ON THE DISTRIBUTION OF THE CHARACTERISTIC ROOTS OF NORMAL SECOND-MOMENT MATRICES[1]

By A. M. Mood

The Rand Corporation

1. Summary. Distributions of characteristic roots have been obtained by Girshick [1], Fisher [2], Hsu [3], and Roy [4]. The present paper outlines an alternative derivation of these distributions which is somewhat more elementary than those that have been published and which may have some pedagogical utility. The primary object of the paper, however, is to obtain the normalizing constants for these distributions; though the correct values of the constants have been published in the references cited above, no convincing derivation seems to have been recorded.

2. The problem. Let

$$a_{ij} = \sum_{\alpha=1}^{m} (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j) \qquad (i, j = 1, 2, \cdots, k) \tag{1}$$

be sums of squares and products for samples of size $m(>k)$ from a k-variate normal distribution with covariance matrix $\| \sigma_{ij} \| (= \| \sigma^{ij} \|^{-1})$ having a k-fold characteristic root λ. The a_{ij} are distributed by the Wishart density function

$$f(a_{ij}; m - 1, \sigma^{ij}) = \frac{| \frac{1}{2} \sigma^{ij} |^{\frac{1}{2}(m-1)} | a_{ij} |^{\frac{1}{2}(m-k-2)} e^{-\frac{1}{2}\Sigma \sigma^{ij} a_{ij}}}{\pi^{\frac{1}{4}k(k-1)} \prod_{i=1}^{k} \Gamma[\frac{1}{2}(m - i)]} \tag{2}$$

with $m - 1$ degrees of freedom. Let b_{ij} be similarly distributed with $n - 1$ degrees of freedom and independently of the a_{ij}.

We are concerned with the distribution of the roots $w_1, \cdots, w_k$ of

$$| a_{ij} - w\sigma_{ij} | = 0, \tag{3}$$

which roots form a natural multivariate analogue of chi-square. Similarly the roots of

$$| a_{ij} - vb_{ij} | = 0 \tag{4}$$

provide an analogue for the variance ratio, and the roots of

$$| a_{ij} - u(a_{ij} + b_{ij}) | = 0 \tag{5}$$

an analogue for the intraclass correlation. More important, the roots of (5)

[1] This work was done during the academic year 1939–40 when the author was a graduate student at Princeton University; it was completed just as the Hsu and Fisher papers appeared, and was therefore never submitted for publication. Recently the author learned from Hotelling that a derivation of the normalizing constants would be of interest.

are directly related to Hotelling's canonical correlations [5] for two sets of variates. For all these problems it is necessary only to obtain the distribution for the roots of (5) since the roots of (4) are

$$v_i = u_i/(1 - u_i) \tag{6}$$

and the distribution of the w_i may be obtained by letting $n \to \infty$ in the distribution of the u_i.

3. Density function for the u_i. It is no essential simplification to suppose, as we shall do, that

$$\begin{aligned} \sigma_{ij} = \delta_{ij} &= 1 \quad \text{for} \quad i = j, \\ &= 0 \quad \text{for} \quad i \neq j. \end{aligned} \tag{7}$$

The joint density for a_{ij} and b_{ij} is

$$f(a_{ij}, m - 1, \sigma^{ij}) f(b_{ij}, n - 1, \sigma^{ij}), \tag{8}$$

where f is defined by (2). If $u_i (u_1 < u_2 < \cdots < u_k)$ are the roots of (5), there exists [6] a nonsingular linear transformation $\| q^{ij} \|$ such that

$$\| q^{ij} \|' \, \| a_{ij} + b_{ij} \| \, \| q^{ij} \| = \| \delta_{ij} \|, \tag{9}$$

$$\| q^{ij} \|' \, \| a^{ij} \| \, \| q^{ij} \| = \| u_i \delta_{ij} \|, \tag{10}$$

$$\| q^{ij} \|' \, \| b_{ij} \| \, \| q^{ij} \| = \| (1 - u_i) \delta_{ij} \|, \tag{11}$$

where the prime denotes the transpose.

We shall transform the $k^2 + k$ variates a_{ij} and b_{ij} of (8) to the $k^2 + k$ variates q_{ij} and u_i where

$$\| q_{ij} \| = \| q^{ij} \|^{-1}.$$

The transformed density is

$$D_1(u_i, q_{ij}) = K_1 F(q_{ij}) \left[\prod_1^k u_i \right]^{\frac{1}{2}(m-k-2)} \left[\prod_1^k (1 - u_i) \right]^{\frac{1}{2}(n-k-2)} J, \tag{12}$$

where J is the Jacobian $\partial(a_{ij}, b_{ij})/\partial(u_i, q_{ij})$ and K_1 is the normalizing constant.

