On Bivariate and Mixture of Bivariate Birnbaum-Saunders Distributions

Mohsen Khosravi[1], Debasis Kundu[2], Ahad Jamalizadeh[1]

Abstract

Univariate Birnbaum-Saunders distribution has received a considerable amount of attention during the last few years. Recently, Kundu et al. [9] introduced a bivariate Birnbaum-Saunders distribution. It is observed that the mixture of bivariate Birnbaum-Saunders distributions can be written as the weighted mixture of bivariate inverse Gaussian distribution and its reciprocals. In this paper further we introduce a mixture of two bivariate Birnbaum-Saunders distributions and discuss its different properties. The mixture model has eleven parameters, hence it is a very flexible model. The maximum likelihood estimators cannot be obtained in explicit forms. We propose to use EM algorithm to compute the maximum likelihood estimators. It is observed that it saves computational time significantly. We performed some simulation experiments, and one data analysis has been performed to illustrate the EM algorithm. It is observed that the performance of the EM algorithm is quite satisfactory.

Key Words and Phrases: Birnbaum-Saunders distribution; bivariate Birnbaum-Saunders distribution; maximum likelihood estimators; EM algorithm; asymptotic distribution; bivariate inverse Gaussian distribution; Fisher information matrix.

[1]Department of Statistics, Faculty of Mathematics & Computer, Shahid Bahonar University of Kerman, Kerman,Iran,76169-14111.

[2] 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

Birnbaum and Saunders [2, 3] introduced a two-parameter lifetime distribution mainly for modelling the failure time distribution for fatigue failure caused under cyclic loading. The cumulative distribution function (CDF) and the corresponding probability density function (PDF) of a two-parameter Birnbaum-Saunders random variable T are

$$F_T(t;\alpha,\beta) = \Phi(a(t;\alpha,\beta)); \quad 0 < t < \infty, \quad \alpha,\beta > 0, \tag{1}$$

and

$$f_T(t;\alpha,\beta) = \phi(a(t;\alpha,\beta))A(t;\alpha,\beta), \tag{2}$$

respectively, where

$$a(t;\alpha,\beta) = \left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{1/2} - \left(\frac{\beta}{t}\right)^{1/2}\right\}\right], \tag{3}$$

$$A(t;\alpha,\beta) = \frac{d}{dt}a(t;\alpha,\beta) = \frac{t+\beta}{2\alpha\sqrt{\beta}t^{3/2}}, \tag{4}$$

$\Phi(\cdot)$ is the CDF and $\phi(\cdot)$ is the PDF of a standard normal random variable. Here α is the shape parameter and β is the scale parameter. From now on, a Birnbaum-Saunders random variable with the CDF (1) or PDF (2) will be denoted by BS(α,β).

Since the introduction of the model considerable amount of work has been taken place on the development of Birnbaum-Saunders distribution and its different generalizations. Recently, Kundu et al. [9] introduced a bivariate Birnbaum-Saunders distribution with five parameters. The bivariate Birnbaum-Saunders distribution has been obtained from the bivariate normal distribution, using similar transformation as the univariate Birnbaum-Saunders distribution. They discussed different properties of the bivariate Birnbaum-Saunders distribution, and provided several inferential issues.

The main aim of this paper is to consider a mixture of two bivariate Birnbaum-Saunders distributions with different set of parameters. Mixture distribution plays an important role in different data analysis purposes, see for example McLachlan and Peel [10]. If the data are coming from different subpopulations, and their identifications are not known, then the mixture distribution can be used quite effectively to analyze that data set. Unlike normal distribution, Birnbaum-Saunders distribution is a skewed distribution. Therefore, if the subpopulations do not have symmetric distribution, mixture of Birnbaum-Saunder distributions may be used to analyze these data. Moreover, a multimodal distribution can be approximated very well by a mixture distribution. Although, extensive work has been done on a mixture multivariate normal distribution, not much work has been done on a mixture of multivariate non-normal distributions. This is an attempt towards that direction.

Recently, Balakrishnan et al. [1] studied different mixtures of univariate Birnbaum-Saunders models. In this paper we study different properties of a mixture two bivariate Birnbaum-Saunders distributions. It is observed that the joint PDF can be unimodal or bimodal depending on the parameter values. Different moments and product moments are obtained. Desmond [5] established that univariate Birnbaum-Saunders distribution can be written as a mixture of inverse Gaussian distribution and its reciprocals. Using that result, Balakrishnan et al. [1] observed that the mixtures of univariate Birnbaum-Saunders distributions can be written as the mixture of inverse Gaussian distributions and their reciprocals. Similar results have been established even for bivariate case also.

The bivariate Birnbaum-Saunders distribution has eleven parameters. The maximum likelihood estimators (MLEs) of the unknown parameters can be obtained by solving eleven dimensional optimization problem, although they cannot be obtained in closed form. We propose to use EM algorithm to compute the MLEs. While implementing the EM algorithm, at each 'E'-step the corresponding 'M'-step can be performed by solving two two-dimensional (2-D) optimization problems. Therefore, it can be performed quite conveniently. We have performed some simulation experiments to verify the performances of the MLEs. One real data set has been analyzed to illustrate different aspects of the EM algorithm, and observe its performances. It is observed that the proposed model and the EM algorithm work quite well.

The rest of the paper is organized as follows. In Section 2, we provide some preliminaries. The mixture of two bivariate Birnbaum-Saunders distributions is introduced in Section 3 and discuss its different properties in the same section. We discuss the EM algorithm in Section 4. In Section 5, we provide the simulation results, and the analysis of the real data set. Finally we conclude the paper in Section 6.

2 Preliminaries

2.1 Univariate Birnbaum-Saunders Distribution

A positive random variable T is said to have a two-parameter Birnbaum-Saunders distribution if it has the CDF and PDF as given in (1) and (2), respectively. It is known that if

$$T \sim \text{BS}(\alpha, \beta) \quad \Leftrightarrow \quad X = a(T; \alpha, \beta) \sim N(0, 1). \tag{5}$$

The following connection between Birnbaum-Saunders and inverse Gaussian distributions has been introduced by Desmond [5]. Suppose $X_1 \sim \text{IG}(\mu, \lambda)$ with parameters $\mu > 0$ and $\lambda > 0$. The PDF of X_1 is given by

$$f_{X_1}(x; \mu, \lambda) = \sqrt{\frac{\lambda}{x^3}} \phi\left(a(x; \sqrt{\mu/\lambda}, \mu)\right); \qquad x > 0. \tag{6}$$

Furthermore, let X_2 be a random variable such that $1/X_2 \sim \text{IG}(1/\mu, \lambda/\mu^2)$, which is independent of X_1. Consider, the random variable T, where

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

The PDF of T becomes

$$f_T(x;\mu,\lambda) = \frac{1}{2} f_{X_1}(x;\mu,\lambda) + \frac{1}{2} f_{X_2}(x;\mu,\lambda); \qquad x > 0. \tag{7}$$

Here $f_{X_1}(x;\mu,\lambda)$ and $f_{X_2}(x;\mu,\lambda)$ are the PDFs of X_1 and X_2, respectively. Moreover, the PDF of X_2 can be written as

$$f_{X_2}(x;\mu,\lambda) = \frac{x f_{X_1}(x;\mu,\lambda)}{\mu}; \quad x > 0, \tag{8}$$

where $\mu = E(X_1)$. Hence X_2 is the length biased version of X_1.

2.2 Bivariate Birnbaum-Saunders Distribution

Kundu et al. [9] introduced a bivariate extension of the BS(α,β) model. The bivariate random vector $\boldsymbol{T} = (T_1, T_2)^T$ is said to have a bivariate Birnbaum-Saunders distribution with parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho) = (\alpha_1, \alpha_2, \beta_1, \beta_2, \rho)$, denoted by BBS$(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$, if the joint CDF of $\boldsymbol{T}$, is of the following form;

$$F_{\boldsymbol{T}}(t_1, t_2) = P(T_1 \le t_1, T_2 \le t_2) = \Phi_2(a(t_1;\alpha_1,\beta_1), a(t_2;\alpha_2,\beta_2);\rho); \quad t_1 > 0, t_2 > 0, \tag{9}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$, $-1 < \rho < 1$, and $\Phi_2(u, v; \rho)$ is the CDF of a bivariate standard normal random vector with correlation coefficient ρ.

If $(T_1, T_2)^T \sim$ BBS$(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$, then it has the PDF

$$f_{\boldsymbol{T}}(t_1, t_2) = \phi_2(a(t_1;\alpha_1,\beta_1), a(t_2;\alpha_2,\beta_2);\rho) \times \prod_{i=1}^{2} A(t_i;\alpha_i,\beta_i), \tag{10}$$

where

$$\phi_2(u, v; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(u^2 + v^2 - 2\rho uv)\right\}, \tag{11}$$

is the standard bivariate normal probability density function with correlation coefficient ρ, and $A(.)$ is same as defined in (4). The density surface of a bivariate Birnbaum-Saunders distribution is unimodal and it can take different shapes, see for example [9].

Kundu et al. [9] developed several interesting properties of the bivariate Birnbaum-Saunders distribution. It is observed that if $(T_1, T_2)^T$ has a bivariate Birnbaum-Saunders distribution, the marginals have univariate Birnbaum-Saunders distribution. T_1 and T_2 are positively correlated if $\rho > 0$, and for $\rho < 0$, they are negatively correlated. One important aspect of the bivariate Birnbaum-Saunders distribution is that it can have correlation over the entire range namely from -1 to 1. Moreover, the generation from a bivariate Birnbaum-Saunders distribution can be easily performed using bivariate normal distribution.

