ON THE DISTRIBUTION OF ROOTS OF CERTAIN DETERMINANTAL EQUATIONS

By P. L. HSU

THE extension (Fisher, 1938) of Fisher's discriminant analysis to more than two multivariate samples has directed attention to the problem of the exact distribution of the roots of a certain type of determinantal equation, which are required for various significance tests. While Fisher was solving this problem he submitted it to me for its interest in relation to matrix algebra. The purpose of the present paper is to give a complete demonstration of the analytic solution, including the case in which the number of variates, p, exceeds one of the sample numbers n_1.

Consider a set of $p(n_1+n_2)$ random variables

$$y_{ir},\ z_{it} \quad (i = 1, 2, \ldots, p;\ r = 1, 2, \ldots, n_1;\ t = 1, 2, \ldots, n_2),$$

following the distribution law

$$\text{const.}\ \exp\left[-\tfrac{1}{2}\sum_{i,j=1}^{p} \alpha_{ij}(a_{ij}+b_{ij})\right]\prod dy\,dz, \qquad \ldots\ldots(1)$$

where

$$a_{ij} = \sum_{r=1}^{n_1} y_{ir}y_{jr}, \quad b_{ij} = \sum_{t=1}^{n_2} z_{it}z_{jt}.$$

We shall assume that $n_2 \geqslant p$, so that the matrix $\|b_{ij}\|$ is almost always positively definite. Thus the determinantal equation

$$|\,a_{ij}-\theta(a_{ij}+b_{ij})\,| = 0 \qquad \ldots\ldots(2)$$

is known to possess exactly p or n_1 (whichever is smaller) real, not identically vanishing, roots, each lying between 0 and 1. Denoting them by $\theta_1, \theta_2, \ldots$, in the order of descending magnitude, we shall now establish the simultaneous distribution of these θ's.

THEOREM 1. *The simultaneous distribution of the roots of* (2) *is given by*

$$CP\left\{\prod_{i=1}^{p}\theta_i\right\}^{\frac{1}{2}(n_1-p-1)}\left\{\prod_{i=1}^{p}(1-\theta_i)\right\}^{\frac{1}{2}(n_2-p-1)}\prod_{i=1}^{p} d\theta_i,\ \textit{if}\ p\leqslant n_1, \qquad \ldots\ldots(3)$$

and

$$C_1P_1\left\{\prod_{i=1}^{n_1}\theta_i\right\}^{\frac{1}{2}(p-n_1-1)}\left\{\prod_{i=1}^{n_1}(1-\theta_i)\right\}^{\frac{1}{2}(n_2-p-1)}\prod_{i=1}^{n_1} d\theta_i,\ \textit{if}\ n_1\leqslant p, \qquad \ldots\ldots(4)$$

where

$$C = \pi^{\frac{1}{2}p}\prod_{i=1}^{p}\frac{\Gamma\frac{1}{2}(n_1+n_2-i+1)}{\Gamma\frac{1}{2}(n_1-i+1)\,\Gamma\frac{1}{2}(n_2-i+1)\,\Gamma\frac{1}{2}(p-i+1)},$$

$$C_1 = \pi^{\frac{1}{2}n_1}\prod_{i=1}^{n_1}\frac{\Gamma\frac{1}{2}(n_1+n_2-i+1)}{\Gamma\frac{1}{2}(n_1-i+1)\,\Gamma\frac{1}{2}(n_1+n_2-p-i+1)\,\Gamma\frac{1}{2}(p-i+1)},$$

and

$$P = \prod_{i=1}^{p}\prod_{j=i+1}^{p}(\theta_i-\theta_j), \quad P_1 = \prod_{i=1}^{n_1}\prod_{j=i+1}^{n_1}(\theta_i-\theta_j).$$

I. *Proof of* (3), case $n_1 \geqslant p$.

Instead of the θ_i, consider new variables ϕ_i, where

$$\phi_i = \theta_i(1-\theta_i)^{-1},$$

so that $\phi_1, \phi_2, \ldots, \phi_p$ are the roots, in descending order, of the equation

$$|a_{ij} - \phi b_{ij}| = 0. \qquad \ldots\ldots(5)$$

Obviously we can set, in (1), $\alpha_{ii} = 1$ and $\alpha_{ij} = 0$ for $i \neq j$. Hence, instead of (1) we may regard the parent population distribution as

$$\{\sqrt{(2\pi)}\}^{-p(n_1+n_2)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p}(a_{ii}+b_{ii})\right]\Pi\, dy\, dz. \qquad \ldots\ldots(6)$$

