A BETA-FUNCTION APPROXIMATION TO THE DISTRIBUTION OF THE TRACE OF A MULTIVARIATE MATRIX

by

DR. TITO A. MIJARES *

The Joint Distribution of the Roots.

The theory of canomical roots came as a natural outgrowth of extending Fisher's discriminant function analysis (1936) to the case of more than two populations. The distribution of the non-zero roots of certain determinantal equations of the form

(1) $$|\underline{T}_1 - \theta(\underline{T}_1 + \underline{T}_2)| = 0$$

arising in multivariate analysis appeared, almost simultaneously, with the publications of papers by R. A. Fisher (1939), M. A. Girshick (1939), P. L. Hsu (1939), A. M. Mood (1951) and S. N. Roy (1939). $\underline{T}_1$ and $\underline{T}_2$ are $p \times p$ matrices having each a Wishart distribution $W(\underline{T}_1 \mid n_1)$ and $W(\underline{T}_2 \mid n_2)$ with n_1 and n_2 degrees of freedom, respectively.

If $\theta_1, \theta_2, \ldots, \theta_s$ are the $s \leq p$ nonzero roots of (1), then the joint distribution of the roots is given by[1/]

---

* Associate Professor, U.P. Statistical Center.

1/ The form in current literature appears varied largely because of notation and because the roots are ordered differently. Some books like those of S. S. Wilks, *Mathematical Statistics* (Princeton University Press, 1950) an T. W. Anderson, *Introduction to Multivariate Statistical Analysis* (John Wiley and Sons, Inc., 1958) give some standard forms. In this paper, we shall essentially follow those of P. L. Hsu, "On the distribution of roots of a determinantal equation," Ann. Eug., vol. 9. See also S. N. Roy, *Some Aspects of Multivariate Analysis*, (John Wiley and Sons, Indian Statistical Institute, 1957) and K. C. S. Pillai, *Concise Tables for Statisticians* (University of the Philippines, 1957).

$$(2) \quad \pi^{s/2} \prod_{i=1}^{s} \frac{\Gamma\left(\frac{2m+2n+s+i+2}{2}\right)}{\Gamma\left(\frac{2m+i+2}{2}\right)\Gamma\left(\frac{2n+i+2}{2}\right)\Gamma\left(\frac{i}{2}\right)} \prod_{i=1}^{s} \theta_i^m (1-\theta_i)^n \prod_{i>j} (\theta_i - \theta_j) \prod_{i=1}^{s} d\theta_i ,$$

$$0 < \theta_1 \leq \theta_2 \leq \dots \leq \theta_s < 1 ,$$

where the parameters m, n, s depend on the multivariate situations and null hypotheses (see Hsu (1939); Roy (1957); Pillai (1957); Anderson (1958)).

The Moments of the Sum of the Roots.

One aspect of the problem which we wish to discuss in this paper relates to the derivation of the moments of the sum[2/]

$$(3) \qquad V_1^{(s)} = \theta_1 + \theta_2 + \dots + \theta_s$$

of the roots whose joint density is given by (2). Another aspect is the fitting of an exact beta-function to the distribution of this sum of trace of the multivariate matrix.

A number of workers may be associated with the first aspect of the problem. D. N. Nanda (1950) appeared to have worked first on the moments of the roots. He obtained the first three moments of (3) for the case of s = 2 under the condition m = 0. The moments were obtained by directly expanding the moment generating function for the sum of two roots, then collecting terms and integrating term by term.

K. C. S. Pillai (1956) obtained a recursion formula for the moments of the sum of s roots in terms of the sum of s — 2 and used it to obtain the first four moments for the sum of 2,

2/ In this paper, we shall interchangeably use "trace" of the multivariate matrix, sum of the roots and first elementary symmetric functions of the roots without distinction.

3 and 4 roots. M. M. Orense (1958) obtained the first three moments of the sum of 5 roots. He also obtained the fourth moment using the method adopted by T. A. Mijares (1958) which is entirely diffrerent from Pillai's method. Mijares also gave the generalization for the first 4 raw moments of the sum of s roots. These expressions are given by Pillai and Mijares (1960). A. Ting (1959), using generalizations of the raw moments, obtained the 4th central moment. Mijares (1961) showed an inverse method of deriving the moments of the trace of the multivariate matrix and outlined an extension of the method to the moments of other elementary symmetric functions of the roots. A systematic method and proof for obtaining any moment of any elementary symmetric function of any number of roots is given in his thesis (1962).

