26-31 July 2015

60th World Statistics Congress – ISI2015

Special Topics Session 062

Statistical Modeling of Multi-level Multivariate Data Using Kronecker Product Structures in Covariance Matrices

Developing near-exact distributions for likelihood ratio statistics to test for Kronecker product structures in covariance matrices

Coelho, C. A. and Roy, A.

Developing near-exact distributions for likelihood ratio statistics to test for Kronecker product structures in covariance matrices

Carlos A. Coelho*
Departamento de Matemática and Centro de Matemática e Aplicações
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
cmac@fct.unl.pt

Anuradha Roy
Department of Management Science and Statistics
The University of Texas at San Antonio, San Antonio, Texas, U.S.A.
aroy@utsa.edu

Abstract

The authors show how by using an adequate split of null hypotheses of the type $H_0 : V \otimes \Sigma$, where Σ is a positive definite matrix which may or not bear some particular structure like circularity or compound symmetry, into a sequence of nested hypotheses which generate conditionally independent tests, it is possible to easily obtain not only the likelihood ratio test statistics as well as to derive the expressions for their moments. The authors also show how this split of the null hypothesis induces on the characteristic function of the logarithm of the statistic a much useful factorization and how this factorization may be used to lay down the basis for the development of very sharp near-exact distributions for the statistic, which enable an easy computation of very sharp near-exact p-values and quantiles.

Keywords: asymptotic distributions; hypotheses testing; likelihood ratio tests; mixtures of distributions.

1. Introduction

Covariance matrices with a Kronecker product structure are actually almost omnipresent and in recent times they sprout the attention of many researchers. They find applications in a wide range of areas, from time series to studies on microarrays (Teng & Huang, 2009), from data imputation (Allen & Tibshirani, 2010) to magnetoencephalography (Bijma et al., 2005). Many authors have developed and studied inferential processes for such covariance matrices, as for example Lu & Zimmerman (2005), Roy & Khattree (2005), Roy & Leiva (2008, 2011), Allen & Tibshirani (2012). However, finding good approximations for the distributions of the test statistics for large numbers of variables or for small sample sizes, mainly when compared with the number of variables involved, still remains a problem, as Lu & Zimmerman (2005) recognize when they write "Note that when N is relatively small, the null distribution of T is quite different from the limiting χ^2 distribution, and that this disparity increases (relative to the percentiles magnitude) as mn increases".

It is exactly about finding very sharp approximations for the distributions of the test statistics for these situations of very small samples and large numbers of variavles involved that this paper is about.

2. Splitting the null hypotheses

We may split any null hypothesis H_0 into a set of nested null hypotheses as

$$H_0 \equiv H_{0z|a\ldots y} \circ H_{0y|a\ldots x} \circ \ldots \circ H_{0c|ab} \circ H_{0b|a} \circ H_{0a} \tag{1}$$

as long as

$$\Omega_{H_0} \equiv \Omega_{H_{0z|a\ldots y}} \subset \Omega_{H_{1z|a\ldots y}} \equiv \Omega_{H_{0z|a\ldots x}} \subset \ldots \subset \Omega_{H_{1d|abc}} \equiv \Omega_{H_{0c|ab}} \\ \subset \Omega_{H_{1c|ab}} \equiv \Omega_{H_{0b|a}} \subset \Omega_{H_{1b|a}} \equiv \Omega_{0a} \subset \Omega_{H_{1a}} \equiv \Omega_{H_1}$$

where H_{1*} represents the alternative hypothesis to H_{0*} and where $\Omega_{H_{0*}}$ represents the parameter space under the null hypothesis H_{0*} and $\Omega_{H_{1*}}$ the union of the parameter spaces under H_{0*} and H_{1*}. We will not reject

H_0 if we do not reject any of the null hypotheses $H_{0z|a...y}$ through H_{0a} in (1). In (1), the symbol "$\circ$" is to be read as "after", and $H_{0c|ab}$ represents some null hypothesis, assuming $H_{0b|a}$ and H_{0a}, and $H_{0z|a...y}$ represents some null hypothesis, assuming all the other null hypotheses that in (1) appear after this same null hypothesis.

