A STUDY OF

THE DISTRIBUTIONS OF THE T^2 & WILKS' STATISTICS

BY

K:V.R. SASTRY

1950

2564
043:217

Thesis submitted for the Diploma in Statistics of the Indian Council of Agricultural Research, Ministry of Agriculture, Government of India, New Delhi.

Supervisor:

Dr. R.D. Narain

Contents

A STUDY OF THE DISTRIBUTIONS OF T^2 AND WILKS' 'STATISTIC'S

I. Introduction

Much work has been done by different authors in the case of the univariate analysis of variance and when the parent population is assumed to be normal, this study is almost complete. But in practice it appears necessary to measure the individual for a large number of correlated characters and thus the univariate theory needs to be appropriately generalised. The study of multivariate samples dates back to as early as the time of Pearson's investigations into the problem of racial likeness with the aid of a number of measurements on the skulls. Suppose there is a sample of skulls each of which is measured for certain p-characters which are supposed to be the discriminating characteristics of a race. It may be of interest to test the set of p sample means against a hypothetical set or the corresponding mean values calculated from another sample, which might have been taken at a different occasion, or different locality. The problem may be made more general by considering more than two such independent samples, the homogeneity of the mean values of which is the problem under question.

An instance of a different type of multivariate problem in economic studies may be posed somewhat like this:- Consider the demand for a particular commodity; the prices of alternate or complementary commodities, the population and income levels at the particular point of the time interval, besides the price of the commodity in question are responsible for variations that may be observed in the volume of demand. This leads to the study of the relation of one variate with a number of variates. One cannot overlook the presence of significant correlations within the set of these predicting or independent variates and thus the study of individual correlations of the single dependent variate with each of the independent variates separately does not seem to give an overall picture of the situation. The problem considered may be further generalised by considering a set of characters on the dependent side which the set of independent variates together effect and the variates within which by themselves may be correlated with each other.

When there is a single multivariate sample known to have emanated from a multivariate normal population with unknown covariance matrix, Hotelling's T^2 (1) is the appropriate statistic to test the set of sample means against a set of p-values. When there are two samples (p-variate) known to have emanated from two multivariate normal populations having identical, but unspecified covariance matrix, then Mahalanobis or Studentized D^2 statistics (2) is the appropriate statistic to test the hypothesis of equality of the two sets of means or the hypothesis of the identity of the populations from which the samples have emanated.

The quantity defined to measure the relationship of one variate with a set of auxiliary or independent variates is the multiple correlation coefficient (R^2) and an overall test of the relation will thus be a test of significance of the sample multiple correlation coefficient.

When it comes to the problem of tests of significance, as in univariate case, one is confronted with the consideration of the two kinds of error (viz.)

(i) rejecting and hypothesis when it is true,

and (ii) accepting the hypothesis when an alternative is true. While evaluating the two, there comes the need for the probability distributions of the test statistic and both when the hypothesis is true and false or the null and the non-null distributions of which always the former could be easily obtained. But to study the other important characters of the test, e.g., the unbiassedness, the non-null distribution will be required. The non-null distributions of T^2 and D^2 were first obtained in 1938 (2, 3). The non-null distributions of the multiple correlation coefficient when the set of independent variates is ~~fixed~~ either random or fixed, ~~fix~~ were firs obtained in 1928 by Fisher (4), who called them the A and C-distributions respectively of the multiple correlation coefficient.

The multiple correlation coefficient when the set of independent variates follows a normal law, and when the dependent variate is an artificially constructed pseudo-variate taking fixed values, happens to be an algebraical derivate of T^2 and D^2 (5) and we shall denote this by R^2 throughout this thesis. The similarity or the duality between the distributions of this R^2 and C-distributions was later established and one of the most useful results that followed is the derivation of the T^2 and D^2 distributions from the C-distribution. The distribution of this R^2 which thus appears to be playing quite an important role in the multivariate theory was obtained by various authors from time to time (2, 3, 6, 7, 8 & 9). But it should however be noted that all the authors in attacking the problem whether through an analytic approach or through the concept of hyperspace geometry have been invariably assuming the invariance property of this multiple correlation coefficient under any linear transformation of the set of p variates, and employing the same property advantageously to simplify the otherwise unsolved problem. But to know if some interesting properties are covered under the simplicity effected by the above property, the distribution of R^2 is derived without assuming the invariance property in this thesis. T^2 and D^2 distributions are then derived as particular cases of this distribution.

