



Testing the hypothesis of a doubly exchangeable covariance matrix

Carlos A. Coelho¹ · Anuradha Roy²

Received: 25 June 2018 / Published online: 5 July 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

In this paper the authors study the problem of testing the hypothesis of a doubly exchangeable covariance matrix for three-level multivariate observations, taken on m variables over u sites and over v time/space points. Through the decomposition of the main hypothesis into a set of three sub-hypotheses, the likelihood ratio test statistic is defined, its exact moments are determined, and its exact distribution is studied. Because this distribution is very much intricate, a 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 simulation study is also conducted and two real-data examples are presented.

Keywords Characteristic function · Composition of hypotheses · Distribution of likelihood ratio statistics · Mixtures · Near-exact distributions · Product distribution

Mathematics Subject Classification 62H15 · 62H10 · 62E15 · 62E20 · 62E10 · 60E10

This research was partially supported by FCT–Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology), projects UID/MAT/00297/2013 and UID/MAT/00297/2019, through Centro de Matemática e Aplicações (CMA/FCT/UNL). The second author also thanks the support from the summer research grant from the College of Business at the University of Texas at San Antonio.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00184-019-00724-7>) contains supplementary material, which is available to authorized users.

✉ Carlos A. Coelho
cmac@fct.unl.pt

¹ Departamento de Matemática and Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

² Department of Management Science and Statistics, The University of Texas at San Antonio, San Antonio, TX, USA

1 Introduction

Advances in computing power in the past few decades greatly encouraged the collection of multi-level multivariate data in all fields of science: biomedical, medical, environmental, social sciences and engineering. And, with these data sets complex multivariate testing problems occur frequently. It is common in clinical trial studies to collect measurements on more than one response variable at several locations taken repeatedly over time on one experimental unit to test the effectiveness of some medication, diet or treatment. These are called three-level multivariate data and the doubly exchangeable covariance structure (defined below) is a suitable variance-covariance matrix for this kind of data.

Good examples of such data may be found in osteopenia and osteoporosis studies, being estimated that one of every four post-menopausal women may suffer from osteoporosis. Although it is more common in white or Asian women older than 50 years, osteoporosis can occur in almost any person at any age, being estimated that in fact, more than 2 million American men have osteoporosis, the national cost for osteoporosis and related injuries being assessed at \$14 billion each year in the United States.

Let \mathbf{y} be the muv -variate real-valued random vector of all measurements. We partition this vector \mathbf{y} as follows:

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_v \end{pmatrix}, \quad \text{where } \mathbf{y}_t = \begin{pmatrix} \mathbf{y}_{t1} \\ \vdots \\ \mathbf{y}_{tu} \end{pmatrix}, \quad \text{with } \mathbf{y}_{ts} = \begin{pmatrix} \mathbf{y}_{ts1} \\ \vdots \\ \mathbf{y}_{tsm} \end{pmatrix},$$

for $t = 1, \dots, v$, $s = 1, \dots, u$. The m -dimensional vector of measurements \mathbf{y}_{ts} represents the replicate on the s th location and at the t th time point.

Let $\Theta = \text{Cov}[\mathbf{y}]$ be the $(muv \times muv)$ -dimensional partitioned covariance matrix. We say that the covariance matrix Θ has a doubly exchangeable covariance structure (Roy and Leiva 2007) if it can be written as

$$\begin{aligned} \Theta &= \mathbf{I}_{uv} \otimes \mathbf{U}_0 + [\mathbf{I}_v \otimes (\mathbf{J}_u - \mathbf{I}_u)] \otimes \mathbf{U}_1 + [\mathbf{J}_{uv} - (\mathbf{I}_v \otimes \mathbf{J}_u)] \otimes \mathbf{W} \\ &= \mathbf{I}_{uv} \otimes (\mathbf{U}_0 - \mathbf{U}_1) + \mathbf{I}_v \otimes \mathbf{J}_u \otimes (\mathbf{U}_1 - \mathbf{W}) + \mathbf{J}_{uv} \otimes \mathbf{W}, \\ &= \mathbf{I}_v \otimes \mathbf{U} + (\mathbf{J}_v - \mathbf{I}_v) \otimes \mathbf{W}^* \end{aligned} \quad (1)$$

where

$$\mathbf{U} = \mathbf{I}_u \otimes \mathbf{U}_0 + (\mathbf{J}_u - \mathbf{I}_u) \otimes \mathbf{U}_1$$

and

$$\mathbf{W}^* = \mathbf{J}_u \otimes \mathbf{W},$$

with \mathbf{U}_0 a positive definite symmetric $m \times m$ matrix, and \mathbf{U}_1 and \mathbf{W} symmetric $m \times m$ matrices. The matrices \mathbf{U}_0 , \mathbf{U}_1 and \mathbf{W} are all unstructured. That is, the doubly

exchangeable covariance structure can be written as a matrix having diagonal blocks of the same matrix \mathbf{U} with component matrices \mathbf{U}_0 and \mathbf{U}_1 and otherwise being filled-up with matrices \mathbf{W} .

Thus, the vectors $\mathbf{y}_{11}, \dots, \mathbf{y}_{1u}, \dots, \mathbf{y}_{v1}, \dots, \mathbf{y}_{vu}$ are doubly exchangeable if

$$\text{Cov}[\mathbf{y}_{ts}; \mathbf{y}_{t^*s^*}] = \begin{cases} \mathbf{U}_0 & \text{if } t = t^* \text{ and } s = s^*, \\ \mathbf{U}_1 & \text{if } t = t^* \text{ and } s \neq s^*, \\ \mathbf{W} & \text{if } t \neq t^*, \end{cases}$$

with $\text{Cov}[\mathbf{y}_{ts}; \mathbf{y}_{t^*s^*}] = \text{Cov}[\mathbf{y}_{t^*s^*}; \mathbf{y}_{ts}]$.

The $m \times m$ diagonal blocks \mathbf{U}_0 in (1) represent the variance-covariance matrix of the m response variables at any given location and at any given time point, whereas the $m \times m$ off-diagonal blocks \mathbf{U}_1 in (1) represent the covariance matrix of the m response variables between any two different locations, at any given time point. We assume \mathbf{U}_0 is the same for all locations and time points, and \mathbf{U}_1 is the same for all time points. The $m \times m$ off-diagonal blocks \mathbf{W} represent the covariance matrix of the m response variables between any two different time points. It is assumed to be the same for any pair of different time points, irrespective of location.

A good and simple example of a three-level multivariate dataset on osteoporosis may be obtained from Johnson and Wichern (2007) who report data on a study where an investigator measures the mineral content of three bones, radius, humerus and ulna ($m = 3$) by photon absorptiometry to examine whether a particular dietary supplement increases bone mineral content and mass in older women. All three measurements are recorded on the dominant and non-dominant sides ($u = 2$) for each woman. These two-level multivariate measurements are then taken again one year after their first participation in the experimental program. Thus, this whole dataset has a three-level multivariate structure, with $m = 3$ variables, for $u = 2$ locations, over $v = 2$ time points and given what the variables involved represent, testing for a doubly exchangeable covariance structure may be an adequate goal. Another simple example is from a bone densitometry study where the interest may be to test the doubly exchangeable covariance structure for an osteopenia study on twelve patients. Bone mineral density were obtained from the femoral neck and trochanter ($m = 2$), for the right and left femur ($u = 2$), taken at two different times ($v = 2$), separated by about two years. This latter example is addressed in Sect. 6, while the example of the osteoporosis data from Johnson and Wichern (2007) is placed in the supplementary material.

Roy and Fonseca (2012) fitted a general linear model to three-level multivariate data, with a doubly exchangeable covariance structure for the error vector, and Leiva and Roy (2011, 2012) used this doubly exchangeable covariance structure for classification of three-level multivariate data, while hypothesis testing on three-level multivariate data was first studied by Roy and Leiva (2008), where these two authors introduced parametrically parsimonious models for hypotheses testing of Kronecker product covariance structures.

However, none of these authors addressed directly the problem of testing the doubly exchangeable covariance structure.

