}(y_i;\mu_j^{(m)},\lambda_j^{(m)})}{\sum_{l=1}^{2} p_l^{(m)} f_{X_1}(y_i;\mu_l^{(m)},\lambda_l^{(m)}) + \sum_{l=1}^{2} p_l^{(m)} f_{X_2}(y_i;\mu_l^{(m)},\lambda_l^{(m)})}, \\ b_{ij}^{(m)} &= \frac{p_j^{(m)} f_{X_2}(y_i;\mu_j^{(m)},\lambda_j^{(m)})}{\sum_{l=1}^{2} p_l^{(m)} f_{X_1}(y_i;\mu_l^{(m)},\lambda_l^{(m)}) + \sum_{l=1}^{2} p_l^{(m)} f_{X_2}(y_i;\mu_l^{(m)},\lambda_l^{(m)})}. \end{aligned}$$

Therefore, at the m-th iteration, the E-step of the EM algorithm can be obtained by computing the pseudo log-likelihood function as follows:

$$l^{(ps)}(\underline{\theta}|y, \underline{\theta}^{(m)}) = \sum_{i=1}^{n}\sum_{j=1}^{2} a_{ij}^{(m)} \log\left(\frac{p_j}{2} f_{X_1}(y_i; \mu_j, \lambda_j)\right) + \sum_{i=1}^{n}\sum_{j=1}^{2} b_{ij}^{(m)} \log\left(\frac{p_j}{2} f_{X_2}(y_i; \mu_j, \lambda_j)\right). \tag{55}$$

The M-step of the EM algorithm involves the maximization of (55) with respect to the unknown parameters, and for $j = 1$ and 2, they are obtained as follows:

$$\begin{aligned}
\widehat{p}_j^{(m+1)} &= \frac{a_{.j}^{(m)} + b_{.j}^{(m)}}{n}, \\
\widehat{\mu}_j^{(m+1)} &= \frac{B_j^{(m)}(A_j^{(m)} - B_j^{(m)}) + \sqrt{(B_j^{(m)}(A_j^{(m)} - B_j^{(m)}))^2 + A_j^{(m)} C_j^{(m)} D_j^{(m)}(2B_j^{(m)} - A_j^{(m)})}}{D_j^{(m)} A_j^{(m)}}, \\
\widehat{\lambda}_j^{(m+1)} &= \frac{[a_{.j}^{(m)} + v_{.j}^{(m)}][\widehat{\mu}_j^{(m+1)}]^2}{\sum_{i=1}^{n}(a_{ij}^{(m)} + b_{ij}^{(m)})(y_i - \mu_j^{(m+1)})^2/y_i}.
\end{aligned}$$

Here, $a_{.j}^{(m)} = \sum_{i=1}^{n} a_{ij}^{(m)}$, $b_{.j}^{(m)} = \sum_{i=1}^{n} b_{ij}^{(m)}$, $A_j^{(m)} = b_{.j}^{(m)}$, $B_j^{(m)} = (a_{.j}^{(m)} + b_{.j}^{(m)})/2$, $C_j^{(m)} = \sum_{i=1}^{n} y_i(a_{ij}^{(m)} + b_{ij}^{(m)})y_i/2$, and $D_j^{(m)} = \sum_{i=1}^{n} y_i(a_{ij}^{(m)} + b_{ij}^{(m)})/(2y_i)$.

8 Bivariate BS Distribution

Until now, we have focused only on univariate BS distribution. In this section, we introduce the bivariate BS (BVBS) distribution, discuss several properties and address associated inferential issues.

8.1 PDF and CDF

Kundu et al. [119] introduced the BVBS distribution by using the same idea as in the construction of the univariate BS distribution. The bivariate random vector (T_1, T_2) is said to have a BVBS, with parameters $\alpha_1 > 0$, $\beta_1 > 0$, $\alpha_2 > 0$, $\beta_2 > 0$, $-1 < \rho < 1$, if the joint cumulative distribution function of T_1 and T_2 can be expressed as

$$F_{T_1,T_2}(t_1,t_2) = P(T_1 \le t_1, T_2 \le t_2) = \Phi_2\left[\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\beta_2}} - \sqrt{\frac{\beta_2}{t_2}}\right); \rho\right]$$

for $t_1 > 0$, $t_2 > 0$, where $\Phi_2(u, v; \rho)$ is the cumulative distribution function of a standard bivariate normal vector (Z_1, Z_2) with correlation coefficient ρ. The corresponding joint PDF is

$$\begin{aligned} f_{T_1,T_2}(t_1,t_2) &= \phi_2\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}}\right), \frac{1}{\alpha_2}\left(\sqrt{\frac{t_2}{\beta_2}} - \sqrt{\frac{\beta_2}{t_2}}\right); \rho\right) \\ &\quad \times \frac{1}{2\alpha_1\beta_1}\left\{\left(\frac{\beta_1}{t_1}\right)^{1/2} + \left(\frac{\beta_1}{t_1}\right)^{3/2}\right\} \times \frac{1}{2\alpha_2\beta_2}\left\{\left(\frac{\beta_2}{t_2}\right)^{1/2} + \left(\frac{\beta_2}{t_2}\right)^{3/2}\right\}, \end{aligned}$$

where $\phi_2(u, v; \rho)$ denotes the joint PDF of Z_1 and Z_2 given by

$$\phi_2(u, v, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2 + v^2 - 2\rho uv)\right\}. \tag{56}$$

From here on, we shall denote this distribution by $\mathrm{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$. As expected, the joint PDF can take on different shapes depending on the shape parameters α_1 and α_2 and the correlation parameter ρ. The surface plots of $f_{T_1,T_2}(t_1, t_2)$, for different parameter values, can be found in Kundu et al. [119]. It has been observed that the joint PDF of the BS distribution is unimodal. It is an absolutely continuous distribution and due to its flexibility, it can be used quite effectively to analyze bivariate data provided there are no ties.

8.2 Copula Representation

Recently, Kundu and Gupta [121] observed that the CDF of $\mathrm{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$ has the following copula representation:

$$F_{T_1,T_2}(t_1, t_2) = C_G(F_{T_1}(t_1; \alpha_1, \beta_1), F_{T_2}(t_2; \alpha_2, \beta_2); \rho).$$

Here, $F_{T_1}(t_1)$ and $F_{T_2}(t_2)$ denote the CDFs of $\mathrm{BS}(\alpha_1, \lambda_1)$ and $\mathrm{BS}(\alpha_2, \lambda_2)$, respectively, and $C_G(u, v; \rho)$ denotes the Gaussian copula defined by, for $0 \le u, v \le 1$,

$$C_G(u, v; \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \phi_2(x, y; \rho)\mathrm{d}x\mathrm{d}y = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho). \tag{57}$$

Using the above copula structure in (57), it has been shown that a BVBS distribution has a total positivity of order two (TP_2) property if $\rho > 0$ and reverse rule of order two (RR_2) property for $\rho < 0$ for all values of $\alpha_1, \beta_1, \alpha_2$ and β_2. It has also been shown, using the above copula structure, that for a $BS_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, for all values of $\alpha_1, \beta_1, \alpha_2, \beta_2$, the Blomqvist's beta ($\beta$), Kendall's tau ($\tau$) and Spearman's rho ($\rho_S$) become

$$\beta = \frac{2}{\pi}\arcsin(\rho), \quad \tau = \frac{2}{\pi}\arcsin(\rho), \quad \rho_S = \frac{6}{\pi}\arcsin\left(\frac{\rho}{2}\right),$$

respectively.

8.3 Generation and Other Properties

It is very simple to generate bivariate BS distribution using univariate normal random number generators. The following algorithm has been suggested by Kundu et al. [119] to generate bivariate BS distribution.

Algorithm:

Step 1: Generate independent U_1 and U_2 from $N(0,1)$;

Step 2: Compute

$$Z_1 = \frac{\sqrt{1+\rho}+\sqrt{1-\rho}}{2}U_1 + \frac{\sqrt{1+\rho}-\sqrt{1-\rho}}{2}U_2,$$

$$Z_2 = \frac{\sqrt{1+\rho}-\sqrt{1-\rho}}{2}U_1 + \frac{\sqrt{1+\rho}+\sqrt{1-\rho}}{2}U_2;$$

Step 3: Set

$$T_i = \beta_i\left[\frac{1}{2}\alpha_i Z_i + \sqrt{\left(\frac{1}{2}\alpha_i Z_i\right)^2 + 1}\right]^2, \quad \text{for} \quad i = 1, 2.$$

Several properties of the bivariate BS distribution have been established by these authors. It has been observed that if $(T_1, T_2) \sim BS_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$, then $T_1 \sim BS(\alpha_1, \beta_1)$ and $T_2 \sim$

BS(α_2, β_2). T_1 and T_2 are independent if and only if $\rho = 0$. The correlation between T_1 and T_2 ranges from -1 to 1, depending on the value of ρ. Using the joint and marginal PDFs, the conditional PDF of T_1, given T_2, can be easily obtained. Different moments and product moments can be obtained involving infinite series expressions. It has also been observed that for $\rho > 0$ and for all values of $\alpha_1, \beta_1, \alpha_2, \beta_2$, T_1 (T_2) is stochastically increasing in T_2 (T_1). If $(h_1(t_1, t_2), h_2(t_1, t_2))$ denotes the bivariate hazard rate of Johnson and Kotz [108], then for $\rho > 0$ and for fixed t_2 (t_1), $h_1(t_1, t_2)(h_2(t_1, t_2))$ is an unimodal function of t_1 (t_2). It may also be noted that for all values of α_1, α_2 and ρ, if $\beta_1 = \beta_2$, then the stress-strength parameter $R = P(T_1 < T_2) = 1/2$. Different reliability properties of a BVBS distribution can be found in Gupta [98].

8.4 ML and MM Estimators

We shall now discuss the estimation of the parameters $\alpha_1, \beta_1, \alpha_2, \beta_2$ and ρ based on a random sample, say $\{(t_{1i}, t_{2i}), i = 1, \cdots, n\}$, of size n from $\text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2; \rho)$. Based on the random sample, the log-likelihood function is given by

$$\begin{aligned} l(\underline{\theta}) \;=\; & -n\ln(\alpha_1) - n\ln(\beta_1) - n\ln(\alpha_2) - n\ln(\beta_2) - \frac{n}{2}\ln(1-\rho^2) + \sum_{i=1}^{n} \ln\left\{ \left(\frac{\beta_1}{t_{1i}}\right)^{1/2} + \left(\frac{\beta_1}{t_{1i}}\right)^{3/2} \right\} \\ & + \sum_{i=1}^{n} \ln\left\{ \left(\frac{\beta_2}{t_{2i}}\right)^{1/2} + \left(\frac{\beta_2}{t_{2i}}\right)^{3/2} \right\} - \frac{1}{2(1-\rho^2)} \left\{ \sum_{i=1}^{n} \frac{1}{\alpha_1^2} \left(\left(\frac{t_{1i}}{\beta_1}\right)^{1/2} - \left(\frac{\beta_1}{t_{1i}}\right)^{1/2} \right)^2 \right. \\ & \left. + \sum_{i=1}^{n} \frac{1}{\alpha_2^2} \left(\left(\frac{t_{2i}}{\beta_2}\right)^{1/2} - \left(\frac{\beta_2}{t_{2i}}\right)^{1/2} \right)^2 - \frac{2\rho}{\alpha_1\alpha_2} \sum_{i=1}^{n} \left(\frac{t_{2i}}{\beta_2}\right)\left(\frac{\beta_2}{t_{2i}}\right) \right\}. \end{aligned} \tag{58}$$

The ML estimators of the parameters can be obtained by maximizing (58) with respect to α_1, α_2, β_1, β_2 and ρ. Evidently, the ML estimators cannot be obtained in explicit form. From the observation that

$$\left\{ \left(\sqrt{\frac{T_1}{\beta_1}} - \sqrt{\frac{\beta_1}{T_1}} \right), \left(\sqrt{\frac{T_2}{\beta_2}} - \sqrt{\frac{\beta_2}{T_2}} \right) \right\} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2\rho \\ \alpha_1\alpha_2\rho & \alpha_2^2 \end{pmatrix} \right),$$

it is clear that for fixed β_1 and β_2, the ML estimators of α_1, α_2 and ρ are

$$\widehat{\alpha}_i = \left(\frac{s_i}{\beta_i} + \frac{\beta_i}{r_i} - 2\right)^{1/2}, \quad i = 1, 2,$$

and

$$\widehat{\rho}(\beta_1, \beta_2) = \frac{\sum_{i=1}^{n}\left(\left(\frac{t_{1i}}{\beta_1}\right)^{1/2} - \left(\frac{\beta_1}{t_{1i}}\right)^{1/2}\right)\left(\left(\frac{t_{2i}}{\beta_2}\right)^{1/2} - \left(\frac{\beta_2}{t_{2i}}\right)^{1/2}\right)}{\sum_{i=1}^{n}\left(\left(\frac{t_{1i}}{\beta_1}\right)^{1/2} - \left(\frac{\beta_2}{t_{2i}}\right)^{1/2}\right)^2\left(\left(\frac{t_{1i}}{\beta_1}\right)^{1/2} - \left(\frac{\beta_2}{t_{2i}}\right)^{1/2}\right)^2},$$

where

$$s_i = \frac{1}{n}\sum_{k=1}^{n} t_{ik} \quad \text{and} \quad r_i = \left[\frac{1}{n}\sum_{k=1}^{n} t_{ik}^{-1}\right]^{-1}, \quad \text{for} \quad i = 1, 2.$$

It may be observed that $\widehat{\alpha}_1(\beta_1, \beta_2)$ is a function of β_1 only, while $\widehat{\alpha}_2(\beta_1, \beta_2)$ is a function of β_2 only. The ML estimators of β_1 and β_2 can then be obtained by maximizing the profile log-likelihood function of β_1 and β_2 given by

$$\begin{aligned} l_{\text{profile}}(\beta_1, \beta_2) &= l(\widehat{\alpha}_1(\beta_1), \beta_1, \widehat{\alpha}_2(\beta_2), \beta_2, \widehat{\rho}(\beta_1, \beta_2)) \\ &= -n\ln(\widehat{\alpha}_1(\beta_1)) - n\ln(\beta_1) - n\ln(\widehat{\alpha}_2(\beta_2)) - n\ln(\beta_2) - \frac{n}{2}\ln(1 - \widehat{\rho}^2(\beta_1, \beta_2)) \\ &\quad + \sum_{i=1}^{n}\ln\left\{\left(\frac{\beta_1}{t_{1i}}\right)^{1/2} + \left(\frac{\beta_1}{t_{1i}}\right)^{3/2}\right\} + \sum_{i=1}^{n}\ln\left\{\left(\frac{\beta_2}{t_{2i}}\right)^{1/2} + \left(\frac{\beta_2}{t_{2i}}\right)^{3/2}\right\}. \end{aligned}$$

Clearly, no explicit solutions exist, and so some numerical techniques like Newton-Raphson algorithm or some of its variants need to be used for computing the ML estimators of β_1 and β_2. Once the ML estimators of β_1 and β_2 are obtained, the ML estimators of α_1, α_2 and ρ can be easily obtained. The explicit expression of the Fisher information matrix, although complicated, has been provided by Kundu et al. [119], which can be used to obtain the asymptotic variance-covariance matrix of the ML estimators.

Since the ML estimators do not have explicit forms, Kundu et al. [119] proposed modified moment estimators which do have explicit expressions. Using the same idea as in the case of univariate modified moment estimators, in the bivariate case, the modified moment estimators take on the following forms:

$$\widetilde{\alpha}_i = \left\{ 2\left[\left(\frac{s_i}{r_i} \right)^{1/2} - 1 \right] \right\}^{1/2} \quad \text{and} \quad \widetilde{\beta}_i = (s_i r_i)^{1/2}, \quad \text{for} \quad i = 1, 2,$$

and

$$\widetilde{\rho} = \frac{\sum_{i=1}^{n} \left(\left(\frac{t_{1i}}{\widetilde{\beta}_1} \right)^{1/2} - \left(\frac{\widetilde{\beta}_1}{t_{1i}} \right)^{1/2} \right) \left(\left(\frac{t_{2i}}{\widetilde{\beta}_2} \right)^{1/2} - \left(\frac{\widetilde{\beta}_2}{t_{2i}} \right)^{1/2} \right)}{\sum_{i=1}^{n} \left(\left(\frac{t_{1i}}{\widetilde{\beta}_1} \right)^{1/2} - \left(\frac{\widetilde{\beta}_2}{t_{2i}} \right)^{1/2} \right)^2 \left(\left(\frac{t_{1i}}{\widetilde{\beta}_1} \right)^{1/2} - \left(\frac{\widetilde{\beta}_2}{t_{2i}} \right)^{1/2} \right)^2}.$$

Since the modified moment estimators have explicit forms, they can be used as initial values in the numerical computation of the ML estimators. Kundu et al. [119] performed some simulations to investigate the performance of the ML estimators for the bivariate BS distribution. It has been observed by these authors that even for small sample sizes, say 10, the ML estimators of all the parameters are almost unbiased. Interestingly, the bias and the mean squared errors of the ML estimators do not depend on the true value of ρ.

Kundu et al. [119] also constructed confidence intervals based on pivotal quantities associated with all these estimators obtained from the empirical Fisher information matrix. Extensive Monte Carlo simulations have been performed by these authors to examine the effectiveness of the proposed method. It has been observed that the asymptotic normality does not work satisfactorily in the case of small sample sizes. The coverage probabilities turn out to be quite low if the sample size is small. Parametric or non-parametric bootstrap confidence intervals are recommended in this case. Recently, Kundu and Gupta [121] provided a two-step estimation procedure using the copula structure, mainly based on the idea of Joe [107], and provided the asymptotic distribution of the two-step estimators. It has been observed, based on extensive simulations carried out by Kundu and Gupta [121], that the two-step estimators behave almost as efficiently as the ML estimators. Hence, they may be used effectively in place of the ML estimators to avoid solving a high-dimensional optimization problem.

9 Multivariate BS Distribution

In the last section, we have introduced the bivariate BS distribution and discussed its properties. In this section, we introduce the multivariate BS distribution along the same lines. We then present several properties of this multivariate BS distribution and also discuss the ML estimation of the model parameters.

9.1 PDF and CDF

The multivariate BS distribution has been introduced by Kundu et al. [120] along the same lines as the bivariate BS distribution, discussed in the last section.

