Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 430 (2009) 2592–2606
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

The advantage of decomposing elaborate hypotheses on covariance matrices into conditionally independent hypotheses in building near-exact distributions for the test statistics

Carlos A. Coelho*, Filipe J. Marques

Mathematics Department, Faculty of Sciences and Technology, The New University of Lisbon, Quinta da Torre, 2829-516 Caparica, Portugal

Received 30 November 2007; accepted 11 July 2008
Available online 2 September 2008
Submitted by G. Trenkler

Abstract

The aim of this paper is to show how the decomposition of elaborate hypotheses on the structure of covariance matrices into conditionally independent simpler hypotheses, by inducing the factorization of the overall test statistic into a product of several independent simpler test statistics, may be used to obtain near-exact distributions for the overall test statistics, even in situations where asymptotic distributions are not available in the literature and adequately fit ones are not easy to obtain.

Keywords: Near-exact distributions; Conditionally independent hypotheses; Likelihood ratio test statistics; Generalized sphericity tests; Generalized near-integer Gamma distribution; Mixtures

1. Introduction

The concept of a near-exact distribution has already been introduced in a number of papers [6,7,11,1,15,16,9,22]. In a nutshell, near-exact distributions are new asymptotic distributions that lay closer to the exact distribution than the usual asymptotic distributions. They correspond to what we call near-exact c.f.s (characteristic functions) which are c.f.s obtained from the exact c.f.

* Corresponding author. Tel.: +351 21 294 8388; fax: +351 21 294 8391.
E-mail addresses: cmac@fct.unl.pt (C.A. Coelho), fjm@fct.unl.pt (F.J. Marques).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2008.07.003

by leaving most of it unchanged and replacing the remaining smaller part by an asymptotic result (which is intended to be asymptotic both in terms of sample size and overall number of variables, in a manner that will be more precisely stated in this section). This replacement is done in such a way that the resulting c.f. corresponds to a distribution with a manageable c.d.f. (cumulative distribution function) so that the computation of near-exact quantiles is rendered feasible, easy and precise.

There are mainly two ways in which near-exact distributions, or rather, near-exact c.f.s may be obtained. One is when the exact distribution is an infinite mixture, in which case we may cut short the series that corresponds to the exact c.f. at a given number of terms and replace the remainder by the c.f. of just one, or a mixture of two or three distributions of the same kind of the ones in the infinite mixture. The parameters in these replacing parts of the c.f. are computed in such a way that the first few derivatives at the origin match the derivatives of the part being replaced. More precisely, if the exact c.f. of the statistic W may be written as

$$\Phi_W(t) = \sum_{i=0}^{\infty} p_i\, \Phi_i(t),$$

where $p_i > 0$ $(i = 0, 1, \ldots)$ and $\sum_{i=0}^{\infty} p_i = 1$, and where $\Phi_i(t)$ are c.f.s, then we will use instead of $\Phi_W(t)$

$$\Phi_W^*(t) = \sum_{i=0}^{n^*} p_i\, \Phi_i(t) + \theta\, \Phi_2(t),$$

where $\theta = 1 - \sum_{i=0}^{n^*} p_i$ and $\Phi_2(t)$ is usually either a c.f. of the type of $\Phi_i(t)$ or the c.f. of the mixture of two or three such c.f.s, defined in such a way that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t^h}\, \theta\, \Phi_2(t)\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t^h} \sum_{i=n^*+1}^{\infty} p_i\, \Phi_i(t)\right|_{t=0}$$

for $h \in H$ (where usually we will have $H = \{1, 2\}$, $H = \{1, \ldots, 4\}$ or $H = \{1, \ldots, 6\}$).

Another way to obtain a near-exact c.f. is through the factorization of the exact c.f. If we may write

$$\Phi_W(t) = \Phi_1(t)\, \Phi_2(t),$$

where we recognize $\Phi_1(t)$ as corresponding to a known and manageable distribution and $\Phi_2(t)$ as corresponding to a non-manageable distribution, then we will use as a near-exact c.f. for W the c.f.

$$\Phi_W^*(t) = \Phi_1(t)\, \Phi_2^*(t), \tag{1}$$

where for $h \in H$ (usually with $H = \{1, 2\}$, $H = \{1, \ldots, 4\}$ or $H = \{1, \ldots, 6\}$)

$$\left.\frac{\mathrm{d}}{\mathrm{d}t^h}\, \Phi_2^*(t)\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t^h}\, \Phi_2(t)\right|_{t=0} \tag{2}$$

in such a way that, if we write $\Phi_2^*(t)$ as a function of the sample size (say n) and the overall number of variables involved (say p), that is, if we write $\Phi_2^*(t) \equiv \Phi_2^*(t; n, p)$, we want

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} \left| \frac{\Phi_2(t) - \Phi_2^*(t; n, p)}{t} \right| \mathrm{d}t = 0 \quad \text{and} \quad \lim_{p\to\infty} \int_{-\infty}^{+\infty} \left| \frac{\Phi_2(t) - \Phi_2^*(t; n, p)}{t} \right| \mathrm{d}t = 0$$

with

$$\int_{-\infty}^{+\infty}\left|\frac{\Phi_2(t)-\Phi_2^*(t;n,p)}{t}\right|\mathrm{d}t \geqslant \int_{-\infty}^{+\infty}|\Phi_1(t)|\left|\frac{\Phi_2(t)-\Phi_2^*(t;n,p)}{t}\right|\mathrm{d}t$$
$$\geqslant \max_w \left|F_W(w)-F_W^*(w;n,p)\right|, \qquad (3)$$

where $F_W(\cdot)$ is the exact c.d.f. of W and $F_W^*(\cdot\,;n,p)$ is the near-exact c.d.f. of W, corresponding to the c.f. $\Phi_W^*(t)$ in (1). The relation in (3) will be used at the end of Section 3 in the definition of a measure of proximity between distributions (see expression (6)).

