Multivariate log-Birnbaum-Saunders regression models

Guillermo Martínez-Flórez*
Departamento de Matemáticas y Estadística, Universidad de Córdoba, Montería, Colombia - gmartinez@correo.unicordoba.edu.co

Germán Moreno-Arenas
Escuela de Matemáticas, Universidad Industrial de Santander, Bucaramanga, Colombia - gmorenoa@uis.edu.co

Rafael Bráz Azevedo Farias
Depart. de Estatística e Matemática Aplicada, Universidade Federal do Ceará, Ceará, Brazil - rafael@dema.ufc.br

Abstract

In this paper we present a multivariate version of the skewed log-Birnbaum-Saunders regression model. This new family of distributions holds good properties, such as the marginal distributions of the dependent variables are univariate skewed log-Birnbaum-Saunders distribution and have the usual log-Birnbaum-Saunders distribution as a particular case. Furthermore, the model parameters are estimated by using maximum-likelihood methods and a closed-form expression for the Fisher's information matrix is presented, which aid us to perform testing hypothesis for model parameters by using approximations obtained from the asymptotic normality of maximum-likelihood estimators. Two real data set are analysed and the results are discussed, illustrating the usefulness of the extension considered.

Keywords: log-linear Birnbaum-Saunders model, sinh-normal distribution, univariate log-Birnbaum-Saunders distribution

1 Introduction

The univariate family of distributions proposed by Birnbaum and Saunders (1969), also known as the fatigue life distributions, has been widely applied for describing fatigue lifetimes. This family was originally derived from a model for which failure follows from the development and growth of a dominant crack. The random variable T is said to be Birnbaum-Saunders distribution (BS) if T is of the form

$$T = \frac{\eta}{4}\left[\alpha Z + \sqrt{\alpha^2 Z^2 + 4}\right]^2,$$

where $Z \sim \mathrm{N}(0,1)$ is a standard normal distribution, $\alpha > 0$ is the shape parameter and $\eta > 0$ is the scale parameter and the median of the distribution. The density function of a BS random variable T with parameters α and η is given by

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

where $\alpha > 0$ is the parameter that controls the skewness. We denote this distribution by $T \sim \mathrm{BS}(\alpha, \eta)$. It is well known that the BS distribution becomes asymmetric as α increases and symmetric around η as α gets close to zero. Some other properties include $kT \sim \mathrm{BS}(\alpha, k\eta)$, for $k > 0$, and $T^{-1} \sim \mathrm{BS}(\alpha, \eta^{-1})$. For more detail, see Birnbaum and Saunders (1969) and Díaz-García and Leiva (2005), among others.

Asymmetric extensions of Birnbaum-Saunders model were studied by Vilca-Labra and Leiva-Sánchez (2006), who introduced the generalized BS distribution for elliptical distributions. Subsequently Castillo et al. (2011) introduced the epsilon Birnbaum-Saunders distribution and Gomez et al. (2009) study an extension of the generalized BS with slash-elliptical distributions. A bivariate BS distribution which is absolutely continuous and has five parameters is proposed by Kundu et al. (2010), this extension is based on transformations of a random vector with bivariate normal distribution. Lemonte et al. (2015) introduced a skewed multivariate Birnbaum-Saunders (SMVBS) distribution. This multivariate extension is based on the skew-normal distribution of Arnold et al. (2002). The probability density function of the SMVBS distribution defined in Lemonte et al. (2015) is given by

$$f_{T_1,\ldots,T_p}(t_1,\ldots,t_p) = 2\prod_{j=1}^{p}\phi(a_{t_j})\Phi\left(\lambda\prod_{j=1}^{p}a_{t_j}\right)\prod_{j=1}^{p}\frac{t_j^{-3/2}(t_j+\eta_j)}{2\alpha_j\sqrt{\eta_j}}, \quad t_1,\ldots,t_j>0, \tag{2}$$

where $\alpha_j > 0$ and $\eta_j > 0$ are shape and scale parameters, respectively, and

$$a_{t_j} = a_{t_j}(\alpha_j,\eta_j) = \frac{1}{\alpha_j}\left(\sqrt{\frac{t}{\eta_j}} - \sqrt{\frac{\eta_j}{t}}\right), \tag{3}$$

for $j = 1,2,\ldots,p$, while $\lambda \in \mathbb{R}$ is a skewness parameter. This distribution is denoted by SMVBS$(\boldsymbol{\alpha},\boldsymbol{\eta},\lambda)$.

Lemonte et al. (2015) show that the marginals of the SMVBS distribution are univariate Birnbaum-Saunders distributions with parameters α_j and η_j. The estimation by the method of modified moments and maximum likelihood are also discussed in this paper.

Besides, Rieck and Nedelman (1991) studied the sinh-normal distribution which is obtained as a transformation of the standard normal distribution after considering the random variable

$$Y = \operatorname{arcsinh}(\alpha Z/2)\,\sigma + \gamma, \tag{4}$$

where $Z \sim \mathrm{N}(0,1)$, with $\alpha > 0$, $\gamma \in \mathbb{R}$ and $\sigma > 0$, denoting shape, location and scale parameters, respectively. We denote by $Y \sim \mathrm{SHN}(\alpha,\gamma,\sigma)$. This distribution is also known as log-Birnbaum Saunders distribution, since that if $Y \sim \mathrm{SHN}(\alpha,\gamma,\sigma=2)$ then $T = \exp(Y)$ follow a Birnbaum Saunders distribution, which we denote by $T \sim \mathrm{BS}(\alpha,\beta)$ where $\alpha > 0$ is a shape parameter and $\beta = \exp(\gamma)$ is a scale parameter. For more detail, see, for example Birnbaum and Saunders (1969) and Díaz-García and Leiva (2005).
An asymmetrical extension of the SHN distribution is defined by Leiva et al. (2010) and it is obtained considering that the variable Z presented in (4) follows the skew-normal (SN) distribution defined in Azzalini (1985). The resulting density function is

$$f(z) = 2b'_z\phi(b_z)\Phi(\lambda b_z), \tag{5}$$

where $b_z = \frac{2}{\alpha}\sinh\left(\frac{z-\gamma}{\sigma}\right)$ and $b'_z = \frac{2}{\alpha\sigma}\cosh\left(\frac{z-\gamma}{\sigma}\right)$ is the derivative of b_z with respect to the variable z. This distribution is called log-skew-Birnbaum-Saunders distribution.
Díaz-García and Domínguez-Molina (2006) introduced multivariate versions of the SHN distributions. They define the random vector $\boldsymbol{Y}$ with multivariate sinh-normal distribution, which is denoted by $\boldsymbol{Y} \sim \mathrm{SHN}_q(\alpha\mathbf{1}_q,\boldsymbol{\mu},\boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = 2\boldsymbol{I}_q$ is the identity matrix $q \times q$, as

$$f(\boldsymbol{y}) = (2\pi\alpha^2)^{-q/2}\left(\prod_{j=1}^{q}\cosh\left(\frac{y_j-\mu_j}{2}\right)\right)\exp\left\{-\frac{2}{\alpha^2}\sum_{j=1}^{q}\sinh^2\left(\frac{y_j-\mu_j}{2}\right)\right\}, \quad \boldsymbol{y} \in \mathbb{R}^q.$$

