On the moments of the trace of a matrix and approximations to its non-central distribution

by

C. G. Khatri and K.C.S. Pillai*

Gujarat University and Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 68

April, 1966

* The work of this author was supported in part by the National Science Foundation, Grant No. GP-4600.

On the moments of the trace of a matrix

and approximations to its non-central distribution

by

C. G. Khatri and K.C.S. Pillai*

Gujarat University and Purdue University

1. Introduction and Summary. Let $\underset{\sim}{A}_1$ and $\underset{\sim}{A}_2$ be two symmetric matrices of order p, $\underset{\sim}{A}_1$, positive definite and having a Wishart distribution [2,18] with f_1 degrees of freedom and $\underset{\sim}{A}_2$, at least positive semi-definite and having a (pseudo) non-central (linear) Wishart distribution [1,3,5,18,19] with f_2 degrees of freedom. Now let

$$\underset{\sim}{A}_2 = \underset{\sim}{C}\,\underset{\sim}{Y}\,\underset{\sim}{Y}'\,\underset{\sim}{C}'$$

where $\underset{\sim}{Y}$ is $p \times f_2$ and $\underset{\sim}{C}$ is a lower triangular matrix such that

$$\underset{\sim}{A}_1 + \underset{\sim}{A}_2 = \underset{\sim}{C}\,\underset{\sim}{C}' \quad .$$

Now consider the $s(=\text{minimum}\ (f_2, p))$ non-zero characteristic roots of the matrix $\underset{\sim}{Y}\,\underset{\sim}{Y}'$. It can be shown that the density function of the characteristic roots of $\underset{\sim}{Y}'\underset{\sim}{Y}$ for $f_2 \leq p$ can be obtained from that of the characteristic roots of $\underset{\sim}{Y}\,\underset{\sim}{Y}'$ for $f_2 \geq p$ if in the latter case the following changes are made: [6,18]

$$(1.1) \qquad (f_1, f_2, p) \longrightarrow (f_1 + f_2 - p, p, f_2) \ .$$

Now define $U^{(s)} = tr(\underset{\sim}{I}_p - \underset{\sim}{Y}\,\underset{\sim}{Y}')^{-1} - p = tr(\underset{\sim}{I}_{f_2} - \underset{\sim}{Y}'\underset{\sim}{Y})^{-1} - f_2$. In view of (1.1), we only consider $U^{(s)}$ when $s = p$, i.e. $U^{(p)}$, based on the density function [9] of $\underset{\sim}{L} = \underset{\sim}{Y}\,\underset{\sim}{Y}'$ for $f_2 \geq p$. The first four moments of $U^{(s)}$ have been studied by Pillai in the central case [11,12,13,14,17] those for $U^{(2)}$ also by Pillai [15] in the non-central (linear) case and the first two moments of $U^{(p)}$ by the authors [7]. These results are extended in the present paper, obtaining the third

* The work of this author was supported in part by the National Science Foundation, Grant No. GP-4600.

and fourth moments of $U^{(p)}$ and further, two approximations to the distribution of $U^{(p)}$ are suggested in the linear case.

2. Moments of $U^{(p)}$. In the previous paper by the authors [7] it has been shown that

$$(2.1) \qquad 1 + U^{(p)} = \{(1-\ell_{11})(1-\underset{\sim}{u}'\underset{\sim}{u})\}^{-1} + (1-\underset{\sim}{u}'\underset{\sim}{u})^{-1}(\underset{\sim}{u}'\underset{\sim}{M}\,\underset{\sim}{u}) + \operatorname{tr} \underset{\sim}{M}$$

where ℓ_{11}, $\underset{\sim}{u}:(p-1) \times 1$ and $\underset{\sim}{M}$ are independently distributed and their respective distributions are given by

