Inference for the bivariate Birnbaum-Saunders distribution

Luis Benites[1], Filidor Vilca[1], Víctor Leiva[2*]

[1]Departamento de Estatística, Universidade Estadual de Campinas, Brazil
[2]Instituto de Estadística, Universidad de Valparaíso, Chile

Abstract

Multivariate distributions are a topic largely studied and, particularly, because of its applicability, the bivariate case is often taken into account. Birnbaum-Saunders distributions have been widely considered due to their good properties and useful for modeling different types of phenomena. We investigate estimation and hypothesis testing in the bivariate Birnbaum-Saunders distribution. About estimation, modified moment and maximum likelihood methods are employed. We prove that the modified moment estimators are consistent and asymptotically normal distributed. Regarding hypothesis testing, likelihood ratio, score and Wald statistics are analyzed. We obtain the Fisher information in a matrix form, which facilitates the implementation of the score and Wald statistics. We validate our approach with simulated and real-world data. Our study provides new findings and improves the results proposed until now on this topic.

Keywords: asymptotic tests; data analysis; moment and maximum likelihood estimators; Monte Carlo simulation.

Mathematics Subject Classification: Primary: 62E15; Secondary: 62F03/5/10/12

1 Introduction

Diverse real-world phenomena can be described by multivariate distributions. Particularly, because of its applicability, bivariate methods are often taking into account; see Johnson & Wichern (1999) and Kotz et al (2000). Birnbaum & Saunders (1969a) constructed a distribution from a physical setting related to material fatigue, which has several good properties and a close relation to the normal distribution; see Johnson et al (1994, Vol. II pp. 651-663), Balakrishnan et al (2009a) and Kotz et al (2010). Due to its construction, applications of the Birnbaum-Saunders (BS) model to lifetime data are carried out in a natural way; see, e.g., Villegas et al (2011) and Leiva et al (2014b). However, its applications have been also conducted in business, environment, industry and medicine; see, e.g., Barros et al (2008, 2014), Balakrishnan et al (2009b), Vilca et al (2010, 2011), Leiva et al (2011, 2014a,c,d), Ferreira et al (2012), Paula et al (2012) and Marchant et al (2013).

[*]Corresponding author: Víctor Leiva. Instituto de Estadística, Universidad de Valparaíso, Avda. Gran Bretaña 1111, Playa Ancha, Valparaíso, Chile. Email: victor.leiva@uv.cl. URL: www.deuv.cl/leiva

Bivariate extensions of univariate distributions are natural ways to be considered and have been obtained for several continuous asymmetric distributions, such as beta, exponential, gamma, uniform and Weibull; see Johnson et al (1994) for univariate distributions and Kotz et al (2000) for their bivariate cases. Recently, Kundu et al (2010) presented an extension of the univariate BS distribution to the bivariate case (BS_2 in short). Other multivariate extensions of the BS distribution are attributed to Caro-Lopera et al (2012), Kundu et al (2013) and Sanchez et al (2014).

Maximum likelihood (ML) estimation is widely used and it has been considered for BS distributions; see Birnbaum & Saunders (1969b) and Kundu et al (2010). Moment estimation is simpler than the ML estimation but, as is well known, the moment estimators may not be unique and not always exist. This occurs when the parameters of the BS_2 distribution are estimated. An alternative way to estimate these parameters is to use the modified moment (MM) method, which provides estimators that always exist uniquely; see Ng et al (2003), Leiva et al (2008) and Kundu et al (2010).

The main objectives of this paper are to estimate the parameters of the BS_2 distribution and to construct hypothesis tests for them. First, we estimate these parameters with the MM method and prove these estimators are consistent and asymptotically normal distributed. Second, we derive the corresponding ML estimators under restrictions imposed for the BS_2 parameters considering some postulated hypotheses. Third, we construct hypothesis tests for these parameters by means of the likelihood ratio (LR), score and Wald statistics. This latter statistic is simpler to compute than LR and score statistics. We obtain the Fisher information in a matrix form, which improves the results reached by Kundu et al (2010) and facilitates the implementation of the score and Wald statistics. Fourth, we conduct numerical studies by Monte Carlo (MC) simulations to evaluate the performance of the proposed results and use them to analyze real-world bivariate data illustrating their applicability.

The paper is organized as follows. In Section 2, we present the BS_2 distribution and some of its properties. In Section 3, we provide parameter estimation based on the MM and ML methods. Also, the asymptotic distribution of the MM estimator is derived and closed form expressions for the corresponding expected Fisher information matrix are obtained. In Section 4, we propose hypothesis tests based on LR, score and Wald statistics. In Section 5, we conduct the numerical part of this study including simulated and real-world data. Finally, in Section 6, we sketch the conclusions of this work.

2 Birnbaum-Saunders distributions

In this section, we provide a background about univariate and bivariate BS distribution.

2.1 The univariate BS distribution

The univariate BS distribution is related to the univariate normal distribution by

$$T = \beta(\alpha\, Z/2 + \sqrt{(\alpha\, Z/2)^2 + 1})^2, \tag{1}$$

where $Z \sim \mathrm{N}(0,1)$, $\alpha > 0$ and $\beta > 0$. The random variable T given in (1) is said to have a BS distribution with shape and scale parameters, α and β, respectively, which is denoted by $T \sim \mathrm{BS}(\alpha,\beta)$. The cumulative distribution function (CDF) of T is given by $F_T(t) = \Phi(a(t;\alpha,\beta))$, for $t > 0$, where $\Phi(\cdot)$ is the standard normal CDF and $a(t;\alpha,\beta) = (\sqrt{t/\beta} - \sqrt{\beta/t})/\alpha$.

2.2 The bivariate BS distribution

A bivariate random vector $\boldsymbol{T} = (T_1, T_2)^\top$ is said to have a BS_2 distribution with shape ($\alpha_1 > 0$, $\alpha_2 > 0$), scale ($\beta_1 > 0$, $\beta_2 > 0$) and correlation ($-1 < \rho < 1$) parameters, if the joint CDF of T_1 and T_2 can be expressed as $F_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \rho) = \Phi_2(a(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho)$, for $\boldsymbol{t} = (t_1, t_2)^\top \in \mathbb{R}_+^2$, where $\Phi_2(\cdot; \rho)$ is the joint CDF of $\boldsymbol{Z} = (Z_1, Z_2)^\top \sim \mathrm{N}_2(\boldsymbol{0}, \boldsymbol{\Sigma})$, with $\sigma_{11} = \sigma_{22} = 1$ and correlation coefficient ρ, $a(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = (a(t_1; \alpha_1, \beta_1), a(t_2; \alpha_2, \beta_2))^\top$, with $a(t_j; \alpha_j, \beta_j) = (\sqrt{t_j/\beta_j} - \sqrt{\beta_j/t_j})/\alpha_j$, for $j = 1, 2$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top$. In this case, we use the notation $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$.

From Kundu et al (2010), $\boldsymbol{T}$ has a joint probability density function (PDF) given by

$$f_{\boldsymbol{T}}(\boldsymbol{t}) = \tfrac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-d(\boldsymbol{t})/2\right) A(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \phi_2\big(a(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho\big) A(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \boldsymbol{t} \in \mathbb{R}_+^2, \tag{2}$$

where $d(\boldsymbol{t}) = d(\boldsymbol{t}, \boldsymbol{\theta}) = a(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} a(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta})$ is the Mahalanobis distance (MD), with $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)^\top$, and $\phi_2(.; \rho)$ is the PDF of $\boldsymbol{Z}$ given by

$$\phi_2(\boldsymbol{z}; \rho) = \phi_2(z_1, z_2; \rho) = \tfrac{1}{2\pi\sqrt{1-\rho^2}} \exp\Big(-\tfrac{1}{2(1-\rho^2)}(z_1^2 + z_2^2 - 2\rho z_1 z_2)\Big), \ \boldsymbol{z} = (z_1, z_2)^\top \in \mathbb{R}^2,$$

and $A(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = A(t_1; \alpha_1, \beta_1) A(t_2; \alpha_2, \beta_2)$, with $A(t_j; \alpha_1, \beta_1) = t_j^{-3/2}(t_j + \beta_j)/(2\alpha_j \beta_j^{1/2})$, for $j = 1, 2$. Some properties of the BS_2 distribution can be derived by using the close relationship between the distributions of $\boldsymbol{T}$ and $\boldsymbol{Z}$; see Kundu et al (2010). Let $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$. Then,
(i) $T_j \sim \mathrm{BS}(\alpha_j, \beta_j)$, for $j = 1, 2$;
(ii) $\boldsymbol{c} \odot \boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{c} \odot \boldsymbol{\beta}, \rho)$, where $\boldsymbol{c} = (c_1, c_2)^\top \in \mathbb{R}_+^2$ and $\odot$ denotes the Hadamard product;
(iii) $\boldsymbol{T}^{-1} = (T_1^{-1}, T_2^{-1})^\top \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}, \rho)$, where $\boldsymbol{\beta}^{-1} = (1/\beta_1, 1/\beta_2)^\top$;
(iv) $\boldsymbol{T}_1^{-1} = (T_1^{-1}, T_2)^\top \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}_1^{-1}, -\rho)$, where $\boldsymbol{\beta}_1^{-1} = (1/\beta_1, \beta_2)^\top$;
(v) $\boldsymbol{T}_2^{-1} = (T_1, T_2^{-1})^\top \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}_2^{-1}, -\rho)$, where $\boldsymbol{\beta}_2^{-1} = (\beta_1, 1/\beta_2)^\top$;
(vi) T_1 and T_2 are independent if and only if $\rho = 0$;
(vii) The conditional PDF of T_1 given $T_2 = t_2$ is

