METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL JOURNAL OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT
REVISTA INTERNACIONAL DE ESTADISTICA

Vol. XLIII - N. 1-2 **30-VI-1985**

S. KOUROUKLIS - P.G. MOSCHOPOULOS

On the distribution of the trace of a noncentral Wishart matrix

DIPARTIMENTO DI STATISTICA, PROBABILITA' E STATISTICHE APPLICATE
UNIVERSITA' DEGLI STUDI DI ROMA « LA SAPIENZA »

S. KOUROUKLIS (*) - P.G. MOSCHOPOULOS (**)

On the distribution of the trace of a noncentral Wishart matrix

1. INTRODUCTION

The trace T say of a central or noncentral Wishart matrix appears in connection with a number of problems in multivariate analysis. These include principal components analysis, where the trace of a sample covariance matrix is used in estimating the total variance of the population principal components, and problems with symmetries in the covariance structure. For example, when the covariance matrix Σ of a multivariate normal distribution has the intraclass-correlation structure $\Sigma = \sigma^2 [(1-\rho) I + \rho \, 1 \, 1']$, it is well known that the maximum likelihood estimators of σ^2 and ρ involve the trace of a Wishart matrix. Also, the moments and distribution of the likelihood ratio statistic for testing sphericity, i.e., $\Sigma = \sigma^2 I$, depend on representations of the distribution of the trace of a Wishart matrix.

In the central case Khatri and Srivastava [1] have expressed the density of T in terms of zonal polynomials, while in the noncentral case, three expressions are available from Mathai and Pillai [5]. One of these expressions is in terms of zonal polynomials, another in terms of a confluent hypergeometric function, and a third in terms of finite sums of gamma densities when the degrees of freedom is odd.

(*) Department of Statistics The Pennsylvania State University.
(**) Mathematical Sciences The University of Texas at Dallas.

In the general case (i.e., when the degrees of freedom is either even or odd), these representations, while exact, are numerically intractable in view of the computational problems involved in computing relevant zonal polynomials or multiple series.

One could of course use the moments of T given in Mathai [3] to obtain various moment-based approximations; but the accuracy of any such approximation would have to be assessed in terms of errors from the exact values.

In this paper we obtain a single gamma series representation for the exact density of T. The coefficients of the series are easily computed by a simple, stable recurrence relation, thus yielding an expression which is computationally more sound than the above ones (even in the case that the degrees of freedom is odd). This representation can reduce to a mixture representation, an important advantage of which is the ease with which bounds for the truncation error can be derived. Finally, the mixing distribution is identified as that of a finite sum of independent random variables each of which is the sum of a Poisson and a Poisson mixture of a negative binomials. This might be thought of as the multivariate analog of a well known property of the noncentral chi-square distribution.

The motivation in our approach has been the fact that, as noted in Mathai [3], the distribution of T is associated with that of a linear combination of (central or noncentral) chi-square variables, and thus it admits single series representations, see Ruben [8] and Kotz *et al.* [2]. To obtain our series however, we employ a direct method of inverting the moment generating function of T, that is closer in spirit to that in Mathai [4] and Moschopoulos [6].

2. The exact density of the trace

Let $X_1, \ldots, X_n$ be independent random vectors with X_i following a p-variate normal distribution $N_p(\mu_i, \Sigma)$, where μ_i is a $p \times 1$ (real) vector, Σ is a $p \times p$ positive definite matrix, and $n > p$. Set $X = (X_1, \ldots, X_n)$ and $M = (\mu_1, \ldots, \mu_n)$. Then the $p \times p$ random matrix XX' (where X' denotes the transpose of X) has a noncentral Wishart distribution $W_p(n, \Sigma, \Omega)$ with n degrees of freedom, parameter matrix Σ, and noncentrality parameter $\Omega = 1/2\ \Sigma^{-1/2} M M' \Sigma^{-1/2}$. When $M = 0$, the distribution of XX' is a central Wishart $W_p(n, \Sigma)$.

Let now Q be a $p \times p$ orthogonal matrix that diagonalizes Σ, that is

$$Q' \Sigma Q = \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p),$$

where λ_i $i = 1, \ldots, p$ are the eigenvalues of Σ. Then, denoting by T the trace of XX' we obtain

$$T = tr(XX') = tr(Q' XX' Q) = tr(YY') = \sum_{i=1}^{p} \sum_{j=1}^{n} Y_{ij}^2, \tag{2.1}$$

where $Y = Q' X = (Y_1, \ldots, Y_n)$, $Y_j = (Y_{1j}, \ldots, Y_{pj})'$, $j = 1, \ldots, n$. Since $Y_1, \ldots, Y_n$ are independent and $Y_j \sim N_p(Q' \mu_j, \Lambda)$, we have that Y_{ij} $i = 1, \ldots, p$, $j = 1, \ldots, n$ are independent and normally distributed. Hence, (2.1) entails that

$$T = \sum_{i=1}^{p} \sum_{j=1}^{n} Y_{ij}^2 \sim \sum_{i=1}^{p} \lambda_i \chi_n^2(b_i), \tag{2.2}$$

where $\chi_n^2(b_1), \ldots, \chi_n^2(b_p)$ are independent random variables and $\chi_n^2(b_i)$ has a noncentral chi-square distribution with n degrees of freedom and noncentrality parameter b_i, the ith diagonal element of $Q' \Omega Q$.