2.3 Bivariate Inverse Gaussian Distribution

Kocherlakota [8] introduced a bivariate generalization of the inverse Gaussian distribution. The bivariate random vector $\boldsymbol{X} = (X_1, X_2)^T$ is said to have the bivariate inverse Gaussian distribution with parameters $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^T$ and ρ, if the joint PDF of $\boldsymbol{X}$ is given by

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\lambda}, \rho) &= \frac{1}{2}\sqrt{\frac{\lambda_1\lambda_2}{x_1^3 x_2^3}}\Big[\phi_2\left(a(x_1; \sqrt{\mu_1/\lambda_1}, \mu_1), a(x_2; \sqrt{\mu_2/\lambda_2}, \mu_2); \rho\right) \\ &\quad + \phi_2\left(a(x_1; \sqrt{\mu_1/\lambda_1}, \mu_1), a(x_2; \sqrt{\mu_2/\lambda_2}, \mu_2); -\rho\right)\Big] \\ &= \frac{1}{2}\sqrt{\frac{\lambda_1\lambda_2}{x_1^3 x_2^3}}\phi_2\left(a(x_1; \sqrt{\mu_1/\lambda_1}, \mu_1), a(x_2; \sqrt{\mu_2/\lambda_2}, \mu_2); \rho\right) \\ &\quad \times \left(1 + exp\left(-\frac{2\rho}{1-\rho^2} a(x_1; \sqrt{\mu_1/\lambda_1}, \mu_1), a(x_2; \sqrt{\mu_2/\lambda_2}, \mu_2)\right)\right) \end{aligned} \tag{12}$$

for $x_1 \geq 0$, $x_2 \geq 0$, where $\mu_1 > 0$, $\mu_2 > 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, $-1 < \rho < 1$. From now on a bivariate distribution with PDF (12) will be denoted by $\text{BIG}(\lambda_1, \lambda_2, \mu_1, \mu_2, \rho)$.

It is a natural generalization of the univariate inverse Gaussian distribution to the bivariate case, and it corrects the PDF originally proposed by Wasan [14]. This five-parameter distribution has unimodal PDF, the marginals X_1 and X_2 are $\text{IG}(\mu_1, \lambda_1)$ and $\text{IG}(\mu_2, \lambda_2)$, respectively. It is known that if $X \sim \text{IG}(\mu, \lambda)$, then $Z = \dfrac{\lambda(X-\mu)^2}{\mu^2 X}$ has a chi-square distribution with one degree of freedom. Similarly, if $(X_1, X_2)^T \sim \text{BIG}(\lambda_1, \lambda_2, \mu_1, \mu_2, \rho)$, then $(Z_1, Z_2)^T$, where $Z_i = \dfrac{\lambda_i (X_i - \mu_i)^2}{\mu_i^2 X_i}$; $i = 1, 2$, has a bivariate chi-square distribution with one degree of freedom.

3 Mixture of Bivariate Birnbaum-Saunders Distributions

3.1 Definition and Some Properties

A random vector $(X_1, X_2)^T$ is said to have a mixture of two bivariate Birnbaum-Saunders distributions, if the joint joint PDF of $(X_1, X_2)^T$ can be written as

$$f_{X_1,X_2}(x_1, x_2) = p f_{U_1,U_2}(x_1, x_2; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1) + (1-p) f_{V_1,V_2}(x_1, x_2; \boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \rho_2), \tag{13}$$

here $(U_1, U_2)^T \sim \text{BBS}(\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1)$, $(V_1, V_2)^T \sim \text{BBS}(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \rho_2)$, $\boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{12})^T$, $\boldsymbol{\alpha}_2 = (\alpha_{21}, \alpha_{22})^T$, $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12})^T$, $\boldsymbol{\beta}_2 = (\beta_{21}, \beta_{22})^T$, $\alpha_{ij} > 0$, $\beta_{ij} > 0$, for all $i, j = 1, 2$, $-1 < \rho_1, \rho_2 < 1$, and $0 \le p \le 1$. From now on a bivariate random vector with joint PDF (13), will be called a mixture of bivariate Birnbaum-Saunders distributions (MBBS), and it will be denoted by $\text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho_1, \rho_2, p)$. Therefore, the mixture has eleven parameters. The joint CDF of $(X_1, X_2)^T$, with PDF (13) can be written as

$$\begin{aligned} P(X_1 \le x_1, X_2 \le x_2) &= p\Phi\left(a(x_1; \alpha_{11}, \beta_{11}), a(x_2; \alpha_{12}, \beta_{12}); \rho_1\right) \\ &\quad + (1-p)\Phi\left(a(x_1; \alpha_{21}, \beta_{21}), a(x_2; \alpha_{22}, \beta_{22}); \rho_2\right). \end{aligned} \tag{14}$$

Due to presence of eleven parameters, the joint CDF of MBBS can take different shapes. In Figure 1, we provide the PDF surface of MBBS, for different parameter values. It is clear that it can take different shapes. The PDF surface of MBBS can be unimodal or bimodal, therefore, it is clearly more flexible than the bivariate Birnbaum-Saunders distribution.

Using the mixture representation, it may be noted that the MBBS model can be expressed as follows

$$(X_1, X_2)^T = \begin{cases} (U_1, U_2)^T & \text{with probability} \quad p \\ (V_1, V_2)^T & \text{with probability} \quad (1-p), \end{cases} \tag{15}$$

where $(U_1, U_2)^T$ and $(V_1, V_2)^T$ are same as defined above. Kundu et al. [9] provided a simple method to generate samples from a BBS distribution, hence using the mixture representation (15), random samples from a MBBS can be generated quite conveniently.

The following results can be easily obtained using the representation (15). We will use the following notations. $\boldsymbol{\beta}_1^{-1} = (\beta_{11}^{-1}, \beta_{12}^{-1})^T$ and $\boldsymbol{\beta}_2^{-1} = (\beta_{21}^{-1}, \beta_{22}^{-1})^T$.

Theorem 1: If $(X_1, X_2)^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho_1, \rho_2, p)$, then

(a) $(X_1, X_2^{-1})^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2^{-1}, \rho_1, -\rho_2, p)$.

(b) $(X_1^{-1}, X_2)^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1^{-1}, \boldsymbol{\beta}_2, -\rho_1, \rho_2, p)$.

(c) $(X_1^{-1}, X_2^{-1})^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1^{-1}, \boldsymbol{\beta}_2^{-1}, -\rho_1, -\rho_2, p)$.

Proof: The proof mainly follows using Theorem 3.2 of Kundu et al. [9] and the mixture representation (15). ∎

Theorem 2: If $(X_1, X_2)^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho_1, \rho_2, p)$, then

$$\begin{aligned}
E(X_1) &= p\beta_{11}\left(1 + \frac{1}{2}\alpha_{11}^2\right) + (1-p)\beta_{12}\left(1 + \frac{1}{2}\alpha_{12}^2\right) \\
E(X_1^{-1}) &= p\beta_{11}^{-1}\left(1 + \frac{1}{2}\alpha_{11}^2\right) + (1-p)\beta_{12}^{-1}\left(1 + \frac{1}{2}\alpha_{12}^2\right) \\
E(X_2) &= p\beta_{21}\left(1 + \frac{1}{2}\alpha_{21}^2\right) + (1-p)\beta_{22}\left(1 + \frac{1}{2}\alpha_{22}^2\right) \\
E(X_2^{-1}) &= p\beta_{21}^{-1}\left(1 + \frac{1}{2}\alpha_{21}^2\right) + (1-p)\beta_{22}^{-1}\left(1 + \frac{1}{2}\alpha_{22}^2\right) \\
E(X_1X_2) &= pE(U_1U_2) + (1-p)E(V_1V_2),
\end{aligned}$$

where

$$\begin{aligned}
E(U_1U_2) &= \beta_{11}\beta_{12}\left[1 + \frac{1}{2}(\alpha_{11}^2 + \alpha_{12}^2) + \frac{1}{4}\alpha_{11}^2\alpha_{12}^2(1+\rho_1^2) + \alpha_{11}\alpha_{12}I_1\right] \\
E(V_1V_2) &= \beta_{21}\beta_{22}\left[1 + \frac{1}{2}(\alpha_{21}^2 + \alpha_{22}^2) + \frac{1}{4}\alpha_{21}^2\alpha_{22}^2(1+\rho_2^2) + \alpha_{21}\alpha_{22}I_2\right].
\end{aligned}$$

and the exact expressions of I_1 and I_2 are presented in the Appendix A.

Proof: The proof mainly follows using Theorem 1, and the expressions of $E(U_1U_2)$, $E(V_1V_2)$ obtained in Kundu et al. [9].

