On Elementary Symmetric Functions of the Roots of Two Matrices in Multivariate Analysis

Ъу

K. C. Sreedharan Pillai

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 23

September, 1964

On Elementary Symmetric Functions of the Roots of Two Matrices in Multivariate Analysis*

Ъу

K. C. Sreedharan Pillai Purdue University

1. Summary

A lemma is proved to show that the moments of the (s-i)th elementary symmetric function (esf) in s non-null characteristic roots, $\lambda_{i}(i=1,2,\ldots,s)$, of a matrix in multivariate analysis can be derived from those of the ith esf. Using this lemma the first four moments of the (s-l)th esf have been obtained from those of the first esf already known (Pillai, 1954, 1960; Pillai and Samson, 1959). Further, a second lemma is given showing that the moments of the (s-l)th esf in the s characteristic roots, $\theta_{i} = \lambda_{i}/(1+\lambda_{i})$, are derivable from those of the first esf in the λ 's. Upper percentage points (5°/o and $1^{\circ}/o$) are obtained for the distribution of the (s-l)th esf in the λ 's for s=3 using the moment quotients. An example is given to illustrate the use of this criterion.

2. Introduction

Many of the distribution problems in multivariate analysis are based on the distribution of the non-null characteristic roots of a matrix derived from sample observations taken from multivariate normal populations. This distribution, given by Roy (1939), is of the form

^{*}This research was supported in part by Contract NONR-1100(26) with the Office of Naval Research. Reproduction in whole or in part permitted for any purposes of the United States Government.

$$f(\lambda_{1},\lambda_{2},\ldots,\lambda_{s};m,n) = C(s,m,n) \prod_{i=1}^{s} \lambda_{i}^{m}/(1+\lambda_{i})^{m+n+s+1} \int_{1>j}^{\pi} (\lambda_{i}-\lambda_{j})$$

$$0 < \lambda_{1} \le \lambda_{2} \le \cdots \le \lambda_{s} < \infty,$$

$$(2.1)$$

where

$$C(s,m,n) = \pi^{s/2} \prod_{i=1}^{s} \left\{ \Gamma(\frac{2m+2n+s+i+2}{2})/\Gamma(\frac{2m+i+1}{2}) \Gamma(\frac{2n+i+1}{2}) \Gamma(i/2) \right\}$$
 (2.2)

and m and n are defined differently for various situations described by Pillai (1955, 1960).

The studies on the first esf in the λ 's have been carried out by Pillai (1954, 1956, 1960) and Pillai and Samson (1959). In this paper, moments of some of the other esf's are considered and in particular those of the (s-1)th, (s-2)th and the second esf's. Further, the esf's in the characteristic roots, $\theta_i = \lambda_i/(1+\lambda_i)$, are studied in a similar manner.

3. Moments of the (s-1)th esf in the λ 's

First consider the following lemma:

Lemma 1. If $U_{i,m,n}^{(s)}$ and $\mu_r^*\{U_{i,m,n}^{(s)}\}$ denote the ith esf in the s λ 's following the distribution (2.1), and its rth moment respectively, then

$$\mu_{\mathbf{r}}^{!}\left\{U_{\mathbf{s-i},\mathbf{m},\mathbf{n}}^{(\mathbf{s})}\right\} = \frac{C(\mathbf{s},\mathbf{m},\mathbf{n})}{C(\mathbf{s},\mathbf{n-r},\mathbf{m+r})} \quad \mu_{\mathbf{r}}^{!}\left\{U_{\mathbf{i},\mathbf{n-r},\mathbf{m+r}}^{(\mathbf{s})}\right\}. \tag{3.1}$$

Proof. Let $\gamma_i = 1/\lambda_i$ (i=1,2,...,s). Then the ith esf in the γ 's equals the quotient of the (s-i)th esf and the sth esf in the λ 's or, reciprocally, the (s-i)th esf in the λ 's equals the quotient of the ith esf and the sth esf in the γ 's. Now note that the distribution of the γ 's follows the distribution (2.1) with m and n interchanged. Hence the lemma follows.

