ON THE SAMPLING THEORY OF ROOTS OF DETERMINANTAL EQUATIONS

By M. A. Girshick[1]

In a recent paper[2] Hotelling has considered two functions of the covariances of two sets of variates (having a multivariate normal distribution with s variates in the first set, t variates in the second, $s \leq t$) which he designates by Q and Z and which he defines as follows:

$$Q^2 = \frac{(-1)^s C}{AB} \quad \text{and}^3 \quad Z = \frac{D}{AB} \tag{1.1}$$

where A is the determinant of the covariances among the variates of the first set, B the determinant of the covariances among the variates of the second set, D the determinant of covariances of the two sets taken together, and C a determinant obtained from D by replacing the covariances among the variates of the first set by zeros. Both Q^2 and Z are shown to be invariant under internal linear transformations of either set of variates.

In solving the problem of determining linear functions of the two sets of variates for which the multiple correlation is a maximum, Hotelling arrives at a set of parameters $\rho_1, \rho_2, \cdots, \rho_s$ which he names "canonical correlations" and which are the positive or zero roots of the determinantal polynomial

$$D(\lambda) = \begin{vmatrix} -\lambda\sigma_{11} & \cdots & -\lambda\sigma_{1s} & \sigma_{1,s+1} & \cdots & \sigma_{1,s+t} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -\lambda\sigma_{s1} & \cdots & -\lambda\sigma_{ss} & \sigma_{s,s+1} & \cdots & \sigma_{s,s+t} \\ \sigma_{s+1,1} & \cdots & \sigma_{s+1,s} & -\lambda\sigma_{s+1,s+1} & \cdots & -\lambda\sigma_{s+1,s+t} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_{s+t,1} & \cdots & \sigma_{s+t,s} & -\lambda\sigma_{s+t,s+1} & \cdots & -\lambda\sigma_{s+t,s+t} \end{vmatrix}. \tag{1.2}$$

The ρ's are equal in number to the variates of the first set and bear the following relations to Q and Z:

$$Q^2 = \rho_1^2\rho_2^2 \cdots \rho_s^2 \tag{1.3}$$

$$Z = (1 - \rho_1^2)(1 - \rho_2^2) \cdots (1 - \rho_s^2). \tag{1.4}$$

The corresponding functions for the sample covariances Hotelling designates by q and z, and the sample canonical correlations by $r_1, r_2, \cdots, r_s$. Under the assumption of complete independence between the two sets of variates and

---

[1] Most of this Research was accomplished at Columbia University under a Grant-in-Aid from the Carnegie Corporation of New York.

[2] Harold Hotelling, "Relations Between Two Sets of Variates," *Biometrika*, Vol. XXVIII, Dec. 1936.

[3] The function Z was first considered by S. S. Wilks in *Biometrika*, Vol. XXIV, Nov. 1932.

in the case $s = 2$ and $t = 2$, he shows that the joint distribution of q and z is of the form

$$(1.5) \qquad \tfrac{1}{2}(n-2)(n-3)z^{\frac{1}{2}(n-5)}\,dq\,dz$$

q and z satisfying the inequalities

$$0 \le z \le 1, \qquad 0 \le q \le 1, \qquad z \le (1-q)^2$$

and the joint distribution of the canonical correlations r_1 and r_2 is of the form

$$(1.6) \qquad (n-2)(n-3)(r_1^2 - r_2^2)(1-r_1^2)^{\frac{1}{2}(n-5)}(1-r_2^2)^{\frac{1}{2}(n-5)}\,dr_1\,dr_2$$

where n is one less than the number in the sample for each variate.

I

In Part I of this paper we shall, assuming independence between the two sets, find the joint moments of q and z for a general value of s and t and extend the joint distribution of q and z and hence of the canonical correlations to the case where there are two variates in the first set and any number of variates in the second, i.e. $s = 2$ and $t > 2$.[4]

1. **Joint Moments of q and z.** Since we are assuming complete independence between the two sets of variates we may without any loss of generality represent the sample values of the second set as points on the first t axes of unit distance from the origin in a space of n dimensions. The matrix of observations in the case of s variates in the first set and t variates in the second set will take the form

$$(1.7) \qquad \begin{Vmatrix} x_{11} & x_{12} & x_{13} \cdots & x_{1t} \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} \cdots & x_{2t} \cdots & x_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{s1} & x_{s2} & x_{s3} \cdots & x_{st} \cdots & x_{sn} \\ 1 & 0 & 0 \cdots & 0 \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdots & 1 \cdots & 0 \end{Vmatrix}$$

The polynomial $D(\lambda)$ of (1.2) in terms of sample variances and covariances calculated from (1.7) then becomes

$$(1.8) \qquad D(\lambda) = \begin{vmatrix} -\lambda a_{11} \cdots & -\lambda a_{1s} & x_{11} \cdots & x_{1t} \\ \cdot & \cdot & \cdot & \cdot \\ -\lambda a_{s1} \cdots & -\lambda a_{ss} & x_{s1} \cdots & x_{st} \\ x_{11} \cdots & x_{s1} & -\lambda \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ x_{1t} \cdots & x_{st} & 0 \cdots & -\lambda \end{vmatrix}$$

where $a_{ij} = \sum_1^n x_i x_j$.

[4] This extension is a generalization of Hotelling's method loc. cit.

We multiply the first s rows of (1.8) by λ and factor out λ from the last t columns. This yields

$$(1.9) \qquad D(\lambda) = \lambda^{t-s} \begin{vmatrix} -\lambda^2 a_{11} & \cdots & -\lambda^2 a_{1s} & x_{11} & \cdots & x_{1t} \\ \cdot & & \cdot & \cdot & & \cdot \\ -\lambda^2 a_{s1} & \cdots & -\lambda^2 a_{ss} & x_{s1} & \cdots & x_{st} \\ x_{11} & \cdots & x_{s1} & -1 & \cdots & 0 \\ \cdot & & \cdot & \cdot & & \cdot \\ x_{1t} & \cdots & x_{st} & 0 & \cdots & -1 \end{vmatrix}.$$

As a further simplification, we multiply the $(s + j)^{\text{th}}$ column by x_{ij} for all j from 1 to t and add the result to the i^{th} column. When this is done for every value of i from 1 to s and the resulting determinant expanded by means of the last t columns, the determinantal polynomial (1.9) becomes

$$D(\lambda) = \lambda^{t-s} \begin{vmatrix} b_{11} - \lambda^2 a_{11} & b_{12} - \lambda^2 a_{12} & \cdots & b_{1s} - \lambda^2 a_{1s} \\ \cdot & \cdot & \cdots & \cdot \\ b_{s1} - \lambda^2 a_{s1} & b_{s2} - \lambda^2 a_{s2} & \cdots & b_{ss} - \lambda^2 a_{ss} \end{vmatrix}$$

or symbolically

$$(1.10) \qquad D(\lambda) = \lambda^{t-s} \mid b_{ij} - \lambda^2 a_{ij} \mid$$

where $b_{ij} = \sum_1^t x_i x_j$.

Hence the s roots of $D(\lambda)$ which do not necessarily vanish may be obtained from the polynomial

$$(1.11) \qquad Q(\lambda) = \mid b_{ij} - \lambda^2 a_{ij} \mid.$$

The coefficient of the highest power of λ in $Q(\lambda)$ is given by $\mid a_{ij} \mid$, the determinant of the elements a_{ij}. Taking this in conjunction with (1.3) and (1.4) we see that

$$(1.12) \qquad \begin{aligned} q^2 &= \frac{Q(0)}{\mid a_{ij} \mid} = \frac{\mid b_{ij} \mid}{\mid a_{ij} \mid} \\ z &= \frac{Q(1)}{\mid a_{ij} \mid} = \frac{\mid c_{ij} \mid}{\mid a_{ij} \mid} \end{aligned}$$

where $c_{ij} = \sum_{t+1}^n x_i x_j$.