Definition 1: Let $\underline{\alpha}, \underline{\beta} \in \mathbb{R}^p$, where $\underline{\alpha} = (\alpha_1, \cdots, \alpha_p)^\top$ and $\underline{\beta} = (\beta_1, \cdots, \beta_p)^\top$, with $\alpha_i > 0, \beta_i > 0$ for $i = 1, \cdots, p$. Let $\mathbf{\Gamma}$ be a $p \times p$ positive-definite correlation matrix. Then, the random vector $\underline{T} = (T_1, \cdots, T_p)^\top$ is said to have a p-variate BS distribution with parameters $(\underline{\alpha}, \underline{\beta}, \mathbf{\Gamma})$ if it has the joint CDF as

$$\begin{aligned} P(\underline{T} \leq \underline{t}) &= P(T_1 \leq t_1, \cdots, T_p \leq t_p) \\ &= \Phi_p \left[\frac{1}{\alpha_1} \left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}} \right), \cdots, \frac{1}{\alpha_p} \left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}} \right); \mathbf{\Gamma} \right], \end{aligned} \tag{59}$$

for $t_1 > 0, \cdots, t_p > 0$. Here, for $\underline{u} = (u_1, \cdots, u_p)^\top$, $\Phi_p(\underline{u}; \mathbf{\Gamma})$ denotes the joint CDF of a standard normal vector $\underline{Z} = (Z_1, \cdots, Z_p)^\top$ with correlation matrix $\mathbf{\Gamma}$.

The joint PDF of $\underline{T} = (T_1, \cdots, T_p)^\top$ can be obtained readily from (59) as

$$\begin{aligned} f_{\underline{T}}(\underline{t}; \underline{\alpha}, \underline{\beta}, \mathbf{\Gamma}) &= \phi_p \left(\frac{1}{\alpha_1} \left(\sqrt{\frac{t_1}{\beta_1}} - \sqrt{\frac{\beta_1}{t_1}} \right), \cdots, \frac{1}{\alpha_p} \left(\sqrt{\frac{t_p}{\beta_p}} - \sqrt{\frac{\beta_p}{t_p}} \right); \mathbf{\Gamma} \right) \\ &\quad \times \prod_{i=1}^{p} \frac{1}{2\alpha_i \beta_i} \left\{ \left(\frac{\beta_i}{t_i} \right)^{\frac{1}{2}} + \left(\frac{\beta_i}{t_i} \right)^{\frac{3}{2}} \right\}, \end{aligned} \tag{60}$$

for $t_1 > 0, \cdots, t_p > 0$; here, for $\underline{u} = (u_1, \cdots, u_p)^\top$,

$$\phi_p(u_1, \cdots, u_p; \mathbf{\Gamma}) = \frac{1}{(2\pi)^{\frac{p}{2}} |\mathbf{\Gamma}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} \underline{u}^\top \mathbf{\Gamma}^{-1} \underline{u} \right\},$$

is the PDF of the standard normal vector with correlation matrix $\mathbf{\Gamma}$. Hereafter, the p-variate BS distribution, with joint PDF as in (60), will be denoted by $\mathrm{BS}_p(\underline{\alpha}, \underline{\beta}, \mathbf{\Gamma})$.

9.2 Marginal and Conditional Distributions

The marginal and conditional distributions of $\mathrm{BS}_p(\underline{\alpha}, \underline{\beta}, \mathbf{\Gamma})$ are as follows; the proofs can be seen in the work of Kundu et al. [120].

Result: Let $\underline{T} \sim \mathrm{BS}_p(\underline{\alpha}, \underline{\beta}, \mathbf{\Gamma})$, and let $\underline{T}, \underline{\alpha}, \underline{\beta}, \mathbf{\Gamma}$ be partitioned as follows:

$$\underline{T} = \begin{pmatrix} \underline{T}_1 \\ \underline{T}_2 \end{pmatrix}, \quad \underline{\alpha} = \begin{pmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \end{pmatrix}, \quad \underline{\beta} = \begin{pmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{pmatrix}, \quad \mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} \\ \mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} \end{pmatrix}, \tag{61}$$

where $\underline{T}_1, \underline{\alpha}_1, \underline{\beta}_1$ are all $q \times 1$ vectors and $\mathbf{\Gamma}_{11}$ is a $q \times q$ matrix. The remaining elements are all defined suitably. Then, we have:

(a) $\underline{T}_1 \sim \mathrm{BS}_q(\underline{\alpha}_1, \underline{\beta}_1, \mathbf{\Gamma}_{11})$ and $\underline{T}_2 \sim \mathrm{BS}_{p-q}(\underline{\alpha}_2, \underline{\beta}_2, \mathbf{\Gamma}_{22})$;

(b) The conditional CDF of $\underline{T}_1$, given $\underline{T}_2 = \underline{t}_2$, is

$$P[\underline{T}_1 \le \underline{t}_1 | \underline{T}_2 = \underline{t}_2] = \Phi_q(\underline{w}; \mathbf{\Gamma}_{11.2});$$

(c) The conditional PDF of $\underline{T}_1$, given $\underline{T}_2 = \underline{t}_2$, is

$$f_{\underline{T}_1|(\underline{T}_2=\underline{t}_2)}(\underline{t}_1) = \phi_q(\underline{w}; \mathbf{\Gamma}_{11.2}) \prod_{i=1}^{q} \frac{1}{2\alpha_i\beta_i} \left\{ \left(\frac{\beta_i}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta_i}{t_i}\right)^{\frac{3}{2}} \right\}, \tag{62}$$

where

$$\underline{w} = \underline{v}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\underline{v}_2, \qquad \underline{v} = (v_1 \cdots, v_p)^\top, \qquad v_i = \frac{1}{\alpha_i}\left(\sqrt{\frac{t_i}{\beta_i}} - \sqrt{\frac{\beta_i}{t_i}}\right) \text{ for } i = 1, \cdots, p,$$

$$\boldsymbol{\Gamma}_{11.2} = \boldsymbol{\Gamma}_{11} - \boldsymbol{\Gamma}_{12}\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\Gamma}_{21}, \quad \underline{v} = \begin{pmatrix} \underline{v}_1 \\ \underline{v}_2 \end{pmatrix},$$

and $\underline{v}_1$, $\underline{v}_2$ are vectors of dimensions $q \times 1$ and $(p-q) \times 1$, respectively;

(d) $\underline{T}_1$ and $\underline{T}_2$ are independent if and only if $\boldsymbol{\Gamma}_{12} = \mathbf{0}$.

9.3 Generation and Other Properties

The following algorithm can be adopted to generate $\underline{T} = (T_1, \cdots, T_p)^\top$ from $\text{BS}_p(\underline{\alpha}, \underline{\beta}, \boldsymbol{\Gamma})$ in (60).

Algorithm

Step 1: Make a Cholesky decomposition of $\boldsymbol{\Gamma} = \boldsymbol{A}\boldsymbol{A}^\top$ (say);

Step 2: Generate p independent standard normal random numbers, say, $U_1, \cdots, U_p$;

Step 3: Compute $\underline{Z} = (Z_1, \cdots, Z_p)^\top = \boldsymbol{A}\ (U_1, \cdots, U_p)^\top$;

Step 4: Make the transformation

$$T_i = \beta_i \left[\frac{1}{2}\alpha_i Z_i + \sqrt{\left(\frac{1}{2}\alpha_i Z_i\right)^2 + 1} \right]^2 \quad \text{for} \quad i = 1, \cdots, p.$$

Then, $\underline{T} = (T_1, \cdots, T_p)^\top$ has the required $\text{BS}_p(\underline{\alpha}, \underline{\beta}, \boldsymbol{\Gamma})$ distribution.

Now, let use the notation $1/\underline{a} = (1/a_1, \cdots, 1/a_k)^\top$ for a vector $\underline{a} = (a_1, \cdots, a_k)^\top \in \mathbf{R}^k$, where $a_i \neq 0$, for $i = \cdots, k$. Then, we have the following result for which a detailed proof can be found in Kundu et al. [120].

Result: Let $\underline{T} \sim \text{BS}_p(\underline{\alpha}, \underline{\beta}, \boldsymbol{\Gamma})$, and $\underline{T}_1$, $\underline{T}_2$ be the same as defined in Section 9.2. Then:

(a)

$$\begin{pmatrix} \underline{T}_1 \\ \frac{1}{\underline{T}_2} \end{pmatrix} \sim \text{BS}_p\left(\begin{pmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \end{pmatrix}, \begin{pmatrix} \underline{\beta}_1 \\ \frac{1}{\underline{\beta}_2} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma}_{11} & -\boldsymbol{\Gamma}_{12} \\ -\boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{pmatrix} \right),$$

(b)

$$\begin{pmatrix} \frac{1}{\underline{T}_1} \\ \underline{T}_2 \end{pmatrix} \sim \text{BS}_p \left(\begin{pmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \end{pmatrix}, \begin{pmatrix} \frac{1}{\underline{\beta}_1} \\ \underline{\beta}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{\Gamma}_{11} & -\mathbf{\Gamma}_{12} \\ -\mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} \end{pmatrix} \right),$$

(c)

$$\begin{pmatrix} \frac{1}{\underline{T}_1} \\ \frac{1}{\underline{T}_2} \end{pmatrix} \sim \text{BS}_p \left(\begin{pmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \end{pmatrix}, \begin{pmatrix} \frac{1}{\underline{\beta}_1} \\ \frac{1}{\underline{\beta}_2} \end{pmatrix}, \begin{pmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} \\ \mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} \end{pmatrix} \right).$$

9.4 ML Estimation

Now, we discuss the ML estimators of the model parameters based on the data $\{(t_{i1}, \cdots, t_{ip})^\top; i = 1, \cdots, n\}$. The log-likelihood function, without the additive constant, is given by

$$\begin{aligned} l(\underline{\alpha}, \underline{\beta}, \mathbf{\Gamma}|\text{data}) &= -\frac{n}{2} \ln(|\mathbf{\Gamma}|) - \frac{1}{2} \sum_{i=1}^{n} \underline{v}_i^\top \mathbf{\Gamma}^{-1} \underline{v}_i - n \sum_{j=1}^{p} \ln(\alpha_j) - n \sum_{j=1}^{p} \ln(\beta_j) \\ &\quad + \sum_{i=1}^{n} \sum_{j=1}^{p} \ln \left\{ \left(\frac{\beta_{ij}}{t_{ij}} \right)^{\frac{1}{2}} + \left(\frac{\beta_{ij}}{t_{ij}} \right)^{\frac{3}{2}} \right\}, \end{aligned} \tag{63}$$

where

$$\underline{v}_i^\top = \left[\frac{1}{\alpha_1} \left(\sqrt{\frac{t_{i1}}{\beta_1}} - \sqrt{\frac{\beta_1}{t_{i1}}} \right), \cdots, \frac{1}{\alpha_p} \left(\sqrt{\frac{t_{ip}}{\beta_p}} - \sqrt{\frac{\beta_p}{t_{ip}}} \right) \right].$$

Then, the ML estimators of the unknown parameters can be obtained by maximizing (63) with respect to the parameters $\underline{\alpha}, \underline{\beta}$ and $\mathbf{\Gamma}$, which would require a $2p + \binom{p}{2}$ dimensional optimization process. For this purpose, the following procedure can be adopted for reducing the computational effort significantly. Observe that

$$\left[\left(\sqrt{\frac{T_1}{\beta_1}} - \sqrt{\frac{\beta_1}{T_1}} \right), \cdots, \left(\sqrt{\frac{T_p}{\beta_p}} - \sqrt{\frac{\beta_p}{T_p}} \right) \right]^\top \sim N_p \left(\underline{0}, \boldsymbol{D} \mathbf{\Gamma} \boldsymbol{D}^\top \right), \tag{64}$$

where $\boldsymbol{D}$ is a diagonal matrix given by $\boldsymbol{D} = \text{diag}\{\alpha_1, \cdots, \alpha_p\}$. Therefore, for given $\underline{\beta}$, the ML estimators of $\underline{\alpha}$ and $\mathbf{\Gamma}$ become

$$\widehat{\alpha}_j(\underline{\beta}) = \left(\frac{1}{n} \sum_{i=1}^{n} \left(\sqrt{\frac{t_{ij}}{\beta_j}} - \sqrt{\frac{\beta_j}{t_{ij}}} \right)^2 \right)^{\frac{1}{2}} = \left(\frac{1}{\beta_j} \left\{ \frac{1}{n} \sum_{i=1}^{n} t_{ij} \right\} + \beta_j \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{t_{ij}} \right\} - 2 \right)^{\frac{1}{2}},$$

$$j = 1, \cdots, p, \tag{65}$$

and

$$\widehat{\mathbf{\Gamma}}(\underline{\beta}) = \boldsymbol{P}(\underline{\beta})\boldsymbol{Q}(\underline{\beta})\boldsymbol{P}^{\top}(\underline{\beta}); \tag{66}$$

here, $\boldsymbol{P}(\underline{\beta})$ is a diagonal matrix given by $\boldsymbol{P}(\underline{\beta}) = \text{diag}\{1/\widehat{\alpha}_1(\underline{\beta}), \cdots, 1/\widehat{\alpha}_p(\underline{\beta})\}$, and the elements $q_{jk}(\underline{\beta})$ of the matrix $\boldsymbol{Q}(\underline{\beta})$ are given by

$$q_{jk}(\underline{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\left(\sqrt{\frac{t_{ij}}{\beta_j}} - \sqrt{\frac{\beta_j}{t_{ij}}}\right)\left(\sqrt{\frac{t_{ik}}{\beta_k}} - \sqrt{\frac{\beta_k}{t_{ik}}}\right), \quad \text{for} \quad j,k = 1,\cdots,p. \tag{67}$$

Thus, we obtain the p-dimensional profile log-likelihood function $l(\underline{\widehat{\alpha}}(\underline{\beta}), \underline{\beta}, \widehat{\mathbf{\Gamma}}(\underline{\beta})|data)$. The ML estimator of $\underline{\beta}$ can then be obtained by maximizing the p-dimensional profile log-likelihood function, and once we get the ML estimator of $\underline{\beta}$, say $\widehat{\underline{\beta}}$, the ML estimators of $\underline{\alpha}$ and $\mathbf{\Gamma}$ can be obtained readily by substituting $\widehat{\underline{\beta}}$ in place of $\underline{\beta}$ in Eqs. (65) and (66), respectively.

However, for computing the ML estimators of the parameters, we need to maximize the profile log-likelihood function of $\underline{\beta}$ and we may use the Newton-Raphson iterative process for this purpose. Finding a proper p-dimensional initial guess for $\underline{\beta}$ becomes quite important in this case. Modified moment estimators, similar to those proposed by Ng et al. [195] and described earlier, can be used effectively as initial guess, and they are as follows:

$$\beta_j^{(0)} = \left(\frac{1}{n}\sum_{i=1}^{n} t_{ij} \bigg/ \frac{1}{n}\sum_{i=1}^{n}\frac{1}{t_{ij}}\right)^{\frac{1}{2}}, \quad j = 1,\cdots,p.$$

Note that if $\underline{\beta}$ is known, then the ML estimators of $\underline{\alpha}$ and $\mathbf{\Gamma}$ can be obtained explicitly.

If $\underline{\alpha}$ and $\underline{\beta}$ are known, then the ML estimator of $\mathbf{\Gamma}$ is $\widehat{\mathbf{\Gamma}} = \boldsymbol{D}^{-1}\boldsymbol{Q}(\underline{\beta})\boldsymbol{D}^{-1}$, where the elements of the matrix $\boldsymbol{Q}(\underline{\beta})$ are as in (67) and the matrix $\boldsymbol{D}$ is as defined earlier. From (64), it immediately follows in this case that $\hat{\mathbf{\Gamma}}$ has a Wishart distribution with parameters p and $\mathbf{\Gamma}$. Furthermore, if only $\underline{\beta}$ is known, it is clear that $\widehat{\alpha}_j^2(\underline{\beta})$, defined in (65), is distributed as χ_1^2, for $j = 1,\cdots,p$.

10 Some Related Distributions

In this section, we describe various distributions related to the BS distribution that are derived from the univariate and bivariate BS distributions. We discuss their properties and also address inferential issues associated with them. We focus here mainly on sinh-normal or log-BS, bivariate sinh-normal, length-biased BS and epsilon-BS distributions.

10.1 Sinh-Normal or log-BS Distribution

Let Y be a real-valued random variable with cumulative distribution function $F_Y(\cdot)$ given by

$$P(Y \le y) = F_Y(y;\alpha,\gamma,\sigma) = \Phi\left\{\frac{2}{\alpha}\sinh\left(\frac{y-\gamma}{\sigma}\right)\right\}, \quad \text{for} \quad -\infty < y < \infty. \tag{68}$$

Here, $\alpha > 0$, $\sigma > 0$, $-\infty < \gamma < \infty$, and $\sinh(x)$ is the hyperbolic sine function of x, defined as $\sinh(x) = (e^x - e^{-x})/2$. In this case, Y is said to have a sinh-normal distribution, and is denoted by $\text{SN}(\alpha,\gamma,\sigma)$. The PDF of sinh-normal distribution is given by

$$f_Y(y;\alpha,\gamma,\sigma) = \frac{2}{\alpha\sigma\sqrt{2\pi}} \times \cosh\left(\frac{y-\gamma}{\sigma}\right) \times \exp\left[\left(-\frac{2}{\alpha^2}\sinh^2\left(\frac{y-\gamma}{\sigma}\right)\right)\right].$$

Here, $\cosh(x)$ is the hyperbolic cosine function of x, defined as $\cosh(x) = (e^x + e^{-x})/2$. In this case, α is the shape parameter, σ is the scale parameter and γ is the location parameter. Note that if $T \sim \text{BS}(\alpha,\beta)$, $\ln(T) \sim \text{SN}(\alpha,\ln(\beta),2)$. For this reason, this distribution is also known as log BS distribution.

It is clear from (68) that if $Y \sim \text{SN}(\alpha,\gamma,\sigma)$, then

$$Z = \frac{2}{\alpha}\sinh\left(\frac{Y-\gamma}{\sigma}\right) \sim \text{N}(0,1). \tag{69}$$

From (69), it follows that if $Z \sim \text{N}(0,1)$, then

$$Y = \sigma \ \text{arcsinh}\left(\frac{\alpha Z}{2}\right) + \gamma \sim \text{SN}(\alpha,\gamma,\sigma). \tag{70}$$

The representation in (70) of the sinh-normal distribution can be used for generation purposes. Using this representation, connecting with the standard normal distribution, sinh-normal distribution can be easily generated, and consequently it can be used for the generation of BS distribution as well.