$$f_{T_1|(T_2=t_2)}(t_1) = \phi\big(a(t_1; \alpha_{1\rho}, \beta_1) - \mu_1(t_2)\big)\, A(t_1; \alpha_{1\rho}, \beta_1), \tag{3}$$

where $\phi(\cdot)$ is the PDF of the N(0,1) distribution, $\alpha_{1\rho} = \sqrt{1-\rho^2}\,\alpha_1$ and $\mu_1(t_2) = \rho\, a(t_2; \alpha_{2\rho}, \beta_2)$, with $\alpha_{2\rho} = \sqrt{1-\rho^2}\,\alpha_2$. From the PDF given in (3), note that the conditional distribution of T_1 given $T_2 = t_2$ is univariate non-central BS (NBS), which is denoted by $T_1|(T_2 = t_2) \sim \mathrm{NBS}(\alpha_{1\rho}, \beta_1, \mu_1(t_2))$; see Guiraud et al (2009).

Properties above mentioned allow us to obtain the following: $(\alpha_1 T_1/\beta_1, \alpha_2 T_2/\beta_2)^\top \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\alpha}, \rho)$ and $(T_1/\beta_1, T_2/\beta_2)^\top \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{1}_2, \rho)$, where $\boldsymbol{1}_2 = (1, 1)^\top$. Consequently, the distribution of any function of T_1/β_1 and T_2/β_2 does not depend on β_1 and β_2, such as $T_1T_2/(\beta_1\beta_2)$ and $\sqrt{T_1T_2}/\sqrt{\beta_1\beta_2}$. Thus, according to Kundu et al (2010), their expectations are given by

$$\begin{aligned}
\psi_1^*(\boldsymbol{\alpha}, \rho) &= \mathrm{E}[T_1T_2/(\beta_1\beta_2)] = 1 + (\alpha_1^2 + \alpha_2^2)/2 + \alpha_1^2\alpha_2^2(1+\rho^2)/4 + \alpha_1\alpha_2 I_1(\rho), && (4)\\
\psi_1(\boldsymbol{\alpha}, \rho) &= \mathrm{E}[\sqrt{T_1T_2}/\sqrt{\beta_1\beta_2}] = \mathrm{E}[\sqrt{\beta_1\beta_2}/\sqrt{T_1T_2}] = \alpha_1\alpha_2\rho/4 + I_2(\rho), && (5)
\end{aligned}$$

where $I_1(\rho) = \mathrm{E}[Z_1Z_2\, Z_{12}]$ and $I_2(\rho) = \mathrm{E}[Z_{12}]$, with $Z_{12} = \sqrt{(\alpha_1Z_1/2)^2 + 1}\,\sqrt{(\alpha_2Z_2/2)^2 + 1}$, and $\boldsymbol{Z} = (Z_1, Z_2)^\top \sim \mathrm{N}_2(\boldsymbol{0}, \boldsymbol{\Sigma})$. It can be seen that $\psi_1(\boldsymbol{\alpha}, -\rho) = \mathrm{E}[\sqrt{\beta_1T_2}/\sqrt{T_1\beta_2}] = \mathrm{E}[\sqrt{\beta_2T_1}/\sqrt{T_2\beta_1}]$.

Expressions for $I_1(\rho)$ and $I_2(\rho)$ given in (4) and (5) are defined by (see Kundu et al, 2010)

$$\begin{aligned} I_1(\rho) &= a_{0,0}+\tfrac{1}{2^3}a_{0,1}(\alpha_1^2+\alpha_2^2)+\tfrac{1}{2^6}\alpha_1^2\alpha_2^2a_{1,1}+\sum_{i=2}^{\infty}(-1)^{i-1}\tfrac{1\cdot3\cdots(2i-3)}{2^{3i}i!}a_{0,i}(\alpha_1^{2i}+\alpha_2^{2i}) \qquad (6)\\ &+\sum_{i=2}^{\infty}(-1)^{i-1}\tfrac{1\cdot3\cdots(2i-3)}{2^{3i+3}i!}a_{1,i}(\alpha_1^2\alpha_2^{2i}+\alpha_2^2\alpha_1^{2i})+\sum_{i=2}^{\infty}\sum_{j=2}^{\infty}(-1)^{i+j}\tfrac{1\cdot3\cdots(2i-3)}{2^{3i}i!}\tfrac{1.3\ldots(2j-3)}{2^{3j}j!}\alpha_1^{2i}\alpha_2^{2j}a_{i,j},\\ I_2(\rho) &= 1+\frac{1}{2^3}(\alpha_1^2+\alpha_2^2)+\frac{1}{2^6}\alpha_1^2\alpha_2^2(1+2\rho^2)+\sum_{i=2}^{\infty}(-1)^{i-1}\frac{1\cdot3\cdots(2i-3)}{2^{3i}i!}b_{0,i}(\alpha_1^{2i}+\alpha_2^{2i}) \qquad (7)\\ &+\sum_{i=2}^{\infty}(-1)^{i-1}\tfrac{1\cdot3\cdots(2i-3)}{2^{3i+3}i!}b_{1,i}(\alpha_1^2\alpha_2^{2i}+\alpha_2^2\alpha_1^{2i})+\sum_{i,j=2}^{\infty}(-1)^{i+j}\tfrac{1\cdot3\cdots(2i-3)}{2^{3i}i!}\tfrac{1\cdot3\cdots(2j-3)}{2^{3j}j!}\alpha_1^{2i}\alpha_2^{2j}b_{i,j}, \end{aligned}$$

where $a_{m,n}=\mathrm{E}[Z_1^{2m+1}Z_2^{2n+1}]$, $b_{m,n}=\mathrm{E}[Z_1^{2m}Z_2^{2n}]$, for non-negative integers m and n.

3 Estimation

In this section, we estimate the model parameters based on MM and ML methods with n independent observations extracted from the BS_2 distribution. Kundu et al (2010) developed ML estimation of these parameters, but we provide expressions for the ML estimates under the null hypotheses, whose implementation is easier than those presented by these authors.

3.1 Modified moment estimation

Let $\boldsymbol{T}_1,\ldots,\boldsymbol{T}_n$ be a sample of size n from $\boldsymbol{T}\sim\mathrm{BS}_2(\boldsymbol{\alpha},\boldsymbol{\beta},\rho)$. Then, the MM estimators of the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, denoted respectively by $\widehat{\boldsymbol{\alpha}}_M$ and $\widehat{\boldsymbol{\beta}}_M$, have elements obtained by equating the population moments $\mathrm{E}[T_j]$ and $\mathrm{E}[T_j^{-1}]$ to sample moments $S_j=\sum_{i=1}^n T_{ji}/n$ and $R_j=((1/n)\sum_{i=1}^n T_{ji}^{-1})^{-1}$, for $j=1,2$, respectively, given by

$$\widehat{\alpha}_{Mj}=\big(2\big((S_j/R_j)^{1/2}-1\big)\big)^{1/2} \quad \text{and} \quad \widehat{\beta}_{Mj}=(S_jR_j)^{1/2}. \qquad (8)$$

Because the MM estimators have an explicit form, they can be used in practical situations, for example, as initial values in the iterative procedure for computing the ML estimates.

The following lemmas provide tools to study the asymptotic distribution of the estimators presented in (8) and the ML estimators, and to obtain the Fisher information matrix of the BS_2 distribution. According to Fang & Zhang (1990, p. 40), we use $\mathrm{Cov}[\boldsymbol{T},\boldsymbol{V}]$ for denoting the covariance matrix between the vectors $\boldsymbol{T}$ and $\boldsymbol{V}$. If $\boldsymbol{T}=\boldsymbol{V}$, we use $\mathrm{Var}[\boldsymbol{T}]=\mathrm{Cov}[\boldsymbol{T},\boldsymbol{T}]$ for denoting the variance-covariance matrix of $\boldsymbol{T}$.

Lemma 1. *Let $\boldsymbol{T}=(T_1,T_2)^\top\sim\mathrm{BS}_2(\boldsymbol{\alpha},\boldsymbol{\beta},\rho)$ and $a(\boldsymbol{T};\boldsymbol{\alpha},\boldsymbol{\beta})=(a(T_1;\alpha_1,\beta_1),a(T_2;\alpha_2,\beta_2))^\top$. Then, for non-negative integers m and n, we have*

(a) $\mathrm{E}[a^{2m+1}(T_1,1,\beta_1)a^{2n+1}(T_2,1,\beta_2)]=\alpha_1^{2m+1}\alpha_2^{2n+1}a_{m,n}$;

(b) $\mathrm{E}[a^{2m}(T_1,1,\beta_1)a^{2n}(T_2,1,\beta_2)]=\alpha_1^{2m}\alpha_2^{2n}\,b_{m,n}$,

where $a_{m,n}$ and $b_{m,n}$ are as given in (6) *and* (7), *respectively. A special case of (a) and (b) is*

$$\mathrm{E}[a(T_1,1,\beta_1)a(T_2,1,\beta_2)]=\alpha_1\alpha_2\rho.$$

Proof. It is obtained from the relation between the BS and normal distributions, and the approximations provided in Kundu et al (2010).