Recalling that the moment generating function of a $\chi_n^2(b)$ distribution is $(1 - 2t)^{-n/2} \exp\{2bt(1-2t)^{-1}\}$, it follows immediately from (2.2) that the moment generating function of T is

$$M_T(t) = \exp\left\{-\sum_{j=1}^{p} b_j\right\} \prod_{j=1}^{p} (1 - 2\lambda_j t)^{-n/2} \exp\left\{\sum_{j=1}^{p} b_j (1 - 2\lambda_j t)^{-1}\right\} \tag{2.3}$$

for $t < \min\{1/(2\lambda_j): j = 1, \ldots, p\}$.

Below, we derive a series expansion for $M_T(t)$ in powers of $(1 - 2\lambda t)^{-1}$ for an arbitrary $\lambda > 0$. The density of T is then obtained by term by term inversion.

Using the identity

$$1 - 2\lambda_j t = (1 - 2\lambda t)(\lambda_j / \lambda)[1 - (1 - \lambda/\lambda_j)(1 - 2\lambda t)^{-1}]$$

and letting for convenience $z = (1 - 2\lambda t)^{-1}$, (2.3) becomes

$$M_T(t) = \exp\left\{-\sum_{j=1}^{p} b_j\right\} z^{np/2} \prod_{j=1}^{p} (\lambda_j/\lambda)^{-n/2}$$
$$\prod_{j=1}^{p} [1 - (1 - \lambda/\lambda_j) z]^{-n/2} \exp\left\{\sum_{j=1}^{p} b_j (\lambda/\lambda_j) z [1 - (1 - \lambda/\lambda_j) z]^{-1}\right\}.$$

Hence for $t < 1/(2\lambda)$ and $\max\{|1-\lambda/\lambda_j| : j = 1,\ldots,p\} < 1 - 2\lambda t$ we have

$$\begin{aligned}\log M_T(t) &= C + \sum_{j=1}^{p} (n/2) \log [1 - (1-\lambda/\lambda_j) z]^{-1} \\ &+ \sum_{j=1}^{p} b_j (\lambda/\lambda_j) z [1 - (1-\lambda/\lambda_j) z]^{-1} \\ &= C + \sum_{j=1}^{p} (n/2) \sum_{k=1}^{\infty} (1-\lambda/\lambda_j)^k z^k / k \\ &+ \sum_{j=1}^{p} b_j (\lambda/\lambda_j) \sum_{k=0}^{\infty} (1-\lambda/\lambda_j)^k z^{k+1} = C + \sum_{k=1}^{\infty} d_k z^k, \end{aligned} \qquad (2.4)$$

where $C = -\sum_{j=1}^{p} b_j + \log\{z^{np/2} \prod_{j=1}^{p} (\lambda_j/\lambda)^{-n/2}\}$ and

$$d_k = [n/(2k)] \sum_{j=1}^{p} (1-\lambda/\lambda_j)^k + \lambda \sum_{j=1}^{p} (b_j/\lambda_j)(1-\lambda/\lambda_j)^{k-1} \qquad (2.5)$$

for $k = 1, 2, \ldots$.

Now, by expanding $\exp\left\{\sum_{k=1}^{\infty} d_k z^k\right\}$ in a power series we have

$$\exp\left\{\sum_{k=1}^{\infty} d_k z^k\right\} = 1 + \sum_{k=1}^{\infty} \delta_k z^k,$$

where the coefficients δ_k satisfy the recurrence relation

$$\delta_{k+1} = 1/(k+1) \sum_{j=1}^{k+1} j\, d_j\, \delta_{k+1-j}, \qquad k = 0, 1, \ldots \qquad (2.6)$$

with $\delta_0 = 1$ (see also Moschopoulos [6]). Then (2.4) gives

$$M_T(t) = \exp\left\{-\sum_{j=1}^{p} b_j\right\} \prod_{j=1}^{p} (\lambda_j/\lambda)^{-n/2} \sum_{k=0}^{\infty} \delta_k (1-2\lambda t)^{-(np/2+k)}. \qquad (2.7)$$

