A NEW APPROACH TO SAMPLING DISTRIBUTIONS OF THE MULTIVARIATE NORMAL THEORY—PART II

BY R. D. NARAIN

Indian Council of Agricultural Research, New Delhi

1. THE MULTIPLE CORRELATION COEFFICIENT DISTRIBUTION

IN Part I, to derive the distribution of the multiple correlation coefficient ω, we expressed

$$\frac{\omega^2}{1 - \omega^2}$$

as the ratio of two sums of squares. The alternative derivation which we give now makes use of the fact that

$$1 - \omega^2$$

can be expressed as a product. In fact we have

$$\frac{1}{1 - \omega^2} = \prod_{r=1}^{p-1} \left(1 + \frac{b^2_{rp \cdot (012 \cdots r-1)} \sum_{i=1}^{N} x_{ip}{}^2{}_{\cdot (012 \cdots r-1)}}{\sum_{i=1}^{N} x_{ir}{}^2{}_{\cdot (p012 \cdots r-1)}} \right) \qquad (1.0)$$

where x_p is the dependent variate, $x_1, x_2, \ldots x_{p-1}$ are the independent variates, and x_0, as before stands for unity. Since the coefficient of multiple correlation is invariant for any linear transformations of the independent variates, the distribution problem may be simplified at the very outset by assuming all the correlations between $x_1, x_2, \ldots x_p$ to vanish except that between x_p and x_1 which would be equivalent to the multiple correlation coefficient Ω in the population. Without loss of generality we may further assume all the variances unity. The probability function of the primary variates can then be written as

$$\text{Const.} \times e^{-\frac{1}{2} \sum_{i=1}^{N} \sum_{i=1}^{p} \xi_{ir}^2} \prod_{i,r} dx_{ir} \qquad (1.1)$$

where

$$\xi_{i1} = \frac{(x_{i1} - m_1) - \Omega\,(x_{ip} - m_p)}{\sqrt{1 - \Omega^2}} \qquad (1.2)$$

and

$$\xi_{ir} = x_{ir} - m_r \qquad r = 2, 3, \ldots p \qquad (1.3)$$

As before we shall express each x_r as sum of its orthogonal constituents, but in a different order; we have

$$x_{ir} = x_{ir \cdot (p\,012 \cdots r-1)} + b_{rp \cdot (012 \cdots r-1)}\, x_{p \cdot (012 \cdots r-1)}$$
$$+ \sum_{s=0}^{r-1} b_{rs \cdot (012 \cdots s-1)}\, x_{is \cdot (012 \cdots s-1)} \qquad (1.4)$$

and

$$x_{ip} = x_{ip\cdot 0} + b_{p0} \tag{1.5}$$

Arguing exactly on the same lines as those for obtaining the distribution (2.1) of the earlier paper (1948) the distribution of

$$\Sigma\, x^2_{ir\cdot(p012\cdots r-1)} \{= \chi_r^2\}$$

$$b^2_{rp\cdot(012\cdots r-1)} \sum_{i=1}^{N} x^2_{ip\cdot 012\cdots r-1} \quad \{= t_r^2\} \tag{1.6}$$

and

$$\sum_{i=1}^{N} x^2_{ip\cdot 0} \qquad \{= h^2\}$$

may be written as

$$\prod_{r=2}^{p-1}\left[\frac{1}{2^{\frac{N-r-1}{2}}\,\Gamma\left(\frac{N-r-1}{2}\right)\sqrt{2\pi}}\, e^{-\frac{1}{2}\{\chi_r^2+t_r^2\}}\,(\chi_r^2)^{\frac{N-r-3}{2}}\, d(\chi_r^2)\, dt_r\right]$$

$$\times \frac{1}{\sqrt{2\pi}\, 2^{\frac{N-2}{2}}\,\Gamma\left(\frac{N-2}{2}\right)(1-\Omega^2)^{\frac{N-1}{2}}}\, e^{-\frac{1}{2(1-\Omega^2)}\{\chi_1^2+(t_1-\Omega h)^2\}}\,.$$

