Testing of hypothesis of a block compound symmetric covariance matrix

Carlos A. Coelho[a,1,*], Anuradha Roy[b]

[a] *Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal*

[b] *Department of Management Science and Statistics, The University of Texas at San Antonio, San Antonio, Texas, U.S.A.*

Abstract

In this paper the authors study the problem of testing the hypothesis of a block compound symmetry covariance matrix with two-level multivariate observations, taken for m variables over u sites or time points. Through the decomposition of the main hypothesis into two sub-hypotheses, the likelihood ratio test statistic is easily obtained as well as its exact moments. Its exact distribution is then analyzed. Because this distribution is quite elaborate, yielding a non-manageable distribution function, a manageable but very precise near-exact distribution is developed. Numerical studies conducted to evaluate the closeness between this near-exact distribution and the exact distribution show the very good performance of this approximation even for very small sample sizes. A real data example is studied and a simulation is also conducted.

Keywords: characteristic function, composition of hypothesis, distribution of likelihood ratio statistics, near-exact distributions, product of independent Beta random variables, sum of independent Gamma random variables.

1. Introduction

We say that a covariance matrix has a BCS (block compound symmetry) structure (Rao, 1945, 1953) if it can be written as

$$\begin{aligned}\boldsymbol{\Theta} &= \begin{bmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \dots & \boldsymbol{\Sigma}_1 \\ \vdots & & \ddots & \vdots \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_0 \end{bmatrix} \\ &= \mathbf{I}_u \otimes (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1) + \mathbf{J}_u \otimes \boldsymbol{\Sigma}_1, \end{aligned} \tag{1.1}$$

*Corresponding author: Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Departamento de Matemática, Quinta da Torre, 2829-516 Caparica; phone: +351 21 294 8388, Fax: +351 21 294 8391.

Email addresses: cmac@fct.unl.pt (Carlos A. Coelho), Anuradha.Roy@utsa.edu (Anuradha Roy)

January 23, 2013

where, for $u \geq 2$, $\mathbf{I}_u$ is the $u \times u$ identity matrix, $\mathbf{1}_u$ is a $u \times 1$ vector of ones, $\mathbf{J}_u = \mathbf{1}_u \mathbf{1}_u'$ and $\otimes$ represents the Kronecker product. We assume $\mathbf{\Sigma}_0$ to be a positive definite symmetric $m \times m$ matrix, and $\mathbf{\Sigma}_1$ to be a symmetric $m \times m$ matrix, subject to the constraints $-\frac{1}{u-1}\mathbf{\Sigma}_0 < \mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_1 < \mathbf{\Sigma}_0$, which mean that $\mathbf{\Sigma}_0 - \mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_0 + (u-1)\mathbf{\Sigma}_1$ are positive definite matrices, so that the $mu \times mu$ matrix $\mathbf{\Theta}$ is also positive definite (for a proof, see Lemma 2.1 in Roy and Leiva (2011)). The $m \times m$ diagonal blocks $\mathbf{\Sigma}_0$ in $\mathbf{\Theta}$ represent the variance-covariance matrix of the m response variables at any given site or time point, whereas the $m \times m$ off diagonal blocks $\mathbf{\Sigma}_1$ in $\mathbf{\Theta}$ represent the covariance matrix of the m response variables between any two sites or time points. We assume $\mathbf{\Sigma}_0$ is constant for all sites and time points. Also, $\mathbf{\Sigma}_1$ is the same for any two different sites or time points.

In this paper our goal is to develop a different approach for the l.r.t. (likelihood ratio test) of the BCS structure and to develop near-exact distributions for the associated statistic, in order to make this test easy to implement in practice, since its practical application has been hindered by the complexity of the exact distribution of its l.r.t. statistic.

The need to test for BCS structure arises in many situations, namely those in which it is assumed as a structure for the covariance matrices involved in further analyses such as in many biomedical and medical researches. One has to be very careful when assuming this structure for two-level multivariate data, since an incorrect assumption may result in wrong conclusions. Thus, testing the validity of this BCS structure is of vital importance before assuming it, for any statistical analysis.

2. Formulation of the hypothesis and the likelihood ratio test

Let us assume that $\boldsymbol{y} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$. We are interested in testing the hypothesis

$$H_0 : \mathbf{\Sigma} = \mathbf{\Theta}\,, \tag{2.1}$$

where $\mathbf{\Theta}$ is defined in (1.1).

In Lemma 3.1 in Roy and Fonseca (2012), it is shown that we may write

$$\mathbf{\Gamma\Theta\Gamma'} = \begin{bmatrix} \mathbf{\Delta}_2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{u-1} \otimes \mathbf{\Delta}_1 \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{\Delta}_1 &= \mathbf{\Sigma}_0 - \mathbf{\Sigma}_1, \\ \mathbf{\Delta}_2 &= \mathbf{\Sigma}_0 + (u-1)\,\mathbf{\Sigma}_1, \end{aligned}$$

and $\mathbf{\Gamma} = \underset{u\times u}{\boldsymbol{C}^{*\prime}} \otimes \boldsymbol{I}_m$, with $\boldsymbol{C}^*$ an orthogonal Helmert matrix whose first column is proportional to a vector of $\mathbf{1}$'s. We should note that $\mathbf{\Gamma}$ is not a function of either $\mathbf{\Sigma}_0$, nor $\mathbf{\Sigma}_1$.