The random variable Y is said to have a standard sinh-normal distribution when $\gamma = 0$ and $\sigma = 1$. Therefore, if $Y \sim \text{SN}(\alpha, 0, 1)$, then the corresponding PDF becomes

$$f_Y(y; \alpha, 0, 1) = \frac{2}{\alpha\sqrt{2\pi}} \cosh(y) \exp\left[\left(-\frac{2}{\alpha^2}\sinh^2(y)\right)\right]. \tag{71}$$

It is immediate from (71) that the $\text{SN}(\alpha, 0, 1)$ is symmetric about 0. The PDF in (71) of Y, for different values of α when $\gamma = 0$, and $\sigma = 1$, has been presented by Rieck [221]. It has been observed by Rieck [221] that for $\alpha \le 2$, the PDF is strongly unimodal and for $\alpha > 2$, it is bimodal. Furthermore, if $Y \sim \text{SN}(\alpha, \gamma, \sigma)$, then $U = 2\alpha^{-1}(Y - \gamma)/\sigma$ converges to the standard normal distribution as $\alpha \to 0$

If $Y \sim \text{SN}(\alpha, \gamma, \sigma)$, then the MGF of Y can be obtained as [see Rieck [221] or Leiva et al. [131]]

$$M(s) = E(e^{sY}) = e^{\mu s}\left[\frac{K_a(\delta^{-2}) + K_b(\delta^{-1})}{2K_{1/2}(\delta^{-2})}\right], \tag{72}$$

where $a = (\sigma s + 1)/2$, $b = (\sigma s - 1)/2$, and $K_\lambda(\cdot)$ is the modified Bessel function of the third kind, given by

$$K_\lambda(w) = \frac{1}{2}\left(\frac{w}{2}\right)^\lambda \int_0^\infty y^{-\lambda-1} e^{-y-(w^2/4y)} \text{d}y;$$

see Gradshteyn and Randzhik ([95], p. 907). Thus, by differentiating the moment generating function in (72), moments of Y can be obtained readily. It is clear that the variance or the fourth moment cannot be expressed in closed-form. However, it has been observed that as α increases, the kurtosis of Y increases. Moreover, for $\alpha \le 2$, the kurtosis of Y is smaller than that of the normal. For $\alpha > 2$, when α increases, the SN distribution begins to display bimodality, with modes that are more separated, and the kurtosis is greater than that of the normal.

We shall now discuss the estimation of the model parameters based on a random sample of size n, say $\{y_1, \cdots, y_n\}$, from SN(α, γ, σ). The log-likelihood function of the observed sample, without the additive constant, is given by

$$l(\alpha, \gamma, \sigma|\text{data}) = -n\ln(\alpha) - n\ln(\sigma) + \sum_{i=1}^{n} \cosh\left(\frac{y_i - \gamma}{\sigma}\right) - \frac{2}{\alpha^2}\sum_{i=1}^{n} \sinh^2\left(\frac{y_i - \gamma}{\sigma}\right).$$

It is clear that when all the parameters are unknown, the ML estimators of the unknown parameters cannot be obtained in closed-form. But, for a given γ and σ, the ML estimator of α, say $\widehat{\alpha}(\gamma, \sigma)$, can be obtained as

$$\widehat{\alpha}(\gamma, \sigma) = \left[\frac{4}{n}\sum_{i=1}^{n} \sinh^2\left(\frac{y_i - \gamma}{\sigma}\right)\right]^{1/2}. \tag{73}$$

Then, the ML estimators of γ and σ can be obtained by maximizing the profile log-likelihood function of γ and σ, namely, $l(\widehat{\alpha}(\gamma, \sigma), \gamma, \sigma)$. It may be noted that for $\sigma = 2$, $\gamma = \ln(\beta)$ and $y_i = \ln x_i$, $\widehat{\alpha}$ obtained from (73) is the same as that obtained from (79) presented later in Section 13.1.

10.2 Bivariate Sinh-Normal distribution

Proceeding in the same way as we did with the univariate sinh-normal distribution, we can introduce bivariate sinh-normal distribution as follows. The random vector (Y_1, Y_2) is said to have a bivariate sinh-normal distribution, with parameters $\alpha_1 > 0$, $-\infty < \gamma_1 < \infty$, $\sigma_1 > 0$, $\alpha_2 > 0$, $-\infty < \gamma_2 < \infty$, $\sigma_2 > 0$, $-1 < \rho < 1$, if the joint CDF of (Y_1, Y_2) is

$$P(Y_1 \le y_1, Y_2 \le y_2) = F_{Y_1,Y_2}(y_1, y_2; \underline{\theta}) = \Phi_2\left\{\frac{2}{\alpha_1}\sinh\left(\frac{y_1 - \gamma_1}{\sigma_1}\right), \frac{2}{\alpha_2}\sinh\left(\frac{y_2 - \gamma_2}{\sigma_2}\right); \rho\right\}$$

for $-\infty < y_1,\ y_2 < \infty$. Here, $\underline{\theta} = (\alpha_1, \gamma_1, \sigma_1, \alpha_2, \gamma_2, \sigma_2, \rho)$ and $\Phi_2(u, v; \rho)$ is the joint cumulative distribution function of a standard bivariate normal vector (Z_1, Z_2) with correlation coefficient ρ. The joint PDF of Y_1 and Y_2 is

$$f_{Y_1,Y_2}(y_1, y_2; \underline{\theta}) \quad = \quad \frac{2}{\alpha_1\alpha_2\sigma_1\sigma_2\pi}\phi_2\left\{\frac{2}{\alpha_1}\sinh\left(\frac{y_1 - \gamma_1}{\sigma_1}\right), \frac{2}{\alpha_2}\sinh\left(\frac{y_2 - \gamma_2}{\sigma_2}\right); \rho\right\}$$

$$\cosh\left(\frac{y_1-\gamma_1}{\sigma_1}\right)\cosh\left(\frac{y_2-\gamma_2}{\sigma_2}\right), \tag{74}$$

where $\phi_2(u, v, \rho)$ is as in (56). The contour plots of (74), for different values of α_1 and α_2, can be found in Kundu [117] from which it is evident that it can take on different shapes. It is symmetric along the axis $x = y$, but it can be both unimodal and bimodal, and so it can be used effectively to analyze bivariate data. From here on, if a bivariate random vector (Y_1, Y_2) has the joint PDF as in (74), we will denote it by $\text{BSN}(\alpha_1, \gamma_1, \sigma_1, \alpha_2, \gamma_2, \sigma_2, \rho)$.

Some simple properties can be easily observed for the bivariate sinh-normal distribution. For example, if $(Y_1, Y_2) \sim \text{BSN}(\alpha_1, \gamma_1, \sigma_1, \alpha_2, \gamma_2, \sigma_2, \rho)$, then $Y_1 \sim \text{SN}(\alpha_1, \gamma_1, \sigma_1)$ and $Y_2 \sim \text{SN}(\alpha_2, \gamma_2, \sigma_2)$. The conditional PDF of Y_1, given Y_2, can be obtained easily as well. Note that by using an algorithm similar to the one described in Section 8.3, we can first generate (Z_1, Z_2) and then make the transformation

$$Y_i = \sigma_i \ \text{arcsinh}\left(\frac{\alpha_i Z_i}{2}\right) + \gamma_i, \quad i = 1, 2.$$

Then, $(Y_1, Y_2) \sim \text{BSN}(\alpha_1, \gamma_1, \sigma_1, \alpha_2, \gamma_2, \sigma_2, \rho)$, as required.

It may be easily seen that if $(X_1, X_2) \sim \text{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2, \rho)$, then $(Y_1, Y_2) = (\ln(X_1), \ln(X_2)) \sim \text{BSN}(\alpha_1, \ln(\beta_1), 2, \alpha_2, \ln(\beta_2), 2, \rho)$. It will be interesting to develop other properties of this distribution along with inferential procedures. Multivariate sinh distribution can also be defined in an analogous manner. Much work remains to be done in this direction.

10.3 Length-Biased BS distribution

The length-biased distribution plays an important role in various applications. Length-biased distribution is a special case of weighted distributions. The length-biased distribution of a random variable X can be described as follows. If Y is a positive random variable with PDF

$f_Y(\cdot)$ and $E(Y) = \mu < \infty$, then the corresponding length-biased random variable T has its PDF $f_T(\cdot)$ as

$$f_T(t) = \frac{t f_X(t)}{\mu}, \quad t > 0.$$

Therefore, it is clear that the length-biased version of a distribution does not introduce any extra parameter and has the same number of parameters as the original distribution.

The length-biased version of different distributions have found applications in diverse fields, such as biometry, ecology, reliability and survival analysis. A review of different length-biased distributions can be found in Gupta and Kirmani [101]. Patil [208] has provided several applications of length-biased distributions in environmental science.

Recently, Leiva et al. [142] studied the length-biased BS (LBS) distribution. If $Y \sim \text{BS}(\alpha, \beta)$, then the LBS of Y, say T, has its PDF as

$$f_T(t) = \phi\left(\frac{1}{\alpha}\left(\left[\frac{t}{\beta}\right]^{1/2} - \left[\frac{\beta}{y}\right]^{1/2}\right)\right) \times \frac{1}{(\alpha^3 + 2\alpha)\beta} \times \left(\left[\frac{t}{\beta}\right]^{1/2} + \left[\frac{\beta}{t}\right]^{1/2}\right), \quad t > 0.$$

From here on, this distribution will be denoted by $\text{LBS}(\alpha, \beta)$. Here also, β plays the role of the scale parameter while α plays the role of a shape parameter. The PDFs of $\text{LBS}(\alpha, 1)$, for different values of α, indicate that the PDF is unimodal. The mode can be obtained as a solution of the cubic equation

$$t^3 + \beta(1 - \alpha^2)t^2 - \beta^2(1 + \alpha^2)t - \beta^2 = 0.$$

Moreover, the HF of LBS is always unimodal just as in the case of the BS distribution. The moments of T depend on the moments of Y and they are related as follows:

$$\begin{aligned} \text{E}(T^r) \quad &= \quad \frac{1}{\alpha^2 + 2}\left\{\frac{2^{r+1}\Gamma(r + \frac{3}{2})\alpha^{2r+2}\beta^r}{\sqrt{\pi}} - \sum_{k=1}^{r+1}(-1)^k\binom{2r+1}{k}\frac{E(Y^{r-k+1})}{\beta^{1-k}}\right. \\ &\qquad \left. - \sum_{k=r+2}^{2r+1}(-1)^k\binom{2r+2}{k}\frac{E(Y^{k-r-1})}{\beta^{k-2r-1}}\right\}. \end{aligned}$$

We shall now briefly discuss the estimation of the model parameters α and β based on a sample of size n, say $\{t_1, \cdots, t_n\}$, from a LBS(α, β). The log-likelihood function of the data, without the additive constant, is given by

$$l(\alpha, \beta|\text{data}) = -\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right) - n\ln(2\alpha + \alpha^2) - \frac{3}{2}n\ln(\beta) + \sum_{i=1}^{n}\ln(t_i + \beta). \quad (75)$$

The ML estimators of α and β can then be obtained by maximizing the log-likelihood function in (75) with respect to α and β. Since they do not have explicit forms, they need to be obtained numerically. For constructing confidence intervals of the parameters, since the exact distributions of the ML estimators are not possible to obtain, the asymptotic distribution of the ML estimators can be used. It has been shown by Leiva et al. [142] that $(\widehat{\alpha}, \widehat{\beta})$, the ML estimator of (α, β), is asymptotically bivariate normal with mean (α, β) and variance-covariance matrix $\widehat{\Sigma}$, where the elements of $\widehat{\Sigma}^{-1}$ are as follows:

$$\widehat{\Sigma}^{-1} = -\begin{bmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{bmatrix},$$

with

$$\begin{aligned}
\sigma^{11} &= \sum_{i=1}^{n}\frac{3\left(2 - \frac{t_i}{\beta} - \frac{\beta}{t_i}\right)}{\alpha^4} - \frac{n(4 + 3\alpha^4)}{(2\alpha + \alpha^3)^2}, \\
\sigma^{12} &= \sigma^{21} = \sum_{i=1}^{n}\frac{\left(\frac{1}{t_i} - \frac{t_i}{\beta^2}\right)}{\alpha^3}, \\
\sigma^{22} &= \frac{3n}{2\beta^2} - \sum_{i=1}^{n}\left\{\frac{t_i}{\alpha^2\beta^3} - \frac{1}{(t_i + \beta)^2}\right\}.
\end{aligned}$$

The asymptotic distribution of the ML estimators can be used for constructing approximate confidence intervals. Alternatively, bootstrap confidence intervals can also be constructed and this may be especially suitable for small sample sizes.

10.4 Epsilon BS Distribution

Mudholkar and Hutson [194] introduced the epsilon skew-normal distribution defined on the entire real line as follows. A random variable X is said to have an epsilon skew-normal distribution if it has the PDF

$$g(x;\epsilon) = \begin{cases} \phi\left(\frac{x}{(1+\epsilon)}\right) & \text{if} \quad x < 0, \\ \phi\left(\frac{x}{(1-\epsilon)}\right) & \text{if} \quad x \geq 0, \end{cases}$$

for $-1 \leq \epsilon \leq 1$. It is clear that for $\epsilon = 0$, the epsilon skew-normal distribution becomes the standard normal distribution. For $\epsilon = 1$, it is a negative half normal distribution and for $\epsilon = -1$, it is a positive half normal distribution.

Arellano-Valle et al. [8] generalized the epsilon skew-normal distribution to any symmetric distribution as follows. A random variable X is said to have a epsilon skew-symmetric distribution if it has the density function

$$h(x;\epsilon) = \begin{cases} f\left(\frac{x}{(1+\epsilon)}\right) & \text{if} \quad x < 0, \\ f\left(\frac{x}{(1-\epsilon)}\right) & \text{if} \quad x \geq 0, \end{cases}$$

where $f(\cdot)$ is any symmetric density function and $-1 < \epsilon < 1$. It will be denoted by $X \sim \text{ES}f(\epsilon)$.

Using the same transformation as in (5), Castillo et al. [49], see also Vilca et al. [259] in this aspect, defined the epsilon generalized-BS distribution as follows. A random variable W is said to have an epsilon generalized-BS distribution, with parameters $\alpha > 0$, $\beta > 0$, $-1 < \epsilon < 1$, if

$$W = \frac{\beta}{4}\left[\alpha X + \sqrt{\{\alpha X\}^2 + 4}\right]^2,$$

where $X \sim \text{ES}f(\epsilon)$. This distribution will be denoted by EGBS(α,β,ϵ). If $W \sim$ EGBS(α,β,ϵ), then the PDF of W is

$$f_W(w|\alpha,\beta,\epsilon) = \frac{w^{-3/2}(w+\beta)}{2\alpha\sqrt{\beta}} \times \begin{cases} f\left(\frac{t(w)}{1+\epsilon}\right) & \text{if} \quad w < \beta, \\ f\left(\frac{t(w)}{1-\epsilon}\right) & \text{if} \quad w \geq \beta, \end{cases}$$

where

$$t(w) = \frac{1}{\alpha}\left(\sqrt{\frac{w}{\beta}} - \sqrt{\frac{\beta}{w}}\right).$$

These authors then discussed different properties of the epsilon generalized-BS distribution. For example, it can be easily seen that if $W \sim \text{EGBS}(\alpha, \beta, \epsilon)$, then (a) $aW \sim \text{EGBS}(\alpha, a\beta, \epsilon)$, for $a > 0$, and (b) $W^{-1} \sim \text{EGBS}(\alpha, \beta^{-1}, \epsilon)$. Moreover, the moments of W, for $k = 1, 2, \cdots$, can be expressed as follows :

$$\text{E}(W^k) = \beta^k \sum_{r=0}^{2k} \left(\frac{\alpha}{2}\right)^{2k-r} \binom{2k}{r} \text{E}\left(X^{2k-r}\left((\alpha X/2)^2 + 1\right)^{r/2}\right).$$

When $f(\cdot) = \phi(\cdot)$, the standard normal PDF, the epsilon generalized-BS distribution is known as the epsilon BS distribution. Castillo et al. [49] derived the first and second moments explicitly in terms of incomplete moments of the standard normal distribution. These authors also discussed the estimation of the model parameters of epsilon BS distribution. The ML estimators are obtained by solving three non-linear equations, and explicit expressions of the elements of the Fisher information matrix have also been provided.

11 Log-BS Regression

In this section we introduce log-BS regression model which has been used quite extensively in practice in recent years. First, we provide the basic formulation of the model, and then we discuss about different inferential issues related to this model.

11.1 Basic Formulation

Rieck and Nedelman [226] first introduced the log-BS regression model as follows. Let $T_1, \cdots, T_n$ be n independent random variables, and the distribution of T_i be a BS distribution with shape parameter α_i and scale parameter β_i. Let us further assume that the distribution of T_i depends on a set of p explanatory variables $\underline{x}_i = (x_{i1}, \cdots, x_{ip})$ as follows:

1. $\beta_i = \exp(\underline{x}_i^\top \underline{\theta})$, for $i = 1, \cdots, n$, $\underline{\theta} = (\theta_1, \cdots, \theta_p)$ is a set of p unknown parameters to be estimated;

2. The shape parameter is independent of the explanatory vector $\underline{x}_i$, and they are constant, i.e. $\alpha_1 = \cdots = \alpha_n = \alpha$.

Now, from the properties of the BS distribution, it immediately follows that $T_i = \exp(\underline{x}_i^\top \underline{\theta}) U_i$, where U_i is distributed as BS with scale parameter one and shape parameter α. Therefore, if we denote $Y_i = \ln(T_i)$, for $i = 1, \cdots, n$, then

$$Y_i = \underline{x}_i^\top \underline{\theta} + \epsilon_i, \qquad i = 1, \cdots, n. \tag{76}$$

Here, $\epsilon_i = \ln(U_i)$ has a log-BS or sinh distribution, *i.e.*, $\text{SN}(\alpha, 0, 2)$. Because of the form, the model in (76) is known as the log-BS regression model. Here also, the problem is same as the standard regression problem, *i.e.*, to estimate the unknown parameters, namely $\underline{\theta}$ and α, based on the sample t_i and the associated covariates $\underline{x}_i$, for $i = 1, \cdots, n$.

11.2 ML Estimators

Based on independent observations $y_1, \cdots, y_n$, from the regression model in (76), the log-likelihood function, without the additive constant, is given by [see Rieck and Nedelman [226]]

$$l(\underline{\theta}, \alpha | \text{data}) = \sum_{i=1}^{n} \ln(w_i) - \frac{1}{2} \sum_{i=1}^{n} z_i^2, \tag{77}$$

where

$$w_i = \frac{2}{\alpha} \cosh\left(\frac{y_i - \underline{x}_i^\top \underline{\theta}}{2}\right) \quad \text{and} \quad z_i = \frac{2}{\alpha} \sinh\left(\frac{y_i - \underline{x}_i^\top \underline{\theta}}{2}\right).$$

The ML estimators of α and $\underline{\theta}$ can be obtained by maximizing (77) with respect to the parameters. The normal equations are

$$\frac{\partial l(\underline{\theta}, \alpha | \text{data})}{\partial \theta_j} = \left(\frac{1}{2}\right) \sum_{i=1}^{n} x_{ij} \left(z_i w_i - \frac{z_i}{w_i}\right) = 0, \tag{78}$$

$$\frac{\partial l(\underline{\theta}, \alpha | \text{data})}{\partial \alpha} = -\frac{n}{\alpha} + \left(\frac{1}{\alpha}\right) \sum_{i=1}^{n} z_i^2 = 0. \tag{79}$$

From (79), it is clear that, for given $\underline{\theta}$, the ML estimator of α can be obtained as

$$\widehat{\alpha}(\theta) = \left\{ \frac{4}{n} \sum_{i=1}^{n} \sinh^2 \left(\frac{y_i - \underline{x}_i^\top \underline{\theta}}{2} \right) \right\}^{1/2},$$

and the ML estimators of θ_j's can then be obtained from (78), or by maximizing the profile log-likelihood function of $\underline{\theta}$, *i.e.*, $l(\underline{\theta}, \widehat{\alpha}(\underline{\theta})|\text{data})$, with respect to $\underline{\theta}$. No explicit solutions are available in this case, and so suitable numerical techniques like Newton-Raphson or the Fisher's scoring method may be applied to determine the ML estimates. Rieck [221] compared, by extensive Monte Carlo simulation studies, the performances of the Newton-Raphson and the Fisher's scoring algorithm, and observed that the Newton-Raphson method performed better than the Fisher's scoring algorithm particularly for small sample sizes and when the number of parameters is small.