In this paper we shall indicate how completely homogeneous symmetric functions are used in the derivation of the moments of the trace of the multivariate matrix. A general fitting of the distribution to the beta distribution will be given.

Notations.

Denote the completely homogeneous symmetric function of the pth degree in k arguments by

$$(4) \qquad \phi_p(x_1, \ldots, x_k) = \sum_{P_{(p)}} x_1^{p_1} x_2^{p_2} \ldots x_k^{p_k} ,$$

where $\sum$ extends over all partitions $P_{(p)}$ of a non-negative integer $p = \sum_{i=1} p_i$. Define $\phi_0 = 1$ and $\phi_{p'} = 0$ for $p' < 0$.

If $r_1 \neq r_2 \neq \ldots \neq r_k$ be non-negative powers of the x's in the successive columns of the k-order determinant given in (5) below and if ϕ_j be the completely homogeneous symmetric function of the jth degree in all k arguments, then the following relation exists:

(5) $$|x_{k-i+1}^{r_j}| = D\,|\phi_{r_j-k+i}|\ , \quad (i,\ j = 1, \ldots, k),$$

where

(6) $$D = |x_{k-i+1}^{k-j}|\ .$$

A proof of (5) is given in Mijares (1961).

Note that D in (6) is simply equal to $\prod_{i>j} (x_i - x_j)$. By changing the arguments from x to θ and the number of arguments from k to s, we can write (2) in the form

(7) $$f(\theta_1, \ldots, \theta_s) \prod d\theta_i$$
$$= C(s,m,n)\ |\theta_{s-i+1}^{s-j'}|\ \cdot \prod_{i=1}^{s} \theta_i^m (1-\theta_i)^n \prod_{i=1}^{s} d\theta_i\ ,$$
$$(i,\ j' = 1, \ldots, s)$$

where C(s,m,n) is given by

(8) $$C(s,m,n) = \pi^{s/2} \prod_{i=1}^{s} \frac{\Gamma(\frac{2m+2n+s+i+2}{2})}{\Gamma(\frac{2m+i+2}{2})\ \Gamma(\frac{2n+i+2}{2})\ \Gamma(\frac{i}{2})}\ .$$

Hence, the mathematical expectation of the rth power $[V_1^{(s)}]^r$ of the sum of the roots may be given by

(9) $$E[V_1^{(s)}]^r = C(s,m,n) \int\int \ldots \int [V_1^{(s)}]^r\ |\theta_{s-i+1}^{s-j'}| \prod_{i=1}^{s} \theta_i^m (1-\theta_i)^n\, d\theta_i,$$
$$i,\ j' = 1, \ldots, s.$$

By denoting the pth completely homogenous symmetric of the k arguments in θ by ϕ_p, we have, from (5),

$$(10) \qquad |\theta_{s-i+1}^{r_{j'}}| = |\theta_{s-i+1}^{s-j'}| \cdot |\phi_{r_j - s + i}| \ , \quad i, j' = 1, \ldots, s.$$

Hereafter, we shall refer to the last determinant as the ϕ-determinant and the determinant on the left side of (10) as the ϕ-determinant. (10) gives us then a unique representation of the θ-determinant in terms of ϕ-determinant.

We shall denote further the θ-determinant corresponding to the ϕ-determinant by

$$(11) \qquad |\theta_{s-i+1}^{s-j'}| \cdot |\phi_q| = U'(q_s, q_{s-1}, \ldots, q_1) \ ,$$

whose $q = q_{s-j'+1} - s + i$ is the suffix of the (i,j')th element of the ϕ-determinant. We refer to (11) as U'-determinant. If (11) is multiplied by $C(s,m,n) \prod_i \theta_i^m (1-\theta_i)^n$ and integrated over the entire range of the variables θ, we denote it by U without the prime, i.e.,

$$(12) \qquad U(q_s, q_{s-1}, \ldots, q_1) = C(s,m,n) \int \ldots \int U'(q_s, q_{s-1}, \ldots, q_1) \prod_i \theta_i^m (1-\theta_i)^n d\theta_i .$$

A Derivation of the Moments.

By the nature of the structure of the ϕ-determinant, it is possible to find a linear combination of types of U'-determinants for any power r of $\phi_1 = V_1^{(s)}$.