Then, if we consider the l.r.t. (likelihood ratio test) statistic for each of the sub-hypotheses in (1), the l.r.t. statistic to test the overall null hypothesis H_0 will be, by Lemma 10.3.1 in Anderson (2003), the product of these statistics. Moreover, if the choice of these sub-hypotheses is adequately done, it is usually possible to show that the set of the l.r.t. statistics to test the null sub-hypotheses form a set of independent statistics (see Coelho & Marques (2009), Coelho et al. (2010)). These facts give us not only a much simpler way to obtain the l.r.t. statistic for the overall null hypothesis H_0 as well as a much simpler way to obtain an expression for the moments of this overall l.r.t. statistic, then the common formal derivation of the statistic and its moments, since we may then derive the expression for the l.r.t. statistic for the test of the overall null hypothesis H_0 and the expression for its moments from the expressions of the l.r.t. statistics for the null sub-hypotheses and corresponding moments, which are in most cases either already known or much easier to derive than the expressions for the overall l.r.t. statistic and its moments.

Furthermore, if the split of the null hypothesis is done in such a way that the l.r.t. statistics used to test the resulting sub-hypotheses are independent, this split of the null hypothesis is also much useful by generating a factorization of the characteristic function (c.f.) of the logarithm of the overall l.r.t. statistic which will be much useful in developing very sharp near-exact distributions for the l.r.t. statistic itself.

This technique may be successfully applied to the development of likelihood ratio tests for covariance matrices with a Kronecker product structure, as it is shown in Marques et al. (2013).

Since even the exact distributions of the l.r.t. statistics to test the sub-hypotheses are already generally quite elaborate, with the expressions for their exact p.d.f.'s (probability density functions) and c.d.f.'s (cumulative distribution functions) having expressions that are too elaborate to be used in practice, the exact distributions of the l.r.t. statistics to test hypotheses on covariance matrices that have a Kronecker product structure have usually a very elaborate structure, being not possible to use them in practice.

As such, the development of near-exact distributions arises as a desirable goal. Near-exact distributions are asymptotic distributions built using a different approach. Usually working from an adequate factorization of the c.f. of the logarithm of the l.r.t. statistic, we leave unchanged the set of factors that correspond to a manageable distribution and approximate asymptotically the remaining set of factors, in such a way that the resulting c.f., which we will call a near-exact c.f., corresponds to a known manageable distribution, from which p-values and quantiles may be easily computed. These near-exact distributions lie much closer to the exact distribution than any common asymptotic distribution and, when correctly built for statistics used in Multivariate Analysis, will show a marked asymptotic behavior not only for increasing sample sizes but also for increasing number of variables involved.

3. A quite simple example

As a quite simple example of application of the technique described in the previous section we will use the block-matrix sphericity test. This is a test to the null hypothesis

$$H_0 : \Sigma = I_m \otimes \Delta \,, \tag{2}$$

where Δ is a $p{\times}p$ symmetric positive-definite unspecified matrix.

This test is addressed in Marques et al. (2013). We will use the same decomposition of the null hypothesis used in that reference, but we will use a different approach in building the near-exact distributions, which yields even better approximations.

The null hypothesis in (2) may be written as

$$H_{02|1} \circ H_{01}$$

where

$$H_{01} : \Sigma = bdiag(\Sigma_1, \dots, \Sigma_k)\,, \qquad (\text{with } \Sigma_k \text{ of dimensions } p{\times}p \ (k = 1, \dots, m))\,,$$

is the null hypothesis of the test of independence of m sets of p variables each, and

$$H_{02|1} : \begin{array}{c} \Sigma_1 = \dots = \Sigma_m \\ \text{assuming } H_{01} \end{array}$$

is the null hypothesis of the test of equality of m covariance matrices.