It has been observed that the likelihood function may have more than one maxima mainly due to the fact that sinh-normal is bimodal for $\alpha > 2$. It has been shown by Rieck [221] that if $\alpha < 2$ is known, then the ML estimator of $\underline{\theta}$ exists and is unique. Under the assumptions that $|x_{ij}|$ are uniformly bounded and as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} \underline{x}_i \underline{x}_i^\top = \mathbf{A} > 0,$$

Rieck and Nedelman [225] showed that there exists at least one solution of the normal equations that is a consistent estimate of the true parameter vector $(\underline{\theta}^\top, \alpha)$. Moreover, among all possible solutions of the normal equations, one and only one tends to $(\underline{\theta}^\top, \alpha)$ with probability one. They also proved that $\sqrt{n}((\widehat{\underline{\theta}} - \underline{\theta})^\top, \widehat{\alpha} - \alpha)$ converges in distribution to a multivariate normal distribution with mean vector $\underline{0}$, and the asymptotic variance-covariance matrix Σ, where

$$\Sigma = \begin{bmatrix} \frac{4\mathbf{A}^{-1}}{C(\alpha)} & \underline{0} \\ \underline{0}^\top & \frac{\alpha^2}{2} \end{bmatrix},$$

with

$$C(\alpha) = 2 + \frac{4}{\alpha^2} - \left(\frac{2\pi}{\alpha^2} \right)^{1/2} \left\{ 1 - \text{erf}[(2/\alpha^2)^{1/2}] \right\} \exp(2/\alpha^2)$$

and $\text{erf}(x) = 2\Phi(\sqrt{2}x) - 1$; see Rieck and Nedelman [225] for details. Thus, it is immediate that if $\widehat{\alpha}$ is the ML estimator of α, then the asymptotic variance-covariance matrix of $(\widehat{\underline{\theta}}^\top, \widehat{\alpha})$ can be estimated by

$$\begin{bmatrix} \frac{4}{C(\widehat{\alpha})} \sum_{i=1}^{n} \underline{x}_i \underline{x}_I^\top & \mathbf{0} \\ \mathbf{0}^\top & \frac{\widehat{\alpha}^2}{2n} \end{bmatrix},$$

which can be used for constructing confidence intervals.

Although the ML estimator of $\underline{\theta}$ cannot be obtained in closed-form, the least squares estimator of $\underline{\theta}$ can be obtained in closed-form as

$$\widehat{\underline{\theta}} = (\underline{X}^\top \underline{X})^{-1} \underline{X}^\top \underline{Y},$$

where $\underline{Y}$ is the column vector of the y_i's and $\boldsymbol{X}^\top = (\underline{x}_1, \cdots, \underline{x}_n)$. But, it has been shown, using Monte Carlo simulations, by Rieck and Nedelman [226] that the LSE of $\underline{\theta}$ is not as efficient as the ML estimator of $\underline{\theta}$. Yet, it is an unbiased estimator of $\underline{\theta}$ and is quite efficient for small values of α.

11.3 Testing of Hypotheses

Lemonte et al. [169] considered the following testing problem for log-BS model:

$$\text{H}_{a0} : \beta_r = \beta_r^0 \quad vs. \quad \text{H}_{a1} : \beta_r \neq \beta_r^0,$$

$$\text{H}_{b0} : \beta_r \geq \beta_r^0 \quad vs. \quad \text{H}_{b1} : \beta_r < \beta_r^0,$$

$$\text{H}_{c0} : \beta_r \leq \beta_r^0 \quad vs. \quad \text{H}_{c1} : \beta_r > \beta_r^0.$$

They then discussed test procedures for the above hypotheses and their performance characteristics.

12 Generalized BS Distribution

The univariate, bivariate and multivariate BS distributions have been derived by making suitable transformations on the univariate, bivariate and multivariate normal distributions, respectively. Attempts have been made to generalize BS distributions by replacing normal distribution with the elliptical distribution. In this section, we introduce univariate, multivariate and matrix-variate generalized BS distributions, discuss their properties and also address some inferential issues.

12.1 Univariate Generalized BS Distribution

Díaz-García and Leiva-Sánchez [69] generalized the univariate BS distribution by using an elliptical distribution in place of normal distribution. A random variable X is said to have an elliptical distribution if the PDF of X is given by

$$f_X(x) = cg\left[\frac{(x-\mu)^2}{\sigma^2}\right], \quad x \in \mathbb{R}, \tag{80}$$

where $g(u)$, with $u > 0$, is a real-valued function and corresponds to the kernel of the PDF of X, and c is the normalizing constant, such that $f_X(x)$ is a valid PDF. If the random variable X has the PDF in (80), it will be denoted by $\text{EC}(\mu, \sigma^2, g)$. Here, μ and σ are the location and scale parameters, respectively, and g is the corresponding kernel. Several well-known distributions such as the normal distribution, t-distribution, Cauchy distribution, Pearson VII distribution, Laplace distribution, Kotz type distribution and logistic distribution are all members of the family of elliptically contoured distributions.

Based on the elliptically contoured distribution, Díaz-García and Leiva-Sánchez [69] defined the generalized BS (GBS) distribution as follows. The random variable T is said to have a GBS distribution, with parameters α, β and kernel g, if

$$U = \frac{1}{\alpha}\left[\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}}\right] \sim \mathrm{EC}(0, 1, g),$$

and is denoted by $T \sim \mathrm{GBS}(\alpha, \beta, g)$. It can be shown, using the standard transformation technique, that if $T \sim \mathrm{GBS}(\alpha, \beta, g)$, then the PDF of T is as follows:

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

where c is the normalizing constant as mentioned in (80). For different special cases, c can be explicitly obtained and these can be found in Díaz-García and Leiva-Sánchez [69]. These authors have derived the moments of the GBS distribution in terms of the moments of the corresponding elliptically contoured distribution. Inferential procedures have not yet been developed in general, but only in some special cases. It will, therefore, be of interest to develop efficient inferential procedures for different kernel functions. In fact, in practice, choosing a proper kernel is quite an important issue that has not been addressed in detail.

It is important to mention here that Fang and Balakrishnan [76] recently extended some of the stochastic ordering results for the univariate BS model detailed earlier in Section 8 to the case of generalized BS models with associated random shocks. There seems to be a lot of potential for further exploration in this direction!

12.2 Multivariate Generalized BS Distribution

Along the same lines as the univariate GBS distribution, the multivariate GBS distribution can be defined. First, let us recall [see Fang and Zhang [79]] the definition of the multivariate elliptically symmetric distribution. A p-dimensional random vector $\underline{X}$ is said to have an elliptically symmetric distribution with p-dimensional location vector $\underline{\mu}$, a $p \times p$ positive definite dispersion matrix $\boldsymbol{\Sigma}$ and the density generator (kernel) $h^{(p)} : \mathbb{R}_+ \to \mathbb{R}_+$, if the PDF of $\underline{X}$ is of the form

$$f_{EC_p}(\underline{x};\underline{\mu},\mathbf{\Sigma},h^{(p)}) = |\mathbf{\Sigma}|^{-1/2}h^{(p)}(w(\underline{x}-\underline{\mu})^\top\mathbf{\Sigma}^{-1}(\underline{x}-\underline{\mu})), \tag{81}$$

where $w(\underline{x}):\mathbb{R}^p \to \mathbb{R}_+$ with $w(\underline{x}) = (\underline{x}-\underline{\mu})^\top\mathbf{\Sigma}^{-1}(\underline{x}-\underline{\mu})$, $h:\mathbb{R}_+\to\mathbb{R}_+$, and

$$\int_{\mathbb{R}^p} f_{EC_p}(\underline{x};\underline{\mu},\mathbf{\Sigma},h^{(p)})\mathrm{d}\underline{x} = 1.$$

In what follows, we shall denote $f_{EC_p}(\underline{x};\underline{0},\mathbf{\Sigma},h^{(p)})$ simply by $f_{EC_p}(\underline{x};\mathbf{\Sigma},h^{(p)})$ and the corresponding random vector by $\mathrm{EC}_p(\mathbf{\Sigma},h^{(p)})$.

Based on the elliptically symmetric distribution, Kundu et al. [120] introduced the generalized multivariate BS distribution as follows.

Definition: Let $\underline{\alpha},\underline{\beta}\in\mathbb{R}^p$, where $\underline{\alpha} = (\alpha_1,\cdots,\alpha_p)^\top$ and $\underline{\beta} = (\beta_1,\cdots,\beta_p)^\top$, with $\alpha_i > 0, \beta_i > 0$, for $i = 1,\cdots,p$. Let $\mathbf{\Gamma}$ be a $p\times p$ positive-definite correlation matrix. Then, the random vector $\underline{T} = (T_1,\cdots,T_p)^\top$ is said to have a generalized multivariate BS distribution, with parameters $(\underline{\alpha},\underline{\beta},\mathbf{\Gamma})$ and density generator $h^{(p)}$, denoted by $\underline{T}\sim \mathrm{GBS}_p(\underline{\alpha},\underline{\beta},\mathbf{\Gamma},h^{(p)})$, if the CDF of $\underline{T}$, i.e., $P(\underline{T}\le\underline{t}) = P(T_1\le t_1,\cdots,T_p\le t_p)$, is given by

$$P(\underline{T}\le\underline{t}) = F_{EC_p}\left[\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}}-\sqrt{\frac{\beta_1}{t_1}}\right),\cdots,\frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}}-\sqrt{\frac{\beta_p}{t_p}}\right);\mathbf{\Gamma},h^{(p)}\right]$$

for $\underline{t} > \underline{0}$, where $F_{EC_p}(\cdot;\mathbf{\Gamma},h^{(p)})$ denotes the CDF of $EC_p(\mathbf{\Gamma},h)$. The corresponding joint PDF of $\underline{T} = (T_1,\cdots,T_p)^\top$ is, for $\underline{t} > \underline{0}$,

$$\begin{aligned} f_{\underline{T}}(\underline{t};\underline{\alpha},\underline{\beta},\mathbf{\Gamma}) &= f_{EC_p}\left(\frac{1}{\alpha_1}\left(\sqrt{\frac{t_1}{\beta_1}}-\sqrt{\frac{\beta_1}{t_1}}\right),\cdots,\frac{1}{\alpha_p}\left(\sqrt{\frac{t_p}{\beta_p}}-\sqrt{\frac{\beta_p}{t_p}}\right);\mathbf{\Gamma},h^{(p)}\right) \\ &\quad\times\prod_{i=1}^{p}\frac{1}{2\alpha_i\beta_i}\left\{\left(\frac{\beta_i}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta_i}{t_i}\right)^{\frac{3}{2}}\right\}, \end{aligned}$$

where $f_{EC_p}(\cdot)$ is as given in (81).

The density generator $h^{(p)}(\cdot)$ can take on different forms resulting in multivariate normal distribution, symmetric Kotz type distribution, multivariate t-distribution, symmetric multivariate Pearson type VII distribution, and so on. These authors have then discussed different properties of the generalized multivariate BS distribution including the distributions of the marginals, distributions of reciprocals, total positivity of order two property etc. Special attention has been paid to two particular cases, namely, (i) multivariate normal and (ii) multivariate t. Generation of random samples and different inferential issues have also been discussed by these authors in detail.

Recently, Fang et al. [78] proceeded along the lines of Section 8 and discussed stochastic comparisons of minima and maxima arising from independent and non-identically distributed bivariate BS random vectors with respect to the usual stochastic order. Specifically, let $(X_1, X_2) \sim \mathrm{BS}_2(\alpha_1, \beta_1, \alpha_2, \beta_2, \rho)$ and $(X_1^*, X_2^*) \sim \mathrm{BS}_2(\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*, \rho)$ independently, and let further $0 < v \leq 2$. Then, Fang et al. [78] established that when $\alpha_1 = \alpha_2 = \alpha_1^* = \alpha_2^*$, $(\beta_1^{-1/v}, \beta_2^{-1/v}) \geq_m (\beta_1^{*-1/v}, \beta_2^{*-1/v})$ implies $X_{2:2} \geq_{st} X_{2:2}^*$ and $(\beta_1^{1/v}, \beta_2^{1/v}) \geq_m (\beta_1^{*1/v}, \beta_2^{*1/v})$ implies $X_{1:2}^* \geq_{st} X_{1:2}$. In a similar manner, they also proved that when $\beta_1 = \beta_2 = \beta_1^* = \beta_2^*$, $(1/\alpha_1, 1/\alpha_2) \geq_m (1/\alpha_1^*, 1/\alpha_2^*)$ implies $X_{2:2} \geq_{st} X_{2:2}^*$ and $X_{1:2}^* \geq_{st} X_{1:2}$. Generalizations of these results to bivariate generalized BS distributions, analogous to those mentioned in the last subsection, remain open!

12.3 Matrix-Variate Generalized BS Distribution

We have already detailed the univariate and multivariate generalized BS distributions, and they are natural extensions of the univariate and multivariate BS distributions. Along the same lines, Caro-Lopera et al. [48] defined the matrix-variate generalized BS distribution using an elliptic random matrix. First, we introduce a matrix-variate elliptic distribution. Let $\boldsymbol{X} = (X_{ij})$ be an $n \times k$ random matrix. It is said to have a matrix-variate elliptic distribution with location matrix $\boldsymbol{M} \in \mathbb{R}^{n\times k}$, scale matrices $\boldsymbol{\Omega} \in \mathbb{R}^{k\times k}$ with $\mathrm{rank}(\boldsymbol{\Omega}) = k$, and $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ with rank($\boldsymbol{\Sigma}$) $= n$, and a density generator g, if the PDF of $\boldsymbol{X}$ is

$$f_{\boldsymbol{X}}(\boldsymbol{X}) = c|\boldsymbol{\Omega}|^{-n/2}|\boldsymbol{\Sigma}|^{-k/2}g(\mathrm{tr}(\boldsymbol{\Omega}^{-1}(\boldsymbol{X}-\boldsymbol{M})^{top}\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{M}))), \quad \boldsymbol{X} \in \mathbb{R}^{n\times k},$$

where c is the normalizing constant for the density generator g. It will be denoted by $\mathrm{EC}_{n\times k}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}; g)$. Now, we are in a position to define the matrix-variate generalized BS distribution.

Definition: Let $\boldsymbol{Z} = (Z_{ij}) \sim \mathrm{EC}_{n\times k}(\boldsymbol{0}, \boldsymbol{I}_k, \boldsymbol{I}_n; g)$ and $\boldsymbol{T} = (T_{ij})$, where

$$T_{ij} = \beta_{ij}\left[\frac{\alpha_{ij}Z_{ij}}{2} + \sqrt{\left\{\frac{\alpha_{ij}Z_{ij}}{2}\right\}^2 + 1}\right]^2, \quad \alpha_{ij} > 0,\ \beta_{ij} > 0,\ i = 1,\ldots,n, j = 1,\ldots,k.$$

Then, the random matrix $\boldsymbol{T}$ is said to have a generalized matrix-variate BS distribution, denoted by $\mathrm{GBS}_{n\times k}(\boldsymbol{A}, \boldsymbol{B}, g)$, where $\boldsymbol{A} = (\alpha_{ij})$ and $\boldsymbol{B} = (\beta_{ij})$. It has been shown by Caro-Lopera et al. [48] that if $T \sim \mathrm{GBS}_{n\times k}(\boldsymbol{A}, \boldsymbol{B}, g)$, then the PDF of $\boldsymbol{T}$ is given by

$$f_{\boldsymbol{T}}(\boldsymbol{T}) = \frac{c}{2^{n+k}}g\left(\sum_{i=1}^{n}\sum_{j=1}^{k}\frac{1}{\alpha_{ij}^2}\left[\frac{T_{ij}}{\beta_{ij}} + \frac{\beta_{ij}}{T_{ij}} - 2\right]\right)\prod_{i=1}^{n}\prod_{j=1}^{k}\frac{T_{ij}^{-3/2}[T_{ij}+\beta_{ij}]}{\alpha_{ij}\sqrt{\beta_{ij}}}, \quad T_{ij} > 0,$$

for $i = 1,\ldots,n$ and $j = 1,\ldots,k$.

Note that if we take

$$g(u) = \exp\left\{-\frac{u}{2}\right\}, \quad u > 0, \quad \text{and} \quad c = \frac{2^{nk/2}}{\pi^{nk/2}},$$

then we obtain the matrix-variate BS distribution. Therefore, a random matrix $\boldsymbol{T} \in \mathbb{R}^{n\times k}$ is said to have a matrix-variate BS distribution if the PDF of $\boldsymbol{T}$ is

$$f_{\boldsymbol{T}}(\boldsymbol{T}) = \frac{c}{2^{n+k}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{k}\frac{1}{\alpha_{ij}^2}\left[\frac{T_{ij}}{\beta_{ij}} + \frac{\beta_{ij}}{T_{ij}} - 2\right]\right)\prod_{i=1}^{n}\prod_{j=1}^{k}\frac{T_{ij}^{-3/2}[T_{ij}+\beta_{ij}]}{\alpha_{ij}\sqrt{\beta_{ij}}}, \quad T_{ij} > 0,$$

for $i = 1,\ldots,n$ and $j = 1,\ldots,k$.

These authors have provided an alternate representation of the PDF of the matrix-variate generalized BS distribution in terms of Hadamard matrix products, and also indicated some open problems. Interested readers may refer to their article. An important open problem is the inference for the model parameters in this case.

13 Illustrative Examples

In this section, we present some numerical examples in order to illustrate some of the inferential results described in the preceding sections as well as the usefulness of the models introduced.

13.1 Example 1: Fatigue Data

This data set was originally analyzed by Birnbaum and Saunders [44] and it represents the fatigue life of 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per seconds (cps). It consists of 101 observations with maximum stress per cycle 31,000 psi. The data are as follows:

70 90 96 97 99 100 103 104 104 105 107 108 108 108 109 109 112 112 113 114 114 114 116 119 120 120 120 121 121 123 124 124 124 124 124 128 128 129 129 130 130 130 131 131 131 131 131 132 132 132 133 134 134 134 134 134 136 136 137 138 138 138 139 139 141 141 142 142 142 142 142 142 144 144 145 146 148 148 149 151 151 152 155 156 157 157 157 157 158 159 162 163 163 164 166 166 168 170 174 196 212

In summary, we have in this case $n = 101$, $s = (1/n)\sum_{i=1}^{n} t_i = 133.73267$, and $r = n/\sum_{i=1}^{n} t_i^{-1}$ $= 129.93321$. The ML, MM, UML and UMM estimates are presented in Table 1; see Ng et al. [195]. The associated 90% and 95% confidence intervals are presented in Table 2.