Lemma 2. *Let* $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ *and* $\boldsymbol{T}^{-1} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}, \rho)$*. Then, we have*

$$
\begin{aligned}
\mathrm{Var}[\boldsymbol{T}] &= D^2(\boldsymbol{\alpha})D^2(\boldsymbol{\beta})\left(\boldsymbol{I}_2 + \tfrac{5}{4}D^2(\boldsymbol{\alpha})\right) + \beta_1\beta_2 C(\boldsymbol{\alpha}, \rho)\boldsymbol{J}_{12},\\
\mathrm{Var}[\boldsymbol{T}^{-1}] &= D^2(\boldsymbol{\alpha})D^{-2}(\boldsymbol{\beta})\left(\boldsymbol{I}_2 + \tfrac{5}{4}D^2(\boldsymbol{\alpha})\right) + \tfrac{1}{\beta_1\beta_2} C(\boldsymbol{\alpha}, \rho)\boldsymbol{J}_{12},\\
\mathrm{Cov}[\boldsymbol{T}, \boldsymbol{T}^{-1}] &= -D^2(\boldsymbol{\alpha})\left(\tfrac{1}{4}D^2(\boldsymbol{\alpha}) + \boldsymbol{I}_2\right) + C(\boldsymbol{\alpha}, -\rho)D(\boldsymbol{\beta})\boldsymbol{J}_{12}D^{-1}(\boldsymbol{\beta}),
\end{aligned}
$$

where $D(\boldsymbol{a}) = \mathrm{diag}(a_1, a_2)$ *and* $D^{-1}(\boldsymbol{a}) = \mathrm{diag}(1/a_1, 1/a_2)$*, that is,* $D(\boldsymbol{a})$ *denotes a diagonal matrix formed by the vector* $\boldsymbol{a} = (a_1, a_2)^\top$*. In addition,* $C(\boldsymbol{\alpha}, \rho) = \psi_1^*(\boldsymbol{\alpha}, \rho) - (1 + \alpha_1^2/2)(1 + \alpha_2^2/2)$*, with* $\psi_1^*(\boldsymbol{\alpha}, \rho)$ *being as given in* (4) *and* $\boldsymbol{J}_{12} = \boldsymbol{1}_2\boldsymbol{1}_2^\top - \boldsymbol{I}_2$ *being a* 2×2 *symmetric matrix.*

Proof. From (4) and by considering $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ and the distributions of $\boldsymbol{T}^{-1}$, $\boldsymbol{T}_1^{-1}$ and $\boldsymbol{T}_2^{-1}$ provided in properties (iii)-(v) of Subsection 2.2, we have

$$
\begin{aligned}
\mathrm{E}[T_1T_2] &= \beta_1\beta_2\big(1 + (\alpha_1^2 + \alpha_2^2)/2 + \alpha_1^2\alpha_2^2(1 + \rho^2)/4 + \alpha_1\alpha_2 I_1(\rho)\big),\\
\mathrm{E}[1/(T_1T_2)] &= \big(1/(\beta_1\beta_2)\big)\big(1 + (\alpha_1^2 + \alpha_2^2)/2 + \alpha_1^2\alpha_2^2(1 + \rho^2)/4 + \alpha_1\alpha_2 I_1(\rho)\big),\\
\mathrm{E}[T_1/T_2] &= (\beta_1/\beta_2)\big(1 + (\alpha_1^2 + \alpha_2^2)/2 + \alpha_1^2\alpha_2^2(1 + \rho^2)/4 + \alpha_1\alpha_2 I_1(-\rho)\big),\\
\mathrm{E}[T_2/T_1] &= (\beta_2/\beta_1)\big(1 + (\alpha_1^2 + \alpha_2^2)/2 + \alpha_1^2\alpha_2^2(1 + \rho^2)/4 + \alpha_1\alpha_2 I_1(-\rho)\big).
\end{aligned}
$$

Thus, using the definition of the covariance matrix between vectors and the covariance matrices of $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$, $\mathrm{Var}[\boldsymbol{T}_j]$ say, for $j = 1, 2$, and performing some matrix algebra, the proof is complete. ■

The following theorem provides the asymptotic distribution of the estimators $\widehat{\boldsymbol{\alpha}}_M$ and $\widehat{\boldsymbol{\beta}}_M$, with elements given in (8), which can be used for constructing confidence regions and hypothesis tests for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Theorem 1. *Let* $\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n$ *be a sample of size* n *from* $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ *and* $\widehat{\boldsymbol{\alpha}}_M$ *and* $\widehat{\boldsymbol{\beta}}_M$ *be the MM estimators of* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\beta}$*, respectively. Then, for* ρ *fixed, we have*

$$
\sqrt{n}\left(\begin{pmatrix} \widehat{\boldsymbol{\alpha}}_M \\ \widehat{\boldsymbol{\beta}}_M \end{pmatrix} - \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix}\right) \xrightarrow{\mathrm{d}} \mathrm{N}_4\left(\boldsymbol{0}, \begin{pmatrix} \boldsymbol{\Upsilon}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Upsilon}_{\boldsymbol{\beta}\boldsymbol{\beta}} \end{pmatrix}\right), \quad \text{as } n \to \infty,
$$

where $\xrightarrow{\mathrm{d}}$ *denotes "convergence in distribution to",* $\boldsymbol{\Upsilon}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} = \frac{1}{2}\big(D(\boldsymbol{\alpha})\boldsymbol{\Sigma}_\rho^* D(\boldsymbol{\alpha}) + (I_1(\rho) + I_1(-\rho))\boldsymbol{J}_{12}\big)$*, with* $I_1(\cdot)$ *being as given in* (6) *and* $\boldsymbol{\Sigma}_\rho^* = \boldsymbol{I}_2 + (\rho^2/2)\boldsymbol{J}_{12}$*, and* $\boldsymbol{\Upsilon}_{\boldsymbol{\beta}\boldsymbol{\beta}}$ *has elements*

$$
\begin{aligned}
\gamma_{11} &= (\alpha_1\beta_1)^2\frac{(3\alpha_1^2 + 4)}{(\alpha_1^2 + 2)^2},\\
\gamma_{12} &= \frac{2\beta_1\beta_2}{(\alpha_1^2 + 2)(\alpha_2^2 + 2)}\left(I_1(\rho) - I_1(-\rho)\right),\\
\gamma_{22} &= (\alpha_2\beta_2)^2\frac{(3\alpha_2^2 + 4)}{(\alpha_2^2 + 2)^2}.
\end{aligned}
$$

Proof. It follows from Lemma 2 and the delta method. Specifically, let

$$\boldsymbol{S} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{P} = \begin{pmatrix} 1/R_1 \\ 1/R_2 \end{pmatrix},$$

with S_j and R_j being given in (8). Thus, by using the central limit theorem, we have

$$\sqrt{n}\left(\begin{pmatrix} \boldsymbol{S} \\ \boldsymbol{P} \end{pmatrix} - \begin{pmatrix} \mathrm{E}[\boldsymbol{T}] \\ \mathrm{E}[\boldsymbol{T}^{-1}] \end{pmatrix}\right) \xrightarrow{\mathrm{d}} \mathrm{N}_4\left(\boldsymbol{0}, \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{\Psi}_{12} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} \end{pmatrix}\right), \quad \text{as } n \to \infty,$$

where $\boldsymbol{\Psi}_{11} = \mathrm{Var}[\boldsymbol{T}]$, $\boldsymbol{\Psi}_{12} = \mathrm{Cov}[\boldsymbol{T}, \boldsymbol{T}^{-1}]$, $\boldsymbol{\Psi}_{21} = \boldsymbol{\Psi}_{12}^\top$ and $\boldsymbol{\Psi}_{22} = \mathrm{Var}[\boldsymbol{T}^{-1}]$ given in Lemma 2. Now, from (8), we have

$$\sqrt{n}\left(\begin{pmatrix} \widehat{\boldsymbol{\alpha}}_M \\ \widehat{\boldsymbol{\beta}}_M \end{pmatrix} - \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix}\right) = \sqrt{n}\left(g\begin{pmatrix} \boldsymbol{S} \\ \boldsymbol{P} \end{pmatrix} - g\begin{pmatrix} \mathrm{E}[\boldsymbol{T}] \\ \mathrm{E}[\boldsymbol{T}^{-1}] \end{pmatrix}\right),$$

where $g(\cdot)$ is a differentiable function defined by

$$g(x_1, x_2, x_3, x_4) = \big(g_1(x_1, x_3), g_2(x_2, x_4), g_3(x_1, x_3), g_4(x_2, x_4)\big)^\top,$$