The doubly exchangeable covariance structure is actually a very general and rich covariance structure which generalizes at the same time both the compound symmetry

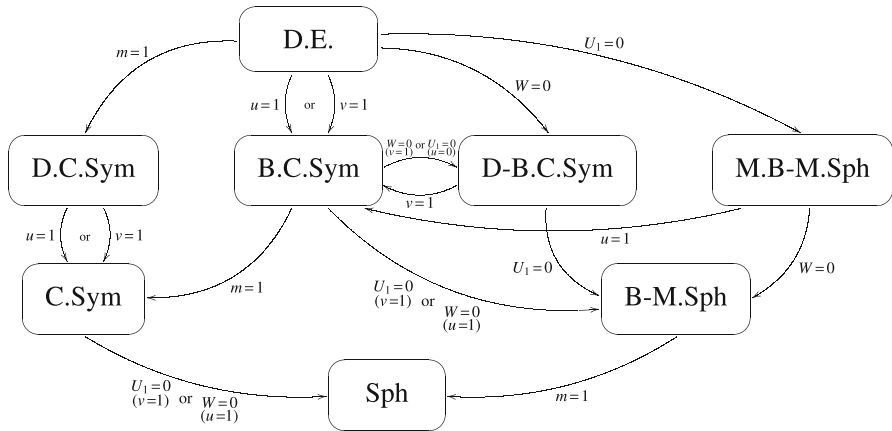


Fig. 1 Web of relations of covariance structures which are particular cases of the doubly exchangeable structure. Covariance structures depicted and their acronyms: *D.E.*—Double exchangeable, *D.C.Sym*—Double compound symmetry, *B.C.Sym*—Block Compound Symmetry (or block exchangeability), *D-B.C.Sym*—Diagonal-block compound symmetry, *M.B-M.Sph*—Multi-Block-Matrix Sphericity, *C.Sym*—Compound symmetry, *B-M.Sph.*—Block-matrix sphericity, *Sph*—Sphericity

and the sphericity structures. As such it has many interesting covariance structures as particular cases. Some of these are depicted in Fig. 1. While some of these structures are rather well-known, as the above mentioned compound symmetry and sphericity structures, and also the block compound symmetry (Votaw 1948; Szatrowski 1976, 1982; Coelho and Roy 2017) and the block-matrix sphericity structure (Moschopoulos 1992; Cardeno and Nagar 2001; Marques and Coelho 2012), others are not even known in the literature as it is the case of the structure derived from the doubly exchangeable covariance structure when $U_1 = \mathbf{0}$ but $W \neq \mathbf{0}$, which we call multi-block-matrix sphericity (see Fig. 1). We may note how for either $u = 1$ or $v = 1$ we have the block compound symmetric structure for which a likelihood ratio test (l.r.t.) was developed in Coelho and Roy (2017), being thus the test in the present paper a generalization of the test in this reference.

Therefore, once the test for double exchangeability is developed, it may be used to test for (see Fig. 1)

- (i) block compound symmetry or block exchangeability, for $u = 1$ or $v = 1$,
- (ii) double compound symmetry, for $m = 1$, or
- (iii) compound symmetry, for $m = 1$ and either $u = 1$ or $v = 1$.

And, as such, in case we reject the doubly exchangeable structure for a given covariance matrix, we may then take our testing procedure one step further by testing for example for block compound symmetry or block exchangeability of submatrices, in trying to evaluate which may be the reason why the doubly exchangeable structure was rejected, and this can be done just by using one of the particular cases of this same test.

The use of the doubly exchangeable structure will also enable the use of a much smaller number of parameters to model the covariance structure, since for an unstructured covariance matrix of dimensions $mu\nu \times mu\nu$ there are $mu\nu(mu\nu + 1)/2$ unknown parameters, whereas the doubly exchangeable covariance matrix has only $3m(m + 1)/2$

unknown parameters, which does not depend on either u or v . Perlman (1987) stresses the importance of carrying out tests for symmetry structures by saying that “if symmetries are known to be present, then sharper statistical inferences can be obtained”.

These facts motivated the authors to develop a l.r.t. for the doubly exchangeable covariance structure for three-level multivariate data. The method used allows for an easy way to obtain the l.r.t. statistic as well as the characterization of its exact distribution, which is done in Sects. 2 and 3 of the paper. Given the extremely complicated structure of the exact distribution of the l.r.t. statistic, the development of sharp but manageable approximations stands up as a most desirable goal, in order to enable the practical application of the test. It happens that the approach undertaken also enables to obtain an extremely useful factorization of the characteristic function of the logarithm of the l.r.t. statistic, which in turn opens the way for the development of extremely sharp and manageable near-exact approximations for the distribution of the likelihood ratio statistic. This work is done in Sect. 4 of the paper. In Sect. 5 the authors carry out a number of numerical studies which show the extremely good performance of the near-exact distributions developed, even for very small samples and/or very large numbers of variables involved. In Sect. 6 two real-data examples are used to exemplify the practical application of the l.r.t. and of the near-exact approximations developed. Conclusions are drawn in Sect. 7.

2 Formulation of the hypothesis and the likelihood ratio test

Let $\mathbf{y} \sim N_{muv}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We are interested in testing the hypothesis

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Theta} , \tag{2}$$

where $\boldsymbol{\Theta}$ is defined in (1).

In Lemma 3.1 in Roy and Fonseca (2012), it is shown that for $\boldsymbol{\Gamma}^* = \underset{v \times v}{\mathbf{C}'} \otimes \mathbf{I}_{mu}$ and $\boldsymbol{\Gamma}^\bullet = \mathbf{I}_v \otimes (\underset{u \times u}{\mathbf{C}^{*'}} \otimes \mathbf{I}_m)$, where \mathbf{C} and \mathbf{C}^* are orthogonal Helmert matrices whose first columns are proportional to $\mathbf{1}$'s,

$$\boldsymbol{\Gamma}^\bullet \boldsymbol{\Gamma}^* \boldsymbol{\Theta} \boldsymbol{\Gamma}^{*'} \boldsymbol{\Gamma}^{\bullet'} = \text{diag}(\boldsymbol{\Delta}_3, \mathbf{I}_{u-1} \otimes \boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \mathbf{I}_{u-1} \otimes \boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \dots, \mathbf{I}_{u-1} \otimes \boldsymbol{\Delta}_1)$$

where

$$\boldsymbol{\Delta}_1 = \mathbf{U}_0 - \mathbf{U}_1,$$

$$\boldsymbol{\Delta}_2 = \mathbf{U}_0 + (u-1)\mathbf{U}_1 - u\mathbf{W} = (\mathbf{U}_0 - \mathbf{U}_1) + u(\mathbf{U}_1 - \mathbf{W}),$$

$$\text{and } \boldsymbol{\Delta}_3 = \mathbf{U}_0 + (u-1)\mathbf{U}_1 + u(v-1)\mathbf{W} = (\mathbf{U}_0 - \mathbf{U}_1) + u(\mathbf{U}_1 - \mathbf{W}) + uv\mathbf{W}.$$

Since $\boldsymbol{\Gamma}^\bullet$ and $\boldsymbol{\Gamma}^*$ are not function of either \mathbf{U}_0 , or \mathbf{U}_1 or \mathbf{W} , to test H_0 in (2) is equivalent to test

$$H_0 : \boldsymbol{\Sigma}^* = \boldsymbol{\Omega} \tag{3}$$

where

$$\Sigma^* = \Gamma^* \Gamma^* \Sigma \Gamma^{*'} \Gamma^{*'} \quad \text{and} \quad \Omega = \Gamma^* \Gamma^* \Theta \Gamma^{*'} \Gamma^{*'}.$$

We may split the null hypothesis in (3) as

$$H_0 \equiv (H_{0c|a} \parallel H_{0b|a}) \circ H_{0a}, \tag{4}$$

where ‘ \circ ’ means ‘after’ and ‘ \parallel ’ means ‘parallel’, meaning ‘either after or before’.

In (4),

$$H_{0a} : \Sigma^* = \text{block-diag}(\Sigma_i^*, i = 1, \dots, uv), \tag{5}$$

is the hypothesis of independence of the uv diagonal blocks Σ_i^* ($i = 1, \dots, uv$) of size $m \times m$ of Σ^* ;

$$H_{0b|a} : \underbrace{\Sigma_2^* = \dots = \Sigma_u^*}_{u-1} = \underbrace{\Sigma_{u+2}^* = \dots = \Sigma_{2u}^*}_{u-1} = \dots = \underbrace{\Sigma_{(v-1)u+2}^* = \dots = \Sigma_{vu}^*}_{u-1},$$

assuming H_{0a} ,

(6)

is the hypothesis of equality of $v(u - 1)$ covariance matrices of dimension $m \times m$, assuming H_{0a} , and

$$H_{0c|a} : \Sigma_{u+1}^* = \Sigma_{2u+1}^* = \dots = \Sigma_{(v-1)u+1}^* \tag{7}$$

assuming H_{0a} ,

is the hypothesis of equality of the covariance matrices $\Sigma_{u+1}^*, \Sigma_{2u+1}^*, \dots, \Sigma_{(v-1)u+1}^*$, assuming H_{0a} .

The l.r.t. statistic to test H_{0a} in (5) is (Anderson 2003, Sect. 9.2)

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

where $\mathbf{A} = \Gamma^* \Gamma^* \mathbf{A}^+ \Gamma^{*'} \Gamma^{*'}$ is the maximum likelihood estimator (m.l.e.) of Σ^* , and \mathbf{A}_j its j -th diagonal $m \times m$ block, being \mathbf{A}^+ the m.l.e. of Σ .