Let us assume we have a sample of size n from the distribution of $\underline{X} \sim N_{p^*}(\underline{\mu}, \Sigma)$, where $p^* = mp$. Then, based on Lemma 10.3.1 in Anderson (2003), the l.r.t. statistic to test H_0 in (2) will be the product of the l.r.t. statistics $\Lambda_{2|1}$ and Λ_1, used to test respectively $H_{02|1}$ and H_{01}, which is then the statistic

$$\Lambda = \underbrace{\frac{|A|^{n/2}}{\prod_{k=1}^{m}|A_k|^{n/2}}}_{\Lambda_1} \underbrace{\frac{\prod_{k=1}^{m}|A_k|^{n/2}}{\left|\frac{1}{m}\,A^*\right|^{nm/2}}}_{\Lambda_{2|1}} = \frac{|A|^{n/2}}{\left|\frac{1}{m}\,A^*\right|^{nm/2}} \tag{3}$$

where A is the maximum likelihood estimator (m.l.e.) of Σ, A_k its k-th diagonal block of dimensions $p{\times}p$, and $A^* = \sum_{k=1}^{m} A_k$.

4. Building near-exact distributions for the l.r.t. statistic

Then, if we take $W = -\log\Lambda$, $W_1 = -\log\Lambda_1$ and $W_2 = -\log\Lambda_{2|1}$, given the independence of Λ_1 and $\Lambda_{2|1}$, which is easy to prove, based on Lemma 10.4.1 in Anderson (2003) and Theorem 5 in Jensen (1988), we may write the c.f. of W as (see further details in Marques et al. (2013))

$$\begin{aligned}
\Phi_W(t) \;&=\; E\left(e^{\mathrm{i}tW}\right) = E\left(e^{\mathrm{i}t(W_1+W_2)}\right) = E\left(e^{\mathrm{i}tW_1}\right)E\left(e^{\mathrm{i}tW_2}\right) = E\left(\Lambda_1^{-\mathrm{i}t}\right)E\left(\Lambda_{2|1}^{-\mathrm{i}t}\right)\\
&= \underbrace{\left\{\prod_{j=2}^{p^*-1}\left(\frac{n-j}{n}\right)^{r_j^*}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-r_j^*}\right\}}_{\Phi_{W_1,1}(t)} \underbrace{\left\{\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-1}{2}-\frac{n}{2}\mathrm{i}t\right)}{\Gamma\left(\frac{n}{2}-\frac{n}{2}\mathrm{i}t\right)\Gamma\left(\frac{n-1}{2}\right)}\right\}^{m^*}}_{\Phi_{W_1,2(t)}}\\
&\quad\times \underbrace{\left\{\prod_{j=1}^{p-1}\left(\frac{n-j}{n}\right)^{r_j}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-r_j}\right\}}_{\Phi_{W_2,1}(t)} \underbrace{\left\{\prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{m}\frac{\Gamma(a_j+b_{jk})\,\Gamma(a_j+b_{jk}^*-n\mathrm{i}t)}{\Gamma(a_j+b_{jk}^*)\,\Gamma(a_j+b_{jk}-n\mathrm{i}t)}\right\} \times\left(\prod_{k=1}^{m}\frac{\Gamma(a_p+b_{pk})\,\Gamma\left(a_p+b_{pk}^*-\frac{n}{2}\mathrm{i}t\right)}{\Gamma(a_p+b_{pk}^*)\,\Gamma\left(a_p+b_{pk}-\frac{n}{2}\mathrm{i}t\right)}\right)^{p\perp\!\!\!\perp 2}}_{\Phi_{W_2,2}(t)}\\
&= \underbrace{\left\{\prod_{j=1}^{p-1}\left(\frac{n-j}{n}\right)^{v_j}\left(\frac{n-j}{n}-\mathrm{i}t\right)^{-v_j}\right\}}_{\Phi_1(t)}\;\Phi_2(t)
\end{aligned} \tag{4}$$