Theorem 3: If $(X_1, X_2)^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho_1, \rho_2, p)$, where $\beta_{11} = \beta_{21} = \beta_1$ and $\beta_{12} = \beta_{22} = \beta_2$, then

$$\begin{pmatrix} a(X_1, 1, \beta_1) \\ \\ a(X_2, 1, \beta_2) \end{pmatrix} \stackrel{d}{=} pN_2(\mathbf{0}, \boldsymbol{\Sigma}_1) + (1-p)N_2(\mathbf{0}, \boldsymbol{\Sigma}_2),$$

where $\stackrel{d}{=}$ means equal in distribution and

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} \alpha_{11}^2 & \alpha_{11}\alpha_{12}\rho_1 \\ \alpha_{12}\alpha_{11}\rho_1 & \alpha_{12}^2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} \alpha_{21}^2 & \alpha_{21}\alpha_{22}\rho_2 \\ \alpha_{21}\alpha_{22}\rho_2 & \alpha_{21}^2 \end{bmatrix}.$$

Proof: The result follows using expressions (5) and (15). ∎

Theorem 4: If $(X_1, X_2)^T \sim \text{MBBS}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho_1, \rho_2, p)$, where $\beta_{11} = \beta_{21} = \beta_1$ and $\beta_{12} = \beta_{22} = \beta_2$, then $P(X_1 < X_2) = 1/2$.

Proof: From (15), it is immediate that

$P(X_1 < X_2) = pP(U_1 < U_2) + (1-p)P(V_1 < V_2)$. Since $P(U_1 < U_2) = P(V_1 < V_2) = 1/2$, the result follows. ∎

3.2 Connection Between BBS and MBIG

Desmond [5] established a connection between a BS distribution and an IG distribution. Using that result, Balakrishnan et al. [1] showed that a mixture of BS distribution can be expressed as a mixture of IG distribution and their reciprocals. In this section, we introduce a new bivariate IG distribution and we call this as Modified Bivariate Inverse Gaussian (MBIG) distribution. This MBIG model has a IG marginal, and the other marginal may be considered as a weighted BS or mixture of IG and Length-Biased IG (LIG), see for example

Balakrishnan et al. [1]. In this section we want to establish a similar connection between BBS and MBIG distributions. First we will define a MBIG distribution.

DEFINITION: A bivariate random vector $\boldsymbol{X} = (X_1, X_2)^T$ is said to have a MBIG distribution with parameters $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^T$ and ρ, if the joint PDF of X_1 and X_2 is given by

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\lambda}, \rho) = \frac{1}{2}\sqrt{\frac{\lambda_1\lambda_2}{x_1^3 x_2^3}}\phi_2\left(a(x_1; \sqrt{\mu_1/\lambda_1}, \sqrt{\mu_2/\lambda_2}; \rho)\right) \times \left(1 + \frac{x_1}{\mu_1}\right); \quad x_1 \geq 0, x_2 \geq 0. \quad (16)$$

Here $\mu_1 > 0$, $\mu_2 > 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, $-1 < \rho < 1$. From now on a random vector with the joint PDF (16) will be denoted by MBIG$(\boldsymbol{\mu}, \boldsymbol{\lambda}, \rho)$.

THEOREM 5: The function $f : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ as defined in (16) is a proper PDF.

PROOF: It is clear that $f(x_1, x_2; \boldsymbol{\mu}, \boldsymbol{\lambda}, \rho) \geq 0$, for all $x_1, x_2 \geq 0$. We will show that $c = 1$, where

$$\int_0^\infty \int_0^\infty f(x_1, x_2; \boldsymbol{\mu}, \boldsymbol{\lambda}, \rho) dx_1 dx_2 = c^{-1}.$$

From now on we will be using the following notations $\alpha_1 = \sqrt{\mu_1/\lambda_1}$ and $\alpha_2 = \sqrt{\mu_2/\lambda_2}$. Consider the following transformation

$$W_1 = a(X_1; \alpha_1, \mu_1) \quad W_2 = a(X_2; \alpha_2, \mu_2).$$

Since the Jacobian of transformation is

$$|J| = A(x_1; \alpha_1, \mu_1) \times A(x_2; \alpha_2, \mu_2),$$

it follows that the joint PDF of $\boldsymbol{W} = (W_1, W_2)$ is

$$f_{\boldsymbol{W}}(w_1, w_2) = c\left(1 - \frac{\alpha_2 w_2}{\sqrt{4 + (\alpha_2 w_2)^2}}\right)\phi_2(w_1, w_2; \rho). \quad (17)$$

Now

$$\int_{-\infty}^\infty \int_{-\infty}^\infty f_{\boldsymbol{W}}(w_1, w_2) dw_1 dw_2 \quad = \quad c\int_{-\infty}^\infty \int_{-\infty}^\infty \phi_2(w_1, w_2; \rho) dw_1 dw_2$$

$$-c \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\alpha_2 w_2}{\sqrt{4+(\alpha_2 w_2)^2}} \phi_2(w_1, w_2; \rho) dw_1 dw_2$$
$$= c - 0 = c = 1. \qquad \blacksquare$$

The following result will be useful for further development.

THEOREM 6: If $(X_1, X_2)^T \sim \text{MBIG}(\boldsymbol{\mu}, \boldsymbol{\lambda}, \rho)$, then

(i) $X_2 \sim \text{IG}(\mu_2, \lambda_2)$, where the PDF of $\text{IG}(\mu, \lambda)$ is as given in (6).

(ii) The PDF of the random variable X_1, for $x > 0$, is given by

$$f_{X_1}(x_1; \boldsymbol{\mu}, \boldsymbol{\lambda}, \rho) = \frac{\sqrt{\lambda_1}(x_1 + \mu_1)}{2\mu_1 \sqrt{x_1^3}} \times \phi\left(a(x_1; \alpha_1, \mu_1)\right) \times \psi(x_1),$$

where

$$\psi(x_1) = \left(1 - E\left[\frac{\alpha_2 Z}{\sqrt{4 + (\alpha_2 Z)^2}}\right]\right)$$

and

$$Z \sim N\left(\rho a(x_1; \alpha_1, \mu_1), 1 - \rho^2\right).$$

PROOF: Consider the same transformation W_1 and W_2 as in the Theorem 5, and it has the joint PDF given in (17). The PDF of W_2 becomes

$$f_{W_2}(w_2) = \left(1 - \frac{\alpha_2 w_2}{\sqrt{4 + (\alpha_2 w_2)^2}}\right) \phi(w_2). \tag{18}$$

Now consider the inverse transformation

$$X_2 = a^{-1}(W_2; \alpha_2, \mu_2) = \frac{\mu_2}{4}\left(\alpha_2 W_2 + \sqrt{4 + (\alpha_2 W_2)^2}\right)^2.$$

The PDF of X_2 becomes

$$f_{X_2}(x_2) = \sqrt{\frac{\lambda_2}{x_2^3}} \phi\left(a(x_2; \alpha_2, \lambda_2)\right),$$

hence the result (i) is proved. To prove (ii), let us re write (17) as

$$f_{\boldsymbol{W}}(w_1, w_2) = \left(1 - \frac{\alpha_2 w_2}{\sqrt{4 + (\alpha_2 w_2)^2}}\right) \times \phi(w_1)\phi(w_2; \rho w_1, 1 - \rho^2).$$

The PDF of W_1 becomes

$$f_{W_1}(w_1) = \int_{-\infty}^{\infty} f_{\boldsymbol{W}}(w_1, w_2) dw_2 = \phi(w_1) \int_{-\infty}^{\infty} \left(1 - \frac{\alpha_2 w_2}{\sqrt{4 + (\alpha_2 w_2)^2}}\right) \times \phi(w_2; \rho w_1, 1-\rho^2) dw_2.$$

The result (ii) follows using the inverse transformation

$$X_1 = a^{-1}(W_1; \alpha_1, \mu_1) = \frac{\mu_1}{4} \left(\alpha_1 W_1 + \sqrt{4 + (\alpha_1 W_1)^2}\right)^2.$$

∎

COMMENTS: If $(X_1, X_2)^T \sim \text{MBIG}(\boldsymbol{\mu}, \boldsymbol{\lambda}, \rho)$, then X_1 and X_2 are independent iff $\rho = 0$. In that case $\psi(x_1) = 1$.

COMMENTS: The random variable X_1 is a weighted mixture of the inverse Gaussian and length biased inverse Gaussian distribution, where the weighting proportion is $\psi(x_1)/2$, since

$$f_{X_1}(x_1; \boldsymbol{\mu}, \boldsymbol{\lambda}, \rho) = \frac{\psi(x_1)}{2} \left(f_1(x_1) + f_2(x_1)\right)$$

where

$$f_1(x_1) = \sqrt{\frac{\lambda_1}{x_1^3}} \phi\left(a(x_1; \alpha_1, \lambda_1)\right) \quad \text{and} \quad f_2(x_1) = \frac{x_1}{\mu_1} \sqrt{\frac{\lambda_1}{x_1^3}} \phi\left(a(x_1; \alpha_1, \lambda_1)\right).$$

Clearly, $f_1(x_1)$ is the PDF of an inverse Gaussian distribution, and $f_2(x_1)$ is the PDF of length biased inverse Gaussian distribution. Moreover, when X_1 and X_2 are independent, i.e. $\rho = 0$, then X_1 has a Birnbaum-Saunders distribution.

Although when $\rho \neq 0$, the function $\psi(\cdot)$ cannot be obtained in explicit form, it can be expressed in terms of confluent hypergeometric function. The details are provided in the Appendix A, for completeness purposes.