The l.r.t. statistic to test $H_{0b|a}$ in (6) is (Anderson 2003, Sect. 10.2)

$$\Lambda_b = \left((v(u - 1))^{mv(u-1)} \frac{\prod_{\ell=1}^v \prod_{k=1}^{u-1} |\mathbf{A}_{(\ell-1)u+1+k}|}{|\mathbf{A}^*|^{v(u-1)}} \right)^{n/2}, \tag{8}$$

where

$$\mathbf{A}^* = \sum_{\ell=1}^v \sum_{k=1}^{u-1} \mathbf{A}_{(\ell-1)u+1+k}.$$

The l.r.t. statistic to test $H_{0c|a}$ in (7) is (Anderson 2003, Sect. 10.2)

$$\Lambda_c = \left((v-1)^{m(v-1)} \frac{\prod_{k=1}^{v-1} |\mathbf{A}_{ku+1}|}{|\mathbf{A}^{**}|^{v-1}} \right)^{n/2} \tag{9}$$

where

$$\mathbf{A}^{**} = \sum_{k=1}^{v-1} \mathbf{A}_{ku+1}.$$

Then, through an extension of Lemma 10.3.1 in (Anderson 2003, Sect. 10.3), the l.r.t. statistic to test H_0 in (3) will be

$$\begin{aligned} \Lambda &= \Lambda_a \Lambda_b \Lambda_c \\ &= \left((v(u-1))^{mv(u-1)} (v-1)^{m(v-1)} \frac{|\mathbf{A}|}{|\mathbf{A}_1| |\mathbf{A}^{*}|^{v(u-1)} |\mathbf{A}^{**}|^{v-1}} \right)^{n/2}, \end{aligned} \tag{10}$$

with

$$E(\Lambda^h) = E(\Lambda_a^h) E(\Lambda_b^h) E(\Lambda_c^h), \tag{11}$$

since on one hand, under H_{0a} Λ_a is independent of $\prod_{j=1}^{uv} |\mathbf{A}_j|$ (Marques and Coelho 2012; Coelho and Marques 2012b), which makes Λ_a independent of Λ_b and Λ_c , while on the other hand, the \mathbf{A}_j ($j = 1, \dots, uv$), under H_{0a} , are independent among themselves, which makes Λ_b and Λ_c independent because they are built on different \mathbf{A}_j 's.

In the following section we obtain the expressions for the moments of all three l.r.t. statistics Λ_a , Λ_b and Λ_c , as well as their distributions.

3 On the exact distribution of the l.r.t. statistic

Using the results in Coelho (2004), Coelho et al. (2010) and Marques et al. (2011) we may write the h -th moment of Λ_a as

$$\begin{aligned} E(\Lambda_a^h) &= \prod_{k=1}^{uv-1} \prod_{j=1}^m \frac{\Gamma\left(\frac{n-j}{2}\right) \Gamma\left(\frac{n-(uv-k)m-j}{2} + \frac{nh}{2}\right)}{\Gamma\left(\frac{n-(uv-k)m-j}{2}\right) \Gamma\left(\frac{n-j}{2} + \frac{nh}{2}\right)} \\ &= \underbrace{\left\{ \prod_{j=3}^{muv} \binom{n-j}{n}^{r_j} \binom{n-j}{n}^{-r_j} \right\}}_{\Phi_{a,1}(h)} \underbrace{\left\{ \left(\frac{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{n-2}{2} + \frac{nh}{2}\right)}{\Gamma\left(\frac{n-1}{2} + \frac{nh}{2}\right) \Gamma\left(\frac{n-2}{2}\right)} \right)^{k^*} \right\}}_{\Phi_{a,2}(h)} \end{aligned} \tag{12}$$

with

$$k^* = \begin{cases} \left\lfloor \frac{uv}{2} \right\rfloor, & m \text{ odd} \\ 0, & m \text{ even,} \end{cases} \quad (13)$$

where $\lfloor \cdot \rfloor$ represents the integer function, that is, the largest integer that does not exceed its argument, and

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

with

$$h_j = \begin{cases} uv - 1, & j = 1, \dots, m \\ -1, & j = m + 1, \dots, muv - 2. \end{cases}$$

Now, using the results in (Coelho and Marques 2012a; Coelho et al. 2010; Marques et al. 2011) we obtain the expression for the h -th moment of Λ_b as

$$\begin{aligned} E\left(\Lambda_b^h\right) &= \prod_{j=1}^m \prod_{k=1}^{v(u-1)} \frac{\Gamma\left(\frac{n-1}{2} - \frac{j-1}{2v(u-1)} + \frac{k-1}{v(u-1)}\right) \Gamma\left(\frac{n-j}{2} + \frac{n}{2}h\right)}{\Gamma\left(\frac{n-1}{2} - \frac{j-1}{2v(u-1)} + \frac{k-1}{v(u-1)} + \frac{n}{2}h\right) \Gamma\left(\frac{n-j}{2}\right)} \\ &= \underbrace{\left\{ \prod_{j=2}^m \binom{n-j}{n}^{s_j} \binom{n-j}{n}^{-s_j} \right\}}_{\Phi_{b,1}(h)} \\ &\quad \times \left\{ \prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v(u-1)} \frac{\Gamma\left(n-1 + \frac{k-2j}{v(u-1)}\right) \Gamma\left(n-1 + \left\lfloor \frac{k-2j}{v(u-1)} \right\rfloor + nh\right)}{\Gamma\left(n-1 + \frac{k-2j}{v(u-1)} + nh\right) \Gamma\left(n-1 + \left\lfloor \frac{k-2j}{v(u-1)} \right\rfloor\right)} \right\} \\ &\quad \times \underbrace{\left\{ \prod_{k=1}^{v(u-1)} \frac{\Gamma\left(\frac{n-m}{2} + \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)}\right) \Gamma\left(\frac{n-m}{2} + \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} \right\rfloor + \frac{n}{2}h\right)}{\Gamma\left(\frac{n-m}{2} + \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} + \frac{n}{2}h\right) \Gamma\left(\frac{n-m}{2} + \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} \right\rfloor\right)} \right\}^{\tilde{m}}}_{\Phi_{b,2}(h)} \end{aligned} \quad (14)$$

where $\tilde{m} = m - 2\lfloor m/2 \rfloor$ and s_j ($j = 2, \dots, m$) are given in Appendix A in the supplementary material.

By looking at (8) and (9) we may see that the h -th moment of Λ_c may be obtained, as follows, from the h -th moment of Λ_b by first replacing v by 1 and then replacing u by v ,

$$\begin{aligned}
 E\left(\Lambda_c^h\right) &= \prod_{j=1}^m \prod_{k=1}^{v-1} \frac{\Gamma\left(\frac{n-1}{2}-\frac{j-1}{2(v-1)}+\frac{k-1}{v-1}\right) \Gamma\left(\frac{n-j}{2}+\frac{n}{2} h\right)}{\Gamma\left(\frac{n-1}{2}-\frac{j-1}{2(v-1)}+\frac{k-1}{v-1}+\frac{n}{2} h\right) \Gamma\left(\frac{n-j}{2}\right)} \\
 &= \underbrace{\left\{ \prod_{j=2}^m \left(\frac{n-j}{n}\right)^{\delta_j} \left(\frac{n-j}{n}+h\right)^{-\delta_j} \right\}}_{\Phi_{c,1}(h)} \\
 &\quad \times \left\{ \prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v-1} \frac{\Gamma\left(n-1+\frac{k-2j}{v-1}\right) \Gamma\left(n-1+\left\lfloor \frac{k-2j}{v-1} \right\rfloor+n h\right)}{\Gamma\left(n-1+\frac{k-2j}{v-1}+n h\right) \Gamma\left(n-1+\left\lfloor \frac{k-2j}{v-1} \right\rfloor\right)} \right\} \\
 &\quad \times \underbrace{\left\{ \prod_{k=1}^{v-1} \frac{\Gamma\left(\frac{n-m}{2}+\frac{m-1}{2}+\frac{2k-m-1}{2(v-1)}\right) \Gamma\left(\frac{n-m}{2}+\left\lfloor \frac{m-1}{2}+\frac{2k-m-1}{2(v-1)} \right\rfloor+\frac{n}{2} h\right)}{\Gamma\left(\frac{n-m}{2}+\frac{m-1}{2}+\frac{2k-m-1}{2(v-1)}+\frac{n}{2} h\right) \Gamma\left(\frac{n-m}{2}+\left\lfloor \frac{m-1}{2}+\frac{2k-m-1}{2(v-1)} \right\rfloor\right)} \right\}^{\tilde{m}}}_{\Phi_{c,2}(h)}
 \end{aligned}
 \tag{15}$$

where $\tilde{m} = m - 2\lfloor m/2 \rfloor$ and the shape parameters δ_j ($j = 2, \dots, m$) are given in Appendix A in the supplementary material.

Since the supports of Λ_a , Λ_b and Λ_c are delimited, their distributions are defined by their moments, and as such, from the first expression in (12) we may write

$$\Lambda_a \stackrel{st}{\sim} \prod_{j=1}^m \prod_{k=1}^{uv-1} \left(X_{jk}\right)^{n/2}, \quad \text{where } X_{jk} \sim \text{Beta}\left(\frac{n-(uv-k)m-j}{2}, \frac{(uv-k)m}{2}\right),
 \tag{16}$$

where ‘ $\stackrel{st}{\sim}$ ’ means ‘stochastically equivalent to’ and X_{jk} ($j = 1, \dots, m$; $k = 1, \dots, uv - 1$) are independent random variables, while from the first expression in (14) we may write

$$\Lambda_b \stackrel{st}{\sim} \prod_{j=1}^m \prod_{k=1}^{v(u-1)} \left(X_{jk}^*\right)^{n/2}, \quad \text{where } X_{jk}^* \sim \text{Beta}\left(\frac{n-j}{2}, \frac{j-1}{2}+\frac{2k-j-1}{2v(u-1)}\right),
 \tag{17}$$

where X_{jk}^* ($j = 1, \dots, m$; $k = 1, \dots, v(u - 1)$) are independent, and from the first expression in (15) we may write

$$\Lambda_c \stackrel{st}{\sim} \prod_{j=1}^m \prod_{k=1}^{v-1} \left(X_{jk}^{**}\right)^{n/2}, \quad \text{where } X_{jk}^{**} \sim \text{Beta}\left(\frac{n-j}{2}, \frac{j-1}{2}+\frac{2k-j-1}{2(v-1)}\right),
 \tag{18}$$

where X_{jk}^{**} ($j = 1, \dots, m; k = 1, \dots, v - 1$) are independent, so that we may write for the overall l.r.t. statistic for H_0 in (4)

$$\Lambda \stackrel{st}{\sim} \prod_{j=1}^m \left\{ \left(\prod_{k=1}^{uv-1} X_{jk} \right)^{n/2} \times \left(\prod_{k=1}^{v(u-1)} X_{jk}^* \right)^{n/2} \times \left(\prod_{k=1}^{v-1} X_{jk}^{**} \right)^{n/2} \right\}, \tag{19}$$

where all random variables are independent.