$$\times (\chi^2)^{\frac{N-4}{2}}\, d(\chi_1^2)\, dt_1 \times f(h^2)\; d(h^2) \tag{1.7}$$

where

$$f(h^2) = \frac{1}{2^{\frac{N-1}{2}}\,\Gamma\frac{N-1}{2}}\, e^{-\frac{1}{2}h^2}\,(h^2)^{\frac{N-3}{2}} \tag{1.8}$$

Setting

$$t_r = l_r \sin\theta_r$$

$$\chi_r = l_r \cos\theta_r$$

and integrating out l_r we obtain

$$\left(\frac{1}{\sqrt{2\pi}}\right)^{p-1} \frac{1}{\Gamma\left(\frac{N-p}{2}\right)\{2(1-\Omega^2)\}^{\frac{N-1}{2}}}$$

$$\times \int_0^\infty e^{-\frac{1}{2(1-\Omega^2)}\{l_1^2+\Omega^2 h^2-2l_1\,\Omega\, h\sin\theta_1\}}\,(l_1^2)^{\frac{N-3}{2}}\; d(l_1^2)$$

$$\times \prod_{r=1}^{p-1}\{(\cos\theta_r)^{N-r-2}\; d\theta_r\}\, f(h^2)\; d(h^2) \tag{1.9}$$

The range of variation of each θ_r is from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$. Restricting the variation to range from zero to $\frac{\pi}{2}$, the distribution reduces to

$$\left(\frac{2}{\sqrt{\pi}}\right)^{p-1} \frac{1}{\Gamma\left(\frac{N-p}{2}\right)\{2(1-\Omega^2)\}^{\frac{N-1}{2}}}$$

$$\times \int_0^\infty e^{-\frac{1}{2(1-\Omega^2)}\{l_1^2+\Omega^2 h^2\}} (l_1^2)^{\frac{N-3}{2}} \cosh\left(\frac{\Omega}{1-\Omega^2} l_1 h \sin\theta_1\right) d(l_1^2)$$

$$\times \prod_{r=1}^{p-1} \{(\cos\theta_r)^{N-r-2} d\theta_r\} \times f(h^2)\, d(h^2) \tag{2.0}$$

Expanding the hyperbolic cosine and integrating for l_1 term by term, we obtain

$$\frac{1}{\Gamma\left(\frac{N-p}{2}\right)} \left(\frac{2}{\sqrt{\pi}}\right)^{p-1} e^{-\frac{1}{2}\frac{\Omega^2 h^2}{1-\Omega^2}} \sum_{m=0}^{\infty} \left[\frac{2^m \Gamma\left(\frac{N-1}{2}+m\right)}{(2m)!} \left(\frac{\Omega^2 h^2}{1-\Omega^2}\sin^2\theta_1\right)^m\right]$$

$$\times \prod_{r=1}^{p-1} \{(\cos\theta_r)^{N-r-2}\, d\theta_r\} \times f(h^2)\, d(h^2) \tag{2.1}$$

We note that

$$1-\omega^2 = \prod_{r=1}^{p-1} \cos^2\theta_r$$

We shall find, therefore, the distribution of

$$z_{p-1} = \cos\theta_1 \cos\theta_2 \ldots \cos\theta_{p-1}$$

Set generally

$$z_r = \cos\theta_1 \cos\theta_2 \ldots \cos\theta_r$$

The distribution (2.1) thus transforms to

$$\left(\frac{2}{\sqrt{\pi}}\right)^{p-1} \frac{1}{\Gamma\left(\frac{N-p}{2}\right)} e^{-\frac{1}{2}\frac{\Omega^2 h^2}{1-\Omega^2}}$$

$$\times \sum_0^\infty \left[\frac{(1-z_1^2)^{\frac{2m-1}{2}} z_1, z_2, \ldots z_{p-2} z_{p-1}^{N-p-1}}{\sqrt{(z_1^2-z_2^2),(z_2^2-z_3^2),\ldots(z_{p-2}^2-z_{p-1}^2)}}\right.$$

$$\left.\times \frac{2^m \Gamma\left(\frac{N-1}{2}+m\right)}{(2m)!} \left(\frac{\Omega^2 h^2}{1-\Omega^2}\right)^m\right] \times f(h^2)\, d(h^2) \tag{2.2}$$