When dealing with l.r.t. (likelihood ratio test) statistics, mainly those more commonly used in Multivariate Statistics, $\Phi_1(t)$ is the c.f. of a sum of independent Logbeta r.v.s (random variables), that are r.v.s whose exponential has a Beta distribution. This c.f. may be alternatively written as the c.f. of a GIG (Generalized Integer Gamma) distribution, that is, the distribution of the sum of a given number of independent Gamma distributions, all with different rate parameters and integer shape parameters (see [5]), while $\Phi_2(t)$ is the c.f. of a sum of other independent Logbeta r.v.s, which is not possible to be written under the form of the c.f. of a GIG distribution and which will be asymptotically replaced by the c.f. of a single Gamma distribution or the c.f. of a mixture of two or three Gamma distributions with the same rate parameters. This replacement was already used in [22] and is indeed well justified since a single Logbeta distribution may be represented under the form of an infinite mixture of Exponential distributions (see [8]). A sum of independent Logbeta random variables may thus be represented under the form of an infinite mixture of sums of Exponential distributions, which are themselves mixtures of Exponential or Gamma distributions.

The near-exact distributions we are interested in this paper are exactly the ones of this second kind.

2. How may the decomposition of a complex hypothesis into more elementary hypotheses help in building near-exact distributions for the overall test statistic?

Let us suppose we have Λ, the l.r.t. statistic to test a given null hypothesis H_0, and that we want to write, for $W = -\log\Lambda$

$$\Phi_W(t) = \Phi_1(t)\,\Phi_2(t)$$

(the reason why we usually want to handle the c.f. of $W = -\log\Lambda$ instead of the c.f. of Λ is that while the moments of Λ may be relatively easy to obtain and they commonly exist for any positive integer order, and even for any order not necessarily integer just above a given negative value, the expression for the c.f. of Λ may be too hard to obtain and to handle; while on the other hand once obtained a near-exact distribution for W it will then be easy to obtain the corresponding near-exact distribution for $\Lambda = \mathrm{e}^{-W}$).

It may happen that this factorization may be too hard to obtain from scratch, given the complexity of Λ itself.

But, let us suppose we may write

$$\Lambda = \prod_{j=1}^{m}\Lambda_j, \qquad (4)$$

where Λ_j is the l.r.t. statistic to test $H_{0j|1,\ldots,j-1}$, the jth nested conditionally independent null hypothesis we may split H_0 into. That is, we are assuming we may write

$$H_0 \equiv H_{0m|1,\ldots,m-1} \circ \cdots \circ H_{03|1,2} \circ H_{02|1} \circ H_{01},$$

which means that to test H_0 is equivalent to test sequentially the m independent hypotheses $H_{0j|1,\ldots,j-1}$ $(j = 1, \ldots, m)$, starting with H_{01} and testing then $H_{02|1}$ and then $H_{03|1,2}$, and so on, and where to test $H_{0j|1,\ldots,j-1}$ is to test H_{0j} assuming that H_{01} through $H_{0,j-1}$ are true. The hypotheses $H_{0j|1,\ldots,j-1}$ $(j = 1, \ldots, m)$ are all independent in the sense that under the null hypothesis H_0 it is possible to prove that the Λ_j in (4) are independent. This is the case with all the likelihood ratio statistics used and mentioned in this paper (see [4,18]), although in some non-regular cases independence may not hold as [17] shows. Let us further suppose that, for

$$W_j = -\log \Lambda_j \quad (j = 1, \ldots, m),$$

we have available the factorizations

$$\Phi_{W_j}(t) = \Phi_{j1}(t)\,\Phi_{j2}(t)\,. \tag{5}$$

Then, under H_0, given the independence of the Λ_j and thus also of the W_j, we may easily write

$$\begin{aligned}\Phi_W(t) &= \prod_{j=1}^{m} \Phi_{W_j}(t)\\ &= \underbrace{\left\{\prod_{j=1}^{m} \Phi_{j1}(t)\right\}}_{\Phi_1(t)} \underbrace{\left\{\prod_{j=1}^{m} \Phi_{j2}(t)\right\}}_{\Phi_2(t)}\\ &= \Phi_1(t)\,\Phi_2(t),\end{aligned}$$

where for the most common l.r.t. statistics used in Multivariate Statistics the c.f.s $\Phi_{j1}(t)$ may be obtained under the form of c.f.s of GIG distributions [6,7,22,9] and all the $\Phi_{j2}(t)$ are c.f.s of the sum of independent Logbeta r.v.s. Then, $\Phi_1(t)$ itself will be the c.f. of a GIG distribution, and $\Phi_2(t)$ will be itself the c.f. of the sum of independent Logbeta r.v.s, being thus adequately asymptotically replaced by the c.f. of a single Gamma distribution or the c.f. of the mixture of two or three Gamma distributions, verifying (2) (see [22]) and yielding this way either a GNIG (Generalized Near-Integer Gamma) distribution or a mixture of two or three GNIG distributions (the GNIG distribution is the distribution of the sum of a r.v. with a GIG distribution with an independent r.v. with a Gamma distribution with a non-integer shape parameter – for details on this distribution see [7]).

3. Examples of application

Since decompositions of the c.f.s of the type in (5) are already available for the Wilks Λ statistic, or the l.r.t. statistic to test the independence of several sets of variables (see [6,7]), and also for the l.r.t. statistic to test sphericity (see [22]) and for the l.r.t. statistic to test the equality of several variance–covariance matrices (see [9]), we may think of l.r.t.s whose statistic may be factorized as in (4) and where the Λ_j are the above mentioned statistics.

This was indeed what was somehow done when obtaining either the exact distribution for the generalized Wilks Λ statistic, under the form of a GIG distribution, when at most one of the sets of variables has an odd number of variables (see [5]) or when obtaining a near-exact distribution for the same statistic, under the form of a GNIG distribution, for the general case when two or more sets have an odd number of variables (see [6,7]).