From the equations (1.12) we obtain

$$(1.13) \qquad E\{\mid a_{ij} \mid^{\frac{1}{2}(\alpha+2\beta)} q^\alpha z^\beta\} = E\{\mid b_{ij} \mid^{\frac{1}{2}\alpha} \mid c_{ij} \mid^\beta\}$$

where E stands for the mathematical expectation of the expressions in the $\{\ \ \}$.

It is obvious from the definition of b_{ij} and c_{ij} that the two determinants $\mid b_{ij} \mid$ and $\mid c_{ij} \mid$ are independently distributed. Moreover, the joint distribution of q and z does not depend on the determinant $\mid a_{ij} \mid$. The truth of the latter statement can be seen from the following geometrical considerations. If we consider the sample values of each variate as a point in an n-dimensional space, then the two sets of variates determine two flat spaces, one of s dimensions and one of t dimensions in that space. A sample canonical correlation can then be considered as the cosine of a certain minimum or stationary angle between two lines, one line lying in the flat s space and the other in the flat t space. Since q and z are functions of the canonical correlations, they therefore depend only on lines and angles between two planes. The quantities a_{ij} on the other hand, depend on lines and angles lying entirely within one of these planes.

From the above considerations we see that equation (1.13) can be written as

$$E\{|a_{ij}|^{\frac{1}{2}(\alpha+2\beta)}\}E(q^{\alpha}z^{\beta}) = E(|b_{ij}|^{\frac{1}{2}\alpha})E(|c_{ij}|^{\beta})$$

or

$$E(q^{\alpha}z^{\beta}) = \frac{E(|b_{ij}|^{\frac{1}{2}\alpha})E(|c_{ij}|^{\beta})}{E(|a_{ij}|^{\frac{1}{2}(\alpha+2\beta)})} \tag{1.14}$$

The m^{th} moment of a determinant $|d_{ij}|$ of sums of sample cross products of p variates is given by the formula[5]

$$E(|d_{ij}|^{m}) = \frac{2^{pm}}{|D_{ij}|^{m}}\prod_{i=1}^{p}\left[\frac{\Gamma\left(\frac{n+2m+1-i}{2}\right)}{\Gamma\left(\frac{n+1-i}{2}\right)}\right], \tag{1.15}$$

where D_{ij} denotes the cofactor corresponding to σ_{ij} divided by the determinant $|\sigma_{ij}|$. Substituting (1.15) in (1.14) and simplifying, we get for the joint moments of q and z

$$E(q^{\alpha}z^{\beta}) = \prod_{i=1}^{s}\left[\frac{\Gamma\left(\frac{t+\alpha+1-i}{2}\right)\Gamma\left(\frac{n-t+2\beta+1-i}{2}\right)\Gamma\left(\frac{n+1-i}{2}\right)}{\Gamma\left(\frac{t+1-i}{2}\right)\Gamma\left(\frac{n-t+1-i}{2}\right)\Gamma\left(\frac{n+\alpha+2\beta+1-i}{2}\right)}\right]. \tag{1.16}$$

2. **Joint Distribution of q and z for $s = 2$, $t > 2$.** In order to determine the joint distribution of q and z for $s = 2$ and $t > 2$, we shall first prove the following lemma.

LEMMA: *Let q and z be defined as in* (1.1) *for two sets of variates having s variates in either set and let q' and z' be similarly defined with $s < t$ where s is the number of variates in the first set and t the number of variates in the second set, then for $n = t + s$, the joint distribution of q^2 and z is identical with that of z' and q'^2.*

PROOF. If the number of variates in either set are the same and $n = t + s$, then by (1.12)

$$q^2 = \frac{|b_{ij}|}{|a_{ij}|}, \qquad z = \frac{|c_{ij}|}{|a_{ij}|}$$

[5] Cf. S. S. Wilks, "Certain Generalizations in the Analysis of Variance," *Biometrika*, Vol. XXIV, Nov. 1932.

where

$$(1.17)\qquad b_{ij} = \sum_1^t x_i x_j, \qquad c_{ij} = \sum_{s+1}^{t+s} x_i x_j, \qquad a_{ij} = \sum_1^{t+s} x_i x_j$$

and $s = t$.

However, for $s < t$, and $n = t + s$, we take for the second set of t variates points on the t axes at unit distance from the origin in the $(t + s)$-dimensional space *perpendicular* to the first s axes. The matrix of observations in this case takes the form

$$(1.18)\qquad \left\| \begin{array}{cccccccc} x_{11} & x_{12} & \cdots & x_{1s} & x_{1,s+1} & \cdots & x_{1,s+t} \\ x_{21} & x_{22} & \cdots & x_{2s} & x_{2,s+1} & \cdots & x_{2,s+t} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ x_{s1} & x_{s2} & \cdots & x_{ss} & x_{s,s+1} & \cdots & x_{s,s+t} \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{array} \right\| .$$

Employing the same arguments as in equations (1.8) (1.9) and (1.10) we find that

$$(1.19)\qquad Q(\lambda) = | c_{ij} - \lambda^2 a_{ij} |, \qquad q'^2 = \frac{| c_{ij} |}{| a_{ij} |}, \qquad z' = \frac{| b_{ij} |}{| a_{ij} |}$$

where

$$b_{ij} = \sum_1^t x_i x_j, \qquad c_{ij} = \sum_{s+1}^{t+s} x_i x_j, \qquad a_{ij} = \sum_1^{t+s} x_i x_j.$$

Comparing these equations with (1.17) we see that

$$(1.20)\qquad z = q'^2, \qquad q^2 = z'.$$

This proves the lemma.

Now let $s = 2$. Setting $n = t + 2$ in equation (1.5) and using the transformation (1.20) we get for the joint distribution of q' and z'

$$(1.21)\qquad \tfrac{1}{2}t(t - 1)q'^{t-2}z'^{-\frac{1}{2}}\, dq'\, dz'.$$

Let r be the correlation between the two variates of the first set. The distribution of r in samples for which $n = t + 2$ when the population correlation is zero is known to be

$$(1.22)\qquad \frac{\Gamma\left(\frac{t+2}{2}\right)}{\Gamma\left(\frac{t+1}{2}\right)\sqrt{\pi}} (1 - r^2)^{\frac{1}{2}(t-1)}\, dr.$$

The distribution of r is independent of q and z. Hence, the joint distribution of q', z', and r is given by the product of (1.21) and (1.22). Dropping the primes from q' and z' in (1.21), we get for the joint distribution of the three quantities in the case $n = t + 2$,

$$(1.23) \qquad \frac{1}{2}t(t-1)\frac{\Gamma\left(\frac{t+2}{2}\right)}{\Gamma\left(\frac{t+1}{2}\right)\sqrt{\pi}} q^{t-2}z^{-\frac{1}{2}}(1-r^2)^{\frac{1}{2}(t-1)}\,dq\,dz\,dr.$$

We shall now derive the joint distribution of q and z for a general value of n for $s = 2$, $t > 2$. We set $x_1 = x$, $x_2 = y$ and take the t sample variates of the second set to be points on the first t axes at unit distance from the origin in a space of n dimensions. As in (1.12) calculate q and z.

$$(1.24) \qquad q^2 = \frac{\sum_1^t x^2 \sum_1^t y^2 - \left(\sum_1^t xy\right)^2}{1-r^2}, \qquad z = \frac{\sum_{t+1}^n x^2 \sum_{t+1}^n y^2 - \left(\sum_{t+1}^n xy\right)^2}{1-r^2}.$$