Since $n_1 \geqslant p$ we can write down the distribution law (Wishart & Bartlett, 1933) of the a_{ij} and b_{ij}:

$$C_2\,|a_{ij}|^{\frac{1}{2}(n_1-p-1)}\,|b_{ij}|^{\frac{1}{2}(n_2-p-1)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p}(a_{ii}+b_{ii})\right]\Pi\, da\, db, \qquad \ldots\ldots(7)$$

where $$C_2 = 2^{-\frac{1}{2}p(n_1+n_2)}\pi^{-\frac{1}{4}p(p-1)}\left\{\prod_{i=1}^{p}\Gamma\tfrac{1}{2}(n_1-i+1)\,\Gamma\tfrac{1}{2}(n_2-i+1)\right\}^{-1}.$$

In the space of the a_{ij} and b_{ij} we shall consider only those points for which the following conditions are satisfied:

(i) that the matrix $\|b_{ij}\|$ is positively definite,

(ii) that equation (5) has no multiple roots.

The remaining points in the space will form a set whose measure is zero. Consider now the transformation

$$\|a_{ij}\| = WD_\phi W', \quad \|b_{ij}\| = WW', \qquad \ldots\ldots(8)$$

where

$$W = \begin{Vmatrix} w_{11} & w_{12} & \cdots & w_{1p} \\ w_{21} & w_{22} & \cdots & w_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ w_{p1} & w_{p2} & \cdots & w_{pp} \end{Vmatrix}$$

is a matrix of real elements, W denotes the transposed matrix of W, and D_ϕ is the diagonal matrix

$$D_\phi = \begin{Vmatrix} \phi_1 & & & \\ & \phi_2 & & \\ & & \cdots & \\ & & & \phi_p \end{Vmatrix}$$

Since the ϕ_i are all distinct and arranged in fixed order of descending magnitude, it can easily be shown that the matrix W is uniquely determined by the a_{ij} and b_{ij} *except only that any entire column of W may change its sign.* Hence, if we impose the condition that the first row of W may assume positive values only, the equations (8) will establish a (1, 1)-

correspondence between the quantities a_{ij}, b_{ij} and the quantities ϕ_i, w_{ij}. The domain of the new variables will accordingly be specified as follows:

$$\infty > \phi_1 > \phi_2 > \ldots > \phi_p \geqslant 0,$$

$$0 \leqslant w_{1j} < \infty \quad (j = 1, 2, \ldots, p),$$

$$-\infty < w_{ij} < \infty \quad (i = 2, \ldots, p;\ j = 1, 2, \ldots, p).$$

We may now proceed to compute the Jacobian, say Δ, of the transformation (8).

We shall establish the formula

$$\Delta = \pm 2^p \mid W \mid^{p+2} P', \qquad \ldots\ldots(9)$$

where

$$P' = \prod_{i=1}^{p} \prod_{j=i+1}^{p} (\phi_i - \phi_j),$$

for the case $p = 3$. It is seen that the method of proof will hold good for any p, but owing to cumbersome notation we restrict ourselves to the particular case.

For $p = 3$ let us arrange the variables so that

$$\Delta = \frac{\partial(a_{11}, a_{12}, a_{13}, a_{22}, a_{23}, a_{33}, b_{11}, b_{12}, b_{13}, b_{22}, b_{23}, b_{33})}{\partial(\phi_1, \phi_2, \phi_3, w_{11}, w_{12}, w_{13}, w_{21}, w_{22}, w_{23}, w_{31}, w_{32}, w_{33})}.$$