D^2 statistic, as already referred to, is appropriate only in the case of two p-variate samples to test the hypothesis of homogeneity of set of means under the assumption of identical covariance matrix and the problem of k-multivariate samples is not yet completely solved even under the assumption of identical but unspecified population covariance matrix. There have been two approaches suggested neither of which is fully investigated as yet. One of them is by the method of discriminant function leading to the study of the (k-1) canonical correlations, which are defined as the roots of a certain determinantal equation; the joint distribution of (null distribution) of these roots was obtained by Hsu and others (10, 11, 12) and the null distribution of the individual roots was later obtained by Roy and others (13, 14). One drawback from which this approach suffers, as already pointed out by Bartlett (15), is that the order of roots in the sample may not correspond to the order in the population. Another difficulty is the lack of knowledge as to which root or what particular functions of the roots provides the most powerful test; for though the nonnull distributions of these roots have been to some extent studied by Bartlett, Roy and others (16, 17, 18), these attempts hardly throw much light on the comparative power functions. The other approach is through the likelihood ratio method which gives the ratio of the within sample covariance matrix to the total covariance matrix as the appropriate statistic to test the homogeneity of the means in the k-saple case. This approach was due to Wilks and others (19, 20, 21) and the the criterion defined by him to measure the relation between two sets of variates, one set containing (k-1) variates and the other containing p variates, also reduces to the above likelihood ratio when the set of (k-1) variates is fixed and the other set follows a multivariate normal law. Another interesting property of this criterion, that makes this approach rather preferable, is its being a symmetric function of the above referred (k-1) roots. The null distribution of this statistic was partly studies by Wald and Brookner (22). The nonnull distribution of this statistic was not obtained before even for any value of k 2. Anderson (9) has given the moments of the statistic for $k = 3$ when the set of $k-1=2$ variates is fixed and the other set follows a normal law; i.e., for the likelihood ratio appropriate to test the homogeneity of means for three multivariate samples. In the present paper the derivation of the nonnull distribution for $k = 3$ will be given

(i) when the set of (k-1) variates is fixed - the statistic is a generalisation of R^2 and

(ii) when the set of p variates if fixed - which thus gives the statistic relevant to what is called the general regression problem by Hsu (23).

One particular object of this study is to investigate if some sort of duality (as in the case of the distribution of R^2 and Fisher's z distribution) exists between the above two nonnull distributions. In the case of the distribution of the canonical correlations, ~~the~~ duality has been shown by Bartlett (16) to hold only if the number of the population non-vanishing correlations is one or zero.

2. Distribution of the multiple correlation coefficient when the dependent variate is fixed and the set of independent variates follows a multivariate normal law.

Let $[\, x_{io} ; \{ x_{iu} \} \,]$ $(u = 1, 2, \cdots p ; i = 1, 2, \cdots N)$ be a sample of size N, x_o being the dependent variate and $\{x_u\}$ the set of independent variates.

Given that i) $\{x_u\}$ follow a normal law

ii) $E(x_r) = m_r x_o$; $r = 1, 2, \cdots p$

and iii) $|\alpha_{rs}|$ is the covariance matrix of $\{x_r\}$

then the probability of the sample is given by

$$\text{Const. } \operatorname{Exp} -\frac{1}{2}\left[\sum_{r,s=1}^{p} \alpha^{rs} \left\{ \sum_{i=1}^{N} (x_{ir} - m_r x_{io})(x_{is} - m_s x_{io}) \right\} \right] \cdot \prod_{i=1}^{N} \prod_{r=1}^{p} \{ dx_{ir} \} \qquad (2.1)$$

where $\quad |\alpha^{rs}| = |\alpha_{rs}|^{-1}$

Under these conditions, the likelihood ratio λ appropriate to test the hypothesis $H (m_1 = 0, \cdots m_p = 0)$ is then given by

$$(\lambda)^{\frac{2}{N}} = 1 - R^2 = \frac{\begin{vmatrix} v_{oo} & v_{o1} & \cdots & v_{op} \\ v_{1o} & v_{11} & \cdots & v_{1p} \\ \vdots & \vdots & & \vdots \\ v_{po} & v_{p1} & \cdots & v_{pp} \end{vmatrix}}{v_{oo} \cdot \begin{vmatrix} v_{11} & v_{12} & \cdots & v_{1p} \\ v_{21} & v_{22} & \cdots & v_{2p} \\ \vdots & \vdots & & \vdots \\ v_{p1} & v_{p2} & \cdots & v_{pp} \end{vmatrix}} \qquad (2.2)$$

where $\quad v_{rs} = \sum_{i=1}^{N} x_{ir} x_{is}$.