We transform the points $(x_1, \cdots, x_n)$ and $(y_1, \cdots, y_n)$ to hyperspherical coordinates, the transformation to be represented parametrically by the equations

$$(1.25) \qquad \begin{aligned}
x_1 &= \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{t-1}\sin\theta_t \\
x_2 &= \cos\theta_1 \sin\theta_2 \cdots \sin\theta_{t-1}\sin\theta_t \\
x_3 &= \cos\theta_2 \cdots \sin\theta_{t-1}\sin\theta_t \\
&\cdot \quad \cdots \quad \cdot \quad \cdot \\
x_t &= \cos\theta_{t-1}\sin\theta_t \\
x_{t+1} &= \cos\theta_t \cos\theta_{t+1} \\
x_{t+2} &= \cos\theta_t \sin\theta_{t+1}\cos\theta_{t+2} \\
&\cdot \quad \cdot \quad \cdot \\
x_{n-1} &= \cos\theta_t \sin\theta_{t+1}\sin\theta_{t+2}\cdots\cos\theta_{n-1} \\
x_n &= \cos\theta_t \sin\theta_{t+1}\sin\theta_{t+2}\cdots\sin\theta_{n-1}
\end{aligned}$$

with the same representation for the y's in terms of parameters $\phi_1, \phi_2, \cdots \phi_{n-1}$. It is to be observed that in (1.24) and (1.25) $\Sigma x^2 = 1$, $\Sigma y^2 = 1$. This we may assume since q and z are invariant under such transformations.

In this new coordinate system, our samples $(x_1, \cdots, x_n)$ and $(y_1, \cdots, y_n)$ are taken as random points on a unit hypersphere about the origin in n dimensions. There is no loss of generality in this since x and y are assumed to be uncorrelated in the population and hence possess spherical symmetry of the density distribution in a space of n dimensions.

The element of probability for the x points on this hypersphere is proportional to the $(n-1)$-dimensional area on this sphere. Now the $n-1$ dimensional area is given by

$$\sqrt{g}\,d\theta_1\,d\theta_2 \cdots d\theta_{n-1}$$

where g is a determinant of order $n - 1$ in which the element in the i^{th} row and j^{th} column is

$$\sum_{\alpha=1}^{n} \frac{\partial x_\alpha \partial x_\alpha}{\partial \theta_i \partial \theta_j}.$$

When $i \neq j$, all these quantities vanish as can be seen by inspection from (1.25). When $i = j$, we have

$$\sum_1^n \left(\frac{\partial x_\alpha}{\partial \theta_1}\right)^2 = \sin^2 \theta_2 \sin^2 \theta_3 \cdots \sin^2 \theta_t$$

$$\sum_1^n \left(\frac{\partial x_2}{\partial \theta_2}\right)^2 = \sin^2 \theta_3 \cdots \sin^2 \theta_t$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sum_1^n \left(\frac{\partial x_\alpha}{\partial \theta_t}\right)^2 = 1$$

$$\sum_1^n \left(\frac{\partial x_\alpha}{\partial \theta_{t+1}}\right)^2 = \cos^2 \theta_t$$

$$\sum_1^n \left(\frac{\partial x_\alpha}{\partial \theta_{t+2}}\right)^2 = \cos^2 \theta_t \sin^2 \theta_{t+1}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sum_1^n \left(\frac{\partial x_\alpha}{\partial \theta_{n-1}}\right)^2 = \cos^2 \theta_j \sin^2 \theta_{j+1} \cdots \sin^2 \theta_{n-2}.$$

Therefore

$$g = \sin^2 \theta_2 \sin^4 \theta_3 \cdots \sin^{2(t-1)} \theta_t \cos^{2(n-t-1)} \theta_t \cdots \sin^2 \theta_{n-2}$$

and hence the element of generalized area is given by

$$\begin{aligned}(1.26)\qquad &\sin \theta_2 \sin^2 \theta_3 \cdots \sin^{t-1} \theta_t \cos^{n-t-1} \theta_t \\ &\qquad\qquad \sin^{n-t-2} \theta_{t+1} \cdots \sin \theta_{n-2}\, d\theta_1\, d\theta_2 \cdots d\theta_{n-1}.\end{aligned}$$

Similarly we can show that the element of generalized area for the y point is

$$\begin{aligned}(1.27)\qquad &\sin \phi_2 \sin^2 \phi_3 \cdots \sin^{t-1} \phi_t \cos^{n-t-1} \phi_t \\ &\qquad\qquad \sin^{n-t-2} \phi_{t+1} \cdots \sin \phi_{n-2}\, d\phi_1\, d\phi_2 \cdots d\phi_{n-1}.\end{aligned}$$

The joint distribution of $\theta_1, \theta_2, \cdots, \theta_{n-1}$ and $\phi_1, \phi_2, \cdots, \phi_{n-1}$ (since the θ's are independent of the ϕ's) is proportional to the product of (1.26) and (1.27).

We now introduce four new sets of variables, u, v, u', v', defined by the following equations

$$(1.28)\qquad x_i = u_i \sin \theta_t, \qquad y_i = v_i \sin \phi_t \qquad (i = 1, 2, \cdots, t)$$

$$(1.29)\qquad x_j = u_j' \cos \theta_t \qquad y_j = v_j' \cos \phi_t \qquad (j = t + 1, \cdots, n).$$

The u_i and v_i can be regarded as two points on a sphere in a space of t dimensions and u_j' and v_j' as two points on a sphere in a space of $n - t$ dimensions.

Let λ be the angle between the two points u and v and μ the angle between the two points u' and v'. Then

$$\cos \lambda = \sum_{i=1}^{t} u_i v_i ; \qquad \cos \mu = \sum_{j=t+1}^{n} u'_j v'_j .$$

The probability element for λ is proportional to $\sin^{t-2} \lambda \, d\lambda$, and that for μ is proportional to $\sin^{n-t-2} \mu \, d\mu$.

From the definition of u_i and v_i, we see that they depend only on $\theta_1, \theta_2, \cdots, \theta_{t-1}$; $\phi_1, \phi_2, \cdots, \phi_{t-1}$ respectively, and u'_j and v'_j depend only on θ_{t+1}, $\theta_{t+2}, \cdots, \theta_{n-1}$; $\phi_{t+1}, \phi_{t+2}, \cdots, \phi_{n-1}$ respectively. It follows that the quantities λ, μ, θ_t, and ϕ_t are independently distributed.

The joint distribution of the θ's and ϕ's we integrate between constant limits with respect to all the variates except θ_t and ϕ_t. This gives for the joint distribution of θ_t and ϕ_t

$$A_n \sin^{t-1} \theta_t \sin^{t-1} \phi_t \cos^{n-t-1} \theta_t \cos^{n-t-1} \phi_t \, d\theta_t \, d\phi_t$$

where A_n is a constant depending only on n.

Multiplying this by the distributions of λ and μ and dropping the subscript t from θ and ϕ we get for the joint distribution of λ, μ, θ, and ϕ

$$k_n \sin^{t-1} \theta \sin^{t-1} \phi \cos^{n-t-1} \theta \cos^{n-t-1} \phi \sin^{t-2} \lambda \cos^{n-t-2} \mu \, d\theta \, d\phi \, d\lambda \, d\mu \tag{1.30}$$

where k_n is a constant depending on n. The limits of integration for θ and ϕ are 0 and $\pi/2$; for λ and μ they are 0 and π.

Expressing q and z in terms of the new quantities as defined in (1.25), (1.28) and (1.29) we get

$$q^2 = \frac{\left(\sum_1^t x^2\right)\left(\sum_1^t y^2\right) - \left(\sum_1^t xy\right)^2}{1 - r^2} = \frac{\sin^2 \theta \sin^2 \phi \sin^2 \lambda}{1 - r^2} \tag{1.31}$$

$$z = \frac{\left(\sum_{t+1}^n x^2\right)\left(\sum_{t+1}^n y^2\right) - \left(\sum_{t+1}^n xy\right)^2}{1 - r^2} = \frac{\cos^2 \theta \cos^2 \phi \sin^2 \mu}{1 - r^2} \tag{1.32}$$

where

$$r = \Sigma xy = \sin \theta \sin \phi \cos \lambda + \cos \theta \cos \phi \cos \mu \tag{1.33}$$

is the sample correlation between x and y.