Then

$$\Delta = \begin{vmatrix}
w_{11}^2 & w_{12}^2 & w_{13}^2 & 2\phi_1 w_{11} & 2\phi_2 w_{12} & 2\phi_3 w_{13} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
w_{11}w_{21} & w_{12}w_{22} & w_{13}w_{23} & \phi_1 w_{21} & \phi_2 w_{22} & \phi_3 w_{23} & \phi_1 w_{11} & \phi_2 w_{12} & \phi_3 w_{13} & \cdot & \cdot & \cdot \\
w_{11}w_{31} & w_{12}w_{32} & w_{13}w_{33} & \phi_1 w_{31} & \phi_2 w_{32} & \phi_3 w_{33} & \cdot & \cdot & \cdot & \phi_1 w_{11} & \phi_2 w_{12} & \phi_3 w_{13} \\
w_{21}^2 & w_{22}^2 & w_{23}^2 & \cdot & \cdot & \cdot & 2\phi_1 w_{21} & 2\phi_2 w_{22} & 2\phi_3 w_{23} & \cdot & \cdot & \cdot \\
w_{21}w_{31} & w_{22}w_{32} & w_{23}w_{33} & \cdot & \cdot & \cdot & \phi_1 w_{31} & \phi_2 w_{32} & \phi_3 w_{33} & \phi_1 w_{21} & \phi_2 w_{22} & \phi_3 w_{23} \\
w_{31}^2 & w_{32}^2 & w_{33}^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2\phi_1 w_{31} & 2\phi_2 w_{32} & 2\phi_3 w_{33} \\
\cdot & \cdot & \cdot & 2w_{11} & 2w_{12} & 2w_{13} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & w_{21} & w_{22} & w_{23} & w_{11} & w_{12} & w_{13} & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & w_{31} & w_{32} & w_{33} & \cdot & \cdot & \cdot & w_{11} & w_{12} & w_{13} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2w_{21} & 2w_{22} & 2w_{23} & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & w_{31} & w_{32} & w_{33} & w_{21} & w_{22} & w_{23} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2w_{31} & 2w_{32} & 2w_{33}
\end{vmatrix} \qquad \ldots\ldots(10)$$

Denote by W_{ij} the co-factor of w_{ij}. In order to evaluate the determinant Δ, multiply it by another determinant, Δ_1, according to the column-by-column law:

$$\Delta_1 = \begin{vmatrix}
W_{11}^2 & W_{12}^2 & W_{13}^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
2W_{11}W_{21} & 2W_{12}W_{22} & 2W_{13}W_{23} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
2W_{11}W_{31} & 2W_{12}W_{32} & 2W_{13}W_{33} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
W_{21}^2 & W_{22}^2 & W_{23}^2 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
2W_{21}W_{31} & 2W_{22}W_{32} & 2W_{23}W_{33} & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
W_{31}^2 & W_{32}^2 & W_{33}^2 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
-\phi_1 W_{11}^2 & -\phi_2 W_{12}^2 & -\phi_3 W_{13}^2 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
-2\phi_1 W_{11}W_{21} & -2\phi_2 W_{12}W_{22} & -2\phi_3 W_{13}W_{23} & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
-2\phi_1 W_{11}W_{31} & -2\phi_2 W_{12}W_{32} & -2\phi_3 W_{13}W_{33} & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
-\phi_1 W_{21}^2 & -\phi_2 W_{22}^2 & -\phi_3 W_{23}^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
-2\phi_1 W_{21}W_{31} & -2\phi_2 W_{22}W_{32} & -2\phi_3 W_{23}W_{33} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\
-\phi_1 W_{31}^2 & -\phi_2 W_{32}^2 & -\phi_3 W_{33}^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1
\end{vmatrix}$$

The result of this multiplication is the determinant

$$
\Delta_1\Delta = \begin{vmatrix}
|W|^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & |W|^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & |W|^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
w_{21}^2 & w_{22}^2 & w_{23}^2 & \cdot & \cdot & \cdot & 2\phi_1 w_{21} & 2\phi_2 w_{22} & 2\phi_3 w_{23} & \cdot & \cdot & \cdot \\
w_{21}w_{31} & w_{22}w_{32} & w_{23}w_{33} & \cdot & \cdot & \cdot & \phi_1 w_{31} & \phi_2 w_{32} & \phi_3 w_{33} & \phi_1 w_{21} & \phi_2 w_{22} & \phi_3 w_{23} \\
w_{31}^2 & w_{32}^2 & w_{33}^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2\phi_1 w_{31} & 2\phi_2 w_{32} & 2\phi_3 w_{33} \\
\cdot & \cdot & \cdot & 2w_{11} & 2w_{12} & 2w_{13} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & w_{21} & w_{22} & w_{23} & w_{11} & w_{12} & w_{13} & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & w_{31} & w_{32} & w_{33} & \cdot & \cdot & \cdot & w_{11} & w_{12} & w_{13} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2w_{21} & 2w_{22} & 2w_{23} & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & w_{31} & w_{32} & w_{33} & w_{21} & w_{22} & w_{23} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2w_{31} & 2w_{32} & 2w_{33}
\end{vmatrix}
$$