Similar to the univariate case, we have the following relation between BBS and MBIG distributions. Let $\boldsymbol{X} = (X_1, X_2)^T \sim \text{BBS}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$, $\boldsymbol{V} = (V_1, V_2)^T \sim \text{MBIG}(\boldsymbol{\mu}, \boldsymbol{\lambda}, \rho)$ with $\boldsymbol{\mu} = (\beta_1, \beta_2)^T$, $\boldsymbol{\lambda} = (\beta_1/\alpha_1^2, \beta_2/\alpha_2^2)^T$, and $\boldsymbol{W} = (W_1, W_2)^T$ be a random vector such that $(W_1^{-1}, W_2^{-1})^T \sim \text{MBIG}(\boldsymbol{\mu}', \boldsymbol{\lambda}', \rho)$, with $\boldsymbol{\mu}' = (1/\beta_1, 1/\beta_2)^T$, $\boldsymbol{\lambda}' = (1/(\beta_1\alpha_1^2), 1/(\beta_2\alpha_2^2)^T$, which is independent of $\boldsymbol{V}$. Consider a new random variable given by

$$\boldsymbol{X} = \begin{cases} \boldsymbol{V} & \text{with probability } \frac{1}{2} \\ \boldsymbol{W} & \text{with probability } \frac{1}{2}. \end{cases} \tag{19}$$

The PDF of $\boldsymbol{X}$ is

$$f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\rho) = \frac{1}{2} f_{\boldsymbol{V}}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\lambda},\rho) + \frac{1}{2} f_{\boldsymbol{W}}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\lambda},\rho), \tag{20}$$

where $f_{\boldsymbol{V}}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\lambda},\rho)$ and $f_{\boldsymbol{W}}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\lambda},\rho)$, are the densities of $\boldsymbol{V}$ and $\boldsymbol{W}$, respectively.

Note that it is a generalization of the one dimensional result to the two dimensional result. It matches with the one dimensional result for both the marginals when $\rho = 0$. Therefore, it is immediate that a MBBS as given in (13) can be expressed as a mixture of two MBIG's and their reciprocals.

4 Inference

4.1 Maximum Likelihood Estimation

In this section we discuss the estimation of the unknown parameters based on a random sample of size n, say, $\mathcal{D} = \{(x_{11}, x_{12}), \ldots, (x_{n1}, x_{n2})\}$ from a MBBS$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho_1, \rho_2, p)$. Using the joint PDF (13), the log-likelihood function can be written as

$$l(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left(p f_{U_1,U_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1) + (1-p) f_{V_1,V_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \rho_2) \right), \tag{21}$$

where the eleven dimensional dimensional vector $\boldsymbol{\theta} = (\boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \rho_2, p)$. The MLEs of the unknown parameters can be obtained by maximizing (21) with respect to the unknown parameters. The MBBS model has eleven unknown parameters, and we need to solve an eleven dimensional optimization problem to compute the MLEs.

It is well known that a mixture model can be treated as a missing value problem, and so the expectation maximization (EM) algorithm can be used quite effectively to compute the unknown model parameters, see for example McLachlan and Peel ([10], pp 47 - 50). In this case also we propose to use the EM algorithm, which reduces the computational burden significantly. For this purpose using the mixture representation (15), we assume that the complete observations are random samples from (X_1, X_2, Δ), where $\Delta = 1$, if $(X_1, X_2)^T = (U_1, U_2)^T$ and $\Delta = 0$, if $(X_1, X_2)^T = (V_1, V_2)^T$. Therefore,

$$P(\Delta = 1) = p \quad \text{and} \quad P(\Delta = 0) = (1 - p).$$

The complete samples are as follows:

$$\{(x_{11}, x_{12}, \delta_1), \ldots, (x_{n1}, x_{n2}, \delta_n)\}. \tag{22}$$

Based on the complete observations, the complete likelihood function is

$$L_c(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left(p f_{U_1,U_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1)\right)^{\delta_i} \times \left((1-p) f_{V_1,V_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \rho_2)\right)^{1-\delta_i}, \tag{23}$$

hence the log-likelihood function becomes

$$\begin{aligned} l_c(\boldsymbol{\theta}) \;&=\; \sum_{i=1}^{n} \delta_i \ln\left(f_{U_1,U_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1)\right) + \sum_{i=1}^{n} (1-\delta_i) \ln\left(f_{V_1,V_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_2, \boldsymbol{\beta}_2, \rho_2)\right) \\ &\quad + \ln p \sum_{i=1}^{n} \delta_1 + \ln(1-p) \sum_{i=1}^{n} (1-\delta_i). \end{aligned} \tag{24}$$

The EM algorithm consists of maximizing the conditional expectation of the complete data log-likelihood function, based on the observed data and the current value of $\boldsymbol{\theta}$, say $\widetilde{\boldsymbol{\theta}}$ in an iterative two-step algorithm process, see for example Dempster et al. [4]. The E-step is to compute the conditional expectation denoted by $Q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}})$, and the M step is maximizing $Q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\theta}$.

E-Step: It consists of calculating $Q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}})$, $\widetilde{\boldsymbol{\theta}}$ being the current parameter value.

$$Q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}}) = E(l(\boldsymbol{\theta}|\mathcal{D}, \widetilde{\boldsymbol{\theta}})) \;=\; \sum_{i=1}^{n} E(\delta_i|\mathcal{D}, \widetilde{\boldsymbol{\theta}}) \ln\left(f_{U_1,U_2}(x_{i1}, x_{i2}; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1, \rho_1)\right)$$

$$+\sum_{i=1}^{n} E((1-\delta_i)|\mathcal{D},\widetilde{\boldsymbol{\theta}}) \ln\left(f_{V_1,V_2}(x_{i1},x_{i2};\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2,\rho_2)\right)$$
$$+\ln p \sum_{i=1}^{n} E(\delta_i|\mathcal{D},\widetilde{\boldsymbol{\theta}}) + \ln(1-p)\sum_{i=1}^{n} E((1-\delta_i)|\mathcal{D},\widetilde{\boldsymbol{\theta}}). \quad (25)$$

In this case

$$\begin{aligned}
E(\delta_i|\mathcal{D},\widetilde{\boldsymbol{\theta}}) &= \frac{f_{U_1,U_2}(x_{i1},x_{i2};\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\beta}}_1,\widetilde{\rho}_1)}{f_{U_1,U_2}(x_{i1},x_{i2};\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\beta}}_1,\widetilde{\rho}_1)+f_{V_1,V_2}(x_{i1},x_{i2};\widetilde{\boldsymbol{\alpha}}_2,\widetilde{\boldsymbol{\beta}}_2,\widetilde{\rho}_2)} = \widetilde{a}_i \quad \text{(say)},\\
E((1-\delta_i)|\mathcal{D},\widetilde{\boldsymbol{\theta}}) &= \frac{f_{V_1,V_2}(x_{i1},x_{i2};\widetilde{\boldsymbol{\alpha}}_2,\widetilde{\boldsymbol{\beta}}_2,\widetilde{\rho}_2)}{f_{U_1,U_2}(x_{i1},x_{i2};\widetilde{\boldsymbol{\alpha}}_1,\widetilde{\boldsymbol{\beta}}_1,\widetilde{\rho}_1)+f_{V_1,V_2}(x_{i1},x_{i2};\widetilde{\boldsymbol{\alpha}}_2,\widetilde{\boldsymbol{\beta}}_2,\widetilde{\rho}_2)} = \widetilde{b}_i \quad \text{(say)}.
\end{aligned}$$

M-Step: It involves maximizing the $Q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\theta}$, to obtain $\widetilde{\widetilde{\boldsymbol{\theta}}}$, where

$$\widetilde{\widetilde{\boldsymbol{\theta}}} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}}).$$

From (25), it is clear that

$$\widetilde{\widetilde{p}} = \frac{\sum_{i=1}^{n}\widetilde{a}_i}{n},$$
$$(\widetilde{\widetilde{\boldsymbol{\alpha}}}_1,\widetilde{\widetilde{\boldsymbol{\beta}}}_1,\widetilde{\widetilde{\rho}}_1) = \arg\max_{\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1,\rho_1} \sum_{i=1}^{n}\widetilde{a}_i \ln\left(f_{U_1,U_2}(x_{i1},x_{i2};\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1,\rho_1)\right)$$

and

$$(\widetilde{\widetilde{\boldsymbol{\alpha}}}_2,\widetilde{\widetilde{\boldsymbol{\beta}}}_2,\widetilde{\widetilde{\rho}}_2) = \arg\max_{\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2,\rho_2} \sum_{i=1}^{n}\widetilde{b}_i \ln\left(f_{V_1,V_2}(x_{i1},x_{i2};\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2,\rho_2)\right).$$

It has been shown in Appendix B, that $(\widetilde{\widetilde{\boldsymbol{\alpha}}}_1,\widetilde{\widetilde{\boldsymbol{\beta}}}_1,\widetilde{\widetilde{\rho}}_1)$ and $(\widetilde{\widetilde{\boldsymbol{\alpha}}}_2,\widetilde{\widetilde{\boldsymbol{\beta}}}_2,\widetilde{\widetilde{\rho}}_2)$ can be obtained by solving two separate two-dimensional optimization problems.

We propose the following algorithm to compute the MLEs of the unknown parameters.

Algorithm:

- Step 1: Choose some initial guess values of $\boldsymbol{\theta}$, say $\boldsymbol{\theta}^{(0)}$.
- Step 2: Obtain

$$\boldsymbol{\theta}^{(1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(0)})$$

- Step 3: Continue the process until convergence takes place.

4.2 Testing of Hypotheses

In this section we discuss different testing of hypotheses problems which will have some practical importance.

