On Multivariate Log Birnbaum-Saunders Distribution

Debasis Kundu[1]

Abstract

Univariate Birnbaum-Saunders distribution has received a considerable attention in recent years. Rieck and Nedelman (1991) introduced a log Birnbaum-Saunders distribution. We introduce a multivariate log Birnbaum-Saunders distribution and discuss its different properties. It is observed that the proposed multivariate model can be obtained from the multivariate Gaussian copula. We have proposed the maximum likelihood estimators of the unknown parameters. Since it is a computationally challenging problem, particularly if the dimension is high, we have considered the approximate maximum likelihood estimators based on the Copula structure using two step procedure. The asymptotic distributions of both these estimators have been obtained. We compare their performances using Monte Carlo simulations, and it is observed that their performances are very similar in nature. One data set has been analyzed for illustrative purposes.

Key Words and Phrases: Birnbaum-Saunders distribution; Gaussian copula; Fisher information matrix; maximum likelihood estimators; Shannon entropy.

AMS Subject Classifications: 62F10, 62F03, 62H12.

[1] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

A two-parameter failure time model was proposed by Birnbaum and Saunders (1969a, 1969b) to analyze fatigue failure time data. It has been obtained by taking a monotone transformation of the Gaussian distribution. Since then, Birnbaum-Saunders distribution has received a considerable attention in the statistical literature. It has several desirable properties, and an interesting physical interpretation also. Birnbaum-Saunders distribution is closely related to the inverse Gaussian distribution. An extensive review on Birnbaum-Saunders distribution till mid 90's can be obtained in Johnson et al. (1994). For some recent references the readers are referred to Cordeiro and Lemonte (2011), Balakrishnan et al. (2010), Kundu et al. (2010), Lemonte (2012, 2013a, 2013b, 2013c) and the references cited therein.

Rieck and Nedelman (1991) introduced log Birnbaum-Saunders distribution, which has a support on the whole real line and it is a symmetric distribution. It can be both unimodal and bimodal, and it has several other interesting properties also. Log Birnbaum-Saunders distribution is more flexible than many other symmetric distributions. This model can be used quite successfully in the regression set up. Lemonte (2012) proposed a log Birnbaum-Saunders regression model with asymmetric error. Kundu (2015a) proposed a bivariate log Birnbaum-Saunders distribution and established its different properties. Bivariate log Birnbaum-Saunders distribution has been obtained from bivariate Birnbaum-Saunders distribution. Different inferential procedures have also been established. Bivariate sinh-normal distribution and bivariate sinh-elliptical distributions have been introduced by Kundu (2015b) and Vilca et al. (2014), respectively, along the same line, of which log Birnbaum-Saunders distribution is a special case.

It is quite common in practice to fit normal distribution to a symmetric data set. Log Birnbaum-Saunders distribution is a symmetric distribution and it is more flexible than the normal distribution. Due to this reason it has been used in the literature to fit a symmetric data when the normal distribution does not provide a very good fit. Similar situation is quite likely to occur even in the multivariate case also. In this case the marginals may be symmetric, but normal distribution may not provide a very good fit to all the marginals. Therefore, using multivariate normal distribution to analyze these data sets may not be appropriate. If it is observed that log Birnbaum-Saunders distribution provides a good fit to the marginals, then the multivariate log Birnbaum-Saunders distribution may be used in place of the normal distribution. The practitioner has one more option to choose from the class of multivariate distribution functions for analyzing a symmetric multivariate data set.

In this paper we introduce multivariate log Birnbaum-Saunders (MLBS) distribution. We establish several non-trivial properties of this proposed MLBS distribution. The multivariate log Birnbaum-Saunders distribution can be obtained from the multivariate Gaussian copula. Generation from a MLBS distribution is quite simple using a multivariate normal generator, hence simulation experiments can be performed quite conveniently in this case. Using the copula structure dependency properties of the MLBS distribution can be easily established. We have provided a result based on one dimensional integration only to compute the Shannon entropy of any multivariate distribution, which can be obtained by taking a monotone transformation of a multivariate normal distribution. The result has been used to compute the Shannon entropy of a MLBS distribution. It is observed that the Shannon entropy is an increasing function of the scale parameters, and it is increasing to a constant.

The maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in closed form. For a p-dimensional MLBS distribution, the MLEs of the unknown parameters can be obtained by solving a p-dimensional optimization problem. For large p, it is a computationally challenging problem. To avoid that, we propose an approximate maximum likelihood estimators (AMLEs) which can be obtained in two steps, utilizing the

Gaussian copula structure. The asymptotic properties of the MLEs and AMLEs have been established. Simulation experiments have been performed to compare the performances of the MLEs and AMLEs, and it is observed that their performances are very similar in nature. One real data set of 231 students of the marks of three different subjects has been analyzed, and it is observed that the proposed MLBS fits quite well to the above data set. It fits marginally better than the multivariate normal distributions.

Rest of the paper is organized as follows. In Section 2, first we provide some preliminaries, then we introduce the multivariate log Birnbaum-Saunders distribution and discuss its several properties. In Section 3, we discuss different estimation procedures. The simulation experiments and the analysis of one data set has been provided in Section 4. Finally we conclude the paper in Section 5.

2 Multivariate log Birnbaum-Saunders Distribution

2.1 Multivariate Birnbaum-Saunders Distribution

Kundu et al. (2012) introduced a p-variate Birnbaum-Saunders distribution as follows. The random vector $\boldsymbol{T} = (T_1, \ldots, T_p)^T$ is said to have a p-variate Birnbaum-Saunders distribution, if it has the following joint CDF;

$$F_{\boldsymbol{T}}(\boldsymbol{t}) = P(T_1 \le t_1, \ldots, T_p \le t_p) = \Phi_p(a(t_1; \alpha_1, \beta_1), \ldots, a(t_p; \alpha_p, \beta_p)), \qquad (1)$$

where $\alpha_1 > 0, \ldots, \alpha_p > 0, \beta_1 > 0, \ldots, \beta_p > 0$, $\Phi_p(\cdot; \boldsymbol{\Sigma})$ is the CDF of a multivariate standard normal distribution function with correlation matrix $\boldsymbol{\Sigma} > \boldsymbol{0}$, i.e. $\boldsymbol{\Sigma}$ is a positive definite matrix with all the diagonal elements to be one, and

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

Therefore, a p-variate Birnbaum-Saunders distribution function has the PDF

$$f_{\boldsymbol{T}}(\boldsymbol{t}) = \phi_p(a(t_1;\alpha_1,\beta_1),\ldots,a(t_p;\alpha_p,\beta_p);\boldsymbol{\Sigma})\prod_{i=1}^{p} A(t_i;\alpha_i,\beta_i), \tag{2}$$

where $\phi_p(\cdot,\boldsymbol{\Sigma})$ is the PDF of a standard p-variate normal distribution, i.e.

$$\phi_p(\boldsymbol{t};\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{t}}; \quad \boldsymbol{t} \in \mathbb{R}^p, \tag{3}$$

and

$$A(t;\alpha,\beta) = \frac{d}{dt}a(t;\alpha,\beta) = \frac{1}{2\alpha\beta}\left\{\left(\frac{\beta}{t}\right)^{1/2} + \left(\frac{\beta}{t}\right)^{3/2}\right\} = \frac{t+\beta}{2\alpha\sqrt{\beta}t^{3/2}}.$$

From now on, a p-variate Birnbaum-Saunders random variable with the CDF (2) will be denoted by MBS$(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma},p)$. Here $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_p)^T$ and $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^T$.

2.2 MLBS: Joint, Conditional and Marginal Distributions

A p-variate random vector $\boldsymbol{Y} = (Y_1,\ldots,Y_P)^T$ is said to have a p-variate log Birnbaum-Saunders distribution, if it has the following joint CDF

$$F_{\boldsymbol{Y}}(\boldsymbol{y}) = P(Y_1 \le y_1,\ldots,Y_p \le y_p) = \Phi_p\left\{\frac{2}{\alpha_1}\sinh\left(\frac{y_1 - \ln\beta_1}{2}\right),\ldots,\frac{2}{\alpha_p}\sinh\left(\frac{y_p - \ln\beta_p}{2}\right);\boldsymbol{\Sigma}\right\}. \tag{4}$$

The associated joint PDF becomes

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{\prod_{i=1}^{p}\alpha_i}\phi_p\left\{\frac{2}{\alpha_1}\sinh\left(\frac{y_1 - \ln\beta_1}{2}\right),\ldots,\frac{2}{\alpha_p}\sinh\left(\frac{y_p - \ln\beta_p}{2}\right);\boldsymbol{\Sigma}\right\}\times \prod_{i=1}^{p}\cosh\left(\frac{y_i - \ln\beta_i}{2}\right), \tag{5}$$

and it will be denoted by MLBS$(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma},p)$, where $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_p)^T$ and $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^T$.