In the subsections ahead we will use the following notation:

- $\Lambda_1(q, p; N_j)$, with $\Lambda_1(q, p; N_j) = \frac{\prod_{j=1}^{q} |A_j|^{N_j/2}}{|A|^{N/2}} \frac{N^{Np/2}}{\prod_{j=1}^{q} N_j^{N_j p/2}}$, to denote the l.r.t. statistic used to test $H_{01} : \Sigma_1 = \cdots = \Sigma_q$, based on samples of size N_j $(j = 1, \ldots, q)$ from $N_p(\underline{\mu}_j, \Sigma_j)$, with $A_j = \hat{\Sigma}_j$, $A = A_1 + \cdots + A_q$ and $N = N_1 + \cdots + N_q$;
- $\Lambda_2(N; p_1, \ldots, p_k)$, with $\Lambda_2(N; p_1, \ldots, p_k) = \left(\frac{|A|}{\prod_{i=1}^{k} |A_{ii}|}\right)^{N/2}$, to denote the l.r.t. statistic used to test $H_{02} : \Sigma = \mathrm{diag}(\Sigma_{11}, \ldots, \Sigma_{ii}, \ldots, \Sigma_{kk})$, based on a sample of size N from $N_p(\underline{\mu}, \Sigma)$, with $A = \hat{\Sigma}$, $A_{ii} = \hat{\Sigma}_{ii}$ and

$$\underline{\mu} = \left[\underline{\mu}_1, \ldots, \underline{\mu}_i, \ldots, \underline{\mu}_k\right]', \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \cdots & \Sigma_{1i} & \cdots & \Sigma_{1k} \\ \vdots & \ddots & \vdots & & \vdots \\ \Sigma_{i1} & \cdots & \Sigma_{ii} & \cdots & \Sigma_{ik} \\ \vdots & & \vdots & \ddots & \vdots \\ \Sigma_{k1} & \cdots & \Sigma_{ki} & \cdots & \Sigma_{kk} \end{bmatrix};$$

- $\Lambda_3(p; N)$, with $\Lambda_3(p; N) = \left(\frac{|A|}{\left(\mathrm{tr}\,\frac{1}{p} A\right)^p}\right)^{N/2}$, to denote the l.r.t. statistic used to test $H_{03} : \Sigma = \sigma^2 I_p$, based on a sample of size N from $N_p(\underline{\mu}, \Sigma)$, with $A = \hat{\Sigma}$.

To evaluate the proximity of the approximations obtained, to the exact distribution, we will use the measure

$$\Delta = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| \frac{\Phi_W(t) - \Phi_W^*(t)}{t} \right| \mathrm{d}t, \tag{6}$$

where $\Phi_W(t)$ and $\Phi_W^*(t)$ represent respectively the exact and the approximate (near-exact) c.f. of W, the logarithm of the l.r.t. statistic under study.

This measure of proximity has already been used in [15,22] and its values are an upper bound on the absolute value of the difference between the corresponding c.d.f.s, with

$$\Delta \geqslant \max_{x \in S_W} |F_W(x) - F_W^*(x)|, \tag{7}$$

where S_W is the support of the r.v. W and $F_W(x)$ and $F_W^*(x)$ represent respectively the exact c.d.f. of W and the approximate c.d.f. corresponding to $\Phi_W^*(t)$. This measure is related with the Berry–Esseen bound (see [3,14,19,21, Chapter VI, Section 21]).

3.1. The test of equality of several multivariate Normal distributions

It may be the best well known example of the situation we are trying to illustrate, since we may write in this case [2,20,25]

$$H_0 : \underline{\mu}_1 = \cdots = \underline{\mu}_q, \quad \Sigma_1 = \cdots = \Sigma_q \tag{8}$$

with

$$H_0 \equiv H_{02|1} \circ H_{01},$$

where

$$H_{01} : \Sigma_1 = \cdots = \Sigma_q$$

and

$$H_{02|1} : \quad \begin{array}{l} \underline{\mu}_1 = \cdots = \underline{\mu}_q \\ \text{given that } \Sigma_1 = \cdots = \Sigma_q, \end{array}$$

so that, using the notation defined at the beginning of this section, we may write the l.r.t. statistic to test H_0 in (8) as

$$\Lambda = \Lambda_2(N; p, q-1)\Lambda_1(q, p; N_j) \quad (\text{with } N = N_1 + \cdots + N_q), \tag{9}$$

where p is the order of the matrices $\Sigma_1, \ldots, \Sigma_q$ and where, under H_0 in (8), $\Lambda_2(N; p, q-1)$ and $\Lambda_1(q, p; N_j)$ are independent.

Since factorizations of the type in (5) for the c.f. of $-\log\,\Lambda_2(N; p, q-1)$ and $-\log\,\Lambda_1(q, p; N_j)$ are available [6,7,9], the process of obtaining near-exact distributions for the l.r.t. statistic for this test may be quite easily implemented by using the approach proposed in this paper.

For $N_1 = \cdots = N_q = n$, and $\alpha = 1, \ldots, \lfloor p/q \rfloor$, the exact c.f. of $W = -\log\,\Lambda$, where Λ is the l.r.t. statistic in (9), may be written either as (see [2,20,25])

$$\begin{aligned} \Phi_W(t) = & \prod_{j=1}^{p} \frac{\Gamma\left(\frac{nq-j}{2}\right)\Gamma\left(\frac{(n-1)q+1-j}{2} - \frac{nq}{2}it\right)}{\Gamma\left(\frac{nq-j}{2} - \frac{nq}{2}it\right)\Gamma\left(\frac{(n-1)q+1-j}{2}\right)} \\ & \times \prod_{j=1}^{p}\prod_{k=1}^{q} \frac{\Gamma\left(\frac{n-1}{2} + \frac{1-j}{2q} + \frac{k-1}{q}\right)\Gamma\left(\frac{n-j}{2} - \frac{n}{2}it\right)}{\Gamma\left(\frac{n-1}{2} + \frac{1-j}{2q} + \frac{k-1}{q} - \frac{n}{2}it\right)\Gamma\left(\frac{n-j}{2}\right)} \end{aligned} \tag{10}$$