Problem 1: Test the following hypothesis:

$$H_0 : \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}, \quad \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}, \rho_1 = \rho_2 = \rho \quad \text{vs.} \quad H_1 : \text{At least one is different} \tag{26}$$

The problem is of interest as it tests whether the data are coming from one population or from two different populations. We use the likelihood ratio test for this purpose. Fit a bivariate Birnbaum-Saunders distribution to the given data set and obtain the maximum likelihood estimators of the unknown parameters using the method proposed by Kundu et al. [9]. Suppose the maximized log-likelihood value becomes l_{01}, and say l_1 denote the maximized log-likelihood value based on the fitted model (13). Under H_0

$$-2(l_{01} - l_1) \sim \chi^2_6.$$

Therefore, reject the null hypothesis at the α% level of significance if

$$-2(l_0 - l_1) > \chi^2_{6,\alpha}.$$

Here χ^2_k denotes the chi-square distribution with k degrees of freedom, and $\chi^2_{k,\alpha}$ denote the upper α-th percentile point of a chi-square distribution with k degrees of freedom.

Problem 2: Test the following hypothesis:

$$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta} = (\beta_1, \beta_2) \quad \text{vs.} \quad H_1 : \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2. \tag{27}$$

This is an important problem, as it tests whether the scale parameter vectors of the two clusters are same or not. In this case, under H_0, the MLEs of the unknown parameters can be obtained using an EM algorithm as described before. Under H_0, at each M-step the maximization with respect to different parameters can be obtained as follows. First compute

$$
\begin{aligned}
(\widetilde{\widetilde{\beta}}_1, \widetilde{\widetilde{\beta}}_2) = \widetilde{\widetilde{\boldsymbol{\beta}}} \; &= \; \arg\max_{\beta_1,\beta_2} \left\{ \sum_{i=1}^{n} \widetilde{a}_i \ln\Big(f_{U_1,U_2}(x_{i1}, x_{i2}; \widetilde{\widetilde{\boldsymbol{\alpha}}}_1(\boldsymbol{\beta}), \boldsymbol{\beta}, \widetilde{\widetilde{\rho}}_1(\boldsymbol{\beta})\Big) \right. \\
&\left. \qquad + \sum_{i=1}^{n} \widetilde{b}_i \ln\Big(f_{V_1,V_2}(x_{i1}, x_{i2}; \widetilde{\widetilde{\boldsymbol{\alpha}}}_2(\boldsymbol{\beta}), \boldsymbol{\beta}, \widetilde{\widetilde{\rho}}_2(\boldsymbol{\beta})\Big) \right\}, \qquad (28)
\end{aligned}
$$

where the elements $\widetilde{\widetilde{\boldsymbol{\alpha}}}_1(\boldsymbol{\beta})$, $\widetilde{\widetilde{\boldsymbol{\alpha}}}_2(\boldsymbol{\beta})$, $\widetilde{\widetilde{\rho}}_1(\boldsymbol{\beta})$ and $\widetilde{\widetilde{\rho}}_2(\boldsymbol{\beta})$ are same as defined in Appendix B. The other elements can be obtained as described in Appendix B. If l_{02} denotes the maximized log-likelihood value under H_0 of problem 2, then

$$-2(l_{02} - l_1) \sim \chi^2_6.$$

Therefore, reject the null hypothesis at the α% level of significance if

$$-2(l_{02} - l_1) > \chi^2_{6,\alpha}.$$

5 Simulation Experiments and Data Analysis

5.1 Simulation Experiments

In this section we perform some simulation experiments to see performances of the MLEs and the proposed EM algorithm for different parameter values and for different sample sizes. We have used two different models namely: Model 1: $\alpha_{11} = \alpha_{12} = \beta_{11} = \beta_{12} = 1.0$, $\rho_1 = 0.5$, $\alpha_{21} = \alpha_{22} = \beta_{21} = \beta_{22} = 2.0$, $\rho_2 = 0.5$ and $p = 0.5$, and Model 2: $\alpha_{11} = \alpha_{12} = \beta_{11} = \beta_{12} = 1.0$, $\rho_1 = 0.5$, $\alpha_{21} = \alpha_{22} = \beta_{21} = \beta_{22} = 2.0$, $\rho_2 = 0.5$ and $p = 0.25$. Note that in Model 1 and Model 2, all the α and β parameters are same for both the groups, although their proportions are different. We have used different sample sizes namely: n = 25, 50, 75, 100. In each case we have generated the sample from the proposed MBBS distribution, and computed the MLEs of the unknown parameters based on the proposed EM algorithm.

At each E-Step the corresponding M-Step has been performed by solving two separate 2-D optimization problems. The 2-D optimization problems have been solved using grid search method with a grid size 0.0001. The iteration of the EM algorithm stops, when the difference of the relative log-likelihood values at the two consecutive iterations, is less than 10^{-4}. In all the cases considered the iteration stops before 25 iterations. We replicate the process 1000 times, and report the average estimates and the associated mean squared errors (within bracket) in each case. All the results are reported in Tables 1 and 2.

Parameter	$n = 25$	$n = 50$	$n = 75$	$n = 100$
α_{11}	1.1523 (0.2241)	1.1167 (0.1341)	1.0412 (0.0742)	1.0101 (0.0265)
α_{12}	1.1467 (0.2311)	1.1239 (0.1389)	1.0487 (0.0699)	1.0228 (0.0221)
β_{11}	1.1441 (0.2151)	1.1098 (0.1311)	1.0578 (0.0775)	1.0200 (0.0211)
β_{12}	1.1512 (0.2226)	1.1009 (0.1309)	1.0447 (0.0612)	1.0111 (0.0227)
ρ_1	0.5091 (0.0165)	0.5012 (0.0061)	0.5005 (0.0031)	0.5000 (0.0016)
α_{21}	2.1578 (0.4248)	2.1212 (0.2145)	2.0561 (0.1289)	2.0198 (0.0429)
α_{22}	2.1498 (0.4176)	2.1137 (0.2145)	2.0561 (0.1498)	2.0098 (0.0465)
β_{21}	2.1601 (0.4265)	2.1067 (0.2187)	2.0345 (0.1235)	2.0113 (0.0422)
β_{22}	2.1316 (0.4209)	2.1176 (0.2227)	2.0492 (0.1265)	2.0249 (0.0490)
ρ_2	0.5078 (0.0144)	0.5011 (0.0059)	0.5008 (0.0028)	0.5000 (0.0013)
p	0.5117 (0.0113)	0.5055 (0.0052)	0.5010 (0.0032)	0.5003 (0.0009)

Table 1: Average estimates and the associated mean squared errors of the MLEs of the unknown parameters based on 1000 replications for Model 1.

Some of the points are quite clear from the simulation experiments. In all the cases the EM algorithm converges within 25 iterations, and it indicates that the proposed EM algorithm works well in this case. It is observed that as sample size increases the average biases and mean squared errors decrease in each case. It verifies the consistency properties of the MLEs. Comparing Tables 1 and 2, it is observed that the average biases and mean squared errors of $\alpha_{11}(\alpha_{21})$, $\alpha_{12}(\alpha_{22})$, $\beta_{11}(\beta_{21})$, $\beta_{12}(\beta_{22})$, $\rho_1(\rho_2)$ are more (less) for Model 1 (Model 2) than Model 2 (Model 1). It is not very surprising, it is mainly due to that fact that p is smaller in Model 2 than Model 1.

Parameter	$n = 25$	$n = 50$	$n = 75$	$n = 100$
α_{11}	1.1911 (0.4154)	1.1641 (0.2312)	1.0979 (0.1432)	1.0542 (0.0445)
α_{12}	1.1754 (0.3978)	1.1543 (0.2276)	1.0886 (0.1387)	1.0529 (0.0451)
β_{11}	1.1667 (0.4134)	1.1600 (0.2287)	1.1010 (0.1401)	1.0498 (0.0311)
β_{12}	1.1512 (0.2226)	1.1009 (0.1309)	1.0447 (0.0612)	1.0111 (0.0399)
ρ_1	0.5182 (0.0245)	0.5113 (0.0123)	0.5019 (0.0067)	0.5006 (0.0023)
α_{21}	2.1067 (0.2156)	2.0178 (0.1231)	2.0019 (0.0787)	2.0009 (0.0245)
α_{22}	2.1102 (0.2245)	2.0098 (0.1178)	2.0023 (0.0665)	2.0008 (0.0225)
β_{21}	2.0198 (0.2278)	2.0223 (0.1156)	2.0020 (0.0698)	2.0000 (0.0200)
β_{22}	2.1100 (0.2016)	2.0054 (0.1136)	2.0073 (0.0542)	2.0004 (0.0198)
ρ_2	0.5078 (0.0144)	0.5011 (0.0059)	0.5008 (0.0028)	0.5000 (0.0013)
p	0.2563 (0.0075)	0.2523 (0.0036)	0.2515 (0.0021)	0.2503 (0.0007)

Table 2: Average estimates and the associated mean squared errors of the MLEs of the unknown parameters based on 1000 replications for Model 2.

5.2 Real Data

Researchers are interested to assess pulmonary function in non-pathological populations. For this purpose, 50 subjects (25 male and 25 female) have been chosen to run on a trade mill until exhaustion. Samples of air were collected at definite intervals and gas contents analyzed. Two measures of oxygen consumptions X_1 (resting volume O_2 in mL/kg/min) and X_2 (maximum volume O_2 in L/min) are reported in Table 6.8 of Johnson and Wichern ([7], page 266). We would like to analyze this data set without identifying the sex of the subjects. The scatter plot of both the groups is provided in Figure 2.