It follows that if $Y_i = \ln T_i$, for $i = 1,\ldots,p$, then $\boldsymbol{T} = (T_1,\ldots,T_p)^T \sim$ MBS$(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma},p)$. The following relation can be easily established. Suppose $\boldsymbol{Z} = (Z_1,\ldots,Z_P)^T \sim N_p(\boldsymbol{0},\boldsymbol{\Sigma})$, a p-variate normal distribution with the mean vector $\boldsymbol{0}$, and the dispersion matrix $\boldsymbol{\Sigma}$. If

$$Y_i = 2\text{arcsinh}\left(\frac{\alpha_i Z_i}{2}\right) + \ln\beta_i; \quad i = 1,\ldots,p, \tag{6}$$

where $\text{arcsinh}(x) = \ln(x + \sqrt{x^2+1})$, for $-\infty < x < \infty$, then $\boldsymbol{Y} = (Y_1, \ldots, Y_p)^T \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$. Therefore, using (6), generation from a MLBS can be easily performed. Since $\text{arcsinh}(x)$ is an odd function of x, from (6) it immediately follows that Y_i is symmetric about $\ln \beta_i$, hence the mean and median of Y_i are both $\ln \beta_i$.

The following result will provide the marginals, the conditional distributions and one characterization of the MLBS distribution. We use the following notations.

$$u_i = \frac{2}{\alpha_1} \sinh\left(\frac{y_1 - \ln \beta_1}{2}\right), \quad i = 1, \ldots, p, \quad \boldsymbol{u} = (u_1, \ldots, u_p)^T.$$

For $0 < q < p$, we make the following partitions.

$$\boldsymbol{Y}^{(1)} = (Y_1, \ldots, Y_q)^T, \quad \boldsymbol{Y}^{(2)} = (Y_{q+1}, \ldots, Y_p)^T,$$

similarly, $\boldsymbol{\alpha}^{(1)}$, $\boldsymbol{\alpha}^{(2)}$, $\boldsymbol{\beta}^{(1)}$, $\boldsymbol{\beta}^{(2)}$, $\boldsymbol{u}^{(1)}$ and $\boldsymbol{u}^{(2)}$ are also defined. The matrix $\boldsymbol{\Sigma}$ is partitioned as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix},$$

here $\boldsymbol{\Sigma}_{11}$ is a $q \times q$ matrix, other matrices are defined so that they are compatible. Further we define

$$\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

THEOREM 1: Suppose $\boldsymbol{Y} = (Y_1, \ldots, Y_p)^T \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$. Then

(a) $\boldsymbol{Y}^{(1)} \sim \text{MLBS}(\boldsymbol{\alpha}^{(1)}, \boldsymbol{\beta}^{(1)}, \boldsymbol{\Sigma}_{11}, q)$.

(b) $\boldsymbol{Y}^{(2)} \sim \text{MLBS}(\boldsymbol{\alpha}^{(2)}, \boldsymbol{\beta}^{(2)}, \boldsymbol{\Sigma}_{22}, q - p)$.

(c) The conditional PDF of $\boldsymbol{Y}^{(1)}|\boldsymbol{Y}^{(2)} = \boldsymbol{y}^{(2)}$, $f(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)})$, can be written as follows:

$$f(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)}) = \frac{1}{\prod_{i=1}^{q} \alpha_i} \prod_{i=1}^{q} \cosh\left(\frac{y_i - \ln \beta_i}{2}\right) g(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)}),$$

where

$$g(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)}) = \frac{1}{(2\pi)^{q/2}\sqrt{|\boldsymbol{\Sigma}_{11.2}|}} e^{-\frac{1}{2}\left[\boldsymbol{u}^{(1)} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{u}^{(2)}\right]^T \boldsymbol{\Sigma}_{11.2}^{-1}\left[\boldsymbol{u}^{(1)} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{u}^{(2)}\right]}.$$

(d) The conditional CDF of $\boldsymbol{Y}^{(1)}|\boldsymbol{Y}^{(2)} = \boldsymbol{y}^{(2)}$, $F(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)})$, can be written as follows:

$$F(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)}) = \Phi_q(\boldsymbol{u}^{(1)} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{u}^{(2)}; \boldsymbol{\Sigma}_{11.2}).$$

(e) Suppose $\mathcal{A}$ is the set of all $p \times p$ matrices, whose each row and each column has exactly one non-zero element either 1 or -1. Then $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$, if and only if $\boldsymbol{AY} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{A\beta}, \boldsymbol{A\Sigma A}^T, p)$, for all $\boldsymbol{A} \in \mathcal{A}$.

Proof: (a) and (b) can be obtained from the definition. (c) can be obtained by using the conditional PDF of a multivariate normal distribution. To prove (d), we use the following notations $\boldsymbol{v}^{(1)} = (v_1, \ldots, v_q)^T$ and $\boldsymbol{w}^{(1)} = (w_1, \ldots, w_q)^T$ where

$$v_i = \frac{2}{\alpha_i}\sinh\left(\frac{w_i - \ln\beta_i}{2}\right); i = 1, \ldots, q \quad \text{and} \quad a_i = \frac{2}{\alpha_i}\sinh\left(\frac{y_i - \ln\beta_i}{2}\right); i = 1, \ldots, q.$$

Therefore,

$$\begin{aligned} F(\boldsymbol{y}^{(1)}|\boldsymbol{y}^{(2)}) &= \int_{-\infty}^{y_1} \cdots \int_{-\infty}^{y_q} f(\boldsymbol{w}^{(1)}|\boldsymbol{y}^{(2)}) dw_1 \cdots dw_p \\ &= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_q} g(\boldsymbol{v}^{(1)}|\boldsymbol{y}^{(2)}) dv_1 \cdots dv_p. \end{aligned}$$

Now make the following one to one transformation from $\boldsymbol{v}^{(1)}$ to $\boldsymbol{z}^{(1)} = (z_1, \ldots, z_q)^T$, where $\boldsymbol{z}^{(1)} = \boldsymbol{v}^{(1)} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{u}^{(2)}$, and the result follows. (e) 'if part': suppose $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$. First observe that if $\boldsymbol{Z} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$, then $\boldsymbol{AZ} \sim N_p(\boldsymbol{0}, \boldsymbol{A\Sigma A}^T)$, hence it follows that $\boldsymbol{AY} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{A\beta}, \boldsymbol{A\Sigma A}^T, p)$ from the relation (6). Since $\boldsymbol{I} \in \mathcal{A}$, 'only if' part is obvious. ∎

2.3 Copula Representation and Dependency Properties

Now we will discuss the multivariate total positivity of order two (MTP_2) and the multivariate reverse rule of order two (MRR_2) properties of the joint PDF of the MLBS distribution, in the sense of Karlin and Rinott (1980). We shall be using the following notations now. For any two real numbers, a and b, let $a \vee b = \min\{a, b\}$ and $a \wedge b = \max\{a, b\}$. For $\boldsymbol{x} = (x_1, \ldots, x_p)^T$

and $\boldsymbol{y} = (y_1, \ldots, y_p)^T$, $\boldsymbol{x} \vee \boldsymbol{y} = (x_1 \vee y_1, \ldots, x_p \vee y_p)$ and $\boldsymbol{x} \wedge \boldsymbol{y} = (x_1 \wedge y_1, \ldots, x_p \wedge y_p)$. Let us recall that a function $g : \mathbb{R}^p \to \mathbb{R}^+$ is said to be MTP$_2$ (MRR$_2$) in the sense of Karlin and Rinott (1980), if $g(\boldsymbol{x})g(\boldsymbol{y}) \leq (\geq) g(\boldsymbol{x} \vee \boldsymbol{y})g(\boldsymbol{x} \wedge \boldsymbol{y})$ for all $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{y} \in \mathbb{R}^p$. We have the following result.

THEOREM 2: Let $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$. If $\alpha_1 = \ldots = \alpha_p$, $\beta_1 = \ldots = \beta_p$, and all the off diagonal elements of $\boldsymbol{\Sigma}^{-1}$ are less (greater) than or equal to zero, then the PDF of $\boldsymbol{Y}$ has MTP$_2$ (MRR$_2$) property.