or as

$$\begin{aligned} \Phi_W(t) = & \left.\begin{array}{l} \displaystyle\prod_{\substack{j=2\\ j\neq\alpha q}}^{p+q-1} \left(\frac{n-j/q}{n}\right)^{r_j}\left(\frac{n-j/q}{n} - it\right)^{-r_j} \\ \displaystyle\times \prod_{j=2}^{p}\left(\frac{n-j}{n}\right)^{u_j}\left(\frac{n-j}{n} - it\right)^{-u_j} \end{array}\right\} \Phi_1(t) \\ & \left.\begin{array}{l} \displaystyle\times\left(\frac{\Gamma\left(\frac{nq-1}{2}\right)\Gamma\left(\frac{nq-2}{2} - \frac{nq}{2}it\right)}{\Gamma\left(\frac{nq-1}{2} - \frac{nq}{2}it\right)\Gamma\left(\frac{nq-2}{2}\right)}\right)^{p\perp\!\!\!\perp 2} \\ \displaystyle\times \prod_{j=1}^{\lfloor p/2\rfloor}\prod_{k=1}^{q} \frac{\Gamma(a_j + b_{jk})\,\Gamma(a_j + b^*_{jk} - nit)}{\Gamma(a_j + b^*_{jk})\,\Gamma(a_j + b_{jk} - nit)} \\ \displaystyle\times\left(\prod_{k=1}^{q} \frac{\Gamma(a_p + b_{pk})}{\Gamma(a_p + b^*_{pk})}\,\frac{\Gamma\left(a_p + b^*_{pk} - \frac{n}{2}it\right)}{\Gamma\left(a_p + b_{pk} - \frac{n}{2}it\right)}\right)^{p\perp\!\!\!\perp 2} \end{array}\right\} \Phi_2(t) \end{aligned} \tag{11}$$

for suitably defined shape parameters r_j $(j = 2, \ldots, p+q-1;\ j \neq \alpha q)$ and u_j $(j = 2, \ldots, p)$ (see [23], also for parameters $a_j, b_{jk}, a_p,\ b_{pk},\ b^*_{jk}$ and b^*_{pk}).

But then, according to the explanation in the previous Sections 1 and 2, we may replace $\Phi_2(t)$ in (11) by

$$\Phi^*_2(t) = \sum_{k=1}^{h/2} p_k\,\lambda^{s_k}(\lambda - it)^{-s_k} \tag{12}$$

with

$$\sum_{k=1}^{h/2} p_k = 1, \quad \text{and} \quad 0 < p_k \leqslant 1;\ k = 1, \ldots, h/2, \tag{13}$$

where $h = 2$, $h = 4$ or $h = 6$, is the number of derivatives of $\Phi_2(t)$ at $t = 0$ that will be equated by the corresponding derivatives of $\Phi_2^*(t)$ at $t = 0$. This way we obtain

$$\Phi_W(t) = \left.\begin{array}{l} \prod\limits_{\substack{j=2\\ j\neq \alpha q}}^{p+q-1} \left(\frac{n-j/q}{n}\right)^{r_j} \left(\frac{n-j/q}{n} - it\right)^{-r_j} \\ \times \prod\limits_{j=2}^{p} \left(\frac{n-j}{n}\right)^{u_j} \left(\frac{n-j}{n} - it\right)^{-u_j} \end{array}\right\} \Phi_1(t) \\ \times \Phi_2^*(t) \tag{14}$$

as a near-exact c.f. for W, yielding thus near-exact distributions for W which are, according to the number of moments of $\Phi_2(t)$ being equated and the concomitant definition of $\Phi_2^*(t)$, either a GNIG distribution or a mixture of two or three GNIG distributions of depth $2p + q - 2 - \lfloor p/q \rfloor$, with integer shape parameters r_j $(j = 2, \ldots, p + q - 1;\ j \neq \alpha q)$ and u_j $(j = 2, \ldots, p)$ and non-integer shape parameter s_k, for the kth mixture term $(k = 1, \ldots, h/2)$. These near-exact distributions will be respectively designated in the tables by the acronyms GNIG, M2GNIG and M3GNIG.

Once obtained a well-fit near-exact distribution for W under the form of a GNIG distribution or a mixture of GNIG distributions, the corresponding c.d.f.s are easily available (see [7] and also [22,23]), being the corresponding near-exact c.d.f.s for Λ very easily derived by simple exponentiation of the r.v. W. From these c.d.f.s the computation of near-exact quantiles is quick and easy, when using one of the several available softwares able to handle extended precision computations.

The computation of near-exact quantiles is not carried out in this subsection since this would only be a matter of interest in order to compare them with the exact ones, but the computation of these quantiles with enough precision (at least 12–16 decimal places) would be a difficult and time consuming task, almost rendered useless, given the very good approximations that the near-exact distributions developed above exhibit, which is possible to be assessed through the computation of the values of the measure Δ using the c.f.s in (10) and (14) for different values of n, p and q. To establish a comparison, we have also used the asymptotic Box-style distribution for $W = -\log \Lambda$ in [2, Section 10.5]. Results are shown in Tables 1–6.

The results in Tables 1–6 show not only the extreme closeness to the exact distribution that all the near-exact distributions exhibit, with, as expected, the M3GNIG distribution displaying always the best fit and the single GNIG distribution a less good fit, although always much better than the asymptotic distribution. Moreover, the near-exact distributions display not only an asymptotic behaviour for increasing sample sizes but also an asymptotic behaviour for increasing values of p and q, which is a much desirable feature. Apart from a couple of values which came out a little too good for the M3GNIG distribution, as for example for $n = 20$, $p = 10$ and $q = 5$ and for $n = 13$, $p = 11$ and $q = 5$, the evolution of the values of the measure Δ is pretty steady. Indeed we should bear in mind that for either even p or odd q we have the exact distribution for the part of the l.r.t. statistic which is a Wilks Lambda statistic, given under the form of a GIG distribution.

Also, even for very small sample sizes, that is, for sample sizes which are very close to the number of variables being used, all the near-exact distributions show a very good performance,

Table 1
Values of Δ for the asymptotic and near-exact distributions for the l.r.t statistic of equality of q multivariate Normal distributions, for $n=20$, $p=10$

	$q=4$	$q=5$	$q=6$	$q=7$
Asymp.	3.61×10^{-2}	3.49×10^{-2}	3.64×10^{-2}	3.92×10^{-2}
GNIG	1.19×10^{-6}	6.13×10^{-7}	4.61×10^{-7}	2.89×10^{-7}
M2GNIG	3.39×10^{-10}	8.16×10^{-11}	5.72×10^{-11}	2.69×10^{-11}
M3GNIG	1.08×10^{-13}	2.56×10^{-15}	6.82×10^{-15}	2.86×10^{-15}

Table 2
Values of Δ for the asymptotic and near-exact distributions for the l.r.t statistic of equality of q multivariate Normal distributions, for $n=30$, $p=10$

	$q=4$	$q=5$	$q=6$	$q=7$
Asymp.	1.06×10^{-2}	8.98×10^{-3}	8.30×10^{-3}	8.10×10^{-3}
GNIG	5.95×10^{-7}	3.16×10^{-7}	2.37×10^{-7}	1.50×10^{-7}
M2GNIG	8.08×10^{-11}	1.94×10^{-11}	1.46×10^{-11}	6.94×10^{-12}
M3GNIG	1.26×10^{-14}	1.88×10^{-15}	9.62×10^{-16}	4.60×10^{-16}

still with a clear asymptotic behaviour for increasing values of p and q, even in situations where the asymptotic distribution breaks down (indeed not being a distribution any more).