Thus, to test H_0 in (2.1), is equivalent to test

$$H_0 : \mathbf{\Sigma}^* = \mathbf{\Omega} \tag{2.2}$$

where

$$\mathbf{\Sigma}^* = \mathbf{\Gamma\Sigma\Gamma}' \qquad \text{and} \qquad \mathbf{\Omega} = \mathbf{\Gamma\Theta\Gamma}'\,,$$

where we may split the null hypothesis in (2.2) as

$$H_0 \equiv H_{0b|a} \circ H_{0a}\,,$$

where '$\circ$' means 'after', and where

$$H_{0a} : \mathbf{\Sigma}^* = \text{block-diag}(\mathbf{\Sigma}_i^*,\, i = 1, \dots, u)\,, \tag{2.3}$$

is the hypothesis of independence of the u diagonal blocks of size $m{\times}m$ of $\mathbf{\Sigma}^*$,

$$\begin{array}{ll} H_{0b|a} : & \mathbf{\Sigma}_2^* = \cdots = \mathbf{\Sigma}_u^*\,, \\ & \text{assuming } H_{0a} \end{array} \tag{2.4}$$

is the null hypothesis corresponding to the test of equality of the $u-1$ covariance matrices $\mathbf{\Sigma}_2^*, \dots, \mathbf{\Sigma}_u^*$, assuming H_{0a}.

The l.r.t. statistic to test H_{0a} in (2.3) is, for a sample of size n, (Anderson, 2003, Sec. 9.2)

$$\Lambda_a = \left(\frac{|\boldsymbol{A}|}{\prod_{j=1}^{u} |\boldsymbol{A}_j|} \right)^{n/2},$$

where $\boldsymbol{A}$ is the m.l.e. (maximum likelihood estimator) of $\mathbf{\Sigma^*}$, and $\boldsymbol{A}_j$ its $m{\times}m$ j-th diagonal block.

The l.r.t. statistic to test $H_{0b|a}$ in (2.4) is (Anderson, 2003, Sec. 10.2)

$$\Lambda_b = \left((u-1)^{m(u-1)} \frac{\prod_{j=2}^{u} |\boldsymbol{A}_j|}{|\boldsymbol{A}^*|^{u-1}} \right)^{n/2}, \tag{2.5}$$

where

$$\boldsymbol{A}^* = \sum_{j=2}^{u} \boldsymbol{A}_j\,.$$

Then the l.r.t. statistic to test H_0 in (2.2) will be

$$\Lambda = \Lambda_a \Lambda_b = \left((u-1)^{m(u-1)} \frac{|\boldsymbol{A}|}{|\boldsymbol{A}_1||\boldsymbol{A}^*|^{u-1}} \right)^{n/2}, \tag{2.6}$$

with

$$E\left(\Lambda^h\right) = E\left(\Lambda_a^h\right) E\left(\Lambda_b^h\right), \tag{2.7}$$

since, under H_{0a}, Λ_a is independent of $\boldsymbol{A}_1, \dots, \boldsymbol{A}_u$ (Marques and Coelho, 2012; Coelho and Marques, 2012b), which makes Λ_a independent of Λ_b, given that this latter one is only function of $\boldsymbol{A}_2, \dots, \boldsymbol{A}_u$.

In (2.7) we have (Marques et al., 2011)

$$E\left(\Lambda_a^h\right) = \prod_{k=1}^{u-1}\prod_{j=1}^{m} \frac{\Gamma\left(\frac{n-j}{2}\right)\Gamma\left(\frac{n-(uv-k)m-j}{2}+\frac{n}{2}h\right)}{\Gamma\left(\frac{n-(uv-k)m-j}{2}\right)\Gamma\left(\frac{n-j}{2}+\frac{n}{2}h\right)}$$

$$= \underbrace{\left\{\prod_{j=3}^{mu}\left(\frac{n-j}{n}\right)^{r_j}\left(\frac{n-j}{n}+h\right)^{-r_j}\right\}}_{\Phi_{a,1}(h)} \underbrace{\left(\frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}+\frac{n}{2}h\right)}{\Gamma\left(\frac{n-1}{2}+\frac{n}{2}h\right)\Gamma\left(\frac{n-2}{2}\right)}\right)^{k^*}}_{\Phi_{a,2}(h)} \tag{2.8}$$

where

$$k^* = \begin{cases} \left\lfloor \frac{u}{2} \right\rfloor, & m \text{ odd} \\ 0, & m \text{ even,} \end{cases}$$

and

$$r_j = \begin{cases} h_{j-2} + (-1)^j k^*, & j = 3, 4 \\ r_{j-2} + h_{j-2}, & j = 5, \ldots, mu \end{cases} \tag{2.9}$$

with

$$h_j = \begin{cases} uv-1, & j = 1, \ldots, m \\ -1, & j = m+1, \ldots, mu-2\,, \end{cases} \tag{2.10}$$

while for Λ_b we have

$$E\left(\Lambda_b^h\right) = \prod_{j=1}^{m}\prod_{k=1}^{u-1} \frac{\Gamma\left(\frac{n-1}{2}-\frac{j-1}{2(u-1)}+\frac{k-1}{u-1}\right)\Gamma\left(\frac{n-j}{2}+\frac{n}{2}h\right)}{\Gamma\left(\frac{n-1}{2}-\frac{j-1}{2(u-1)}+\frac{k-1}{u-1}+\frac{n}{2}h\right)\Gamma\left(\frac{n-j}{2}\right)}$$

$$= \underbrace{\left\{\prod_{j=2}^{m}\left(\frac{n-j}{n}\right)^{s_j}\left(\frac{n-j}{n}+h\right)^{-s_j}\right\}}_{\Phi_{b,1}(h)}$$