Table 1: Point estimates of α and β for Example 1

Estimate	α	β
ML	0.170385	131.818792
MM	0.170385	131.819255
UML	0.172089	131.809130
UMM	0.172089	131.809593

Table 2: 90% and 95% confidence intervals for α and β for Example 1

Estimate	α		β	
	90%	95%	90%	95%
ML	(0.1527,0.1927)	(0.1497,0.1976)	(128.2552,135.5861)	(127.5944,136.3325)
MM	(0.1527,0.1927)	(0.1497,0.1976)	(128.2556,135.5866)	(127.5948,136.3330)
UML	(0.1541,0.1949)	(0.1511,0.1999)	(128.2116,135.6143)	(127.5448,136.3685)
UMM	(0.1541,0.1949)	(0.1511,0.1999)	(128.2121,135.6148)	(127.5452,136.3690)

13.2 Example 2: Insurance Data

The following data represent Swedish third party motor insurance for 1977 for one of several geographical zones. The data were compiled by a Swedish committee on the analysis of risk premium in motor insurance. The data points are the aggregate payments by the insurer in thousand Skr (Swedish currency). The data set was originally reported in Andrews and Herzberg [5], and is as as follows:

5014 5855 6486 6540 6656 6656 7212 7541 7558 7797 8546 9345 11762 12478 13624 14451 14940 14963 15092 16203 16229 16730 18027 18343 19365 21782 24248 29069 34267 38993

We would like to see whether the two-parameter BS distribution fits these data well or not. We divide all the data points by 10000 and obtain the ML, MM, UML and UMM estimates, and these results are presented in Table 3. The associated 90% and 95% confidence

Table 3: Point estimates of α and β for Example 2

Estimate	α	β
ML	0.559551	1.255955
MM	0.559551	1.256602
UML	0.540899	1.253993
UMM	0.540899	1.253346

intervals are presented in Table 4. Now to check whether the BS distribution fits the data or not, we have computed the Kolmogorov-Smirnov (KS) distance between the fitted CDF

Table 4: 90% and 95% confidence intervals for α and β for Example 2

Estimate	α		β	
	90%	95%	90%	95%
ML	(0.4407,0.6783)	(0.4176,0.7014)	(0.9854,1.5278)	(0.9326,1.5806)
MM	(0.4407,0.6783)	(0.4176,0.7014)	(0.9848,1.5271)	(0.9321,1.5798)
UML	(0.4221,0.6597)	(0.3989,0.6828)	(0.9827,1.5252)	(0.9299,1.5780)
UMM	(0.4220,0.6597)	(0.3989,0.6828)	(0.9822,1.5244)	(0.9295,1.5771)

based on ML, MM, UML and UMM estimates and the empirical CDF. We have reported these KS distances and the associated p-values in Table 5. It is clear from the results in Table 5 that the BS distribution fits the insurance data quite well.

Table 5: The KS distances between the fitted CDF and the empirical CDF for Example 2

Estimate	KS Distance	p-value
ML	0.1385	0.6130
MM	0.1387	0.6106
UML	0.1457	0.5470
UMM	0.1455	0.5494

13.3 Example 3: Ball Bearings Data

The following data set is from McCool [190], and it provides the fatigue life in hours of ten ball bearings of a certain type:

152.7 172.0 172.5 173.3 193.0 204.7 216.5 234.9 262.6 422.6

Cohen et al. [56] first used this data set as an illustrative example for the fit of a three-parameter Weibull distribution. Ng et al. [196] used the first 8 order statistics and fitted the BS distribution, based on the assumption that it is a Type-II right censored sample with $n = 10$ and $r = 8$. The ML estimates of α and β are found to be 0.1792 and 200.7262, respectively. The biased-corrected estimate of α turns out to be 0.2108. The 90% and 95%

confidence intervals for α and β based on ML and UML estimates are presented in Table 6.

Table 6: 90% and 95% confidence intervals for α and β for Example 3

Estimate	α		β	
	90%	95%	90%	95%
ML	(0.1017,0.2566)	(0.0868,0.2715)	(183.1828,221.9857)	(180.1662,226.5831)
UML	(0.0925,0.3290)	(0.0698,0.3517)	(180.4109,226.1973)	(176.9795,231.8331)

13.4 Example 4: Progressively Censored Data

Pradhan and Kundu [214] considered the same data set as in Example 3, and generated three different Type-II progressively censored samples as follows:

MCS-1: $n = 10$, $m = 6$, $R_1 = 4$, $R_2 = \cdots = R_6 = 0$;

MCS-2: $n = 10$, $m = 6$, $R_1 = \cdots = R_5 = 0$, $R_6 = 4$;

MCS-3: $n = 10$, $m = 7$, $R_1 = 2$, $R_2 = 1$, $R_3 \cdots = R_7 = 0$.

The ML estimates of α and β were obtained using the EM algorithm. The number of iterations needed for the convergence of the EM algorithm for the three schemes are 30, 44 and 24, respectively. The ML estimates of α and β, along with their standard errors, and associated 95% confidence intervals are presented in Table 7.

Table 7: Estimates of α and β, along with their standard errors, and 95% confidence intervals for the data set of McCool [190]

Censoring	α			β		
schemes	Estimate	s.e.	95% CI	Estimate	s.e.	95% CI
MCS-1	0.1639	0.0367	[0.0921, 0.2358]	194.0795	9.9935	[174.4922, 213.6669]
MCS-2	0.1484	0.0332	[0.0833, 0.2134]	195.4253	9.1186	[177.5528, 213.2978]
MCS-3	0.1570	0.0351	[0.0882, 0.2258]	195.8228	9.6617	[176.8859, 214.7597]

13.5 Example 5: Bivariate Bone Mineral Data

We now provide the data analysis of a bivariate data set. The data in this case, obtained from Johnson and Wichern [110], represent the bone mineral density (BMD) measured in gm/cm^2 for 24 individuals. The first figure represents the BMD of the bone Dominant Radius before starting the study and the second figure represents the BMD of the same bone after one year:

(1.103 1.027), (0.842 0.857), (0.925 0.875), (0.857 0.873), (0.795 0.811), (0.787 0.640), (0.933 0.947), (0.799 0.886), (0.945 0.991), (0.921 0.977), (0.792 0.825), (0.815 0.851), (0.755 0.770), (0.880 0.912), (0.900 0.905), (0.764 0.756), (0.733 0.765), (0.932 0.932), (0.856 0.843), (0.890 0.879), (0.688 0.673), (0.940 0.949), (0.493 0.463), (0.835 0.776).

The sample means and sample variances of the two components are (0.8408, 0.8410) and (0.0128, 0.0149), respectively. The sample correlation coefficient is 0.9222. Kundu et al. [119] used the BVBS distribution to model these bivariate data. From the observations, it is observed that $s_1 = 0.8408$, $s_2 = 0.8410$, $r_1 = 0.8225$, $r_2 = 0.8179$, and so the modified moment estimates are

$$\widetilde{\alpha}_1 = 0.1491, \quad \widetilde{\alpha}_2 = 0.1674, \quad \widetilde{\beta}_1 = 0.8316, \quad \widetilde{\beta}_2 = 0.8294, \quad \widetilde{\rho} = 0.9343.$$

Using these as initial values, the ML estimates are determined as

$$\widehat{\alpha}_1 = 0.1491, \quad \widehat{\alpha}_2 = 0.1674, \quad \widehat{\beta}_1 = 0.8312, \quad \widehat{\beta}_2 = 0.8292, \quad \widetilde{\rho} = 0.9343.$$

The 95% confidence intervals for α_1, α_2, β_1, β_2 and ρ, based on the empirical Fisher information matrix, become (0.1069, 0.1913), (0.1200, 0.2148), (0.7818, 0.8806), (0.7739, 0.8845), (0.8885, 0.9801), respectively.

Table 8: The sample mean, variance and coefficient of skewness of T_i and T_i^{-1}, for $i = 1, \cdots, 4$, for Example 6.

Variables → Statistics ↓	T_1	T_1^{-1}	T_2	T_2^{-1}	T_3	T_3^{-1}	T_4	T_4^{-1}
Mean	0.844	1.211	0.818	1.245	0.704	1.452	0.694	1.474
Variance	0.012	0.041	0.011	0.034	0.011	0.050	0.010	0.052
Skewness	-0.793	2.468	-0.543	1.679	-0.022	0.381	-0.133	0.755

13.6 Example 6: Multivariate Bone Mineral Data

We now provide the analysis of a multivariate data, taken from Johnson and Wichern ([110], page 34), representing the mineral contents of four major bones of 25 new born babies. Here, T_1, T_2, T_3 and T_4 represent dominant radius, radius, dominant ulna and ulna, respectively. The data are not presented here, but the sample mean, variance and skewness of the individual T_i's and their reciprocals are all presented in Table 8.

Kundu et al. [120] fitted a 4-variate BS distribution to this data set. First, the MM estimates are obtained from the marginals and they are provided in Table 9. Using MM

Table 9: The MM estimates of α_i and β_i, the KS distance between the empirical distribution function and the fitted distribution function, and the corresponding p values.

	α	β distance	KS	p
T_1	0.1473	0.8347	0.161	0.537
T_2	0.1372	0.8107	0.145	0.671
T_3	0.1525	0.6963	0.109	0.929
T_4	0.1503	0.6861	0.094	0.979

estimates as initial guess, the ML estimates of β_1, β_2, β_3 and β_4 are obtained as 0.8547, 0.7907, 0.7363 and 0.8161, respectively, and the corresponding maximized log-likelihood value (without the additive constant) is 402.61. Finally, the corresponding ML estimates of α_1,

α_2, α_3 and α_4 are obtained as 0.1491, 0.1393, 0.1625 and 0.2304, respectively. The 95% non-parametric bootstrap confidence intervals of β_1, β_2, β_3 and β_4 are then obtained as (0.8069, 0.9025), (0.7475, 0.8339), (0.6950, 0.7776) and (0.7760, 0.8562), respectively. Similarly, the 95% non-parametric bootstrap confidence intervals of α_1, α_2, α_3 and α_4 are obtained as (0.1085, 0.1897), (0.1015, 0.1771), (0.1204, 0.2046) and (0.1890, 0.2718), respectively. The ML estimate of $\mathbf{\Gamma}$ is obtained as

$$\widehat{\mathbf{\Gamma}} = \begin{bmatrix} 1.000 & 0.767 & 0.715 & 0.515 \\ 0.767 & 1.000 & 0.612 & 0.381 \\ 0.715 & 0.612 & 1.000 & 0.693 \\ 0.515 & 0.381 & 0.693 & 1.000 \end{bmatrix}. \tag{82}$$

13.7 Example 7: Multivariate Bone Mineral Data (Revisited)

Kundu et al. [120] analyzed the data set in Example 6 by using generalized multivariate BS distribution with multivariate t-kernel. The degrees of freedom ν was varied from 1 to 20 for a profile analysis with respect to ν. The ML estimates of all the unknown parameters and the corresponding maximized log-likelihood values, for different choices of ν, are obtained and are presented in Table 10. It is observed that the maximized log-likelihood values first increase and then decrease. The maximum occurs at $\nu = 9$, with the associated log-

Table 10: The maximized log-likelihood value vs. degrees of freedom $\nu = 1(1)20$ for Example 7.

ν	Maximized log-likelihood	ν	Maximized log-likelihood	ν	Maximized log-likelihood	ν	Maximized log-likelihood
1	428.315491	2	433.549164	3	436.112915	4	436.911774
5	437.847198	6	438.225861	7	439.095734	8	439.184052
9	443.246613	10	442.009432	11	441.526855	12	441.068146
13	439.946442	14	438.317993	15	437.837219	16	437.024994
17	436.234161	18	435.532867	19	434.860138	20	434.289551

likelihood value (without the additive constant) being 443.2466. It is important to mention here that the selection of the best t-kernel function through the maximized log-likelihood value is equivalent to selecting by the Akaike Information Criterion since the number of model parameters remains the same when ν varies. Furthermore, this maximized log-likelihood value of 443.246 for the multivariate t kernel with $\nu = 9$ degrees of freedom is significantly larger than the corresponding value of 402.61 for the multivariate normal kernel, which does provide a strong evidence to the fact that the multivariate t kernel provides a much better fit for these data.

Now, we provide detailed results for the case $\nu = 9$. In this case, the ML estimates of β_1, β_2, β_3 and β_4 are found to be 35.1756, 30.1062, 31.9564 and 40.1928, respectively. The corresponding 95% confidence intervals, obtained by the use of non-parametric bootstrap method, are (25.216, 45.134), (20.155, 40.057), (22.740, 41.172), and (28.822, 51.562), respectively. The ML estimates of α_1, α_2, α_3 and α_4 are 0.7746, 1.0585, 0.9457 and 0.9193, and the associated 95% non-parametric bootstrap confidence intervals are (0.6940, 0.8551), (0.9538, 1.1632), (0.8488, 1.0425), and (0.8418, 0.9968), respectively. Finally, the ML estimate of $\mathbf{\Gamma}$ is obtained as

$$\widehat{\mathbf{\Gamma}} = \begin{bmatrix} 1.000 & 0.796 & 0.696 & 0.583 \\ 0.796 & 1.000 & 0.735 & 0.813 \\ 0.696 & 0.735 & 1.000 & 0.693 \\ 0.583 & 0.813 & 0.693 & 1.000 \end{bmatrix}. \tag{83}$$

14 Concluding Remarks and Further Reading

In this paper, we have considered the two-parameter BS distribution, which was introduced almost fifty years ago. The BS model has received considerable attention since then for various reasons. The two-parameter BS distribution has a shape and a scale parameter. Due to the presence of the shape parameter, the PDF of a BS distribution can take on different shapes. It has non-monotone HF and it has a nice physical interpretation. Several generalizations of the BS distribution have been proposed in the literature and they have found numerous applications in many different fields. Recently, bivariate, multivariate and matrix-variate BS distributions have also been introduced in the literature. We have provided a detailed review of various models and methods available to date with regard to these models, and have also mentioned several open problems for future work.

Extensive work has been done on different issues with regard to BS distribution during the last 15 years. Due to limited space, we are unable to provide detailed description of all the work. Efficient *R* packages have been developed by Leiva et al. [134] and Barros et al. [37]. Interested readers are referred to the following articles for further reading. For different applications of the univariate BS, multivariate BS and related distributions, one may look at Ahmed et al. [4], Aslam et al. [6], Aslam and Kantam [7], Baklizi and El Masri [13], Balakrishnan et al. [19], Balamurali et al. [31], Castillo et al. [49], Desousa et al. [67], Garcia-Papani et al. [91], Gomes et al. [94], Ismail [104], Kotz [115], Leiva and Marchant [135], Leiva and Saulo [148], Leiva et al. [132, 143, 153, 130, 151, 138, 140, 146, 149, 137, 136, 152, 147, 133], Lio and Park [178], Lio et al. [179], Marchant et al. [182, 183, 185], Paula et al. [210], Podlaski [213], Rojas et al. [228], Saulo et al. [240, 239, 238], Upadhyay et al. [249], Vilca et al. [259], Villegas et al. [260], Wanke et al. [269], Wanke and Leiva [270], Wu and Tsai [271], Zhang et al. [278, 279], and the references cited therein. In a recent paper, Mohammadi et al. [193] have discussed the modeling of wind speed and wind power distributions by BS distribution. Also, in the reliability analysis of nano-materials, Leiva et al. [141] have recently made use of the BS lifetime distribution. Garcia-Papani et al. [89, 90] have considered a spatial modelling of the BS distribution and applied it to agricultural engineering data.

For different inference related issues, one may refer to Ahmed et al. [3], Arellano-Valle et al. [8], Athayde [9], Audrey et al. [11], Azevedo et al. [12], Balakrishnan et al. [21, 25, 24], Balakrishnan and Zhu [26, 29, 30], Barros et al. [35], Chang and Tang [53], Cordeiro et al. [60], Cysneiros et al. [61], Desmond and Yang [66], Farias and Lemonte [80], Guo et al. [97], Jeng [106], Lachos et al. [123], Lemonte [155, 156, 157, 158, 159, 160], Lemonte and Cordeiro [161], Lemonte et al. [169, 163, 170], Lemonte and Ferrari [164, 165, 166, 167], Lemonte and Patriota [172], Li et al. [174], Li and Xu [175], Lillo et al. [176], Lu and Chang [181], Meintanis [191], Moala et al. [192], Niu et al. [197], Padgett and Tomlinson [206], Pérez and Correa [211], Qu and Xie [216], Riquelme [227], Sánchez et al. [230], Santana et al. [233], Santos-Neto et al. [234, 235], Saulo et al. [237], Sha and Ng [243], Teimouri et al. [246], Tsionas [247], Upadhyay and Mukherjee [248], Vanegas and Paula [250], Vanegas et al. [251], Vilca et al. [258, 252, 253, 254, 255, 257], Wang [262], Wang and Fei [264, 265], Wang et al. [266, 268], Xiao et al. [273], Xie and Wei [274], Xu and Tang [276], Xu et al. [277], Zhu and Balakrishnan [280], and the references cited therein.

Several different models relating to univariate and multivariate BS distributions can be found in Athayde et al. [10], Balakrishnan and Saulo [23], Barros et al. [36], Bhatti [42], Cancho et al. [47], Cordeiro et al. [57], Cordeiro and Lemonte [58, 59], Desmond et al. [64], Díaz-García and Domínguez-Molina [70, 71], Díaz-García and Leiva [68], Ferreira et al. [81], Fierro et al. [82], Fonseca and Cribari-Neto [84], Genç [92], Gomes et al. [93], Guiraud et al. [96], Hashimoto [103], Jamalizadeh and Kundu [105], Khosravi et al. [112, 113], Kundu [116, 117, 118], Leiva et al. [145], Martínez-Flórez et al. [187, 188], Marchant et al. [184, 186], Olmos et al. [198], Onar and Padgett [199], Ortega et al. [200], Owen [201], Owen and Padgett [202, 203, 204], Park and Padgett [207], Patriota [209], Pescim et al. [212], Pourmousa et al. [215], Raaijmakers [217, 218], Romeiro et al. [229], Reina et al. [219], Reyes et al. [220], Sanhueza et al. [232], Vilca and Leiva [256], Volodin and Dzhungurova [261] and Ziane et al. [281]. Recently, some survival analytic methods have been developed based on BS and related models. For example, one may refer to Leão [125], Leão et al. [126, 127, 128] and Balakrishnan and Liu [22]. This is one direction in which there is a lot more research work that could be carried out.

Acknowledgements:

The authors would like to thank the associate editor and three anonymous reviewers for their constructive comments, which have helped to improve the manuscript significantly.

References

[1] Achcar, J.A. (1993), "Inferences for the Birnbaum-Saunders fatigue life model using Bayesian methods", *Computational Statistics and Data Analysis*, vol. 15, 367 - 380.