On the other hand, Lemonte (2013) introduced the multivariate BS regression model that generalizes the univariate regression model introduced by Rieck and Nedelman (1991). In this model is supposed to have n multivariate random variables $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_n$ and in each one a number of responses q is measured. Thus, it is

obtained the following multivariate log-linear regression model $\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i$, with $\boldsymbol{\epsilon}_i \sim \text{SHN}_q(\alpha\mathbf{1}_q, \mathbf{0}_q, 2\boldsymbol{I}_q)$, for $i = 1, 2, \ldots, n$, where $\mathbf{1}_q$ and $\mathbf{0}_q$ are vectors of ones and zeros, respectively, of dimension q.
In other cases, the multivariate distributions can be built based on the theory of copulas with known marginal distributions, see Joe (1997) and Nelsen (1999), among others. Kundu et al. (2010) used the Clayton's copula to propose a bivariate extension of the power-normal distribution. Kundu (2014) proposes the bivariate sinh-normal distribution based on the bivariate normal copula, specifically, the bivariate random variable (X_1, X_2) is said to have a bivariate sinh-normal distribution, which is denoted by $\text{CBSHN}(\alpha_1, \alpha_2, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, for $\alpha_1, \sigma_1, \alpha_2, \sigma_2 > 0$, $\mu_1, \mu_2 \in \mathbb{R}$ and $-1 < \rho < 1$. The joint cumulative distribution function (CDF) of (X_1, X_2) has the form

$$F_{X_1,X_2}(x_1, x_2) = \Phi_2\left(b(x_1, \alpha_1, \mu_1, \sigma_1), b(x_2, \alpha_2, \mu_2, \sigma_2), \rho\right), \tag{6}$$

where $\Phi_2(\cdot)$ denotes the CDF of the bivariate normal distribution and ρ is the parameter that controls the dependence in the bivariate normal copula.
As can be perceived, few studies on multivariate versions of the BS, sinh-normal distributions and the log-linear-BS regression models have been carried out. Further we highlight the work of Lemonte (2013) in the asymmetric field, where the skewed multivariate BS model is studied. In this paper, we examine an extension to asymmetric multivariate log-BS model and study the asymmetric multivariate log-BS regression model.
The paper is outlined as follows. In Section 2 we define the skewed multivariate sinh-normal distribution and it is shown some general results about the shape of the density function. In Section 3 we introduce the skewed multivariate sinh-normal regression model and we discuss maximum likelihood estimation in Section 4. The potentiality of the new model is illustrated by means of an application in two real data sets in Section 5.

2 Skewed Multivariate Sinh-Normal Distribution

For the case in which $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_p)$ is p-dimensional, the following notation is convenient for describing such an extension. For each $j = 1, 2, \ldots, p$, define the vector $\boldsymbol{Z}_{(j)}$ to be the $(p-1)-$dimensional random vector obtained from $\boldsymbol{Z}$ by deleting Z_j. In parallel fashion, for a real vector $\boldsymbol{z} = (z_1, z_2, \ldots, z_p)$, $\boldsymbol{z}_{(j)}$ is obtained by deleting the j-th coordinate z_j of $\boldsymbol{z}$.
Supported on the results of Arnold et al. (2002), suppose that, for each $j = 1, 2, \ldots, p$, the conditional distribution of Z_j given $\boldsymbol{Z}_{(j)} = \boldsymbol{z}_{(j)}$ is a skew-normal distribution with a parameter which is a function of $\boldsymbol{z}_{(j)}$. We assume, for each j, that

$$Z_j|\boldsymbol{Z}_{(j)} = \boldsymbol{z}_{(j)} \sim \text{SN}\left(\lambda \prod_{j' \neq j} z_{j'}\right). \tag{7}$$

Thus the joint density function of $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_p)$ is given by

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2\left(\prod_{j=1}^{p} \phi(z_j)\right)\Phi\left(\lambda \prod_{j=1}^{p} z_j\right),$$

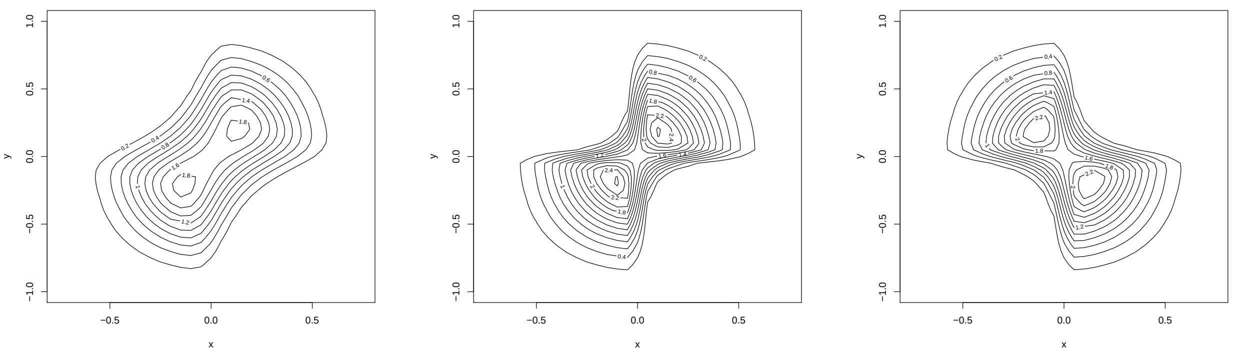
and under the transformation

$$Y_j = \sigma_j \text{arcsinh}\left(\frac{\alpha_j Z_j}{2}\right) + \gamma_j, \quad Z_j \sim \text{N}(0, 1) \tag{8}$$

is obtained the skewed multivariate sinh-normal distribution with density function:

$$f_{Y_1,\ldots,Y_p}(y_1, \ldots, y_p) = 2\left(\prod_{j=1}^{p} b'_j\right)\left(\prod_{j=1}^{p} \phi(b_j)\right)\Phi\left(\lambda \prod_{j=1}^{p} b_j\right), \; y_1, \ldots, y_p \in \mathbb{R}, \tag{9}$$

where $b_j = \frac{2}{\alpha_j}\sinh\left(\frac{Y_j - \gamma_j}{\sigma_j}\right)$, $b'_j = \frac{2}{\alpha_j\sigma_j}\cosh\left(\frac{Y_j - \gamma_j}{\sigma_j}\right)$ is the derivative of b_j, $\alpha_j > 0$ and $\sigma_j > 0$ are shape and scale parameters, respectively, γ_j and $\lambda \in \mathbb{R}$ are location and skewness parameters, respectively. We denote by SMVSHN$(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \lambda)$, with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_p)'$ and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_p)'$. It is worth to note that the random variables $Y_1, Y_2, \ldots, Y_n$ become independent for $\lambda = 0$ in (9) and hence the proposed bivariate model reduces to the bivariate model considered by Díaz-García and Domínguez-Molina (2006). Some contour plots of the joint probability density function (9) are presented in Figure (1) for certain values of the parameters. From these figures we observe that the joint probability density function (9) can take different forms and will therefore be useful in analyzing bivariate survival data.



(a) $\lambda = 2$ (b) $\lambda = 7.5$ (c) $\lambda = -5$

Figure 1: Contour plots SBVSHN for $\alpha_1 = 0.5$, $\alpha_2 = 0.75$, $\gamma_1 = \gamma_2 = 0$ and $\sigma_1 = \sigma_2 = 1$.

Theorem 2.1 provides the marginal and conditional distributions of the SMVSHN distribution.