To integrate out $z_1, z_2, \ldots z_{p-2}$ we put

$$\sin^2\phi_r = \frac{1-z_r^2}{1-z_{r+1}^2} \qquad r = 1, 2, \ldots p-2$$

So that the Jacobian

$$\frac{\partial(z_1, z_2, \ldots z_{p-2})}{\partial(\phi_1, \phi_2, \ldots \phi_{p-2})} = \prod_{r=1}^{p-2} \frac{(1-z_{r+1}^2)\sin\varphi_r \cos\phi_r}{z_r}$$

The distribution (2.2) therefore reduces to

$$\left(\frac{2}{\sqrt{\pi}}\right)^{p-1} \frac{1}{\Gamma\left(\frac{N-p}{2}\right)} e^{-\frac{1}{2}\frac{\Omega^2 h^2}{1-\Omega^2}}$$

$$\times \sum_0^\infty \left[\frac{2^m \Gamma\left(\frac{N-1}{2}+m\right)}{(2m)!}\left(\frac{\Omega^2 h^2}{1-\Omega^2}\right)^m \left\{\prod_{r=1}^{p-2} (\sin\phi_r)^{2m+r-1}\right\}(1-z^2_{p-1})^m\right]$$

$$\times z_{p-1}^{N-p-1}(1-z^2_{p-1})^{\frac{p-3}{2}} dz_{p-1} \prod_{r=1}^{p-2} d\phi_r \times f(h^2)\, d(h^2) \qquad (2.3)$$

Integrating out the ϕ's, we have

$$\frac{2}{\sqrt{\pi}} \frac{1}{\Gamma\left(\frac{N-p}{2}\right)} e^{-\frac{1}{2}\frac{\Omega^2 h^2}{1-\Omega^2}}$$

$$\times \sum_0^\infty \left[\frac{2^m}{(2m)!} \frac{\Gamma\left(\frac{N-1}{2}+m\right)\Gamma(m+\frac{1}{2})}{\Gamma\left(\frac{p-1}{2}+m\right)} \left(\frac{\Omega^2 h^2}{1-\Omega^2}\right)^m (1-z^2_{p-1})^m\right]$$

$$\times (1-z^2_{p-1})^{\frac{p-3}{2}} z_p^{N-p-1} dz_{p-1} f(h^2)\, d(h^2) \qquad (2.4)$$

Making use of the relation

$$\frac{1}{\sqrt{\pi}} \frac{2^m \Gamma(m+\frac{1}{2})}{(2m)!} = \frac{1}{2^m (m)!}$$

We have, since

$$\omega^2 = 1 - z^2_{p-1}$$

$$\frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N-p}{2}\right)\Gamma\left(\frac{p-1}{2}\right)} e^{-\frac{1}{2}\frac{\Omega^2 h^2}{1-\Omega^2}}\ {}_1F_1\left[\frac{N-1}{2}, \frac{p-1}{2}, \frac{\Omega^2 h^2 \omega^2}{2(1-\Omega^2)}\right]$$

$$\times (\omega^2)^{\frac{p-3}{2}} (1-\omega^2)^{\frac{N-p-2}{2}} d(\omega^2) f(h^2)\, d(h^2) \qquad (2.5)$$