$$(2.2) \qquad \exp(-\lambda^2) \sum_{j=0}^{\infty} \frac{(\lambda^2)^j}{j!} \; \frac{\ell_{11}^{\frac{1}{2}f_2+j-1}(1-\ell_{11})^{\frac{1}{2}f_1-1}}{\beta[\frac{1}{2}f_2+j, \frac{1}{2}f_1]} \; d\,\ell_{11}$$

$$(2.3) \qquad \left[\Gamma(\tfrac{1}{2}f_1)/\{\Pi^{\frac{1}{2}(p-1)}\Gamma[\tfrac{1}{2}(f_1-p+1)]\}\right](1-\underset{\sim}{u}'\underset{\sim}{u})^{\frac{1}{2}(f_1-p+1)-1} \; d\,\underset{\sim}{u}$$

and

$$(2.4) \qquad \prod_{i=1}^{p-1} \{\Gamma[\tfrac{1}{2}(f_1+f_2-i)]/\Gamma[\tfrac{1}{2}(f_1-i+1)]\Gamma[\tfrac{1}{2}(f_2-i)]\} \; \frac{|\underset{\sim}{M}|^{\frac{1}{2}[f_2-1-(p-1)-1]} \; d\,\underset{\sim}{M}}{\Pi^{\frac{1}{4}(p-1)(p-2)}|\underset{\sim}{I}_{p-1}+\underset{\sim}{M}|^{\frac{1}{2}(\nu-1)}} ,$$

where $\underset{\sim}{M} = (\underset{\sim}{I}_{p-1}-\underset{\sim}{L}_{22})^{-1}-\underset{\sim}{I}_{p-1}$, $\underset{\sim}{L}_{22} = \underset{\sim}{L}_{11} - \underset{\sim}{\ell}\,\underset{\sim}{\ell}'/\ell_{11}$,

$$\underset{\sim}{L} = \begin{pmatrix} \ell_{11} & \underset{\sim}{\ell}' \\ \underset{\sim}{\ell} & \underset{\sim}{L}_{11} \end{pmatrix}\begin{matrix} 1 \\ p-1 \end{matrix} \quad \text{and} \quad \nu = f_1 + f_2 \; .$$
$$\begin{matrix} 1 & p-1 \end{matrix}$$

Now note that (2.3) is invariant under an orthogonal transformation of $\underset{\sim}{u}$, $x_i = u_i^2/(1-u_1^2- \dots - u_{i-1}^2)$, $i = 1,2,\dots, p-1$, $u_o = 0$, is distributed as [7]

$$(2.5) \qquad g_i(x_i) = \{\beta[\tfrac{1}{2}, \tfrac{1}{2}(f_1-i)]\}^{-1} x_i^{\frac{1}{2}-1}(1-x_i)^{\frac{1}{2}(f_1-i)-1} ,$$

and $x_1,\dots, x_{p-1}$ are independent. Further, define $\alpha = 1/(1-\underset{\sim}{u}'\underset{\sim}{u})$ and $\beta = \operatorname{tr} \underset{\sim}{M} + \underset{\sim}{u}'\underset{\sim}{M}\,\underset{\sim}{u}/(1-\underset{\sim}{u}'\underset{\sim}{u})$. Then, computing $E(\alpha^i)$, $i = 1,2,3,4$, $E(\alpha^i\beta)$, $i = 1,2,3$,

$E(\alpha^i\beta^2)$, $i = 1,2$, $E(\alpha\beta^3)$ and $E(1-\ell_{11})^{-i} - E(1-\ell_{11,0})^{-i}$, $i = 1,2,3,4$ (where $\ell_{11,0}$ is a variate whose distribution is the same as that of ℓ_{11} when $\lambda = 0$ and independently distributed of $\underset{\sim}{u}$ and $\underset{\sim}{M}$), we can obtain the first four moments of $U^{(p)}$. It may be pointed out that $E(\alpha^i\beta)$ involves $E(\mathrm{tr}\,\underset{\sim}{M})$, $E(\alpha^i\beta^2)$ involves $E(\mathrm{tr}\,\underset{\sim}{M})^2$ and $E(\mathrm{tr}_2\underset{\sim}{M})$, $E(\alpha\beta^3)$, $E(\mathrm{tr}\,\underset{\sim}{M})^3$, $E[(\mathrm{tr}\,\underset{\sim}{M})(\mathrm{tr}_2\underset{\sim}{M})]$ and $E(\mathrm{tr}_3\underset{\sim}{M})$, where $\mathrm{tr}_i\,\underset{\sim}{M}$ denotes the ith elementary symmetric function in the $(p-1)$ characteristic roots of $\underset{\sim}{M}$. All these results are available in [8].