On the other hand, based on the results in Appendix B in the supplementary material and from the second expressions in (12), (14) and (15) we may respectively write,

$$\Lambda_a \stackrel{st}{\sim} \left(\prod_{j=3}^{muv} e^{-Z_j} \right) \times \left(\prod_{j=1}^{k^*} (Y_j)^{n/2} \right) \tag{20}$$

where

$$Z_j \sim \Gamma \left(r_j, \frac{n-j}{n} \right) \text{ and } Y_j \sim \text{Beta} \left(\frac{n-2}{2}, \frac{1}{2} \right) \tag{21}$$

are all independent random variables, while for Λ_b we may write

$$\Lambda_b \stackrel{st}{\sim} \left(\prod_{j=2}^m e^{-Z_j^*} \right) \times \left(\prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v(u-1)} (Y_{1jk}^*)^n \right) \times \left(\prod_{k=1}^{v(u-1)} (Y_{2k}^*)^{n/2} \right)^{m \perp\!\!\!\perp 2} \tag{22}$$

where $m \perp\!\!\!\perp 2 = \text{mod}(m, 2)$,

$$Z_j^* \sim \Gamma \left(s_j, \frac{n-j}{n} \right), \quad Y_{1jk}^* \sim \text{Beta} \left(n-1 + \left\lfloor \frac{k-2j}{v(u-1)} \right\rfloor, \frac{k-2j}{v(u-1)} - \left\lfloor \frac{k-2j}{v(u-1)} \right\rfloor \right), \tag{23}$$

and

$$Y_{2k}^* \sim \text{Beta} \left(\frac{n-m}{2} + \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} \right\rfloor, \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} - \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} \right\rfloor \right) \tag{24}$$

are all independent random variables, while for Λ_c we may write

$$\Lambda_c \stackrel{st}{\sim} \left(\prod_{j=2}^m e^{-Z_j^{**}} \right) \times \left(\prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v-1} (Y_{1jk}^{**})^n \right) \times \left(\prod_{k=1}^{v-1} (Y_{2k}^{**})^{n/2} \right)^{m \perp\!\!\!\perp 2} \tag{25}$$

where

$$Z_j^{**} \sim \Gamma\left(\delta_j, \frac{n-j}{n}\right), \quad Y_{1jk}^{**} \sim \text{Beta}\left(n-1 + \left\lfloor \frac{k-2j}{v-1} \right\rfloor, \frac{k-2j}{v-1} - \left\lfloor \frac{k-2j}{v-1} \right\rfloor\right), \tag{26}$$

and

$$Y_{2k}^{**} \sim \text{Beta}\left(\frac{n-m}{2} + \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2(v-1)} \right\rfloor, \frac{m-1}{2} + \frac{2k-m-1}{2(v-1)} - \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2(v-1)} \right\rfloor\right) \tag{27}$$

are all independent random variables.

Thus, we have the following Theorem.

Theorem 1 *The exact distribution of the overall l.r.t. statistic Λ in (10), to test H_0 in (3) or (4) is, for general u, v and m , the same as that of*

$$\left(\prod_{j=2}^{muv} e^{-T_j}\right) \times \left(\prod_{j=1}^{k^*} Y_j\right)^{n/2} \times \left(\prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v(u-1)} Y_{1jk}^*\right)^n \times \left(\prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v-1} Y_{1jk}^{**}\right)^n \times \left\{ \left(\prod_{k=1}^{v(u-1)} Y_{2k}^*\right)^{n/2} \times \left(\prod_{k=1}^{v-1} Y_{2k}^{**}\right)^{n/2} \right\}^{\tilde{m}} \tag{28}$$

where $\tilde{m} = m - 2\lfloor m/2 \rfloor$ and, for $j = 2, \dots, muv$,

$$T_j \sim \Gamma\left(\mu_j, \frac{n-j}{n}\right),$$

with

$$\mu_j = \sum_{j=2}^{muv} (r_j^+ + s_j^+ + \delta_j^+) \tag{29}$$

where

$$r_j^+ = \begin{cases} 0, & j = 2 \\ r_j, & j = 3, \dots, muv \end{cases}$$

and

$$s_j^+ = \begin{cases} s_j, & j = 2, \dots, m \\ 0, & j = m + 1, \dots, muv, \end{cases} \quad \delta_j^+ = \begin{cases} \delta_j, & j = 2, \dots, m \\ 0, & j = m + 1, \dots, muv, \end{cases}$$

where the shape parameters s_j and δ_j are defined in Appendix A in the supplementary material. The distributions of $Y_j, Y_{1jk}^*, Y_{2k}^*, Y_{1jk}^{**}$ and Y_{2k}^{**} are defined in (21), (23), (24), (26) and (27).

Proof The proof of the above theorem is rather trivial, from the previously established results. We may only remark that the random variables T_j are the sum of the random variables Z_j, Z_j^* and Z_j^{**} , which are independent Gamma distributed random variables, all with the same rate parameters $\frac{n-j}{n}$, for a given j . As such, for a given j , their sum is a Gamma distributed random variable with that same rate parameter and a shape parameter which is the sum of the original shape parameters. \square

The following Corollary refers to the particular cases for $v = 1$ and $v = 2$, which are particular cases of interest.

Corollary 1 For both $v = 1$ and $v = 2$, the hypothesis $H_{0c|a}$ vanishes and as such the distribution of Λ is in this case the same as that of

$$\left(\prod_{j=2}^{muv} e^{-T_j} \right) \times \left(\prod_{j=1}^{k^*} Y_j \right)^{n/2} \times \left(\prod_{j=1}^{\lfloor m/2 \rfloor} \prod_{k=1}^{v(u-1)} Y_{1jk}^* \right)^n \times \left\{ \left(\prod_{k=1}^{v(u-1)} Y_{2k}^* \right)^{n/2} \right\}^{m \lfloor 2 \rfloor} \tag{30}$$

where, for $j = 2, \dots, muv$,

$$T_j \sim \Gamma \left(\mu_j^*, \frac{n-j}{n} \right),$$

with

$$\mu_j^* = \sum_{j=2}^{muv} (r_j^+ + s_j^+) \tag{31}$$

where

$$r_j^+ = \begin{cases} 0, & j = 2 \\ r_j, & j = 3, \dots, muv, \end{cases} \quad \text{and} \quad s_j^+ = \begin{cases} s_j, & j = 2, \dots, m \\ 0, & j = m + 1, \dots, muv, \end{cases}$$

where the shape parameters s_j are defined in Appendix A in the supplementary material. The distributions of Y_j, Y_{1jk}^* and Y_{2k}^* , are defined in (21), (23) and (24).

The case for $v = 1$ is equivalent to the test for block compound symmetric covariance structure, addressed in Coelho and Roy (2017).

4 The characteristic function of $W = -\log \Lambda$ and the development of near-exact distributions

4.1 The characteristic function of $W = -\log \Lambda$

In the previous section, the reason we obtain two equivalent representations for the exact distribution of Λ , one from (19) and the other one in Theorem 1, is because the first one of these representations neither yields a manageable cumulative distribution function nor it adequately leads to a sharp approximation to the exact distribution of Λ .

In order to obtain a very sharp and manageable approximation to the exact distribution of Λ , we will base our developments on the representations given by Theorem 1 and Corollary 1.

From Theorem 1 and expressions (12)–(15) we may write the characteristic function (c.f.) of $W = -\log \Lambda$ as

$$\begin{aligned} \Phi_W(t) &= E\left(e^{itW}\right) = E\left(\Lambda^{-it}\right) \\ &= \underbrace{\left\{ \prod_{j=2}^{muv} \left(\frac{n-j}{n}\right)^{\mu_j} \left(\frac{n-j}{n} - it\right)^{-\mu_j} \right\}}_{\Phi_{W,1}(t)} \times \underbrace{\Phi_{a,2}(-it) \Phi_{b,2}(-it) \Phi_{c,2}(-it)}_{\Phi_{W,2}(t)} \end{aligned} \tag{32}$$

where μ_j is given by (29), $\Phi_{a,2}(\cdot)$, $\Phi_{b,2}(\cdot)$ and $\Phi_{c,2}(\cdot)$ are defined in (12)–(15), and $\Phi_{W,1}(t)$ is actually equal to $\Phi_{a,1}(-it)\Phi_{b,1}(-it)\Phi_{c,1}(-it)$.