PROOF: Let us assume $\alpha_1 = \ldots = \alpha_p = \alpha$ and $\beta_1 = \ldots = \beta_p = \beta$. We take $\boldsymbol{y}_1 = (y_{11}, \ldots, y_{1p})^T$ and $\boldsymbol{y}_2 = (y_{21}, \ldots, y_{2p})^T$. To show that the PDF of $\boldsymbol{Y}$ has MTP$_2$ (MRR$_2$) property, it is enough to show that

$$\boldsymbol{u}_1^T \Sigma^{-1} \boldsymbol{u}_1 + \boldsymbol{u}_2^T \Sigma^{-1} \boldsymbol{u}_2 \geq (\leq) (\boldsymbol{u}_1 \vee \boldsymbol{u}_2)^T \Sigma^{-1} (\boldsymbol{u}_1 \vee \boldsymbol{u}_2) + (\boldsymbol{u}_1 \wedge \boldsymbol{u}_2)^T \Sigma^{-1} (\boldsymbol{u}_1 \wedge \boldsymbol{u}_2), \quad (7)$$

where $\boldsymbol{u}_1 = (u_{11}, \ldots, u_{1p})^T$ and $\boldsymbol{u}_2 = (u_{21}, \ldots, u_{2p})^T$, and

$$u_{ij} = \frac{2}{\alpha} \sinh\left(\frac{y_{ij} - \beta}{2}\right); \quad i = 1, 2, j = 1, \ldots, p.$$

If the elements of $\boldsymbol{\Sigma}^{-1}$ are denoted by $((\sigma^{kj}))$, for $k, j = 1, \ldots, p$, then proving (7) is equivalent to proving

$$\sum_{\substack{k,j=1 \\ k \neq j}}^{p} (u_{1k} u_{1j} + u_{2k} u_{2j}) \sigma^{kj} \geq (\leq) \sum_{\substack{k,j=1 \\ k \neq j}}^{p} ((u_{1k} \wedge u_{2k})(u_{1j} \wedge u_{2j}) + (u_{1k} \vee u_{2k})(u_{2j} \vee u_{2j})) \sigma^{kj}. \quad (8)$$

For all $k, j = 1, \ldots, p$, by taking any ordering of $u_{1j}, u_{1k}, u_{2j}, u_{2k}$, it easily follows that

$$(u_{1k} u_{1j} + u_{2k} u_{2j}) \leq (u_{1k} \wedge u_{2k})(u_{1j} \wedge u_{2j}) + (u_{1k} \vee u_{2k})(u_{2j} \vee u_{2j}),$$

hence the result follows as $\sigma^{jk} \leq (\geq) 0$. ∎

If $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$, then it immediately follows that

$$F_{\boldsymbol{Y}}(y_1, \ldots, y_p) = C_G(F_1(y_1), \ldots, F_p(y_p)); \quad (y_1, \ldots, y_p) \in (-\infty, \infty) \times \cdots (-\infty, \infty), \quad (9)$$

where $C_G(u_1,\ldots,u_p)$ is the multivariate Gaussian copula and it has the following form

$$\begin{aligned} C_G(u_1,\ldots,u_p) &= \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_p)} \phi_p(x_1,\ldots,x_p;\boldsymbol{\Sigma})dx_1\ldots dx \\ &= \Phi_p(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_p);\boldsymbol{\Sigma}). \end{aligned}$$

The multivariate Gaussian copula density takes the form;

$$c_G(u_1,\ldots,u_p) = \frac{\partial^p}{\partial u_1,\ldots,\partial u_p} C_G(u_1,\ldots,u_p) = \frac{\phi_p(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_p);\boldsymbol{\Sigma})}{\phi(\Phi^{-1}(u_1))\ldots\phi(\Phi^{-1}(u_p))}. \qquad (10)$$

We have the following result on the multivariate Gaussian copula density.

Lemma 1: If all the off diagonal elements of $\boldsymbol{\Sigma}^{-1}$ are less (greater) than or equal to zero, then the multivariate Gaussian copula density has MTP_2 (MRR_2) property.

Proof: It follows along the same line as Theorem 2.

Corollary 1: Since MTP_2 (MRR_2) property is a copula property, Nelsen (2006), it follows that if $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma},p)$ and if all the off diagonal elements of $\boldsymbol{\Sigma}^{-1}$ are less (greater) than or equal to zero, then the PDF of $\boldsymbol{Y}$ has MTP_2 (MRR_2) property. It is interesting to observe that using the copula property more general results than Theorem 2, can be easily obtained.

Corollary 2: Similar result can be obtained in case of MBS distribution also. If $\boldsymbol{Y} \sim \text{MBS}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma},p)$ and if all the off diagonal elements of $\boldsymbol{\Sigma}^{-1}$ are less (greater) than or equal to zero, then the PDF of $\boldsymbol{Y}$ has MTP_2 (MRR_2) property. This is a stronger result than Theorem 3 of Kundu et al. (2012), where the same result has been established under the equality assumptions on the elements of $\boldsymbol{\alpha}$'s and $\boldsymbol{\beta}$'s.

2.4 Shannon Entropy and Mutual Information Index

The entropy is an important concept in information theory, and it has been very well studied in case of multivariate normal distribution. In this section we present the Shannon entropy based on one dimensional integration only for a multivariate distribution, which can be obtained by taking a monotone transformation of a multivariate normal distribution. Let us recall the following. Suppose $\boldsymbol{X} \in \mathbb{R}^m$ and $\boldsymbol{Y} \in \mathbb{R}^k$, are two random vectors, with joint and marginal PDFs, $p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})$, $p_{\boldsymbol{X}}(\boldsymbol{x})$ and $p_{\boldsymbol{Y}}(\boldsymbol{y})$, respectively. Then the mutual information index between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined by

$$I_{\boldsymbol{X}\boldsymbol{Y}} = E\left[\ln\left\{\frac{p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{X},\boldsymbol{Y})}{p_{\boldsymbol{X}}(\boldsymbol{X})p_{\boldsymbol{Y}}(\boldsymbol{Y})}\right\}\right] = \int_{\mathbb{R}^m}\int_{\mathbb{R}^k} \ln\left\{\frac{p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})}{p_{\boldsymbol{X}}(\boldsymbol{x})p_{\boldsymbol{Y}}(\boldsymbol{x})}\right\} p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y}. \tag{11}$$

The entropy of the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ are defined as

$$H_{\boldsymbol{X}} = -E\left(\ln\{p_{\boldsymbol{X}}(\boldsymbol{X})\}\right) = -\int_{\mathbb{R}^m} \ln\{p_{\boldsymbol{X}}(\boldsymbol{x})\}p_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}, \tag{12}$$

and

$$H_{\boldsymbol{Y}} = -E\left(\ln\{p_{\boldsymbol{Y}}(\boldsymbol{Y})\}\right) = -\int_{\mathbb{R}^k} \ln\{p_{\boldsymbol{Y}}(\boldsymbol{y})\}p_{\boldsymbol{Y}}(\boldsymbol{y})d\boldsymbol{y}, \tag{13}$$

respectively. It is immediate that the mutual information index $I_{\boldsymbol{X}\boldsymbol{Y}}$ between $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be computed as

$$I_{\boldsymbol{X}\boldsymbol{Y}} = H_{\boldsymbol{X}} + H_{\boldsymbol{Y}} - H_{\boldsymbol{X}\boldsymbol{Y}}. \tag{14}$$

Here $H_{\boldsymbol{X}}$ and $H_{\boldsymbol{Y}}$ are same as defined before, and $H_{\boldsymbol{X}\boldsymbol{Y}}$ is the joint entropy of $(\boldsymbol{X},\boldsymbol{Y})$. It is well known that if $\boldsymbol{X}$ is a m-variate normal distribution with mean vector $\boldsymbol{\mu}$ and dispersion matrix $\boldsymbol{\Sigma}$, then

$$H_{\boldsymbol{X}} = \frac{1}{2}\ln|\boldsymbol{\Sigma}| + \frac{m}{2}(1+\ln(2\pi)). \tag{15}$$

We present the following general results, which can be used to compute Shannon entropy and mutual information index for MLBS, MBS and some other related distributions. It may have some independent interest also. Suppose for fixed parameter vector δ, $a(y;\delta)$ is an increasing function of $y \in \mathbb{R}$, and there exits a unique function $b(y;\delta)$, such that $\frac{d}{dy}a(y;\delta) = a'(y,\delta) = b(a(y;\delta),\delta)$. Suppose the random vectors $\boldsymbol{X}^{(1)} \in \mathbb{R}^m$, $\boldsymbol{X}^{(2)} \in \mathbb{R}^k$,