From lemma 1, putting i=1 in (3.1), the rth moment of the (s-1)th esf in the λ 's is obtained as

$$\mu_{r}^{i} \left\{ U_{s-1,m,n}^{(s)} \right\} = \frac{C(s,m,n)}{C(s,n-r,m+r)} \quad \mu_{r}^{i} \left\{ U_{1,n-r,m+r}^{(s)} \right\}. \tag{3.2}$$

The first three central moments of $U_{1,m,n}^{(s)}$ are available in (Pillai, 1954, Pillai and Samson, 1959) and the fourth central moment in (Pillai, 1960). Using these results, the following first four moments of the (s-1)th esf in the λ 's are obtained:

$$\mu_{1}^{i} \left\{ U_{s-1,m,n}^{(s)} \right\} = \frac{5(2n!s-1)}{2(m+1)} \prod_{i=1}^{s} \frac{(2m+i+1)}{(2n+i-1)}, \qquad (3.3)$$

$$\mu_{2}^{*}\left\{\begin{array}{l}U_{s-1,m,n}^{(s)}\right\} = \left\{\frac{s(2n+s-3)}{4(m+2)^{2}}\right\}\left\{\frac{(2m+2n+s+1)(2m+s+4)}{(m+1)(2m+5)} + s(2n+s-3)\right\}A \qquad (3.4)$$

where

$$A = \frac{s}{\pi} \frac{(2m+i+3)(2m+i+1)}{(2n+i-1)(2n+i-3)},$$

$$\mu_{3}^{i} \left\{ U_{s-1,m,n}^{(s)} \right\} = \left\{ \frac{s(2n+s-5)}{8(m+3)^{3}} \right\} \left\{ \frac{(2m+2n+s+1)(2m+s+6)}{(m+2)(2m+7)} \right\}$$
(3.5)

$$\left[\frac{\frac{1}{(m+2n+s-2)(2m+2s+6)}}{(m+1)(2m+8)} + 3s(2n+s-5)\right] + s^{2}(2n+s-5)^{2}$$

where

$$B = \frac{s}{\pi} \frac{(2m+i+5)(2m+i+3)(2m+i+1)}{(2n+i-1)(2n+i-3)(2n+i-5)}$$

۲.

and

$$\mu_{\mu}^{\dagger} \left\{ U_{s-1,m,n}^{(s)} \right\} = \left\{ \frac{s(2n+s-7)}{16(m+4)^{4}} \right\} \left\{ \frac{\frac{1}{8}(2m+2n+s+1)[C+D(n-4)(m+n+s+1)]}{(m+1)(m+2)(m+3)(m+5)(2m+7)(2m+9)(2m+11)} + E \right\} F$$
..... (3.6)

where

$$c = \sum_{j=0}^{6} (-1)^{j} f_{j} p^{6-j},$$

$$D = \sum_{j=0}^{5} (-1)^{j} g_{j} p^{5-j}$$

$$\mathbf{p} = (\mathbf{m} + \mathbf{s} + 5),$$

$$E = s(2n+s-7) \left\{ \frac{2(2m+2n+s+1)(2m+s+8)}{(m+3)(2m+9)} \left[\frac{1+(2n+m+s-3)(2m+2s+8)}{(m+2)(m+5)} + 3s(2n+s-7) \right] + s^{2}(2n+s-7)^{2} \right\},$$

$$F = \begin{cases} s & (2m+i+7)(2m+i+5)(2m+i+3)(2m+i+1) \\ (2n+i-1)(2n+i-3)(2n+i-5)(2n+i-7) \end{cases}$$

and f's and g's are polynomials in s given below:

$$f_0 = s^2 + s + 4,$$

$$f_1 = 4.5s^3 + 12.5s^2 + 10s + 16,$$

$$f_2 = 8.25s^4 + 37.5s^3 + 42.5s^2 + 15.25s + 21,$$

$$f_3 = 7.875s^5 + 49.625s^4 + 93.75s^3 + 44.25s^2 - 10.25s + 7,$$

$$f_4 = 4.125s^6 + 33.1875s^5 + 89.625s^4 + 85.0625s^3 - 3.25s^2 - 34.25s - 5,$$

$$f_5 = 1.125s^7 + 11s^6 + 38.5625s^5 + 56.125s^4 + 20.9375s^3 - 24.25s^2 - 20.75s - 3,$$

$$f_6 = 0.125s^8 + 1.4375s^7 + 6.1875s^6 + 12.0625s^5 + 8.9375s^4 - 3s^3 - 7.75s^2 - 3s,$$

$$g_0 = 2s$$
,
 $g_1 = 8s^2 + 15s - 20$,
 $g_2 = 12.5s^3 + 47s^2 + 9s - 92$,
 $g_3 = 9.5s^4 + 53.75s^3 + 81s^2 - 55s - 156$,
 $g_4 = 3.5s^5 + 26.5s^4 + 68.75s^3 + 47s^2 - 87s - 116$,
 $g_5 = 0.5s^6 + 4.75s^5 + 16.75s^4 + 26s^3 + 5s^2 - 36s - 32$.

4. Moments of the (s-1)th esf in the 0's

As in the previous section a lemma may be stated.

Lemma 2. If $V_{i,m,n}^{(s)}$ and $\mu_r^i \left\{ V_{i,m,n}^{(s)} \right\}$ denote the ith esf in the s θ 's and its rth moment respectively, then

$$\mu_{r}^{i} \left\{ V_{s-1,m,n}^{(s)} \right\} = \frac{C(s,m,n)}{C(s,n,m+r)} \sum_{j=0}^{r} {r \choose j} s^{r-j} \mu_{j}^{i} \left\{ U_{1,m,m+r}^{(s)} \right\}. \tag{4.1}$$

Proof. For proving (4.1), it is enough if we note that

$$V_{s-1,m,n}^{(s)} = (s + \sum_{i=1}^{s} \gamma_i) / \frac{s}{n} (1 + \gamma_i) \text{ where } \gamma^i s$$
 (4.2)

as stated above, follow the distribution (2.1) with m and n interchanged. The rest easily follows.

The first four moments of the (s-1)th esf in the θ 's thus may be obtained by using lemma 2. However they are not given here since it has been shown (Pillai, 1964) that the moments of $V_{i,m,n}^{(s)}$ can be derived in a simple manner from the respective moments of $U_{i,m,n}^{(s)}$ (and vice versa).

5. Moments of other esf's in the θ 's and λ 's

It has been shown (Pillai, 1954, 1956; Pillai and Mijares, 1959) that the distribution of θ 's obtainable from (2.1) by the transformation $\lambda_{\bf i} = \theta_{\bf i}/(1-\theta_{\bf i})({\bf i}=1,2,\ldots,s) \quad \text{can be expressed in a determinantal form and when integrated within the limits, } 0 < \theta_{\bf i} \leq \ldots \leq \theta_{\bf s} < 1, \text{ is given by}$

$$C(s,m,n) \ V(s-1,s-2,...,1,0) = C(s,m,n)$$

$$\begin{vmatrix} \int_{0}^{1} \theta_{s}^{m+s-1} (1-\theta_{s})^{n} d\theta_{s} \int_{0}^{1} \theta_{s}^{m} (1-\theta_{s})^{n} d\theta_{s} \\ \vdots & \vdots & \vdots \\ \int_{0}^{\theta_{s}} \theta_{1}^{m+s-1} (1-\theta_{1})^{n} d\theta_{1} \int_{0}^{\theta_{s}} \theta_{1}^{m} (1-\theta_{1})^{n} d\theta_{1} \\ \end{vmatrix}$$

$$(5.1)$$

Further,

$$\mu_{1}^{i} \left\{ v_{1,m,n}^{(s)} \right\} = C(s,m,n) V(s,s-1,...,s-i+1,s-i-1,...,1,0).$$
 (5.2)