Writing $\quad h^2 = \sum_{i=1}^{N} x_{io}^2 , \quad h^2 w_r = \sum_{i=1}^{N} x_{ir} x_{io}$

and $\quad a_{rs} = \sum_{i=1}^{N} (x_{ir} - w_r x_{io})(x_{is} - w_s x_{io})$

$$v_{rs} = a_{rs} + h^2 w_r w_s \qquad (2.3)$$

(2.2) can be rewritten as

$$1 - R^2 = \frac{|a_{rs}|}{|v_{rs}|} \qquad (r, s = 1, 2, \cdots p)$$

$$\text{or } \quad R^2 = h^2 \sum v^{rs} w_r w_s \qquad (2.4)$$

We will then proceed to derive the sampling distribution of λ or R^2 for fixed x_0 when

i) H is true

and ii) H is not true.

The distribution of $[\{a_{rs}\}, \{w_r\}]$; $(r, s = 1, 2, \ldots p)$ derived from (2.1) is given by

$$\text{Const. } \mathcal{E}xp -\frac{1}{2}\left[\sum_{r,s=1}^{p} \alpha^{rs}\{a_{rs} + (w_r - m_r)(w_s - m_s)h^2\}\right]$$
$$|a_{rs}|^{\frac{N-p-2}{2}} \prod_{r,s}\{da_{rs}\}\prod_{r}\{dw_r\}. \quad (2.5)$$

Writing $\Delta^2 = \sum_{r,s=1}^{p} \alpha^{rs} m_r m_s$ and using (2.3), (2.5) leads to the distribution of $[\{v_{rs}\}, \{w_r\}]$ given by

$$\text{Const. } \mathcal{E}xp -\frac{1}{2}\left[\sum_{r,s} \alpha^{rs} v_{rs} + h^2\Delta^2 + \sum_{r,s} \alpha^{rs} m_r w_s h^2\right]$$
$$|v_{rs}|^{\frac{N-p-2}{2}} \left(1 - h^2 \sum v^{rs} w_r w_s\right)^{\frac{N-p-2}{2}} \prod_{r,s}(dv_{rs})\prod_{r} dw_r$$
$$(2.6)$$

Because Δ^2 is a + ve definite quadratic form in m_r's, $\Delta^2 = 0$ if and only if each m_r vanishes. Thus $\Delta^2 = 0$ implies that H is true and vice versa. Thus when H is true, the distribution of R^2 which is a quadratic in w_r of rank p follows immediately from (2.6) as

$$\frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-p}{2}\right)\Gamma\left(\frac{p}{2}\right)} (1 - R^2)^{\frac{N-p-2}{2}} (R^2)^{\frac{p}{2}-1} \, dR^2 \quad (2.7)$$

which is the well known Beta distribution.

ii) Distribution when H is not true.

Without loss of generality $\alpha^{rs} = 0$ for $r \neq s$

$\alpha^{rr} = 1$

The former derivations involve the assumption of all the m's except m_1 vanishing. We will not assume any m to vanish so that

$$\Delta^2 = m_1^2 + m_2^2 + \cdots + m_p^2$$

and so let us write $m_r = \Delta c_r$

where $c_1^2 + c_2^2 + \cdots + c_p^2 = 1$ (2.8)