$$\underbrace{\times\left\{\prod_{j=1}^{\lfloor m/2\rfloor}\prod_{k=1}^{u-1} \frac{\Gamma\left(n-1+\frac{k-2j}{u-1}\right)\Gamma\left(n+\left\lfloor\frac{k-2j}{u-1}-1\right\rfloor+nh\right)}{\Gamma\left(n-1+\frac{k-2j}{u-1}+nh\right)\Gamma\left(n+\left\lfloor\frac{k-2j}{u-1}-1\right\rfloor\right)}\right\} \times\left\{\prod_{k=1}^{u-1} \frac{\Gamma\left(\frac{n-m}{2}+\frac{m(u-1)-u-m+2k}{2(u-1)}\right)\Gamma\left(\frac{n-m}{2}+\left\lfloor\frac{m(u-1)-u-m+2k}{2(u-1)}\right\rfloor+\frac{n}{2}h\right)}{\Gamma\left(\frac{n-m}{2}+\frac{m(u-1)-u-m+2k}{2(u-1)}+\frac{n}{2}h\right)\Gamma\left(\frac{n-m}{2}+\left\lfloor\frac{m(u-1)-u-m+2k}{2(u-1)}\right\rfloor\right)}\right\}^{m \perp\!\!\!\perp 2}}_{\Phi_{b,2}(h)} \tag{2.11}$$

where s_j $(j = 2, \ldots, m)$ are given in Appendix A and where $m \perp\!\!\!\perp 2$ is the remainder of the integer division of m by 2.

Since the supports of Λ_a and Λ_b are delimited, their distributions are defined by their moments, and as such, from the first expression in (2.8) we may write

$$\Lambda_a \sim \prod_{k=1}^{u-1} \prod_{j=1}^{m} X_{jk}\,, \qquad \text{where} \qquad X_{jk} \sim Beta\left(\frac{n-(u-k)m-j}{2}, \frac{(u-k)m}{2}\right), \tag{2.12}$$

where X_{jk} $(j = 1, \ldots, m; k = 1, \ldots, u-1)$ are independent, while from the first expression in (2.11) we may write

$$\Lambda_b \sim \prod_{j=1}^{m} \prod_{k=1}^{u-1} X^*_{jk}\,, \qquad \text{where} \qquad X^*_{jk} \sim Beta\left(\frac{n-j}{2}, \frac{2k+(u-2)j-u}{2}\right), \tag{2.13}$$

where X^*_{jk} $(j = 1, \ldots, m; k = 1, \ldots, u-1)$ are independent so that we may say that

$$\Lambda \sim \prod_{j=1}^{m} \left\{ \left(\prod_{k=1}^{u-1} X_{jk} \right) \left(\prod_{k=}^{u-1} X^*_{jk} \right) \right\}, \tag{2.14}$$

where all random variables are independent.

On the other hand, based on the results in Appendix B and from the second expressions in (2.8) and (2.11) we may write, for Λ_a,

$$\Lambda_a \sim \left(\prod_{j=3}^{mu} e^{-Z_j} \right) \left(\prod_{j=1}^{k^*} W_j \right) \tag{2.15}$$

where

$$Z_j \sim \Gamma\left(r_j, \frac{n-j}{n}\right) \quad \text{and} \quad W_j \sim Beta\left(\frac{n-2}{2}, \frac{1}{2}\right)$$

are all independent r.v.'s (random variables), while for Λ_b we may write

$$\Lambda_b \sim \left(\prod_{j=2}^{m} e^{-Z^*_j} \right) \left(\prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{u-1} W^*_{1jk} \right) \left(\prod_{k=1}^{u-1} W^*_{2k} \right)^{m \perp\!\!\!\perp 2} \tag{2.16}$$

where

$$Z^*_j \sim \Gamma\left(v s_j, \frac{n-j}{n}\right), \quad W^*_{1jk} \sim Beta\left(n + \left\lfloor \frac{k-2j}{u-1} - 1 \right\rfloor, \frac{k-2j}{u-1} - \left\lfloor \frac{k-2j}{u-1} \right\rfloor\right),$$

and

$$W_{2k}^* \sim Beta\left(\frac{n-m}{2}+\left\lfloor\frac{m(u-1)-u-m+2k}{2(u-1)}\right\rfloor, \frac{m(u-1)-u-m+2k}{2(u-1)}-\left\lfloor\frac{m(u-1)-u-m+2k}{2(u-1)}\right\rfloor\right)$$

are all independent r.v.'s.

From (2.15) and (2.16) we may write,

$$\Lambda \sim \left(\prod_{j=2}^{mu} e^{-T_j}\right)\left(\prod_{j=1}^{k^*} W_j\right)\left(\prod_{j=1}^{\lfloor m/2\rfloor}\prod_{k=1}^{u-1} W_{1jk}^*\right)\left(\prod_{k=1}^{u-1} W_{2k}^*\right)^{m\perp\!\!\!\perp 2} \qquad (2.17)$$

where

$$T_j \sim \Gamma\left(\gamma_j, \frac{n-j}{n}\right), \quad (j=2,\ldots,mu)$$

with

$$\gamma_j = \sum_{j=2}^{mu}\left(r_j^+ + s_j^+\right) \qquad (2.18)$$

where

$$r_j^+ = \begin{cases} 0 & j=2 \\ r_j & j=3,\ldots,mu \end{cases} \qquad \text{and} \qquad s_j^+ = \begin{cases} s_j & j=2,\ldots,m \\ 0 & j=m+1,\ldots,mu \end{cases} \qquad (2.19)$$

where r_j are given by (2.9) and (2.10), s_j are given by (A.1)–(A.5) and all the other variables are defined as above.

The form of the distribution of Λ in (2.17), although it may look more complicated than the one in (2.14), is more useful for the development of near-exact distributions, as we will see in the next section.