[2] Achcar, J.A., Moala, F.A. (2010), "Use of MCMC methods to obtain Bayesian inferences for the Birnbaum-Saunders distribution in the presence of censored data and covariates", *Advances and Applications in Statistics*, vol. 17, 1 - 27.

[3] Ahmed, S.E., Budsaba, K., Lisawadi, S., Volodin, A. (2008), "Parametric estimation for the Birnbaum-Saunders lifetime distribution based on a new parameterization", *Thailand Statistics*, vol. 6, 213 - 240.

[4] Ahmed, S.E., Castro-Kuriss, C., Flores, E., Leiva, V., Sanhueza, A. (2010), "A truncated version of the Birnbaum-Saunders distribution with an application in financial risk", *Pakistan Journal of Statistics*, vol. 26, 293 - 310.

[5] Andrews, D.F., Herzberg, A.M. (1985), *Data*, Springer-Verlag, New York.

[6] Aslam, M., Jun, C-H, Ahmad, M. (2011), "New acceptance sampling plans based on life tests for Birnbaum-Saunders distributions", *Journal of Statistical Computation and Simulation*, vol. 81, 461 - 470.

[7] Aslam, M., Kantam, R.R.L. (2008), "Economic reliability acceptance sampling based on truncated life tests in the Birnbaum-Saunders distribution", *Pakistan Journal of Statistics*, vol. 24, 269 - 276.

[8] Arellano-Valle, R.B., Gómez, H.W., Quintana, F.A. (2005), "Statistical inference for a general class of asymmetric distributions", *Journal of Statistical Planning and Inference*, vol. 128, 427 - 443.

[9] Athayde, E., Azevedo, A., Barros, M., Leiva, V., (2017), "Failure rate of Birnbaum-Saunders distributions: shape, change-point, estimation and robustness", *Brazilian Journal of Probability and Statistics*, (in press).

[10] Athayde, E., Azevedo, C., Leiva, V., Sanhueza, A. (2012), "About Birnbaum-Saunders distributions based on the Johnson system", *Communications in Statistics - Theory and Methods*, vol. 41, 2061 - 2079.

[11] Audrey, H.M.A.C., Francisco, C-B, Carlos, A.G.A. Jr. (2008), "On Birnbaum-Saunders inference", *Computational Statistics and Data Analysis*, vol. 52, 4939 - 4950.

[12] Azevedo, C., Leiva, V., Athayde, E., Balakrishnan, N. (2012), "Shape and change point analyses of the Birnbaum-Saunders-t hazard rate and associated estimation", *Computational Statistics and Data Analysis*, vol. 56, 3887 - 3897.

[13] Baklizi, A., El Masri, A.E. (2004), "Acceptance sampling based on truncated life tests in the Birnbaum-Saunders model", *Risk Analysis*, vol. 24, 1453 - 1457.

[14] Balakrishnan, N. (2007), "Progressive censoring methodology; an appraisal" (with discussions), *Test*, vol. 16, 211 - 296.

[15] Balakrishnan, N., Aggarwala, R. (2000), *Progressive Censoring: Theory, Methods and Applications*, Birkhäuser, Boston.

[16] Balakrishnan, N., Cohen, A.C. (1991), *Order Statistics and Inference: Estimation Methods*, Academic Press, San Diego.

[17] Balakrishnan, N., Cramer, E. (2014), *The Art of Progressive Censoring*, Birkhäuser, Boston.

[18] Balakrishnan, N., Gupta, R.C., Kundu, D., Leiva, V., Sanhueza, A. (2010), "On some mixture models based on the Birnbaum-Saunders distribution and associated inference", *Journal of Statistical Planning and Inference*, vol. 141, 2175 - 2190.

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

[20] Balakrishnan, N., Leiva, V., Sanhueza, A., Cabrera, E. (2009), "Mixture inverse Gaussian distributions and its transformations, moments and applications", *Statistics*, vol. 43, 91 - 104.

[21] Balakrishnan, N., Leiva, V., Sanhueza, A., Vilca, F. (2009), "Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm", *Statistics and Operations Research Transactions*, vol. 33, 171 - 191.

[22] Balakrishnan, N., Liu, K. (2018), "Semi-parametric likelihood inference for Birnbaum-Saunders frailty model", *REVSTAT-Statistical Journal*, vol. 16, 231 - 255.

[23] Balakrishnan, N. and Saulo, H. (2017), "Skew scale-mixture Birnbaum-Saunders distributions and associated inference", Under review.

[24] Balakrishnan, N., Saulo, H., Bourguignon, M., Zhu, X. (2017), "On moment-type estimators for a class of log-symmetric distributions", *Computational Statistics*, vol. 32, 1339 - 1355.

[25] Balakrishnan, N., Saulo, H., Leão, J. (2017), "On a new class of skewed Birnbaum-Saunders models", *Journal of Statistical Theory and Practice*, vol. 11, 573 - 593.

[26] Balakrishnan, N., Zhu, X. (2013), "A simple and efficient method of estimation of the parameters of a bivariate Birnbaum-Saunders distribution based on type-II censored samples", In: *Multivariate Statistics*, pp. 34 - 47, World Scientific Publication, Hackensack, NJ.

[27] Balakrishnan, N., Zhu, X. (2014a), "On the existence and uniqueness of the maximum likelihood estimates of parameters of Birnbaum-Saunders distribution based on Type-I, Type-II and hybrid censored samples", *Statistics*, vol. 48, 1013 - 1032.

[28] Balakrishnan, N., Zhu, X. (2014b), "An improved method of estimation for the parameters of the Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 84, 2285 - 2294.

[29] Balakrishnan, N., Zhu, X. (2015a), "Inference for the bivariate Birnbaum-Saunders lifetime regression model and associated inference", *Metrika*, vol. 78, 853 - 872.

[30] Balakrishnan, N., Zhu, X. (2015b), "Inference for the Birnbaum-Saunders lifetime regression model with applications", *Communications in Statistics - Simulation and Computation*, vol. 44, 2073 - 2100.

[31] Balamurali, S., Jeyadurga, P., Usha, M. (2018), "Optimal design of repetitive group sampling plans for Weibull and gamma distributions with applications and comparison to the Birnbaum-Saunders distribution", *Journal of Applied Statistics*, (to appear).

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

[33] Barndorff-Neilsen, O.E. (1991), "Modified signed log likelihood ratio", *Biometrika*, vol. 78, 557 - 563.

[34] Barreto, L.S., Cysneiros, A.H.M.A., Cribari-Neto, F. (2013), "Improved Birnbaum-Saunders inference under type-II censoring", *Computational Statistics and Data Analysis*, vol. 57, 68 - 81.

[35] Barros, M., Leiva, V., Ospina, R., Tsuyuguchi, A. (2014), "Goodness-of-fit tests for the Birnbaum-Saunders distribution with censored reliability data", *IEEE Transactions on Reliability*, vol. 63, 543 - 554.

[36] Barros, M., Paula, G.A., Leiva, V. (2008), "A new class of survival regression models with heavy-tailed errors: robustness and diagnostics", *Lifetime Data Analysis*, vol. 14, 316 - 332.

[37] Barros, M., Paula, G.A., Leiva, V. (2009), "An *R* implementation for generalized Birnbaum-Saunders distributions", *Computational Statistics and Data Analysis*, vol. 53, 1511 - 1528.

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

[39] Berger, J.O. and Bernardo, J.M. (1992), "On the development of reference priors (with discussions)", In: *Bayesian Statistics*, vol. 4, pp. 35 - 60, Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Clarendon Press, Oxford, England.

[40] Bernardo, J.M. (1979), “Reference posterior distributions for Bayesian inference (with discussions)”, *Journal of the Royal Statistical Society, Ser. B*, vol. 41, 113 - 147.

[41] Bhattacharyya, G. K., Fries, A. (1982), “Fatigue failure models - Birnbaum-Saunders versus inverse Gaussian”, *IEEE Transactions on Reliability*, vol. 31, 439 - 440.

[42] Bhatti, C.R. (2010), “The Birnbaum-Saunders auto-regression conditional duration model”, *Mathematics and Computers in Simulation*, vol. 80, 2062 - 2078.

[43] Birnbaum, Z.W., Saunders, S.C. (1969a), “A new family of life distribution”, *Journal of Applied Probability*, vol. 6, 319 - 327.

[44] Birnbaum, Z.W., Saunders, S.C. (1969b), “Estimation for a family of life distributions with applications to fatigue”, *Journal of Applied Probability*, vol. 6, 328 - 347.

[45] Bourguingon, M., Leão, J., Leiva, V., Santos-Neto, M. (2017), “The transmuted Birnbaum-Saunders distribution”, *REVSTAT-Statistical Journal*, vol. 15, 601 - 628.

[46] Cancho, V.G., Louzada, F., Barriga, G.D.C. (2012), “The geometric Birnbaum-Saunders regression model with cure rate” *Journal of Statistical Planning and Inference*, vol. 142, 993 - 1000.

[47] Cancho, V.G., Ortega, E.M.M., Paula, G.A. (2010), “On estimation and influence diagnostics for log-Birnbaum-Saunders Student-t regression models- full Bayesian analysis”, *Journal of Statistical Planning and Inference*, vol. 140, 2486 - 2496.

[48] Caro-Lopera, F.J., Leiva, V., Balakrishnan, N. (2012), "Connection between the Hadamard and matrix products with an application to matrix-variate Birnbaum-Saunders distributions", *Journal of Multivariate Analysis*, vol. 104, 126 - 139.

[49] Castillo, N.O., Gomez, H.W., Bolfarine, H. (2011), "Epsilon Birnbaum-Saunders distribution family: properties and inference", *Statistical Papers*, vol. 52, 871 - 883.

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

[51] Chang, D.S., Tang, L.C. (1994a), "Percentile bounds and tolerance limits for the Birnbaum-Saunders distribution", *Communications in Statistics - Theory and Methods*, vol. 23, 2853 - 2863.

[52] Chang, D.S., Tang, L.C. (1994b), "Random number generator for the Birnbaum-Saunders distribution", *Computers and Industrial Engineering*, vol. 27, 345 - 348.

[53] Chang, D.S., Tang, L.C. (1994c), "Graphical analysis for Birnbaum-Saunders distribution", *Microelectronics and Reliability*, vol. 34, 17 - 22.

[54] Chhikara, R.S., Folks, J.L. (1989), *The Inverse Gaussian Distribution*, Marcel Dekker, New York.

[55] Cohen, A.C. (1963), "Progressively censored samples in life testing", *Technometrics*, vol. 5, 327 - 329.

[56] Cohen, A.C., Whitten, B.J., Ding, Y. (1984), "Modified moment estimation for three-parameter Weibull distribution", *Journal of Quality Technology*, vol. 16, 159 - 167.

[57] Cordeiro, G. M., Cancho, V. G., Ortega, E. M. M., Barriga, G. D. C. (2016), "A model with long-term survivors: negative binomial Birnbaum-Saunders", *Communications in Statistics - Theory and Methods*, vol. 45, 1370 - 1387.

[58] Cordeiro, G.M., Lemonte, A.J. (2011), "The β Birnbaum-Saunders distribution: An improved distribution for fatigue life model", *Computational Statistics and Data Analysis*, vol. 55, 1445 - 1461.

[59] Cordeiro, G.M., Lemonte, A.J. (2014), "The exponentiated generalized Birnbaum-Saunders distribution", *Applied Mathematics and Computation*, vol. 247, 762 - 779.

[60] Cordeiro, G. M., Lima, M. C., Cysneiros, A., Pascoa, M.A.R., Pescim, R.R., Ortega, E.M.M. (2016), "An extended Birnbaum-Saunders distribution: Theory, estimation and applications", *Communications in Statistics - Theory and Methods*, vol. 45, 2268 - 2297.

[61] Cysneiros, A., Cribari-Neto, F., Araújo Jr., C. (2008), "On Birnbaum-Saunders inference", *Computational Statistics and Data Analysis*, vol. 52, 4939 - 4950.

[62] Desmond, A.F. (1985), "Stochastic models of failure in random environments", *Canadian Journal of Statistics*, vol. 13, 171 - 183.

[63] Desmond, A.F. (1986), "On the relationship between two fatigue-life models", *IEEE Transactions on Reliability*, vol. 35, 167 - 169.

[64] Desmond, A.F., Cintora Gonzalez, C.L., Singh, R.S., Lu, X.W. (2012), "A mixed effect log-linear model based on the Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 56, 399 - 407.

[65] Desmond, A.F., Rodríguea-Yam, G.A., Lu, X. (2008), "Estimation of parameters for a Birnbaum-Saunders regression model with censored data", *Journal of Statistical Computation and Simulation*, vol. 78, 983 - 997.

[66] Desmond, A.F., Yang, Z. (1996), "Shortest prediction intervals for the Birnbaum-Saunders distribution", *Communications in Statistics - Theory and Methods*, vol. 24, 1383 - 1401.

[67] Desousa, M., Saulo, H., Leiva, V., Scalco, P. (2017), "On a tobit-Birnbaum-Saunders model with an application to medical data", *Journal of Applied Statistics*, vol. 45, 932 - 955.

[68] Díaz-García, J. A., Leiva, V. (2002), "A new family of life distributions based on the Birnbaum-Saunders distribution", *Technical Report I-02-17*, CIMAT, Mexico. http://www.cimat.mx/biblioteca/RepTec/.

[69] Díaz-García, J. A., Leiva, V. (2005), "A new family of life distributions based on elliptically contoured distributions", . *Journal of Statistical Planning and Inference*, vol. 128, 445 - 457; Erratum, *Journal of Statistical Planning and Inference*, vol. 137, 1512 - 1513.

[70] Díaz-García, J. A., Domínguez-Molina, J. R. (2006a), "Some generalizations of Birnbaum-Saunders and sinh-normal distributions", *International Mathematical Forum*, vol. 1, 1709 - 1727.

[71] Díaz-García, J. A., Domínguez-Molina, J. R. (2006b), "A new family of life distributions for dependent data: Estimation", *Computational Statistics and Data Analysis*, vol. 51, 5927 - 5939.

[72] Draper, N.R., Guttman, I. (1979), "Bayesian inference on parameters associated with families closed under reciprocation", *Sankhya*, vol. 41, B, 77 - 90.

[73] Dupuis, D.J., Mills, J.E. (1998), "Robust estimation of the Birnbaum-Saunders distribution", *IEEE Transactions on Reliability*, vol. 47, 88 - 95.

[74] Efron, B. (1982), *The Jacknife, the Bootstrap and Other Resampling Plans*, Society of Industrial and Applied Mathematics, Philadelphia, PA.

[75] Engelhardt, M., Bain, L.J., Wright, F.T. (1981), "Inferences on the parameters of the Birnbaum-Saunders fatigue life distribution based on maximum likelihood estimation", *Technometrics*, vol. 23, 251 - 255.

[76] Fang, L., Balakrishnan, N. (2018), "Ordering properties of the smallest order statistics from generalized Birnbaum-Saunders models with associated random shocks", *Metrika*, vol. 81, 19 - 35.

[77] Fang, L., Zhu, X., Balakrishnan, N. (2016), "Stochastic comparisons of parallel and series systems with heterogeneous Birnbaum-Saunders components", *Statistics and Probability Letters*, vol. 112, 131 - 136.

[78] Fang, L., Zhu, X., Balakrishnan, N. (2018), "Stochastic ordering of minima and maxima from heterogeneous bivariate Birnbaum-Saunders random vectors", *Statistics*, vol. 52, 147 - 155.

[79] Fang, K. T., Zhang, Y. T. (1990), *Generalized Multivariate Analysis*, Sciences Press, Beijing; Springer-Verlag, Berlin.

[80] Farias, R. B. A., Lemonte, A. J. (2011), "Bayesian inference for the Birnbaum-Saunders nonlinear regression model", *Statistical Methods and Applications*, vol. 20, 423 - 438.

[81] Ferreira, M., Gomes, M. I., Leiva, V. (2012), "On an extreme value version of the Birnbaum-Saunders distribution", *REVSTAT-Statistical Journal*, vol. 10, 181 - 210.

[82] Fierro, R., Leiva, V., Ruggeri, F., Sanhueza, A. (2013), "On a Birnbaum-Saunders distribution arising from a non-homogeneous Poisson process", *Statistics and Probability Letters*, vol. 83, 1233 - 1239.

[83] Fletcher, H. (1911), "A verification of the theory of Brownian movements and a direct determination of the value of NE for gaseous ionization", *The Physical Review*, vol. 33, 81 - 110.

[84] Fonseca, R.V., Cribari-Neto, F. (2018), "Bimodal Birnbaum-Saunders generalized autoregressive score model", *Journal of Applied Statistics*, (to appear).

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

[86] Freudenthal, A.M., Shinozuka, M. (1961), "Structural safety under conditions of ultimate load failure and fatigue", *Technical Report WADD-TR-61-77*, Wright Air Development Division.

[87] From, S.G., Li, L. (2006), "Estimation of the parameters of the Birnbaum-Saunders distribution", *Communications in Statistics - Theory and Methods*, vol. 35, 2157 - 2169.

[88] Galea, M., Leiva-Sánchez, V., Paula, G. (2004), "Influence diagnostics in log-Birnbaum-Saunders regression models", *Journal of Applied Statistics*, vol. 31, 1049 - 1064.

[89] Garcia-Papani, F., Uribe-Opazo, M, Leiva, V., Aykroyd, R.G. (2017), "Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data", *Stochastic Environmental Research and Risk Assessment*, vol. 31, 105 - 124.

[90] Garcia-Papani F., Uribe-Opazo M., Leiva V., Aykroyd R. G. (2018), "Birnbaum-Saunders spatial regression models: Diagnostics and application to georeferenced data", *Chemometrics and Intelligent Laboratory Systems* (in press).

[91] Garcia-Papani F., Uribe-Opazo M., Leiva V., Ruggeri, F. (2018), "Kriging with external drift in a Birnbaum-Saunders geostatistical model", *Stochastic Environmental Research and Risk Assessment* (in press).

[92] Genç, A. (2013), "The generalized T Birnbaum-Saunders family", *Statistics*, vol. 47, 613 - 625.

[93] Gomes, H.W., Olivares-Pacheco, J.F., Bolfarine, H. (2009), "An extension of the generalized Birnbaum-Saunders distribution", *Statistics and Probability Letters*, vol. 79, 331 - 338.

[94] Gomes, M.I., Ferreira, M., Leiva, V. (2012), "The extreme value Birnbaum-Saunders model and its moments and applications in biometry", *Biometrical Letters*, vol. 49, 81 - 94.

[95] Gradshteyn, I.S., Randzhik, I.M. (2000), *Tables of Integrals, Series and Products*, Academic Press, New York.

[96] Guiraud, P., Leiva, V., Fierro, R. (2009), "A non-central version of the Birnbaum-Saunders distribution for reliability analysis", *IEEE Transactions on Reliability*, vol. 52, 152 - 160.