The right hand side of (5.2) can be shown to be equal to

$$\binom{s}{i} = \frac{i}{\pi} \frac{(2m+s-j+2)}{(2m+2n+2s-j+3)}$$
 (5.3)

based on some particular cases evaluated by Pillai and Mijares (1959). From this result, using the method given by Pillai (1964) i.e. by attaching negative signs to all terms except that in n in each linear factor involving n in (5.3) and further changing n to m+n+s+1, the first moment of the ith esf in the λ 's is given by

$$\mu_{1}^{i} \left\{ U_{i,m,n}^{(s)} \right\} = {s \choose i} \frac{i}{\pi} \frac{(2m+s-j+2)}{(2n+j-1)} .$$
 (5.4)

Now consider
$$\mu_2^* \left\{ v_{2,m,n}^{(s)} \right\}$$
 . It can be shown that (Pillai (1964))

$$\mu_{2}^{!} \left\{ V_{2,m,n}^{(s)} \right\} = C(s,m,n) \left[V(s+1,s,s-3,...,1,0) + V(s+1,s-1,s-2,s-4,...,1,0) + V(s,s-1,s-2,s-3,s-5,s-6,...,1,0) \right].$$

$$+ V(s,s-1,s-2,s-3,s-5,s-6,...,1,0) \right].$$

$$....(5.5)$$

Now substituting the values of the determinants in (5.5), already evaluated by Pillai and Mijares (1959),

$$\mu_{2} \left\{ V_{2,m,n}^{(s)} \right\} = \frac{s(s-1)(2m+s)(2m+s+1)}{6} G_{1}$$

$$3! \pi (2m+2n+2s-j+5)$$

$$j=1$$
(5.6)

where
$$G_1 = 6n^2 \left\{ 4 \text{ s(s-1)m}^2 + 2(\text{s-1})(2\text{s}^2+\text{s+8})\text{m} + \text{s}^4 + 7\text{s}^2 - 8\text{s} + 12 \right\}$$

+ $3n \left\{ 16\text{s(s-1)m}^3 + 4(\text{s-1})(8\text{s}^2+5\text{s+8})\text{m}^2 + 2(\text{s-1})(10\text{s}^3+12\text{s}^2+27\text{s}+24)\text{m} + 4\text{s}^5+3\text{s}^4+12\text{s}^3+5\text{s}^2-24\text{s}+36} \right\} + \text{s(s+1)}(2\text{m}+\text{s+1})(2\text{m}+\text{s+2})(\text{m+s})(2\text{m}+2\text{s}+1)$
+ $(\text{s-2})(2\text{m}+2\text{s}+3)(2\text{m}+\text{s-1})\left\{ 4\text{sm}^2 + 2\text{s}(3\text{s}+2)\text{m} + 2\text{s}^3 + 3\text{s}^2 + \text{s} + 6 \right\}.$

The second moment of the second esf in the λ 's can be obtained from (5.6) using the method given by Pillai (1964). We get

$$\mu_{2} \left\{ U_{2,m,n}^{(s)} \right\} = \frac{s(s-1)}{3!} \frac{(2m+s)(2m+s+1)}{6} \cdot G \qquad (5.7)$$

$$\pi (2n+j-3)$$

$$j=1$$

where G is obtained from G_1 by attaching negative sign to the first degree term in n and then changing n to m+n+s+1 in all the terms.