We now consider a transformation of the variables θ, ϕ, and μ in (1.30) to the new variables q, z, and r. Without troubling to compute the Jacobian J of the transformation, we know that it is independent of n since the relations (1.31), (1.32) and (1.33) do not involve n. Substituting from (1.31) and (1.32) into (1.30) we get for the joint distribution of q, z, r, and λ

$$k_n \psi q^{t-1} z^{\frac{1}{2}(n-t-1)} (1 - r^2)^{\frac{1}{2}(n-2)} \, dq \, dz \, dr \, d\lambda$$

where ψ is independent of n. Integrating with respect to λ between limits which are independent of n, we get for the joint distribution of q, z, and r

$$k_n \Psi q^{t-1} z^{\frac{1}{2}(n-t-1)} (1 - r^2)^{\frac{1}{2}(n-2)} \, dq \, dz \, dr. \tag{1.34}$$

But, for $n = t + 2$, this joint distribution reduces to (1.23). Therefore

$$k_{t+2}\Psi = \frac{1}{2} t(t-1) \frac{\Gamma\left(\frac{t+2}{2}\right)}{\Gamma\left(\frac{t+1}{2}\right)\sqrt{\pi}} z^{-1} q^{-1} (1 - r^2)^{-\frac{1}{2}}$$

so that (1.34) can be written as

$$k'_n q^{t-2} z^{\frac{1}{2}(n-t-3)} (1 - r^2)^{\frac{1}{2}(n-3)} \, dq \, dz \, dr.$$

However, since the distribution of r is known to be

$$\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\sqrt{\pi}} (1 - r^2)^{\frac{1}{2}(n-3)} \, dr$$

we finally get for the joint distribution of q and z

$$h_n q^{t-2} z^{\frac{1}{2}(n-t-3)} \, dq \, dz$$

where h_n depends on n. The integral over the entire region defined by the inequalities

$$0 \leq q \leq 1, \qquad 0 \leq z \leq 1, \qquad z \leq (1 - q)^2$$

must equal unity; the constant h_n is therefore readily found to be $\frac{(n-2)!}{2(t-2)!\,(n-t-2)!}$. Thus the joint distribution in the final form is

$$\frac{(n-2)!}{2(t-2)!\,(n-t-2)!} q^{t-2} z^{\frac{1}{2}(n-t-3)} \, dq \, dz. \tag{1.35}$$

Now by (1.3) and (1.4), $q = r_1 r_2$, $z = (1 - r_1^2)(1 - r_2^2)$, and hence the Jacobian

$$\frac{\partial(q, z)}{\partial(r_1, r_2)} = 2(r_1^2 - r_2^2). \tag{1.36}$$

Making the transformation in (1.35) we get the joint distribution of the canonical correlations r_1 and r_2 (for the case $s = 2$ and a general value of t) in the form

$$\frac{(n-2)!}{(t-2)!\,(n-t-2)!} (r_1^2 - r_2^2)(r_1 r_2)^{t-2} [(1 - r_1^2)(1 - r_2^2)]^{\frac{1}{2}(n-t-3)} \, dr_1 \, dr_2. \tag{1.37}$$

II. JOINT LIMITING DISTRIBUTIONS OF CANONICAL CORRELATIONS AND LATENT ROOTS

In formula (1.37) we set

$$k_1 = nr_1^2, \qquad k_2 = nr_2^2$$

and get for the joint distribution of k_1 and k_2

$$(2.1) \qquad \frac{(n-2)!}{4(t-2)!\,(n-t-2)!\,n^t}(k_1 - k_2)(k_1k_2)^{\frac{1}{2}(t-3)} \left[\left(1 - \frac{k_1}{n}\right)\left(1 - \frac{k_2}{n}\right)\right]^{\frac{1}{2}(n-t-3)} dk_1\,dk_2 .$$

When $n \to \infty$, the quantity $\frac{(n-2)!}{n^t(n-t-2)!}$ approaches 1 and $\left(1 - \frac{k}{n}\right)^{\frac{1}{2}(n-t-3)}$ approaches $e^{-\frac{1}{2}k^2}$. Hence the limiting distribution of the two canonical correlations is given by

$$(2.2) \qquad \frac{1}{4(t-2)!}(k_1 - k_2)(k_1k_2)^{\frac{1}{2}(t-3)}\,e^{-\frac{1}{2}(k_1+k_2)}\,dk_1\,dk_2 .$$

We shall call (2.2) the "generalized chi-square" distribution and show that the roots of the characteristic polynomial

$$(2.3) \qquad \varphi(k) = \begin{vmatrix} a_{11} - k & a_{12} \\ a_{21} & a_{22} - k \end{vmatrix}$$

are distributed in precisely this form. Here $a_{ij} = \Sigma x_ix_j$ where x_1 and x_2 are normally and independently[6] distributed with unit variance in the population and zero mean in the sample.

Let k_1 and k_2 be the roots of (2.3). That is, k_1 and k_2 are the two roots of the quadratic equation

$$(2.4) \qquad k^2 - p_1k + p_2 = 0$$

where

$$(2.5) \qquad p_1 = k_1 + k_2 = a_{11} + a_{22}$$

$$(2.6) \qquad p_2 = k_1k_2 = a_{11}a_{22} - a_{12}^2 .$$

In the absence of correlation in the population, the joint distribution of a_{11}, a_{22} and a_{12} is known to be

$$(2.7) \qquad h_n \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}(a_{11}+a_{22})}\,da_{11}\,da_{22}\,da_{12}$$

where h_n is a constant depending only on n.

[6] The part of the assumption relating to independence may be removed without loss of generality. See last paragraph below.

We consider a transformation to the variables p_1, p_2 and a_{12}. From (2.5) and (2.6) we calculate the Jacobian J of the transformation,

$$J = \frac{1}{a_{11} - a_{22}} \tag{2.8}$$

and since

$$2a_{ii} = p_1 \pm (p_1^2 - 4p_2 - 4a_{12}^2)^{\frac{1}{2}} \tag{2.9}$$

$$J = \frac{1}{(p_1^2 - 4p_2 - 4a_{12}^2)^{\frac{1}{2}}}. \tag{2.10}$$

Substituting from (2.5) and (2.6) into (2.7) and multiplying by J, we get for the joint distribution of k_1, k_2 and a_{12}

$$h_n p_2^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}p_1} \frac{dp_1\, dp_2\, da_{12}}{(p_1^2 - 4p_2 - 4a_{12}^2)^{\frac{1}{2}}}. \tag{2.11}$$

We make the transformation $u = a_{12}^2$ and get for the joint distribution of k_1, k_2 and u

$$\frac{h_n}{2} p_2^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}p_1} \frac{dp_1\, dp_2\, du}{(bu - 4u^2)^{\frac{1}{2}}} \tag{2.12}$$

where $b = p_1^2 - 4p_2$.

Since both a_{11} and a_{22} are real, equation (2.9) shows that $b - 4u \geq 0$. Hence the limits of integration for u are 0 and $\frac{b}{4}$. Integrating out u in (2.12) between the above limits we obtain the joint distribution of p_1 and p_2.

Now the integral

$$\int_0^{b/4} \frac{du}{(bu - 4u^2)^{\frac{1}{2}}} = -\frac{1}{2} \sin^{-1}\left(\frac{-8u + b}{b}\right)\Bigg]_0^{b/4} = c \tag{2.13}$$

where c is some constant. Hence the joint distribution of p_1 and p_2 is given by

$$H_n p_2^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}p_1}\, dp_1\, dp_2. \tag{2.14}$$

By integrating (2.14) over the region $0 \leq p_2 \leq \left(\frac{p_1}{2}\right)^2$ and $0 \leq p_1 \leq \infty$ we get $H_n = \frac{1}{4}(n - 2)!$.