with $g_1(x_1, x_3) = \sqrt{2}\sqrt{\sqrt{x_1 x_3} - 1}$, $g_2(x_2, x_4) = \sqrt{2}\sqrt{\sqrt{x_2 x_4} - 1}$, $g_3(x_1, x_3) = \sqrt{x_1/x_3}$ and $g_4(x_2, x_4) = \sqrt{x_2/x_4}$. For instance, $\widehat{\alpha}_{M1} = g_1(S_1, 1/R_1)$, $\widehat{\alpha}_{M2} = g_2(S_2, 1/R_2)$, $\widehat{\beta}_{M1} = g_3(S_1, 1/R_1)$ and $\widehat{\beta}_{M2} = g_4(S_2, 1/R_2)$. Therefore, by applying the delta method, we have

$$\sqrt{n}\left(\begin{pmatrix} \widehat{\boldsymbol{\alpha}}_M \\ \widehat{\boldsymbol{\beta}}_M \end{pmatrix} - \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix}\right) \xrightarrow{\mathrm{d}} \mathrm{N}_4\left(\boldsymbol{0}, \boldsymbol{G}^\top(\boldsymbol{\alpha}, \boldsymbol{\beta}) \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{\Psi}_{12} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} \end{pmatrix} \boldsymbol{G}(\boldsymbol{\alpha}, \boldsymbol{\beta})\right), \quad \text{as } n \to \infty,$$

where

$$\boldsymbol{G}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{2}\begin{pmatrix} D^{-1}(\boldsymbol{\alpha}) D^{-1}(\boldsymbol{\beta}) & D(\boldsymbol{\gamma}) \\ D^{-1}(\boldsymbol{\alpha}) D(\boldsymbol{\beta}) & -D^2(\boldsymbol{\beta}) D(\boldsymbol{\gamma}) \end{pmatrix},$$

with $\boldsymbol{\gamma} = (2/(2+\alpha_1^2), 2/(2+\alpha_2^2))^\top$. Thus, after some algebra, the theorem is proven. ■

Corollary 1. *Let $\widehat{\boldsymbol{\alpha}}_M$ and $\widehat{\boldsymbol{\beta}}_M$ be the MM estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively. Then, we have*

$$\sqrt{n}(\widehat{\boldsymbol{\alpha}}_M - \boldsymbol{\alpha}) \xrightarrow{\mathrm{d}} \mathrm{N}_2(\boldsymbol{0}, \boldsymbol{\Upsilon}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}) \ \text{ and } \ \sqrt{n}(\widehat{\boldsymbol{\beta}}_M - \boldsymbol{\beta}) \xrightarrow{\mathrm{d}} \mathrm{N}_2(\boldsymbol{0}, \boldsymbol{\Upsilon}_{\boldsymbol{\beta}\boldsymbol{\beta}}), \quad \text{as } n \to \infty.$$

Corollary 1 can be used for constructing confidence intervals (CIs) for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, with ρ fixed.

Note that the MM method only provides estimators for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of the BS_2 distribution, given in (8), but not for ρ. One way to solve it is to obtain a moment-type estimator for ρ by using the relationship between the BS and normal distributions established in (1), which is naturally extended to the bivariate case. If consistent estimators of β_1 and β_2 are available, then a moment-type estimator of ρ and its consistency and asymptotic distribution are presented in the following theorem.

Theorem 2. *Let $\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n$ be a sample of size n from $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ and $\breve{\beta}_1$ and $\breve{\beta}_2$ be consistent estimators of β_1 and β_2, respectively. Then, the moment-type estimator of ρ given by*

$$\breve{\rho} = \frac{\sum_{j=1}^{n} a(T_{1i}; 1, \breve{\beta}_1) a(T_{2i}; 1, \breve{\beta}_2)}{\sqrt{\sum_{j=1}^{n} a^2(T_{1i}; 1, \breve{\beta}_1) \, \sum_{j=1}^{n} a^2(T_{2i}; 1, \breve{\beta}_2)}}$$

is consistent and asymptotically normal distributed, that is, by applying properties of almost sure convergence for $n \to \infty$, $\breve{\rho} \xrightarrow{\text{a.s.}} \rho$ and $\sqrt{n}(\breve{\rho} - \rho) \xrightarrow{\text{d}} \mathrm{N}(0, (1-\rho^2)^2)$.

Proof. First, let β_1 and β_2 to be fixed and $\boldsymbol{W}_i = a(\boldsymbol{T}_i; \boldsymbol{1}_2, \boldsymbol{\beta})$ to follow a bivariate normal distribution, that is, $\boldsymbol{W}_i \sim \mathrm{N}_2(\boldsymbol{0}, D(\boldsymbol{\alpha})\boldsymbol{\Sigma} D(\boldsymbol{\alpha}))$, for $i = 1, \ldots, n$. Thus, we define the 2×2 sample variance-covariance matrix as $\boldsymbol{S} = S_n(\beta_1, \beta_2) = (1/n)\sum_{i=1}^{n} \boldsymbol{W}_i \boldsymbol{W}_i^{\top}$, with its entries denoted by S_{11}, S_{12} and S_{22}. We can study the asymptotic properties of $\boldsymbol{S}$ and, through them, the properties of $\breve{\rho} = g_3(\boldsymbol{S}) = S_{12}/\sqrt{S_{11}\, S_{22}}$. Then,

$$\sqrt{n} \begin{pmatrix} S_{11} - \alpha_1^2 \\ S_{12} - \alpha_1\alpha_2\rho \\ S_{22} - \alpha_2^2 \end{pmatrix} \xrightarrow{\text{d}} \mathrm{N}_3(\boldsymbol{0}, \boldsymbol{\Upsilon}), \quad \text{as } n \to \infty,$$

where

$$\boldsymbol{\Upsilon} = \begin{pmatrix} 2\alpha_1^4 & 2\alpha_1^3\alpha_2\rho & 2\alpha_1^2\alpha_2^2\rho^2 \\ 2\alpha_1^3\alpha_2\rho & \alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_2^2\rho^2 & 2\alpha_1\alpha_2^3\rho \\ 2\alpha_1^2\alpha_2^2\rho^2 & 2\alpha_1\alpha_2^3\rho & 2\alpha_2^4 \end{pmatrix}.$$

The consistency of the estimator $\breve{\rho}$ follows from the fact that $\boldsymbol{S}$ tends to $\boldsymbol{\Upsilon}$ with probability one. Now, applying the delta method, we obtain the asymptotic distribution of $\breve{\rho}$. The required result then follows for β_1 and β_2 fixed. Therefore, by substituting β_1 and β_2 with their respective consistent estimators $\breve{\beta}_1$ and $\breve{\beta}_2$ into the reached result and then applying properties of almost sure convergence and the delta method, the proof of the theorem is complete. ∎

3.2 Maximum likelihood estimation

Let $\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n$ be a sample of size n from $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$. Then, the log-likelihood function for $\boldsymbol{\theta}$, with $\boldsymbol{\theta}$ as given below (2), may be expressed as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}), \tag{9}$$

where $\ell(\boldsymbol{\theta}) = -n\log(2\pi) - \frac{n}{2}\log(1-\rho^2) + \sum_{i=1}^{n} \log\big(A(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})\big) - \frac{1}{2}\sum_{i=1}^{n} d(\boldsymbol{t}_i)$, with $d(\boldsymbol{t}_i)$ being as given in (2). By deriving $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and ρ, the score function for $\boldsymbol{\theta}$ is expressed as $U(\boldsymbol{\theta}) = (U_{\boldsymbol{\alpha}}^{\top}(\boldsymbol{\theta}), U_{\boldsymbol{\beta}}^{\top}(\boldsymbol{\theta}), U_{\rho}(\boldsymbol{\theta}))^{\top}$, with

$$\begin{aligned} U_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) &= D^{-1}(\boldsymbol{\alpha})\Big(-n\boldsymbol{1}_2 + \sum_{i=1}^{n} D\big(a(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})\big)\boldsymbol{\Sigma}^{-1} a(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})\Big), \\ U_{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= -\tfrac{1}{2} D^{-1}(\boldsymbol{\beta}) \sum_{i=1}^{n} B(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\Sigma}^{-1} a(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}), \\ U_{\rho}(\boldsymbol{\theta}) &= \tfrac{n\rho}{1-\rho^2} - \tfrac{1}{2} \sum_{i=1}^{n} a^{\top}(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\Sigma}_{\rho}^{-1} a(\boldsymbol{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}), \end{aligned}$$

where $B(\boldsymbol{t}_i;\boldsymbol{\alpha},\boldsymbol{\beta}) = D^{-1}(\boldsymbol{t}_i+\boldsymbol{\beta})(\boldsymbol{t}_i-\boldsymbol{\beta}) - D^{-1/2}(\boldsymbol{\beta})D^{-1}(\boldsymbol{\alpha})D^{-1/2}(\boldsymbol{t}_i)D(\boldsymbol{t}_i+\boldsymbol{\beta})$ and $\boldsymbol{\Sigma}_\rho^{-1} = \partial\boldsymbol{\Sigma}^{-1}/\partial\rho = -\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_\rho\boldsymbol{\Sigma}^{-1}$. The ML estimate of $\boldsymbol{\theta}$ is obtained from the log-likelihood equations; see Kundu et al (2010). Next, we present the expected Fisher information matrix, which is required for the implementation of the score and Wald statistics. We obtain an explicit expression for this matrix of the BS_2 distribution and its inverse, being obtained by using the results provided in (5) and Lemma 1, and given by