$(\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}) \in \mathbb{R}^{m+k}$, have the PDFs, respectively, as follows:

$$f_{\boldsymbol{X}^{(1)}}(\boldsymbol{x}^{(1)}) = \phi_m(a(x_1, \delta_1), \ldots, a(x_m, \delta_m); \boldsymbol{\Sigma}_{11}) \prod_{i=1}^{m} a'(x_i; \delta_i), \tag{16}$$

$$f_{\boldsymbol{X}^{(2)}}(\boldsymbol{x}^{(2)}) = \phi_k(a(x_{m+1}, \delta_{m+1}), \ldots, a(x_{m+k}, \delta_{m+k}); \boldsymbol{\Sigma}_{22}) \prod_{i=1}^{k} a'(x_{m+i}; \delta_{m+i}), \tag{17}$$

$$f_{\boldsymbol{X}^{(1)},\boldsymbol{X}^{(2)}}((\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}) = \phi_m(a(x_1, \delta_1), \ldots, a(x_{m+k}, \delta_{m+k}); \boldsymbol{\Sigma}) \prod_{i=1}^{m+k} a'(x_i; \delta_i), \tag{18}$$

here $\boldsymbol{x}^{(1)} = (x_1, \ldots, x_m)$ and $\boldsymbol{x}^{(2)} = (x_{m+1}, \ldots, x_{m+k})$, $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{22}$ and $\boldsymbol{\Sigma}_{22}$ are correlation matrices, with the following decomposition;

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}.$$

Theorem 3: Suppose, $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ are random vectors as defined above, and $Z \sim N(0, 1)$, then

(a) $H_{\boldsymbol{X}^{(1)}} = \frac{1}{2} \ln |\boldsymbol{\Sigma}_{11}| + \frac{m}{2}(1 + \ln(2\pi) - \sum_{i=1}^{m} E\left(\ln b(Z, \delta_i)\right)$

(b) $H_{\boldsymbol{X}^{(2)}} = \frac{1}{2} \ln |\boldsymbol{\Sigma}_{22}| + \frac{k}{2}(1 + \ln(2\pi) - \sum_{i=1}^{k} E\left(\ln b(Z, \delta_{m+i})\right)$

(c) $I_{\boldsymbol{X}^{(1)}\boldsymbol{X}^{(2)}} = \frac{1}{2}\left[\ln |\boldsymbol{\Sigma}_{11}| + \ln |\boldsymbol{\Sigma}_{22}| - \ln |\boldsymbol{\Sigma}|\right]$

Proof: To prove (a), observe that

$$\begin{aligned} H_{\boldsymbol{X}^{(1)}} &= -\int_{\mathbb{R}^m} \ln f_{\boldsymbol{X}^{(1)}}(\boldsymbol{x}^{(1)}) f_{\boldsymbol{X}^{(1)}}(\boldsymbol{x}^{(1)}) d\boldsymbol{x}^{(1)} \\ &= \int_{\mathbb{R}^m} \left\{ -\ln \phi_m(z_1, \ldots, z_m; \boldsymbol{\Sigma}_{11}) - \sum_{i=1}^{m} \ln b(z_i; \delta_i) \right\} \phi_m(z_1, \ldots, z_m; \boldsymbol{\Sigma}_{11}) dz_1 \ldots dz_m. \end{aligned}$$

The second equality follows by taking the transformation $z_i = a(x_i, \delta_i)$, for $i = 1, \ldots, m$. Now the result follows using (15). Proof of (b) follows along the same line. Proof of (c) can be obtained using (a) and (b). ∎

We will use the following notation. If W is a χ^2 random variable with one degree of freedom, and $c > 0$, is a constant, then $E(\ln(W/4 + c)) = \gamma(c)$. Now we will present the entropy and the mutual information index of the MLBS distribution.

THEOREM 4: Let $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$, $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$ are same as in Theorem 1, then

$$H_{\boldsymbol{Y}} = \frac{1}{2}\left[\ln|\boldsymbol{\Sigma}| + p(1 + \ln(2\pi)) - \sum_{i=1}^{p} \gamma(1/\alpha_i^2)\right]$$

and

$$I_{\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}} = \frac{1}{2}\left[\ln|\boldsymbol{\Sigma}_{11}| + \ln|\boldsymbol{\Sigma}_{22}| - \ln|\boldsymbol{\Sigma}|\right].$$

PROOF: It follows from Theorem 3, by observing the fact $a(y_i, \delta_i) = \frac{2}{\alpha_i}\sinh\left(\frac{y_i - \ln\beta_i}{2}\right)$, for $i = 1, \ldots, p$. ∎

Similarly, we can present the results for MBS distribution also. Let for $Z \sim N(0, 1)$,

$$\delta(\alpha, \beta) = E\ln\left[\frac{Z^2\alpha^2 + 4 + Z\alpha\sqrt{Z^2\alpha^2 + 4}}{\sqrt{2}\alpha\beta\left[Z^2\alpha^2 + 2 + Z\alpha\sqrt{Z^2\alpha^2 + 4}\right]^{3/2}}\right]. \tag{19}$$

THEOREM 5: Let $\boldsymbol{Y} \sim \text{MBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$, $\boldsymbol{\Sigma}$, $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$ are decomposed same as in Theorem 1, then

$$H_{\boldsymbol{Y}} = \frac{1}{2}\left[\ln|\boldsymbol{\Sigma}| + p(1 + \ln(2\pi)) - \sum_{i=1}^{p} \delta(\alpha_i, \beta_i)\right]$$

and

$$I_{\boldsymbol{Y}^{(1)}\boldsymbol{Y}^{(2)}} = \frac{1}{2}\left[\ln|\boldsymbol{\Sigma}_{11}| + \ln|\boldsymbol{\Sigma}_{22}| - \ln|\boldsymbol{\Sigma}|\right].$$

PROOF: It follows from Theorem 3, by observing the fact $a(y_i, \delta_i) = \frac{1}{\alpha_i}\left[\left(\frac{t}{\beta}\right)^{1/2} - \left(\frac{\beta}{t}\right)^{1/2}\right]$, for $i = 1, \ldots, p$. ∎

Similar results can be obtained for multivariate log-normal distribution also.

3 Inference

3.1 Maximum Likelihood Estimators

In this section we discuss the maximum likelihood estimators of the unknown parameters and the associated inference based on the observed data $\{(y_{i1}, \ldots, y_{ip}); i = 1, \ldots, n\}$. The log-likelihood function without the additive constant can be written as

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}|data) = -\frac{n}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{v}_i^T\boldsymbol{\Sigma}^{-1}\boldsymbol{v}_i - n\sum_{j=1}^{p}\ln\alpha_j + \sum_{i=1}^{n}\sum_{j=1}^{p}\ln w_{ij}(\beta_j), \qquad (20)$$

here for $i = 1, \ldots, n$, $j = 1, \ldots, p$,

$$\boldsymbol{v}_i^T = \left[\frac{2}{\alpha_1}\sinh\left(\frac{y_{i1} - \ln\beta_1}{2}\right), \ldots, \frac{2}{\alpha_p}\sinh\left(\frac{y_{ip} - \ln\beta_p}{2}\right)\right] \quad \text{and} \quad w_{ij}(\beta_j) = \cosh\left(\frac{y_{ij} - \ln\beta_j}{2}\right).$$

The MLEs of the unknown parameters can be obtained by maximizing (20) with respect to the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$, which would require a $2p + \dfrac{p(p-1)}{2}$ dimensional optimization process. We use the profile likelihood approach to reduce the computational burden significantly. Observe that

$$\left[2\sinh\left(\frac{Y_1 - \ln\beta_1}{2}\right), \ldots, 2\sinh\left(\frac{Y_p - \ln\beta_p}{2}\right)\right]^T \sim N_p\left(\mathbf{0}, \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}^T\right). \qquad (21)$$

Here $\boldsymbol{D}$ is a diagonal matrix given by $\boldsymbol{D} = \text{diag}(\alpha_1, \ldots, \alpha_p)$. Therefore, for a given $\boldsymbol{\beta}$, the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ can be obtained respectively, as

$$\widehat{\alpha}_j(\beta_j) = \left[\frac{4}{n}\sum_{i=1}^{n}\sinh^2\left(\frac{y_{ij} - \ln\beta_j}{2}\right)\right]^{1/2}; \quad j = 1, \ldots, p, \qquad (22)$$

and

$$\widehat{\boldsymbol{\Sigma}}(\boldsymbol{\beta}) = \boldsymbol{P}(\boldsymbol{\beta})\boldsymbol{Q}(\boldsymbol{\beta})\boldsymbol{P}^T(\boldsymbol{\beta}). \qquad (23)$$