Theorem 2.1. *If* $(Y_1, Y_2, \ldots, Y_n) \sim$ SMVSHN$(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\sigma}, \lambda)$ *then*

1. $Y_j \sim$ SHN$(\alpha_j, \gamma_j, \sigma_j)$ *for* $j = 1, 2, \ldots, n$.
2. *The conditional density function of* $Y_j|\boldsymbol{Y}_{(j)} = \boldsymbol{y}_{(j)}$ *is given by*

$$f_{Y_j|\boldsymbol{Y}_{(j)}}(y_j|\boldsymbol{Y}_{(j)} = \boldsymbol{y}_{(j)}) = 2b_j'\phi(b_j)\Phi\left(\lambda\prod_{j=1}^{n} b_j\right). \tag{10}$$

3. *The cumulative distribution function of* $Y_j|\boldsymbol{Y}_{(j)} = \boldsymbol{y}_{(j)}$ *is given by*

$$P(Y_j \leq y_j|\boldsymbol{Y}_{(j)} = \boldsymbol{y}_{(j)}) = \Phi(b_j) - 2T\left(b_j, \lambda\prod_{j' \neq j} b_{j'}\right), \tag{11}$$

where $T(\cdot)$ *is the Owen function.*

3 Skewed Multivariate Sinh-Normal Regression Model

We now introduce the asymmetric version of the model studied by Lemonte (2013). We consider the situation where n multivariate random variables $y_1, y_2, \ldots, y_n$ are collected and the number of responses measured at each observation is q, with $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iq})'$ a vector $q \times 1$ of observed dependent variables, where y_{ij} is the log-survival time corresponding to the j-th experimental unit. Suppose also that there are p-variables explanatory $X_1, X_2, \ldots, X_p$ where $\boldsymbol{X}_i = (x_{i1}, x_{i2}, \ldots, x_{iq})$, is a $q \times p$ matrix model associated to the ith observable response $\boldsymbol{y}_i$, with $\boldsymbol{x}_{ij} = (x_{ij1}, x_{ij2}, \ldots, x_{ijp})'$ for $j = 1, 2, \ldots, q$ a p-dimensional vector with values of the explanatory variables. Such that the matrices of observed values are $\boldsymbol{y} = \text{vec}(y_1, y_2, \ldots, y_n)$, with

dimension $N \times 1$, $\boldsymbol{X} = (\boldsymbol{X}_1', \ldots, \boldsymbol{X}_n')'$ of dimension $N \times p$, with $N = nq$ and $\mathrm{vec}(\cdot)$ the vec operator which transforms a matrix into a column vector.
Thus the skewed multivariate Birnbaum-Saunders log-linear regression model is defined by

$$\boldsymbol{y} = \boldsymbol{X}'\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{12}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is a $p \times 1$ vector of unknown parameters, and $\boldsymbol{\epsilon}$ is a $N \times 1$ vector of random errors such that $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n)' \sim \mathrm{SMVSHN}(\alpha \mathbf{1}_N, \mathbf{0}_N, 2\boldsymbol{I}_N, \lambda)$, with $\mathbf{0}_N$ denoting a $N \times 1$ vector of zeros. From Theorem 2.1 follows

$$\boldsymbol{y}_i = \boldsymbol{X}_i'\boldsymbol{\beta} + \boldsymbol{\epsilon}_i, \quad i = 1, 2, \ldots, n, \tag{13}$$

where ϵ_i's are independent and identically distributed random variables such that $\boldsymbol{\epsilon}_i \sim \mathrm{SHN}_q(\alpha \mathbf{1}_q, \mathbf{0}_q, 2\boldsymbol{I}_q)$, with $\mathbf{0}_q$ denoting a $q \times 1$ vector of zeros.
Therefore, it follows that $\boldsymbol{y}_i \sim \mathrm{SHN}_q(\alpha \mathbf{1}_q, \boldsymbol{X}_i'\boldsymbol{\beta}, 2\boldsymbol{I}_q)$, that is, each marginal is a multivariate log-BS distribution with dimension q. For $p = 2$, $\mathbb{E}(Y_1) = x_{1j}'\boldsymbol{\beta}$, $\mathbb{E}(Y_2) = x_{2j}'\boldsymbol{\beta}$ and $\mathbb{E}(Y_1 Y_2) = x_{1j}'\boldsymbol{\beta} x_{2j}'\boldsymbol{\beta} + c(\alpha, \lambda)$, where

$$c(\alpha, \lambda) = 8 \int_{-\infty}^{\infty} \sinh^{-1}(\alpha x_1/2) \sinh^{-1}(\alpha x_2/2) \phi(x_1) \phi(x_2) \Phi(\lambda x_1 x_2) dx_1 dx_2.$$

Then the correlation coefficient between Y_1 and Y_2 is $\rho(Y_1, Y_2) = \frac{c(\alpha,\lambda)}{4Q(\alpha)}$, where $Q(\alpha)$ is the variance of the random variable $\mathrm{arcsinh}(\alpha Z_j/2)$, which must be calculated by numerical methods.

4 Inference

Initially for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, q$ and $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}', \lambda)$ we define the following expressions, $\boldsymbol{\xi}_1 = \boldsymbol{\xi}_1(\boldsymbol{\theta}) = \mathrm{vec}(\boldsymbol{\xi}_{11}, \boldsymbol{\xi}_{21}, \ldots, \boldsymbol{\xi}_{n1})$, $\boldsymbol{\xi}_2 = \boldsymbol{\xi}_2(\boldsymbol{\theta}) = \mathrm{vec}(\boldsymbol{\xi}_{12}, \boldsymbol{\xi}_{22}, \ldots, \boldsymbol{\xi}_{n2})$ and $\boldsymbol{\xi}_3 = \boldsymbol{\xi}_3(\boldsymbol{\theta}) = \mathrm{vec}(\boldsymbol{\xi}_{13}, \boldsymbol{\xi}_{23}, \ldots, \boldsymbol{\xi}_{n3})$ where

$$\boldsymbol{\xi}_{i1} = (\xi_{i11}, \xi_{i12}, \ldots, \xi_{i1q})' = \frac{2}{\alpha} \cosh\left(\frac{\boldsymbol{y}_i - \boldsymbol{x}_{ij}'\boldsymbol{\beta}}{2} \right),$$

$$\boldsymbol{\xi}_{i2} = (\xi_{i21}, \xi_{i22}, \ldots, \xi_{i2q})' = \frac{2}{\alpha} \sinh\left(\frac{\boldsymbol{y}_i - \boldsymbol{x}_{ij}'\boldsymbol{\beta}}{2} \right)$$

and

$$\boldsymbol{\xi}_{i3} = (\xi_{i31}, \xi_{i32}, \ldots, \xi_{i3q})' = \Phi\left\{ \lambda \frac{2^p}{\alpha^p} \prod_{i=1}^{n} \sinh\left(\frac{\boldsymbol{y}_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{2} \right) \right\} = \Phi\left(\lambda \prod_{i=1}^{n} \boldsymbol{\xi}_{i2} \right),$$

with $\xi_{i1j} = \xi_{i1j}(\boldsymbol{\theta}) = \frac{2}{\alpha} \cosh\left(\frac{y_{ij} - \boldsymbol{x}_{ij}'\boldsymbol{\beta}}{2} \right)$, $\xi_{i2j} = \xi_{i2j}(\boldsymbol{\theta}) = \frac{2}{\alpha} \sinh\left(\frac{y_{ij} - \boldsymbol{x}_{ij}'\boldsymbol{\beta}}{2} \right)$ and $\xi_{i3j} = \xi_{i3j}(\boldsymbol{\theta}) = \Phi\left\{ \lambda \frac{2^p}{\alpha^p} \prod_{i=1}^{n} \sinh\left(\frac{y_{ij} - \boldsymbol{x}_{ij}'\boldsymbol{\beta}}{2} \right) \right\} =$ $\Phi\left(\lambda \prod_{i=1}^{n} \xi_{i2j} \right)$.
The log-likelihood function for the vector $\boldsymbol{\theta}$ can be written as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}) = -\frac{N}{2} \log(8\pi) + \mathbf{1}_N' \log(\boldsymbol{\xi}_1) - \frac{\boldsymbol{\xi}_2'\boldsymbol{\xi}_2}{2} + \mathbf{1}_N' \log(\boldsymbol{\xi}_3), \tag{14}$$