3. The characteristic function of $W = -\log\Lambda$ and the near-exact approximation

From the developments in the previous section and the second parts of (2.8) and (2.11) we may write the c.f. (characteristic function) of $W = -\log\Lambda$ as

$$\begin{aligned} \Phi_W(t) &= E\left(e^{\mathrm{i}tW}\right) = E\left(\Lambda^{-\mathrm{i}t}\right) \\ &= \underbrace{\left\{\prod_{j=2}^{mu}\left(\frac{n-j}{n}\right)^{\gamma_j}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-\gamma_j}\right\}}_{\Phi_{W,1}(t)}\ \underbrace{\Phi_{a,2}(-\mathrm{i}t)\Phi_{b,2}(-\mathrm{i}t)}_{\Phi_{W,2}(t)} \end{aligned} \qquad (3.1)$$

where γ_j is given by (2.18) and $\Phi_{a,2}(\,\cdot\,)$ and $\Phi_{b,2}(\,\cdot\,)$ are defined in (2.8) and (2.11), and $\Phi_{W,1}(t)$ is actually equal to $\Phi_{a,1}(-\mathrm{i}t)\Phi_{b,1}(-\mathrm{i}t)$, being these two functions also defined in (2.8) and (2.11).

Then, in building the near-exact distributions we will keep $\Phi_{W,1}(t)$ untouched and we will approximate $\Phi_{W,2}(t)$ asymptotically by the c.f. of a finite mixture of Gamma distributions.

While $\Phi_{W,1}(t)$ is the c.f. of a GIG (Generalized Integer Gamma) distribution (Coelho, 1998) of depth $mu-1$, which is the distribution of the sum of mu independent Gamma distributed random variables, all with integer shape parameters, $\Phi_{W,1}(t)$ is the c.f. of the sum of $k^*+\lfloor m/2\rfloor(u-1)+(m \perp\!\!\!\perp 2)$ independent Logbeta distributed random variables. For $u=2$ and even m, $\Phi_{W,1}(t)$ yields indeed the exact c.f. for W, which means that in this case we have the exact p.d.f. and c.d.f. of W and Λ in a simple closed form, in the form of the p.d.f. and c.d.f. of a GIG distribution of depth $2m$, with shape parameters γ_j and rate parameters $(n-j)/n$ $(j=1,\dots,2m)$ for W, or the form of a p.d.f. or c.d.f. of an EGIG (Exponentiated Generalized Integer Gamma) distribution (Arnold et al., 2013) for Λ.

It will be based on the results in Sections 5 and 6 of Tricomi and Erdélyi (1951), which show that we can asymptotically approximate the c.f. of a Logbeta(a,b) distribution by the c.f. of an infinite mixture of $\Gamma(b+j,a)$ $(j=0,1,\dots)$ distributions, that we will replace $\Phi_{W,2}(t)$ by

$$\Phi_2(t)=\sum_{k=0}^{m^*}\pi_k\,\lambda^r(\lambda-\mathrm{i}t)^{-r}\,, \tag{3.2}$$

which is the c.f. of a finite mixture of Gamma distributions, all with the same rate parameter λ. In (3.2) we will take λ as the rate parameter in

$$\Phi^*(t)=\theta\lambda^{\tau_1}(\lambda-\mathrm{i}t)^{-\tau_1}+(1-\theta)\lambda^{\tau_2}(\lambda-\mathrm{i}t)^{-\tau_2}$$

where θ, λ, τ_1 and τ_2 are determined in such a way that

$$\left.\frac{\partial^h}{\partial t^h}\Phi^*(t)\right|_{t=0}=\left.\frac{\partial^h}{\partial t^h}\Phi_{W,2}(t)\right|_{t=0}\,,\qquad h=1,\dots,4\,,$$

and

$$\begin{aligned} r &= \frac{k^*}{2}+\sum_{j=1}^{\lfloor m/2\rfloor}\sum_{k=1}^{u-1}\frac{k-2j}{u-1}-\left\lfloor\frac{k-2j}{u-1}\right\rfloor+\sum_{k=1}^{u-1}\frac{m(u-1)-u-m+2k}{2(u-1)}-\left\lfloor\frac{m(u-1)-u-m+2k}{2(u-1)}\right\rfloor \\ &= \left\{\begin{array}{ll}\frac{m}{4}(u-2) & m \text{ even}\\ \left\lfloor\frac{u}{2}\right\rfloor+\frac{m+1}{4}(u-2) & m \text{ odd}\end{array}\right. \qquad (u\geq 2)\,,\end{aligned} \tag{3.3}$$

which is the sum of the second parameters of all the Beta r.v.'s in (2.17). Then the weights $\pi_0,\dots,\pi_{m^*-1}$ in (3.2) will be determined in such a way that

$$\left.\frac{\partial^h}{\partial t^h}\Phi_2(t)\right|_{t=0}=\left.\frac{\partial^h}{\partial t^h}\Phi_{W,2}(t)\right|_{t=0}\,,\qquad h=1,\dots,m^*\,,$$

and with $\pi_{m^*} = 1 - \sum_{k=0}^{m^*-1} \pi_k$.

The near-exact distributions built in this way will match m^* exact moments and will have c.f.

$$\Phi^*_W(t) = \Phi_{W,1}(t)\Phi_2(t)\,, \tag{3.4}$$

which, for non-integer r, is the c.f. of a finite mixture, with weights π_k ($k = 0, \ldots, m^*$), of GNIG (Generalized Near-Integer Gamma) distributions of depth mu, with integer shape parameters γ_j, given by (2.18) and (2.19) and non-integer shape parameter r given by (3.3) and corresponding rate parameters $(n-j)/n$ ($j = 2, \ldots, mu$) and λ (Coelho, 2004; Coelho and Marques, 2012a, Appendix 1). Using the notation in Appendix 1 in Coelho and Marques (2012a), these near-exact distributions will yield for $W = -\log\Lambda$ p.d.f.'s and c.d.f.'s of the form

$$f_W(w) = \sum_{k=0}^{m^*} \pi_k\, f^{GNIG}\left(w \mid \gamma_2, \ldots, \gamma_{mu}, r+k; \frac{n-2}{n}, \ldots, \frac{n-mu}{n}, \lambda; mu\right)\,, \quad (w > 0)$$