Here $\boldsymbol{P}(\boldsymbol{\beta})$ is a diagonal matrix given by $\boldsymbol{P}(\boldsymbol{\beta}) = \text{diag}\{1/\widehat{\alpha}_1(\beta_1), \ldots, 1/\widehat{\alpha}_p(\beta_p)\}$, and the elements of the matrix $\boldsymbol{Q}(\boldsymbol{\beta}) = ((q_{jk}(\boldsymbol{\beta}))$ are given by

$$q_{jk}(\boldsymbol{\beta}) = \frac{\sum_{i=1}^{n}\sinh\left(\frac{y_{ij} - \ln\beta_j}{2}\right)\sinh\left(\frac{y_{ik} - \ln\beta_k}{2}\right)}{\sqrt{\sum_{i=1}^{n}\sinh^2\left(\frac{y_{ij} - \ln\beta_j}{2}\right)}\sqrt{\sum_{i=1}^{n}\sinh^2\left(\frac{y_{ik} - \ln\beta_k}{2}\right)}}. \qquad (24)$$

Therefore, the MLE of $\boldsymbol{\beta}$, say $\widehat{\boldsymbol{\beta}}$ can be obtained first by maximizing the profile log-likelihood function of $\boldsymbol{\beta}$, namely $l(\widehat{\boldsymbol{\alpha}}(\boldsymbol{\beta}), \boldsymbol{\beta}, \widehat{\boldsymbol{\Sigma}}(\boldsymbol{\beta})|data)$, and once $\widehat{\boldsymbol{\beta}}$ is obtained the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ can be obtained by replacing $\boldsymbol{\beta}$ with $\widehat{\boldsymbol{\beta}}$ in (22) and (23), respectively. Hence by using profile log-likelihood method, instead of solving $2p + \frac{p(p-1)}{2}$ dimensional optimization problem, the MLE can be obtained by solving a p dimensional optimization problem only. For large p, it is a significant gain.

Now we discuss the asymptotic properties of the MLEs.

THEOREM 6: If $\boldsymbol{\theta} = (\alpha_1, \beta_1, \ldots, \alpha_p, \beta_p, \boldsymbol{\Sigma})$ is the parameter vector, and $\widehat{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$, then

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \stackrel{d}{\longrightarrow} N_m(\mathbf{0}, \boldsymbol{I}^{-1}). \tag{25}$$

Here m denotes the dimension of $\boldsymbol{\theta}$ and $m = 2p + \frac{p(p-1)}{2}$. The $m \times m$ matrix $\boldsymbol{I}$ is the expected Fisher information matrix, and the explicit expressions of the elements of the matrix $\boldsymbol{I}$ are provided in the Appendix A.

PROOF: Since the MLBS distribution is a regular family, it satisfies all the regularity conditions for the MLEs to be consistent and asymptotically normally distributed. Hence the result follows from the standard asymptotic properties of the MLEs.

If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are known, then the MLE of $\boldsymbol{\Sigma}$ is

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{D}^{-1}\boldsymbol{Q}(\boldsymbol{\beta})\boldsymbol{D}^{-1},$$

here $\boldsymbol{D}$ and $\boldsymbol{Q}(\boldsymbol{\beta})$ are same as defined before. Using (21), it follows that $\widehat{\boldsymbol{\Sigma}}$ has a Wishart distribution with parameter p and $\boldsymbol{\Sigma}$. Moreover, if only $\boldsymbol{\beta}$ is known, then using the characteristic function of the Wishart distribution, it can be shown that $\widehat{\alpha}_j^2(\beta_j)/\alpha_j^2$ follows a χ^2, distribution with one degree of freedom, for $j = 1, \ldots, p$. Moreover, if $\boldsymbol{\Sigma} = \boldsymbol{I}$, then $\widehat{\alpha}_j(\beta_j)$'s are independently distributed.

3.2 Two-Step Estimators

In this previous section we have discussed about the MLEs of the unknown parameters, and it is observed that the MLEs of the unknown parameters can be obtained by maximizing profile likelihood function and it solves a p-dimensional optimization problem. Although, the method can be implemented very conveniently for small p, but for large p, there are some issues regarding the choice of the initial values and the convergence of the algorithm. We propose to use the two-step estimators, using the idea of Joe (2005). The basic idea about the two-step estimators is the following. First estimate the parameters of all the marginal distributions, using the marginal data, and then obtain estimates of the copula parameters using the copula structure. We also obtain the asymptotic distribution of the proposed estimators.

First Step: Using the data $(y_{1j}, \ldots, y_{nj})$ first estimate (α_j, β_j), for $j = 1, \ldots, p$. This can be obtained by maximizing the log-likelihood function of the j-th marginal namely

$$l_j(\alpha, \beta) = -n \ln \alpha + \sum_{i=1}^{n} \cosh\left(\frac{y_{ij} - \ln \beta_j}{2}\right) + \sum_{i=1}^{n} \ln \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y_{ij} - \ln \beta_j}{2}\right)\right). \tag{26}$$

If we denote the estimators of α_j and β_j as $\widetilde{\alpha}_j$ and $\widetilde{\beta}_j$, respectively, then using the same procedure as the previous section, it follows that $\widetilde{\beta}_j$ can be obtained by maximizing $l_j(\widetilde{\alpha}_j(\beta_j), \beta_j)$, with respect to β_j, where $\widehat{\alpha}_j(\beta_j)$, is same as defined in (22). Once $\widetilde{\beta}_j$ is obtained, $\widetilde{\alpha}_j$ can be obtained as $\widetilde{\alpha}_j(\widetilde{\beta}_j)$.

Second Step: Let us use the following notations

$$u_{ij} = \frac{2}{\alpha_j} \sinh\left(\frac{y_{ij} - \ln \beta_j}{2}\right); \quad i = 1, \ldots, n, \quad j = 1, \ldots, p.$$

Now maximizing the copula density function with respect to $\mathbf{\Sigma}$ for known marginals, we can obtain the estimator of $\mathbf{\Sigma}$, say $\widetilde{\mathbf{\Sigma}} = ((\widetilde{\sigma}_{st}))$, and it will be

$$\widetilde{\sigma}_{st} = \frac{\sum_{i=1}^{n} u_{is} u_{it}}{\sqrt{\sum_{i=1}^{n} u_{is}^2} \sqrt{\sum_{i=1}^{n} u_{it}^2}}; \qquad s, t = 1, \ldots, p. \tag{27}$$

It is clear that the two-step procedure reduces the computational burden significantly particularly if p is large. It involves solving p one-dimensional optimization problems. These estimates may be used for initial guesses to compute the MLEs of the unknown parameters. We have the following asymptotic properties of the two-step estimators.

THEOREM 7: If $\boldsymbol{\theta} = (\alpha_1, \beta_1, \ldots, \alpha_p, \beta_p, \boldsymbol{\Sigma})$ is the parameter vector, and $\widetilde{\boldsymbol{\theta}}$ is the two-step estimator of $\boldsymbol{\theta}$, then

$$\sqrt{n}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \xrightarrow{d} N_m(\mathbf{0}, \boldsymbol{V}). \tag{28}$$

Here m denotes the dimension of $\boldsymbol{\theta}$ as in Theorem 3, and the explicit expressions of the different elements of the $p \times p$ dispersion matrix $\boldsymbol{V}$ is provided in the Appendix B.

PROOF: It mainly follows using the inference function approach of Godambe (1991). In this particular case, the proof can be obtained along the same line as in Joe (2005).

4 Simulation Results and Data Analysis

4.1 Simulation Results

In this section we present some simulation results mainly to compare the performances of the MLEs and AMLEs as proposed in the previous section. We have considered the following 4-variate log Birnbaum-Saunders model:

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1 \qquad \beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$$

and

$$\boldsymbol{\Sigma}^{1/2} = \frac{1}{\sqrt{1+\rho^2}} \begin{bmatrix} 1 & \rho & 0 & 0 \\ \rho & 1 & \rho & 0 \\ 0 & \rho & 1 & \rho \\ 0 & 0 & \rho & 1 \end{bmatrix}, \quad i.e. \quad \boldsymbol{\Sigma} = \frac{1}{1+\rho^2} \begin{bmatrix} 1+\rho^2 & 2\rho & 0\rho^2 & 0 \\ 2\rho & 1+\rho^2 & 2\rho & \rho^2 \\ \rho^2 & 2\rho & 1+\rho^2 & 2\rho \\ 0 & \rho^2 & 2\rho & 1+\rho^2 \end{bmatrix},$$

with $\rho = 0.25$. We have taken different sample sizes, and compute the MLEs and AMLEs in each case. We report the average biases and the mean squared errors (MSEs) of all the parameter values based on 1000 replications in Tables 1 to 4.