$$\boldsymbol{I}_F(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \boldsymbol{0} & \boldsymbol{I}_{\boldsymbol{\alpha}\rho} \\ \boldsymbol{0} & \boldsymbol{I}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{0} \\ \boldsymbol{I}_{\rho\boldsymbol{\alpha}} & \boldsymbol{0} & 1_{\rho\rho} \end{pmatrix}, \tag{10}$$

where

$$\begin{aligned} \boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} &= \tfrac{2}{1-\rho^2}D^{-2}(\boldsymbol{\alpha}) - \tfrac{\rho^2}{1-\rho^2}D^{-1}(\boldsymbol{\alpha})\boldsymbol{1}_2\boldsymbol{1}_2^\top D^{-1}(\boldsymbol{\alpha}), \\ \boldsymbol{I}_{\boldsymbol{\alpha}\rho} &= -\tfrac{\rho}{1-\rho^2}D^{-1}(\boldsymbol{\alpha})\boldsymbol{1}_2, \\ 1_{\rho\rho} &= \tfrac{1+\rho^2}{(1-\rho^2)^2}, \\ \boldsymbol{I}_{\boldsymbol{\beta}\boldsymbol{\beta}} &= D^{-2}(\boldsymbol{\beta})\left(\tfrac{1}{1-\rho^2}\left(\tfrac{\rho^2}{4}\boldsymbol{I}_2 + D^{-2}(\boldsymbol{\alpha})\right) + D(\boldsymbol{\alpha}_J)\right) \\ &\quad -\tfrac{\rho}{2(1-\rho^2)\alpha_1\alpha_2\beta_1\beta_2}\big(\psi_1(\boldsymbol{\alpha},\rho)+\psi_1(\boldsymbol{\alpha},-\rho)\big)(\boldsymbol{1}_2\boldsymbol{1}_2^\top - \boldsymbol{I}_2), \end{aligned}$$

with $\psi_1(\boldsymbol{\alpha},\rho)$ and $\psi_1(\boldsymbol{\alpha},-\rho)$ being as given in (4) and (5), respectively, and $\boldsymbol{\alpha}_J = (J(\alpha_1), J(\alpha_2))^\top$, with $J(\alpha_j) = \mathrm{E}[T_j/\beta_j+1]^{-2}$, for $j=1,2$; see Kundu et al (2010). To obtain the elements $\boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}$, $\boldsymbol{I}_{\boldsymbol{\alpha}\rho}$ and $1_{\rho\rho}$ of $\boldsymbol{I}_F(\boldsymbol{\theta})$ given in (10), we use $\mathrm{E}[a(T_1,1,\beta_1)a(T_2,1,\beta_2)] = \alpha_1\alpha_2\rho$ given in Lemma 1, whereas $\boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}$ is reached by using (5). Then, the inverse Fisher information matrix can be written as

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \boldsymbol{I}_F^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\Omega}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \boldsymbol{0} & \boldsymbol{\Omega}_{\boldsymbol{\alpha}\rho} \\ \boldsymbol{0} & \boldsymbol{\Omega}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{0} \\ \boldsymbol{\Omega}_{\rho\boldsymbol{\alpha}} & \boldsymbol{0} & w_{\rho\rho} \end{pmatrix}, \tag{11}$$

where

$$\begin{aligned} \boldsymbol{\Omega}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} &= \left(\boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} - \frac{1}{1_{\rho\rho}}\boldsymbol{I}_{\boldsymbol{\alpha}\rho}\boldsymbol{I}_{\rho\boldsymbol{\alpha}}\right)^{-1}, \\ \boldsymbol{\Omega}_{\boldsymbol{\alpha}\rho} &= -\frac{1}{1_{\rho\rho}}\left(\boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} - \frac{1}{1_{\rho\rho}}\boldsymbol{I}_{\boldsymbol{\alpha}\rho}\boldsymbol{I}_{\rho\boldsymbol{\alpha}}\right)^{-1}\boldsymbol{I}_{\boldsymbol{\alpha}\rho}, \\ w_{\rho\rho} &= \frac{1}{1_{\rho\rho} - \boldsymbol{I}_{\rho\boldsymbol{\alpha}}\boldsymbol{I}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}^{-1}\boldsymbol{I}_{\boldsymbol{\alpha}\rho}}. \end{aligned}$$

Now, we obtain the ML estimates of the parameters of the BS_2 distribution considering the following restrictions imposed by the null hypotheses:

H_{01}: $\alpha_1 = \alpha_2 = \alpha$;

H_{02}: $\beta_1 = \beta_2 = \beta$;

H_{03}: $\alpha_1 = \alpha_2 = \alpha$, $\beta_1 = \beta_2 = \beta$;

H_{04}: $\rho = \rho_0$.

Let $\boldsymbol{T}_1 = (T_{11}, T_{21})^\top, \ldots, \boldsymbol{T}_n = (T_{1n}, T_{2n})^\top$ be a sample of size n from $\boldsymbol{T} \sim \mathrm{BS}_2(\boldsymbol{\alpha},\boldsymbol{\beta},\rho)$ and $\ell(\boldsymbol{\theta})$ be the log-likelihood function for the parameter vector $\boldsymbol{\theta}$ as given in (9). Let $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\alpha}}^\top, \widetilde{\boldsymbol{\beta}}^\top, \widetilde{\rho})^\top$ be the ML estimate of $\boldsymbol{\theta}$ under the null hypotheses. Then, *Case 1.* Under H_{01},

$$\begin{aligned}
\widetilde{\alpha}^2 &= \frac{1}{2n}\sum_{i=1}^{n}\left(a^2(t_{1i};1,\widetilde{\beta}_1)+a^2(t_{2i};1,\widetilde{\beta}_2)\right),\\
\widetilde{\rho} &= \frac{2\sum_{i=1}^{n} a(t_{1i};1,\widetilde{\beta}_1)a(t_{2i};1,\widetilde{\beta}_2)}{\sum_{i=1}^{n}(a^2(t_{1i};1,\widetilde{\beta}_1)+a^2(t_{2i};1,\widetilde{\beta}_2))},\\
\widetilde{\boldsymbol{\beta}} &= \arg\max_{\boldsymbol{\beta}}\ \ell(\widetilde{\alpha}\,\mathbf{1}_2^\top,\boldsymbol{\beta},\widetilde{\rho}).
\end{aligned}$$

Case 2. Under H_{02},

$$\begin{aligned}
\widetilde{\alpha}_j^2 &= \frac{1}{n}\sum_{i=1}^{n} a^2_{\boldsymbol{t}_{ji}}(1,\widetilde{\beta}), \quad j=1,2,\\
\widetilde{\rho} &= \frac{\sum_{i=1}^{n} a(t_{1i};1,\widetilde{\beta})a(t_{2i};1,\widetilde{\beta})}{\sqrt{\sum_{i=1}^{n}(a^2(t_{1i};1,\widetilde{\beta})}\sqrt{\sum_{i=1}^{n}a^2(t_{2i};1,\widetilde{\beta}))}},\\
\widetilde{\beta} &= \arg\max_{\beta}\ \ell(\widetilde{\boldsymbol{\alpha}},\beta\,\mathbf{1}_2,\widetilde{\rho}).
\end{aligned}$$

Case 3. Under H_{03},

$$\begin{aligned}
\widetilde{\alpha}^2 &= \frac{1}{2n}\sum_{i=1}^{n}\left(a^2(t_{1i};1,\widetilde{\beta})+a^2(t_{2i};1,\widetilde{\beta})\right),\\
\widetilde{\rho} &= \frac{2\sum_{i=1}^{n} a(t_{1i};1,\widetilde{\beta})a(t_{2i};1,\widetilde{\beta})}{\sum_{i=1}^{n}(a^2(t_{1i};1,\widetilde{\beta})+a^2(t_{2i};1,\widetilde{\beta}))},\\
\widetilde{\beta} &= \arg\max_{\beta}\ \ell(\widetilde{\alpha}\,\mathbf{1}_2,\beta\,\mathbf{1}_2,\widetilde{\rho}).
\end{aligned}$$

Case 4. Under H_{04},

$$\begin{aligned}
\widetilde{\alpha_j}^2 &= \frac{1}{n}\sum_{i=1}^{n} a^2(t_{ji};1,\widetilde{\beta}_j), \quad j=1,2,\\
\widetilde{\boldsymbol{\beta}} &= \arg\max_{\boldsymbol{\beta}}\ \ell(\widetilde{\boldsymbol{\alpha}},\boldsymbol{\beta},\rho_0).
\end{aligned}$$

In all of Cases 1-4 mentioned above, to estimate $\boldsymbol{\beta}$, it is necessary to use a numerical iterative procedure, but the ML estimates of $\boldsymbol{\alpha}$ and ρ have closed form expressions.