where $\mathbf{1}_N$ denotes a $N \times 1$ vector of ones. The log-likelihood function for the i-th random variable $\boldsymbol{y}_i$ is

$$\ell_i(\boldsymbol{\theta}) = -\frac{q}{2} \log(8\pi) + \mathbf{1}_q' \log(\boldsymbol{\xi}_{i1}) - \frac{\boldsymbol{\xi}_{i2}'\boldsymbol{\xi}_{i2}}{2} + \mathbf{1}_q' \log(\boldsymbol{\xi}_{i3}) \tag{15}$$

and the score function obtained by differentiating the log-likelihood function with respect to the unknown parameters is:

$$U_\alpha = -\frac{N}{\alpha} + \frac{\boldsymbol{\xi}_2'\boldsymbol{\xi}_2}{\alpha} - \frac{n\lambda}{\alpha} \boldsymbol{w}'\boldsymbol{p}, \qquad U_{\boldsymbol{\theta}} = \boldsymbol{X}'\boldsymbol{s} - \frac{\lambda}{2} (\boldsymbol{w}'\boldsymbol{Z}M)', \qquad U_\lambda = \boldsymbol{w}'\boldsymbol{p}, \tag{16}$$

where $\boldsymbol{w} = \boldsymbol{w}(\boldsymbol{\theta}) = \mathrm{vec}(\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n)$ with $\boldsymbol{w}_i = \boldsymbol{w}_i(\boldsymbol{\theta}) = (w_{i1}, w_{i2}, \ldots, w_{iq})$ where $w_{ij} = w_{ij}(\boldsymbol{\theta}) = \frac{\phi(\lambda \prod_{k=1}^{n} \xi_{k2j})}{\Phi(\lambda \prod_{k=1}^{n} \xi_{k2j})}$, for each $i = 1, 2, \ldots, p$; $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{\theta}) = \mathrm{vec}(\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n)$, with $\boldsymbol{p}_i = \boldsymbol{p}_i(\boldsymbol{\theta}) = (p_{i1}, p_{i2}, \ldots, p_{iq})$, where $p_{ij} = p_{ij}(\boldsymbol{\theta}) = \prod_{k=1}^{n} \xi_{k2j}$, for each $i = 1, 2, \ldots, p$; $\boldsymbol{Z} = \text{block-diag}\{\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_n\}$, where $\boldsymbol{Z}_i = \mathrm{diag}\{p_{i1}, p_{i2}, \ldots, p_{iq}\}$, $\boldsymbol{M} = (\boldsymbol{M}_1, \boldsymbol{M}_2, \ldots, \boldsymbol{M}_n)'$ of size $N \times p$ where $\boldsymbol{M}_i = (m_{i1}, m_{i2}, \ldots, m_{iq})$ with $m_{ij} = (m_{ij1}, m_{ij2}, \ldots, m_{ijp})$ where $m_{ijr} = \sum_{k=1}^{n} x_{kjr} \coth\left(\frac{y_{kj} - x_{kj}\boldsymbol{\beta}}{2}\right)$ and $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\theta}) = \mathrm{vec}(\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n)$, with $\boldsymbol{s}_i = \boldsymbol{s}_i(\boldsymbol{\theta}) = (s_{i1}, s_{i2}, \ldots, s_{iq})$ where $s_{ij} = s_{ij}(\boldsymbol{\theta}) = (\xi_{i1j}\xi_{i2j} - \xi_{i2j}/\xi_{i1j})/2$.
The maximum likelihood estimators (MLEs) of the parameters are the solutions of the equations $U_\alpha = 0$, $U_{\boldsymbol{\theta}} = \boldsymbol{0}$, $U_\lambda = 0$, which can be solved by using iterative numerical methods such as Newton Raphson. From the above system we have that

$$\hat{\alpha} = \sqrt{\frac{4}{N} \sum_{i=1}^{n} \sum_{j=1}^{q} \sinh^2\left(\frac{y_{ij} - \boldsymbol{x}_{ij}\hat{\boldsymbol{\beta}}}{2}\right)}. \tag{17}$$

We suggest taking as a starting point for the parameters $\boldsymbol{\beta}$ the estimates obtained by the method of ordinary least squares (OLS) $\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}'X)^{-1}\boldsymbol{X}'y$, while for α can be taken by Equation (17). Finally, the starting point for λ can be the solution of the following the nonlinear system $\sum_{i=1}^{n} \sum_{j=1}^{q} \tilde{w_{ij}} \tilde{p_{ij}} = 0$, where $\tilde{w_{ij}}$ and $\tilde{p_{ij}}$ are obtained by replacing w_{ij} and p_{ij} by α and $\boldsymbol{\theta}$ for $\tilde{\alpha}$ and $\tilde{\boldsymbol{\beta}}$, respectively. Replacing α for $\hat{\alpha}(\boldsymbol{\beta})$ we obtain the profile log-likelihood function (apart from an unimportant constant):

$$\ell_p(\boldsymbol{\beta}, \lambda) = \sum_{i=1}^{n} \ell_i(\hat{\alpha}(\boldsymbol{\beta}), \boldsymbol{\beta}, \lambda). \tag{18}$$

The MLE of $\boldsymbol{\beta}, \lambda$ can be obtained from the profile log-likelihood function by maximizing $\ell_p(\boldsymbol{\beta}, \lambda)$ with regard to $\boldsymbol{\beta}, \lambda$. These solutions can not be obtained explicitly and optimization algorithms such as Newton-Raphson type should be applied in order to obtain the estimates. The profile log-likelihood function $\ell_p(\boldsymbol{\beta}, \lambda)$ is not a real log-likelihood function and some of the properties that hold for a genuine log-likelihood do not hold for its profiled version. In particular, there exist score and information biases, both of order $O(1)$.

4.1 Information Matrices

In this section we present and discuss some properties of the information matrices of the multivariate log-Birnbaum-Saunders regression models. The Hessian matrix is given by

$$\ddot{L}_{\boldsymbol{\theta\theta}} = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{pmatrix} \ddot{L}_{\alpha\alpha} & \ddot{L}_{\alpha\boldsymbol{\beta}} & \ddot{L}_{\alpha\lambda} \\ \ddot{L}_{\alpha\boldsymbol{\beta}}^T & \ddot{L}_{\boldsymbol{\beta\beta}} & \ddot{L}_{\boldsymbol{\beta}\lambda} \\ \ddot{L}_{\alpha\lambda}^T & \ddot{L}_{\boldsymbol{\beta}\lambda}^T & \ddot{L}_{\lambda\lambda} \end{pmatrix},$$

with

$$\ddot{L}_{\alpha'\alpha} = \frac{N}{\alpha^2} - \frac{3}{\alpha^2}\boldsymbol{\xi}_2'\boldsymbol{\xi}_2 - \frac{n(n+1)\lambda}{\alpha^2} \sum_{i=1}^{n} \sum_{j=1}^{q} w_{ij} p_{ij} - \frac{n^2\lambda^2}{\alpha^2} \sum_{i=1}^{n} \sum_{j=1}^{q} [\lambda w_{ij} p_{ij}^3 - w_{ij}^2 p_{ij}^2],$$