We next transform p_1 and p_2 in terms of k_1 and k_2 from (2.5) and get for the joint distribution of k_1 and k_2

$$\frac{1}{4(n-2)!} (k_1 - k_2)(k_1 k_2)^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}(k_1+k_2)}\, dk_1\, dk_2. \tag{2.15}$$

This distribution is identical with that of (2.2) with $n = t$.

The above is an example of a more general

THEOREM: *Let $r_1, r_2, \cdots, r_s$ be a set of simple finite canonical roots of the two independent sets of variates $x_1, \cdots, x_s$, and $x_{s+1}, \cdots, x_{s+t}$. Let $k_i = nr_i^2$ $(i = 1, 2, \cdots, s)$. Then the joint limiting distribution of the k's approaches the exact joint sampling distribution of the latent roots of a matrix of sample product sums with t degrees of freedom of s normally distributed variates having unit variance in the population.*

PROOF: The proof follows from equation (1.11). For let us multiply and divide a_{ij} in (1.11) by n and set $n\lambda^2 = k$. The determinantal polynomial becomes

$$\varphi(k) = | b_{ij} - ks_{ij} |. \tag{2.16}$$

Without any loss of generality, we so transform the first set of variates that they become of zero correlation and unit variance in the population. Then it follows that

$$E(s_{ij}) = E\left(\sum \frac{x_i x_j}{n}\right) = \delta_{ij}$$

where δ_{ij} equals zero for $i \neq j$ and 1 for $i = j$.

Now let $P(x \geq a)$ stand for the probability that the variate x be greater than or equal to some constant a. Then, by the Strong Law of Large Numbers we can state that, given an $\epsilon > 0$ and a $\delta > 0$ there exists a positive integer n_0 such that for $n > n_0$

$$P\{| s_{ij} - \delta_{ij}| \geq \delta\} \leq \epsilon.$$

If then we let n increase indefinitely, the quantity $b_{ij} = \sum_1^t x_i x_j$ remains fixed while s_{ij} approaches, in the probability sense, δ_{ij}. Since the roots of a polynomial are continuous functions of the coefficients, we can, by an extension of the Law of Large Numbers, show that in the limit the roots of (2.16) will be distributed like the roots of the polynomial

$$\varphi(k) = | b_{ij} - k\delta_{ij} |.$$

This proves the theorem.

COROLLARY 1. *The limiting distribution of q^2 in case of complete independence between the two sets of variates approaches the exact distribution of a generalized sample variance (i.e. a determinant of sample variances and covariances) with t degrees of freedom. The proof follows from the fact that q^2 is a product of the roots of* (1.11) *and therefore by the above theorem, is distributed in the limit like* $| b_{ij} |$.

COROLLARY 2. *The distribution of the sum of the squares of the canonical correlations approach in the limit a χ^2 distribution with st degrees of freedom. This is obvious since in the limit the sum of the squares of the roots, by the above theorem, has the distribution of $b_{11} + b_{22} + \cdots + b_{ss}$ and each b_{ii} is distributed like χ^2 with t degrees of freedom.*

While the canonical roots of (1.2) are invariant under any non-singular linear transformations, the latent roots of a determinant of sample covariances are invariant only under an orthogonal transformation. But there exists an orthogonal transformation which reduces a set of variates having a multivariate normal distribution to a set which are normally and independently distributed with variances equal to the latent roots of the population generalized variances of the original variates. Hence, in dealing with the distribution of latent roots, we may assume independence in the population without any loss of generality but the assumption of equal variance leads only to a special case. Moreover, the above consideration also explains the form of the asymptotic error of the sample latent root given in Part III of this paper.

III. ASYMPTOTIC STANDARD ERRORS OF LATENT ROOTS AND COEFFICIENTS OF PRINCIPAL COMPONENTS

1. Many statisticians have had occasion to use in their statistical analyses characteristic roots (or as they are sometimes called "latent" roots) of determinants of correlations or covariances. Especially has this become true since the publication of Hotelling's paper on principal components.[7] It is therefore of great importance to find, if not their sampling distributions, at least their limiting distributions and their asymptotic standard errors. This we shall do in this paper for the case of non-vanishing simple roots and by the same method[8] get the asymptotic variances and covariances of the coefficients of principal components. We have already derived in Part II the sampling distribution of the two latent roots of a determinant of covariances obtained from two normally distributed variates having equal variance in the population. This distribution is of no great importance in itself except that it gives us some idea as to the form of the distribution in the general case.

In what follows, we shall use the convention that a repeated subscript in the same term stands for summation. If repeated subscripts appearing in a term are not to be summed, we shall place them in brackets following the expression in which they appear. Thus in the equation (3.1) below, we sum with respect to j but not with respect to q even though on the right hand side q appears twice.

Let $x_1, x_2, \cdots, x_s$ be a set of variates which have a multi-variate normal distribution. We assume that these variates have been resolved into components by Hotelling's method.[9] Let $\gamma_1, \gamma_2, \cdots, \gamma_s$ be the principal components. Then $x_i = a_{ij}\gamma_j$. The a_{ij}'s satisfy the following equations:

$$a_{jq}\sigma_{ij} = \lambda_q a_{iq}, \; [q] \tag{3.1}$$

$$a_{ip}a_{iq} = \lambda_q \delta_{pq} \tag{3.2}$$

---

[7] "Analysis of a Complex of Statistical Variables into Principal Components," *The Journal of Educational Psychology*, Sept. and Oct. 1933. See also M. A. Girshick, "Principal Components," *Journal of the American Statistical Association*, Vol. 31, Sept. 1936.

[8] The method here employed is parallel to the one used by Hotelling in his paper of 1936 in deriving asymptotic standard errors for canonical correlations.

[9] Loc. cit.

where the symbol δ_{pq} has the value zero for $p \neq q$ and 1 for $p = q$, λ_q is a root of the characteristic equation

$$|\sigma_{ij} - \lambda\delta_{ij}| = 0 \tag{3.3}$$

and σ_{ij} is the population covariance of x_i and x_j.

If we multiply (3.1) by a_{ip}, sum with respect to i and use (3.2), we get

$$a_{ip}a_{jq}\sigma_{ij} = \lambda_p^2\delta_{pq}. \tag{3.4}$$

When a root of (3.3) is simple and not equal to zero, the corresponding a_{ij}'s and the root itself are definite analytic functions of the σ_{ij}'s over a region without singularities. A set of sampling errors $d\sigma_{ij}$ in the covariances will then determine a corresponding set of sampling errors in the a_{ij}'s and in the root.

We assume then, that the roots $\lambda_1, \lambda_2, \cdots, \lambda_t$ of (3.3) we are considering are simple and non-vanishing. In terms of the derivatives of the analytic functions we define

$$da_{rk} = \frac{\partial a_{rk}}{\partial\sigma_{pq}}d\sigma_{pq}, \qquad d\lambda_r = \frac{\partial\lambda_r}{\partial\sigma_{pq}}d\sigma_{pq} \tag{3.5}$$

where $d\sigma_{pq} = s_{pq} - \sigma_{pq}$, s_{pq} being the corresponding sample covariance.