We next show that J factors into a function of q_{ij} only and a function of u_i only. Let the earlier variables be ordered

$$a_{11}, a_{12}, \cdots, a_{1k}, a_{22}, a_{23}, \cdots, a_{2k}, a_{33}, \cdots, a_{3k}, \cdots, a_{kk},$$
$$b_{11}, \cdots, b_{1k}, b_{22}, \cdots, b_{2k}, \cdots, b_{kk},$$

and the new variables will be ordered

$$u_1, u_2, \cdots, u_k, q_{11}, q_{21}, \cdots, q_{k1}, q_{12}, q_{22}, \cdots, q_{k2}, \cdots, q_{1k}, \cdots, q_{kk}.$$

On differentiating the relations

$$a_{ij} = \sum_r q_{ir} q_{jr} u_r , \tag{13}$$

$$b_{ij} = \sum_r q_{ir} q_{jr} (1 - u_r), \tag{14}$$

the Jacobian can be written down directly. Supposing this to have been done (with the a_{ij} and b_{ij} corresponding to columns and the u_i and q_{ij} to rows), the result can be simplified by adding the first column of the left half to the first column of the right half, the second column of the left half to the second column of the right half, etc. The first row of the resulting determinant then has elements

$$q_{11}^2 , q_{11}q_{21} , q_{11}q_{31} , \cdots , q_{11}q_{k1} , q_{21}^2 , q_{21}q_{31} , \cdots , q_{21}q_{k1} , \cdots , q_{k1}^2$$

in the left half, and zeros in the right half. The $(k + 1)$th row, for example, has elements

$$2q_{11}u_1 , q_{21}u_1 , \cdots , q_{k1}u_1 , 0, 0, \cdots , 0$$

in the left half, and the same set with the u_1's omitted in the right half.

Now we show that J vanishes if $u_1 = u_2$. It will be easy to follow the argument if one writes down the complete Jacobian for $k = 3$. Assuming $u_1 = u_2$, the following steps produce a row of zeros in J:

1) Multiply the columns of the right half by u_1 and subtract from the corresponding columns of the left half. This makes the elements of the left half of rows $k + 1$ through $3k$ all zero.
2) Make all elements of the $(k + 1)$th row zero except the element $2q_{11}$ (in the b_{11} column) by subtracting proper multiples of the b_{11} column from the columns having nonzero elements in that row.
3) Make all elements of the $(k + 2)$th row zero except that in the b_{12} column.
4) Make all elements of the $(k + 3)$th row zero except that in the b_{13} column.

....

....

....

$k + 1$) Make all elements of the $(2k)$th row zero except that in the b_{1k} column.

$k + 2$) Make all elements of the $(2k + 1)$th row zero by subtracting proper multiples of the k rows above it from that row.

It follows therefore that J has the factor $(u_2 - u_1)$.

Similarly J must have all factors of the form $u_i - u_j$; hence J has the factor

$$\prod_{i>j} (u_i - u_j), \tag{15}$$

and since J is of total degree $k(k - 1)/2$ in the u's the other factor of J must involve only the q's.

Thus it follows that (12) factors into a function of the q_{ij} only and a function of the u_i only, say

$$(16)\qquad D_2(u_i) = K_2\left[\prod_1^k u_i\right]^{\frac{1}{2}(m-k-2)}\cdot\left[\prod_1^k (1-u_i)\right]^{\frac{1}{2}(n-k-2)}\prod_{i>j}(u_i - u_j).$$

4. Normalizing constant. Let us define

$$(17)\qquad L(\alpha,\beta) = \int_0^1\int_0^{u_k}\cdots\int_0^{u_2}[\prod u_i]^{\alpha}[\prod(1-u_i)]^{\beta}[\prod_{i>j}(u_i-u_j)]\prod du_i;$$

then the normalizing constant of (16) is

$$(18)\qquad K_2 = 1/L[\tfrac{1}{2}(m-k-2), \tfrac{1}{2}(n-k-2)].$$

Our procedure will be to first express $L(\alpha, \beta)$ as a multiple of $L(0, 0)$ and then to evaluate the latter factor directly.