For $v = 1$ and $v = 2$, according to Corollary 1 in the previous section $\Phi_W(t)$ reduces to

$$\Phi_W(t) = \underbrace{\left\{ \prod_{j=2}^{muv} \left(\frac{n-j}{n}\right)^{\mu_j^*} \left(\frac{n-j}{n} - it\right)^{-\mu_j^*} \right\}}_{\Phi_{W,1}(t)} \times \underbrace{\Phi_{a,2}(-it) \Phi_{b,2}(-it)}_{\Phi_{W,2}(t)} \tag{33}$$

for μ_j^* given by (31) and where now $\Phi_{W,1}(t)$ is equal to $\Phi_{a,1}(-it)\Phi_{b,1}(-it)$.

Expressions (32) and (33), together with expressions (28) and (30), show that the exact distribution of $W = -\log \Lambda$ is the same as that of the sum of $muv - 1$ independent Gamma random variables with an independent random variable which itself is distributed as a sum of independent Logbeta random variables.

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

4.2 Development of near-exact distributions

Based on the result in Section 5 of Tricomi and Erdélyi (1951), we know that we may, for increasing values of a , asymptotically replace the distribution of any *Logbeta*(a, b) distributed random variable by an infinite mixture of $\Gamma(b + \ell, a)$ distributions ($\ell = 0, 1, \dots$). As such, we could replace $\Phi_{W,2}(t)$ in either (32) or (33) by the c.f. of the sum of infinite mixtures of Gamma distributions, which would be the same as the c.f. of an infinite mixture of sums of Gamma distributions. Although it happens that these Gamma distributions would have different rate parameters, these parameters would anyway be of comparable magnitude. As such, in building our near-exact distributions for $W = -\log \Lambda$ and Λ , while we will leave $\Phi_{W,1}(t)$ unchanged, we will replace $\Phi_{W,2}(t)$ in either (32) or (33), by

$$\Phi^*(t) = \sum_{\ell=0}^{m^*} \pi_{\ell} v^{r+\ell} (v - it)^{-(r+\ell)} \quad (34)$$

which is the c.f. of a finite mixture of $\Gamma(r + \ell, v)$ distributions, where for the general case in (32) and for k^* in (13) and $v > 1$, we will take

$$\begin{aligned} r &= \frac{k^*}{2} + \sum_{j=1}^{\lfloor m/2 \rfloor} \sum_{k=1}^{v(u-1)} \frac{k-2j}{v(u-1)} - \left\lfloor \frac{k-2j}{v(u-1)} \right\rfloor + \sum_{j=1}^{\lfloor m/2 \rfloor} \sum_{k=1}^{v-1} \frac{k-2j}{v(u-1)} - \left\lfloor \frac{k-2j}{v-1} \right\rfloor \\ &+ (m \perp 2) \left(\sum_{k=1}^{v(u-1)} \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} - \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} \right\rfloor \right. \\ &\left. + \sum_{k=1}^{v-1} \frac{m-1}{2} + \frac{2k-m-1}{2(v-1)} - \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2(v-1)} \right\rfloor \right) \\ &= \begin{cases} \frac{m}{4}(uv-3), & \text{even } m \\ \frac{1}{2} \left\lfloor \frac{uv}{2} \right\rfloor + \frac{m+1}{4}(uv-3), & \text{odd } m, \end{cases} \end{aligned} \quad (35)$$

which is the sum of all the second parameters of the Logbeta distributions in $\Phi_{W,2}(t)$ in (32), while for the particular case in (33) we will take

$$\begin{aligned} r &= \frac{k^*}{2} + \sum_{j=1}^{\lfloor m/2 \rfloor} \sum_{k=1}^{v(u-1)} \frac{k-2j}{v(u-1)} - \left\lfloor \frac{k-2j}{v(u-1)} \right\rfloor \\ &+ (m \perp 2) \sum_{k=1}^{v(u-1)} \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} - \left\lfloor \frac{m-1}{2} + \frac{2k-m-1}{2v(u-1)} \right\rfloor \\ &= \begin{cases} \frac{m}{4}(v(u-1)-1), & \text{even } m \\ \frac{1}{2} \left\lfloor \frac{uv}{2} \right\rfloor + \frac{m+1}{4}(v(u-1)-1), & \text{odd } m, \end{cases} \end{aligned} \quad (36)$$

which is the sum of all the second parameters of the Logbeta distributions in $\Phi_{W,2}(t)$ in (33).

The parameter ν in (34) is then taken as the rate parameter in

$$\Phi^{**}(t) = \theta \nu^{s_1} (\nu - it)^{-s_1} + (1 - \theta) \nu^{s_2} (\nu - it)^{-s_2}$$

where θ, ν, s_1 and s_2 are determined in such a way that

$$\left. \frac{d\Phi_{W,2}(t)}{dt^h} \right|_{t=0} = \left. \frac{d\Phi^{**}(t)}{dt^h} \right|_{t=0} \quad \text{for } h = 1, \dots, 4,$$

while the weights π_ℓ ($\ell = 0, \dots, m^* - 1$) in (34) will then be determined in such a way that

$$\left. \frac{d\Phi_{W,2}(t)}{dt^h} \right|_{t=0} = \left. \frac{d\Phi^*(t)}{dt^h} \right|_{t=0} \quad \text{for } h = 1, \dots, m^*,$$

with $\pi_{m^*} = 1 - \sum_{\ell=0}^{m^*-1} \pi_\ell$.

This procedure yields near-exact distributions for W which will match the first m^* exact moments of W and which have c.f.

$$\Phi_{W,1}(t)\Phi^*(t),$$

with $\Phi_{W,1}(t)$ given by (32) or (33) and $\Phi^*(t)$ by (34), where r , given by (35) or (36) is always either an integer or a half-integer.

As such, the near-exact distributions developed yield, for W , distributions which, for non-integer r , are mixtures, with weights π_ℓ ($\ell = 0, \dots, m^*$), of $m^* + 1$ Generalized Near-Integer Gamma (GNIG) distributions of depth $mu\nu$ with integer shape parameters μ_j ($j = 2, \dots, mu\nu$) and real shape parameter r , in the general case, or shape parameters μ_j^* ($j = 2, \dots, mu\nu$) for the case of $\nu = 1$ or $\nu = 2$, and corresponding rate parameters $(n - j)/n$ ($j = 2, \dots, mu\nu$) and ν , and which, for integer r , are similar mixtures but of Generalized Integer Gamma (GIG) distributions, with the same shape and rate parameters. See Coelho (1998, 2004) and Appendix C in the supplementary material for further details on the GIG and GNIG distributions and their probability density and cumulative distribution functions.

Using the notation in Appendix C in the supplementary material for the probability density and cumulative distribution functions of the GNIG distribution, the near-exact distributions obtained for W , for the general case of $\nu > 2$ and for the case of non-integer r , will have, for $w > 0$, probability density and cumulative distribution functions respectively of the form

$$f_W^*(w) = \sum_{\ell=0}^{m^*} \pi_\ell f^{\text{GNIG}}(w | \underline{s}; \underline{\lambda}; g), \quad \text{and} \quad F_W^*(w) = \sum_{\ell=0}^{m^*} \pi_\ell F^{\text{GNIG}}(w | \underline{s}; \underline{\lambda}; g),$$

where

$$\underline{s} = \{\mu_2, \dots, \mu_{muv}, r + \ell\} \quad \text{and} \quad \underline{\lambda} = \left\{ \frac{n-2}{n}, \dots, \frac{n-muv}{n}, \nu \right\}$$

are respectively the sets of shape and rate parameters and $g = muv$ is the depth of the GNIG distributions, while, for $0 < \lambda < 1$ and $z = -\log \lambda$, the near-exact probability density and cumulative distribution functions of Λ are respectively given by

$$f_{\Lambda}^*(\lambda) = \sum_{\ell=0}^{m^*} \pi_{\ell} f^{\text{GNIG}}\left(z \mid \underline{s}; \underline{\lambda}; g\right) \frac{1}{\lambda}, \quad \text{and} \quad F_{\Lambda}^*(\lambda) = \sum_{\ell=0}^{m^*} \pi_{\ell} \left(1 - F^{\text{GNIG}}\left(z \mid \underline{s}; \underline{\lambda}; g\right)\right). \quad (37)$$

For the case $v = 1$ or $v = 2$ all we have to do is to replace the shape parameters μ_j given by (29) with the shape parameters μ_j^* given by (31) and use r given by (36), instead of r given by (35).

For integer r , all we have to do is to replace the GNIG probability density and cumulative distribution functions with their GIG counterparts (see Appendix C in the supplementary material).

5 Numerical studies

In order to assess the performance of the near-exact distributions developed in the previous section we will use

$$\Delta = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| \frac{\Phi_W(t) - \Phi_{W,1}(t)\Phi^*(t)}{t} \right| dt \quad (38)$$

with

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

as a measure of proximity between the exact and the near-exact distributions, where $\Phi_W(t)$ is the exact c.f. of W in (32) or (33) and $F_W(\cdot)$ and $F_W^*(\cdot)$ represent respectively the exact and near-exact cumulative distribution functions of W , corresponding respectively to $\Phi_W(t)$ and $\Phi_{W,1}(t)\Phi^*(t)$, being $F_{\Lambda}(\cdot)$ and $F_{\Lambda}^*(\cdot)$ the corresponding exact and near-exact cumulative distribution functions of Λ .