4 Hypotheses testing

In this section, we use the LR, score and Wald statistics to test H_{0k}, for $k = 1,2,3,4$, given in Section 3. The exact distributions of these statistics are unknown, so that asymptotic distributions are considered. Due to the relationship between the bivariate BS and normal distributions, the corresponding likelihood function satisfies standard regularity conditions; see Cox & Hinkley (1974). Thus, all of these statistics have a asymptotic χ^2 distribution with r_k degrees of freedom, under the null hypothesis, denoted by $\chi^2_{r_k}$. Here, $\widehat{\boldsymbol{\theta}}$ denotes the ML estimator of $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top,\boldsymbol{\beta}^\top,\rho)$ under the unrestricted model and $\widehat{\boldsymbol{\theta}}_k$ is the ML estimator for the restricted model under H_{0k}.

4.1 Likelihood ratio statistic

For each hypothesis H_{0k}, the LR statistic is expressed as

$$\mathrm{LR}_{0k} = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}}_k)) = n \log\left(\tfrac{1-\widetilde{\rho}_k^2}{1-\widehat{\rho}^2}\right) - \sum_{i=1}^{n} \left(\widehat{d}(\boldsymbol{t}_i) - \widetilde{d}_k(\boldsymbol{t}_i)\right) + 2\sum_{i=1}^{n} \log\left(\tfrac{A(\boldsymbol{t}_i;\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\beta}})}{A(\boldsymbol{t}_i;\widetilde{\boldsymbol{\alpha}}_k,\widetilde{\boldsymbol{\beta}}_k)}\right), \quad (12)$$

where $\widehat{d}(\boldsymbol{t}_i) = d(\boldsymbol{t}_i, \widehat{\boldsymbol{\theta}})$ and $\widetilde{d}_k(\boldsymbol{t}_i) = d(\boldsymbol{t}_i, \widetilde{\boldsymbol{\theta}}_k)$. Note that the two first terms of the right side of (12) are analogous to the bivariate normal case.

4.2 Score statistic

The score statistic for each hypothesis H_{0k} is given by

$$\mathrm{S}_{0k} = \tfrac{1}{n}\widetilde{\boldsymbol{U}}_k^\top \boldsymbol{I}_F^{-1}(\widetilde{\boldsymbol{\theta}}_k)\widetilde{\boldsymbol{U}}_k = \tfrac{1}{n}\widetilde{\boldsymbol{U}}_k^\top \widetilde{\boldsymbol{\Omega}}_k \widetilde{\boldsymbol{U}}_k, \quad k = 1,2,3,4, \quad (13)$$

where $\widetilde{\boldsymbol{\Omega}}_k = \boldsymbol{\Omega}(\widetilde{\boldsymbol{\theta}}_k)$ and $\widetilde{\boldsymbol{U}}_k = \boldsymbol{U}(\widetilde{\boldsymbol{\theta}}_k) = (\widetilde{U}_{\boldsymbol{\alpha},k}^\top, \widetilde{U}_{\boldsymbol{\beta},k}^\top, \widetilde{U}_{\rho,k})^\top$. By considering the structure of the inverse Fisher information matrix given in (11), we can express the statistic given in (13) as

$$\mathrm{S}_{0k} = \tfrac{1}{n}\left(\widetilde{U}_{\rho,k}^2\, \widetilde{w}_{\rho\rho,k} + 2\widetilde{U}_{\rho,k}\widetilde{\boldsymbol{\Omega}}_{\rho\boldsymbol{\alpha},k}\widetilde{U}_{\boldsymbol{\alpha},k} + \widetilde{U}_{\boldsymbol{\alpha},k}^\top\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\alpha}\boldsymbol{\alpha},k}\widetilde{U}_{\boldsymbol{\alpha},k} + \widetilde{U}_{\boldsymbol{\beta},k}^\top\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\beta}\boldsymbol{\beta},k}\widetilde{U}_{\boldsymbol{\beta},k}\right), \quad k = 1,2,3,4.$$

For the score statistic, we only need to determine the estimated values under H_{0k}, and in some situations is easier to obtain the parameter estimates under H_{0k} than under the unrestricted model. Other hypotheses of interest are H_{05}: $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$ and H_{06}: $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, and in these cases the score statistics are $\mathrm{S}_{05} = \frac{1}{n}\widetilde{U}_{\boldsymbol{\alpha},5}^\top\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\alpha}\boldsymbol{\alpha},5}\widetilde{U}_{\boldsymbol{\alpha},5}$ and $\mathrm{S}_{06} = \frac{1}{n}\widetilde{U}_{\boldsymbol{\beta},6}^\top\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{\beta}\boldsymbol{\beta},6}\widetilde{U}_{\boldsymbol{\beta},6}$, respectively.

4.3 Wald statistics

Note that the four hypothesis of interest may be rewritten as H_{0k}: $\boldsymbol{A\theta} = \boldsymbol{q}_{0k}$, for $k = 1,2,3,4$, where $\boldsymbol{A}_k$ and $\boldsymbol{q}_{0k}$ are adequate matrices. Specifically, for the indicated hypothesis, notice that:

(H_{01}) $\boldsymbol{A}_1 = (1,-1,0,0,0)$ and $q_{01} = 0$;

(H_{02}) $\boldsymbol{A}_2 = (0,0,1,-1,0)$ and $q_{02} = 0$;

(H_{03}) $\boldsymbol{A}_3 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}$ and $\boldsymbol{q}_{03} = \boldsymbol{0}_2$;

(H_{04}) $\boldsymbol{A}_4 = (0,0,0,0,1)$ and $q_{04} = \rho_0$.

Thus, the Wald statistic for testing H_{0k} is given by

$$\mathrm{W}_{0k} = n\big(\boldsymbol{A}_k\widehat{\boldsymbol{\theta}}\big)^\top\big(\boldsymbol{A}_k\boldsymbol{I}_F^{-1}(\widehat{\boldsymbol{\theta}})\boldsymbol{A}_k^\top\big)^{-1}\big(\boldsymbol{A}_k\widehat{\boldsymbol{\theta}}\big), \quad k = 1,2,3,4,$$

where $\boldsymbol{I}_F^{-1}(\boldsymbol{\theta}) = \boldsymbol{\Omega}$ is the inverse of the expected Fisher information matrix given in (10). Hence, the Wald statistics to test each H_{0k} can be respectively expressed as $\mathrm{W}_{01} = n(\alpha_1 - \alpha_2)^2(\boldsymbol{a}^\top\widehat{\boldsymbol{\Omega}}_{\alpha\alpha}\boldsymbol{a})^{-1}$, $\mathrm{W}_{02} = n(\beta_1 - \beta_2)^2(\boldsymbol{a}^\top\widehat{\boldsymbol{\Omega}}_{\beta\beta}\boldsymbol{a})^{-1}$, $\mathrm{W}_{03} = \mathrm{W}_{01} + \mathrm{W}_{02}$ and $\mathrm{W}_{04} = n(\widehat{\rho} - \rho_0)^2(\widehat{w}_{\rho\rho})^{-1}$, where $\boldsymbol{a} = (1,-1)^\top$.

4.4 Critical region

Each of the statistics LR_{0k}, S_{0k} and W_{0k}, for $k = 1, 2, 3, 4$, have an asymptotic $\chi^2_{r_k}$ distribution under the null hypothesis. Therefore, we reject H_{0k} with a significance level α, if LR_{0k}, S_{0k} and W_{0k} are greater than $\chi^2_{r_k,1-\alpha}$, where $\chi^2_{r_k,1-\alpha}$ is the $100(1-\alpha)$th percentile of the chi-squared distribution with r_k degrees of freedom, for $k = 1, 2, 3, 4$.

5 Numerical applications

In this section, we provide the numerical applications of our work. First, a simulation study is conducted to evaluate the performance of the LR, score and Wald statistics to test the hypotheses H_{0k}. Second, an illustration with a real-world data set is presented to apply some of the obtained results.

5.1 Simulated data

The simulation study is conducted considering all the four hypotheses to compare the three asymptotic test statistics for different sample sizes and parameter values.

By using 1000 MC replications, we determinate the empirical significance level of the tests based on the three statistics under null hypotheses for the BS_2 distribution.

The scenario for the parameter ρ in the simulation study is $\rho \in \{-0.9, -0.5, 0.5, -0.9\}$ under the null hypotheses:

H_{01}: $\alpha_1 = \alpha_2 = 1$, assuming $\beta_1 = 1, \beta_2 = 2$,

H_{02}: $\beta_1 = \beta_2 = 2$, assuming $\alpha_1 = 1, \alpha_2 = 1.5$,

H_{03}: $\alpha_1 = \alpha_2 = 1, \beta_1 = \beta_2 = 2$, and

H_{04}: $\rho = 0$, assuming $\alpha_1 = 0.5, \alpha_2 = 2, \beta_1 = 1$ and $\beta_2 = 2$.

For these scenarios, we study the behavior of the statistics LR_{0k}, S_{0k} and W_{0k}, for $k = 1, 2, 3, 4$, with sample size $n \in \{25, 50, 150\}$.