Δ_1 has the value $2^2 W_{11} W_{12} W_{13} | W |^2$. The product determinant can be reduced in order with the help of Laplace's theorem, and cancelling out the common factors $2^2 | W |^2$, we get the equation

$$
W_{11} W_{12} W_{13} \Delta = 2^3 | W |^5 \begin{vmatrix}
\phi_1 w_{21} & \phi_2 w_{22} & \phi_3 w_{23} & \cdot & \cdot & \cdot \\
\phi_1 w_{31} & \phi_2 w_{32} & \phi_3 w_{33} & \phi_1 w_{21} & \phi_2 w_{22} & \phi_3 w_{23} \\
\cdot & \cdot & \cdot & \phi_1 w_{31} & \phi_2 w_{32} & \phi_3 w_{33} \\
w_{21} & w_{22} & w_{23} & \cdot & \cdot & \cdot \\
w_{31} & w_{32} & w_{33} & w_{21} & w_{22} & w_{23} \\
\cdot & \cdot & \cdot & w_{31} & w_{32} & w_{33}
\end{vmatrix}. \qquad \ldots\ldots(11)
$$

Now (11) is an identity of which both sides are polynomials in the w_{ij}. Hence the right-hand side of (11) must be identically divisible by the polynomial $W_{11} W_{12} W_{13}$. But the determinant $| W |$ is an irreducible polynomial of its elements.* Therefore the other determinant in the right-hand side of (11) is identically divisible by $W_{11} W_{12} W_{13}$. Further, the quotient is of degree zero in the w_{ij} because the dividend and the divisor have the same degree. We conclude, therefore, that

$$
\Delta = k | W |^5, \qquad \ldots\ldots(12)
$$

where k is some quantity depending on the ϕ_i only.

We may determine the value of k from the reduced equation (11). But, keeping the general case in mind, we should go back to the original expression (10) of Δ. If we put therein $w_{ii} = 1$ and $w_{ij} = 0$ for $i \neq j$, we get the value $\pm 2^p P'$. Owing to the identity (12) this is the value of k. The formula (9) is thus established.

By direct substitution into (6) and multiplication by the Jacobian, we get the joint distribution of the new variables, ϕ_i and w_{ij}:

$$
2^p C_2 \, | W'W |^{\frac{1}{2}(n_1+n_2-p)} \left\{ \prod_{i=1}^{p} \phi_i \right\}^{\frac{1}{2}(n_1-p-1)} P' \exp\left[-\tfrac{1}{2} \sum_{i,j=1}^{p} (1+\phi_i) \, w_{ij}^2 \right] \prod d\phi \, dw, \qquad \ldots\ldots(13)
$$

wherein we use $| W'W |^{\frac{1}{2}} = \sqrt{| W |^2}$ instead of $| W |$ because only the absolute value of $| W |$ matters.

It remains to integrate with respect to all the w_{ij} to get the required distribution of the ϕ_i. Instead of the restriction $0 \leqslant w_{1i} (i = 1, 2, \ldots, p)$ we may integrate with respect to all the

* Cf. Bôcher (1929), p. 176, Theorem 1, also p. 213, Theorem 5 and p. 216, Exercise 1.

w_{ij} from $-\infty$ to ∞, and divide the result by 2^p. Using Wilks's method (1932) of evaluating the moments of the generalized variance we easily obtain

$$\int | W'W |^{\frac{1}{2}(n_1+n_2-p)} \exp\left[-\tfrac{1}{2}\sum_{i,j=1}^{p}(1+\phi_i)w_{ij}^2\right]\Pi\, dw = C_3\left\{\prod_{i=1}^{p}(1+\phi_i)\right\}^{-\frac{1}{2}(n_1+n_2)} \quad \text{......(14)}$$

where
$$C_3 = 2^{\frac{1}{2}p(n_1+n_2)}\pi^{\frac{1}{2}p^2}\prod_{i=1}^{p}\frac{\Gamma\frac{1}{2}(n_1+n_2-i+1)}{\Gamma\frac{1}{2}(p-i+1)}.$$

Substituting (14) into (13) and dividing the result by 2^p we get the distribution of the ϕ_i:

$$CP'\left\{\prod_{i=1}^{p}\phi_i\right\}^{\frac{1}{2}(n_1-p-1)}\left\{\prod_{i=1}^{p}(1+\phi_i)\right\}^{-\frac{1}{2}(n_1+n_2)}\prod_{i=1}^{p}d\phi_i, \quad \text{......(15)}$$

because $C_2C_3 = C$.