where $\Phi_2(t) = \Phi_{W_1,2}(t)\Phi_{W_2,2}(t)$, the shape parameters r_j^* and r_j are respectively given by (3.7)–(3.9) in Coelho et al. (2010) and (5)–(7) in Coelho & Marques (2012),

$$p\perp\!\!\!\perp 2 = \begin{cases} 1 & \text{if } p \text{ is odd}\\ 0 & \text{if } p \text{ is even}\,,\end{cases} \qquad v_j = \begin{cases} r_j\,, & j=1\\ r_j^*+r_j\,, & j=2,\dots,p-1\,,\end{cases} \qquad m^* = \lfloor m/2\rfloor(p\perp\!\!\!\perp 2)$$

and the parameters a_j, b_{jk}, b_{jk}^*, a_p, b_{pk} and b_{pk}^* are give by (3) and (4) in Coelho & Marques (2012).

In (4) $\Phi_1(t)$ is the c.f. of a GIG (Generalized Integer Gamma) distribution (Coelho, 1998) of depth $p-1$ and $\Phi_2(t)$ is the c.f. of a sum of independent Logbeta r.v.'s multiplied by $n/2$ or by n.

Then, following the approach in Coelho et al. (2010), in order to build the near-exact distributions we will leave $\Phi_1(t)$ untouched and we will asymptotically replace $\Phi_2(t)$ by the c.f. of a finite mixture of say $m^{**}+1$ Gamma distributions with shape parameters $r+j$, $j=0,\dots,m^{**}$, with

$$r = \frac{m^*}{2} + \sum_{j=1}^{\lfloor p/2\rfloor}\sum_{k=1}^{m}(b_{jk}-b_{jk}^*) + \left(\sum_{k=1}^{m}(b_{pk}-b_{pk}^*)\right)^{p\perp\!\!\!\perp 2}\,, \tag{5}$$

with a rate parameter λ^* numerically computed following the procedure described in Coelho et al. (2010), and with the weights defined in such a way that the resulting near-exact c.f. equates the first m^{**} exact moments of W. This will yield for Λ near-exact cumulative distribution functions (c.d.f.'s) of the form

$$F^*_\Lambda(\ell) = \sum_{j=0}^{m^{**}} \pi_j \left(1 - F^{GNIG}\left(-\log \ell \,\middle|\, v_1, \dots, v_{p-1}; r+j; \frac{n-1}{2}, \dots, \frac{n-p+1}{2}; \lambda^*; p\right)\right), \tag{6}$$

where $F^{GNIG}(\,\cdot\,)$ denotes the c.d.f. of the GNIG distribution (Coelho, 2004), and which when r in (5) is an integer will be mixtures of GIG distributions.

These near-exact distributions yield very sharp approximations to the exact distribution, even for very small sample sizes, and they are asymptotic not only for increasing sample sizes but also for increasing numbers of variables involved, that is, for increasing values of m and p, as it may be seen from the values of the measure Δ^* in Table 1, being this asymptotic behavior more clear for near-exact distributions that equate more exact moments. The measure Δ^* used in Table 1 is the measure

$$\Delta^* = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| \frac{\Phi_W(t) - \Phi^*_W(t)}{t} \right| dt\,, \qquad \text{with} \qquad \max_w |F_W(w) - F^*_W(w)| \,\le\, \Delta^*\,, \tag{7}$$

where $\Phi_W(t)$ represents the exact c.f. of W and $\Phi^*_W(t)$ the near-exact c.f., and $F_W(\,\cdot\,)$ represents the exact c.d.f. of W and $F^*_W(\,\cdot\,)$ the near-exact c.d.f..