This, it will be seen, is identical in form to the distribution (3.7) obtained in Part I (1948) and therefore, integrating out h^2, we obtain as before the general sampling distribution of ω^2 as

$$\frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N-p}{2}\right)\Gamma\left(\frac{p-1}{2}\right)} (1-\Omega^2)^{\frac{N-1}{2}} (1-\omega^2)^{\frac{N-p-2}{2}} (\omega^2)^{\frac{p-3}{2}}$$

$$\times F\left[\frac{N-1}{2}, \frac{N-1}{2}, \frac{p-1}{2}, \Omega^2 \omega^2\right] d(\omega^2) \qquad (2.6)$$

But the important thing to note is that just as we obtained the distribution of ω^2 for fixed values of the independent variates, we may from (2.5) obtain the distribution of ω^2 for fixed values of the *dependent* variate. Since

$$\frac{h^2}{N}$$

is the variance of x_p, which we have assumed to be unity, we obtain the distribution of ω^2 for fixed values of x_p as

$$\frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N-p}{2}\right)\Gamma\left(\frac{p-1}{2}\right)} e^{-\frac{1}{2}\frac{N\Omega^2}{1-\Omega^2}} (1-\omega^2)^{\frac{N-p-2}{2}} (\omega^2)^{\frac{p-3}{2}}$$

$$\times {}_1F_1\left[\frac{N-1}{2}, \frac{p-1}{2}, \frac{N\Omega^2\omega^2}{2(1-\Omega^2)}\right] d(\omega^2) \qquad (2.7)$$

which establishes the result that the distribution of ω^2 is of the same form irrespective of the fact whether the independent variates are considered fixed or the dependent variate is considered fixed. This duality in the distribution theory has been noted by Fisher (1938, 1940), Bartlett (1939) and others. For the case of no real association it can be visualized from purely geometrical considerations of symmetry, but as pointed out by Bartlett (1947) the duality is not quite obvious for the non-null case.

2. Distribution of the D^2-Statistic

Let samples of n_1 and n_2 observations be drawn from two $p-1$ variate normal populations having the same dispersion matrix. Denote these observations by

$$x_{ir} \qquad r = 1, 2, \ldots p-1$$
$$i = 1, 2, \ldots n_1, n_1+1, n_1+2, \ldots n_1+n_2$$

$i > n_1$, referring to observations from the second population. Let

$$E(x_{ir}) = m_r \qquad \text{for } i \leqslant n_1$$
$$\text{and} \quad = m_r' \qquad \text{for } i > n_1$$
$$\sum_{i=1}^{n_1} x_{ir} = n_1 a_r, \quad \sum_{n_1+1}^{n_1+n_2} x_{ir} = n_2 a_r'$$

and

$$(n_1+n_2)\, c_{rs} = \sum_{i=1}^{n_1} (x_{ir} - a_r)(x_{is} - a_s) + \sum_{i=n_1+1}^{n_1+n_2} (x_{ir} - a'_r)(x_{is} - a_s')$$
$$r, s = 1, 2, \ldots p-1$$

The D^2 statistic is defined by (Bose and Roy, 1938)

$$(p-1)\,D^2 = \sum_{r,\,s=1}^{p-1} c^{rs}\,(a_r - a_r')\,(a_s - a_s')$$

where $\| c^{rs} \|$ is the matrix reciprocal to $\| c_{rs} \|$.

The corresponding population parameter $\triangle^2$ is defined by

$$(p-1)\,\triangle^2 = \sum_{r,s=1}^{p-1} \alpha^{rs}\,(m_r - m_r')\,(m_s - m_s')$$

where $\| \alpha^{rs} \|$ is the matrix reciprocal to the common population dispersion matrix $\| \alpha_{rs} \|$