The transformation
$$\phi_i = \theta_i(1-\theta_i)^{-1}$$
constitutes the last step for the establishment of (3).

II. *Proof of* (4), case $n_1 < p$.

For this case the previous method of proof cannot be used, for now there is no joint probability distribution for the set of variables a_{ij}. Starting with the distribution (6) and writing

$$c_{ij} = a_{ij} + b_{ij},$$

we have, by a straightforward application of Wishart & Bartlett's distribution, the joint probability law of the c_{ij} and the y_{ir}:

$$C_4 \,| c_{ij} - a_{ij} |^{\frac{1}{2}(n_2-p-1)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p} c_{ii}\right]\Pi\, dc\, dy,$$

where
$$C_4 = 2^{-\frac{1}{2}p(n_1+n_2)}\pi^{-\frac{1}{2}pn_1-\frac{1}{4}p(p-1)}\left\{\prod_{i=1}^{p}\Gamma\tfrac{1}{2}(n_2-i+1)\right\}^{-1}.$$

Since $\| c_{ij} \|$ is a positively definite matrix, it can be expressed as a matrix product TT', where T has all its elements real. Writing

$$Y = \left\| \begin{matrix} y_{11} & y_{12} & \cdots & y_{1n_1} \\ y_{21} & y_{22} & \cdots & y_{2n_1} \\ \cdots & \cdots & \cdots & \cdots \\ y_{p1} & y_{p2} & \cdots & y_{pn_1} \end{matrix} \right\|,$$

it follows that
$$\| a_{ij} \| = YY'.$$

We now introduce a new set of variables, namely the elements of the matrix

$$U = \left\| \begin{matrix} u_{11} & u_{12} & \cdots & u_{1n_1} \\ u_{21} & u_{22} & \cdots & u_{2n_1} \\ \cdots & \cdots & \cdots & \cdots \\ u_{p1} & u_{p2} & \cdots & u_{pn_1} \end{matrix} \right\|$$

by means of the transformation $Y = TU$. The Jacobian of this transformation being $|T|^{n_1} = \pm |c_{ij}|^{\frac{1}{2}n_1}$, the joint distribution law of the c_{ij} and u_{ir} is

$$C_4 \, | I - UU' |^{\frac{1}{2}(n_2-p-1)} \, | c_{ij} |^{\frac{1}{2}(n_1+n_2-p-1)} \exp\left[-\tfrac{1}{2} \sum_{i=1}^{p} c_{ii} \right] \prod dc\,du. \qquad \ldots\ldots(16)$$

The roots of (2) are precisely the latent roots of the matrix UU', i.e. the roots of the equation

$$| UU' - \theta I | = 0. \qquad \ldots\ldots(17)$$

Therefore from (16) we may integrate with respect to all the c_{ij}. After the integrations have been performed, what remains is the distribution law of the u_{ir}:

$$C_5 \, | I - UU' |^{\frac{1}{2}(n_2-p-1)} \prod du, \qquad \ldots\ldots(18)$$

where

$$C_5 = \pi^{-\frac{1}{2}pn_1} \prod_{i=1}^{p} \frac{\Gamma \frac{1}{2}(n_1+n_2-i+1)}{\Gamma \frac{1}{2}(n_2-i+1)}.$$

The distribution of the non-vanishing roots of (17) is obtained by the following argument. In the case where $n_1 \geqslant p$ the distribution of $\theta_1, \theta_2, \ldots, \theta_p$, which would be derived from (17) and (18), is already known to be given by the formula (3). Now suppose that $n_1 < p$. Then, as is easily shown, the non-vanishing roots of (17) are precisely the roots, all non-vanishing, of the equation

$$| U'U - \theta I | = 0. \qquad \ldots\ldots(19)$$

Further, it may be seen that (18) is identically the same as

$$C_5 \, | I - U'U |^{\frac{1}{2}(n_2-p-1)} \prod du. \qquad \ldots\ldots(20)$$

Now if we write

$$p' = n_1, \quad n_1' = p, \quad n_2' = n_1 + n_2 - p, \qquad \ldots\ldots(21)$$

$$V = \left\| \begin{matrix} v_{11} & v_{12} & \cdots & v_{1n_1'} \\ v_{21} & v_{22} & \cdots & v_{2n_1'} \\ \cdots & \cdots & \cdots & \cdots \\ v_{p'1} & v_{p'2} & \cdots & v_{p'n_1'} \end{matrix} \right\| = U',$$