[97] Guo, X., Wu, H., Li, G., Li, Q. (2017), "Inference for the common mean of several Birnbaum-Saunders populations", *Journal of Applied Statistics*, vol. 44, 941 - 954.

[98] Gupta, R. C. (2015), "Reliability studies of bivariate Birnbaum-Saunders distribution", *Probability in the Engineering and Information Science*, vol. 29, 265 - 276.

[99] Gupta, R.C., Akman, H.O. (1995), "On the reliability study of a weighted inverse Gaussian model", *Journal of Statistical Planning and Inference*, vol. 48, 69 - 83.

[100] Gupta, R.C., Akman, H.O. (1997), "Estimation of critical points in the mixture of inverse Gaussian distribution", *Statistical Papers*, vol. 38, 445 - 452.

[101] Gupta, R.C., Kirmani, S.N.U.A. (1990), "The role of weighted distributions in stochastic modelling", *Communications in Statistics - Theory and Methods*, vol. 19, 3147 - 3162.

[102] Hampel, F.R. (1974), "The influence curve and its role in robust estimation", *Journal of the American Statistical Association*, vol. 69, 383 - 393.

[103] Hashimoto, E.M., Ortega, E.M.M., Cordeiro, G.M., Cancho, V.G. (2014), "The Poisson Birnbaum-Saunders model with long-term survivors", *Statistics*, vol. 48, 1394 - 1413.

[104] Ismail, S. (2005), "Reliability estimation for accelerated Birnbaum-Saunders model with censoring", *Far East Journal of Theoretical Statistics*, vol. 17, 43 - 68.

[105] Jamalizadeh, A., Kundu, D. (2015), "A multivariate Birnbaum-Saunders distribution based on the multivariate skew normal distribution", *Journal of the Japan Statistical Society*, vol. 45, 1 - 20.

[106] Jeng, S-L. (2003), "Inferences for the fatigue life model based on the Birnbaum-Saunders distribution result", *Communications in Statistics - Simulation and Computation*, vol. 32, 43 - 60.

[107] Joe, H. (2005), "Asymptotic efficiency of the two-stage estimation method for copula-based models", *Journal of Multivariate Analysis*, vol. 94, 401 - 419.

[108] Johnson, N.L., Kotz, S. (1975), "A vector multivariate hazard rate", *Journal of Multivariate Analysis*, vol. 5, 53 - 66.

[109] Johnson, N.L., Kotz, S., Balakrishnan, N. (1995), *Continuous Univariate Distributions*, vol. 2, Second Edition, John Wiley & Sons, New York.

[110] Johnson, R.A., Wichern, D.W. (1999), *Applied Multivariate Analysis*, Fourth edition, Prentice-Hall, New Jersey.

[111] Jorgensen, B., Seshadri, V., Whitmore, G.A. (1991), "On the mixture of the inverse Gaussian distribution with its complementary reciprocals", *Scandinavian Journal Statistics*, vol. 18, 77 - 89.

[112] Khosravi, M., Kundu, D., Jamalizadeh, A. (2015), "On bivariate and a mixture of bivariate Birnbaum-Saunders distributions", *Statistical Methodology*, vol. 23, 1 - 17.

[113] Khosravi, M., Leiva, V., Jamalizadeh, A., Porcu, E. (2016), "On a nonlinear Birnbaum-Saunders model based on a bivariate construction and its characteristics", *Communications in Statistics - Theory and Methods*, vol. 45, 772 - 793.

[114] Konstantinowsky, D. (1914), "Elektrische Ladungen und Brownsche Bewegung schrkleiner Metallteilchen Gasen", *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften*, vol. 123, 1697 - 1752.

[115] Kotz, S., Leiva, V., Sanhueza, A. (2010), "Two new mixture models related to the inverse Gaussian distribution", *Methodology and Computing in Applied Probability* vol. 12, 199 - 212.

[116] Kundu, D. (2015a), "Bivariate log Birnbaum-Saunders distribution", *Statistics*, vol. 49, 900 - 917.

[117] Kundu, D. (2015b), "Bivariate sinh-normal distribution and a related model", *Brazilian Journal of Probability and Statistics*, vol. 29, 590 - 607.

[118] Kundu, D. (2017), "On multivariate log Birnbaum-Saunders distribution", *Sankhya*, Ser. B, vol. 79, 292 - 315.

[119] Kundu, D., Balakrishnan, N., Jamalizadeh, A. (2010), "Bivariate Birnbaum-Saunders distribution and associated inference", *Journal of Multivariate Analysis*, vol. 101, 113 - 125.

[120] Kundu, D., Balakrishnan, N., Jamalizadeh, A. (2013), "Generalized multivariate Birnbaum-Saunders distributions and related inferential issues", *Journal of Multivariate Analysis* , vol. 116, 230 - 244.

[121] Kundu, D., Gupta, R.C. (2017), "On bivariate Birnbaum-Saunders distribution", *American Journal of Mathematical and Management Sciences*, vol. 36, 21 - 33.

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

[123] Lachos, V.H., Dey, D.K., Cancho, V.G., Louzada, F. (2017), "Scale mixtures log-Birnbaum-Saunders regression models with censored data: a Bayesian approach", *Journal of Statistical Computation and Simulation*, vol. 87, 2002 - 2022.

[124] Langlands, A. O., Pocock, S. J., Kerr, G. R., Gore, S. M. (1979), "Long term survival of patients with breast cancer: A study of the curability of the disease", *British Medical Journal*, vol. 17, 1247 - 1251.

[125] Leão, J.S. (2017), "Modeling based on a reparametrized Birnbaum-Saunders distribution for survival data", Ph.D. thesis, Federal University of São Carlos, São Carlos, Brazil.

[126] Leão, J.S., Leiva, V., Saulo, H., Tomazella, V. (2017a), "Birnbaum-Saunders frailty regression models: Diagnostics and application to medical data", *Biometrical Journal*, vol. 59, 291 - 314.

[127] Leão, J.S., Leiva, V., Saulo, H., Tomazella, V. (2017a), "A survival model with Birnbaum-Saunders frailty for uncensored and censored data", *Brazilian Journal of Probability and Statistics* (to appear).

[128] Leão, J.S., Leiva, V., Saulo, H., Tomazella, V. (2017c), "Incorporation of frailties into a cure rate regression model and its diagnostics and applications to melanoma data", *Statistics and Medicine*, (in press).

[129] Leiva, V. (2016), *The Birnbaum-Saunders distribution*, Elsevier/Academic Press, Amsterdam.

[130] Leiva, V., Athayde, E., Azevedo, C., Marchant, C. (2011), "Modeling wind energy flux by a Birnbaum-Saunders distribution with an unknown shift parameter", *Journal of Applied Statistics*, vol. 38, 2819 - 2838.

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

[132] Leiva, V., Barros, M., Paula, G. A., Sanhueza, A. (2008b), "Generalized Birnbaum-Saunders distribution applied to air pollutant concentration", *Environmetrics*, vol. 19, 235 - 249.

[133] Leiva, V., Ferreira, M., Gomes, M.I., Lillo, C. (2016), "Extreme value Birnbaum-Saunders regression models applied to environmental data", *Stochastic Environmental Research and Risk Assessment*, vol. 30, 1045 - 1058.

[134] Leiva, V., Hernández, H., Riquelme, M. (2006), "A new package for the Birnbaum-Saunders distribution", *R Journal*, vol. 6, 35 - 40. http://www.r-project.org/doc/Rnews/Rnews$_{2006}$-4.pdf.

[135] Leiva, V., Marchant, C. (2017), "Multivariate generalized Birnbaum-Saunders models applied to case studies in bio-engineering and industry", In: *Recent Studies on Risk Analysis and Statistical Modeling*, pp. 283 - 302, Oliveira, T., Kitsos, C., Oliveira, A., Grilo, M. (Eds.), Springer, Switzerland.

[136] Leiva, V., Marchant, C., Ruggeri, F., Saulo, H. (2015), "A criterion for environmental assessment using Birnbaum-Saunders attribute control charts", *Environmetrics*, vol. 26, 463 - 476.

[137] Leiva, V., Marchant, C., Saulo, H., Aslam, M., Rojas, F. (2014), "Capability indices for Birnbaum-Saunders processes applied to electronic and food industries", *Journal of Applied Statistics*, vol. 41, 1881 - 1902.

[138] Leiva, V., Ponce, M.G., Marchant, C., Bustos, O. (2012), "Fatigue statistical distributions useful for modeling diameter and mortality of trees", *Colombian Journal of Statistics*, vol. 35, 349 - 367.

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

[140] Leiva, V., Rojas, E., Galea, M., Sanhueza, A. (2014), "Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data", *Applied Stochastic Models in Business and Industry*, vol. 30, 115 - 131.

[141] Leiva, V., Ruggeri, F., Saulo, H., Vivanco, J.F. (2017), "A methodology based on the Birnbaum-Saunders distribution for reliability analysis applied to nano-materials", *Reliability Engineering and System Safety*, vol. 157, 192 - 201.

[142] Leiva, V., Sanhueza, A., Angulo, J. M. (2009), "A length-biased version of the Birnbaum-Saunders distribution with application in water quality", *Stochastic Environmental Research and Risk Assessment*, vol. 23, 299 - 307.

[143] Leiva, V., Sanhueza, A. Kotz, S., Araneda, N. (2010), "A unified mixture model based on the inverse Gaussian distribution", *Pakistan Journal of Statistics*, vol. 26, 445 - 460.

[144] Leiva, V., Sanhueza, A., Sen, P. K., Paula, G.A. (2008), "Random number generators for the generalized Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 78. 1105 - 1118.

[145] Leiva, V., Sanhueza, A., Silva, A., Galea, M. (2008), "A new three-parameter extension of the inverse Gaussian distribution", *Statistics and Probability Letters*, vol. 78, 1266 - 1273.

[146] Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., Barros, M. (2014b), "Birnbaum-Saunders statistical modelling: a new approach", *Statistical Modelling*, vol. 14, 21 - 48.

[147] Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., Barros, M. (2016), "A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution", *Applied Stochastic Models in Business and Industry*, vol. 32, 74 - 89.

[148] Leiva, V., Saulo, H. (2017), "Environmental applications based on Birnbaum-Saunders models", In: *Mathematical and Statistical Applications in Life Sciences and Engineering*, pp 283 - 304, Adhikari, M.R., Chaubey, Y.P., Adhikari, A. (Eds.), Springer, New York.

[149] Leiva, V., Saulo, H., Leao, J., Marchant, C., (2014), "A family of autoregressive conditional duration models applied to financial data", *Computational Statistics and Data Analysis*, vol. 9, 175 - 191.

[150] Leiva, V, Saunders, S.C. (2015), "Cumulative damage models", *Wiley StatsRef:*, Statistics Reference Online 1-10.

[151] Leiva, V., Soto, G., Cabrera, E., Cabrera, G. (2011), "New control charts based on the Birnbaum-Saunders distribution and their implementation", *Colombian Journal of Statistics*, vol. 34, 147 - 176.

[152] Leiva, V., Tejo, M., Guiraud, P., Schmachtenberg, O., Orio, P., Marmolejo, F. (2015), "Modeling neural activity with cumulative damage distributions", *Biological Cybernetics*, vol. 109, 421 - 433.

[153] Leiva, V., Vilca, F., Balakrishnan, N., Sanhueza, A. (2010), "A skewed Sinh-Normal distribution and its properties and application to air pollution", *Communications in Statistics - Theory and Methods*, vol. 39, 426 - 443.

[154] Leiva V., Vivanco J.F. (2016), "Fatigue models", *Wiley StatsRef:* Statistics Reference Online 1-10.

[155] Lemonte, A. J. (2011), "Covariance matrix formula for Birnbaum-Saunders regression models", *Journal of Statistical Computation and Simulation*, vol. 81, 899 - 908.

[156] Lemonte, A. J. (2012), "A log-Birnbaum-Saunders regression model with asymmetric errors", *Journal of Statistical Computation and Simulation*, vol. 82, 1775 - 1787.

[157] Lemonte, A. J. (2013a), "A new extended Birnbaum-Saunders regression model for lifetime modeling", *Computational Statistics and Data Analysis*, vol. 64, 34 - 50.

[158] Lemonte, A. J. (2013b), "Multivariate Birnbaum-Saunders regression model", *Journal of Statistical Computation and Simulation*, vol. 83, 2244 - 2257.

[159] Lemonte, A. J. (2013c), "A new extension of the Birnbaum-Saunders distribution", *Brazilian Journal of Probability and Statistics*, vol. 27, 133 - 149.

[160] Lemonte, A. J. (2016), "Robust inference for Birnbaum-Saunders regressions", *Journal of Statistical Computation and Simulation*, vol. 86, 611 - 622.

[161] Lemonte, A.J., Cordeiro, G.M. (2009), "Birnbaum-Saunders nonlinear regression models", *Computational Statistics and Data Analysis*, vol. 53, 4441 - 4452.

[162] Lemonte, A. J., Cordeiro, G. M. (2010), "Asymptotic skewness in Birnbaum-Saunders nonlinear regression models", *Statistics and Probability Letters*, vol. 80, 892 - 898.

[163] Lemonte, A. J., Cordeiro, G. M., Moreno, G. (2012), "Bartlett corrections in Birnbaum-Saunders nonlinear regression models", *Journal of Statistical Computation and Simulation*, vol. 82, 927 - 935.

[164] Lemonte, A.J., Ferrari, S.L.P. (2011a), "Size and power properties of some tests in the Birnbaum-Saunders regression model", *Computational Statistics and Data Analysis*, vol. 55, 1109 - 1117.

[165] Lemonte, A.J., Ferrari, S.L.P. (2011b), "Small-sample corrections for score tests in Birnbaum-Saunders regression model", *Journal of Statistical Planning and Inference*, vol. 141, 1031 - 1040.

[166] Lemonte, A.J., Ferrari, S.L.P. (2011c), "Improved likelihood inference in Birnbaum-Saunders regression", *Communications in Statistics - Theory and Methods*, vol. 40, 232 - 243.

[167] Lemonte, A.J., Ferrari, S.L.P. (2011d), "Testing hypotheses in the Birnbaum-Saunders distribution under type-II censored samples", *Computational Statistics and Data Analysis*, vol. 55, 2388 - 2399.

[168] Lemonte, A.J., Ferrari, S.L.P. (2011e), "Signed likelihood ratio tests in the Birnbaum-Saunders regression model", *Journal of Statistical Planning and Inference*, vol. 141, 1031 - 1040.

[169] Lemonte, A.J., Ferrari, S.L.P., Silvia L. P., Cribari-Neto, F. (2010), "Improved likelihood inference in Birnbaum-Saunders regressions", *Computational Statistics and Data Analysis*, vol. 54, 1307 - 1316.

[170] Lemonte, A. J., Martínez-Flórez, G., Moreno-Arenas, G. (2015), "Multivariate Birnbaum-Saunders distribution: properties and associated inference", *Journal of Statistical Computation and Simulation*, vol. 85, 374 - 392.

[171] Lemonte, A.J., Cribari-Neto, F., Vasconcellos, K.L.P. (2007), "Improved statistical inference for the two-parameter Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 51, 4656 - 4681.

[172] Lemonte, A. J., Patriota, A. G. (2011), "Influence diagnostics in Birnbaum-Saunders nonlinear regression models", *Journal of Applied Statistics*, vol. 38, 871 - 884.

[173] Lemonte, A.J., Simas, A.B., Cribari-Neto, F. (2007), "Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 78, 37 - 49.

[174] Li, A.P., Chen, Z.X., Xie, F.C. (2012), "Diagnostic analysis for heterogeneous log-Birnbaum-Saunders regression model", *Statistics and Probability Letters*, vol. 82, 1690 - 1698.

[175] Li, Y., Xu, A. (2014), "Fiducial inference for Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 86, 1673 - 1685.

[176] Lillo, C., Leiva, V., Nicolis, O., Aykroyd, R.G. (2017), "L-moments of the Birnbaum-Saunders distribution and its extreme value version: Estimation, goodness of fit and application to earthquake data", *Journal of Applied Statistics*, vol. 45, 187 - 209.

[177] Lindley, D.V. (1980), "Approximate Bayesian method", *Trabajos de Estadistica*, vol. 31, 223 - 237.

[178] Lio, Y. L., Park, C. (2008), "A bootstrap control chart for Birnbaum-Saunders percentiles", *Quality and Reliability Engineering International*, vol. 24, 585 - 600.

[179] Lio, Y.L., Tsai, T-R., Wu, S-J. (2010), "Acceptance sampling plans from truncated life tests based on the Birnbaum-Saunders distribution for percentiles", *Communications in Statistics - Simulation and Computation*, vol. 39, 119 - 136.

[180] Louis, T.A. (1982), “Finding the observed information matrix using the EM algorithm”, *Journal of the Royal Statistical Society*, Ser. B, vol. 44, 226 - 244.

[181] Lu, M. C., Chang, D. S. (1997), “Bootstrap prediction intervals for the Birnbaum-Saunders distribution”, *Microelectronics Reliability*, vol. 37, 1213 - 1216.

[182] Marchant, C., Bertin, K., Leiva, V., Saulo, H. (2013), “Generalized Birnbaum-Saunders kernel density estimators and an analysis of financial data”, *Computational Statistics and Data Analysis*, vol. 63, 1 - 15.

[183] Marchant, C., Leiva, V., Cavieres, M.F., Sanhueza, A. (2013), “Air contaminant statistical distributions with application to PM10 in Santiago, Chile”, *Reviews of Environmental Contamination and Toxicology*, vol. 223, 1 - 31.

[184] Marchant C., Leiva V., Cysneiros, F.J.A. (2016) “A multivariate log-linear model for Birnbaum-Saunders distributions”, *IEEE Transactions on Reliability*, vol. 65, 816 - 827.

[185] Marchant, C., Leiva, V., Cysneiros, F.J.A., Liu, S. (2017), “Robust multivariate control charts based on Birnbaum-Saunders distributions”, *Journal of Statistical Computation and Simulation*, vol. 88, 182 - 202.

[186] Marchant, C., Leiva, V., Cysneiros, F.J.A., Vivanco, J.F. (2016), “Diagnostics in multivariate Birnbaum-Saunders regression models”, *Journal of Applied Statistics*, vol. 43, 2829 - 2849.

[187] Martínez-Flórez, G., Bolfarine, H., Gómez, H. W. (2014) “An alpha-power extension for the Birnbaum-Saunders distribution”, *Statistics*, vol. 48, 896 - 912.

[188] Martínez-Flórez, G., Bolfarine, H., Gómez, H. W. (2017) "The log-linear Birnbaum-Saunders power model", *Methodology and Computing in Applied Probability*, vol. 19, 913 - 933.

[189] Mann, N.R., Schafer, R.E., Singpurwalla, N.D. (1974), *Methods for Statistical Analysis of Reliability and Life Data*, John Wiley and Sons, New York.

[190] McCool, J.L. (1974), "Inferential techniques for Weibull populations", Aerospace Research Laboratories Report ARL TR74-0180, Wright-Patterson Air Force Base, Dayton, Ohio.