3.2. The sphericity test

The sphericity test itself, whose null hypothesis may be written as

$$H_0:\ \Sigma=\sigma^2 I_p \tag{15}$$

may be seen as another example of a l.r.t. whose null hypothesis may be written as the composition of two conditionally independent null hypotheses (see Chapter 10, Section 10.7.3 in [2]), since we may indeed write H_0 in (15) once again as

$$H_0\equiv H_{02|1}\circ H_{01},$$

where

$$H_{01}:\Sigma=\mathrm{diag}(\sigma_1^2,\ldots,\sigma_p^2)$$

and

$$H_{02|1}:\quad \begin{array}{l}\sigma_1^2=\cdots=\sigma_p^2\\ \text{given that } \Sigma=\mathrm{diag}(\sigma_1^2,\ldots,\sigma_p^2),\end{array}$$

so that, using the notation defined at the beginning of this section, we may write the l.r.t. statistic to test H_0 in (15) as

$$\Lambda=\Lambda_2(n;\underbrace{1,\ldots,1}_{p})\Lambda_1(p,1;n),$$

where under H_0 in (15), $\Lambda_2(n;1,\ldots,1)$ and $\Lambda_1(p,1;n)$ are independent (see for example [2, Section 10.7.3]).

Although we may question the usefulness of this approach, moreover, given that very accurate near-exact distributions have already been obtained for the sphericity l.r.t. statistic in [22], the

Table 3
Values of Δ for the asymptotic and near-exact distributions for the l.r.t statistic of equality of q multivariate Normal distributions, for $n=50,\ p=10$

	$q=4$	$q=5$	$q=6$	$q=7$
Asymp.	2.94×10^{-3}	2.34×10^{-3}	1.98×10^{-3}	1.77×10^{-3}
GNIG	2.25×10^{-7}	1.22×10^{-7}	9.17×10^{-8}	5.86×10^{-8}
M2GNIG	1.13×10^{-11}	2.50×10^{-12}	2.15×10^{-12}	1.04×10^{-12}
M3GNIG	9.04×10^{-16}	7.62×10^{-16}	9.53×10^{-17}	5.42×10^{-17}

Table 4
Values of Δ for the asymptotic and near-exact distributions for the l.r.t statistic of equality of q multivariate Normal distributions, for $n=12,\ p=10$

	$q=4$	$q=5$	$q=6$	$q=7$
Asymp.	4.03×10^{-1}	4.73×10^{-1}	5.44×10^{-1}	6.12×10^{-1}
GNIG	1.47×10^{-6}	6.83×10^{-7}	5.11×10^{-7}	3.13×10^{-7}
M2GNIG	6.85×10^{-10}	1.44×10^{-10}	9.47×10^{-11}	4.33×10^{-11}
M3GNIG	4.27×10^{-13}	2.70×10^{-14}	1.90×10^{-14}	6.92×10^{-15}

same authors were able to obtain even more accurate near-exact distributions for moderately large values of p for the same statistic through the use of this approach [10]. These near-exact distributions, besides having a simpler formulation, are even more accurate than the previously developed ones namely for larger numbers of variables ($p \geqslant 10$).

The l.r.t. statistic to test H_0 in (15) is, for a sample of size n [24,2]

$$\Lambda = \left(\frac{|A|}{\left(\operatorname{tr} \frac{1}{p} A\right)^p} \right)^{n/2}, \tag{16}$$

where A is the maximum likelihood estimator of Σ.

The exact c.f. of $W = -\log \Lambda$ may be written as (see [2,20,25])

$$\Phi_W(t) = \prod_{j=2}^{p} \frac{\Gamma\left(\frac{n-1}{2} + \frac{j-1}{p}\right) \Gamma\left(\frac{n-j}{2} - \frac{n}{2}\,it\right)}{\Gamma\left(\frac{n-1}{2} + \frac{j-1}{p} - \frac{n}{2}\,it\right) \Gamma\left(\frac{n-j}{2}\right)} \tag{17}$$

or, alternatively, using the approach outlined in Section 2, as (see [10])

$$\begin{aligned}\Phi_W(t) = &\underbrace{\prod_{j=2}^{p} \left(\frac{n-j}{n}\right)^{r_j} \left(\frac{n-j}{n} - it\right)^{-r_j}}_{\Phi_1(t)} \\ &\times \left\{ \prod_{j=1}^{p-k^*} \frac{\Gamma\left(\frac{n-1}{2} + \frac{j-1}{p}\right) \Gamma\left(\frac{n-1}{2} - \frac{n}{2}it\right)}{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{n-1}{2} + \frac{j-1}{p} - \frac{n}{2}it\right)} \right\} \\ &\underbrace{\times \left\{ \prod_{j=p-k^*+1}^{p} \frac{\Gamma\left(\frac{n-1}{2}+\frac{j-1}{p}\right) \Gamma\left(\frac{n}{2} - \frac{n}{2}it\right)}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n-1}{2}+\frac{j-1}{p} - \frac{n}{2}it\right)} \right\}}_{\Phi_2(t)},\end{aligned} \tag{18}$$

where $k^* = \lfloor p/2 \rfloor$ and $r_j = \lfloor \frac{p-j+2}{2} \rfloor$ $(j = 2, \ldots, p)$.