and $\ddot{L}_{\alpha\boldsymbol{\beta}} = (\ddot{L}_{\alpha'\beta_r}) = -\boldsymbol{X}h - \boldsymbol{B}$, where $h_{ij} = -\frac{\xi_{i1j}\xi_{i2j}}{2}$ and $\boldsymbol{B}$ is a matrix of dimension $p \times 1$ with elements $b_r = \frac{n\lambda}{2\alpha} \sum_{i=1}^{n} \sum_{j=1}^{q} [-w_{ij}p_{ij} - \lambda^2 w_{ij} p_{ij}^3 + w_{ij}^2 p_{ij}^2] m_{kjr}$, for $r = 1, 2, \ldots, p$. $\ddot{L}_{\lambda\boldsymbol{\beta}} = (\ddot{L}_{\lambda\beta_r}) = -\boldsymbol{D}$ where $\boldsymbol{D}$ is a $p \times 1$ matrix with elements $d_r = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{q} [w_{ij}p_{ij} - \lambda^2 w_{ij} p_{ij}^3 - \lambda w_{ij}^2 p_{ij}^2] m_{kjr}$, $\ddot{L}_{\boldsymbol{\theta\theta}'} = (\ddot{L}_{\beta_r'\beta_r}) = -\boldsymbol{X}'VX - \boldsymbol{G}$, with $\boldsymbol{V} = \text{block-diag}\{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n\}$ and $\boldsymbol{V}_i = \mathrm{diag}\{v_{i1}, \ldots, v_{iq}\}$ with $v_{ij} = -\frac{1}{4}\left(2\xi_{i2j}^2 + \frac{4}{\alpha^2} - 1 + \frac{\xi_{i2j}^2}{\xi_{i1j}^2}\right)$ and $\boldsymbol{G}$ is a matrix of dimension $p \times p$ with elements $g_{rr'} = -\frac{\lambda^4}{4} w_{ij} p_{ij}^4 m_{ijr} m_{ijr'} - \frac{3\lambda^3}{4} w_{ij}^2 p_{ij}^3 m_{ijr} m_{ijr'} - \frac{\lambda^2}{4} w_{ij}^3 p_{ij}^2 m_{ijr} m_{ijr'} + \frac{\lambda^2}{2} w_{ij} p_{ij}^2 m_{ijr} m_{ijr'} + \frac{\lambda}{4} w_{ij}^2 p_{ij} m_{ijr} m_{ijr'} - \frac{\lambda^2}{4} w_{ij} p_{ij}^2 m_{ijrr'} - \frac{\lambda}{4} w_{ij}^2 p_{ij} m_{ijrr'}$, $\ddot{L}_{\alpha\lambda} = \frac{n}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{q} w_{ij} p_{ij} - \frac{n\lambda}{\alpha} \sum_{i=1}^{n} \sum_{j=1}^{q} [\lambda w_{ij} p_{ij}^3 + w_{ij}^2 p_{ij}^2]$, $\ddot{L}_{\lambda\lambda} = -\lambda \sum_{i=1}^{n} \sum_{j=1}^{q} w_{ij} p_{ij}^3 - \sum_{i=1}^{n} \sum_{j=1}^{q} w_{ij}^2 p_{ij}^2$, where $m_{ijrr'} = \sum_{k=1}^{n} x_{kjr} x_{kjr'} \mathrm{cosch}^2 \Big($

Thus the observed information matrix can be obtained by doing $I_o(\boldsymbol{\theta}) = -\ddot{L}_{\boldsymbol{\theta\theta}}$. Under certain regularity conditions, the Fisher information matrix may also be written as the expectation of the observed information matrix and expressed by

$$\Sigma_{\boldsymbol{\theta\theta}} = -\mathbb{E}\left[\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right] = \begin{pmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha\boldsymbol{\beta}} & \Sigma_{\alpha\lambda} \\ \Sigma_{\alpha\boldsymbol{\beta}}^T & \Sigma_{\boldsymbol{\beta\theta}} & \Sigma_{\boldsymbol{\beta}\lambda} \\ \Sigma_{\alpha\lambda}^T & \Sigma_{\boldsymbol{\beta}\lambda}^T & \Sigma_{\lambda\lambda} \end{pmatrix},$$

where $\Sigma_{\alpha\alpha} = \frac{2N}{\alpha^2} + \frac{n(n+1)\lambda}{\alpha^2}\sum_{i=1}^n q\mathbb{E}\left[w_{ij}p_{ij}\right] + \frac{n^2\lambda^2}{\alpha^2}\sum_{i=1}^n q\mathbb{E}\left[\lambda w_{ij}p_{ij}^3 - w_{ij}^2p_{ij}^2\right]$, $\Sigma_{\boldsymbol{\beta\beta}} = \frac{\psi(\alpha)}{4}\boldsymbol{X}'\boldsymbol{X} + \mathbb{E}(G)$, $\Sigma_{\alpha\boldsymbol{\beta}} = \mathbb{E}(\boldsymbol{B})$, $\Sigma_{\lambda\boldsymbol{\beta}} = \mathbb{E}(\boldsymbol{D})$, $\Sigma_{\alpha\lambda}^T = -\frac{n}{\alpha}\sum_{i=1}^n q\mathbb{E}(w_{ij}p_{ij}) + \frac{n\lambda}{\alpha}\sum_{i=1}^n q\mathbb{E}[\lambda w_{ij}p_{ij}^3 + w_{ij}^2p_{ij}^2]$ and $\Sigma_{\lambda\lambda} = \lambda\sum_{i=1}^n q\mathbb{E}[w_{ij}p_{ij}^3] + \sum_{i=1}^n q\mathbb{E}[w_{ij}^2p_{ij}^2]$. With $\psi(\alpha) = 2 + (4/\alpha^2) - (2\pi/\alpha^2)^{1/2}\left\{1 - \text{erf}[(2/\alpha^2)^{1/2}]\right\}\exp(2/\alpha^2)$, where erf(.) is the error function (see Ryshik and Grandstein, 1963). The expectations in the above expressions must be computed numerically. For $\lambda = 0$ we have that $\Sigma_{\boldsymbol{\theta\theta}} = \text{diag}\{2N/\alpha^2, \psi(\alpha)\boldsymbol{X}'\boldsymbol{X}/4, 2N/\pi\}$, which determinant is $N^2\psi(\alpha)|\boldsymbol{X}'\boldsymbol{X}|/(\pi\alpha^2) \neq 0$, then the information matrix is invertible in the case of products of univariate sinh-normal distributions (independent case).

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

$$\hat{\boldsymbol{\theta}} \xrightarrow{A} N_{p+2}(\boldsymbol{\theta}, \Sigma_{\boldsymbol{\theta\theta}}^{-1}),$$

resulting that the asymptotic variance of the MLE $\hat{\theta}$ is the inverse of $\Sigma_{\boldsymbol{\theta\theta}}$.

The approximation to $N_{p+2}(\boldsymbol{\theta}, \Sigma_{\boldsymbol{\theta\theta}}^{-1})$ can be used to construct confidence intervals for α, β_r, y λ. These are given by $\hat{\alpha} \mp z_{1-\delta/2}\sqrt{\hat{\sigma}(\hat{\alpha})}$, $\hat{\beta}_r \mp z_{1-\delta/2}\sqrt{\hat{\sigma}(\hat{\beta}_r)}$ and $\hat{\lambda} \mp z_{1-\delta/2}\sqrt{\hat{\sigma}(\hat{\lambda})}$, where $\hat{\sigma}(.)$ is found on the diagonal of the matrix $\Sigma_{\boldsymbol{\theta\theta}}^{-1}$ regarding to each parameter and $z_{1-\delta/2}$ is the quantile $100(\delta/2)\%$ of the standard normal distribution. Once we have proposed an asymmetric model as an alternative for a symmetrical model it is important to create a test for non asymmetry of the data. In our case the test can be formulated as

$$H_0: \quad \lambda = 0 \qquad \text{versus} \qquad \lambda \neq 0$$

since we have the asymptotic normality of the MLE, this hypothesis can be tested by using the test statistic (for large n)

$$-2\{\tilde{\ell}_p(\hat{\boldsymbol{\beta}}, \hat{\lambda}) - \ell_p(\tilde{\boldsymbol{\beta}}, 0)\} \sim \chi_1^2$$

where $\tilde{\boldsymbol{\beta}}$ is the MLE of $\boldsymbol{\beta}$ under the null hypothesis (restricted to H_0).