In view of (9), (10), and (11),

$$(19)\qquad \Pi u_i = |a_{ij}|/|a_{ij}+b_{ij}|,$$

$$(20)\qquad \Pi(1-u_i) = |b_{ij}|/|a_{ij}+b_{ij}|;$$

hence

$$(21)\qquad E\left(\frac{|a_{ij}|^r|b_{ij}|^s}{|a_{ij}+b_{ij}|^{r+s}}\right) = \frac{L[\frac{1}{2}(m-k-2)+r, \frac{1}{2}(n-k-2)+s]}{L[\frac{1}{2}(m-k-2), \frac{1}{2}(n-k-2)]}.$$

But this quantity is determinable from (8) by a method due to Wilks [7]. Since the elements of $\|a_{ij}+b_{ij}\|$ are distributed by

$$(22)\qquad f(a_{ij}+b_{ij}, m+n-2, \sigma^{ij}),$$

we find first that

$$(23)\qquad E(|a_{ij}+b_{ij}|^c) = |\tfrac{1}{2}\sigma^{ij}|^{-c}\prod_{i=1}^k\frac{\Gamma[\frac{1}{2}(m+n-1-i)+c]}{\Gamma[\frac{1}{2}(m+n-1-i)]},$$

as does Wilks in [7]. Thus

$$(24)\qquad \begin{aligned}\int\cdots\int|a_{ij}+b_{ij}|^c f(a_{ij}, m-1, \sigma^{ij})f(b_{ij}, n-1, \sigma^{ij})\Pi da_{ij}\,\Pi db_{ij}\\ = |\tfrac{1}{2}\sigma^{ij}|^{-c}\prod_1^k\frac{\Gamma[\frac{1}{2}(m+n-1-i)+c]}{\Gamma[\frac{1}{2}(m+n-1-i)]}\end{aligned}$$

or in another form

$$(25)\qquad \begin{aligned}\int\cdots\int|a_{ij}+b_{ij}|^c|a_{ij}|^{\frac{1}{2}(m-k-2)}|b_{ij}|^{\frac{1}{2}(n-k-2)}e^{[-\frac{1}{2}\Sigma\sigma^{ij}(a_{ij}+b_{ij})]}\Pi\, da_{ij}\,\Pi\, db_{ij}\\ = \frac{\pi^{\frac{1}{4}k(k-1)}}{|\frac{1}{2}\sigma^{ij}|^{\frac{1}{2}(m+n-2)+c}}\prod_{i=1}^k\frac{\Gamma[\frac{1}{2}(m-i)]\Gamma[\frac{1}{2}(n-i)]\Gamma[\frac{1}{2}(m+n-1-i)+c]}{\Gamma[\frac{1}{2}(m+n-1-i)]}.\end{aligned}$$

In this expression we replace m by $m + 2r$ and n by $n + 2s$ and multiply the whole by

$$\frac{|\tfrac{1}{2}\sigma^{ij}|^{\frac{1}{2}(m+n-2)}}{\pi^{\frac{1}{4}k(k-1)}\prod_{i=1}^{k}\Gamma[\tfrac{1}{2}(m-i)]\Gamma[\tfrac{1}{2}(n-i)]}$$

to get

$$E(|a_{ij}+b_{ij}|^{c}\,|a_{ij}|^{r}\,|b_{ij}|^{s}) = |\tfrac{1}{2}\sigma^{ij}|^{-c-r-s} \tag{26}$$
$$\cdot\prod_{i=1}^{k}\frac{\Gamma[\frac{1}{2}(m-i)+r]\Gamma[\frac{1}{2}(n-i)+s]\Gamma[\frac{1}{2}(m+n-1-i)+c+r+s]}{\Gamma[\frac{1}{2}(m+n-1-i)+r+s]\Gamma[\frac{1}{2}(m-i)]\Gamma[\frac{1}{2}(n-i)]}.$$

In this we put $c = -(r + s)$ to get an expression for the right side of (21). In the resulting relation we put $m = n = k + 2$ to get

$$L(r,s) = L(0,0)\prod_{i=1}^{k}\frac{\Gamma[\frac{1}{2}(k+2-i)+r]\Gamma[\frac{1}{2}(k+2-i)+s]\Gamma[\frac{1}{2}(2k+3-i)]}{\Gamma[\frac{1}{2}(2k+3-i)+r+s]\Gamma[\frac{1}{2}(k+2-i)]\Gamma[\frac{1}{2}(k+2-i)]}. \tag{27}$$