It is well known that

$$1 + \frac{(p-1)\,\bar{n}D^2}{2N} = \frac{\left| Nc_{rs} + \frac{\bar{n}}{2}\,(a_r - a_r')\,(a_s - a_s') \right|}{| Nc_{rs} |}$$

where

$$N = n_1 + n_2$$

and

$$\frac{2}{\bar{n}} = \frac{1}{n_1} + \frac{1}{n_2}$$

Introducing quantities x_0 for unity and x_p defined by

$$x_{ip} = \frac{n_2}{N} \text{ for } i = 1, 2, \ldots n_1$$

and

$$= \frac{-n_1}{N} \text{ for } i = n_1 + 1,\ n_1 + 2, \ldots n_1 + n_2$$

we have

$$1 + \frac{(p-1)\,\bar{n}D^2}{2N} = \frac{1}{1-\omega^2}$$

where ω is the coefficient of multiple correlation between the quantity x_p and the variates $x_1, x_2, \ldots x_{p-1}$ in the pooled sample of size N. The distribution of the D^2-statistic follows immediately from (2.7) as

$$\left\{\frac{(p-1)\,\bar{n}}{2N}\right\}^2 \frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N-p}{2}\right)\Gamma\left(\frac{p-1}{2}\right)}\, e^{-\frac{\bar{n}}{4}(p-1)\triangle^2}$$

$$\times\,(D^2)^{\frac{p-3}{2}} \left\{1 + \frac{(p-1)\,\bar{n}D^2}{2N}\right\}^{-\frac{N-1}{2}}$$

$$\times\,{}_1F_1\left[\frac{N-1}{2}, \frac{p-1}{2}, \frac{(p-1)^2\,\bar{n}^2\,\triangle^2\,D^2}{8N + 4\,(p-1)\,\bar{n}D^2}\right] d\,(D^2) \qquad (2.8)$$

The D^2-test for the equality of the means of two multivariate normal populations is a particular case of the general T^2-test as stated by Hsu (1938, 1941).

Let a sample of N observations of $p+q$ random variates be distributed as

$$(2\pi)^{-\frac{N(p+q)}{2}} \,|\, \alpha^{rs} \,|^{\frac{N}{2}}\, e^{-\frac{1}{2}\{\sum_{r,s} \alpha^{rs} \sum_{i} (x_{ir} - \sum_{h} \beta_{rh}\, x_{ih})\,(x_{is} - \sum_{h} \beta_{sh}\, x_{ih})\}} \times \prod_{i,r} dx_{ir} \qquad (2.9)$$

$$r, s = m+1, m+2, \ldots m+p+q$$
$$h = 1, 2, \ldots m$$
$$i = 1, 2, \ldots N$$
$$N > m+p+q$$

The likelihood ratio statistic appropriate to testing the hypothesis

$$\beta_{rm} = 0 \qquad r = m+1, m+2, \ldots m+p+q \qquad (3.0)$$

is given by

$$1 + T^2 = \frac{|\, w_{rs} + d_r\, d_s \,|}{|\, w_{rs} \,|} = \sum_{r,s=m+1}^{m+p+q} w^{rs}\, d_r\, d_s \qquad (3.1)$$

where $\| w^{rs} \|$ is the matrix reciprocal to $\| w_{rs} \|$ and

$$w_{rs} = \sum_i x_{ir\cdot(12\cdots m)}\; x_{is\cdot(12\cdots m)}$$

and

$$d_r = \frac{\sum_i x_{ir\cdot(12\cdots m-1)}\; x_{im\cdot(12\cdots m-1)}}{\sqrt{\sum x^2_{im\cdot(12\cdots m-1)}}}$$

The general non-null distribution of the above T^2-statistic may be obtained exactly as the disrtibution (2.8). If the hypothesis be true

$$T^2\,(N-m-p-q+1)/p+q$$

will be distributed as a variance ratio with $p+q$ and $N-m-p-q+1$ degrees of freedom. If $m=1$, and $x_{i1} \equiv 1$ the test reduces to the usual T^2-test for the significance of the means of a single sample form a multivariate normal population. The D^2-test relevant to two samples corresponds to

$$N = n_1 + n_2$$

$$x_{i2} = \frac{n_2}{N} \text{ for } i = 1, 2, \ldots n_1$$

and

$$x_{i2} = \frac{-n_1}{N} \text{ for } i > n_1$$

where n_1, n_2 are the two sample sizes.