For the first moment, note that $E(V_1^{(s)})$ contains the factors

(13) $$|\theta^{s-j'}_{s-i+1}| \cdot \Phi_1 \qquad (i, j' = 1, \ldots, s)$$

after replacing $V_1^{(s)}$ by Φ_1. This Φ_1 can be further written as

(14) $$\Phi_1 = \begin{vmatrix} \Phi_1 & 0 & \ldots & 0 \\ \Phi_2 & \Phi_0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \Phi_s & \Phi_{s-2} & & \Phi_0 \end{vmatrix}$$

$$= |\Phi_q| \quad , \quad \text{say,}$$

with the suffixes of the Φ's in the last row of the first equality given by q. It is not difficult to see then that the U'-determinant required for the first moment is $U'(s, s-2, \ldots, 1, 0)$ so that by multiplying this U'-determinant by the factor $C(s,m,n) \prod_i \theta_i^m(1-\theta_i)^n$ and integrating over the entire range of the θ's, we have by (13) and (12),

(15) $$E(V_1^{(s)}) = U(q_s, q_{s-2}, \ldots, q_0)$$

$$= \mu_1' .$$

For the second moment, we have an analogous expression to (13),

(16) $$\left|\theta^{s-j'}_{s-i+1}\right| \cdot \Phi_1^2 .$$

Unlike (13) which is expressible in a single U'-determinant, (16) cannot be expressed as such.

Since $\Phi_1^2 = \Phi_1 \cdot \Phi_1 \cdot \Phi_0 \cdots \Phi_0$ to s factors, we take a

Φ-determinant. with this expression for its principal diagonal. This gives

(17) $$\begin{vmatrix} \Phi_1 & \Phi_0 & 0 & \cdots & 0 \\ \Phi_2 & \Phi_1 & 0 & \cdots & 0 \\ \Phi_3 & \Phi_2 & \Phi_0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \Phi_s & \Phi_{s-1} & \Phi_{s-3} & \cdots & \Phi_0 \end{vmatrix} = \Phi_1^2 - \Phi_2 .$$

We need to eliminate Φ_2 from the right side of (17).

Note that $\Phi_2 = \Phi_2 \ \Phi_0 \cdots \Phi_0$ to s factors. Hence, the

Φ-determinant required with this expression in the principal diagonal is

(18) $$\begin{vmatrix} \Phi_2 & 0 & \cdots & 0 \\ \Phi_3 & \Phi_0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \Phi_{s+1} & \Phi_{s-2} & \cdots & \Phi_0 \end{vmatrix} = \Phi_2 .$$

The sum of (17) and (18) gives Φ_1^2. If we now multiply this sum by (2) and integrate the θ's over their entire range, we have, after noting the suffixes of the Φ's in (17) and (18),

$$(19)\quad E\left[(V_1^{(s)})^2\right] = U(s, s-1, s-3, \ldots, 1, 0) + U(s+1, s-2, \ldots, 1, 0) = \mu_2' \,.$$

Using the same procedure as the first two moments, the Φ-determinants multiplied by $\prod_{i>j} (\theta_i - \theta_j)$ gives the following relations among U'-determinant for the third moment:

$$(20)\quad \left| \theta_{s-i+1}^{s-j} \right| \cdot \Phi_1^3 = U'(s, s-1, s-2, \ldots, 0) + 2U'(s+1, s-1, s-3, \ldots, 0) + U'(s+2, s-2, s-3, \ldots, 0) \,.$$

By (12),

$$(21)\quad E\left[(V_1^{(s)})^3\right] = U(s, s-1, s-2, \ldots, 1, 0) + 2U(s+1, s-1, s-3, \ldots, 1, 0) + U(s+2, s-2, s-3, \ldots, 1, 0) = \mu_3' \,.$$

Similarly, following the same method also gives us

$$(22)\quad E\left[(V_1^{(s)})^4\right] = U(s,\ s-1,\ s-2,\ s-3,\ s-5,\ \ldots,\ 1,\ 0)$$
$$+\ 2U(s+1,\ s,\ s-3,\ s-4,\ \ldots,\ 1,\ 0)$$
$$+\ 3U(s+2,\ s-1,\ s-3,\ \ldots,\ 1,\ 0)$$
$$+\ 3U(s+1,\ s-1,\ s-4,\ s-5,\ \ldots,\ 1,\ 0)$$
$$+\ U(s+3,\ s-2,\ s-3,\ \ldots,\ 1,\ 0)$$
$$=\ \mu_4' .$$

Up to this point we have only given the general form of the expressions for the lower order moments of the sum of the roots. It is obvious that the foregoing method becomes no longer easy for higher-order moments. In [1], expressions (15), (19), (21) and (22) were also obtained using the method of diferentiation. The method there is more convenient in obtaining higher-order moments. But the method presented here, of any elementary symmetric functions of the roots in general, can be tied up with the method of obtaining the moments of which is beyond the scope of this paper.