and

$$F_W(w) = \sum_{k=0}^{m^*} \pi_k\, F^{GNIG}\left(w \mid \gamma_2, \ldots, \gamma_{mu}, r+k; \frac{n-2}{n}, \ldots, \frac{n-mu}{n}, \lambda; mu\right)\,, \quad (w>0)\,,$$

while the near-exact p.d.f. and c.d.f. for Λ are respectively given by

$$f_\Lambda(z) = \sum_{k=0}^{m^*} \pi_k\, f^{GNIG}\left(-\log z \mid r_2^*, \ldots, r_p^*, r+k; \frac{n-2}{n}, \ldots, \frac{n-mu}{n}, \lambda; mu\right)\frac{1}{z}\,, \quad (0<z<1)$$

and

$$F_\Lambda(z) = \sum_{k=0}^{m^*} \pi_k \left(1 - F^{GNIG}\left(-\log z \,|\, r_2^*, \ldots, r_p^*, r+k; \frac{n-2}{n}, \ldots, \frac{n-mu}{n}, \lambda; mu\right)\right)\,, \quad (0<z<1)\,.$$

For integer r the above GNIG distributions of depth mu become GIG (Generalized Integer Gamma) distributions of depth mu (Coelho, 1998; Arnold et al., 2013, App. B), which have even simpler and more manageable expressions, and in this case the near-exact distributions for Λ will be mixtures of what Arnold et al. (2013) call EGIG distributions.

We should note that for $m = 1$, this test yields the equivariance-equicorrelation test in Wilks (1946).

4. Numerical studies

In order to assess the performance of the near-exact distributions developed, that is, their closeness to the corresponding exact distribution, we use the measure

$$\Delta = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|\frac{\Phi_W(t) - \Phi^*_W(t)}{t}\right| dt\,, \tag{4.1}$$

with

$$\max_{w>0} |F_W(w) - F_W^*(w)| = \max_{0<z<1} |F_\Lambda(z) - F_\Lambda^*(z)| \leq \Delta\,,$$

where $\Phi_W(t)$ is the exact c.f. of W in (3.1) and $\Phi_W^*(t)$ is the near-exact c.f. of W in (3.4) and $F_W(\,\cdot\,)$ and $F_W^*(\,\cdot\,)$ are the corresponding c.d.f.'s, that is, the exact and near-exact c.d.f. of W, being $F_\Lambda(\,\cdot\,)$ and $F_\Lambda^*(\,\cdot\,)$ the corresponding c.d.f.'s for Λ.

Table 1: Values of Δ for the chi-square and near-exact distributions, for different values of m, u and n

			m^*					m^*	
	χ^2	4	6	10		χ^2	4	6	10
n			$u=2, m=3$		n			$u=3, m=3$	
8	$7.35{\times}10^{-1}$	$6.61{\times}10^{-9}$	$1.12{\times}10^{-10}$	$1.57{\times}10^{-13}$	11	$9.35{\times}10^{-1}$	$2.64{\times}10^{-12}$	$4.88{\times}10^{-15}$	$1.10{\times}10^{-19}$
36	$4.80{\times}10^{-1}$	$3.86{\times}10^{-12}$	$9.48{\times}10^{-15}$	$2.71{\times}10^{-19}$	39	$2.27{\times}10^{-1}$	$2.34{\times}10^{-15}$	$4.50{\times}10^{-19}$	$1.85{\times}10^{-25}$
136	$3.01{\times}10^{-2}$	$5.04{\times}10^{-15}$	$1.38{\times}10^{-18}$	$5.76{\times}10^{-25}$	109	$7.74{\times}10^{-2}$	$4.42{\times}10^{-18}$	$1.05{\times}10^{-22}$	$6.74{\times}10^{-31}$
			$u=5, m=3$					$u=7, m=3$	
17	$1.15{\times}10^{0}$	$5.81{\times}10^{-13}$	$7.29{\times}10^{-17}$	$2.86{\times}10^{-24}$	23	$1.28{\times}10^{0}$	$4.31{\times}10^{-14}$	$1.59{\times}10^{-18}$	$5.32{\times}10^{-27}$
45	$4.67{\times}10^{-1}$	$2.36{\times}10^{-14}$	$8.10{\times}10^{-19}$	$1.73{\times}10^{-27}$	51	$7.02{\times}10^{-1}$	$8.03{\times}10^{-15}$	$1.43{\times}10^{-19}$	$8.80{\times}10^{-29}$
115	$1.77{\times}10^{-1}$	$2.26{\times}10^{-16}$	$1.29{\times}10^{-21}$	$6.84{\times}10^{-32}$	121	$3.06{\times}10^{-1}$	$1.14{\times}10^{-16}$	$5.17{\times}10^{-22}$	$1.27{\times}10^{-32}$
			$u=3, m=5$					$u=3, m=10$	
17	$1.17{\times}10^{0}$	$5.16{\times}10^{-15}$	$2.63{\times}10^{-18}$	$5.81{\times}10^{-25}$	32	$1.45{\times}10^{0}$	$1.25{\times}10^{-14}$	$4.17{\times}10^{-19}$	$1.15{\times}10^{-27}$
45	$4.75{\times}10^{-1}$	$5.02{\times}10^{-16}$	$3.47{\times}10^{-20}$	$5.50{\times}10^{-28}$	60	$9.82{\times}10^{-1}$	$3.85{\times}10^{-15}$	$9.25{\times}10^{-20}$	$1.14{\times}10^{-28}$
115	$1.80{\times}10^{-1}$	$6.36{\times}10^{-18}$	$5.70{\times}10^{-23}$	$2.13{\times}10^{-32}$	130	$5.49{\times}10^{-1}$	$5.66{\times}10^{-17}$	$3.42{\times}10^{-22}$	$2.62{\times}10^{-32}$
			$u=5, m=10$					$u=10, m=3$	
52	$1.62{\times}10^{0}$	$4.38{\times}10^{-18}$	$6.51{\times}10^{-24}$	$3.11{\times}10^{-35}$	32	$1.42{\times}10^{0}$	$1.53{\times}10^{-15}$	$1.36{\times}10^{-20}$	$1.75{\times}10^{-30}$
80	$1.29{\times}10^{0}$	$6.57{\times}10^{-18}$	$1.29{\times}10^{-23}$	$9.27{\times}10^{-35}$	60	$9.60{\times}10^{-1}$	$9.15{\times}10^{-16}$	$6.29{\times}10^{-21}$	$1.66{\times}10^{-31}$
150	$9.51{\times}10^{-1}$	$4.02{\times}10^{-19}$	$3.31{\times}10^{-25}$	$3.87{\times}10^{-37}$	130	$5.29{\times}10^{-1}$	$3.10{\times}10^{-17}$	$5.63{\times}10^{-23}$	$4.87{\times}10^{-35}$
			$u=10, m=5$					$u=10, m=10$	
52	$1.63{\times}10^{0}$	$9.64{\times}10^{-18}$	$1.26{\times}10^{-23}$	$1.85{\times}10^{-35}$	102	$1.91{\times}10^{0}$	$1.13{\times}10^{-21}$	$2.49{\times}10^{-29}$	$2.52{\times}10^{-42}$
80	$1.28{\times}10^{0}$	$1.88{\times}10^{-17}$	$2.62{\times}10^{-23}$	$5.87{\times}10^{-35}$	130	$1.68{\times}10^{0}$	$5.81{\times}10^{-21}$	$2.33{\times}10^{-28}$	$7.27{\times}10^{-43}$
150	$9.42{\times}10^{-1}$	$1.67{\times}10^{-18}$	$9.18{\times}10^{-25}$	$2.87{\times}10^{-37}$	200	$1.42{\times}10^{0}$	$1.99{\times}10^{-21}$	$5.16{\times}10^{-29}$	$6.64{\times}10^{-44}$
					300	$1.22{\times}10^{0}$	$4.06{\times}10^{-22}$	$5.54{\times}10^{-30}$	$1.96{\times}10^{-45}$
					400	$1.08{\times}10^{0}$	$1.16{\times}10^{-22}$	$9.54{\times}10^{-31}$	$1.23{\times}10^{-46}$