Differentiating equation (3.1) and employing the above formulae we get

$$\sigma_{ij}da_{jq} + a_{jq}d\sigma_{ij} = \lambda_q da_{iq} + a_{iq}d\lambda_q. \qquad [q] \tag{3.6}$$

We now multiply this equation by a_{ip}, sum with respect to i, and use equations (3.1) and (3.2). This yields:

$$\lambda_p a_{jp}da_{jq} + a_{ip}a_{jq}d\sigma_{ij} = \lambda_q a_{ip}da_{iq} + \lambda_q\delta_{pq}d\lambda_q. \qquad [p, q] \tag{3.7}$$

When $p = q$, the term $\lambda_p a_{jp}da_{jp}$ cancels out and equation (3.7) reduces to

$$\lambda_p d\lambda_p = a_{ip}a_{jp}d\sigma_{ij}. \qquad [p] \tag{3.8}$$

We change the subscripts p, i, j, to q, k, m, in (3.8) and multiply together the two equations thus obtained. This gives:

$$\lambda_p\lambda_q d\lambda_p d\lambda_q = a_{ip}a_{jp}a_{kq}a_{mq}d\sigma_{ij}d\sigma_{km}. \qquad [p, q] \tag{3.9}$$

Hence

$$\lambda_p\lambda_q E(d\lambda_p d\lambda_q) = a_{ip}a_{jp}a_{kq}a_{mq}E(d\sigma_{ij}d\sigma_{km}) \qquad [p, q] \tag{3.10}$$

where the symbol E denotes the mathematical expectation or mean value of the expression following.

Now it can be easily shown by means of the characteristic function of a multivariate normal distribution that

$$E(d\sigma_{ij}d\sigma_{km}) = \frac{1}{n}(\sigma_{ik}\sigma_{jm} + \sigma_{im}\sigma_{jk}) \tag{3.11}$$

where n is one less than the number in the sample. Substituting this expression in equation (3.10) and using (3.4) we get the following rather simple result

$$\lambda_p \lambda_q E(d\lambda_p d\lambda_q) = \frac{2\lambda_p^4 \delta_{pq}}{n}. \qquad [p, q] \tag{3.12}$$

Setting $p = q$ in this formula we get

$$E[(d\lambda_p)^2] = \frac{2\lambda_p^2}{n}. \tag{3.13}$$

But when $p \neq q$

$$E[d\lambda_p d\lambda_q] = 0. \tag{3.14}$$

Let $l_1, l_2, \cdots, l_t$, be the corresponding latent roots of a determinant of sample covariances. The sample latent root l_p may be expanded about λ_p in a Taylor series of the form

$$l_p = \lambda_p + \frac{\partial \lambda_p}{\partial \sigma_{rl}} d\sigma_{rl} + \frac{1}{2} \frac{\partial^2 \lambda_p}{\partial \sigma_{rl} \partial \sigma_{uv}} d\sigma_{rl} d\sigma_{uv} + \cdots \tag{3.15}$$

or, by (3.5)

$$l_p - \lambda_p = d\lambda_p + \cdots. \tag{3.16}$$

Squaring both sides of (3.16), taking the expected value, and using (3.13) we find *that the sample variance of a latent root l_p, apart from terms of higher order in n^{-1}, is given by $\frac{2\lambda_p^2}{n}$.*

If in (3.11) we set $i = j = k = m$, we get the variance of a sample variance, and it is interesting to note that its form is identical with the first term of the asymptotic expansion of the variance of a sample latent root.

The sample covariance of any two distinct roots is by (3.14) zero for the first term of the asymptotic expansion. That is, the covariance is at least of order n^{-2}. All the above results also follow from the fact, shown by the author in a previous paper,[10] that the coefficients of the principal components and hence the latent roots are maximum likelihood statistics. This property of the latent roots permits us also to state the following

THEOREM: *Let $\lambda_1, \lambda_2, \cdots, \lambda_t$ be any set of simple non-vanishing roots of* (3.3). *For sufficiently large samples these will be approximated by certain of the latent roots $l_1, l_2, \cdots, l_t$ of the samples. If $l_i - \lambda_i$ is divided by the standard error*

$$\sigma_{l_i} = \lambda_i \sqrt{\frac{2}{n}}$$

the resulting variates have a distribution which, as n increases, approaches the normal distribution of t independent variates of zero mean and unit standard deviation.

---

[10] Loc. cit.

COROLLARY: Let λ_1 be a maximum simple, non-vanishing root of (3.3) and let l_1 be the corresponding maximum sample root. Then, $l_1 - \lambda_1$ divided by its standard error has a distribution approaching normality in the limit.

2. The Variance of Log l. The formula for the standard error of the latent root given above contains a population parameter λ the numerical value of which we usually do not know. It is therefore important to find a transformation of the latent root to a new variate which will have or its leading term of the asymptotic standard error a quantity independent of the population parameter.

Let $k = f(l)$ be such a transformation. Then $K = f(\lambda)$ is the corresponding transformation for the population root.

We now expand k in a Taylor series about $l = \lambda$.

$$dk = f'(\lambda)dl + \tfrac{1}{2}f''(\lambda)(dl)^2 + \cdots \tag{3.17}$$

and get an approximation

$$dk = f'(\lambda)dl. \tag{3.18}$$

Squaring both sides and taking the expectation, we get

$$E(dk)^2 = [f'(\lambda)]^2 E[(dl)^2] = [f'(\lambda)]^2 \frac{2\lambda^2}{n}. \tag{3.19}$$

Now set $E(dk)^2 = 2/n$. Then, from (3.19)

$$f'(\lambda) = 1/\lambda$$

or

$$f(\lambda) = \log \lambda \tag{3.20}$$

Hence, if we set $k = \log l$, then

$$\sigma_k^2 = 2/n \tag{3.21}$$

is an approximation to the variance of k and is independent of any population parameter.

3. The Asymptotic Variances and Covariances of Roots of Determinants of Correlations. While the formulas for the asymptotic standard errors of the latent roots of a determinant of covariances are rather simple, this is not the case with the roots of a determinant of correlations. In deriving the asymptotic standard errors of simple non-vanishing roots of a determinant of correlations, we again assume that the variates $x_1, x_2, \cdots, x_s$, which in this case are of unit variance in the population, have been resolved into principal components. The equations of the previous section, up to and including (3.10), remain the same except that we substitute ρ_{ij} for every σ_{ij}, where ρ_{ij} is the population correlation of x_i with x_j. Thus equation (3.10) becomes

$$\lambda_p \lambda_q E(d\lambda_p d\lambda_q) = a_{ip}a_{jp}a_{kq}a_{mq}E(d\rho_{ij}d\rho_{km}), \qquad [p, q] \tag{3.22}$$

where $d\rho_{ij} = r_{ij} - \rho_{ij}$, r_{ij} being the sample correlation between x_i and x_j. The expected value of $d\rho_{ij}d\rho_{km}$ is not, as in the case of the σ's given in the simple form of (3.11) but rather it is given asymptotically, the leading term in n^{-1} being the following lengthy expression:

$$\begin{aligned} nE(d\rho_{ij}d\rho_{km}) &= \rho_{ik}\rho_{mj} + \rho_{kj}\rho_{mi} - \rho_{ij}\rho_{ki}\rho_{mi} - \rho_{ij}\rho_{kj}\rho_{mj} \\ &\quad - \rho_{km}\rho_{ki}\rho_{kj} + \tfrac{1}{2}\rho_{ij}\rho_{km}\rho_{ki}^2 + \tfrac{1}{2}\rho_{ij}\rho_{km}\rho_{kj}^2 \\ &\quad - \rho_{km}\rho_{mi}\rho_{mj} + \tfrac{1}{2}\rho_{ij}\rho_{km}\rho_{mi}^2 + \tfrac{1}{2}\rho_{ij}\rho_{km}\rho_{mj}^2 . \qquad [i, j, k, m] \end{aligned} \tag{3.23}$$

Substituting this in (3.22) and simplifying by means of equations (3.1) and (3.4) we finally get

$$\begin{aligned} n\lambda_p\lambda_q E(d\lambda_p d\lambda_q) &= 2(\lambda_p^4\delta_{pq} + \lambda_p\lambda_q a_{ip}^2 a_{jq}^2 \rho_{ij}^2) \\ &\quad - 2(\lambda_p\lambda_q^2 a_{ip}^2 a_{iq}^2 + \lambda_p^2\lambda_q a_{jp}^2 a_{jq}^2). \qquad [p, q] \end{aligned} \tag{3.24}$$

When $p = q$, (3.24) becomes

$$E[(d\lambda_p)^2] = \frac{2}{n}\left[\lambda_p^2 + a_{ip}^2 a_{jp}^2 \rho_{ij}^2 - 2\lambda_p \sum_{i=1}^{s} a_{ip}^4\right]. \qquad [p] \tag{3.25}$$

When $p \neq q$,

$$E(d\lambda_p d\lambda_q) = \frac{2}{n}[a_{ip}^2 a_{jq}^2 \rho_{ij}^2 - (\lambda_p + \lambda_q)a_{ip}^2 a_{iq}^2]. \qquad [p, q] \tag{3.26}$$

Hence (3.25) is the leading term of the asymptotic expansion of the variance of λ_p, and (3.26) is the leading term of the asymptotic expansion of the covariance of λ_p and λ_q, where λ_p and λ_q are simple, non-vanishing roots of a determinant of correlations.