Further, applying (3.1) with i=2 and r=2, we get from (5.7)

$$\mu_{2}^{!} \left\{ U_{s-2,m,n}^{(s)} \right\} = \frac{C(s,m,n)}{C(s,n-2,m+2)} \quad \mu_{2}^{!} \left\{ U_{2,n-2,m+2}^{(s)} \right\}. \tag{5.8}$$

6. Upper percentage points of $U_{2,m,n}^{(3)}$

Using (3.3) to (3.6) with s=3, the moment quotients, β_1 and β_2 , were computed on IBM 7090 accurate to five decimals for values of $m=-\frac{1}{2}\,(\frac{1}{2})$ 5,7,10(5)50,60,80,100,130,160,200,300,500 and 1000 and n starting from 20 but otherwise as for m. The upper percentage points were computed manually using tables of ''Percentage points of Pearson curves for given β_1 , β_2 expressed in standardized measure'' (Pearson and Hartley, 1958) and extrapolating in some cases from ''Tables of deviates of Type I curves measured from the mean in terms of the standard deviation'' (Pearson, 1931). Tables 1 and 2 give the percentage points thus computed.

7. An Example

The use of the percentage points for $U_{2,m,n}^{(3)}$ may be illustrated by an example. This criterion is being suggested for testing different kinds of hypotheses in multivariate analysis, for example, i) that of equality of the dispersion matrices of two p-variate normal populations, ii) that of equality of the p-dimensional mean vectors for V p-variate normal populations and, iii) that of independence between a p-set and a q-set of variates in a (p+q)-variate normal population. It may be pointed out that for the last hypothesis, transformation has to be carried out from Θ 's to the λ 's since computationally it is easier to obtain Θ 's first. Now, for illustration, hypothesis ii) is considered here using the following data taken from Rao (1952, p. 263) concerning measurements of 140 schoolboys of almost the same age

belonging to six different schools in an Indian city. Three characters were measured (1) head length, (2) height and (3) weight. The problem is to test for the significance of differences between schools in the mean characters. Let S*, S be the matrices consisting of the sums of squares and products 'between' and 'within' schools for the three characters. The S* and S* matrices as well as the values of S* and S* are given in Pillai and Samson (1959). Using those matrices, S* the inverse of S*, and S* are given below:

$$s^{-1} = \begin{pmatrix} .0^{1}823,947,419 \\ -.0^{1}572,276,500 & .0^{2}148,849,391 \\ .0^{6}756,373,421 & -.0^{3}284,873,914 & .0^{3}103,413,946 \end{pmatrix}$$
(7.1)

and

$$s* s^{-1} = \begin{pmatrix} .050,096,98 & .127,295,4 & -.006,541,510 \\ .009,293,867 & .098,659,55 & -.001,449,733 \\ .021,212,45 & .107,934,8 & .052,878,55 \end{pmatrix} \bullet (7.2)$$

From (7.2), $U_{2,0.5,65}^{(3)} = .01192$. By interpolation from Tables 1 and 2 respectively it may be seen that this value is significant at the upper $5^{\circ}/\circ$ level but not significant at the upper $1^{\circ}/\circ$ level. This agrees with the findings (see Pillai and Samson, 1959) based on the test $U_{1,0.5,65}^{(3)}$ and also with those of Rao, who examined the data using the Λ criterion of Pearson and Wilks (1933). Foster (1957), however, finds that the largest root is not significant at the upper $5^{\circ}/\circ$ level but only at the upper $15^{\circ}/\circ$ level.

It may be pointed out that the monotonicity of the power function of this test has been already established (Das Gupta, Anderson and Mudholkar, 1964).

The author wishes to thank Mr. Peter E. Dress, Instructor in Statistics, Department of Statistics, and Mrs. Louise Mao Lui, Statistical Laboratory, Purdue University, for the excellent programming of this material for the IBM 7090 computer, Purdue University's Computer Sciences Center.