3. Certain Generalizations Associated with the T^2-Test

Instead of (3.0), more generally the hypothesis under test may specify $l \leqslant p+q$ linear relationships between the parameters β_{rm}. For simplicity let us take $m = 1$ only, so that the hypothesis H specifies the following l relationships

$$\sum_{r=2}^{p+q} \lambda_{jr} \beta_{r1} = 0, \qquad j = 2, 3, \ldots l+1 \tag{3.2}$$

To test this we may construct l new variates

$$y_j = \sum_{r=2}^{p+q} a_{jr} x_r \qquad j = 2, 3, \ldots l+1$$

where the coefficients a_{jr} may be chosen arbitrarily subject only to the condition that the $2l$ rowed matrix of the coefficients λ_{jr}, a_{jr} is of rank l.

Calculating

$$d_j = \frac{\sum_{i=1}^{N} y_{ij} x_{i1}}{\sqrt{\sum_{i=1}^{N} x_{i1}^2}}$$

and

$$w_{jk} = \sum_{i=1}^{N} y_{ij} y_{ik} - \frac{d_j d_k}{\sum_{i=1}^{N} x_{i1}^2}$$

the statistic

$$\left(\sum_{j,k=2}^{l+1} d_j d_k w^{jk}\right) (N-l)/l$$

can be used as the variance ratio with l and $N-l$ degrees of freedom. Testing equality of means of $p+q$ correlated variables on the basis of a sample of size N, and testing, if $q = p$, whether

$$E(x_r) = E(x_{r+p}), \qquad r = 1, 2, \ldots p$$

considered by Rao (1948), what he calls generalization of Student's test in two directions, both fall as special cases of the above hypothesis. The use of Hotellings T^2 in testing the equality of means of correlated variables was first shown by Hsu (1938). Another particular case of the above hypothesis will be to test whether the differences in respective mean values of two multivariate samples satisfy a given set of linear relationships. The case of $m > 1$, the hypothesis to be tested being

$$\Sigma \lambda_{jr} \beta_{rm} = 0 \qquad j = m+1, m+2, \ldots m+l \tag{3.3}$$

is distributionally identical. Only in all the calculations of sums of products and squares relevant to the test statistic, the variates x_1, x_2, $\ldots x_{m-1}$ are to be eliminated as concomitant variates and the variance ratio is to be used with l and $N-l-m+1$ degrees of freedom.

Another slightly different class of problems for which the T^2-test is available fall under the following category. The probability function (2.9) may be written as the conditional probability function of x_{iu} $(u = m + p + 1, m + p + 2, \ldots m + p + q)$ for given values of x_{it} $(t = m + 1, m + 2, \ldots m + p)$, multiplied by the probability function for x_{it}, *viz.*,

$$(2\pi)^{-\frac{Nq}{2}} \mid \gamma^{uv} \mid^{\frac{N}{2}} \times \text{Exp}\,[-\tfrac{1}{2}\{\sum_{u,v} \gamma^{uv} \sum_{i} (x_{iu} - \sum_{h} \lambda_{uh}\, x_{ih} - \sum_{t} \lambda_{ut}\, x_{it})\,(x_{iv} - \sum_{h} \lambda_{vh}\, x_{ih} - \sum_{t} \lambda_{vt}\, x_{it})\}]$$
$$\times \prod_{i,u} x_{iu} \times dF\,(x_{im+1}, x_{im+2}, \ldots x_{im+p}) \qquad (3.4)$$

$$\begin{aligned} i &= 1, 2, \ldots N \\ u, v &= m+ p+ 1, m+ p+ 2, \ldots m+ p+ q. \\ t &= m + 1, m + 2, \ldots m + p \\ h &= 1, 2, \ldots m \end{aligned}$$

The hypothesis to be tested is

$$\lambda_{um} = 0, \qquad u = m + p + 1, m + p+ 2, \ldots m + p + q$$

[More generally, the hypothesis may specify a set of linear relationships but which as we have seen above is reducible to the above form by constructing a new set of variables, and hence is mathematically equivalent.]