The exact distribution of T is now available from (2.7) by term by term inversion (which is possible here), noting that $(1-2\lambda t)^{-(np/2+k)}$ is the moment generating function of a gamma variable with parameters $np/2+k,\ 2\lambda$. The density of T is given in the following:

THEOREM 2.1. The density of the trace of a noncentral Wishart matrix can be expressed in the form

$$f(x) = \sum_{k=0}^{\infty} c_k\, g(x; np/2+k, 2\lambda),\ x > 0, \text{ where } 0 < \lambda < \infty \qquad (2.8)$$

is arbitrary, $g(x\,;np/2+k,2\lambda)$ is the gamma density with parameters $np/2+k$, 2λ, and

$$c_k = \exp\left\{-\sum_{j=1}^{p} b_j\right\}\prod_{j=1}^{p}(\lambda_j/\lambda)^{-n/2}\,\delta_k\,, \qquad k=0\ \ 1,\ldots \tag{2.9}$$

with b_j and δ_k as in (2.2) and (2.6) respectively.

We thus see that (2.8) is very suitable for computational purposes since the coefficients δ_k (required for the computation of c_k) are easily computed by the simple recurrence relation (2.6). It also follows from Theorem 2.1 that the exact distribution function of the trace is readily obtainable in terms of incomplete gamma functions. Routines for the computation of the latter are widely available e.g. the IMSL routine MDGAM. The value of λ may be adjusted to make the convergence of the series in (2.8) faster.

In the central case (i.e. when $M=0$), we have $b_j=0\,,j=1\,,\ldots,p$ and thus d_k and hence δ_k and c_k simplify considerably.

3. Mixture representations

For $\lambda \leq \min\{\lambda_j : j=1\,,\ldots,p\}=\lambda_{\min}$ it follows from (2.5), (2.6), and (2.9) that the coefficients c_k are all positive (except in the central case with $\lambda=\lambda_j\,,\ j=1\,,\ldots,p$ which trivially gives $c_0=1\,,\ c_k=0\,,$ $k=1\,,2\,,\ldots)$, and hence (2.8) becomes a mixture representation since necessarily $\sum_{k=0}^{\infty} c_k = 1$. Besides their analytical importance, mixture representations are also useful in practice as they lead to readily obtainable upper bounds for the truncation error. For example, denoting by e_l the error committed in approximating the distribution function $F(x)$ of T by the first l terms of the mixture representation we immediately obtain from (2.8)

$$e_l = \sum_{k=l}^{\infty} c_k \int_0^x g(t\,;np/2+k,\ 2\lambda)\,dt$$

$$\leq \left(\sum_{k=l}^{\infty} c_k\right)\int_0^x g(t\,;np/2+l,2\lambda)\,dt \leq 1-\sum_{k=0}^{l-1} c_k\,,$$

where the first inequality follows from the fact that the chi-square distribution is stochastically increasing in its degrees of freedom. For more details see Robbins and Pitman [7] and Ruben [8].

It is now of interest to identify the mixing distribution giving rise to the coefficients-probabilities c_k, $k = 0, 1, \ldots$ (for $0 < \lambda \leq \lambda_{\min}$). This is done in the sequel. To this end, we first introduce a bit of notation. Let $l_1, \ldots, l_m$, $m \leq p$, be the distinct eigenvalues of Σ, n_i be the multiplicity of l_i, β_i be the sum of those b_j's for which $\lambda_j = l_i$, and $m_i = n n_i$. Then (2.2) can be written as

$$T \sim \sum_{i=1}^{m} l_i \chi^2_{m_i}(\beta_i). \tag{3.1}$$

Let also $P(\beta)$, $NB(s, \alpha)$ denote a Poisson variable with parameter $\beta \geq 0$ (with the understanding that if $\beta = 0$ the variable degenerates at zero) and a negative binomial variable with parameters $s > 0$, $0 \leq \alpha \leq 1$, respectively. The following Theorem 3.1 shows that $c_k = P(K = k)$.

THEOREM 3.1. Let $0 < \lambda \leq \min\{\lambda_j : j = 1, \ldots, p\}$. Then, the density of the trace of a noncentral Wishart matrix can be expressed in the form

$$f(x) = \sum_{k=0}^{\infty} P(K = k)\, g(x; np/2 + k, 2\lambda), \quad x > 0, \tag{3.2}$$

where $K = \sum_{i=1}^{m} (K_i + S_i)$ and (K_i, S_i) are independent random vectors such that $K_i \sim P(\beta_i)$ and conditionally on $K_i = k_i$,