4.2 Testing the homogeneity of the shape parameter

In the previously proposed SMVSHN model we had assumed homogeneity of the shape parameter α. This assumption may be unrealistic since the response $\boldsymbol{y}_i$ may be related to the ith observation. This kind of situation has serious complications in the estimation process. Therefore it is necessary to verify the assumption of homogeneity of the shape parameter. A likelihood ratio test is now used to verify this assumption. Based on the method proposed by Xie and Wei (2007) and Qu and Xie (2011), we assume that the homogeneous model takes the form

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i, \quad i = 1, 2, \ldots, n, \qquad \boldsymbol{\epsilon}_i \sim SN_q(\alpha_i\boldsymbol{1}_q, \boldsymbol{0}_q, 2\boldsymbol{I}_q), \quad \alpha_i = \alpha k_i, \quad k_i = k(\boldsymbol{X}_i, \boldsymbol{\rho}), \tag{19}$$

where α is a factor of the shape parameter with weight function $k(\boldsymbol{X}_i, \boldsymbol{\rho})$ and $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_d)'$ is a d-dimensional vector of parameters which indicates the homogeneity. Under the model (19), the log-likelihood function for the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\rho}', \alpha, \boldsymbol{\beta}', \lambda)$ is expressed by

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}) = -\frac{N}{2}\log(8\pi) + \boldsymbol{1}_N'\log(\boldsymbol{\xi}_1^*) - \frac{\boldsymbol{\xi}_2^{*'}\boldsymbol{\xi}_2^*}{2} + \boldsymbol{1}_N'\log(\boldsymbol{\xi}_3^*), \tag{20}$$

with $\boldsymbol{\xi}_1^* = \boldsymbol{\xi}_1^*(\boldsymbol{\theta}) = \text{vec}(\boldsymbol{\xi}_{11}^*, \boldsymbol{\xi}_{21}^*, \ldots, \boldsymbol{\xi}_{n1}^*)$, $\boldsymbol{\xi}_2^* = \boldsymbol{\xi}_2^*(\boldsymbol{\theta}) = \text{vec}(\boldsymbol{\xi}_{12}^*, \boldsymbol{\xi}_{22}^*, \ldots, \boldsymbol{\xi}_{n2}^*)$ and $\boldsymbol{\xi}_3^* = \boldsymbol{\xi}_3^*(\boldsymbol{\theta}) = \text{vec}(\boldsymbol{\xi}_{13}^*, \boldsymbol{\xi}_{23}^*, \ldots, \boldsymbol{\xi}_{n3}^*)$ where

$$\boldsymbol{\xi}_{i1}^* = (\xi_{i11}^*, \xi_{i12}^*, \ldots, \xi_{i1q}^*)' = \frac{2}{\alpha k_i}\cosh\left(\frac{\boldsymbol{y}_i - \boldsymbol{x}_{ij}'\boldsymbol{\theta}}{2}\right),$$

$$\boldsymbol{\xi}^*_{i2} = (\xi^*_{i21}, \xi^*_{i22}, \ldots, \xi^*_{i2q})' = \frac{2}{\alpha k_i} \sinh\left(\frac{\boldsymbol{y}_i - \boldsymbol{x}'_{ij}\boldsymbol{\theta}}{2}\right)$$

and

$$\boldsymbol{\xi}^*_{i3} = (\xi^*_{i31}, \xi^*_{i32}, \ldots, \xi^*_{i3q})' = \Phi\left(\lambda \prod_{i=1}^{n} \boldsymbol{\xi}^*_{i2}\right),$$

with $\xi^*_{i1j} = \xi^*_{i1j}(\boldsymbol{\theta}) = \frac{2}{\alpha k_i} \cosh\left(\frac{y_{ij} - \boldsymbol{x}'_{ij}\boldsymbol{\theta}}{2}\right)$, $\xi^*_{i2j} = \xi^*_{i2j}(\boldsymbol{\theta}) = \frac{2}{\alpha k_i} \sinh\left(\frac{y_{ij} - \boldsymbol{x}'_{ij}\boldsymbol{\theta}}{2}\right)$ and $\xi^*_{i3j} = \xi^*_{i3j}(\boldsymbol{\theta}) = \Phi\left(\lambda \prod_{i=1}^{n} \xi^*_{i2j}\right)$.
Following the same procedure described in Lemonte (2013), the homogeneity test can be verified using the assumptions $H_0: \boldsymbol{\rho} = \boldsymbol{\rho}_0$ versus $H_1: \boldsymbol{\rho} \neq \boldsymbol{\rho}_0$, where $\boldsymbol{\rho}_0$ is a vector and it is assumed that exist only one $\boldsymbol{\rho}_0$ such as $k_i(\boldsymbol{X}, \boldsymbol{\rho}_0) = 1$. The statistical test to corroborate the above hypothesis is:

$$\omega = 2\left[\mathbf{1}'_N[\log\left(\hat{\boldsymbol{\xi}}^*_1\right) - \log\left(\tilde{\boldsymbol{\xi}}^*_1\right)] - \frac{1}{2}\left[\hat{\boldsymbol{\xi}}^{*'}_2 \hat{\boldsymbol{\xi}}^*_2 - \tilde{\boldsymbol{\xi}}^{*'}_2 \tilde{\boldsymbol{\xi}}^*_2\right] + \mathbf{1}'_N[\log\left(\hat{\boldsymbol{\xi}}^*_3\right) - \log\left(\tilde{\boldsymbol{\xi}}^*_3\right)]\right],$$

where $\hat{\boldsymbol{\xi}}^*_1 = \boldsymbol{\xi}^*_1(\hat{\boldsymbol{\theta}})$, $\hat{\boldsymbol{\xi}}^*_2 = \boldsymbol{\xi}^*_2(\hat{\boldsymbol{\theta}})$, $\hat{\boldsymbol{\xi}}^*_3 = \boldsymbol{\xi}^*_3(\hat{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{\xi}}^*_1 = \boldsymbol{\xi}^*_1(\tilde{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{\xi}}^*_2 = \boldsymbol{\xi}^*_2(\tilde{\boldsymbol{\theta}})$ and $\tilde{\boldsymbol{\xi}}^*_3 = \boldsymbol{\xi}^*_3(\tilde{\boldsymbol{\theta}})$. We can show that under the null hypothesis and usual regularity conditions that $\omega \overset{A}{\sim} \chi^2_d$, where $\overset{A}{\sim}$ denotes convergence in distribution. Therefore, it can br use the asymptotic approximation to the chi-square distribution to perform the test.