Table 1: The average biases and the MSEs of the different parameters for MLEs and AMLEs, when $n = 25$, and $\rho = 0.25$

Parameter	Bias (AMLE)	MSE (AMLE)	Bias (MLE)	MSE (AMLE)
α_1	-0.0102	0.0235	0.0093	0.0230
α_2	0.0171	0.0238	0.0181	0.0241
α_3	0.0143	0.0247	0.0136	0.0238
α_4	-0.0160	0.0232	-0.0168	0.0228
β_1	-0.0200	0.0591	0.0113	0.0601
β_2	-0.0082	0.0622	-0.0090	0.0628
β_3	-0.0154	0.0622	-0.0147	0.0618
β_4	-0.0174	0.0567	0.0153	0.0563
σ_{12}	-0.0288	0.0263	-0.0228	0.0257
σ_{13}	-0.0006	0.0402	-0.0017	0.0416
σ_{14}	-0.0012	0.0390	0.0004	0.0411
σ_{23}	-0.0358	0.0281	-0.0227	0.0254
σ_{24}	-0.0015	0.0392	0.0009	0.0397
σ_{34}	-0.0232	0.0292	-0.0230	0.0288

Some of the points are quite clear from the simulation results. First of all in both the cases the average biases and the MSEs decrease as the sample size increases. It verifies the consistency properties of all the estimates. Moreover, the performances of the AMLEs and MLEs are very similar. Therefore, for all practical purposes, we propose to use the AMLEs in place of MLEs, as it reduces computational burden considerably. Moreover, it does not need to provide the initial estimates for multidimensional optimization problem. Here we have reported the results for a specific correlation matrix, although we have performed some other simulation experiments (not reported here) for other correlation matrices also, and the results are very similar in nature.

Table 2: The average biases and the MSEs of the different parameters for MLEs and AMLEs, when $n = 50$, and $\rho = 0.25$

Parameter	Bias (AMLE)	MSE (AMLE)	Bias (MLE)	MSE (AMLE)
α_1	-0.0073	0.0101	-0.0069	0.0098
α_2	0.0159	0.0114	0.0148	0.0117
α_3	0.0134	0.0115	0.0129	0.0097
α_4	-0.0074	0.0109	-0.0067	0.0101
β_1	0.0200	0.0339	0.0118	0.0336
β_2	0.0077	0.0355	0.0082	0.0345
β_3	0.0148	0.0321	0.0137	0.0328
β_4	0.0171	0.0324	0.0147	0.0319
σ_{12}	-0.0219	0.0135	-0.0200	0.0128
σ_{13}	-0.0074	0.0196	-0.0079	0.0201
σ_{14}	-0.0014	0.0193	0.0007	0.0189
σ_{23}	-0.0339	0.0139	-0.0342	0.0146
σ_{24}	-0.0014	0.0191	-0.0009	0.0185
σ_{34}	-0.0202	0.0135	-0.0199	0.0129

4.2 Data Analysis

In this section we present the analysis of a data set for illustrative purposes. The data set in this case represents the marks of three basic subjects namely Physics (P), Chemistry (C) and Mathematics (M) of 231 students who had qualified the Joint Entrance Examination (JEE) 2009 and are presently studying in the Indian Institute of Technology Kanpur. This particular examination is a very famous nation wide examination in India, and it is being conducted at the class 12-th level for entries in different Centrally Funded Technical Institutes in India. Before 2007, these marks were confidential and were not accessible to public, but after 2007 due to the Right to Information (RTI) these marks are available to everybody. Due to paucity of space we are not presenting the marks but they can be obtained from the author on request.

Before progressing further, we present in Table 5 some basic statistics of the marks

Table 3: The average biases and the MSEs of the different parameters for MLEs and AMLEs, when $n = 75$, and $\rho = 0.25$

Parameter	Bias (AMLE)	MSE (AMLE)	Bias (MLE)	MSE (AMLE)
α_1	-0.0023	0.0067	-0.0015	0.0063
α_2	0.0151	0.0081	0.0142	0.0077
α_3	0.0127	0.0081	0.0121	0.0079
α_4	-0.0030	0.0063	-0.0034	0.0069
β_1	0.0017	0.0224	0.0011	0.0218
β_2	0.0036	0.0221	0.0038	0.0225
β_3	-0.0021	0.0212	-0.0019	0.0215
β_4	-0.0056	0.0214	0.0019	0.0210
σ_{12}	-0.0156	0.0088	-0.0160	0.0083
σ_{13}	-0.0028	0.0140	-0.0032	0.0138
σ_{14}	-0.0014	0.0134	0.0005	0.0132
σ_{23}	-0.0288	0.0094	-0.0274	0.0091
σ_{24}	-0.0012	0.0127	-0.0005	0.0119
σ_{34}	-0.0177	0.0084	-0.0172	0.0079

distribution of the three different subjects. We present the mean, standard deviation (S.D.), median (Q_2), first quartile (Q_1) and third quartile (Q_3) values for the different marginals. It is apparent from the table values that all the three marginals are from symmetric distributions.

We would like to fit MLBS distribution to this data set. First we use the two-step estimation procedure. The profile log-likelihood function of β in all the three cases are unimodal, and they are presented in Figure 1. By maximizing the profile log-likelihood

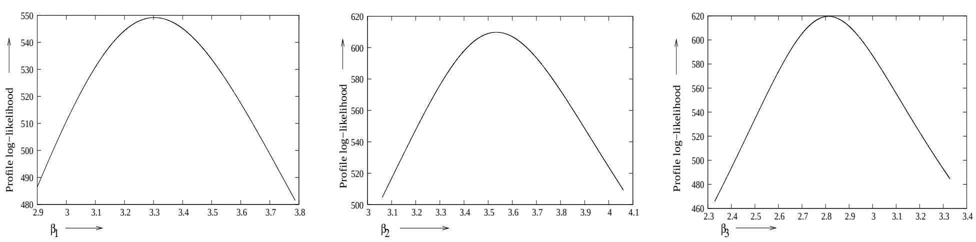

Figure 1: Profile log-likelihood function of (a) β_1, (b) β_2 and (c) β_3.

Table 4: The average biases and the MSEs of the different parameters for MLEs and AMLEs, when $n = 100$, and $\rho = 0.25$

Parameter	Bias (AMLE)	MSE (AMLE)	Bias (MLE)	MSE (AMLE)
α_1	-0.0019	0.0052	-0.0013	0.0050
α_2	0.0150	0.0058	0.0141	0.0059
α_3	0.0117	0.0061	0.0107	0.0059
α_4	-0.0005	0.0050	-0.0006	0.0052
β_1	0.0011	0.0168	0.0009	0.0170
β_2	0.0029	0.0173	0.0021	0.0169
β_3	0.0013	0.0163	0.0015	0.0161
β_4	-0.0046	0.0159	-0.0014	0.0152
σ_{12}	-0.0149	0.0067	-0.0138	0.0062
σ_{13}	-0.0015	0.0096	-0.0012	0.0099
σ_{14}	-0.0008	0.0097	0.0001	0.0093
σ_{23}	-0.0275	0.0068	-0.0218	0.0070
σ_{24}	-0.0007	0.0094	-0.0003	0.0089
σ_{34}	-0.0165	0.0062	-0.0159	0.0061

functions of β_1, β_2 and β_3, we obtain the AMLEs of β's as follows:

$$\widetilde{\beta}_1 = 3.3033, \quad \widetilde{\beta}_2 = 3.5337, \quad \widetilde{\beta}_3 = 2.8125. \tag{29}$$

The associated AMLEs of α are as follows:

$$\widetilde{\alpha}_1 = 0.1535, \quad \widetilde{\alpha}_2 = 0.1178, \quad \widetilde{\alpha}_3 = 0.1129. \tag{30}$$

The AMLE of $\mathbf{\Sigma}$ becomes;

$$\widetilde{\mathbf{\Sigma}} = \begin{bmatrix} 1.0000 & -0.1172 & -0.2420 \\ -0.1172 & 1.0000 & -0.0103 \\ -0.2420 & -0.0103 & 1.0000 \end{bmatrix}.$$

We use the AMLEs of β as the initial estimates to compute MLEs, and we obtain the following MLEs;

$$\widehat{\beta}_1 = 3.3017, \quad \widehat{\beta}_2 = 3.5324, \quad \widehat{\beta}_3 = 2.8092. \tag{31}$$

$$\widehat{\alpha}_1 = 0.1529, \quad \widehat{\alpha}_2 = 0.1201, \quad \widehat{\alpha}_3 = 0.1110. \tag{32}$$

Table 5: Basic statistics of the different marks of the JEE data

Variable	Mean	S.D.	Q_1	Q_2	Q_3
Physics	119.5	15.31	107.5	116	125.5
Chemistry	126.3	11.77	117.0	125.0	131.0
Mathematics	103.4	11.28	96.0	103.0	108.0

and

$$\widehat{\boldsymbol{\Sigma}} = \begin{bmatrix} 1.0000 & -0.1121 & -0.2401 \\ -0.1121 & 1.0000 & -0.0111 \\ -0.2401 & -0.0111 & 1.0000 \end{bmatrix}.$$

The associated log-likelihood value is 1719.4253.