$$S_i \sim NB(m_i/2 + k_i, \lambda/l_i), \quad i = 1, \ldots, m.$$

PROOF. Let (K_i, S_i), $i = 1, \ldots, m$ be as in the statement of the theorem. Let also $V_1, \ldots, V_m$ be independent random variables such that conditionally on $K_i = k_i$ and $S_i = s_i$, V_i is distributed as (central chi-square) $\chi^2_{m_i+2k_i+2s_i}$. Then it follows from results in Robbins and Pitman [7] that the (unconditional) distribution of V_i is that of (l_i/λ) $\chi^2_{m_i}(\beta_i)$, $i = 1, \ldots, m$ and hence on referring to (3.1) we obtain

$$T \sim \lambda(V_1 + \ldots + V_m). \tag{3.3}$$

Now, conditionally on $K_i = k_i$, $S_i = s_i$, $i = 1, \ldots, m$, $V = V_1 + \ldots + V_m$

is distributed as a chi-square with degrees of freedom $np + 2 \sum_{i=1}^{m} (k_i + s_i)$ (by independence of V_i's). Hence the density of V is

$$h(x) = \sum_{k_1=0}^{\infty} \dots \sum_{s_m=0}^{\infty} P(K_i = k_i, S_i = s_i, i = 1, \dots, m)$$

$$g(x; np/2 + \sum_{i=1}^{m} (k_i + s_i), 2)$$

$$= \sum_{k=0}^{\infty} P(K = k)\, g(x; np/2 + k, 2), \; x > 0,$$

which is in view of (3.3) entails (3.2).

Note that for $p = 1$ and $\lambda = \lambda_1$, (3.2) reduces to the well known representation of the noncentral chi-square density as a Poisson mixture of central chi-square densities.

KEY WORDS AND PHRASES:

Noncentral Wishart matrix; exact distribution; series representation; mixture representation.

REFERENCES

[1] KHATRI, C.G. and SRIVASTAVA, M.S. (1971). *On exact nonnull distributions of likelihood ratio criteria for sphericity test and equality of two covariance matrices*, Sankhya A, 33, 201-206.

[2] KOTZ, S., JOHNSON, N.L. and BOYD, D.W. (1967). *Series representations of distributions of quadratic forms in normal variables. I. Central case*, Ann. Math. Statist., 38, 823-837.

[3] MATHAI, A.M. (1980). *Moments of the trace of a noncentral Wishart matrix*, Comm. Statist. - Theor. Meth., A9(8), 795-801.

[4] MATHAI, A.M. (1982). *Storage capacity of a dam with gamma type inputs*, Ann. Inst. Statist. Math. 34, 3, 591-597.

[5] MATHAI, A.M. and PILLAI, K.C.S. (1982). *Further results on the trace of a noncentral Wishart matrix*, Comm. Statist.-Theor. Meth., A 11, 1077-1086.

[6] MOSCHOPOULOS, P.G. (1985). *The distribution of the sum of independent gamma random variables*, Ann. Inst. Statist. Math., Vol 37, No 3.

[7] ROBBINS, H. and PITMAN, E.J. (1949). *Application of the method of mixtures to quadratic forms in normal variables*, Ann. Math. Statist., 20, 552-560.

[8] RUBEN, H. (1962). *Probability content of regions under spherical normal distribution IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables*, Ann. Math. Statist., 33, 542-570.

Summary

It is shown that the density of the trace of a noncentral Wishart matrix has a single gamma series representation, whose coefficients are easily computed by a stable recurrence relation. This representation can reduce to a mixture representation. The mixing distribution is then identified as that of a finite sum of independent random variables, each of which is the sum of a Poisson and a Poisson mixture of negative binomials. Some advantages of the present representation are also indicated.

Riassunto

Si dimostra che la densità della traccia di una matrice non centrale tipo Wishart possiede una figurazione di serie gamma singola, i cui coefficienti vengono facilmente computati tramite un rapporto di ricorrenza stabile. Tale rappresentazione può ridursi a una figurazione di mistura. La distribuzione di mistura è in seguito identificata come quella di una somma finita di variabili casuali indipendenti, ciascuna delle quali è la somma di una Poisson e di una mistura Poisson di binomi negativi. Si indicano poi certi vantaggi della rappresentazione proposta.

Zusammenfassung

Es wird gezeigt, dass die Dichte der Spur der nichtzentrierten Wishart Matrix mit Hilfe einer einfachen Gammareihe dargestellt werden kann, deren Koeffizienten sich mit Hilfe einer stabilen Rekurrenzrelation berechnen lassen. Diese Darstellung reduziertsich zu einer gemischten Representation. Die gemischte Verteilung ist eine endliche Summe von Zufallsvariablen, die sich von einem negative Poisson und einem gemischten negativen Poisson Binom zusammensetzen. Es wird auf gewisse Vorteile dieser Representation hingewiesen.