In Tables 1, 2, and 3 we may analyze values of Δ for different combinations of values of m , u and v and different sample sizes. For each combination of values of m , u and v , at least three different sample sizes n exceeding the total number of variables muv by 2, 30 and 100 are used. For larger combinations of values of m , u and v , some larger values of n are also used to illustrate the asymptotic behavior of the near-exact distributions in what concerns the sample size. For all near-exact distributions, values of m^* equal to 4, 6 and 10 are used, that is, we use for each case near-exact

Table 1 Values of the measure Δ for the near-exact distributions for $m = 2$

n	m^*			n	m^*		
	4	6	10		4	6	10
	$u = 2, v = 2$				$u = 2, v = 5$		
10	4.23×10^{-12}	9.00×10^{-15}	2.64×10^{-19}	22	7.80×10^{-17}	8.63×10^{-22}	8.26×10^{-30}
38	1.68×10^{-14}	3.06×10^{-18}	9.23×10^{-25}	50	2.39×10^{-17}	2.04×10^{-22}	1.42×10^{-31}
108	9.17×10^{-17}	1.97×10^{-21}	8.66×10^{-30}	120	4.65×10^{-19}	9.74×10^{-25}	5.38×10^{-35}
	$u = 5, v = 2$				$u = 2, v = 10$		
22	1.80×10^{-16}	4.11×10^{-21}	1.39×10^{-29}	42	1.73×10^{-19}	1.88×10^{-25}	1.07×10^{-36}
50	6.28×10^{-17}	6.99×10^{-22}	6.21×10^{-31}	70	4.23×10^{-19}	4.80×10^{-25}	2.73×10^{-36}
120	1.30×10^{-18}	2.97×10^{-24}	1.10×10^{-34}	140	3.16×10^{-20}	1.19×10^{-26}	7.31×10^{-39}
	$u = 5, v = 5$				$u = 5, v = 10$		
52	1.78×10^{-20}	8.05×10^{-27}	5.49×10^{-39}	102	6.18×10^{-24}	1.72×10^{-31}	2.71×10^{-46}
80	6.71×10^{-20}	3.77×10^{-26}	3.56×10^{-38}	130	5.38×10^{-23}	2.67×10^{-30}	1.23×10^{-44}
150	7.77×10^{-21}	1.71×10^{-27}	2.37×10^{-40}	200	2.24×10^{-23}	7.15×10^{-31}	1.29×10^{-45}
250	8.33×10^{-22}	7.33×10^{-29}	1.58×10^{-42}	300	4.93×10^{-24}	8.25×10^{-32}	3.96×10^{-47}
	$u = 10, v = 2$				$u = 10, v = 5$		
42	1.10×10^{-19}	1.32×10^{-25}	8.00×10^{-37}	102	6.53×10^{-24}	1.80×10^{-31}	2.77×10^{-46}
70	2.73×10^{-19}	3.35×10^{-25}	1.79×10^{-36}	130	5.69×10^{-23}	2.80×10^{-30}	1.26×10^{-44}
140	2.06×10^{-20}	8.25×10^{-27}	4.20×10^{-39}	200	2.37×10^{-23}	7.50×10^{-31}	1.32×10^{-45}
240	1.83×10^{-21}	2.69×10^{-28}	1.74×10^{-41}	300	5.22×10^{-24}	8.66×10^{-32}	4.06×10^{-47}
	$u = 10, v = 10$						
202	3.92×10^{-28}	2.22×10^{-36}	1.02×10^{-53}				
230	4.39×10^{-27}	5.35×10^{-35}	9.94×10^{-52}				
300	4.53×10^{-27}	5.48×10^{-35}	8.66×10^{-52}				
400	2.04×10^{-27}	1.82×10^{-35}	1.39×10^{-52}				
500	8.98×10^{-28}	5.88×10^{-36}	2.23×10^{-53}				
1000	4.50×10^{-29}	9.29×10^{-38}	2.99×10^{-56}				

distributions matching 4, 6 and 10 exact moments of W . Smaller values of Δ indicate a closer agreement with the exact distribution and as such, a better performance of the corresponding near-exact distribution.

We may see how the near-exact distributions developed provide very sharp approximations to the exact distribution even for very small samples, that is, for sample sizes hardly exceeding the total number of variables involved. Moreover, they also exhibit clear asymptotic behaviors not only for increasing sample sizes, but also for increasing values of m, u and v . The asymptotic behavior in terms of sample size becoming apparent for larger sample sizes as the values of m, u and v get larger.

A Box asymptotic distribution (Box 1949) is developed for $W = -\log \Lambda$ in Appendix D in the supplementary material. This distribution yields a mixture of two Gamma distributions as the asymptotic distribution for $W = -\log \Lambda$. As such, it should be compared with the near-exact distribution that matches only the first exact moment, that is, the near-exact distribution with $m^* = 1$, since it is for this near-exact

Table 2 Values of the measure Δ for the near-exact distributions for $m = 5$

n	m^*			n	m^*		
	4	6	10		4	6	10
	$u = 2, v = 2$				$u = 2, v = 5$		
22	8.71×10^{-16}	1.35×10^{-19}	5.78×10^{-27}	52	7.45×10^{-21}	7.79×10^{-27}	5.53×10^{-38}
50	2.88×10^{-16}	2.22×10^{-20}	1.28×10^{-27}	80	3.98×10^{-20}	1.16×10^{-25}	7.02×10^{-37}
120	5.50×10^{-18}	8.15×10^{-23}	1.98×10^{-31}	150	1.12×10^{-20}	8.20×10^{-27}	1.09×10^{-38}
				250	1.52×10^{-21}	4.06×10^{-28}	9.06×10^{-41}
				350	3.47×10^{-22}	4.79×10^{-29}	3.05×10^{-42}
	$u = 5, v = 2$				$u = 2, v = 10$		
52	2.14×10^{-19}	2.56×10^{-25}	9.77×10^{-38}	102	3.31×10^{-22}	3.53×10^{-29}	4.93×10^{-43}
80	7.41×10^{-19}	1.11×10^{-24}	9.36×10^{-37}	130	2.76×10^{-21}	5.32×10^{-28}	2.25×10^{-41}
150	7.86×10^{-20}	4.64×10^{-26}	1.32×10^{-38}	200	1.09×10^{-21}	1.36×10^{-28}	2.38×10^{-42}
250	8.06×10^{-21}	1.91×10^{-27}	1.04×10^{-40}	300	2.31×10^{-22}	1.53×10^{-29}	7.36×10^{-44}
350	1.66×10^{-21}	2.09×10^{-28}	3.40×10^{-42}	500	2.43×10^{-23}	6.50×10^{-31}	5.05×10^{-46}
	$u = 5, v = 5$				$u = 5, v = 10$		
127	5.22×10^{-23}	2.70×10^{-30}	8.95×10^{-45}	252	4.03×10^{-26}	1.58×10^{-34}	2.67×10^{-51}
155	5.10×10^{-22}	5.07×10^{-29}	5.83×10^{-43}	280	4.72×10^{-25}	3.93×10^{-33}	2.85×10^{-49}
225	2.91×10^{-22}	2.16×10^{-29}	1.35×10^{-43}	350	6.35×10^{-25}	5.45×10^{-33}	4.19×10^{-49}
450	2.07×10^{-23}	5.13×10^{-31}	3.55×10^{-46}	450	3.70×10^{-25}	2.46×10^{-33}	1.13×10^{-49}
650	3.99×10^{-24}	5.10×10^{-32}	9.27×10^{-48}	1000	1.70×10^{-26}	3.15×10^{-35}	1.11×10^{-52}
	$u = 10, v = 2$				$u = 10, v = 5$		
102	3.18×10^{-22}	3.16×10^{-29}	4.45×10^{-43}	252	4.41×10^{-26}	1.76×10^{-34}	3.14×10^{-51}
130	2.70×10^{-21}	4.83×10^{-28}	2.05×10^{-41}	280	5.16×10^{-25}	4.38×10^{-33}	3.34×10^{-49}
200	1.09×10^{-21}	1.26×10^{-28}	2.21×10^{-42}	350	6.94×10^{-25}	6.09×10^{-33}	4.92×10^{-49}
500	2.49×10^{-23}	6.13×10^{-31}	4.78×10^{-46}	450	4.05×10^{-25}	2.75×10^{-33}	1.32×10^{-49}
				1000	1.86×10^{-26}	3.52×10^{-35}	1.30×10^{-52}
	$u = 10, v = 10$						
502	3.22×10^{-29}	1.01×10^{-38}	9.76×10^{-58}				
530	3.52×10^{-28}	2.35×10^{-37}	9.82×10^{-56}				
600	7.50×10^{-28}	6.22×10^{-37}	3.97×10^{-55}				
1000	3.27×10^{-28}	1.76×10^{-37}	4.67×10^{-56}				
5000	3.02×10^{-31}	9.42×10^{-42}	8.31×10^{-63}				

distribution that the part of the distribution that is approximated is asymptotically approximated with a mixture of two Gamma distributions.