Table 5
Values of Δ for the asymptotic and near-exact distributions for the l.r.t statistic of equality of q multivariate Normal distributions, for $n=14$, $p=12$

	$q=4$	$q=5$	$q=6$	$q=7$
Asymp.	5.33×10^{-1}	6.27×10^{-1}	7.17×10^{-1}	7.99×10^{-1}
GNIG	8.58×10^{-7}	4.64×10^{-7}	2.63×10^{-7}	2.03×10^{-7}
M2GNIG	2.45×10^{-10}	8.71×10^{-11}	2.71×10^{-11}	1.87×10^{-11}
M3GNIG	7.85×10^{-14}	2.02×10^{-14}	2.21×10^{-15}	1.79×10^{-15}

Table 6
Values of Δ for the asymptotic and near-exact distributions for the l.r.t statistic of equality of q multivariate Normal distributions, for $n=19$, $p=17$

	$q=4$	$q=5$	$q=6$	$q=7$
Asymp.	8.45×10^{-1}	9.84×10^{-1}	1.11×10^{0}	1.22×10^{0}
GNIG	4.44×10^{-7}	2.64×10^{-7}	2.03×10^{-7}	1.09×10^{-7}
M2GNIG	8.23×10^{-11}	3.38×10^{-11}	2.23×10^{-11}	7.87×10^{-12}
M3GNIG	1.68×10^{-14}	5.26×10^{-15}	3.08×10^{-15}	7.55×10^{-16}

Then, if in (18) we replace $\Phi_2(t)$ by a c.f. $\Phi_2^*(t)$ of the type of the one in (12), using the specification in (13), we will obtain as near-exact distributions for W a GNIG distribution or a mixture of two or three GNIG distributions of depth p with integer shape parameters r_j and non-integer shape parameter s_k, for the kth mixture term ($k = 1, \dots, h/2$). As in the previous subsection, these near-exact distributions will be respectively designated, namely in the tables, by the acronyms GNIGnew, M2GNIGnew and M3GNIGnew (being the acronyms GNIG, M2GNIG and M3GNIG used in Tables 8 and 9 for the corresponding near-exact distributions obtained for the same statistic in [22], without using the decomposition approach proposed in this paper). We will compare these near-exact distributions with the exact distribution as well as in terms of quantiles as through the use of the measure Δ in (6).

Consul [12] obtained representations for the exact p.d.f. of the $(2/n)$th power of the l.r.t. statistic in (16) for $p = 3, 4$ and 6 through the use of the Gauss hypergeometric function (as a side note, there is only a minor typo in expression (3.14), where the constant term should be $K(n)/9!$ instead of $2\,K(n)/9!$), leading to expressions that yield easy computations when using one of the common symbolic softwares with extended precision. However, a common expression for general p, was only obtained by the same author, in [13], where he makes use of the Meijer G-function. But, since the Meijer G-function is indeed only an alternative representation for a Mellin–Barnes type of integral, its computation is still quite heavy, even when using an extended precision software. Such representations for the p.d.f. only allow for the computation of quantiles for the $(2/n)$th power of the l.r.t. statistic with the required precision (that is, up to sixteen decimal places of precision, in order to be able to compare them with the near-exact ones) for moderate values of p ($p \leqslant 10$).

We have computed exact quantiles for the $(2/n)$th power of the l.r.t. statistic in (16) for $p = 5, 6, 7, 8, 10$ and given values of n and we compare them with the near-exact quantiles. The results are presented in Table 7, where a measure

$$\delta = -\log_{10} |\text{exact quantile} - \text{approx. quantile}|$$

Table 7
Quantiles 0.05 and 0.01 for the $(2/n)$th power of the l.r.t. statistic in (16)

	Quantile			
	0.05	δ	0.01	δ
$n = 7,\ p = 5$				
Exact	0.0012622014344466		0.0002183943236075	
Box	0.0018598053836777	3.2	0.0004370905088500	3.7
GNIGnew	0.0012621657432522	7.4	0.0002183808232977	7.9
M2GNIGnew	0.0012622015298382	10.0	0.0002183943197572	11.4
M3GNIGnew	0.0012622014349577	12.3	0.0002183943234605	12.8
$n = 13,\ p = 6$				
Exact	0.0365289320027240		0.0187996936542321	
Box	0.0371562902726480	3.2	0.0193651668489626	3.2
GNIGnew	0.0365288722277811	7.2	0.0187995858186483	7.0
M2GNIGnew	0.0365289321616112	9.8	0.0187996937407390	10.1
M3GNIGnew	0.0365289320026219	13.0	0.0187996936542503	13.7
$n = 9,\ p = 7$				
Exact	0.0001472413497881		0.0000242384697502	
Box	0.0003324977520427	3.7	0.0000879672231885	4.2
GNIGnew	0.0001472404530786	7.2	0.0000242381703782	7.0
M2GNIGnew	0.0001472413507666	9.8	0.0000242384696727	10.1
M3GNIGnew	0.0001472413497887	13.0	0.0000242384697498	13.7
$n = 14,\ p = 8$				
Exact	0.0056126167409998		0.0024755558596933	
Box	0.0060583304947613	3.4	0.0027924174017203	3.5
GNIGnew	0.0056126113751765	8.3	0.0024755489107065	8.2
M2GNIGnew	0.0056126167475432	11.2	0.0024755558612765	11.8
M3GNIGnew	0.0056126167409703	13.5	0.0024755558596937	15.4
$n = 16,\ p = 10$				
Exact	0.0012140286105857		0.0005064717865027	
Box	0.0014135955350563	3.7	0.0006334878610630	3.9
GNIGnew	0.0012140280320840	9.2	0.0005064711210440	9.2
M2GNIGnew	0.0012140286109532	12.4	0.0005064717870345	12.3
M3GNIGnew	0.0012140286105719	13.9	0.0005064717869780	12.3

is used to have a more clear and extended assessment of the number of decimal places matched by the approximate quantiles. In order to establish a comparison, we use Box's asymptotic approximation (see [2, Section 10.7.4]).

We may see how all the near-exact distributions show a very good behaviour for both quantiles, when compared with the Box asymptotic approximation, mainly for the smaller sample sizes (that is, for the smaller differences between n and p). Only for $p = 10$ there seems to be a decline in the performance of the M3GNIG near-exact distribution, but the smaller than expected value for δ may be indeed due to a lack of precision in the computation of the exact quantile.