The evaluation of the U-expressions may be made either using the special case for reducing the "pseudo-determinants" into that of order one less (Roy,1945) or its extention by reducing to the order two less than the original one (Pillai, 1956). Both results reduce also the powers of the o's in the first column. Mijares (1962) gives, for the special case a reduction involving any column.

By evaluating a number of these U-expressions, one observes that the expansions are polynomials homogenous in terms of m and n. It was found that a generalization involving a third parameter s gives for the first and second moments:

$$\mu' = \frac{s(2m + s + 1)}{2(m + n + s + 1)} \, ,$$

$$(23) \quad \mu_2' = \frac{s(2m + s + 1)}{(2m+2n+2s+1)(2m+2n+2s+2)(2m+2n+2s+4)}$$

$$\cdot \Big[2n \{ 2sm + (s^2 + s + 2) \} + 4sm^2 + 2sm(3s + 4) + (2s^3 + 5s^2 + 3s + 2) \Big]$$

For the third and fourth moments, the reader is referred to [1]

The Beta-Function Fitting of the Distribution.

By equating the moments in (23) to the moments of the beta-distribution, an approximation to the distribution of the trace of the multivariate matrix may be given by

$$(24) \qquad \frac{(V_1^{(s)}/s)^{a_1-1} \, (1 - V_1^{(s)}/s)^{b_1-1} \, d(V_1^{(s)}/s)}{B(a_1, b_1)}$$

where $B(a_1, b_1) = \dfrac{\Gamma a_1 \Gamma b_1}{\Gamma(a_1 + b_1)}$,

$$(25) \qquad a_1 = \frac{\mu_1'(\mu_1' - \mu_2')}{\mu_2' - \mu_1'^2} \geq -1 \, ,$$

$$b_1 = \frac{(1 - \mu_1')(\mu_1' - \mu_2')}{\mu_2' - \mu_1'^2} \geq -1$$

and the μ's are given by (23).

Pillai (1957) suggested an approximation also of the beta-form but differs from the beta-approxiation we have above. If we call Pillai's corresponding parameters to our a_1 and b_1 by a' and b' respectively, he has

$$a' = \frac{s(2m + s + 1)}{2}$$

(26)

$$b' = \frac{s(2n + s + 1)}{2},$$

which are simpler expression but the resulting approximation tion is not an exact beta-function fitting of the distribution of the sum of the roots. The expression given by (24) is an exact fitting to the beta-distribution.

The table below compares the values of the beta parameters computed from (25) and (26) for different values of s, m and n.

TABLE 1

VALUES OF $\binom{a'}{b'}$ AND $\binom{a}{b}$ FOR SELECTED VALUES OF m AND n

		n = 5		n = 10		n = 50	
		$\binom{a'}{b'}$	$\binom{a}{b}$	$\binom{a'}{b'}$	$\binom{a}{b}$	$\binom{a'}{b'}$	$\binom{a}{b}$
s = 2	m = 5	13 13	14.08 14.08	13 23	13.76 24.35	13 103	13.23 104.81
	m = 10	23 13	24.35 13.76	23 23	24.05 24.05	23 103	23.37 104.66
	m = 50	103 13	104.81 13.23	103 23	104.66 23.37	103 103	104.01 104.01
s = 5	m = 5	40 40	48.30 48.30	40 65	46.05 78.84	40 265	41.91 277.68
	m = 10	65 40	74.84 46.05	65 65	72.74 72.74	65 265	67.87 276.68
	m = 50	265 40	277.68 41.91	265 65	276.68 67.87	265 265	272.17 272.17
s = 20	m = 5	310 310	464.27 464.27	310 410	434.60 574.79	310 1210	359.08 1401.59
	m = 10	410 310	574.78 434.59	410 410	548.22 548.22	410 1210	470.34 1388.09
	m = 50	1210 310	1401.59 359.03	1210 410	1388.09 470.34	1210 1210	1323.96 1323.96

One may observe from the table that, when the number of roots s is small, the differences in the beta-function parameters appears small. As the number of roots increases, the differences become increasingly larger. This may be observed from the tables s = 5 and s = 20.

It may be further noted that since the exact fitting to the beta-function gives parameters a_1 and b_1 that are greater than the corresponding a'_1 and b'_1 for values given in the above table, one can expect that the curve given by the first approximation is closer to the horizontal axis than the second approximation at both tails. This implies that confidence limits of $V_1^{(s)}/s$ tend to be shorter using the first approximation than using the second approximation.