Expressions for the first two moments of $U^{(p)}$ have been presented in the previous paper by the authors [7]. For the third and fourth moments we get:

$$E(1+U^{(p)})^3 = E(1+U_0^{(p)})^3 + A_1(2\lambda^2)^3 + 3A_2(2\lambda^2)^2 + 3A_3(2\lambda^2) \tag{2.6}$$

where

$$A_1 = \eta_3^{(0)} = [(f_1-p-1)(f_1-p-3)(f_1-p-5)]^{-1} \quad , \tag{2.7}$$

$$A_2 = (\nu-2)\eta_3^{(0)} + \eta_2^{(1)} \quad , \tag{2.8}$$

where

$$\eta_2^{(1)} = (p-1)(f_2-1)(f_1-p-4)\, A_1/(f_1-p) \ , \tag{2.9}$$

$$A_3 = (\nu-2)(\nu-4)\, \eta_3^{(0)} + 2(\nu-2)\, \eta_2^{(1)} + \eta_1^{(2)} \quad , \tag{2.10}$$

where

$$\eta_1^{(2)} = \frac{(p-1)(f_2-1)}{(f_1-p-3)(f_1-p+1)(f_1-p)}\left\{(p-2)(f_2-1)+\frac{2(f_2+1)(f_1-1)}{(f_1-p-2)} + \frac{(p+1)(f_2+1)(f_1-p+1)}{(f_1-p-1)(f_1-p-2)(f_1-p-5)}\right\} . \tag{2.11}$$

Similarly

$$E(1+U^{(p)})^4 = E(1+U_0^{(p)})^4 + B_1(2\lambda^2)^4 + 4B_2(2\lambda^2)^3 + 6B_3(2\lambda^2)^2 + 4B_4(2\lambda^2) \quad , \tag{2.12}$$

where

(2.13) $$B_1 = \eta_4^{(0)} = A_1/(f_1-p-7)$$

(2.14) $$B_2 = (\nu-2)\ \eta_4^{(0)} + \eta_3^{(1)}$$

where

(2.15) $$\eta_3^{(1)} = (p-1)(f_2-1)(f_1-p-6)\ B_1/(f_1-p)$$

(2.16) $$B_3 = (\nu-2)(\nu-4)\ \eta_4^{(0)} + 2(\nu-2)\ \eta_3^{(1)} + \eta_2^{(2)}$$

where

(2.17) $$\eta_2^{(2)} = \{\frac{(f_1-p-4)(f_1-p-6)(p-1)(f_2-1)}{(f_1-p)^2}[\frac{2(f_1-1)(f_1+f_2-p-1)}{(f_1-p+1)(f_1-p-2)} +(p-1)(f_2-1)]-2(p-1)(p-2)(f_2-1)(f_2-2)/\{(f_1-p)(f_1-p+1)\}\}\ B_1$$