In Tables 8 and 9, we may examine the values of the measure Δ in (6), computed using the exact c.f. in (17) and the near-exact distributions GNIG, M2GNIG and M3GNIG in [22], and the near-exact distributions GNIGnew, M2GNIGnew and M3GNIGnew, whose c.f.s are derived from (18) by replacing $\Phi_2(t)$ by c.f.s of the type of $\Phi_2^*(t)$ in (12).

Table 8
Values of Δ for the near-exact distributions for the l.r.t. test statistic of sphericity for several values of n and p

	$p = 10, n = 12$	$p = 12, n = 14$	$p = 14, n = 16$
GNIG	1.03×10^{-6}	5.02×10^{-7}	2.30×10^{-7}
M2GNIG	2.79×10^{-10}	6.18×10^{-11}	1.86×10^{-11}
M3GNIG	2.20×10^{-13}	1.74×10^{-14}	4.82×10^{-15}
GNIGnew	1.20×10^{-7}	5.09×10^{-8}	2.48×10^{-8}
M2GNIGnew	6.59×10^{-11}	1.60×10^{-11}	4.86×10^{-12}
M3GNIGnew	1.23×10^{-14}	2.60×10^{-15}	6.11×10^{-16}

Table 9
Values of Δ for the near-exact distributions for the l.r.t. test statistic of sphericity for several values of n and p

	$p = 10, n = 50$	$p = 12, n = 52$	$p = 14, n = 54$
GNIG	1.94×10^{-7}	1.41×10^{-7}	8.81×10^{-8}
M2GNIG	1.25×10^{-11}	3.84×10^{-12}	2.61×10^{-12}
M3GNIG	1.06×10^{-14}	2.54×10^{-15}	1.26×10^{-15}
GNIGnew	2.66×10^{-8}	1.62×10^{-8}	1.06×10^{-8}
M2GNIGnew	8.90×10^{-12}	3.73×10^{-12}	1.76×10^{-12}
M3GNIGnew	2.69×10^{-15}	7.82×10^{-16}	2.71×10^{-16}

We may see how the near-exact distributions developed using the decomposition technique described in this paper have an even better behaviour than the previously developed and already well-fit near-exact distributions, both for small and large sample sizes, displaying a good asymptotic behaviour both for increasing sample size and increasing number of variables.

3.3. Generalized sphericity tests

Other tests to which the authors intend to apply the approach proposed in this paper are extended versions of the sphericity test, which we may call the 'multi-sample scalar-block sphericity test' and the 'multi-sample matrix-block sphericity test'.

For the multi-sample scalar-block sphericity test we have

$$H_0 : \Sigma_1 = \cdots = \Sigma_q = \begin{bmatrix} \sigma_1^2 I_{p_1} & & 0 \\ & \ddots & \\ 0 & & \sigma_k^2 I_{p_k} \end{bmatrix} \tag{19}$$

with

$$H_0 \equiv H_{03|1,2} \circ H_{02|1} \circ H_{01},$$

where

$$\begin{aligned} &H_{01} : \Sigma_1 = \cdots = \Sigma_q, \\ &H_{02|1} : \quad \Sigma = \mathrm{diag}(\Sigma_{11}, \ldots, \Sigma_{kk}) \\ &\qquad\quad \text{given that } \Sigma_1 = \cdots = \Sigma_q = \Sigma, \end{aligned}$$

where Σ is of dimensions $p\times p$, while Σ_{ii} $(i = 1, \ldots, k)$ of dimensions $p_i \times p_i$ is the ith diagonal block of Σ, with $p = \sum_{i=1}^{k} p_i$, and

$$H_{03|1,2}:\quad \begin{array}{l}\Sigma_{ii}=\sigma_i^2 I_{p_i}\quad \text{for } i=1,\ldots,k\\ \text{given that } \Sigma=\text{diag}(\Sigma_{11},\ldots,\Sigma_{kk})\quad\text{and}\\ \text{given that } \Sigma_1=\cdots=\Sigma_q=\Sigma,\end{array}$$

so that we may write the l.r.t. statistic to test H_0 in (19), for $N=N_1+\cdots+N_q$ as

$$\Lambda=\Lambda_1(q,p;N_j)\Lambda_2(N;p_1,\ldots,p_k)\prod_{i=1}^{k}\Lambda_3(p_i;N)\,. \tag{20}$$

Since, under H_0 in (19), it is possible to prove the independence of all the statistics in (20) and since factorizations of the type in (5) are available for the c.f.s of the logarithms of all the statistics in (20) [6,7,9,22], the process of obtaining near-exact distributions for the l.r.t. statistic for this test may be implemented by using the approach proposed in this paper.

Concerning the multi-sample matrix-block sphericity test, its null hypothesis may be written as

$$H_0:\ \Sigma_1=\cdots=\Sigma_q=\Delta\otimes I_k\left(=\begin{bmatrix}\Delta & & 0\\ & \ddots & \\ 0 & & \Delta\end{bmatrix}\right) \tag{21}$$

with

$$H_0\equiv H_{03|1,2}\circ H_{02|1}\circ H_{01},$$

where

$$H_{01}:\ \Sigma_1=\cdots=\Sigma_q\ (=\Sigma),$$
$$H_{02|1}:\quad \begin{array}{l}\Sigma=\text{diag}(\Sigma_{11},\ldots,\Sigma_{kk})\\ \text{given that } \Sigma_1=\cdots=\Sigma_q\ (=\Sigma),\end{array}$$

where Σ is of dimensions $p\times p$, while Σ_{ii} $(i=1,\ldots,k)$ of dimensions $p^*\times p^*$ is the ith diagonal block of Σ, with $p=kp^*$, and

$$H_{03|1,2}:\quad \begin{array}{l}\Sigma_{11}=\cdots=\Sigma_{kk}(=\Delta)\\ \text{given that } \Sigma=\text{diag}(\Sigma_{11},\ldots,\Sigma_{kk})\quad\text{and}\\ \text{given that } \Sigma_1=\cdots=\Sigma_q=\Sigma,\end{array}$$

so that we may write the l.r.t. statistic to test H_0 in (21), for $N=N_1+\cdots+N_q$ as

$$\Lambda=\Lambda_1(q,p;N_j)\Lambda_2(N;\underbrace{p^*,\ldots,p^*}_{k})\Lambda_1(k,p^*;N)\,. \tag{22}$$

Since, under H_0 in (21), it is possible to prove the independence of all the statistics in (22) and since factorizations of the type in (5) are available for the c.f.s of the logarithms of all the statistics in (22) [6,7,22], the process of obtaining near-exact distributions for the l.r.t. statistic for this test may once again be implemented by using the approach proposed in this paper.