In Tables D.1–D.3 in Appendix D in the supplementary material are displayed the values of the measure Δ in (38) for the Box asymptotic approximation and for the near-exact distributions that match the first $m^* = 1$ and $m^* = 2$ exact moments and we may see how the Box asymptotic distribution cannot match the quality of the approximation provided even by the near-exact distribution that matches the single first exact moment.

Table 3 Values of the measure Δ for the near-exact distributions for $m = 10$

n	m^*			n	m^*		
	4	6	10		4	6	10
$u = 2, v = 2$				$u = 2, v = 5$			
42	1.43×10^{-17}	1.45×10^{-22}	3.20×10^{-32}	102	1.30×10^{-23}	2.68×10^{-31}	1.27×10^{-46}
70	2.03×10^{-17}	2.06×10^{-22}	4.03×10^{-32}	130	8.59×10^{-22}	1.96×10^{-28}	1.48×10^{-41}
140	7.01×10^{-19}	2.26×10^{-24}	4.43×10^{-35}	200	2.98×10^{-22}	4.60×10^{-29}	1.49×10^{-42}
				300	5.74×10^{-23}	4.83×10^{-30}	4.44×10^{-44}
				500	5.53×10^{-24}	1.94×10^{-31}	2.95×10^{-46}
				750	7.93×10^{-25}	1.33×10^{-32}	4.45×10^{-48}
				1000	1.95×10^{-25}	1.90×10^{-33}	2.13×10^{-49}
$u = 5, v = 2$				$u = 2, v = 10$			
102	2.49×10^{-24}	3.43×10^{-31}	2.19×10^{-45}	202	2.19×10^{-26}	5.75×10^{-35}	3.99×10^{-52}
130	1.57×10^{-23}	5.03×10^{-30}	1.02×10^{-43}	230	2.57×10^{-25}	1.40×10^{-33}	4.03×10^{-50}
200	3.21×10^{-24}	1.21×10^{-30}	1.09×10^{-44}	300	2.86×10^{-25}	1.46×10^{-33}	3.73×10^{-50}
300	2.45×10^{-25}	1.30×10^{-31}	3.41×10^{-46}	500	6.25×10^{-26}	1.61×10^{-34}	1.04×10^{-51}
500	1.37×10^{-26}	5.31×10^{-33}	2.34×10^{-48}	1000	3.37×10^{-27}	2.58×10^{-36}	1.49×10^{-54}
$u = 5, v = 5$				$u = 5, v = 10$			
252	6.13×10^{-26}	2.94×10^{-34}	9.95×10^{-51}	502	5.33×10^{-30}	7.80×10^{-40}	1.50×10^{-59}
280	7.19×10^{-25}	7.31×10^{-33}	1.06×10^{-48}	530	5.84×10^{-29}	1.81×10^{-38}	1.51×10^{-57}
350	9.71×10^{-25}	1.02×10^{-32}	1.56×10^{-48}	600	1.25×10^{-28}	4.80×10^{-38}	6.12×10^{-57}
750	8.94×10^{-26}	3.32×10^{-34}	6.30×10^{-51}	1000	5.49×10^{-29}	1.37×10^{-38}	7.25×10^{-58}
1500	4.19×10^{-27}	4.47×10^{-36}	6.99×10^{-54}	2000	3.55×10^{-30}	2.85×10^{-40}	1.55×10^{-60}
$u = 10, v = 2$				$u = 10, v = 5$			
202	1.20×10^{-26}	3.09×10^{-35}	1.91×10^{-52}	502	4.91×10^{-29}	1.64×10^{-38}	2.66×10^{-57}
230	1.37×10^{-25}	7.41×10^{-34}	1.90×10^{-50}	530	5.38×10^{-28}	3.81×10^{-37}	2.67×10^{-55}
300	1.47×10^{-25}	7.51×10^{-34}	1.72×10^{-50}	600	1.15×10^{-27}	1.01×10^{-36}	1.08×10^{-54}
500	3.03×10^{-26}	7.93×10^{-35}	4.68×10^{-52}	1000	5.08×10^{-28}	2.87×10^{-37}	1.27×10^{-55}
1000	1.57×10^{-27}	1.24×10^{-36}	6.54×10^{-55}	2000	3.30×10^{-29}	5.97×10^{-39}	2.70×10^{-58}
				5000	4.77×10^{-31}	1.55×10^{-41}	2.26×10^{-62}
$u = 10, v = 10$							
1002	4.36×10^{-33}	4.29×10^{-44}	4.19×10^{-66}				
1030	4.03×10^{-32}	8.04×10^{-43}	3.14×10^{-64}				
1100	1.03×10^{-31}	2.74×10^{-42}	1.92×10^{-63}				
2000	6.59×10^{-32}	1.25×10^{-42}	4.37×10^{-64}				
5000	1.60×10^{-33}	6.46×10^{-45}	1.03×10^{-67}				

The values used for n, m, u and v in these Tables were the same that were used in Tables 1, 2, and 3. We may notice that some of the values for the measure Δ for the Box asymptotic approximation even go above 1, particularly for cases where the number of variables involved is quite large, and simultaneously the sample size is rather small or not that large. Although this may seem to be something out of the norm, this happens because in these cases the distribution yielded by the Box approximation is not a

legitimate distribution, with its ‘p.d.f.’ and ‘c.d.f.’ going below zero, so that the values reported for the measure Δ which keep being an upper bound between the exact and the asymptotic c.d.f.’s and are indeed correct.

We may also note how the asymptotic Box distribution quickly degrades its quality of approximation as the overall number of variables involved increases, in contrast to what happens with the near-exact distributions.

6 Two real data examples

Example 1: Osteopenia data

To illustrate our proposed testing method, we test the hypothesis (2) on a real data set. The original data consist of bone mineral density values obtained by a technique known as dual X-ray absorptiometry using a GE Lunar Prodigy machine. Measurements were obtained for the femoral neck and the trochanter ($m = 2$), for both the right and left femur ($u = 2$). These four measurements were observed a second time, approximately two years later ($v = 2$). The sample covariance matrix obtained from Roy and Leiva (2011) is matrix A^+ in Appendix E.1 in the supplementary material, which we take as the m.l.e. of Σ .

We see that the variance-covariance matrices (U_0) of the two mineral contents for the femoral neck and trochanter appear very similar for the first as well as for the second year. Also, the covariance matrices (U_1) of the left and right femurs seem to be fairly similar for both years. Finally, the covariance matrices (W) of the two parts of the femur between the two years seem to be similar too. Thus, we will not be much surprised if the hypothesis that the population covariance matrix has a doubly exchangeable covariance structure is not rejected. In fact, as stated in Sect. 2, to test this hypothesis is equivalent to testing the hypothesis in (3). We thus compute the m.l.e. of Σ^* , which is the matrix A in Appendix E.1 in the supplementary material where the orthogonal matrices Γ^\bullet and Γ^* are

$$\Gamma^\bullet = I_v \otimes (C^{*'} \otimes I_m) \quad \text{and} \quad \Gamma^* = C' \otimes I_{mu} = [\gamma_{ij}]$$

where, for $u = 2$,

$$C^* = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \tag{39}$$

and $C = C^*$, given that $v = u = 2$. In particular for $u = 2$ and $m = 2$ the elements γ_{ij} of the symmetric matrix Γ^* are given by

$$\gamma_{ii} = (-1)^{I(i>mu)} 1/\sqrt{2} \quad \text{and} \quad \gamma_{ij} = I(j - i = mu) 1/\sqrt{2} \quad i < j,$$

for all $i, j = 1, \dots, muv$, where $I(\cdot)$ represents the indicator function.

Now, from (10) we have

$$\Lambda = \left(2^4 \frac{|\mathbf{A}|}{|\mathbf{A}_1| |\mathbf{A}_2 + \mathbf{A}_4| |\mathbf{A}_3|} \right)^{n/2},$$

where $\mathbf{A}_1, \dots, \mathbf{A}_4$ denote the four diagonal blocks of dimension 2×2 of \mathbf{A} and where $n = 12$. The computed value of Λ is $\lambda = 4.46637 \times 10^{-12}$.

Using then the near-exact distribution for Λ which matches $m^* = 4$ exact moments, with probability density and cumulative distribution functions given by (37), with $r = 2$, as given by (36), and the other shape parameters μ_j^* ($j = 2, \dots, mu v = 8$) given by (31), with

$$\begin{aligned} r_j^+ &= \{0, 3, 3, 2, 2, 1, 1\} \\ s_j^+ &= \{1, 0, 0, 0, 0, 0, 0\} \\ \mu_j^* &= \{1, 3, 3, 2, 2, 1, 1\} = r_j^+ + s_j^+, \end{aligned}$$

we obtain a near-exact p-value of 0.29468. Thus, we should not reject the null hypothesis that the covariance structure is of the doubly exchangeable type.

In case we had used the common chi-square approximation for the distribution of l.r.t. statistics, we would have $-2 \log \Lambda \overset{a}{\sim} \chi_{\{muv(muv+1)/2\} - \{3m(m+1)/2\}}^2 \equiv \chi_{27}^2$, which would give a p-value of 0.00246. These results show how the chi-square approximation may be completely inadequate for small samples. Indeed, even for quite large samples the chi-square approximation may lead to completely inadequate p-values.