(2.18) $$B_4 = (\nu-2)(\nu-4)(\nu-6)\eta_4^{(0)}+3(\nu-2)(\nu-4)\eta_3^{(1)}+3(\nu-2)\eta_2^{(2)}+\eta_1^{(3)}$$

where

(2.19) $$\eta_1^{(3)} = \{\frac{(f_1-p-2)(f_1-p-4)(f_1-p-6)(p-1)(f_2-1)}{(f_1-p)^3} [\frac{2^3(f_1-1)(f_1+f_2-p-1)(f_1+2f_2-p-2)(f_1+p-2)}{(f_1-p-2)(f_1-p-4)(f_1-p+1)(f_1-p+2)} + \frac{6(f_2-1)(f_1+f_2-p-1)(p-1)(f_1-1)}{(f_1-p-2)(f_1-p+1)} + (p-1)^2(f_2-1)^2]$$
$$- \frac{6(f_1-p-4)(p-1)(p-2)(f_2-1)(f_2-2)}{(f_1-p-2)(f_1-p)(f_1-p+1)(f_1-p+2)}[\{(f_1-p)(p-1)+4\}(f_2-p-1)+2(p+1)(p+2)]$$
$$+ 4(p-1)(p-2)(p-3)(f_2-3)(f_2-2)(f_2-1)/\{(f_1-p)(f_1-p+1)(f_1-p+2)\}\}\ B_1 .$$

3. <u>Approximations to the Distribution of $U^{(p)}$</u>. Pillai [15] has given an approximation to the distribution of $U^{(2)}$ for $f_1 > f_2$ and which is good even for very small values of f_2. The following approximation to the distribution of $U^{(p)}$ for $f_1 > (p-1)\ f_2$, based on its moments discussed in the preceding

section and [7], generalizes Pillai's results for $U^{(2)}$ [15]:

$$(3.1)\qquad g(U^{(p)}) = (U^{(p)})^{p_1-1}/(1+U^{(p)}/k)^{p_1+q_1+1}k^{p_1}\beta(p_1,q_1+1),\ 0 < U^{(p)} < \infty ,$$

where

$$p_1 = 2q_1/\{q_1(h-1)-2h\} ,$$

$$q_1 = 2\{c^2(f_1-p-3)h-(c+d)^2(f_1-p-1)\}/\{c^2(f_1-p-3)(h+1)-2(c+d)^2(f_1-p-1)\}$$

$$k = c\{q_1(h-1)-2h\}/[2(f_1-p-1)] ,$$

$$h = (c+1.99d)^3(f_1-p-1)/\{(c+d)^2(f_1-p-5)c\} ,$$

$$c = pf_2+2\lambda^2 \quad\text{and}\quad d = (f_1+(1-p)f_2-1)/(f_1-p) .$$

It may be pointed out that the case $p = 1$ is that of the non-central F [10]. Hence the accuracy of the approximation may be compared in this case with the approximation to the distribution of non-central F obtained by Patnaik and the exact distribution using Table 7 of [10]. However, it should be pointed out that the approximation to the distribution of $U^{(p)}$ in (3.1) has been suggested in this paper using the first three moments and with consideration of accuracy for $p > 1$. From some numerical comparisons made in [8], the respective exact and approximate moments were observed to be closer as p increased. Table 1 gives some idea of the accuracy of the approximation when $p = 1$.

Table 1

Values of $\int_0^{U^{(1)}} g(t)\,dt$ from approximate and exact distributions

f_1	f_2	λ^2	$U^{(1)}$	Probability: Approximate, Eqn.(3.1)	Probability: Approximate, Patnaik	Probability: Exact
10	3	2	1.1124	.765	.752	.745
10	3	8	1.1124	.154	.203	.206
10	3	8	1.9656	.503	.520	.517
10	5	3	1.663	.738	.731	.731
10	5	3	2.818	.920	.913	.914
20	3	2	0.4647	.708	.706	.700
20	5	3	0.67775	.671	.665	.664
20	5	12	1.02575	.196	.244	.245

It may be observed that the approximation suggested for $U^{(1)}$ is more accurate at the upper tail end than the lower. In this case, the condition $f_1 > (p-2)f_2$ reduces to $f_1 > 0$.