Before progressing further we want to check whether Birnbaum-Saunders distribution can be used or not to analyze the two groups. We have fitted the bivariate Birnbaum-Saunders distribution to each group and the estimates of the unknown parameters are provided in Table 3.

We also provide below the Kolmogorov-Smirnov distance between the empirical and fitted CDFs, and the associated p value (in bracket) for both the marginals for both Males and

Population ↓	α_1	β_1	α_2	β_2	ρ
Males	0.1975	5.2279	0.1733	3.6340	0.2399
Females	0.3316	4.9088	0.1463	2.2912	-0.1296

Table 3: Estimates of the bivariate Birnbaum-Saunders parameters for both Males and Females data.

Population ↓	X_1	X_2
Males	0.0956 (0.9760)	0.1136 (0.9010)
Females	0.1734 (0.4397)	0.0899 (0.9916)

Table 4: Kolmogorov-Smirnov distance and the associated p value in case of bivariate Birnbaum-Saunders model.

Females data in Table 4. From the tables values, it is immediate that BVBS can be used for both Males and Females.

For comparison purposes, we have also fitted bivariate log-normal distribution to both the groups. The estimates of the unknown parameters and the Kolmogorov-Smirnov distances with the associated p values for both the groups are provided in Table 5 and Table 6, respectively. Based on the Kolmogorov-Smirnov distances of the marginals, it is observed

Population ↓	μ_1	σ_1	μ_2	σ_2	ρ
Males	1.6539	0.1968	1.2897	0.1727	0.2394
Females	1.5957	0.3224	0.8288	0.1459	-0.1307

Table 5: Estimates of the bivariate log-normal parameters for both Males and Females data.

that BVBS provides a better fit than bivariate log-normal distribution for both the groups. Hence, it is reasonable to use MBBS model than a mixture of bivariate log-normal model to analyze the data set when the sex are not identified.

Since we do not have any idea about the initial guesses, we have used the method proposed by Balakrishnan et al. [1] to obtain initial estimates of the unknown parameters for both

Population ↓	X_1	X_2
Males	0.1154 (0.8933)	0.1202 (0.8626)
Females	0.1775 (0.4100)	0.0911 (0.9854)

Table 6: Kolmogorov-Smirnov distance and the associated p value for bivariate log-normal distribution.

the marginals, and they are as follows:

$$\alpha_{11}^{(0)} = 0.1869,\ \alpha_{12}^{(0)} = 0.2015,\ \beta_{11}^{(0)} = 5.2800,\ \beta_{12}^{(0)} = 3.4799,\ \rho_1^{(0)} = 0.1834,\ p_1 = 0.5906$$

$$\alpha_{21}^{(0)} = 0.3541,\ \alpha_{22}^{(0)} = 0.1329,\ \beta_{21}^{(0)} = 4.8001,\ \beta_{22}^{(0)} = 2.2297,\ \rho_2^{(0)} = -0.3224,\ p_2 = 0.5660.$$

We start the EM algorithm with the above initial estimates of α's and β's, and for p, we have used $(p_1 + p_2)/2 = 0.5783$. We use the same stopping criterion as before. The log-likelihood values correspond to each iterations are plotted in Figure 2. After 20 iterations, the EM algorithm stops. The MLEs, and the associated standard errors (based on non-parametric bootstrap method of Efron [6]) are provided below.

$$\begin{aligned} &\widehat{\alpha}_{11} = 0.1830(\mp 0.0223), \quad \widehat{\alpha}_{12} = 0.1942(\mp 0.0365), \quad \widehat{\beta}_{11} = 5.3400(\mp 0.7451), \\ &\widehat{\beta}_{12} = 3.5291(\mp 0.4459), \quad \widehat{\rho}_1 = 0.1225(\mp 0.0061), \quad \widehat{\alpha}_{21} = 0.3446(\mp 0.0281), \\ &\widehat{\alpha}_{22} = 0.1388(\mp 0.0154), \quad \widehat{\beta}_{21} = 4.7603(\mp 0.5178), \quad \widehat{\beta}_{22} = 2.2578(\mp 0.2717) \\ &\widehat{\rho}_2 = -0.3225(\mp 0.0215), \qquad \widehat{p} = 0.5468(\mp 0.0381). \end{aligned} \tag{29}$$

It is clear that EM algorithm works quite well in this case. Finally we perform the testing of hypotheses of both Problems 1 and 2. It is observed that at the 5% level of significance both the hypotheses cannot be accepted. Since H_0 as given in (26) has been rejected, it is clear that the date are coming from two sub-populations. Hence, one bivariate Birnbaum-Saunders distribution cannot be used to analyze this data set.

6 Conclusions

In this paper we have considered a bivariate Birnbaum-Saunders distribution and the mixture of two bivariate Birnbaum-Saunders distributions. It has been shown that the bivariate

Birnbaum-Saunders distribution can be written as a mixture of bivariate inverse Gaussian distribution and its reciprocals. We have further considered the mixture of two bivariate Birnbaum-Saunders distributions and established its several properties. We have provided EM algorithm to compute the MLEs of the unknown parameters, and it is observed that the proposed EM algorithm works quite well. Although, in the present paper we have considered only two components, the proposed methodologies can be extended for more than two components also. More work is needed along that direction.

Acknowledgements:

The authors would like to thank two reviewers and also the associate editor for providing many constructive suggestions.

Appendix A

In this Appendix, we provide the expressions of $\psi(x_1)$, I_1 and I_2. We denote $U\ (a,b,z)$ as the Tricomi confluent hypergeometric function introduced by Francesco Tricomi [13] as

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

The following lemmas will be useful for that purpose.

Lemma A.1: Let $X \sim N(\mu, \sigma^2)$, then

$$E\left[X\sqrt{X^2+4}\right] = \frac{8}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\mu^2}{2\sigma^2}\right] \sum_{n=0}^{\infty} \left(\frac{\Gamma(n+\frac{3}{2})}{(2n+1)!} \left[\frac{2\mu}{\sigma^2}\right]^{2n+1} U\left(n+\frac{3}{2}, n+3, \frac{2}{\sigma^2}\right)\right).$$

Proof:

$$E\left[X\sqrt{4+X^2}\right] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\mu^2}{2\sigma^2}\right] \int_{-\infty}^{\infty} x\sqrt{4+x^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \exp\left(\frac{\mu x}{\sigma^2}\right) dx.$$

Now by using the Taylor expansion of $\exp\left(\frac{\mu x}{\sigma^2}\right) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\mu x}{\sigma^2}\right)^n$, and after some simple calculations, we readily obtain the result. ∎

Lemma A.2: Let $X \sim N(\mu, \sigma^2)$, then

$$E\left[\sqrt{X^2+4}\right] = \frac{4}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\mu^2}{2\sigma^2}\right] \sum_{n=0}^{\infty} \left(\frac{\Gamma(n+\frac{1}{2})}{(2n)!}\left[\frac{2\mu}{\sigma^2}\right]^{2n} U\left(n+\frac{1}{2}, n+2, \frac{2}{\sigma^2}\right)\right).$$

Proof: It can be obtained along the same line as in Lemma 1.

Lemma A.3: Let $X \sim \mathrm{N}(\mu, \sigma^2)$, then

$$E\left[\frac{X}{\sqrt{4+X^2}}\right] = \frac{2}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\mu^2}{2\sigma^2}\right] \sum_{n=0}^{\infty} \left(\frac{\Gamma(n+\frac{3}{2})}{(2n+1)!}\left[\frac{2\mu}{\sigma^2}\right]^{2n+1} U\left(n+\frac{3}{2}, n+2, \frac{2}{\sigma^2}\right)\right).$$

Proof: It can be obtained along the same line as in Lemma 1.

Using Lemma 3, we immediately have the following expression for $\psi(\cdot)$

$$\begin{aligned}\psi(x_1) &= 1 - \frac{1}{\sqrt{2\pi(1-\rho^2)}}\sqrt{\frac{\lambda_1}{\mu_1}} \exp\left(-\frac{\rho^2}{2(1-\rho^2)} a^2(x_1; \sqrt{\mu_1/\lambda_1}, \mu_1)\right) \\ &\quad \times 2\sum_{n=0}^{\infty}\left\{\frac{\Gamma(n+\frac{3}{2})}{(2n+1)!}\left(\frac{2\rho}{1-\rho^2}\sqrt{\frac{\lambda_2}{\mu_2}} a(x_1; \sqrt{\mu_1/\lambda_1}, \mu_1)\right)^{2n+1}\right. \\ &\quad \left. \times U\left(n+\frac{3}{2}, n+2, \frac{\lambda_2}{\mu_2}\frac{2}{(1-\rho^2)}\right)\right\}.\end{aligned}$$

To compute I_1 and I_2, we have the following results.