Table 1. – Values of the measure Δ^* in (7) for the near-exact distributions for the statistic Λ in (3), for increasing values of m, p and $n = mp + 1, 100, 500$, and for values of $m^{**} = 2, 4, 6, 10$.

			m^{**}			
m	p	n	2	4	6	10
4	3	13	2.65×10^{-8}	5.82×10^{-12}	2.01×10^{-15}	6.62×10^{-22}
		112	8.90×10^{-11}	7.30×10^{-16}	6.44×10^{-21}	7.92×10^{-31}
		512	8.17×10^{-13}	3.43×10^{-19}	1.47×10^{-25}	4.01×10^{-38}
4	7	29	3.38×10^{-11}	4.69×10^{-16}	4.51×10^{-21}	7.49×10^{-31}
		128	1.17×10^{-11}	2.30×10^{-17}	4.33×10^{-23}	2.51×10^{-34}
		528	2.13×10^{-13}	2.55×10^{-20}	3.05×10^{-27}	7.24×10^{-41}
8	3	25	5.94×10^{-10}	8.57×10^{-15}	1.75×10^{-19}	1.94×10^{-28}
		124	4.14×10^{-11}	1.11×10^{-16}	3.41×10^{-22}	6.21×10^{-33}
		524	5.70×10^{-13}	9.40×10^{-20}	1.74×10^{-26}	1.12×10^{-39}
8	7	57	2.07×10^{-12}	9.22×10^{-19}	5.02×10^{-25}	2.59×10^{-37}
		156	1.13×10^{-12}	4.44×10^{-19}	1.80×10^{-25}	4.00×10^{-38}
		556	2.65×10^{-14}	1.09×10^{-21}	4.34×10^{-29}	8.88×10^{-44}

5. Adding structure to the matrices Δ in (2)

One of the key points about the approach followed is that with this approach it is then very easy to add some structure to the matrices Δ in (2), that is, to further test some structure on these matrices, like circularity or compound symmetry. If for example one would like to further test if the matrices Δ are circular (see Olkin & Press (1969)), then all we have to do is to add a further null hypothesis

$$H_{03|1,2}: \ \begin{array}{l} \Delta \text{ is circular} \\ (\text{assuming } H_{02|1} \text{ and } H_{01}) \end{array}$$

and then the l.r.t. statistic to test $H_0 \equiv H_{03|1,2} \circ H_{02|1} \circ H_{01}$ would be

$$\Lambda = \Lambda_{3|1,2}\Lambda_{2|1}\Lambda_1\,, \tag{8}$$

where $\Lambda_{3|1,2}$ is the l.r.t. statistic to test $H_{03|1,2}$. For the expression of this statistic see Olkin & Press (1969) and Marques & Coelho (2013). Once again it is possible, by using the two results referred before, to show the independence of the three l.r.t. statistics, by showing that $\Lambda_{3|1,2}$ is independent of both $\Lambda_{2|1}$ and Λ_1. Then, the c.f. of $W = -\log \Lambda$ might be written as

$$\Phi_W(t) = \underbrace{\Phi_1(t)\,\Phi_{W_{3,1}}(t)}_{\Phi^*_1(t)}\;\underbrace{\Phi_2(t)\,\Phi_{W_{3,2}}(t)}_{\Phi^*_2(t)} \tag{9}$$

where $\Phi_1(t)$ and $\Phi_2(t)$ are the c.f.'s in (4) and $\Phi_{W_3,1}(t)$ and $\Phi_{W_3,2}(t)$ are the two factors of the c.f. of $W_3 = -\log \Lambda_{3|1,2}$, given in Marques & Coelho (2013). For odd p, $\Phi_{W_3,1}(t)$ is given by (6) in that reference, while $\Phi_{W_3,2}(t)$ is equal to 1, and for even p, $\Phi_{W_3,1}(t)$ and $\Phi_{W_3,2}(t)$ are given by (10) in the same reference.