Again a comparison of the probabilities in Table 1 arouses the natural curiosity to attempt a generalization of Patnaik's approximation [10]. The following is such a generalization equating the first two respective moments of the exact and approximate distributions:

$$(3.2) \qquad g_1(U^{(p)}) = (U^{(p)})^{\frac{1}{2}\nu_1 - 1} / [(1+U^{(p)}/k_1)^{\frac{1}{2}(\nu_1+\nu_2)} k_1^{\frac{1}{2}\nu_1} \beta(\tfrac{1}{2}\nu_1, \tfrac{1}{2}\nu_2)] ,$$

$$0 < U^{(p)} < \infty$$

where

$$k_1 = (pf_2+2\lambda^2)/\nu_1 \quad ,$$

$$\nu_1 = (pf_2+2\lambda^2)^2(f_1-p)/[(4\lambda^2+pf_2)\{f_1+f_2(1-p)-1\}] \quad ,$$

and $$\nu_2 = f_1-p+1 \quad .$$

4. Further accuracy comparisons. For $p = 2$, Pillai and Jayachandran [16] have given the c.d.f. of $U^{(2)}$ in the following form:

$$F(U^{(2)}) = K'[\sum_{j=0}^{6} \sum_{i=0}^{j} (-1)^{i+j} D'_{ij} B_{ij} + \ldots] \tag{4.1}$$

where

$$B_{ij} = \int_0^{U^{(2)}} \int_0^{u^2/4} [v^{m+i}/(1+u+v)^{m+n+j+3}]\, dvdu$$

where $m = (f_2-3)/2$, $n = (f_1-3)/2$, and K' and D'_{ij} are functions of f_1, f_2 and λ^2 given in [16]. Now define

$$B_x(p', q') = \int_0^x z^{p'-1}(1-z)^{q'-1} dz/\beta(p', q') \ .$$

Then the c.d.f. from (3.1) can be written as

$$G(U^{(2)}) = B_{x_1}(p_1, q_1+1) \ , \tag{4.2}$$

where $$x_1 = U^{(2)}/(k+U^{(2)})$$

and the c.d.f. from (3.2) can be written as

$$G_1(U^{(2)}) = B_{x_2}(\tfrac{1}{2}\nu_1, \tfrac{1}{2}\nu_2) \quad , \tag{4.3}$$

where

$$x_2 = U^{(2)}/(k_1+U^{(2)}) \quad .$$

Now $G(U^{(2)}) - F(U^{(2)})$ and $G_1(U^{(2)}) - F(U^{(2)})$ represent respectively the errors of approximations in the c.d.f. from (3.1) and (3.2). Table 2 provides some numberical comparisons in this respect.

Table 2

Values of $G_1(U^{(2)})$, $G_1(U^{(2)})$ and $F(U^{(2)})$

f_1	f_2	λ^2	$U^{(2)}$	$G(U^{(2)})$	$G_1(U^{(2)})$	$F(U^{(2)})$
23	3	1	0.68072	.880	.877	.875
23	3	1.5	0.68072	.843	.833	.829
13	5	0.5	2.17706	.933	.932	.931
23	5	1.5	1.00707	.875	.869	.867
23	7	1	1.31973	.914	.911	.910
23	13	1.5	2.22596	.913	.912	.912

The values of $U^{(2)}$ and $F(U^{(2)})$ in Table 2 are taken from [16]. For $p > 2$, the method of comparison assumes the exact c.d.f. to be a Pearson type with the first four moments the same as those of the exact. Thus using the "Table of percentage points of Pearson curves for given $\sqrt{\beta_1}$ and β_2, expressed in standard measure" [4], upper 5 per cent points are obtained for selected values of f_1, f_2, and λ^2, and similar upper percentage points are obtained for approximations (3.1) and (3.2). These are presented in Table 3.