In Table 1 we may observe the values of Δ for the common chi-square approximation to the distribution of the logarithm of the l.r.t. statistic, which says that $-2\log\Lambda$ is asymptotically a chi-square with $mu(mu+1)/2 - m(m+1)$ degrees of freedom, and for the near-exact distributions developed in the previous section. In this table we may observe values of Δ for different values of u (number of locations or time points), m (number of variables) and n (sample size) and also for different values of m^*, the number of exact moments matched by the near-exact distributions.

As expected, as m^* increases the values of Δ for the near-exact distributions decrease clearly, showing an increasing closeness to the exact distribution. We may also see from Table 1 that the near-exact distributions developed exhibit a very good asymptotic behavior not only for increasing sample sizes, but also for increasing values of both u and m, which is a much desirable feature. For all values of u and m the upper bounds on the difference between the exact and the near-exact c.d.f. exhibit extremely low values. We may note that, for larger values of u and v, the asymptotic behavior for increasing n becomes visible only for larger values of n.

We may see from Table 1 that indeed the chi-square asymptotic distribution may only yield sensible approximations for very large sample sizes and small numbers of variables involved, and that the performance of this approximation worsens much as the number of variables increase, that is, as either u or m increase.

It happens that the measure Δ in (4.1) gives very good upper-bounds on the difference between the exact and the approximate distributions in case these approximations are rather good, and too large values in case these approximations are quite bad. This is the reason why we get some values of Δ above one for the chi-square approximation for the smaller sample sizes for many of the combinations of values of u and m, which indicates that in these cases this approximation has a really very very poor performance, since indeed the values of Δ should always be between zero and one.

5. A real data example and a simulation study

In this section we show the implementation of our new hypothesis testing procedure, using the block-diagonalization of the BCS structure, as a result of the application of Lemma 3.1 in Roy and Fonseca (2012), with a real data set taken from (Johnson and Wichern, 2007, p. 43). An investigator measured the mineral content of bones (radius, humerus and ulna) by photon absorptiometry to examine whether dietary supplements would slow bone loss in 25 older women. Measurements were recorded for the three bones on the dominant and non-dominant sides. As such, data have a two-level multivariate structure, with $u = 2$ and $m = 3$. Thus, we can think about testing the hypothesis that the population covariance matrix has a BCS covariance structure. We need to rearrange the variables in the data set by grouping together the mineral content of the dominant sides of radius, humerus and ulna as the first three variables, that is, the variables in the first location (($u = 1$) — dominant side) and then the mineral contents for the non-dominant side of the same bones (($u = 2$) — non-dominant side). The resulting MLE of $\mathbf{\Sigma}$ is (with five decimal places)

$$\widehat{\mathbf{\Sigma}} = \left[\begin{array}{ccc|ccc} 0.01248 & 0.02146 & 0.00876 & 0.00996 & 0.01928 & 0.00764 \\ 0.02146 & 0.07714 & 0.01616 & 0.01779 & 0.06411 & 0.01233 \\ 0.00876 & 0.01616 & 0.01111 & 0.00819 & 0.01703 & 0.00775 \\ \hline 0.00996 & 0.01779 & 0.00819 & 0.01096 & 0.02026 & 0.00855 \\ 0.01928 & 0.06411 & 0.01703 & 0.02026 & 0.06671 & 0.01612 \\ 0.00764 & 0.01233 & 0.00775 & 0.00855 & 0.01612 & 0.01018 \end{array}\right].$$

We see that the variance-covariance matrices of the three mineral contents for the dominant and non-dominant sides appear very similar. Also, the two covariance matrices (the two block off-diagonal matrices) of the three bones between the dominant and non-dominant sides seem to be fairly similar. Thus, we may hypothesize that this data have a BCS structure.