Theorem A.1: If $\mathbf{Z} = (Z_1, Z_2)^T \sim N_2(0, 0, 1, 1, \rho)$, $|\rho| < 1$, then

$$\begin{aligned} I_1 &= E\left[Z_1 Z_2 \left(\sqrt{\alpha_1^2 Z_1^2 + 4}\right)\left(\sqrt{\alpha_2^2 Z_2^2 + 4}\right)\right] \\ &= \frac{8}{\pi\sqrt{1-\rho^2}} \sum_{n=0}^{\infty}\left(\frac{\left[\Gamma(n+\frac{3}{2})\right]^2}{(2n+1)!} \times \left[\frac{2\rho}{1-\rho^2}\right]^{2n+1}\left[\frac{2}{\alpha_1\alpha_2}\right]^{2n+3}\right. \\ &\quad \left. \times U\left(n+\frac{3}{2}, n+3, \frac{2}{\alpha_1^2(1-\rho^2)}\right) \times U\left(n+\frac{3}{2}, n+3, \frac{2}{\alpha_2^2(1-\rho^2)}\right)\right). \end{aligned}$$

Proof: We have

$$\begin{aligned} I_1 &= E\left[Z_1 Z_2 \left(\sqrt{\alpha_1^2 Z_1^2+4}\right)\left(\sqrt{\alpha_2^2 Z_2^2+4}\right)\right] \\ &= E\left\{E\left[Z_1 Z_2 \left(\sqrt{\alpha_1^2 Z_1^2+4}\right)\left(\sqrt{\alpha_2^2 Z_2^2+4}\right)|Z_1\right]\right\} \\ &= E\left\{Z_1\sqrt{\alpha_1^2 Z_1^2+4}E\left(Z_2\sqrt{\alpha_2^2 Z_2^2+4}\,|Z_1\right)\right\} \end{aligned}$$

Since, $Z_2\,|Z_1=z_1 \sim N\left(\rho z_1, 1-\rho^2\right)$, the required result can be easily obtained by using Lemma A.1. ∎

Theorem A.2 If $\mathbf{Z}=(Z_1,Z_2)^T \sim N_2(0,0,1,1,\rho)^T$, $|\rho|<1$, then

$$\begin{aligned} I_2 &= E\left[\left(\sqrt{\alpha_1^2 Z_1^2+4}\right)\left(\sqrt{\alpha_2^2 Z_2^2+4}\right)\right] \\ &= \frac{4}{\pi\sqrt{1-\rho^2}}\sum_{n=0}^{\infty}\left(\frac{\left[\Gamma(n+\frac{1}{2})\right]^2}{(2n)!}\left[\frac{2\rho}{1-\rho^2}\right]^{2n}\left[\frac{2}{\alpha_1\alpha_2}\right]^{2n+1}\right. \\ &\quad \left.\times U\left(n+\frac{1}{2}, n+2, \frac{2}{\alpha_1^2(1-\rho^2)}\right)\times U\left(n+\frac{1}{2}, n+2, \frac{2}{\alpha_2^2(1-\rho^2)}\right)\right) \end{aligned}$$

Proof: The proof can be obtained using Lemma A.2, and using the conditional distribution similarly as the proof of Theorem A.1. ∎

Appendix B

Now to compute $(\tilde{\tilde{\boldsymbol{\alpha}}}_1, \tilde{\tilde{\boldsymbol{\beta}}}_1, \tilde{\tilde{\rho}}_1)$, we use the following result similarly as in Kundu et al. [9]. If $(U_1,U_2)^T \sim$ BVBS$(\alpha_{11},\beta_{11},\alpha_{12},\beta_{12},\rho_1)$, then

$$\left\{\left(\sqrt{\frac{U_1}{\beta_{11}}}-\sqrt{\frac{\beta_{11}}{U_1}}\right),\left(\sqrt{\frac{U_2}{\beta_{12}}}-\sqrt{\frac{\beta_{12}}{U_1}}\right)\right\} \sim N_2\left((0,0),\begin{pmatrix}\alpha_{11}^2 & \rho_1\alpha_{11}\alpha_{12}\\ \rho_1\alpha_{11}\alpha_{12} & \alpha_{12}^2\end{pmatrix}\right). \quad (30)$$

Moreover, the following result will be useful.

Result: Let $\{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\}$ be a random sample from a bivariate normal distribution with mean vector $(0,0)$, and the covariance matrix $\boldsymbol{\Sigma}$. Further let us denote $f_{\boldsymbol{Z}}(\cdot;\boldsymbol{\Sigma})$, as the

PDF of a bivariate normal distribution with mean vector $(0,0)$, and the covariance matrix $\mathbf{\Sigma}$. For non-negative constants $c_1, \ldots, c_n$, if we denote

$$\widetilde{\mathbf{\Sigma}} = \arg\max_{\mathbf{\Sigma}} \sum_{i=1}^{n} c_i \ln f_{\boldsymbol{Z}}(\boldsymbol{z}_i; \mathbf{\Sigma})$$

then

$$\widetilde{\mathbf{\Sigma}} = \frac{\sum_{i=1}^{n} c_i \boldsymbol{z}_i \boldsymbol{z}_i^T}{\sum_{i=1}^{n} c_i}.$$

Proof: The proof of the above result can be obtained by taking the derivative with respect to the different elements of $\mathbf{\Sigma}$ and equating them to 0, see for example Rao ([11], pp 529 - 531). ∎

Therefore, for a given β_{11} and β_{12}, if

$$(\widetilde{\widetilde{\alpha}}_{11}(\beta_{11},\beta_{12}), \widetilde{\widetilde{\alpha}}_{12}(\beta_{11},\beta_{12}), \widetilde{\widetilde{\rho}}_1(\beta_{11},\beta_{12})) = \arg\max_{\boldsymbol{\alpha}_1,\rho_1} \sum_{i=1}^{n} \widetilde{a}_i \ln\left(f_{U_1,U_2}(x_{i1},x_{i2};\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1,\rho_1)\right),$$

then

$$\begin{aligned}
\widetilde{\widetilde{\alpha}}_{11}(\beta_{11},\beta_{12}) &= \left(\frac{1}{\sum_{i=1}^{n}\widetilde{a}_i}\sum_{i=1}^{n}\widetilde{a}_i\left(\sqrt{\frac{x_{i1}}{\beta_{11}}}-\sqrt{\frac{\beta_{11}}{x_{i1}}}\right)^2\right)^{1/2} \\
\widetilde{\widetilde{\alpha}}_{12}(\beta_{11},\beta_{12}) &= \left(\frac{1}{\sum_{i=1}^{n}\widetilde{a}_i}\sum_{i=1}^{n}\widetilde{a}_i\left(\sqrt{\frac{x_{i2}}{\beta_{12}}}-\sqrt{\frac{\beta_{12}}{x_{i2}}}\right)^2\right)^{1/2} \\
\widetilde{\widetilde{\rho}}_1(\beta_{11},\beta_{12}) &= \frac{1}{\widetilde{\widetilde{\alpha}}_{11}(\beta_{11},\beta_{12}) \times \widetilde{\widetilde{\alpha}}_{12}(\beta_{11},\beta_{12})} \times \\
&\quad \left(\frac{1}{\sum_{i=1}^{n}\widetilde{a}_i}\sum_{i=1}^{n}\widetilde{a}_i\left(\sqrt{\frac{x_{i1}}{\beta_{11}}}-\sqrt{\frac{\beta_{11}}{x_{i1}}}\right)\left(\sqrt{\frac{x_{i2}}{\beta_{12}}}-\sqrt{\frac{\beta_{12}}{x_{i2}}}\right)\right).
\end{aligned}$$

Finally $\widetilde{\widetilde{\beta}}_{11}$ and $\widetilde{\widetilde{\beta}}_{12}$ can be obtained as

$$(\widetilde{\widetilde{\beta}}_{11}, \widetilde{\widetilde{\beta}}_{12}) = \arg\max_{\beta_{11},\beta_{12}} \sum_{i=1}^{n} \widetilde{a}_i \ln\left(f_{U_1,U_2}(x_{i1},x_{i2};\widetilde{\widetilde{\boldsymbol{\alpha}}}_1(\boldsymbol{\beta}_1),\boldsymbol{\beta}_1,\widetilde{\widetilde{\rho}}_1(\boldsymbol{\beta}_1)\right),$$

and this can be performed by solving a two-dimensional optimization problem. Then

$$\widetilde{\widetilde{\boldsymbol{\alpha}}}_1 = \widetilde{\widetilde{\boldsymbol{\alpha}}}_1((\widetilde{\widetilde{\boldsymbol{\beta}}}_1) \quad \text{and} \quad \widetilde{\widetilde{\rho}}_1 = \widetilde{\widetilde{\rho}}_1((\widetilde{\widetilde{\boldsymbol{\beta}}}_1).$$

Similarly, $(\widetilde{\widetilde{\boldsymbol{\alpha}}}_2, \widetilde{\widetilde{\boldsymbol{\beta}}}_2, \widetilde{\widetilde{\rho}}_2)$ can also be obtained along the same line as follows. First obtain $\widetilde{\widetilde{\beta}}_{21}$ and $\widetilde{\widetilde{\beta}}_{22}$, as

$$(\widetilde{\widetilde{\beta}}_{21}, \widetilde{\widetilde{\beta}}_{22}) = \arg\max_{\beta_{21},\beta_{22}} \sum_{i=1}^{n} \widetilde{b}_i \ln \left(f_{V_1,V_2}(x_{i1}, x_{i2}; \widetilde{\widetilde{\boldsymbol{\alpha}}}_2(\boldsymbol{\beta}_2), \boldsymbol{\beta}_2, \widetilde{\widetilde{\rho}}_2(\boldsymbol{\beta}_2) \right),$$

where

$$\begin{aligned}
\widetilde{\widetilde{\alpha}}_{21}(\beta_{21}, \beta_{22}) &= \left(\frac{1}{\sum_{j=1}^{n} \widetilde{b}_j} \sum_{i=1}^{n} \widetilde{b}_i \left(\sqrt{\frac{x_{i1}}{\beta_{21}}} - \sqrt{\frac{\beta_{21}}{x_{i1}}} \right)^2 \right)^{1/2} \\
\widetilde{\widetilde{\alpha}}_{22}(\beta_{21}, \beta_{22}) &= \left(\frac{1}{\sum_{j=1}^{n} \widetilde{b}_j} \sum_{i=1}^{n} \widetilde{b}_i \left(\sqrt{\frac{x_{i2}}{\beta_{22}}} - \sqrt{\frac{\beta_{22}}{x_{i2}}} \right)^2 \right)^{1/2} \\
\widetilde{\widetilde{\rho}}_2(\beta_{21}, \beta_{22}) &= \frac{1}{\widetilde{\widetilde{\alpha}}_{21}(\beta_{21}, \beta_{22}) \times \widetilde{\widetilde{\alpha}}_{22}(\beta_{21}, \beta_{22})} \times \\
&\quad \left(\frac{1}{\sum_{j=1}^{n} \widetilde{b}_j} \sum_{i=1}^{n} \widetilde{b}_i \left(\sqrt{\frac{x_{i1}}{\beta_{21}}} - \sqrt{\frac{\beta_{21}}{x_{i1}}} \right) \left(\sqrt{\frac{x_{i2}}{\beta_{22}}} - \sqrt{\frac{\beta_{22}}{x_{i2}}} \right) \right).
\end{aligned}$$