then the equation (19) will become

$$| VV' - \theta I | = 0, \qquad \ldots\ldots(22)$$

and the expression (20) will take the form

$$C_5' \, | I - VV' |^{\frac{1}{2}(n_2'-p'-1)} \prod dv, \qquad \ldots\ldots(23)$$

where

$$C_5' = \pi^{-\frac{1}{2}p'n_1'} \prod_{i=1}^{p'} \frac{\Gamma \frac{1}{2}(n_1'+n_2'-i+1)}{\Gamma \frac{1}{2}(n_2'-i+1)} = C_5.$$

Since now we have $n_1' > p'$ the distribution of the roots of (22), as derived from (23), must be exactly the formula (3), provided that the integers p, n_1, and n_2 therein are replaced respectively by p', n_1' and n_2'. If we make such changes in (3) and remember (21), we obtain the formula (4).

The proof of Theorem 1 is now complete.

Theorem 2. *If the* $\frac{1}{2}p(p+1)$ *variables* $s_{ij}(i \leqslant j = 1, 2, \ldots, p)$ *have such a domain of existence that the symmetric matrix* $\| s_{ij} \|$ *is always non-singular, and if they are so distributed that their joint probability density function depends only on the latent roots, say* $\lambda_1, \lambda_2, \ldots, \lambda_p$, *arranged in the order of descending magnitude, of* $\| s_{ij} \|$, *i.e. if*

$$df = g(\lambda_1, \lambda_2, \ldots, \lambda_p) \prod ds_{ij},$$

then the joint distribution law of the λ_i *is the following:*

$$\pi^{\frac{1}{4}p(p+1)} \left\{ \prod_{i=1}^{p} \Gamma \tfrac{1}{2}(p-i+1) \right\}^{-1} \left\{ \prod_{i=1}^{p} \prod_{j=i+1}^{p} (\lambda_i - \lambda_j) \right\} g(\lambda_1, \ldots, \lambda_p) \prod d\lambda. \qquad \ldots\ldots(24)$$

Proof. It is a familiar argument that the general formula (24) will follow if we can find a particular example for g for which (24) holds true. This is because the multiplier of g in (24) is entirely independent of the function g itself. The required example is found in the proof of Theorem 1.

Let us take the case $n_1 \geqslant p$ in the distribution law (18). We have seen that the roots of (17) follow the distribution (3). Writing

$$s_{ij} = \sum_{r=1}^{n_1} u_{ir} u_{jr},$$

we get, as a consequence of Wishart's formula, the following distribution for the s_{ij}:

$$C_6 \mid s_{ij} \mid^{\frac{1}{2}(n_1-p-1)} \mid \delta_{ij} - s_{ij} \mid^{\frac{1}{2}(n_2-p-1)} \prod ds, \qquad \ldots\ldots(25)$$

where $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$, and

$$C_6 = \pi^{-\frac{1}{4}p(p-1)} \prod_{i=1}^{p} \frac{\Gamma \frac{1}{2}(n_1 + n_2 - i + 1)}{\Gamma \frac{1}{2}(n_1 - i + 1)\, \Gamma \frac{1}{2}(n_2 - i + 1)}.$$

Now the probability density function appearing in (25) depends on the latent roots of $\| s_{ij} \|$ only. It is in fact equal to

$$C_6 \left\{ \prod_{i=1}^{p} \lambda_i \right\}^{\frac{1}{2}(n_1-p-1)} \left\{ \prod_{i=1}^{p} (1 - \lambda_i) \right\}^{\frac{1}{2}(n_2-p-1)}. \qquad \ldots\ldots(26)$$

On the other hand, the λ_i follow the distribution (3), with the replacement of θ by λ. In other words, the probability density function of the λ_i is (26) multiplied by the expression which coincides with the multiplier of g in (24). Thus Theorem 2 is proved.