Now the natural question is how good the model fits this data set. We have computed the Kolmogorov-Smirnov (KS) distances between the three marginals with the corresponding fitted LBS and the normal distribution for the data set. The KS distances and the associated p-values are reported for both the distributions in Tables 6. It is clear from the table values that LBS distribution fits marginally better than the normal distribution to all the three marginals. Now we would like to check whether the MLBS distribution fits the multivariate data set or not. For that we have used the Gaussian copula test, as suggested by Malevergne and Sornette (2003), and best on the test statistic, we cannot reject the null hypothesis. Hence, based on the marginals and Gaussian copula, we can say that MLBS fits the data reasonably well.

5 Conclusion

In this paper we have proposed a multivariate log Birnbaum-Saunders distribution, and discuss its different properties. It is a symmetric distribution, hence it can be used as

Table 6: Kolmogorov-Smirnov distance between the fitted and empirical distribution functions of the different marginals of the new born babies

Variable	MLBS (KS)	MLBS (p-value)	Normal (KS)	Normal (p-value)
Physics	0.0479	0.6567	0.0510	0.5839
Chemistry	0.0517	0.5476	0.0577	0.4256
Mathematics	0.0621	0.3889	0.0726	0.1746

an alternative to multivariate normal or multivariate t distributions. The proposed MLBS distribution can be obtained from a Gaussian copula, hence many properties of the MLBS distribution can be derived from the copula properties. The MLEs cannot be obtained in closed form, but we have proposed a two-step procedure which can be obtained quite conveniently. We have compared the performances of the MLEs and the two-step estimators, and they are very similar in nature. Hence for all practical purposes, the two-step estimators can be used. One data set has been analyzed, and it is observed that the proposed MLBS model fits the data set quite well.

Acknowledgements:

The author would like to thank one unknown referee and the associate editor for constructive suggestions which have helped to improve the manuscript significantly.

Appendix A: Fisher Information Matrix $\boldsymbol{I}$

We present the elements of the expected Fisher information matrix. For deriving the expected Fisher information matrix the following expressions are useful. Let $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$, where $\boldsymbol{Y} = (Y_1, \ldots, Y_p)^T$, $\boldsymbol{U} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{U} = (U_1, \ldots, U_p)^T$, $\boldsymbol{\Sigma} = ((\sigma_{ik}))$ and $\boldsymbol{\Sigma}^{-1} = ((\sigma^{ik}))$. Then

(a)

$$E\left\{\left(\frac{2}{\alpha_i}\text{sinh}\left(\frac{Y_i - \ln\beta_i}{2}\right) \times \frac{2}{\alpha_k}\text{sinh}\left(\frac{Y_k - \ln\beta_k}{2}\right)\right)\right\} = \sigma_{ik}, \quad i \neq k = 1, \ldots, p. \tag{33}$$

(b)

$$E\left(\frac{2}{\alpha_k}\text{sinh}\left(\frac{Y_k - \ln\beta_k}{2}\right)\right)^2 = 1, \quad k = 1, \ldots, p. \tag{34}$$

We will use the following notations:

$$E\left[\frac{4}{4 + \alpha_j^2 U_j^2}\right] = \psi_1(\alpha_j) \quad \text{and} \quad E\left(\sqrt{4 + \alpha_i^2 U_i^2} \times \sqrt{4 + \alpha_i^2 U_i^2}\right) = \psi_2(\alpha_i, \alpha_k, \sigma_{ik})$$

Moreover,

$$\frac{\partial}{\partial\sigma_{ik}}(\boldsymbol{\Sigma}^{-1}) = -\boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{ik}}\right)\boldsymbol{\Sigma}^{-1} = \boldsymbol{B}^{ik} = ((b^{ik}_{j_1,j_2})), i, k, j_1, j_2 = 1, \ldots, p.$$

$$\frac{\partial^2}{\partial\sigma_{ik}^2}(\boldsymbol{\Sigma}^{-1}) = 2\boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{ik}}\right)\boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{ik}}\right)\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\left(\frac{\partial^2\boldsymbol{\Sigma}}{\partial\sigma_{ik}^2}\right)\boldsymbol{\Sigma}^{-1} = 2\boldsymbol{A}^{ik} = 2((a^{ik}_{j_1,j_2})),$$

for $i, k, j_1, j_2 = 1, \ldots, p$. Furthermore,

$$\begin{aligned}\frac{\partial^2}{\partial\sigma_{ik}\partial\sigma_{st}}(\boldsymbol{\Sigma}^{-1}) &= -\boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{ik}}\right)\boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{st}}\right)\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{st}}\right)\boldsymbol{\Sigma}^{-1}\left(\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{ik}}\right)\boldsymbol{\Sigma}^{-1} \\ &\quad -\boldsymbol{\Sigma}^{-1}\left(\frac{\partial^2\boldsymbol{\Sigma}}{\partial\sigma_{ik}\partial\sigma_{st}}\right)\boldsymbol{\Sigma}^{-1} = -\boldsymbol{C}^{ikst} = ((c^{ikst}_{j_1,j_2})),\end{aligned}$$

for $(i, k) \neq (s, t)$ or $(i, k) \neq (t, s)$, $i, k, s, t, j_1, j_2 = 1, \ldots, p$. Let us denote

$$f(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = -\sum_{i=1}^{p}\ln\alpha_i - \frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\boldsymbol{V}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{V} + \sum_{i=1}^{p}\ln W_j(\beta_j), \tag{35}$$

where

$$\boldsymbol{V}^T = \left[\frac{2}{\alpha_1}\text{sinh}\left(\frac{y_1 - \ln\beta_1}{2}\right), \ldots, \frac{2}{\alpha_p}\text{sinh}\left(\frac{y_p - \ln\beta_p}{2}\right)\right]$$

and for $j = 1, \ldots, p$,

$$W_j(\beta_j) = \cosh\left(\frac{y_j - \ln\beta_j}{2}\right).$$

Then

$$-E\left[\frac{\partial^2 f}{\partial \alpha_i^2}\right] = \frac{1}{\alpha_i^2}\left[3\sigma^{ii} + 2\sum_{k=1,k\neq i}\sigma^{ki}\sigma_{ki} - 1\right], \quad -E\left[\frac{\partial^2 f}{\partial \alpha_i \partial \alpha_k}\right] = -\frac{1}{\alpha_i\alpha_k}\sigma_{ik}\sigma^{ik},$$

$$-E\left[\frac{\partial^2 f}{\partial \beta_i^2}\right] = \frac{1}{\beta_i^2}\left[2\sigma^{ii} + \psi_1(\alpha_i)\right], \quad \text{and} \quad -E\left[\frac{\partial^2 f}{\partial \beta_i \partial \beta_k}\right] = \frac{\sigma^{ik}}{8\alpha_i\alpha_k\beta_i\beta_k}\psi_2(\alpha_i, \alpha_k, \sigma_{ik}).$$

$$-E\left(\frac{\partial^2 f}{\partial \sigma_{ik}^2}\right) = \sum_{j_1=1}^{p_1}\sum_{j_2=1}^{p} a_{j_1,j_2}^{ik}\sigma_{j_1,j_2} + \frac{1}{2}c_{ik},$$

where

$$c_{ik} = \begin{cases} \frac{|\boldsymbol{\Sigma}|-1}{|\boldsymbol{\Sigma}|^2} & \text{if} \quad i = k \\ \frac{2|\boldsymbol{\Sigma}|-4\sigma_{ik}^2}{|\boldsymbol{\Sigma}|^2} & \text{if} \quad i \neq k \end{cases},$$

for $(i,k) \neq (s,t)$, or $(i,k) \neq (t,s)$, and for $i \neq k$, $t \neq s$, we have

$$-E\left(\frac{\partial^2 f}{\partial \sigma_{ik}\partial \sigma_{st}}\right) = -\frac{1}{2}\sum_{j_1=1}^{p_1}\sum_{j_2=1}^{p} c_{j_1,j_2}^{i,k,s,t}\sigma_{j_1,j_2} - d(i,k,s,t),$$

where $d(i,k,s,t)$ is given by

$$d(i,k,s,t) = \frac{1}{|\boldsymbol{\Sigma}|^2}\begin{cases} 2\sigma_{ik}\sigma_{st} & \text{if} \quad i \neq k, s \neq t \\ \sigma_{ik}\sigma_{ss} & \text{if} \quad i \neq k, s = t \\ \frac{1}{2} & \text{if} \quad i = k, s = t \end{cases}$$

$$-E\left(\frac{\partial^2 f}{\partial \sigma_{jk}\partial \alpha_i}\right) = -\frac{2}{\alpha_i}\sum_{m=1}^{p} b_{im}^{jk}\sigma_{im} \quad \text{and} \quad -E\left(\frac{\partial^2 f}{\partial \sigma_{jk}\partial \beta_i}\right) = 0.$$