4. Final remarks

The use of the decomposition approach developed in this paper not only enables us to build very accurate and manageable near-exact approximations to the exact distribution of the overall test statistics but also concomitantly enables us to easily overcome the problems of controlling statistical errors (in particular the error of the first kind), which arise when we have to test sequentially the partial hypotheses. Now we may easily compute near-exact quantiles which enable us to carry the overall test in just one step, avoiding this way the problems brought to our attention by [17] and avoiding also the need for any corrections of the first kind error level, like Sidak's correction [26,27].

We may even plan to go beyond the tests presented in the previous section, namely those in Section 3, with virtually no limit, by adequately nesting the three elementary multivariate likelihood ratio tests (referred at the beginning of the previous section). In every case we will still be able to obtain near-exact distributions for the overall test statistics, in situations where the construction of well-fit asymptotic distributions, using the usual techniques, is too complicated or even virtually impossible.

Acknowledgments

This research was financially supported by the Portuguese Foundation for Science and Technology, through the Center for Mathematics and its Applications (CMA) from the New University of Lisbon.

The authors would like also to express their appreciation for the insightful comments and suggestions made by the referees.

References

[1] R.P. Alberto, C.A. Coelho, Study of the quality of several asymptotic and near-exact approximations based on moments for the distribution of the Wilks Lambda statistic, J. Statist. Plann. Inference, 137 (2007) 1612–1626.
[2] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, third ed., J. Wiley & Sons, New York, 2003
[3] A. Berry, The accuracy of the Gaussian approximation to the sum of independent variates, Trans. Amer. Math. Soc. 49 (1941) 122–136.
[4] D. Basu, On statistics independent of a complete sufficient statistics, Sankhya 15 (1955) 377–380.
[5] C.A. Coelho, The generalized integer gamma distribution – a basis for distributions in multivariate statistics, J. Multivariate Anal. 64 (1998) 86–102.
[6] C.A. Coelho, The generalized integer gamma distribution as an asymptotic replacement for the Logbeta random variable – applications, Amer. J. Math. Manage. Sci. 23 (2003) 383–399.
[7] C.A. Coelho, 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. Multivariate Anal. 89 (2004) 191–218.
[8] C.A. Coelho, R.P. Alberto, L.M. Grilo, A mixture of generalized integer gamma distributions as the exact distribution of the product of an odd number of independent Beta random variables, J. Interdiscip. Math. 9 (2006) 229–248.
[9] C.A. Coelho, F.J. Marques, Near-exact distributions for the likelihood ratio test statistic for testing equality of several variance–covariance matrices, The New University of Lisbon, Mathematics Department, Technical Report #12/2007, submitted for publication.
[10] C.A. Coelho, F.J. Marques, Near-exact distributions for the independence and sphericity likelihood ratio test statistics, The New University of Lisbon, Mathematics Department, Technical Report #20/2007, submitted for publication.
[11] C.A. Coelho, J.T. Mexia, On the exact and near-exact distributions of statistics used in generalized F tests, The New University of Lisbon, Mathematics Department, Technical Report #23/2006, submitted for publication.
[12] P.C. Consul, On the exact distributions of the criterion W for testing sphericity in a p-variate normal distribution, Ann. Math. Statist. 38 (1967) 1170–1174.
[13] P.C. Consul, The exact distributions of likelihood criteria for different hypotheses, in: P.R. Krishnaiah (Ed.), Proceedings of the Second International Symposium on Multivariate Analysis, Dayton, Ohio, vol. II, Academic Press, New York, 1969, pp. 171–181.
[14] C.-G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law, Acta Math. 77 (1945) 1–125.

[15] L.M. Grilo, C.A. Coelho, Development and study of two near-exact approximations to the distribution of the product of an odd number of independent Beta random variables, J. Statist. Plann. Inference 137 (2007) 1560–1575.
[16] L.M. Grilo, C.A. Coelho, The exact and near-exact distributions for the Wilks Lambda statistic used in the test of independence of two sets of variables, The New University of Lisbon, Mathematics Department, Technical Report #15/2007, submitted for publication.
[17] R.V. Hogg, On the resolution of statistical hypotheses, J. Amer. Statist. Assoc. 56 (1961) 978–989.
[18] R.V. Hogg, A.T. Craig, Sufficient statistics in elementary distribution theory, Sankhya 17 (1956) 209–216.
[19] H.-K. Hwang, On convergence rates in the central limit theorems for combinatorial structures, Eur. J. Combin. 19 (1998) 329–343.
[20] A.M. Kshirsagar, Multivariate Analysis, Marcel Dekker Inc., New York, 1972.
[21] M. Loève, Probability Theory, vol. I, fourth ed., Springer-Verlag, New York, 1977.
[22] F.J. Marques, C.A. Coelho, Near-exact distributions for the sphericity likelihood ratio test statistic, J. Statist. Plann. Inference 138 (2008) 726–741.
[23] F.J. Marques, C.A. Coelho, A general near-exact distribution theory for the most common likelihood ratio test statistics used in multivariate statistics, The New University of Lisbon, Mathematics Department, Technical Report #5/2008, submitted for publication.
[24] J.W. Mauchly, Significance test for sphericity of a normal n-variate distribution, Ann. Math. Statist. 11 (1940) 204–209.
[25] R.J. Muirhead, Aspects of Multivariate Statistical Theory, J. Wiley & Sons, New York, 1986
[26] Z. Sidak, Rectangular confidence regions for the means of multivariate normal distributions, J. Amer. Statist. Assoc. 62 (1967) 626–633.
[27] Z. Sidak, On multivariate normal probabilities of rectangles: their dependence on correlations, Ann. Math. Statist. 39 (1968) 1425–1434.