As an application of Theorem 2 let us take the following example. Suppose that the pn random variables y_{ir} $(i = 1, 2, \ldots, p;\ r = 1, 2, \ldots, n)$ follow the distribution

$$\text{const.} \exp\left[-\tfrac{1}{2} \sum_{i,j=1}^{p} \alpha_{ij} s_{ij} \right] \prod dy,$$

where

$$s_{ij} = \sum_{r=1}^{p} y_{ir} y_{jr}.$$

Let $\lambda_1, \lambda_2, \ldots$, in descending order of magnitude, be the not identically vanishing roots of the equation
$$|s_{ij}-\lambda\alpha^{ij}| = 0,$$
where the α^{ij} are the elements of the reciprocal matrix of $\|\alpha_{ij}\|$. In order to obtain the distribution of the λ_i, it is obvious that we may set $\alpha_{ii} = 1$ and $\alpha_{ij} = 0$ for $i \neq j$. Hence we may regard the parent population distribution as

$$(\sqrt{2\pi})^{-pn} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p} s_{ii}\right] \Pi\, dy, \qquad \ldots\ldots(27)$$

and find the distribution of the latent roots $\lambda_1, \lambda_2, \ldots$, of $\|s_{ij}\|$. Now we have, from (27), supposing $n \geqslant p$, the following distribution for the s_{ij}:

$$C_7 |s_{ij}|^{\frac{1}{2}(n-p-1)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p} s_{ii}\right] \Pi\, ds = C_7 \left\{\prod_{i=1}^{p} \lambda_i\right\}^{\frac{1}{2}(n-p-1)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p} \lambda_i\right] \Pi\, ds,$$

where
$$C_7 = 2^{-\frac{1}{2}pn} \pi^{-\frac{1}{4}p(p-1)} \left\{\prod_{i=1}^{p} \Gamma\tfrac{1}{2}(n-i+1)\right\}^{-1}.$$

The probability density function of the $\|s_{ij}\|$ being a function of the λ_i only, we get, by virtue of Theorem 2, the joint distribution of the λ_i when $n \geqslant p$:

$$2^{-\frac{1}{2}pn} \pi^{\frac{1}{2}p} \left\{\prod_{i=1}^{p} \Gamma\tfrac{1}{2}(n-i+1)\, \Gamma\tfrac{1}{2}(p-i+1)\right\}^{-1} \left\{\prod_{i=1}^{p} \prod_{j=i+1}^{p} (\lambda_i - \lambda_j)\right\}$$
$$\times \left\{\prod_{i=1}^{p} \lambda_i\right\}^{\frac{1}{2}(n-p-1)} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{p} \lambda_i\right] \prod_{i=1}^{p} d\lambda_i. \qquad \ldots\ldots(28)$$

In the same manner we get the distribution, when $n \leqslant p$, of $\lambda_1, \lambda_2, \ldots, \lambda_n$ which is the above formula with the letters p and n interchanged. It can be verified that (28) is the limiting form of (15) when we set $\phi_i = n_2^{-1}\lambda_i$ and allow n_2 to approach infinity.

The connexion of the θ's with the problem of several multivariate normal samples is familiar through Fisher's work. We shall proceed to show that the quantities known as the canonical correlations between two sets of random variables, as introduced by Hotelling (1936), are distributed like the square roots of the θ's.

Consider a system of $p+q$ random variables, x_i, and a sample of size n, x_{il} $(i = 1, 2, \ldots, p+q;$ $l = 1, 2, \ldots, n)$. Write

$$\bar{x}_i = \frac{1}{n}\sum_{l=1}^{n} x_{il}, \quad s_{ij} = \sum_{l=1}^{n} (x_{il}-\bar{x}_i)(x_{jl}-\bar{x}_j), \quad (i, j = 1, 2, \ldots, p+q),$$

$$a'_{ij} = -\frac{\begin{vmatrix} 0 & s_{i,p+1} & \cdots & s_{i,p+q} \\ s_{p+1,j} & s_{p+1,p+1} & \cdots & s_{p+1,p+q} \\ \cdots & \cdots & \cdots & \cdots \\ s_{p+q,j} & s_{p+q,p+1} & \cdots & s_{p+q,p+q} \end{vmatrix}}{\begin{vmatrix} s_{p+1,p+1} & \cdots & s_{p+1,p+q} \\ \cdots & \cdots & \cdots \\ s_{p+q,p+1} & \cdots & s_{p+q,p+q} \end{vmatrix}}, \quad b'_{ij} = \frac{\begin{vmatrix} s_{ij} & s_{i,p+1} & \cdots & s_{i,p+q} \\ s_{p+1,j} & s_{p+1,p+1} & \cdots & s_{p+1,p+q} \\ \cdots & \cdots & \cdots & \cdots \\ s_{p+q,j} & s_{p+q,p+1} & \cdots & s_{p+q,p+q} \end{vmatrix}}{\begin{vmatrix} s_{p+1,p+1} & \cdots & s_{p+1,p+q} \\ \cdots & \cdots & \cdots \\ s_{p+q,p+1} & \cdots & s_{p+q,p+q} \end{vmatrix}}$$

$$(i, j = 1, 2, \ldots, p).$$

Then the square roots of the not identically vanishing roots of the equation

$$| a'_{ij} - \theta(a'_{ij} + b'_{ij}) | = 0$$

are Hotelling's canonical correlations between the two sets of variables $(x_1, x_2, \ldots, x_p)$ and $(x_{p+1}, x_{p+2}, \ldots, x_{p+q})$.