Appendix B: Variance Covariance Matrix $\boldsymbol{V}$

In this section we present the elements of the matrix $\boldsymbol{V}$, as defined in Theorem 4. The information matrix $\boldsymbol{I}$ can be decomposed as follows:

$$\boldsymbol{I} = \begin{bmatrix} I_{11} & \dots & I_{1p} & I_{1K} \\ \vdots & \ddots & \vdots & \vdots \\ I_{p1} & \dots & I_{pp} & I_{pK} \\ I_{K1} & \dots & I_{Kp} & I_{KK} \end{bmatrix},$$

here I_{jj} for $j = 1, \ldots, p$ are 2×2 matrices, and I_{KK} is a $K \times K$ matrix where $K = p(p-1)/2$. The matrix $\boldsymbol{V}$ can be written as follows, see Joe (2005).

$$\boldsymbol{V} = (\boldsymbol{B}^{-1})\boldsymbol{M}(-\boldsymbol{B}^{-1}).$$

Here

$$-\boldsymbol{B} = \begin{bmatrix} J_{11} & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & J_{pp} & 0 \\ I_{K1} & \ldots & I_{Kp} & I_{KK} \end{bmatrix} \quad \text{and} \quad \boldsymbol{M} = \begin{bmatrix} J_{11} & \ldots & J_{1p} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ J_{p1} & \ldots & J_{pp} & 0 \\ 0 & \ldots & 0 & I_{KK} \end{bmatrix},$$

J_{jj} is the 2×2 Fisher information matrix from the j-th univariate log-likelihood, and J_{jk} is the 2×2 matrix of the following form

$$J_{jk} = \begin{bmatrix} E\left(\frac{\partial l_j}{\partial \alpha_j}\frac{\partial l_k}{\partial \alpha_k}\right) & E\left(\frac{\partial l_j}{\partial \alpha_j}\frac{\partial l_k}{\partial \beta_k}\right) \\ E\left(\frac{\partial l_j}{\partial \beta_j}\frac{\partial l_k}{\partial \alpha_k}\right) & E\left(\frac{\partial l_j}{\partial \beta_j}\frac{\partial l_k}{\partial \beta_k}\right) \end{bmatrix}.$$

Here l_j and l_k are the log-likelihood function of the j-th and k-th marginal, respectively.

To compute the elements of J_{jj} and J_{jk}, we use the following notations and results. We denote $U(a, b, z)$ as the Triconi confluent hypergeometric function introduced by Triconi (1947) as follows:

$$U(a, b, z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}dt.$$

If $Z \sim N(0, 1)$, then

$$\psi_3(c) = E\sqrt{1 + cZ^2} = \frac{\sqrt{c}}{\sqrt{2}}U(1/2, 2, 1/2c),$$

$$\psi_4 = E(Z^2\sqrt{1 + Z^2}) = \frac{1}{2^{3/2}}U(3/2, 3, 1/2).$$

If (Z_1, Z_2) is a bivariate normal random vector, with $E(Z_1) = E(Z_2) = 0$, $V(Z_1) = V(Z_2) = 1$, and $Cov(Z_1, Z_2) = \rho$, then let us denote

$$\psi_5(\rho) = E\left(Z_1Z_2\sqrt{1 + Z_1^2}\sqrt{1 + Z_2^2}\right).$$

Let $\boldsymbol{Y} \sim \text{MLBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, p)$, where $\boldsymbol{Y} = (Y_1, \ldots, Y_p)^T$, $\boldsymbol{U} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$, $\boldsymbol{U} = (U_1, \ldots, U_p)^T$, $\boldsymbol{\Sigma} = ((\sigma_{ik}))$ and $\boldsymbol{\Sigma}^{-1} = ((\sigma^{ik}))$. Then for $i \neq k = 1, \ldots, p$

(a)

$$E\left\{\left(\frac{2}{\alpha_i}\sinh\left(\frac{Y_i-\ln\beta_i}{2}\right)\times\frac{2}{\alpha_k}\sinh\left(\frac{Y_k-\ln\beta_k}{2}\right)\right)^2\right\}=1+2\sigma_{ik}^2. \quad (36)$$

(b)

$$E\left\{\left(\frac{2}{\alpha_i}\sinh\left(\frac{Y_i-\ln\beta_i}{2}\right)\right)^2\times\frac{2}{\alpha_k}\sinh\left(\frac{Y_k-\ln\beta_k}{2}\right)\times\frac{2}{\alpha_k}\cosh\left(\frac{Y_k-\ln\beta_k}{2}\right)\right\}=0. \quad (37)$$

If the matrix J_{jj} is written as follows:

$$J_{JJ}=\begin{bmatrix} a_{11}^j & a_{12}^j \\ a_{21}^j & a_{22}^j \end{bmatrix},$$

then

$$a_{11}^j=\frac{5}{\alpha_j^2},\quad a_{12}^j=a_{21}^j=0,\quad a_{22}^j=\frac{2}{\alpha_j^2\beta_j^2}+\frac{1}{\beta_j^2}-\frac{1}{4\beta_j^2}\psi_3(\alpha_j^2/4)$$

If the matrix J_{jj} is written as follows:

$$J_{Jk}=\begin{bmatrix} a_{11}^{jk} & a_{12}^{jk} \\ a_{21}^{jk} & a_{22}^{jk} \end{bmatrix},$$

then

$$a_{11}^{jk}=\frac{1}{\alpha_j\alpha_k}-2\left[\frac{1}{\alpha_j\alpha_k^2}+\frac{1}{\alpha_k\alpha_j^2}\right]+\frac{4}{\alpha_j\alpha_k}\left(1+2\sigma_{jk}\right),\qquad a_{12}^{jk}=a_{21}^{jk}=0,$$

$$a_{22}^{jk}=\frac{\alpha_j\alpha_k}{16\beta_j\beta_k}\sigma_{jk}+\frac{\rho_{jk}(\alpha_j+\alpha_k)\psi_4}{4\beta_j\beta_k}+\frac{1}{\beta_j\beta_k}\psi_5(\rho_{jk}).$$

References

[1] Balakrishnan, N., Gupta, R.C., Kundu, D., Leiva, V. and 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.

[2] Birnbaum, Z.W., Saunders, S.C. (1969a), "A new family of life distributions", *Journal of Applied Probability*, vol. 6, 319–327.

[3] 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.

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

[5] Godambe, V.P. (1991), *Estimating functions*, Oxford University Press, Oxford.

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

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994), *Continuous Univariate Distribution*, Volume 1, John Wiley and Sons, New York, USA.

[8] Karlin, S. and Rinott, Y. (1980), "Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions", *Journal of Multivariate Analysis*, vol. 10, 467498.

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

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

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

[12] Kundu, D., Balakrishnan, N. and Jamalizadeh, A. (2012), "Generalized multivariate BirnbaumVSaunders distributions and related inferential issues", *Journal of Multivariate Analysis*, Vol. 116, 230-244.

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

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

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

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

[17] 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.

[18] Lemonte, A.J., Simas, A.B., Cribari Neto, F. (2008), "Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution", *Journal of Statistical Computation and Simulation*, vol. 78, 37 – 49.

[19] Malevergne, Y. and Sornette, D. (2003), 'Testing the Gaussian copula hypothesis for financial assets dependencies", *Quantitative Finance*, vol. 3, 4, 231 - 250.

[20] McLachlan, G.J. and Peel, D. (2000), *Finite mixture models*, Wiley, New York.

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

[22] 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.

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

[24] Triconi, F.G. (1947), “Sulle funzioni ipergeometriche confluent”, *Ann. Mat. Pura Appl.*, vol. 26, 141 - 175. (In Italian).

[25] Vilca, F., Balakrishnan, N. and Zeller, C.B. (2014), “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.