Finally obtain

$$\widetilde{\widetilde{\boldsymbol{\alpha}}}_2 = \widetilde{\widetilde{\boldsymbol{\alpha}}}_2((\widetilde{\widetilde{\boldsymbol{\beta}}}_2) \quad \text{and} \quad \widetilde{\widetilde{\rho}}_2 = \widetilde{\widetilde{\rho}}_2((\widetilde{\widetilde{\boldsymbol{\beta}}}_2).$$

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

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

[6] Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathemaics, Philadelphia.

[7] Johnson, R.A. and Wichern, D.W. (1992), *Applied multivariate statistical analysis*, 3-rd edition, Prentice Hall, New Jersey.

[8] Kocherlakota, S. (1986), "The bivariate inverse Gaussian distribution: an introduction", *Communications in Statistics - Theory and Methods*, vol. 15, 1081 - 1112.

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

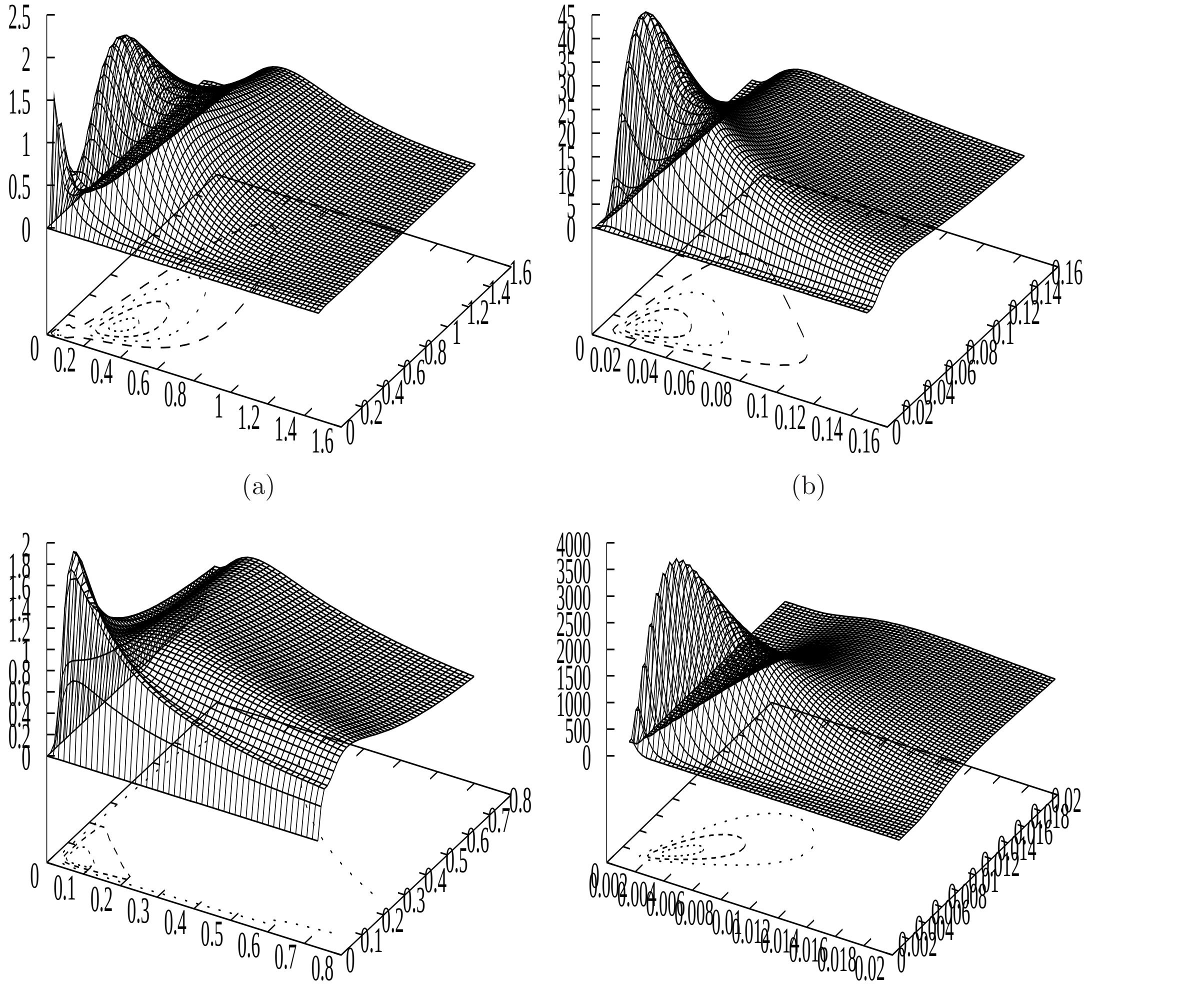
[10] McLachlan, G.J. and Peel, D. (2000), *Finite mixture models*, Wiley, New York

[11] Rao, C.R. (1973), *Linear Statistical Inference and its Applications*, 2nd edition, Wiley Eastern, India.

[12] Shuster, J.J. (1968), "On the inverse Gaussian distribution", *Journal of the American Statistical Association*, vol. 63, 1514 - 1516.

[13] Tricomi, F. G. (1947), "Sulle funzioni ipergeometriche confluenti", (in Italian). *Annali di Matematica Pura ed Applicata*, Serie Quarta 26: 141 - 175.

[14] Wasan, M.L. (1968), "On an inverse Gaussian process", *Scandinavian Acturial Journal*, vol. 51, 69 - 96.



(a) (b)

(c) (d)

Figure 1: Contour plots of the of the joint PDF of the MBBS distribution for different parameter values: $(\alpha_{11}, \alpha_{12}, \beta_{11}, \beta_{12}, \rho_1, \alpha_{21}, \alpha_{22}, \beta_{21}, \beta_{22}, \rho_2, p)$: (a)(2.5,2.5,1.0,1.0,0.5,0.5,1.0,0.5,1.0,0.5,0.1) (b)(2.5,2.5,0.5,0.5,0.5,2.0,2.0,0.5,0.5,0.5,0.5) (c)(2.5,2.5,0.5,0.5,-0.5,1.0,1.0,0.5,0.5,-0.5,0.5) (d)(2.5,2.5,0.25,0.25,0.9,1.0,1.0,1.5,1.5,-0.9,0.5)

.

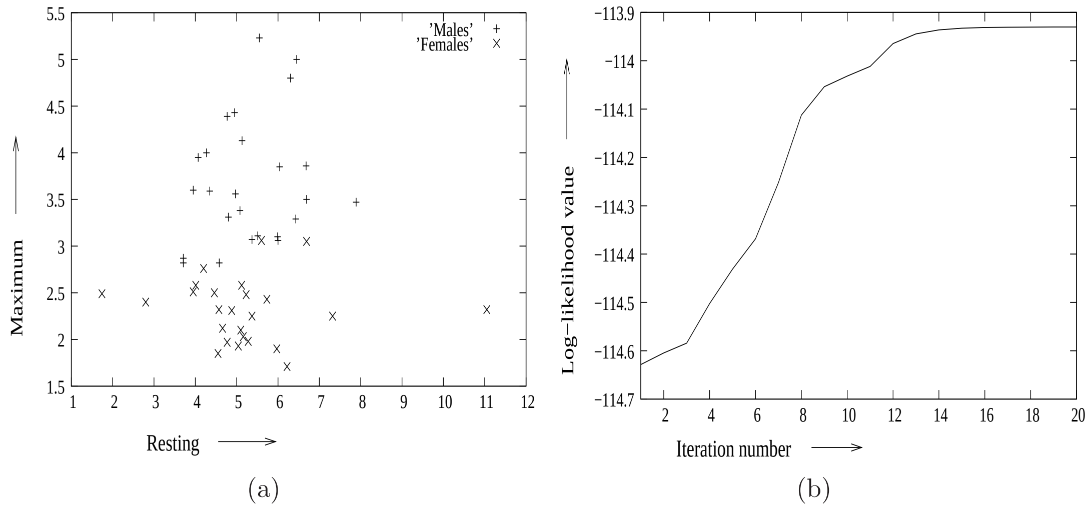


Figure 2: (a) Scatter plot of the real data, (b) log-likelihood value at each iteration of the real data

.