Considering the nominal significance level $\alpha \in \{5\%, 10\%\}$, we obtain the empirical significance levels for the statistics under analysis, which are displayed in Tables 1–4. From these tables, note that when we change the value of the parameter ρ in the null hypotheses H_{0k}, $k = 1, 2, 3$, keeping all the other parameters fixed, as well as the sample size and the different hypotheses, we obtain the same result based on the nominal significance levels, implying that the value of the parameter ρ seems not to change the behavior of the test statistics in the cases H_{0k}, $k = 1, 2, 3$. Now, when we look at the nominal significance levels under the null hypothesis H_{04}: $\rho = 0$, the behavior of the test statistics is similar under H_{0k}, for $k = 1, 2, 3$. In the case of moderate sample sizes, notice that the Wald statistic seems to be liberal, and the score statistic seems to be better than the other two statistics. Therefore, for the BS_2 distribution, we recommend to use the score statistic, because it presents the best behaviour in terms of significance level to test the hypotheses H_{0k}.

Table 1: empirical significance level of the test based on LR, score and Wald statistics for $\rho = 0.5$.

		H_{01}: $\alpha_1 = \alpha_2$			H_{02}: $\beta_1 = \beta_2$			H_{03}: $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$			H_{04}: $\rho = 0$		
α	n	W_{01}	S_{01}	LR_{01}	W_{02}	S_{02}	LR_{02}	W_{03}	S_{03}	LR_{03}	W_{04}	S_{04}	LR_{04}
	25	7.2	5.8	6.6	7.3	5.9	6.2	7.7	5.8	6.6	6.3	5.4	6.0
5%	50	6.4	5.6	6	5.7	5.4	5.4	5.9	5.3	5.4	5.7	5.3	5.3
	150	5.2	5	5.1	5	5.1	5.1	5.2	4.9	5.1	4.9	5	4.9
	25	12.5	10.8	11.7	13.1	10.8	11.9	11.8	10.9	11.6	11.6	10.7	11.2
10%	50	11	10.5	10.8	12.6	10.5	11	11.5	10.6	10.9	11	10.4	10.7
	150	10.3	10	9.9	10.1	10	10.1	10	10.1	10.2	10.1	10	10

Table 2: empirical significance level of the test based on LR, score and Wald statistics for $\rho = -0.5$

		H_{01}: $\alpha_1 = \alpha_2$			H_{02}: $\beta_1 = \beta_2$			H_{03}: $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$		
α	n	W_{01}	S_{01}	LR_{01}	W_{02}	S_{02}	LR_{02}	W_{03}	S_{03}	LR_{03}
	25	7.1	5.9	6.9	6.7	5.8	6.3	7.5	5.8	6.7
5%	50	6.7	5.7	5.8	5.7	5.6	5.5	5.7	5.6	5.5
	150	5.2	4.8	5.2	4.8	4.9	4.8	5.1	5	5
	25	12.4	10.9	11.7	12	10.8	11.7	12.5	10.9	11.6
10%	50	11.5	10.6	10.8	10.9	10.4	10.6	11.6	10.4	10.5
	150	10.3	10	10.1	10.2	10	10.1	10	10	10

Table 3: empirical significance level of the test based on LR, score and Wald statistics for $\rho = 0.9$

		H_{01}: $\alpha_1 = \alpha_2$			H_{02}: $\beta_1 = \beta_2$			H_{03}: $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$		
α	n	W_{01}	S_{01}	LR_{01}	W_{02}	S_{02}	LR_{02}	W_{03}	S_{03}	LR_{03}
	25	8	5.7	7.7	8	5.8	6.8	7.6	5.9	7.1
5%	50	6.2	5.4	6	5.4	5.5	5.7	6.4	5.5	6.1
	150	5.2	5.1	5.2	5.1	5	5.2	5.2	5.1	5.2
	25	12	10.7	11.6	12.9	10.7	12.4	11.8	10.8	11.3
10%	50	11.5	10.5	10.7	11.9	10.4	11.6	11.2	10.5	10.9
	150	10.5	10.1	9.8	10.1	10	10.1	10.2	10.2	10.1

Table 4: empirical significance level of the test based on LR, score and Wald statistics for $\rho = -0.9$

		H_{01}: $\alpha_1 = \alpha_2$			H_{02}: $\beta_1 = \beta_2$			H_{03}: $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$		
α	n	W_{01}	S_{01}	LR_{01}	W_{02}	S_{02}	LR_{02}	W_{03}	S_{03}	LR_{03}
	25	7.1	5.5	6.8	6.5	5.7	6.5	7.3	5.7	6.9
5%	50	6	5.3	5.9	6	5.2	5.5	6.2	5.4	5.8
	150	5.1	5	5.1	5.1	5	4.8	5	5.1	5.2
	25	13.2	10.6	13.1	13.6	10.7	13.5	12.4	10.7	12.2
10%	50	11.8	10.3	11.1	11.5	10.5	11	11.6	10.4	11.3
	150	10.4	10	10.2	10.3	10	10.2	10.1	10	10.1

5.2 Real-world data

An analysis of real-world data is carried out to illustrate the results obtained for the BS_2 distribution. We apply some of these results to the data set presented in Johnson & Wichern (1999), corresponding to the bone mineral density (BMD, in g/cm2) measured before (T_1) and after (T_2) an experimental study in 24 individuals. These bivariate data (t_1, t_2) are: (1.103, 1.027), (0.842, 0.857), (0.925, 0.875), (0.857, 0.873), (0.795, 0.811), (0.787, 0.640), (0.933, 0.947), (0.799, 0.886), (0.945, 0.991), (0.921, 0.977), (0.792, 0.825), (0.815, 0.851), (0.755, 0.770), (0.880, 0.912), (0.900, 0.905), (0.764, 0.756), (0.733, 0.765), (0.932, 0.932), (0.856, 0.843), (0.890, 0.879), (0.688, 0.673), (0.940, 0.949), (0.493, 0.463), (0.835, 0.776).

In order to justify the use of a bivariate distribution, we calculate the sample correlation coefficient between T_1 and T_2, which is $r = 0.92$. Now, how can we justify the use of the BS_2 distribution for modeling these BMD data? First, we compute univariate descriptive statistics for T_1 and T_2 including the standard deviation (SD) and coefficients of variation (CV) and kurtosis (CK), which are reported in Table 5. Second, such as Kundu et al (2010) did, we employ the total time on test (TTT) plot for showing that the data present a univariate hazard rate (HR) similar to that from the univariate BS distribution. Specifically, the HR of a RV T is $h(t) = f(t)/(1 - F(t))$, where $f(\cdot)$ and $F(\cdot)$ are the PDF and CDF of T, respectively. One manner to characterize the HR is by the scaled TTT function. Then, by using it, we can detect the type of HR that the data have. The TTT function is given by $W(u) = H^{-1}(u)/H^{-1}(1)$, for $0 \leq u \leq 1$, where $H^{-1}(u) = \int_0^{F^{-1}(u)}(1 - F(y))\mathrm{d}y$, with $F^{-1}(\cdot)$ being the inverse CDF of T. By plotting the points $(i/n, W_n(i/n))$, with $W_n(i/n) = (\sum_{k=1}^{i} t_{(k)} + (n-i)t_{(i)})/\sum_{k=1}^{n} t_{(k)}$, for $i = 1, \ldots, n$, and $t_{(k)}$ being the ith observed order statistic, it is possible to approximate $W(\cdot)$; see Fig. 2 in Vilca et al (2010) for different theoretical shapes of the scaled TTT function, and more details in Kundu et al (2010). Therefore, based on the descriptive summary and TTT plot, we point out that the BMD data can be suited well with the BS_2 distribution.

Table 5: summary statistics for the indicated data set.

Data set	n	Min	Mean	Median	Max	SD	CV	CK
t_1	24	0.493	0.841	0.849	1.103	0.115	0.137	3.83
t_2	24	0.463	0.841	0.865	1.027	0.125	0.148	4.40

Assuming the BS_2 distribution as a suitable model to describe the BMD data, we estimate its β_1, β_2, α_1, α_2 and ρ parameters with the ML method. The ML estimates of these parameters, along with the corresponding estimated standard errors (SE) approximated by using the expected Fisher information matrix are reported in Table 6. From this table, note that the estimated asymptotic SEs are smaller than those obtained by using the observed Fisher information matrix considered by Kundu et al (2010). In addition, we also report in this table 95% CI based on the normal approximation of the ML estimators of α_j, given by $\widehat{\alpha}_j \pm 1.96\,\mathrm{SE}$, for $j = 1, 2$, which have smaller length than the CIs computed with the observed Fisher information matrix, due, as mentioned, to the use of the expected Fisher information matrix. Table 6 also reports the MM estimates of the parameters, their respective estimated asymptotic SEs and 95% CIs based on the normal approximation of the MM estimators. Note that these CIs have larger length than the CIs obtained by using the ML estimates, but they are easier determined because the MM estimates have a closed form. The results displayed in Table 6 for MM estimates are reached combining Theorems 1 and 2.

Table 6: ML and MM estimates, estimated SE and 95% CI for BMD bivariate data.