In (9), $\Phi_1^*(t)$ is again the c.f. of a GIG distribution, and $\Phi_2^*(t)$ the c.f. of a sum of independent Logbeta r.v.'s. As such, in order to develop near-exact distributions for Λ we will keep $\Phi_1^*(t)$ unchanged and replace asymptotically $\Phi_2^*(t)$ by a finite mixture of Gamma r.v.'s with shape parameters $r^* + j$ $(j = 0, \ldots, m^{**})$, with r^* equal to r in (5) if p is odd, and $r^* = r + 1/2$ if p is even. These near-exact distributions will yield near-exact c.d.f.'s for Λ with a structure similar to that of the c.d.f. in (6). We only have to be aware of the fact that we will have to use a pooled estimator for the m.l.e. of Δ, once assumed $H_{02|1}$ and H_{01}, actually based on nm observations. This pooled estimator will be the matrix A^*.

As such, it is clear that this approach for developing near-exact distributions may lead to even sharper approximations for more elaborate tests, since for example for odd p it is clear that for this new test where we also consider testing the matrices Δ for circularity, more is going into the part of the c.f. of W that is left unchanged, while the part to be asymptotically approximated remains the same as before. As such, the resulting near-exact distributions will lie even closer to the exact distribution than the ones developed for the test with null hypothesis given by (2). We may thus obtain even sharper approximations for more elaborate tests which is usually quite uncommon and unexpected.

In Table 2 may be analyzed the values of the measure Δ^* in (7) for this test

Table 2. – Values of the measure Δ^* in (7) for the near-exact distributions for the statistic Λ in (8), for increasing values of m, $p = 7$ and $n = mp + 1, 100, 500$, and for values of $m^{**} = 2, 4, 6, 10$.

			m^{**}			
m	p	n	2	4	6	10
4	7	29	$3.77{\times}10^{-20}$	$5.23{\times}10^{-25}$	$5.03{\times}10^{-30}$	$8.37{\times}10^{-40}$
		128	$1.55{\times}10^{-24}$	$3.05{\times}10^{-30}$	$5.75{\times}10^{-36}$	$3.33{\times}10^{-47}$
		528	$5.57{\times}10^{-30}$	$6.68{\times}10^{-37}$	$7.99{\times}10^{-44}$	$1.90{\times}10^{-57}$
8	7	57	$5.54{\times}10^{-25}$	$2.46{\times}10^{-31}$	$1.34{\times}10^{-37}$	$6.92{\times}10^{-50}$
		156	$7.01{\times}10^{-28}$	$2.75{\times}10^{-34}$	$1.11{\times}10^{-40}$	$2.48{\times}10^{-53}$
		556	$7.91{\times}10^{-33}$	$3.26{\times}10^{-40}$	$1.30{\times}10^{-47}$	$2.65{\times}10^{-62}$

6. Conclusions

The method used of splitting the null hypothesis into a set of conditionally independent sub-hypotheses and then using the induced factorization of the c.f. of the logarithm of the l.r.t. statistic to build near-exact distributions for the l.r.t. statistic itself is most suited for tests on covariance matrices with a Kronecker product structure, and the more elaborate are the structures studied, the more convenient this method seems to be, giving in many situations even better results for more elaborate hypotheses.

Other tests on Kronecker product covariance structures that may be addressed in a similar way are the tests to the hypotheses of double complete symmetry

$$H_0 : \Sigma = (I_p \otimes R_0) + (J_p - I_p) \otimes R_1\,, \quad \text{with} \quad R_0 = aI_m + b(J_m - I_m) \quad \text{and} \quad R_1 = cJ_m \tag{10}$$

where J_n represents an $n{\times}n$ matrix of 1's, and

$$-\frac{a}{m-1} < b < a \quad \text{and} \quad -\frac{1}{p-1}\left(b + \frac{a-b}{m}\right) < c < b + \frac{a-b}{m}\,,$$