It is seen that the quantities a'_{ij} and b'_{ij} are symmetric bi-linear forms (quadratic forms if $i = j$) in the variables $x_{i1}, x_{i2}, \ldots, x_{in}$ and $x_{j1}, x_{j2}, \ldots, x_{jn}$, with coefficients depending on the values of the second set of variables. It may be shown* that if a certain orthogonal transformation, depending on the values of the second set of variables, is applied separately to each of the systems $(x_{i1}, x_{i2}, \ldots, x_{in})$ for $i = 1, 2, \ldots, p$, the bi-linear forms a'_{ij} and b'_{ij} can be simultaneously transformed into a_{ij} and b_{ij}, where

$$a_{ij} = \sum_{r=1}^{n_1} y_{ir} y_{jr}, \quad b_{ij} = \sum_{t=1}^{n_2} y_{i,\, n_1+t}\, y_{j,\, n_1+t} \quad (i, j = 1, 2, \ldots, p),$$

and
$$n_1 = q, \quad n_2 = n - q - 1.$$

Assume now that the set of variables $(x_1, x_2, \ldots, x_p)$ is normally distributed, and that the set of variables $(x_{p+1}, x_{p+2}, \ldots, x_{p+q})$ is distributed independently of the former set, but otherwise in any manner whatsoever† (provided that the matrix $\| s_{p+i,\, p+j} \|$ is almost always positively definite). Then the joint distribution law of the y_{ir} $(i = 1, 2, \ldots, p;$ $r = 1, 2, \ldots, n_1 + n_2)$ is precisely (1). We have therefore proved the following theorem.

Theorem 3. *Let $(x_1, x_2, \ldots, x_p)$ and $(x_{p+1}, x_{p+2}, \ldots, x_{p+q})$ be two mutually independent sets of variables of which the first set is normally distributed. Let $\theta_1, \theta_2, \ldots$ be the squares of the canonical correlations between the two sets, arranged in the descending order of magnitude. Then the joint distribution of the θ_i is given by*

$$K\left\{\prod_{i=1}^{p} \prod_{j=i+1}^{p} (\theta_i - \theta_j)\right\} \left\{\prod_{i=1}^{p} \theta_i\right\}^{\frac{1}{2}(q-p-1)} \left\{\prod_{i=1}^{p} (1-\theta_i)\right\}^{\frac{1}{2}(n-p-q-2)} \prod_{i=1}^{p} d\theta_i$$

if $p \leqslant q$, where $$K = \pi^{\frac{1}{2}p} \prod_{i=1}^{p} \frac{\Gamma\frac{1}{2}(n-i)}{\Gamma\frac{1}{2}(n-q-i)\, \Gamma\frac{1}{2}(p-i+1)\, \Gamma\frac{1}{2}(q-i+1)}.$$

If $q \leqslant p$, the distribution is represented by the above formula with the letters p and q interchanged.

This is a generalization of Hotelling's result for $p = q = 2$.

REFERENCES

M. Bôcher (1929). *Introduction to Higher Algebra.* New York.

W. G. Cochran (1934). "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance." *Proc. Camb. Phil. Soc.* **30**, 178–91.

R. A. Fisher (1938). "The statistical utilization of multiple measurements." *Ann. Eugen., Lond.*, **8**, 376–86.

H. Hotelling (1936). "Relations between two sets of variates." *Biometrika*, **28**, 321–77.

S. S. Wilks (1932). "Certain generalizations in the analysis of variance." *Biometrika*, **24**, 471–94.

J. Wishart & M. S. Bartlett (1933). "The generalized product moment distribution in a normal system." *Proc. Camb. Phil. Soc.* **29**, 260–70.

* E.g. by means of the theorem of Cochran (1934).

† Here we make a generalization of Hotelling's work as he assumes that both sets of variables are normally distributed.