In order to carry out the test, according to (2.2) we will need to compute the matrix

$$\boldsymbol{A} = \widehat{\boldsymbol{\Sigma}}^* = \boldsymbol{\Gamma}\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Gamma}',$$

where

$$\boldsymbol{\Gamma} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \otimes \boldsymbol{I}_3 = \left[\begin{array}{ccc|ccc} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \hline \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \end{array}\right].$$

Then, from (2.6) we get our calculated value for Λ as 0.0227794. Using the near-exact distributions developed in Section 3, we obtain for this calculated value of Λ the p-values in Table 2.

Table 2: $p-$values from the near-exact approximations for different values of m^* (the number of exact moments matched) for the hypothesis test on bone mineral data

m^*	$p-$value
1	0.2792
2	0.2792168
4	0.2792168718
6	0.279216871862
10	0.279216871862222

Table 2 gives the $p-$values for different values of m^* up to the decimal places which exactly match the decimal places of the $p-$value corresponding to the next m^*. If we just compare the p–value for $m^* = 1$ and $m^* = 2$ we see that the p–value for $m^* = 1$ is exact up to four decimal places. According to the way the near-exact distributions are built, the $p-$values have better precision for increasing values of m^*, the number of exact moments matched by the corresponding near-exact distribution. Thus, we should not reject the null hypothesis that the covariance structure is of the BCS type, with a p–value $= 0.2792$, which is much lower than the p–value $= 0.5786$ obtained when we use the asymptotic χ^2_ν approximation for $-2\log\Lambda$ with $\nu = \frac{mu(mu+1)}{2} - m(m+1) = 9$ degrees of freedom. See (Roy and Leiva, 2011) for detail.

In Figure 1 we have, for $W = -\log\Lambda$, the plots of the p.d.f.'s and c.d.f.'s for the near-exact distribution for $m^* = 1$ and for the asymptotic Gamma distribution with shape parameter 9/2 and rate parameter 1, which corresponds to the chi-square asymptotic distribution with nine degrees of freedom for $-2\log\Lambda$.

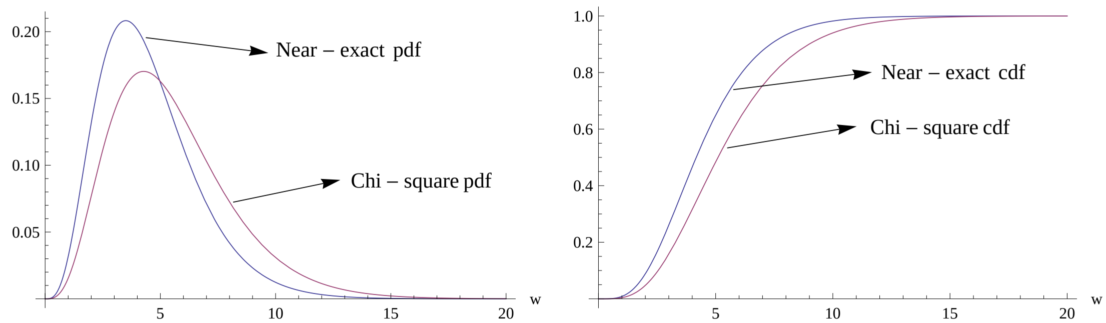

Figure 1: Plots of the p.d.f.'s and c.d.f.'s of the asymptotic $\Gamma(9/2, 1)$ distribution and the near-exact distribution ($m^* = 1$), for $W = -\log\Lambda$.

A simulation study was conducted where 100 000 pseudo-random samples with BCS structure for $u = 2$, $m = 3$ and $n = 25$ were generated. The p-value obtained from this simulation for Λ was 0.28163, which shows that p-values obtained from simulation, even when using quite large simulations, may be not that precise.

6. Conclusions

We have shown that using the approach based on Lemma 3.1 in Roy and Fonseca (2012) and on a decomposition of the null BCS hypothesis we were able to build quite easily the l.r.t. for this hypothesis, as well as to obtain the corresponding test statistic and its moments. From these we were able to obtain the characteristic function for the logarithm of the l.r.t. statistic in a form which is much adequate for developing near-exact distributions both for the l.r.t. statistic itself and its logarithm. With the help of some numerical studies we were able to show that these near-exact distributions yield very good approximations to the exact distribution. These approximations exhibit very good asymptotic behaviors not only in terms of increasing sample sizes, but also in terms of increasing values of the number of variables, and locations or time points. This asymptotic behavior for increasing number of variables is a much desirable feature which common asymptotic distributions do not have. Moreover, by using a measure that gives an upper-bound on the difference between the cumulative distribution functions of the exact and any approximate distribution we were able to show that the common chi-square asymptotic approximation for $-2\log\Lambda$ may only work in practice for very large sample sizes and when the number of variables involved is rather small, or it may indeed not even work at all when the number of variables involved is rather large. An illustration with real data shows the applicability and manageability of the near-exact distributions obtained.

Furthermore, the approach followed in this paper may be extended in order to address more complicated covariance structures arising for multi-level multivariate data.

Acknowledgments

This research was partially supported by CMA/FCT/UNL, under the project PEst-OE/MAT/UI0297/2011. The second author thanks the support for the summer research grant from the College of Business at the University of Texas at San Antonio.