References

- Das Gupta, S., Anderson, T.W. and Mudholkar, G.S. Monotonicity of the power functions of some tests of multivariate linear hypothesis. Ann. Math. Statist., 35, 200-205.
- Foster, F.G. (1957). Upper percentage points of the generalized Beta distribution. Biometrika, 44, 441-53.
- Pearson, E.S. and Hartley, H.O. (1958). Biometrika Tables for Statisticians, l, Cambridge University Press for the Biometrika Trustees.
- Pearson, E.S. and Wilks, S.S. (1933). Methods of analysis appropriate for k samples of two variables. Biometrika, 25, 353-378.
- Pearson, K. (1931). Tables for Statisticians and Biometricians, 2, The Biometric Laboratory, University College, London.
- Pillai, K.C.S. (1954). On some distribution problems in multivariate analysis.

 Mimeograph Series No. 88, Institute of Statistics, University of North
 Carolina.
- Pillai, K.C.S. (1955). Some new test criteria in multivariate analysis. Ann. Math. Statist., 26, pp. 117-121.
- Pillai, K.C.S. (1956). Some results useful in multivariate analysis. Ann. Math. Statist., 27, pp. 1106-1114.
- Pillai, K.C.S. (1960). Statistical Tables for Tests of Multivariate Hypotheses. The Statistical Center, University of the Philippines, Manila.
- Pillai, K.C.S. (1964). On the moments of elementary symmetric functions of the roots of two matrices. Ann. Math. Statist., 35, No. 4.
- Pillai, K.C.S. and Mijares, T.A. (1959). On the moments of the trace of a matrix and approximations to its distribution. Ann. Math. Statist., 30, pp. 1135-1140.
- Pillai, K.C.S. and Samson, Jr., P. (1959). On Hotelling's generalization of T².

 Biometrika, 46, pp. 160-168.
- Rao, C.R. (1952). Advanced Statistical Methods in Biometric Research. John Wiley and Sons, Inc., New York.
- Roy, S.N. (1939). P-statistics or some generalizations in analysis of variance appropriate to multivariate problems. Sankhya, 4, pp. 381-396.

-
υ(3) 2,π, n
ಕ
Points
5%
Upper
Table 1.

1000	.00000142 .00000382 .00000382 .00000113 .0000113 .0001
200	.000057 .000153 .000153 .000153 .00056 .00056 .00056 .00056 .00056 .00057 .0005
300	.000157 .00058 .00058 .000106 .00108
82	00037 00005 00005 00023
970	25000. 25000. 26000.
130	48000 48
300	00000000000000000000000000000000000000
8	00000000000000000000000000000000000000
8	00000000000000000000000000000000000000
જ	000 000 000 000 000 000 000 000 000 00
1-5	0020 0120 0220 0220 0220 0220 0220 0220
O ⁴	00.000.000.000.000.000.000.000.000.000
35	0.00.00.00.00.00.00.00.00.00.00.00.00.0
30	30.00.00.00.00.00.00.00.00.00.00.00.00.0
83	8.888.488.888.888.888.888.888.888.888.8
82	6.000 6.000
п Н	. o . u u u u u u u u u u u u u u u u u

1000	.000028 .000043 .000061 .000062 .000127 .000137 .00027 .000272 .00027 .0
50	
300	
200	2000 2000 2000 2000 2000 2000 2000 200
160	
130	7.000. 7.4000. 7.4000. 7.4000. 7.4000. 7.600
87	6.000 6.000
8	
9	
Ş	17566-48311
4	4 9 4 7 8 7 8
S	1 9 W. 1/6 1 8 6 9
	29.00.00.00.00.00.00.00.00.00.00.00.00.00
	\$0.00000000000000000000000000000000000
	29 20 20 20 20 20 20 20 20 20 20 20 20 20
	28 28 28 28 28 28 28 28 28 28 28 28 28 2
	10,4499999999999999999999999999999999999

Table 2. Upper $1^{\circ}/o$ Points of $U_{2,m,n}^{(3)}$