the hypotheses of block compound symmetry, a structure used by Roy & Leiva (2011),

$$H_0 : \Sigma = I_u \otimes (\Sigma_0 - \Sigma_1) + J_u \otimes \Sigma_1 \tag{11}$$

where $\Sigma_0 - \Sigma_1$ and $\Sigma_0 + (u-1)\Sigma_1$ are positive-definite matrices, or yet, the double exchangeable covariance structure, used by Roy & Fonseca (2012),

$$H_0 : \Sigma = I_{uv} \otimes (U_0 - U_1) + I_v \otimes J_u \otimes (U_1 - W) + J_{uv} \otimes W \tag{12}$$

where, $U_0 - U_1$, $(U_0 - U_1) + u(U_1 - W)$ and $(U_0 - U_1) + u(U_1 - W) + uvW$ are definite-positive matrices, and where the tests in (10) and (11) may be seen as particular cases of the test in (12).

Acknowledgements

This research was supported by FCTFundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology), project UID/MAT/00297/2013, through Centro de Matemática e Aplicações (CMA/FCT/UNL)

References

Allen, G. I. & Tibshirani, R. (2010). Transposable regularized covariance models with an application to missing data imputation. *Annals of Applied Statistics*, 4, 764–790.

Allen, G. I. & Tibshirani, R. (2012). Inference with transposable data: modelling the effects of row and column correlations. *J. R. Statist. Soc. B*, 74, Part 4, .

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

Bijma, F., de Munck, J. C. & Heethaar, R. M. (2005). The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products. *NeuroImage*, 27, 402–415.

Coelho, C.A.(1998). The Generalized Integer Gamma distribution - a basis for distributions 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., Arnold, B. C., & Marques, F. J. (2010). Near-exact distributions for certain likelihood ratio test statistics. *Journal of Statistical Theory and Practice*, 4, 711–725.

Coelho, C. A., & Marques, F. J. (2009). The advantage of decomposing elaborate hypotheses on covariance matrices into conditionally independent hypotheses in building near-exact distributions for the test statistics. *Linear Algebra and its Applications*, 430, 2592–2606.

Coelho, C. A., & Marques, F. J. (2012). 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.

Jensen, S. T. (1988). Covariance hypotheses which are linear in both the covariance and the inverse covariance. *Ann. Statist.*, 16, 302–322.

Lu, N., & Zimmerman, D. L. (2005). The likelihood ratio test for a separable covariance matrix. *Statistics & Probability Letters*, 73, 449–457.

Marques, F. J., & Coelho, C. A. (2013). Obtaining the exact and near-exact distributions of the likelihood ratio statistic to test circular symmetry through the use of characteristic functions. *Computational Statistics*, 28, 2091–2115.

Marques, F. J., Coelho, C. A., & Marques, P. (2013). The Block-Matrix Sphericity Test: Exact and Near-exact Distributions for the Test Statistic, in *Recent Developments in Modeling and Applications in Statistics, Studies in Theoretical and Applied Statistics: Selected Papers of the Statistical Societies*, Oliveira, P. E., Temido, M. G., Henriques, C. and Vichi, M. (eds.), Springer, 169–177.

Olkin I, & Press S. J. (1969). Testing and estimation for a circular stationary model. *Ann. Math. Stat.*, 40, 1358–1373.

Roy, A., & Fonseca, M. (2012). Linear Models with Doubly Exchangeable Distributed Errors, *Communications in Statistics - Theory and Methods*, 41, 2545–2569.

Roy, A., & Khattree, R. (2005). On implementation of a test for Kronecker product covariance structure for multivariate repeated measures data. *Statist. Methodol.*, 2, 297–306.

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, *Communications in Statistics - Theory and Methods*, 40, 1945–1963.

Teng, S. L. & Huang, H. (2009). A statistical framework to infer functional gene relationships from biologically interrelated microarray experiments. *JASA*, 104, 465–473.