Appendix A. Shape parameters in the moment expressions for Λ_b

According to Coelho and Marques (2012a) and Marques et al. (2011) the shape parameters s_j in (2.11) are given by

$$s_j = \begin{cases} s^*_{j-1} & \text{for } j=2,\dots,m, \\ & \text{except for } j=m-2\alpha_1 \\ s^*_{j-1}+(m \perp\!\!\!\perp 2)(\alpha_2-\alpha_1)\left((u-1)-\frac{m-1}{2}+(u-1)\left\lfloor\frac{m}{2(u-1)}\right\rfloor\right) & \text{for } j=m-2\alpha_1 \end{cases} \tag{A.1}$$

with

$$s^*_j = \begin{cases} \gamma_j & \text{for} \quad j=1,\dots,\alpha+1 \\ (u-1)\left(\left\lfloor\frac{m}{2}\right\rfloor-\left\lfloor\frac{j}{2}\right\rfloor\right) & \text{for} \quad j=\alpha+2,\dots,\min(m-2\alpha_1,m-1) \\ & \text{and} \quad j=2+m-2\alpha_1,\dots,2\left\lfloor\frac{m}{2}\right\rfloor-1, \text{ by steps of } 2 \\ (u-1)\left(\left\lfloor\frac{m+1}{2}\right\rfloor-\left\lfloor\frac{j}{2}\right\rfloor\right) & \text{for} \quad j=1+m-2\alpha_1,\dots,m-1, \text{ by steps of } 2\,, \end{cases} \tag{A.2}$$

and

$$\alpha=\left\lfloor\frac{m-1}{u-1}\right\rfloor, \qquad \alpha_1=\left\lfloor\frac{u-2}{u-1}\,\frac{m-1}{2}\right\rfloor, \qquad \alpha_2=\left\lfloor\frac{u-2}{u-1}\,\frac{m+1}{2}\right\rfloor, \tag{A.3}$$

where, for $j=1,\dots,\alpha$,

$$\gamma_j=\left\lfloor\frac{u-1}{2}\right\rfloor\left((j-1)(u-1)-2\,(u \perp\!\!\!\perp 2)\left\lfloor\frac{j}{2}\right\rfloor\right)+\left\lfloor\frac{u-1}{2}\right\rfloor\left\lfloor\frac{u-1+j \perp\!\!\!\perp 2}{2}\right\rfloor \tag{A.4}$$

and

$$\gamma_{\alpha+1}=-\left(\left\lfloor\frac{m}{2}\right\rfloor-\alpha\left\lfloor\frac{u-1}{2}\right\rfloor\right)^2+(u-1)\left(\left\lfloor\frac{m}{2}\right\rfloor-\left\lfloor\frac{\alpha+1}{2}\right\rfloor\right)$$
$$+((u-1) \perp\!\!\!\perp 2)\left(\alpha\left\lfloor\frac{m}{2}\right\rfloor+\frac{\alpha \perp\!\!\!\perp 2}{4}-\frac{\alpha^2}{4}-\alpha^2\left\lfloor\frac{u-1}{2}\right\rfloor\right)\,. \tag{A.5}$$

Appendix B. Gamma distribution and related results

We say that the r.v. X follows a Gamma distribution with shape parameter $r > 0$ and rate parameter $\lambda > 0$, if the p.d.f. of X is

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)}\, e^{-\lambda x}\, x^{r-1}, \qquad (x > 0)$$

and we will denote this fact by $X \sim \Gamma(r, \lambda)$. Then we know that the moment generating function of X is

$$M_X(t) = \lambda^r (\lambda - t)^{-r}\,,$$

so that if we define $Z = e^{-X}$ we will have

$$E(Z^h) = E\left(e^{-hX}\right) = M_X(-h) = \lambda^r (\lambda + h)^{-r}\,.$$

References

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed., Wiley, New Jersey.

Arnold, B. C., Coelho, C. A., Marques, F. J. (2013). The distribution of the product of powers of independent uniform random variables – a simple but useful tool to address and better understand the structure of some distributions. *Journal of Multivariate Analysis*, **113**, 19–36.

Coelho, C. A. (1998). The Generalized Integer Gamma distribution – a basis for distribution in Multivariate Statistics. *Journal of Multivariate Analysis*, **64**, 86-102.

Coelho, C. A. (2004). The Generalized Near-Integer Gamma distribution: a basis for 'near-exact' approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables. *Journal of Multivariate Analysis*, **89**, 191–218.

Coelho, C. A., Marques, F. J. (2012a). Near-exact distributions for the likelihood ratio test statistic to test equality of several variance-covariance matrices in elliptically contoured distributions. *Computational Statistics*, **27**, 627–659.

Coelho, C. A., Marques, F. J. (2012b). The multi-sample block-scalar sphericity test – exact and near-exact distributions for its likelihood ratio test statistic. *Comm. Statist. Theory Methods* (in print).

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

Marques, F. J., Coelho, C. A., Arnold, B. C. (2011). A general near-exact distribution theory for the most common likelihood ratio test statistics used in multivariate analysis. *Test*, **20**, 180–203.

Marques, F. J., Coelho, C. A. (2012). Near-exact distributions for the likelihood ratio test statistic of the multi-sample block-matrix sphericity test. *Appl. Math. Comput.*, **219**, 2861-2874.

Rao, C.R. (1945). Familial correlations or the multivariate generalizations of the intraclass correlation. *Current Science*, **14**, 66-67.

Rao, C.R. (1953). Discriminant functions for genetic differentiation and selection. *Sankhyā*, **12**, 229-246.

Roy, A., Fonseca, M. (2012). Linear models with doubly exchangeable distributed errors, *Comm. Statist. Theory Methods*, **41**, 2545–2569.

Roy, A., Leiva, R. (2011). Estimating and testing a structured covariance matrix for three-level multivariate data. *Communications in Statistics – Theory and Methods*, **40**, 1945–1963.

Tricomi, F. G., Erdélyi, A. (1951). The asymptotic expansion of a ratio of Gamma functions. *Pacific Journal of Mathematics*, **1**, 133–142.

Wilks, S. S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a Normal multivariate distribution. *Annals of Mathematical Statistics*, **17**, 257–281.