Now we are left only with the problem of evaluating

$$L(0,0) = \int\cdots_{R}\int\prod_{i>j}(u_i - u_j)\prod du_i, \tag{28}$$

where R is the region $0 \le u_1 \le u_2 \le \cdots \le u_k \le 1$. We first observe that the integrand may be put in the determinantal form

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ u_1 & u_2 & u_3 & \cdots & u_k \\ u_1^2 & u_2^2 & u_3^2 & \cdots & u_k^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_1^{k-1} & u_2^{k-1} & u_3^{k-1} & \cdots & u_k^{k-1} \end{vmatrix}. \tag{29}$$

Thus the integrand may be written

$$\prod_{i>j}(u_i - u_j) = \sum_{\text{per}}(-1)^{t(\text{per})}\prod_{i=1}^{k}u_i^{\alpha_i-1}, \tag{30}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k$ is a permutation of $1, 2, \cdots k$; where the sum is over all permutations of these integers; and where t(per) is the number of transpositions in the permutation.

On integrating (30) over R it is found that

$$L(0,0) = \sum_{\text{per}}\frac{(-1)^{t(\text{per})}}{\alpha_1(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+\alpha_3)\cdots(\alpha_1+\alpha_2+\cdots+\alpha_k)}. \tag{31}$$

It is shown in the Appendix that this sum has the value

$$(32) \quad L(0, 0) = \frac{1}{k!} \prod_{i>j} \frac{i-j}{i+j}$$

$$(33) \quad = \frac{1}{k!} \frac{(k-1)!(k-2)! \cdots 2!1!}{[3\cdot4\cdot5 \cdots (k+1)][5\cdot6 \cdots (k+2)][7\cdot8 \cdots (k+3)] \cdots [2k-1]}$$

$$(34) \quad = \prod_{i=1}^{k} \frac{\Gamma(k-1+i)\Gamma(2k+1-2i)}{\Gamma(2k+1-i)}.$$

This may also be put in the form

$$(35) \quad L(0, 0) = \frac{1}{\pi^{k/2}} \prod_{i=1}^{k} \frac{\Gamma[\frac{1}{2}(k+2-i)]\Gamma[\frac{1}{2}(k+2-i)]\Gamma[\frac{1}{2}(k+1-i)]}{\Gamma[\frac{1}{2}(2k+3-i)]}.$$

The identity of (34) and (35) is easily shown by induction on k employing the relation

$$(36) \quad \Gamma(h+1)\Gamma(h+\tfrac{1}{2}) = \sqrt{\pi}\Gamma(2h+1)/2^{2h}.$$

The form (35) simplifies the final expression for K_2 which is found by putting (27) and (35) in (18) to get

$$(37) \quad K_2 = \pi^{k/2} \prod_{i=1}^{k} \frac{\Gamma[\frac{1}{2}(m+n-1-i)]}{\Gamma[\frac{1}{2}(m-i)]\Gamma[\frac{1}{2}(n-i)]\Gamma[\frac{1}{2}(k+1-i)]}.$$

Putting (37) in (16) we have the density function for the roots of (5), and the densities for the roots of (3) and (4) can then be obtained as stated at the end of Section 2.

Appendix

We wish to demonstrate that if $g_i (i = 1, 2, \cdots, k)$ are k distinct positive quantities indexed in order of magnitude, then

$$(a) \quad \sum_{\text{per}} \frac{(-1)^{t(\text{per})}}{y_1(y_1+y_2) \cdots (y_1+y_2+\cdots+y_k)} \equiv \left(\prod_{i>j} \frac{g_i - g_j}{g_i + g_j}\right) \Big/ \prod_{1}^{k} g_i,$$

where $y_1, y_2, \cdots, y_k$ is a permutation of $g_1, g_2, \cdots, g_k$, the sum is taken over all permutations of the g_i, and t(per) is the number of transpositions in the permutation $y_1, \cdots, y_k$. This identity was first formulated and proved for $g_i = i$ with considerable aid from J. B. Rosser. Here we give a different and easier argument which handles the more general situation.