4. Asymptotic Variances and Covariances of Coefficients of Principal Components Derived from a Determinant of Covariances. Let $x_i = a_{ij}\gamma_j$ be the equation of transformation of the variates $x_1, x_2, \cdots, x_s$ into principal components. In what follows we assume that the latent roots of the determinantal equation (3.3) are simple and none equal to zero. The last restriction makes the determinant of covariances non-vanishing. The determinant of the a_{ij}'s will therefore be also different from zero. With these assumptions in mind, we now proceed to derive the asymptotic variances and covariances of the a_{ij}'s.

We set $p = q$ in (3.2) and differentiate the result. This yields:

$$d\lambda_p = 2a_{lp}da_{lp}, \qquad [p] \tag{3.27}$$

where the summation index i was replaced by l. Substituting for $d\lambda_p$ from (3.8) we get:

$$a_{ip}a_{jp}d\sigma_{ij} = 2\lambda_p a_{lp}da_{lp}. \qquad [p] \tag{3.28}$$

Now, when $p \neq q$, equation (3.7) reduces to

$$\lambda_p a_{jp}da_{jq} + a_{ip}a_{jq}d\sigma_{ij} = \lambda_q a_{ip}da_{iq}, \qquad [p, q]$$

or

$$a_{ip}a_{jq}d\sigma_{ij} = (\lambda_q - \lambda_p)a_{lp}da_{lq}. \qquad [p, q] \tag{3.29}$$

We combine equations (3.28) and (3.29) into one equation

$$a_{ip}a_{jq}d\sigma_{ij} = (\lambda_q + \epsilon_{pq}\lambda_p)a_{lp}da_{lq}, \qquad [p, q] \tag{3.30}$$

where ϵ_{pq} has the value 1 when $p = q$ and -1 when $p \neq q$. The reciprocal of $\lambda_q + \epsilon_{pq}\lambda_p$, (which is different from zero), we denote by b_{qp}. Then equation (3.30) can be written as

$$a_{ip}b_{qp}a_{jq}d\sigma_{ij} = a_{lp}da_{lq}. \qquad [p, q] \tag{3.31}$$

Since the determinant $| a_{ij} |$ of the a_{ij}'s is different from zero, we can solve this set of homogeneous linear equations for da_{lq}'s, $(l = 1, 2, \cdots, s)$. To do this we multiply equation (3.31) by A^{tp}, where A^{tp} is the element of the t^{th} row and p^{th} column of the inverse of the determinant $| a_{ij} |$, and sum with respect to p. Since $A^{tp}a_{lp} = \delta_{tl}$ we get,

$$A^{tp}a_{ip}b_{qp}a_{jq}d\sigma_{ij} = \delta_{tl}da_{lp} = da_{tq}. \qquad [q] \tag{3.32}$$

We now change the subscripts i, j, t, p, q, in (3.32) to k, m, r, u, v, respectively, multiply the two equations thus obtained, and take the expected value:

$$E(da_{tq}da_{rv}) = A^{tp}A^{ru}a_{ip}a_{ku}b_{qp}b_{vu}a_{jq}a_{mv}E(d\sigma_{ij}d\sigma_{km}). \qquad [q, v] \tag{3.33}$$

Substituting for $Ed\sigma_{ij}d\sigma_{km}$ its values from (3.11) and simplifying by means of (3.4) we get:

$$E(da_{tq}da_{rv}) = \frac{\lambda_v^2\lambda_q^2}{n} A^{tv}A^{rq}b_{vq}b_{qv} + \frac{\lambda_q^2\delta_{qv}}{n}\sum_{u=1}^{s} A^{tu}A^{ru}b_{qu}b_{vu}\lambda_u^2 \tag{3.34}$$

where we sum *only with respect to u*. We may simplify this formula to some extent by employing the relation: $A^{tq} = a_{tq}/\lambda_q$. (This relation is obtained from (3.2) by multiplying each side of that equation by A^{tp} and summing with respect to p). When this is done and the values for the b's are substituted, the final result becomes:

$$\begin{aligned} E(da_{tq}da_{rv}) = {} & \frac{\lambda_v\lambda_q a_{tv}a_{rq}}{n(\lambda_q + \epsilon_{qv}\lambda_v)(\lambda_v + \epsilon_{qv}\lambda_q)} \\ & + \frac{\lambda_q^2\delta_{qv}}{n}\sum_{u=1}^{s}\frac{a_{tu}a_{ru}}{(\lambda_q + \epsilon_{qu}\lambda_u)(\lambda_v + \epsilon_{vu}\lambda_u)}. \end{aligned} \tag{3.35}$$

From this we derive the following specific formulas:

$$\begin{aligned} E(da_{tq}da_{rq}) = {} & \frac{a_{tq}a_{rq}}{4n} \\ & + \frac{\lambda_q^2}{n}\left[\frac{a_{t1}a_{r1}}{(\lambda_q - \lambda_1)^2} + \cdots + \frac{a_{tq}a_{rq}}{4\lambda_q^2} + \cdots + \frac{a_{ts}a_{rs}}{(\lambda_q - \lambda_s)^2}\right] \end{aligned} \tag{3.36}$$

$$(3.37) \qquad E[(da_{tq})^2] = \frac{a_{tq}^2}{4n} + \frac{\lambda_q^2}{n}\left[\frac{a_{t1}^2}{(\lambda_q - \lambda_1)^2} + \cdots + \frac{a_{tq}^2}{4\lambda_q^2} + \cdots + \frac{a_{ts}^2}{(\lambda_q - \lambda_s)^2}\right]$$

$$(3.38) \qquad E(da_{tq}\,da_{rv}) = -\frac{\lambda_q \lambda_v a_{tv} a_{rq}}{n(\lambda_q - \lambda_v)^2}. \qquad (q \neq v)$$

Formulas (3.36), (3.37) and (3.38) give us the leading terms of the asymptotic expansions of the variances and covariances for the principal components. It should be remarked that the coefficients of "mutual regression" equations can be easily shown to be proportional to those of the principal components. Hence their asymptotic standard errors and covariances may be derived in a similar manner and will be of the same form.

5. **Variances and Covariances of Latent Roots when the Population Roots are Equal.** Let $k_1, k_2, \cdots, k_p$ be the latent roots of a generalized sample variance of p normally distributed variates.

Ordinarily the subscripts of the roots designate their ranks, so that $k_1 \geq k_2 \geq \cdots \geq k_p$. We may, however, assign to a root a subscript from 1 to p without any regard to its size.[11] If this is done randomly for every sample of n observations the mathematical expectation of $k_i^r k_j^s k_k^t \cdots$ will be the same for every permutation of the subscripts $i, j, k, \cdots$. This fact permits us to calculate the variances and covariances of the above roots.