	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\rho}$
Estimate	0.1491	0.1674	0.8313	0.8292	0.9343
SE	0.0016	0.0018	0.0046	0.0051	0.0040
CI	(0.1459;0.1522)	(0.1638;0.1709)	(0.8223;0.8402)	(0.8192;0.8392)	(0.9265;0.9422)
	$\widehat{\alpha}_{1M}$	$\widehat{\alpha}_{2M}$	$\widehat{\beta}_{1M}$	$\widehat{\beta}_{2M}$	$\breve{\rho}$
Estimate	0.1491	0.1674	0.8313	0.8292	0.9343
SE	0.0066	0.0072	0.0252	0.0282	0.0259
CI	(0.1361;0.1621)	(0.1533;0.1814)	(0.7821;0.8810)	(0.7740;0.8847)	(0.8835;0.9852)

In order to check the fitting of the BS_2 distribution to the BMD data, we use an interesting property of the MD, that is, $d(\boldsymbol{t}_i) \sim \chi^2_2$, for $i = 1, \ldots, n$. Such distributional result enables us to check the model in practice. Thus, by using the ML estimate of $\boldsymbol{\theta}$ in $d(\boldsymbol{t}_i)$, for $i = 1, \ldots, n$, Figure 1 can be constructed, which shows simulated envelopes, whose lines represent the 5th, 50th and 95th percentiles of 100 simulated points for each value of the BMD data. At the right side of Figure 1, we display the scatterplot with the contours of the fitted PDF of the BS_2 distribution. From this figure, note that such a distribution provides a good fit to the considered data, but a heavy-tailed distribution maybe could be more appropriate, such as pointed out by Vilca et al (2014), because observation #23 appears as a possible atypical case. In fact, by using the measures suggested by Lee et al (2006), we detect changes in the ML estimates of the parameters $\boldsymbol{\alpha}$ and ρ when the observation #23 is deleted. This suggests that a heavy-tailed distribution could fit these data better.

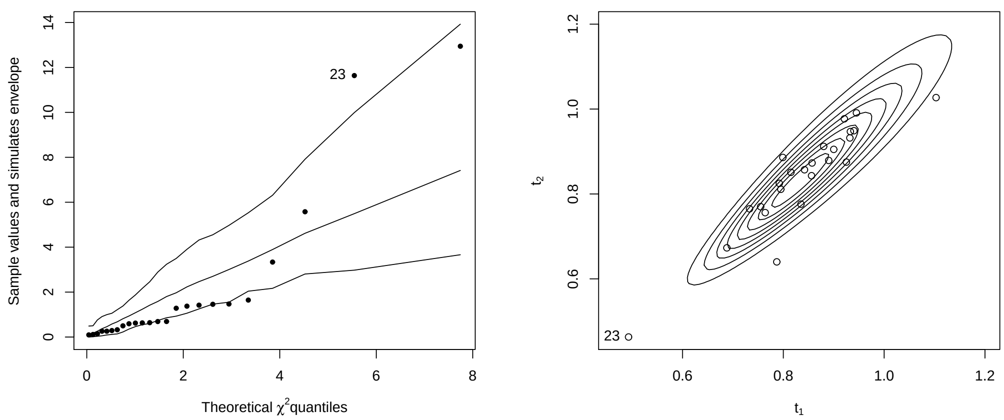

Figure 1: QQ plots and simulates envelope (left) and contour plot of the fitted PDF (right) for t_1 and t_2 data.

Table 7 presents the ML estimates of model parameters under H_{0k}. Note that the estimates of β_1 and β_2 do not show much difference under the four hypotheses, whereas the ML estimates of ρ under H_{01}, H_{02} and H_{03} are large and closed to one. Table 8 presents the LR, score and Wald statistic values and their p-values. Notice that the hypotheses H_{01}, H_{02} and H_{03} are not rejected at 5% level, whereas H_{04} is rejected at 5% level based on these three test statistics, indicating a need for considering the BS_2 distribution to model the BMD bivariate data.

Table 7: ML estimates of the BS_2 distribution parameters for BMD bivariate data.

Hypothesis	$\widetilde{\alpha}_1$	$\widetilde{\alpha}_2$	$\widetilde{\beta}_1$	$\widetilde{\beta}_2$	$\widetilde{\rho}$
H_{01}	0.1585	0.1585	0.8311	0.8291	0.9281
H_{02}	0.1491	0.1674	0.8319	0.8319	0.9342
H_{03}	0.1585	0.1585	0.8301	0.8301	0.9280
H_{04}	0.1491	0.1674	0.8316	0.8293	0

Table 8: Wald, LR and score statistics for BMD bivariate data.

Hypothesis	Wald	p-value	LR	p-value	Score	p-value
H_{01}	2.520	0.284	2.420	0.298	2.301	0.316
H_{02}	0.044	0.978	0.041	0.979	0.040	0.980
H_{03}	2.564	0.633	2.460	0.652	2.336	0.674
H_{04}	20.952	< 0.01	49.526	< 0.01	20.938	< 0.01

6 Concluding remarks

In this paper, due to the importance of the bivariate Birnbaum-Saunders distribution based on its good properties and usefulness for modeling different types of phenomena, we have investigated estimation and hypothesis testing in this distribution. About estimation, we have considered modified moment and maximum likelihood methods. We have proven that the modified moment estimators are consistent and asymptotically normal distributed. Regarding hypothesis testing, we have analyzed likelihood ratio, score and Wald statistics. The Wald statistic has resulted to be simpler to compute than the likelihood ratio and score statistics and it has been shown to be asymptotically chi-squared distributed. We have obtained the Fisher information in a matrix form, facilitating the implementation of the score and Wald statistics. We have conducted a simulation study to compare the three analyzed statistics and evaluated their performance. From this study, we can conclude that the score statistic has presented the best behaviour in terms of nominal level and power. We have applied the obtained results to a real-world data set to illustrate their use. Our study has provided new findings and improved the results proposed until now on this topic.

Acknowledgments This study was partially supported by a CNPq grant from Brazil and FONDECYT 1120879 grant from Chile.

References

Balakrishnan N, Leiva V, Sanhueza A, Cabrera E (2009a) Mixture inverse Gaussian distribution and its transformations, moments and applications. *Statistics* 43:91-104

Balakrishnan N, Leiva V, Sanhueza A, Vilca F (2009b) Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm. *Statistics and Operations Research Transactions* 33:171-192

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

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

Birnbaum ZW, Saunders SC (1969a) A new family of life distributions. *Journal of Applied Probability* 6:319-327

Birnbaum ZW, Saunders SC (1969b) Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* 6:328-347

Caro-Lopera FJ, Leiva V, Balakrishnan N (2012) Connection between the Hadamard and matrix products with an application to a matrix-variate Birnbaum-Saunders distribution. *Journal of Multivariate Analysis* 104:126-139

Cox DR, Hinkley DV (1974) *Theoretical Statistics*. Chapman & Hall:UK

Fang KT, Zhang YT (1990) *Generalized Multivariate Analysis*. Springer:Berlin

Ferreira M, Gomes MI, Leiva V (2012) On an extreme value version of the Birnbaum-Saunders distribution. *REVSTAT Statistical Journal* 10:181-210

Guiraud P, Leiva V, Fierro R (2009) A non-central version of the Birnbaum-Saunders distribution for reliability analysis. *IEEE Transactions on Reliability* 58:152-160

Johnson RA, Wichern DW (1999) *Applied Multivariate Statistical Analysis*. Pearson:US

Johnson NL, Kotz S, Balakrishnan N (1994) *Continuous Univariate Distributions*–Vols 1-2. Wiley:NY

Kotz S, Balakrishnan N, Johnson NL (2000) *Continuous Multivariate Distributions* –Vol 1. Wiley:NY

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

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

Kundu D, Balakrishnan N, Jamalizadeh A (2013) Multivariate generalized Birnbaum-Saunders distributions. *Journal of Multivariate Analysis* 116:230-244

Lee SY, Lu B, Song XY (2006) Assessing local influence for nonlinear structural equation models with ignorable missing data. *Computational Statistics and Data Analysis* 50:1356-1377

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

Leiva V, Barros M, Paula GA, Sanhueza A (2008) Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. *Environmetrics* 19:235-249

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

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

Leiva V, Santos-Neto M, Cysneiros FJA, Barros M (2014c) Birnbaum-Saunders statistical modelling: a new approach. *Statistical Modelling* 14:21-48

Leiva V, Saulo H, Leao J, Marchant C (2014d) A family of autoregressive conditional duration models applied to financial data. *Computational Statistics and Data Analysis* 79:175-191

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

Ng HK, Kundu D, Balakrishnan N (2003) Modified moment estimation for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis* 43:283-298

Paula GA, Leiva V, Barros M, Liu S (2012) Robust statistical modeling using the Birnbau-Saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry* 28:16-34

Sanchez L, Leiva V, Caro-Lopera FJ, Cysneiros FJA (2014) On matrix-variate Birnbaum-Saunders distributions and their estimation and application. *Brazilian Journal of Probability and Statistics* (in press).

Vilca F, Sanhueza A, Leiva V, Christakos G (2010) An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. *Stochastic Environmental Research and Risk Assessment* 24:771-782

Vilca F, Santana L, Leiva V, Balakrishnan N (2011) Estimation of extreme percentiles in Birnbaum-Saunders distributions. *Computational Statistics and Data Analysis* 55:1665-1678

Vilca F, Balakrishnan N, Zeller CB (2014) A robust extension of the bivariate Birnbaum-Saunders distribution and associated inference. *Journal of Multivariate Analysis* 124:418-435

Villegas C, Paula GA, Leiva V (2011) Birnbaum-Saunders mixed models for censored reliability data analysis. *IEEE Transactions on Reliability* 60:748-758