Table 3

Upper 5 per cent points using the exact moment quotients and the approximations (3.1) and (3.2)

p	f_1	f_2	λ^2	Percentage points Eqn.(3.1)	Eqn.(3.2)	Exact
3	20	3	12.5	3.873	4.035	4.028
3	50	10	4.5	1.283	1.304	1.300
4	20	4	12.5	4.883	4.971	4.956
4	50	4	12.5	1.409	1.475	1.470
4	50	10	4.5	1.593	1.604	1.598
5	25	5	12.5	4.377	4.407	4.380
5	25	5	32	7.742	7.786	7.768

Tables 2 and 3 show that approximation (3.1) becomes closer to the exact as p increases. In fact, the moment quotients from (3.1) are closer in general to those of the exact than those from (3.2) even for $p = 1$ as shown by numerical computations in [8]. However, approximation (3.2) still maintains its accuracy noted for $p = 1$ even for larger values of p considered in the tables above. Further, it should be pointed out that the condition $f_1 > (p-1)f_2$ applies for both approximations.

The authors wish to thank Mrs. Louise Mao Lui, Statistical Laboratory, Purdue University, for the excellent programming of the material for the computations in this paper carried out on the IBM 7094 Computer, Purdue University's Computer Science's Center.

REFERENCES

[1] Anderson, T.W. (1946). The non-central Wishart distribution and certain problems of multivariate statistics. Ann. Math. Statist. 17, 409-431.

[2] Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis, Wiley, New York.

[3] Anderson, T.W. and Girshick, M.A. (1944). Some extensions of the Wishart distribution. Ann. Math. Statist., 15, 345-357.

[4] Johnson, N.J., Nixon, Eric, Amos, D.E. and Pearson E.S. (1963). Table of percentage points of Pearson curves, for given $\sqrt{\beta_1}$ and β_2, expressed in standard measure. Biometrika, 50, 459-498.

[5] Khatri, C.G. (1964). Distribution of the 'generalized' multiple correlation matrix in the dual case. Ann. Math. Statist. 35, 1801-1806.

[6] Khatri, C.G. (1965). Classical statistical analysis based on a certain multivariate Complex Gaussian distribution. Ann. Math.Statist. 36, 98-114

[7] Khatri, C.G. and Pillai, K.C.S. (1965) Some results on the non-central multivariate beta distribution and moments of traces of two matrices. Ann. Math. Statist. 36, 1511-1520.

[8] Khatri, C.G. and Pillai, K.C.S. (1965) Further results on the non-central multivariate beta distribution and moments of traces of two matrices. Mimeograph Series.No. 38, Department of Statistics, Purdue University.

[9] Kshirsagar, A.M. (1961). The non-central multivariate beta distribution. Ann. Math. Statist., 32, 104-111.

[10] Patnaik, P.B. (1949). The non-central χ^2- and F- distributions and their applications. Biometrika 34, 202-232.

[11] Pillai, K.C.S. (1954). On some distribution problems in multivariate analysis. Mimeograph Series No. 88, Institute of Statistics, University of North Carolina, Chapel Hill.

[12] Pillai, K.C.S. (1955). Some new test criteria in multivariate analysis. Ann. Math. Statist. 26, 117-121.

[13] Pillai, K.C.S. (1956). Some results useful in multivariate analysis. Ann. Math. Statist. 27, 1106-1114.

[14] Pillai, K.C.S. (1960). Statistical Tables for Tests of Multivariate Hypothesis. The Statistical Center, Manila.

[15] Pillai, K.C.S. (1964). On the non-central multivariate beta distribution and the moments of traces of some matrices. Mimeograph Series No. 11, Department of Statistics, Purdue University. (Presented at the Symposium on Multivariate Analysis, Dayton, Ohio, June 1965).

[16] Pillai, K.C.S. and Jayachandran, K. (1966). Power comparisons of two multivariate hypotheses based on four criteria. Mimeograph Series No. 61, Department of Statistics, Purdue University.

[17] Pillai, K.C.S. and Samson, P. (1959) On Hotelling's generalization of T^2. Biometrika, 46, 160-168.

[18] Roy, S.N. (1957). Some Aspects of Multivariate Analysis, Wiley, New York.

[19] Roy, S.N. and Gnanadesikan, R. (1959). Some contributions to ANOVA in one or two dimensions: II. Ann. Math. Statist., 30, 318-340.