5 Illustrations

5.1 First illustration

Our first illustration of the multivariate sinh-normal model is the model without covariates. It is done by using a data set which contains two different measures of stiffness of each of 30 boards. The stiffness measures used to illustrate the model considered in this paper are "Shock" and "Vibration" of each of 30 boards. The data are reported in Johnson and Wichern (2007). The first measurement involves sending a shock wave down the board and the second measurement is determined while vibrating the board. The data was divided by 1000 for comfort. We initially fitted the model of Lemonte (2013) for independent random variaveis SHN without covariates, BVSHN, namely having $\beta_j = 0$ for $j = 2, 3, \ldots, p$ and $x_{ij1} = 1$ for $i = 1, 2, \ldots, n$, finding the maximum likelihood estimates (with standard errors in parentheses) $\hat{\alpha}_1 = 0.0036(0.0005)$, $\hat{\gamma}_1 = 1.9164(0.0606)$, $\hat{\sigma}_1 = 183.7498(43.1977)$, $\hat{\alpha}_2 = 0.0034(0.0005)$, $\hat{\gamma}_2 = 1.7470(0.0619)$ and $\hat{\sigma}_2 = 198.7043(49.9999)$ with AIC=44.6413.
We also fitted the bivariate sinh-normal model of Kundu (2014), CBSHN, with correlation coefficient ρ. We used the bivariate normal copula. The estimates here were $\hat{\alpha}_1 = 0.0042(0.0004)$, $\hat{\gamma}_1 = 1.9399(0.0740)$, $\hat{\sigma}_1 = 184.4394(51.5562)$, $\hat{\alpha}_2 = 0.0038(0.0003)$, $\hat{\gamma}_2 = 1.7917(0.0721)$, $\hat{\sigma}_2 = 198.5597(50.4597)$ and $\hat{\rho} = 0.9339(0.0389)$. Finally we fitted the skewed bivariate sinh-normal model , SBVSHN, obtaining the estimates $\hat{\alpha}_1 = 0.0037(0.0006)$, $\hat{\gamma}_1 = 1.9273(0.0272)$, $\hat{\sigma}_1 = 182.1839(48.5218)$, $\hat{\alpha}_2 = 0.0034(0.0005)$, $\hat{\gamma}_2 = 1.7457(0.0313)$, $\hat{\sigma}_2 = 197.9438(48.2723)$ and $\hat{\lambda} = -12.3822(5.8359)$.
For a better justification of the model SBVSHN against the model BVSHN, we consider the hypothesis testing of no difference between the model SBVSHN with the bivariate model under normality SHN. The hypothesis of interest is given by $H_0: \lambda = 0$ Vs $H_1: \lambda \neq 0$, which compares the model BVSHN with the model SBVSHN. To perform this test we use the likelihood ratio statistic based on

$$\Lambda = \frac{L_{BVSHN}(\hat{\alpha}_1, \hat{\gamma}_1, \hat{\sigma}_1, \hat{\alpha}_2, \hat{\gamma}_2, \hat{\sigma}_2)}{L_{SBVSHN}(\hat{\alpha}_1, \hat{\gamma}_1, \hat{\sigma}_1, \hat{\alpha}_2, \hat{\gamma}_2, \hat{\sigma}_2, \hat{\lambda})}.$$

We obtained $-2\log(\Lambda) = 27.9746$, which is higher than the value of the percentile of the chi-square distribution with one degree of freedom and 95% of confidence, which value is 3.84. Taking this to the rejection of null hypothesis (H_0) and thus a skewed model best fits the data set. The AIC criterion (see Akaike, 1974) for the model BVSHN is 44.6413, while for the model SBVSHN is 18.6667. Therefore we conclude that the model SBVSHN shows a better fit than the bivariate model SHN under normality.

In order to compare the model CBVSHN against the model SBVSHN, we apply a criterion for non-nested models. We use Vuong (1989) approach (generalized LR statistic) for comparing the SBVSHN and CBSHN non-nested models fitted to the data. A description of this procedure is described next. Being F_θ and G_ζ two non-nested models and $f(y_i|x_i,\theta)$ and $g(y_i|x_i,\zeta)$ the corresponding density functions, the likelihood ratio statistics to compare both models is given by

$$LR\left(\widehat{\theta},\widehat{\zeta}\right) = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \log \frac{f\left(y_i|x_i,\widehat{\theta}\right)}{g\left(y_i|x_i,\widehat{\zeta}\right)} \right\},$$

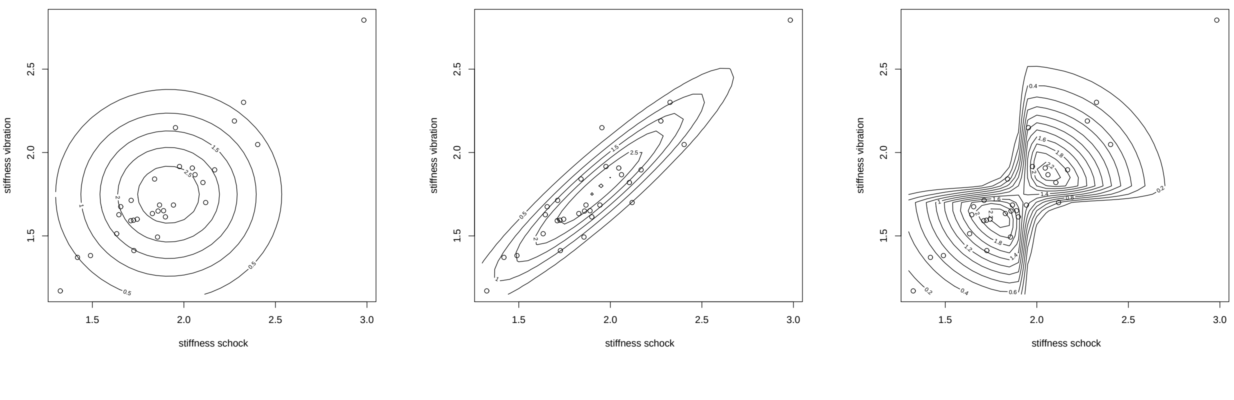
which does not follow a chi-square distribution. To overcome this problem, Vuong (1989) proposed an alternative approach based on the Kullback-Liebler divergence criterion. Based on the divergence between each model and the true process generating the data, namely the model $h^0(y|x)$, one arrives at the statistics

$$T_{LR,NN} = \frac{1}{\sqrt{n}} \frac{LR\left(\widehat{\theta},\widehat{\zeta}\right)}{\widehat{w}}, \tag{21}$$

where

$$\widehat{w}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\log \frac{f\left(y_i|x_i,\widehat{\theta}\right)}{g\left(y_i|x_i,\widehat{\zeta}\right)}\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\log \frac{f\left(y_i|x_i,\widehat{\theta}\right)}{g\left(y_i|x_i,\widehat{\zeta}\right)}\right)^2.$$

For strictly non-nested models, it can be shown that the statistic $T_{LR,NN}$ converges in distribution to a standard normal distribution under the null hypothesis. Thus the null hypothesis is not rejected if $|T_{LR,NN}| \leq z_{p/2}$. On the other hand, we reject at significance level p the null hypothesis of equivalence of the models in favor of model F_θ being better (or worse) than model G_ζ if $T_{LR,NN} > z_p$ (or $T_{LR,NN} < -z_p$). For testing $SBVSHN$ versus $CBSHN$, we obtain $|T_{LR,NN}| = 1.3374$ (p-value =0.1810). Therefore, the $CBSHN$ model not is significantly better than the $SBVSHN$ model according to the generalized LR statistic. It can be concluded that there is no significant difference between models $SBVSHN$ and $CBSHN$. However, in favor of the model $SBVSHN$ we have the fact that this model does not have the complicated estimation methods present in the estimation of models that involve a copula, such as in the model CBSHN, given the structure of this association. The following figures outline shows a good fit of the models CBSHN and SBVSHN.



(a) (b) (c)

Figure 2: Contour plots. (a) BVSHN, (b) CBSHN and (c) SBVSHN.