[191] Meintanis, S.G. (2010), "Inference procedures for the Birnbaum-Saunders distribution and its generalizations", *Computational Statistics and Data Analysis*, vol. 54, 367 - 373.

[192] Moala, F.A., Achcar, J.A., Gimenez, R.P. (2015), "Bayes estimation for the Birnbaum-Saunders distribution in the presence of censored data", *IEEE Latin America Transactions*, vol. 13, 3187 - 3192.

[193] Mohammadi, K., Olavi, O., McGowan, J.G. (2017), "Use of Birnbaum-Saunders distribution for wind speed and wind power probability distributions: A review", *Energy Conversion and Management*, vol. 143, 109 - 122.

[194] Mudholkar, G.S., Hutson, A.D. (2000), "The epsilon skew normal distribution for analyzing near normal data", *Journal of Statistical Planning and Inference*, vol. 83, 291 - 309.

[195] Ng, H.K.T., Kundu, D., Balakrishnan, N. (2003), "Modified moment estimation for the two-parameter Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 43, 283 - 298.

[196] Ng, H.K.T., Kundu, D., Balakrishnan, N. (2006), "Point and interval estimations for the two-parameter Birnbaum-Saunders distribution based on type-II censored samples", *Computational Statistics and Data Analysis*, vol. 50, 3222 - 3242.

[197] Niu, C., Guo, X., Xu, W., Zhu, L. (2013), "Comparison of several BirnbaumSaunders distributions", *Journal of Statistical Computation and Simulation*, vol. 84, 2721 - 2733.

[198] Olmos, N.M., Martínez-Flórez, G., Bolfarine, H. (2017), "Bimodal Birnbaum-Saunders distribution with applications to non negative measurements", *Communications in Statistics - Theory and Methods*, vol. 46, 6240 - 6257.

[199] Onar, A., Padgett, W.J. (2000), "Accelerated test models with the inverse Gaussian distribution", *Journal of Statistical Planning and Inference*, vol. 89, 119 - 133.

[200] Ortega, E. M. M., Cordeiro, G. M., Lemonte, A. J. (2012), "A log-linear regression model for the β-Birnbaum-Saunders distribution with censored data", *Computational Statistics and Data Analysis*, vol. 56, 698 - 718.

[201] Owen, W.J. (2006), "A new three-parameter extension to the Birnbaum-Saunders distribution", *IEEE Transactions on Reliability*, vol. 55, 475 - 479.

[202] Owen, W. J., Padgett, W. J. (1998), "Birnbaum-Saunders-type models for system strength assuming multiplicative damage", In: *Frontiers in Reliability Analysis*, Ed. Basu, A.P., Indian Association for Productivity, Quality and Reliability, Singapore, World Scientific.

[203] Owen, W. J., Padgett, W. J. (1999), "Acceleration models for system strength based on Birnbaum-Saunders distributions", *Lifetime Data Analysis*, vol. 5, 133 - 147.

[204] Owen, W. J., Padgett, W. J. (2000), "A Birnbaum-Saunders accelerated life model", *IEEE Transactions on Reliability*, vol. 49, 224 - 229.

[205] Padgett, W. J. (1982), "On Bayes estimation of reliability for the Birnbaum-Saunders fatigue model using Bayesian methods", *IEEE Transactions on Reliability*, vol. 31, 436 - 438.

[206] Padgett, W. J., Tomlinson, M. A. (2003), "Lower confidence bounds for percentiles of Weibull and Birnbaum-Saunders distributions", *Journal of Statistical Computation and Simulation*, vol 73, 429 - 443.

[207] Park, C., Padgett, W.J. (2006), "A general class of cumulative damage models for materials failure", *Journal of Statistical Planning and Inference*, vol. 136, 3783 - 3801.

[208] Patil, G.P. (2002), "Weighted distributions", In: El-Shaarawi A.H. and Piegorsh, W. W. (eds.) *Encyclopedia of Environmetrics*, vol. 4, John Wiley and Sons, Chichester, pp. 2369 - 2377.

[209] Patriota, A. G. (2012), "On scale-mixture Birnbaum-Saunders distributions", *Journal of Statistical Planning and Inference*, vol. 142, 2221 - 2226.

[210] Paula, G. A., Leiva, V., Barros, M., Liu, S. (2012), "Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance", *Applied Stochastic Models in Business and Industry*, vol. 28, 16 - 34.

[211] Pérez, R. A., Correa, J. C. (2008), "Confidence intervals for the parameters of the Birnbaum-Saunders distribution via relative likelihood", *Revista Colombiana de Estad´stica*, vol. 31, 85 - 94 (in Spanish).

[212] Pescim, R. R., Cordeiro, G. M., Nadarajah, S., Demétrio, C. G. B., Ortega, E. M. M. (2014), "The Kummer beta Birnbaum-Saunders: An alternative fatigue life distribution", *Hacettepe Journal of Mathematics and Statistics*, vol. 43 , 473 - 510.

[213] Podlaski, R. (2008), "Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution", *Canadian Journal of Forest Research*, vol. 38, 518 - 527.

[214] Pradhan, B., Kundu, D. (2013), "Inference and optimal censoring schemes for progressive censored Birnbaum-Saunders distribution", *Journal of Statistical Planning and Inference*, vol. 143, 1098 - 1108.

[215] Pourmousa, R., Jamalizadeh, A., Rezapour, M. (2015), "Multivariate normal mean-variance mixture distribution based on Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 85, 2736 - 2749.

[216] Qu, H., Xie, F. C. (2011), "Diagnostics analysis for log-Birnbaum-Saunders regression models with censored data", *Statistica Neerlandica*, vol. 65, 1 - 21.

[217] Raaijmakers, F. J. M. (1980), "The life time of standby system of units having the Birnbaum-Saunders distribution", *Journal of Applied Statistics*, vol. 17, 490 - 497.

[218] Raaijmakers, F. J. M. (1981), "Reliability of standby system for units with Birnbaum-Saunders distribution", *IEEE Transactions on Reliability*, vol. 30, 198 - 199.

[219] Reina, M.G., Vilca, F., Zeller, C.B., Balakrishnan, N. (2016), "Inverse Gaussian Birnbaum-Saunders distribution", Under review.

[220] Reyes, J., Vilca, F., Gallardo, D. I., Gómez, H. W. (2017), "Modified slash Birnbaum-Saunders distribution", *Hacettepe Journal of Mathematics and Statistics*, vol. 46, 969 - 984.

[221] Rieck, J.R. (1989), "Statistical analysis for the Birnbaum-Saunders fatigue life distribution", Ph.D. thesis, Clemson University, Department of Mathematical Sciences.

[222] Rieck, J.R. (1995), "Parametric estimation for the Birnbaum-Saunders distribution based on symmetrically censored samples", *Communications in Statistics - Theory and Methods*, vol. 24, 1721 - 1736.

[223] Rieck, J.R. (1999), "A moment-generating function with application to the Birnbaum-Saunders distribution", *Communications in Statistics - Theory and Methods*, vol. 28, 2213 - 2222.

[224] Rieck, J.R. (2003), "A comparison of two random number generators for the Birnbaum-Saunders distribution", *Communications in Statistics - Theory and Methods*, vol. 32, 929 - 934.

[225] Rieck, J.R., Nedelman, J.R. (1990), "Maximum-likelihood estimation for log-linear models for the Birnbaum-Saunders distribution", Technical Report 1, State University of New York at Albany, Department of Biometry and Statistics.

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

[227] Riquelme, M., Leiva, V., Galea, M., Sanhueza, A., (2011), "Influence diagnostics on the coefficient of variation of elliptically contoured distributions", *Journal of Applied Statistics*, vol. 38, 513 - 532.

[228] Rojas, F, Leiva, V., Wanke, P., Marchant, C. (2015), "Optimization of contribution margins in food services by modeling independent component demand", *Colombian Journal of Statistics*, vol. 38, 1 - 30.

[229] Romeiro, R.G., Vilca, F., Balakrishnan, N. (2018), "A robust multivariate Birnbaum-Saunders distribution: EM-estimation", *Statistics*, vol. 52, 321 - 344.

[230] Sánchez, L., Leiva, V., Caro-Lopera, F. J., Cysneiros, F. J. A. (2015), "On matrix-variate Birnbaum-Saunders distributions and their estimation and application", *Brazilian Journal of Probability and Statistics*, vol. 29, 790 - 812.

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

[232] Sanhueza, A., Leiva, V., Flores, E. (2009), "On a length-biased life distribution based on the sinh-normal model", *Journal of the Korean Statistical Society*, vol. 38, 323 - 330.

[233] Santana, L., Vilca, F., Leiva, V. (2011), "Influence analysis in skew-Birnbaum-Saunders regression models and applications", *Journal of Applied Statistics*, vol. 38, 1633 - 1649.

[234] Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., Ahmed, S. E. (2012), "On new parameterizations of the Birnbaum-Saunders distribution", *Pakistan Journal of Statistics*, vol. 28, 1 - 26.

[235] Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., Barros, M. (2014), "On a reparameterized Birnbaum-Saunders distribution and its moments, estimation and applications", *REVSTAT-Statistical Journal*, vol. 12, 247 - 272.

[236] Santos-Neto, M., Cysneiros, F. J. A., Leiva, V., Barros, M. (2016), "Reparameterized Birnbaum-Saunders regression models with varying precision", *Electronic Journal of Statistics*, vol. 10, 2825 - 2855.

[237] Saulo, H., Balakrishnan, N., Zhu, X., Gonzales, J.F.B., Leão, J. (2017), "Estimation in generalized bivariate Birnbaum-Saunders models", *Metrika*, vol. 80, 427 - 453.

[238] Saulo, H., Leao, J., Leiva, V., Aykroyd, R.G. (2017), "Birnbaum-Saunders autoregressive conditional duration models applied to high-frequency financial data", *Statistical Papers* (to appear).

[239] Saulo, H., Leiva, V., Ruggeri, F. (2015), "Monitoring environmental risk by a methodology based on control charts", In *Theory and Practice of Risk Assessment*, pp 177 - 197, Kitsos C, Oliveira TA, Rigas A, Gulati S (eds.), Springer Proceedings in Mathematics and Statistics 136. Springer, Switzerland.

[240] Saulo, H., Leiva, V., Ziegelmann, F.A., Marchant, C. (2013), "A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data", *Stochastic Environmental Research and Risk Assessment*, vol. 27, 1479 - 1491.

[241] Seshadri, V. (1993), *The inverse Gaussian distribution; a case study in exponential families*, Oxford University Press, Oxford.

[242] Seshadri, V. (1999), *The inverse Gaussian distribution: statistical theory and applications*, Lecture Notes in Statistics No. 137, Springer, New York.

[243] Sha, N., Ng, T.L. (2017), "Bayesian inference for Birnbaum-Saunders distribution and its generalization", *Journal of Statistical Computation and Simulation*, vol. 87, 2411 - 2429.

[244] Shaked, M., Shantikumar, G. (2007), *Stochastic Orders*, Springer, New York.

[245] Shröedinger, E. (1915), "Zur Theorie der Fall-und Steigversuche an Teilchenn mit Brownscher Bewegung", *Physikalische Zeitschrift*, vol. 16, 289 - 295.

[246] Teimouri, M., Hosseini, S. M., Nadarajah, S. (2013), "Ratios of Birnbaum-Saunders random variables", *Quality Technology and Quantitative Management*, vol. 10, 457 - 481.

[247] Tsionas, E. G. (2001), "Bayesian inference in Birnbaum-Saunders regression", *Communications in Statistics - Theory and Methods*, vol. 30, 179 - 193.

[248] Upadhyay, S.K., Mukherjee, B. (2010), "Bayes analysis and comparison of accelerated Weibull and accelerated Birnbaum-Saunders models", *Communications in Statistics - Theory and Methods*, vol. 39, 195 - 213.

[249] Upadhyay, S. K., Mukherjee, B., Gupta, A. (2009), "Accelerated test system strength models based on Birnbaum-Saunders distribution: A complete Bayesian analysis and comparison", *Lifetime Data Analysis*, vol. 15, 379 - 396.

[250] Vanegas, L.H., Paula, G.A. (2017), "Log-symmetric regression models under the presence of non-informative left- or right-censored observations", *Test*, vol. 26, 405 - 428.

[251] Vanegas, L. H., Rondon, L. M., Cysneiros, F. J. A. (2012), "Diagnostic procedures in Birnbaum-Saunders nonlinear regression models", *Computational Statistics and Data Analysis*, vol. 56, 1662 - 1680.

[252] Vilca, F., Azevedo, C.L.N., Balakrishnan, N. (2017), "Bayesian inference for sinh-normal independent nonlinear regression model", *Journal of Applied Statistics*, vol. 44, 2052 - 2074.

[253] Vilca, F., Balakrishnan, N., Zeller, C. B. (2014a), "The bivariate sinh-elliptical distribution with applications to Birnbaum-Saunders distribution and associated regression and measurement error models", *Computational Statistics and Data Analysis*, vol. 80, 1 - 16.

[254] Vilca, F., Balakrishnan, N., Zeller, C. B. (2014b), "A robust extension of the bivariate Birnbaum-Saunders distribution and associated inference", *Journal of Multivariate Analysis*, vol. 124, 418 - 435.

[255] Vilca, F., Balakrishnan, N., Zeller, C. B. (2015), "Multivariate skew-normal generalized hyperbolic distribution and its properties", *Journal of Multivariate Analysis*, vol. 128, 73 - 85.

[256] Vilca, F., Leiva, V. (2006), "A new fatigue life model based on the family of skew-elliptic distributions", *Communications in Statistics - Theory and Methods*, vol. 35, 229 - 244.

[257] Vilca, F., Romeiro, R. G., Balakrishnan, N. (2016), "A bivariate Birnbaum-Saunders regression model", *Computational Statistics and Data Analysis*, vol. 97, 169 - 183.

[258] Vilca, F., Santana, L., Leiva, V., Balakrishnan, N. (2011), "Estimation of extreme percentiles in Birnbaum-Saunders distributions", *Computational Statistics and Data Analysis*, vol. 55, 1665 - 1678.

[259] Vilca, F., Sanhueza, A., Leiva, V., Christakos, G. (2010), "An extended Birnbaum-Saunders model and its applications in the study of environmental quality in Santiago, Chile", *Stochastic Environmental Research and Risk Analysis*, vol. 24, 771 - 782.

[260] Villegas, C., Paula, G.A., Leiva, V. (2011), "Birnbaum-Saunders mixed models for censored reliability data analysis", *IEEE Transactions on Reliability*, vol. 60, 748 - 758.

[261] Volodin, I. N., Dzhungurova, O. A. (2000), "On limit distribution emerging in the generalized Birnbaum-Saunders model", *Journal of Mathematical Sciences*, vol. 99, 1348 - 1366.

[262] Wang, B. X. (2012), "Generalized interval estimation for the Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 56, 4320 - 4326.

[263] Wang, Z.H., Desmond, A.F., Lu, X. W. (2006), "Modified censored moment estimation for the two-parameter Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 50, 1033 - 1051.

[264] Wang, R.-H., Fei, H.-L. (2004), "Statistical analysis for the Birnbaum-Saunders fatigue life distribution under type-II bilateral censoring and multiply type-II censoring", *Chinese Quarterly Journal of Mathematics*, vol. 19, 126 - 132.

[265] Wang, R.-H., Fei, H.-L. (2006), "Statistical analysis for the Birnbaum-Saunders fatigue life distribution under multiply type-II censoring", *Chinese Quarterly Journal of Mathematics*, vol. 21, 15 - 27.

[266] Wang, M., Park, C., Sun, X. (2015), "Simple robust parameter estimation for the Birnbaum-Saunders distribution", *Journal of Statistical Distributions and Applications*, vol. 2, 1 - 11.

[267] Wang, M., Sun, X., Park, C. (2016), "Bayesian analysis of Birnbaum-Saunders distribution via the generalized ratio-of-uniforms method", *Computational Statistics*, vol. 31, 207 - 225.

[268] Wang, M., Zhao, J., Sun, X., Park, C. (2013), "Robust explicit estimation of the two-parameter Birnbaum-Saunders distribution", *Journal of Applied Statistics*, vol. 40, 2259 - 2274.

[269] Wanke, P., Ewbank, H., Leiva, V., Rojas, F. (2016), "Inventory management for new products with triangularly distributed demand and lead-time", *Computers and Operations Research*, vol. 69, 97 - 108.

[270] Wanke, P., Leiva, V. (2015), "Exploring the potential use of the Birnbaum-Saunders distribution in inventory management", *Mathematical Problems in Engineering*, Art. ID 827 - 846.

[271] Wu, C.-J., Tsai, T.-R. (2005), "Acceptance sampling plans for Birnbaum-Saunders distribution under truncated life tests", *International Journal of Reliability, Quality and Safety Engineering*, vol. 12, 507 - 519.

[272] Wu, J., Wong, A.C.M. (2004), "Improved interval estimation for the two-parameter Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 47, 809 - 821.

[273] Xiao, Q. C., Liu, Z. M., Balakrishnan, N., Lu, X. W. (2010), "Estimation of the Birnbaum-Saunders regression model with current status data", *Computational Statistics and Data Analysis*, vol. 54, 326 - 332.

[274] Xie, F.-C., Wei, B.-C. (2007), "Diagnostics analysis for log-Birnbaum-Saunders regression models", *Computational Statistics and Data Analysis*, vol. 47, 809 - 821.

[275] Xu, A., Tang, Y. (2010), "Reference analysis for Birnbaum-Saunders distribution", *Computational Statistics and Data Analysis*, vol. 54, 185 - 192.

[276] Xu, A. C., Tang, Y. C. (2011), "Bayesian analysis of Birnbaum-Saunders distribution with partial information", *Computational Statistics and Data Analysis*, vol. 55, 2324 - 2333.

[277] Xu, W., Yu, W., Li, Z. (2016), "Confidence interval estimation by a joint pivot method", *Computational Statistics*, vol. 31, 497 - 511.

[278] Zhang, Y., Lu, X., Desmond, A.F. (2016), "Variable selection in a log-linear Birnbaum-Saunders regression model for high dimensional survival data via the elastic-net and stochastic EM", *Technometrics*, vol. 58, 383 - 392.

[279] Zhang, C., Shi, Y., Bai, X., Fu, Q. (2017), "Inference for constant-stress accelerated life tests with dependent competing risks from bivariate Birnbaum-Saunders distribution based on adaptive progressively hybrid censoring", *IEEE Transactions on Reliability*, vol. 66, 111 - 122.

[280] Zhu, X., Balakrishnan, N. (2015), "Birnbaum-Saunders distribution based on Laplace kernel and some properties and inferential issues", *Statistics and Probability Letters*, vol. 101, 1 - 10.

[281] Ziane, Y., Zougab, N., Adjabi, S. (2018), "Birnbaum-Saunders power-exponential kernel density estimation and Bayes local bandwidth selection for nonnegative heavy tailed data", *Computational Statistics*, vol. 33, 299 - 318.