We should note that while the near-exact distribution that we used to compute the near-exact p-value yields a value of 7.73×10^{-12} for the measure Δ in (38), the chi-square approximation yields for this same measure the value of 0.70. These results clearly show that the common chi-square approximation, opposite to the near-exact approach, leads to too many rejections of the null hypothesis, or, equivalently, in general, to too low p-values, clearly inadequate for any practical purposes.

In Fig. 2 we may analyze the plots of the cumulative distribution function for the near-exact distribution for $-\log \Lambda$ and for the $\Gamma(27/2, 1)$ distribution which corresponds to the χ_{27}^2 approximation for $-2 \log \Lambda$, to see how much they differ.

Example 2: Osteoporosis data

See the supplementary material.

7 Discussion and conclusions

We may see how the techniques used to handle the null hypothesis in (2), by first bringing it to the form in (3) and then using the decomposition in (4) enabled the development of very accurate near-exact distributions for the l.r.t. statistic.

From the results of the numerical studies carried out we see that the near-exact distributions developed show an interesting set of nice features. They not only have a good asymptotic behavior for increasing sample sizes, but also an extraordinary per-

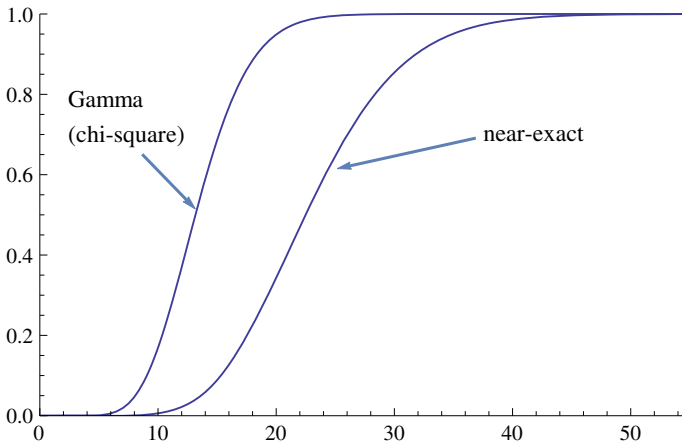


Fig. 2 Plots of the cumulative distribution function for the near-exact distribution for $-\log \Lambda$ and for the $\Gamma(27/2, 1)$ distribution which corresponds to the χ_{27}^2 approximation for $-2 \log \Lambda$

formance for very small sample sizes, as for example for sample sizes exceeding only by two the overall number of variables. Furthermore, opposite to common asymptotic distributions, these near-exact distributions also display a marked asymptotic behavior for increasing values of m , u and v . All these features add up to make the developed near-exact approximations the best choice for practical applications of the test studied.

Moreover, given the discussion in Sections 9.11 and 10.11 of Anderson (2003), the results presented concerning the exact distribution of the l.r.t. statistic as well as the near-exact distributions developed may be made extensive to cases where the vector y has an elliptically contoured distribution.

As future research of interest, stemming out of the work developed in the present paper the authors point out the tests to the hypotheses

$$\begin{aligned}
 &H_0: \mathbf{W} = \mathbf{0}, \text{ assuming } \Sigma = \Theta \text{ (that is, } \Sigma \text{ is doubly-exchangeable)} \\
 &\text{vs} \\
 &H_1: \Sigma = \Theta,
 \end{aligned}$$

which is the test between the doubly exchangeable covariance structure, in H_1 , and the diagonal-block compound symmetry or diagonal-block exchangeable covariance structure, in H_0 , as well as the test to the hypotheses

$$\begin{aligned}
 &H_0: \mathbf{U}_1 = \mathbf{0}, \text{ assuming } \Sigma = \Theta \\
 &\text{vs} \\
 &H_1: \Sigma = \Theta,
 \end{aligned}$$

which is the test between the doubly exchangeable covariance structure, in H_1 , and the multi-block-matrix sphericity covariance structure, in H_0 .

Also the tests to the hypotheses

$$H_0: \mathbf{W} = \mathbf{0}, U_1 = \mathbf{0}, \text{ assuming } \Sigma = \Theta$$

vs

$$H_1: \Sigma = \Theta, \text{ with } U_1 = \mathbf{0}$$

which is the test between the multi-block-matrix sphericity covariance structure, in H_1 , and the block-matrix sphericity covariance structure, in H_0 , and yet the test to the hypotheses

$$H_0: \mathbf{W} = \mathbf{0}, U_1 = \mathbf{0}, \text{ assuming } \Sigma = \Theta$$

vs

$$H_1: \Sigma = \Theta, \text{ with } \mathbf{W} = \mathbf{0}$$

which is the test between the multi-block compound symmetry covariance structure, in H_1 , and the block-matrix sphericity covariance structure, in H_0 , are tests whose development is of interest.

Since all these tests lay upon the assumption of a doubly exchangeable structure, these are a few more reasons why developing a test for the doubly exchangeable structure was an imperative goal. And once these four tests are developed we will be able to test between any two of the covariance structures in Fig. 1.

Acknowledgements This research was partially supported by CMA/FCT/UNL, under projects UID/MAT/00297/2013 and UID/MAT/00297/2019. The second author thanks the support for the summer research Grant from the College of Business at the University of Texas at San Antonio. The authors also want to thank the Editor-in-Chief and the two anonymous reviewers for their careful reading, valuable comments and suggestions that led to a quite improved version of the manuscript.

References

- Anderson TW (2003) An Introduction to multivariate statistical analysis, 3rd edn. Wiley, Hoboken
- Box GEP (1949) A general distribution theory for a class of likelihood criteria. *Biometrika* 36:317–346
- Cardeño L, Nagar DK (2001) Testing block sphericity of a covariance matrix. *Divulg Mat* 9:25–34
- Coelho CA (1998) The generalized integer gamma distribution—a basis for distributions in multivariate statistics. *J Multivar Anal* 64:86–102
- Coelho CA (2004) The generalized near-integer gamma distribution: a basis for ‘near-exact’ approximations to the distributions of statistics which are the product of an odd number of independent Beta random variables. *J Multivar Anal* 89:191–218
- Coelho CA, Arnold BC, Marques FJ (2010) Near-exact distributions for certain likelihood ratio test statistics. *J Stat Theory Pract* 4:711–725
- Coelho CA, Marques FJ (2012) Near-exact distributions for the likelihood ratio test statistic to test equality of several variance-covariance matrices in elliptically contoured distributions. *Comput Statist* 27:627–659
- Coelho CA, Marques FJ (2013) The multi-sample block-scalar sphericity test—exact and near-exact distributions for its likelihood ratio test statistic. *Commun Statist Theory Methods* 42:1153–1175
- Coelho CA, Roy A (2017) Testing the hypothesis of a block compound symmetric covariance matrix for elliptically contoured distributions. *Test* 26:308–330
- Johnson RA, Wichern DW (2007) Applied multivariate statistical analysis, 6th edn. Prentice Hall, Upper Saddle River
- Leiva R, Roy A (2011) Linear discrimination for multi-level multivariate data with separable means and jointly equicorrelated covariance structure. *J Statist Plann Inference* 141:1910–1924

- Leiva R, Roy A (2012) Linear discrimination for three-level multivariate data with separable additive mean vector and doubly exchangeable covariance structure. *Comput Statist Data Anal* 56:1644–1661
- Marques FJ, Coelho CA (2012) Near-exact distributions for the likelihood ratio test statistic of the multi-sample block-matrix sphericity test. *Appl Math Comput* 219:2861–2874
- Marques FJ, Coelho CA, Arnold BC (2011) A general near-exact distribution theory for the most common likelihood ratio test statistics used in multivariate analysis. *Test* 20:180–203
- Moschopoulos P-G (1992) The Hypothesis of multisample block sphericity. *Sankhya* 54:260–270
- Perlman MD (1987) Group symmetry covariance models. *Statist Sci* 2:421–425
- Roy A, Fonseca M (2012) Linear models with doubly exchangeable distributed errors. *Commun Statist Theory Methods* 41:2545–2569
- Roy A, Leiva R (2007) Discrimination with jointly equicorrelated multi-level multivariate data. *Adv Data Anal Classif* 1:175–199
- Roy A, Leiva R (2008) Likelihood ratio tests for triply multivariate data with structured correlation on spatial repeated measurements. *Statist Probab Lett* 78:1971–1980
- Roy A, Leiva R (2011) Estimating and testing a structured covariance matrix for three-level multivariate data. *Comm Statist Theory Methods* 40:1945–1963
- Starnes JW, Neidre DB, Nyman JS, Roy A, Nelson MJ, Gutierrez G, Wang X (2013) Synergistic effect of exercise and statins on femoral strength in rats. *Exp Gerontol* 48:751–755
- Szatrowski TH (1976) Estimation and testing for block compound symmetry and other patterned covariance matrices with linear and nonlinear structure. Tech Rep No. 107, Stanford University
- Szatrowski TH (1982) Testing and estimation in the block compound symmetry problem. *J Edu Statist* 7:3–18
- Tricomi FG, Erdélyi A (1951) The asymptotic expansion of a ratio of Gamma functions. *Pac J Math* 1:133–142
- Votaw D (1948) Testing compound symmetry in a normal multivariate distribution. *Ann Math Statist* 19:447–473
- Wilks SS (1946) Sample criteria for testing equality of means, equality of variances, and equality of covariances in a Normal multivariate distribution. *Ann Math Statist* 17:257–281

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.