Example (i). m = n = 5, s = 2. From the table, a' = b' = 13 using the table for percentage points of the B-ditribution[3] one obtains .3143 for the lower limit at 2-1/2% level and .6857 for the upper limit for the same per cent level. This gives approximately 95% confidence interval of .3714. However, from the table also given above a = b = 48.03. The lower 2-1/2% point is .3350 and the upper 2-1/2 point is .6650 so that the 95% confidence interval is .3300. There is a difference of .0400 in the confidence interval.

Example (ii). m = 5, n = 50, s = 5. For large values of the beta-function parameters α, β Carter's approximation[4] to the lower percent point of the incomplete beta-function is given by $x(I \mid \alpha, \beta)$, where

3 E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Cambridge, Table 16.

4 E. S. Pearson and H. O. Hartley (*loc. cit.*) p. 35. This is convenient when $2\alpha > 60$ and $2\beta > 40$.

(27) $$x(I \mid \alpha, \beta) = \frac{\alpha}{\alpha + \beta e^{2z}}$$

and

(28) $$z = \frac{X\sqrt{h + \lambda}}{h} - \left(\frac{1}{\beta - 1/2} - \frac{1}{\alpha - 1/2}\right)(\tau - A)$$

X = standardized normal deviate

= 1.96, for two-tailed 5% level.

$$A = \frac{1}{12}\left(\frac{1}{\alpha - 1/2} + \frac{1}{\beta - 1/2}\right), \quad h = \frac{1}{3A}$$

λ = .1402 for two-tailed 5% level

τ = .4868 for two-tailed 5% level.

The corresponding upper percentage point is given by the rela-

(29) $x(I \mid \alpha, \beta) = 1 - x'(I \mid \beta, \alpha)$,

i.e., by interchanging α and β and taking $x' = 1 - x$ as the

percentage point.

The computed values are:

	α = a′ (40) β = b′ (265)	α = a (41.91) β = b (277.68)
Upper 2-1/2% pt.	.1711	.1677
Lower 2-1/2% pt.	.0957	.0980
Confidence Interval	.0754	.0697

The two examples are typical and the differences in their confidence interbal appear to be least (although still not negligible) when m = n. Here, the two approximations give symmetrical beta-curves. For large values of s however large discrepancies in the confidence interval may be expected, especiallly in cases where m and n differ much.

Remarks.

One other way that the two approximations may be compared is to compute their moment ratios

$$\beta_1 = \mu_3^2 / \mu_2^3 \quad \text{and} \quad \beta_2 = \mu_4 / \mu_2^2.$$

The author's conjecture is that using a' and b' will give non-negligible departures from the true values of

β_1 and β_2, especially where values of m and n differ very much. This is only to be expected since only the mean of the true distribution was used, in effect, in fitting to a beta distribution.

On the other hand, a and b were obtained using the mean and variance of the true distribution. Values of a and b have been computed by the author at the MIT Computation Center, Cambridge, using the FORTRAN program he wrote for the IBM 709/7090 for values of 2m = −1(1)10(10)60(20)120, 2n = 10(10)200 and s = 2, 3, . . ., 50. In all cases, differences from the true third and fourth moments using the exact beta-function fitting are practically zero.

REFERENCES

[1] Mijares, T.A., "On the Distribution of the Sum of Six Roots of a Matrix in Multivariate Analysis," Library, University of the Philippines (unpublished thesis), 1958.

[2] ————, "On the Moments of Elementary Symmetric Function of the Roots of a Matrix in Multivariate Analysis," **Ann. of Math. Stat.**, Vol. 32, 1961.

[3] Nanda, D. N., "Distribution of the Sum of the Roots of a Determinantal Equation Under a Certain Condition," **Ann. of Math. Stat.**, Vol. 21 (1950)

[4] Orense, M. O., "On the Distribution of the Sum of Five Roots of a Matrix in Multivariate Analysis," Library. University of the Philippines (unpublished thesis), 1958.

[5] Pillai, K. C. S., **Statistical Tables for Tests of Multivariate Hypothesis**, The Statistical Center, University of the Philippines, Manila (1960).

[6] ————, "Some Results Useful in Multivariate Analysis," **Ann. Math. Stat.**, Vol. 27 (1956).

[7] Roy, S. N., "The Indivdual Sampling Distribution of the Maximum, the Minimum and Any Intermediate One of the p-statistics on the Null Hypotheses," **Sankhya**, Vol. 7 (1945).