First we obtain another identity as a lemma, namely,

$$(b) \quad \sum_{i=1}^{k} g_i \prod_{j \neq i} \frac{g_i + g_j}{g_i - g_j} \equiv \sum_{i=1}^{k} g_i.$$

The following argument for (b) was formulated by John Nash. The left side of (b) is a rational function of g_1, say $P(g_1)/Q(g_1)$, which we may suppose to be reduced to its lowest terms. We first argue that the rational function is really a polynomial becuase it does not become infinite for any finite value of g_1. Certainly the only possible roots of $Q(g_1)$ are $g_2, \cdots, g_k$. Suppose $g_1 = g_2 + \epsilon$, then the first two terms (the others do not have $g_1 - g_2$ in their denominators) on the left of (b) may be written

$$\frac{2g_2 + \epsilon}{\epsilon}\left[(g_2 + \epsilon)\prod_3^k \frac{g_2 + \epsilon + g_j}{g_2 + \epsilon - g_j} - g_2 \prod_3^k \frac{g_2 + g_j}{g_2 - g_j}\right],$$

which is clearly bounded as $\epsilon \to 0$. Similarly no other g_i is a root of $Q(g_1)$; hence the left side of (b) is a polynomial $P(g_1)$ in g_1. Now let g_1 become large; the first term on the left of (b) becomes essentially g_1 while the others become constants; hence

$$P(g_1) = g_1 + C_1(g_2, \cdots, g_k).$$

Similarly as a function of g_2 the left of (b) is of the form

$$g_2 + C_2(g_1, g_3, \cdots, g_k),$$

and so forth. Furthermore, the left of (b) is homogeneous of degree one in the g's; it must therefore be $\sum_1^k g_i$.

Having (b) we can prove (a) by induction. It is true for $k = 2$, and we shall show it to be true for $k + 1$ given it to be true for k. Applying (a) to the left side of the following relation, we have

$$\sum_{\text{per}} \frac{(-1)^{t(\text{per})}}{y_1(y_1 + y_2)(y_1 + y_2 + y_3) \cdots (y_1 + y_2 + \cdots + y_{k+1})}$$
$$= \sum_{\substack{\text{per} \\ y_1 < y_2 < \cdots < y_k}}' \frac{1}{\prod_1^k y_i}\left[\prod_{i>j}^k \frac{y_i - y_j}{y_i + y_j}\right]\frac{(-1)^{t(\text{per})}}{y_1 + y_2 + \cdots y_{k+1}},$$

where the sum on the right is over all permutations which have $y_1 < y_2 < y_3 < \cdots < y_k$. This means that the sum has only $k + 1$ terms; these terms arise from putting y_{k+1} equal to $g_1, g_2, \cdots, g_{k+1}$ in turn and arranging the other g's in ascending order. Thus the right side of this last relation may be written

$$\sum_{y=g_1}^{g_{k+1}} \frac{y}{\prod_1^{k+1} g_i}\left[\prod_{i>j}^{k+1} \frac{g_i - g_j}{g_i + g_j}\right]\left[\prod_{\substack{j=1 \\ g_j \neq y}}^{k+1} \frac{y + g_j}{|y - g_j|}\right]\frac{(-1)^{k+1-j}}{\sum_1^{k+1} g_i}$$
$$= \left[\left(\prod_{i>j}^{k+1} \frac{g_i - g_j}{g_i + g_j}\right)\Big/\prod_1^{k+1} g_i\right]\left[\sum_{i=1}^{k+1} g_i\left(\prod_{j\neq i} \frac{g_i + g_j}{g_i - g_j}\right)\frac{1}{\sum g_i}\right],$$

and the final bracket is unity in view of (b).

REFERENCES

[1] M. A. GIRSHICK, "On the sampling theory of the roots of determinantal equations," *Annals of Math. Stat.*, Vol. 10 (1939), pp. 203–224.
[2] R. A. FISHER, "The sampling distribution of some statistics obtained from non-linear equations," *Annals of Eugenics*, Vol. 9 (1939), pp. 238–249.
[3] P. L. HSU, "On the distribution of roots of certain determinantal equations," *Annals of Eugenics*, Vol. 9 (1939), pp. 250–258.
[4] S. N. ROY, "p-statistics, or some generalizations on the analysis of variance appropriate to multivariate problems," *Sankhyā*, Vol. 3 (1939), pp. 341–396.
[5] H. HOTELLING, "Relations between two sets of variates," *Biometrika*, Vol. 28 (1936), pp. 321–377.
[6] M. BÔCHER, *Introduction to Higher Algebra*, Macmillan, 1929, p. 171.
[7] S. S. WILKS, "Certain generalizations in the analysis of variance," *Biometrika*, Vol. 24 (1932), pp. 471–494.