On the likelihood ratio principle the statistic relevant to this test is given by

$$1 + U = \frac{|\, a_{uv\cdot(12\cdots m-1,\, m+1,\, \cdots m+p)} \,|}{|\, a_{uv\cdot(12\cdots m-1,\, m, m+1,\, \cdots m+p)} \,|} \qquad (3.5)$$

where

$$a_{uv\cdot(12\cdots r)} \equiv \sum_{i} x_{iu\cdot(12\cdots r)}\, x_{iv(12\cdots r)}.$$

If the hypothhesis be true, the factor $dF\,(x_{im+1} \ldots x_{im+p})$ is irrelevant to the distribution of U which will therefore follow a T^2 distribution with q and $N - p - q - m + 1$ degrees of freedom. The hypothesis covers the important case of testing for the compatability of any assigned discriminant function formula (Bartlett, 1939; Fisher, 1940). It is easily seen that

$$1 + U = \frac{1+ T^2_{p+q}}{1+ T_p^{\,2}} \qquad (3.6)$$

where T^2_{p+q} has been written for T^2 defined in (3.1) and $T_p^{\,2}$ is its expression for $q = 0$. The use of this statistic has been discussed in general by Rao (1948). He however imposes the restriction that the set $d_{m+1}, d_{m+2}, \ldots d_{m+p}$ should be distributed independently of the set $d_{m+p+1}, d_{m+p+2}, \ldots d_{m+p+q}$. This is unnecessary, as is obvious from the above analysis. The test holds generally irrespective of any real correlations that may exist between the two sets.

Another point of interest in connection with this test is the fact that though under the hypothesis the statistic U is distributed as a T^2, its power function will in general be different from the power function of the usual T^2-test. In other words, the non-null distribution of U will not be similar to the distribution (2.8). This is because of the factor

$$dF\,(x_{im+1}, \ldots x_{im+p})$$

in (3.4) which becomes relevant when the hypothesis tested is not true. This distribution has been worked out and is being given in a separate paper.

Summary

Another exact analytical method of working out the general sampling distribution of the multiple correlation coefficient has been demonstrated, providing incidentally the sampling distribution of the multiple correlation coefficient for fixed values of the dependent variate.

Two general hypothesis have been stated which comprehend all uses of the T^2-test. The power functions of the test in the two cases is shown to be different.

References

Bartlett, M. S. .. "The standard errors of discriminant function coefficients," *Supp. J.R.S.S.*, 1939, **6**, 169.

——— .. "Multivariate analysis," *ibid.*, 1947, **9**, 176.

Bose, R. C., and Roy, S. N. "The distribution of the studentised *D*--Statistic," *Sankhya*, 1938, **4**, 19.

Fisher, R. A. .. "The statistical utilization of multiple measurements," *Ann. Eug. Lond.*, 1938, **8**, 376.

Fisher, R. A. .. "The precision of discriminant functions," *ibid.*, 1940, **10**, 422.

Hsu, P. L. .. "Notes on Hotelling's generalised T," *Ann. Math. Stats.*, 1938, **9**, 231.

——— .. "Canonical reduction of the general regression problem," *Ann. Eug. Lond.*, 1941, **11**, 42.

Narain, R. D. .. "A new approach to sampling distributions of the multivariate normal theory," *Jour. Ind. Soc. Agr. Stat.*, 1948, **1**, 59.

Rao, C. R. .. "Tests of Significance in Multivariate Analysis," *Biometrika*, 1948, **35**, 58.