5.2 Second illustration

Our second illustration consider a real set data available on Lepadatu et al. (2005). These data are from a study where the main objective was to optimize the lifetime of dies and another responses (Von Misses stress, manufacturing force, etc.) in metal forming process, to improve the quality of the process by minimizing the effects of variation without eliminating the causes (since they are too difficult and very expensive to control). The BS distribution appears in this context given that according to the authors "the die of fatigue cracks are caused by repeated application of loads which individual would be too small to cause failure". The models considered here are: $y_{ij} = \log(T_{ij}) = \beta_1 + \beta_2 x_{ij2} + \epsilon_{ij}$, for $i = 1, 2, \ldots, 15$ $j = 1, 2, 3$, where T_1 is the Von Misses stress, T_2 is the manufacturing force, T_3 is the lifetime of die and x_{ij2} is the work temperature. Initially we assume that $\epsilon_{ij} \sim SHN(\alpha, 0, 2)$, for which we find that the MLEs are given by $\hat{\alpha} = 2.4166(0.2566)$, $\hat{\beta}_1 = 12.5295(1.2099)$ and $\hat{\beta}_2 = -0.0047(0.0016)$ with $AIC = -164.3584$. We also fitted the models under the assumption that $\boldsymbol{\epsilon} \sim SMVSHN(\alpha \mathbf{1}_{45}, \mathbf{0}_{45}, \boldsymbol{I}_{45}, \lambda)$. The MLEs under this assumption are $\hat{\alpha} = 2.6321(0.2854)$, $\hat{\beta}_1 = 10.3682(0.2961)$, $\hat{\beta}_2 = -0.0021(0.0002)$ and $\hat{\lambda} = -12.0998(1.1320)$, with $AIC = -154.5038$. As in the previous illustration we consider testing the hypothesis of no difference between the models SBVSHN and SHN under normality. The hypothesis of interest is $H_0 : \lambda = 0$ Vs $H_1 : \lambda \neq 0$, and the value of statistic of the test is

$$\Lambda = \frac{L_{MVSHN}(\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2)}{L_{SBVSHN}(\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2, \hat{\lambda})} = 0.0197.$$

For comparison we use $-2\log(\Lambda) = 7.8546$, which is greater than the chi-square value with one degree of freedom and a significance of 5%. Furthermore, taking into account the value of AIC for both cases, we conclude that the model with skewed errors (SMVSHN) best fits the data set. Finally we verify the homogeneity assumption of the shape parameter α. Then, as suggested by Cook and Weisberg (1983), power model is usually used in practical situations. Thus, we assume for simplicity that $k_i = x_i^{\rho}$, where x_i is the work temperature. Therefore $\rho = 0$ implies that $k_i = 1$ for $\alpha_i = \alpha$ for all i. Thus the homogeneity test of the shape parameter $\alpha_i = \alpha$ is equivalent to hypothesis test $H_0 : \rho = 0$. The value of the likelihood ratio statistic for this test is $\omega = 0.1686$ with $p_{\text{value}} = 0.6813$, which leads to non-rejection of the null hypothesis, concluding that the assumption of homogeneity of shape parameter is acceptable.

6 Conclusions

In this paper we have extended the log-Birnbaum-Saunders model by replacing the sinh-normal distribution by the skewed sinh-skew-normal distribution, which is a special case of the skew-normal distribution. This new family of distributions holds good properties, such as the marginal distributions of the dependent variables are univariate skewed log-Birnbaum-Saunders distribution and have the usual log-Birnbaum-Saunders distribution as a particular case. The usefulness of the extension considered here was checked by analyzing two real data sets. The model parameters of the illustrations were estimated by using maximum-likelihood methods and a closed-form expression for the Fisher's information matrix presented here was used to perform testing hypothesis for model parameters by using approximations obtained from the asymptotic normality of the MLEs. Finally, the multivariate version of the skewed log-Birnbaum-Saunders regression model proposed here shows good properties and a good fit to the illustrations data sets. Thus the model proposed here is a flexible alternative to others versions of the multivariate log-Birnbaum-Saunders regression since has the multivariate model proposed by Lemonte et al. (2015) as a particular case.

References

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

Arellano-Valle, R., Gómez, H., Quintana, A. (2005). Statistical inference for a general class of asymmetric distributions. *J Stat Plan Inference*. Vol. 128(2), 427–443

Arnold, B.C., Castilho, E., Sarabia, J.M. (2002). Conditionally specified multivariate skewed distributions. *Sankhya A*. Vol. 64, 206–226.

Barros, M. (2007). Modelos de Regressão Birnbaum-Saunders Generalizados. Tese de doutorado. *Instituto de Matemática e Estatística, Universidade de São Paulo, Brazil.*

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

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

Castillo, N., Gomez, H. and Bolfarine, H. (2011). Epsilon Birnbaum-Saunders distribution family: properties and inference. *Statistical Papers.* Vol. 52, 871–883.

Cox, D.R., Hinkley, D.V. (1974). *Theoretical Statistics.* Chapman and Hall: London.

Díaz-García, J. and Leiva, V. (2005). A new family of life distributions based on the elliptically contoured distributions. *J. Statist. Plann. Inference..* Vol. 128, 445–457.

Díaz-García, J. and Domínguez-Molina, J.R. (2006). Some generalisations of Birnbaum–Saunders and sinh–normal distributions. *International Mathematical Forum* Vol. 1, 1709–1727.

Galea, M., Leiva, V. and Paula, G. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics.* Vol. 31, 1049–1064.

Gomez, H., Olivares, J. and Bolfarine, H. (2009). An extension of the generalized Birnbaum-Saunders distribution. *Statistics and Probability Letters.* Vol 79, 331–338.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

Johnson, R.A., Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*, 6th ed. Prentice-Hall: New Jersey.

Kundu, D., Balakrishnan, N. and Jamalizadeh, A. (2010). Bivariate Birnbaum-Saunders distribution and associated inference, *Journal of Multivariate Analysis.* Vol. 101, 113–125

Kundu, D. Bivariate sinh-normal distribution and a related model. *Brazilian Journal of Probability and Statistics*, (to appear)

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

Lemonte, A. and Cordeiro, G. (2009). Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis.* Vol. 53, 4441–4452.

Lemonte, A. (2012). A log-Birnbaum-Saunders regression model with asymmetric errors. *Journal of Statistical Computation and Simulation.* Vol. 82, 1775–1787, 2012.

Lemonte, A. (2013). Multivariate Birnbaum–Saunders regression model. *Journal of Statistical Computation and Simulation.* Vol. 12, 2244–2257, 2013.

Lemonte, A., Martínez-Florez, G. and Moreno-Arenas, G. (2013). Multivariate Birnbaum-Saunders distribution: properties and associated inference. *Journal of Statistical Computation and Simulation.* Vol. 85, 374–392.

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

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

Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer, New York.

Owen, D.B. (1956). Tables for computing bivariate normal probabilities. *Annals of Mathematical Statistics*. Vol. 27, 1075–1090.

Qu, H. and Wei, B.C. (2011). Diagnostics analysis for log-Birnbaum-Saunders regression models with censored data. *Statistica Neerlandica*. Vol. 65, 1–21.

Rieck, J.R. and Nedelman, J.R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*. Vol. 33, 51–60.

Vilca-Labra, F. and Leiva-Sánchez, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics-Theory and Methods*. Vol. 35, 229–244.

Vuong, Q. (1989). Likelihood Ratio Tests for Model selection and Non-Nested Hypotheses. *Econometrica*. Vol. 57, 307–333.

Xie, F.C. and Wei, B.C. (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistics and Data Analysis*. Vol. 51, 4692–4706.