We may assume, without any loss of generality, that the p variates are independently distributed,[12] and furthermore we assume the population roots to be all equal to unity. Then equation (3.11) becomes

$$(3.39) \qquad E(s_{ij}s_{km}) = \delta_{ij}\delta_{km} + \frac{1}{n}(\delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}).$$

Where s_{pq} is the sample variance of x_p and x_q and δ_{pq} is the Kronecker delta.

Now it can be easily shown that

$$(3.40) \quad \sum_1^p s_{ii} = \sum_1^p k, \quad \sum_{i<j} (s_{ii}s_{jj} - s_{ij}^2) = \sum_{i<j} k_i k_j, \quad \sum_1^p s_{ii}^2 + 2\sum_{i<j} s_{ij}^2 = \sum_1^p k^2.$$

Hence $E(k) = 1$, and

$$E(\sum k^2) = E(\sum s_{ii}^2 + 2\sum_{i<j} s_{ij}^2)$$

or

$$(3.41) \qquad pEk^2 = pEs_{ii}^2 + p(p-1)Es_{ij}^2. \qquad (i \neq j)$$

Substituting from (3.39) in (3.41) we get

$$E(k^2) = 1 + \frac{p+1}{n}.$$

[11] This approach was suggested to the author by Professor Hotelling.

[12] See Part II, last Paragraph.

The variance of k is therefore given exactly by

$$\sigma_k^2 = E(k^2) - 1 = \frac{p + 1}{n}. \tag{3.42}$$

In a similar manner we find the covariances of k_i and k_j to be

$$\sigma_{k_i k_j} = -\frac{1}{n} \tag{3.43}$$

IV. DISTRIBUTION AND MOMENTS OF QUANTITIES RELATED TO q AND z

From the known distribution of q and z and their expressions in terms of the ratio of determinants given by (1.1) and (1.12), we can derive moments and distributions of several related functions of sample variances and correlations of two independent sets of variates.

$$\text{Let } p = \frac{q^2}{z} = \frac{|b_{ij}|}{|c_{ij}|} \quad \text{by (1.12).} \tag{4.1}$$

Since the two determinants in (4.1) are independently distributed, the sampling distribution of p, given in the above form, can be obtained for a general value of s and t from Wilks'[13] distribution of the ratio of independent generalized variances.

Thus, for $s = 2$ and $t \geq 2$, the distribution of p is given by

$$\frac{\Gamma(n-2)}{2\Gamma(t-1)\Gamma(n-t-1)} p^{\frac{1}{2}(t-3)} \frac{dp}{(1+\sqrt{p})^{n-2}}. \tag{4.2}$$

When the number of variates in each set is the same, the numerator of q^2 in (1.1) becomes the square of the determinant of covariances *between* the two sets of variates. Thus

$$q^2 = \frac{|a_{i\alpha}|^2}{|a_{ij}|\,|a_{\alpha\beta}|} \tag{4.3}$$

where i, j, take on values from 1 to s, α, β take on values from $s + 1$ to $2s$, and $a_{uv} = \sum_1^n x_u x_v$.

If the two sets are independent, the quantities q^2, $|a_{ij}|$, $|a_{\alpha\beta}|$, are independently distributed. Hence

$$E(|a_{i\alpha}|^m) = Eq^m(|a_{ij}|^{\frac{1}{2}m})\,E(|a_{\alpha\beta}|^{\frac{1}{2}m}). \tag{4.4}$$

Setting $\beta = 0$ in (1.16) and employing formula (1.15) we get for the moment of $|a_{i\alpha}|$

$$E(|a_{i\alpha}|^m) = \frac{2^{sm}}{|A_{ij}|^{\frac{1}{2}m}|A_{\alpha\beta}|^{\frac{1}{2}m}} \prod_{i=1}^{s} \left[\frac{\Gamma\left(\frac{s+m-i+1}{2}\right)\Gamma\left(\frac{n+m-i+1}{2}\right)}{\Gamma\left(\frac{s-i+1}{2}\right)\Gamma\left(\frac{n-i+1}{2}\right)}\right] \tag{4.5}$$

---

[13] Loc. cit., pp. 478–479.

where A_{uv} denotes the cofactor corresponding to σ_{uv} divided by the determinant $|\sigma_{uv}|$, σ_{uv} being the population covariance of x_u and x_v.

We may replace the product sums in (4.3) by sample correlations and, with the assumption that all the variates come from independent populations, obtain the m^{th} moment of the determinant of correlations between the two sets as

$$(4.6)\qquad E(|r_{i\alpha}|^m) = \frac{\Gamma^{2s}\left(\frac{n}{2}\right)}{\Gamma^{2s}\left(\frac{n+m}{2}\right)} \prod_{i=1}^{s} \left[\frac{\Gamma\left(\frac{n+m-i+1}{2}\right)\Gamma\left(\frac{s+m-i+1}{2}\right)}{\Gamma\left(\frac{n-i+1}{2}\right)\Gamma\left(\frac{s-i+1}{2}\right)}\right].$$

This follows from the expression for the m^{th} moment of q and the formula

$$(4.7)\qquad E(|r_{uv}|^k) = \frac{\Gamma^{s}\left(\frac{n}{2}\right)}{\Gamma^{s}\left(\frac{n+2k}{2}\right)} \prod_{i=1}^{s} \left[\frac{\Gamma\left(\frac{n+2k-i+1}{2}\right)}{\Gamma\left(\frac{n-i+1}{2}\right)}\right]$$

derived by Wilks.[14]

If we set $s = t = 2$, the numerator of q^2 in (4.3) becomes the square of a determinant of sample covariances (or correlations) known to psychologists as the tetrad. We shall here derive its distribution under the assumptions that the four variates are independently distributed.

We write

$$(4.8)\qquad q = \frac{T}{u_1 u_2}$$

where

$$(4.9)\qquad T = r_{13}r_{24} - r_{14}r_{23}, \qquad u_1 = (1 - r_{12}^2)^{\frac{1}{2}}, \qquad u_2 = (1 - r_{34}^2)^{\frac{1}{2}}$$

and q is taken as positive.

Now the distribution of q for $s = t = 2$ is given by

$$(4.10)\qquad (n-2)(1-q)^{n-3}dq$$

and the distribution of u is known to be

$$(4.11)\qquad \frac{2\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)} u^{n-2}(1-u^2)^{-\frac{1}{2}}\,du.$$

Hence the distribution of u_1, u_2 and q is given by

$$(4.12)\qquad \frac{4(n-2)}{\pi} \frac{\Gamma^2\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)} (1-q)^{n-3}(u_1u_2)^{n-2}[(1-u_1^2)(1-u_2^2)]^{-\frac{1}{2}}\,du_1\,du_2\,dq.$$

---

[14] Loc. cit., p. 492.

Performing the transformation (4.8) and integrating out u_1 and u_2 we get for the distribution of the tetrad

$$\frac{4(n-2)\Gamma^2\left(\frac{n}{2}\right)}{\pi\Gamma^2\left(\frac{n-1}{2}\right)} \int_T^1 \int_{\frac{T}{u_1}}^1 \frac{(u_1u_2 - T)^{n-3}}{\sqrt{(1-u_1^2)(1-u_2^2)}}\, du_1,\, du_2. \tag{4.13}$$

All the moments of T can of course be obtained by setting $s = 2$ in (4.6).[15]

U. S. DEPARTMENT OF AGRICULTURE,
WASHINGTON, D. C.

[15] The limiting distribution of the tetrad was given by J. L. Doob in an article entitled "The Limiting Distributions of Certain Statistics," *Annals of Mathematical Statistics*, Vol. 6, (1935). For a more general distribution of the tetrad and other statistics considered in this paper see W. G. Madow, "Contributions to the Theory of Multivariate Statistical Analysis," *Transactions of the American Mathematical Society*, Nov. 1938.