We then adopt the transformation $\{w_r\}$ to $\{b_r\}$

$$b_r = \frac{h\begin{vmatrix} w_r & w_1 & w_2 & \cdots & w_{r-1} \\ v_{1r} & v_{11} & v_{12} & \cdots & v_{1\,r-1} \\ \vdots & & & & \vdots \\ v_{r-1\,r} & v_{r-1\,1} & v_{r-1\,2} & \cdots & v_{r-1\,r-1} \end{vmatrix}}{\begin{vmatrix} v_{11} & v_{12} & \cdots & v_{1r} \\ v_{21} & v_{22} & \cdots & v_{2r} \\ \vdots & \vdots & & \vdots \\ v_{r1} & v_{r2} & \cdots & v_{rr} \end{vmatrix}} \qquad (2.9)$$

and letting

$$\chi_r^2 = \frac{\begin{vmatrix} v_{11} & v_{12} & \cdots & v_{1r} \\ v_{21} & v_{22} & \cdots & v_{2r} \\ \vdots & \vdots & & \vdots \\ v_{r1} & v_{r2} & \cdots & v_{rr} \end{vmatrix}}{\begin{vmatrix} v_{11} & v_{12} & \cdots & v_{1\,r-1} \\ v_{21} & v_{22} & \cdots & v_{2\,r-1} \\ \vdots & \vdots & & \vdots \\ v_{r-1\,1} & v_{r-1\,2} & \cdots & v_{r-1\,r-1} \end{vmatrix}}$$

the jacobian of the transformation is given by

$$\frac{\partial(w_1, w_2, \cdots w_p)}{\partial(b_1, b_2, \cdots b_p)} = (\chi_1^2 \chi_2^2 \cdots \chi_p^2) = |v_{rs}|$$

We observe that under this tranformation

$$R^2 = \sum_{r=1}^{p} b_r^2 \chi_r^2 \qquad (2.10)$$

$$v_{rr} = \chi_r^2 + \sum_{u=r-1}^{1} z_{ru}^2$$

and $\quad h w_r = b_r \chi_r \; x_r + \sum_{u=1}^{r-1} b_u \chi_u \cdot z_{ru}$

where $\quad z_{ru} = b_{ru} \chi_u$

The joint distribution of

$$\{ b_r \chi_r ,\; z_{r\,r-1}, \cdots z_{r1} ,\; \chi_r^2 \} \quad (r = 1, 2, \cdots p)$$

can then be derived from (2.6) with the use of (2.8), (2.9) and (2.10) as

$$\text{Const.}\; e^{-\frac{h^2 \Delta^2}{2}} \left(1 - \sum b_r^2 \chi_r^2 \right)^{-\frac{N-p-2}{2}}$$

$$\prod_{r=1}^{p} \left[\operatorname{Exp} - \tfrac{1}{2} \left(\chi_r^2 + z_{r\,r-1}^2 + \cdots z_{r1}^2 + h \Delta c_r \sum_{u=1}^{r} b_u \chi_u z_{ru} \right) \right.$$

$$\prod_{r=1}^{p} \left(d \chi_r^2 \; d z_{r\,r-1} \cdots d z_{r1} \;\; d\, \overline{b_r \chi_r} \right) \qquad (2.11)$$

Setting the further transformation

$$\begin{aligned}
b_p \chi_p &= R \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{p-2} \cos\theta_{p-1} \\
b_{p-1} \chi_{p-1} &= R \cos\theta_1 \cdot \cos\theta_2 \cdots \cos\theta_{p-2} \sin\theta_{p-1} \\
&\cdots\cdots\cdots\cdots \\
&\cdots\cdots\cdots\cdots \\
b_r \chi_r &= R \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{r-1} \sin\theta_r \\
&\cdots\cdots\cdots\cdots \\
&\cdots\cdots\cdots\cdots \\
b_2 \chi_2 &= R \cos\theta_1 \sin\theta_2 \\
b_1 \chi_1 &= R \sin\theta_1
\end{aligned} \qquad (2.12)$$

the jacobian is given by

$$\frac{\partial (b_p \chi_p ,\; b_{p-1} \chi_{p-1} , \cdots , b_2 \chi_2 , b_1 \chi_1)}{\partial (R, \theta_1 , \theta_2 , \cdots \cdots \theta_{p-2} , \theta_{p-1})}$$

$$= R^{p-1} \cos^{p-2}\theta_1 \; \cos^{p-3}\theta_2 \cdots \cos\theta_{p-2} \qquad (2.13)$$

Using (2.12) and (2.13), (2.11) leads then to the joint distribution of

$$(R^2, \cos\theta_1, \ldots, \cos\theta_{p-1})$$

$$\{x_r^2, z_{rr-1}, \ldots z_{r1}\} \; (r = 1, 2, \ldots p) \quad \left\{ \begin{array}{l} 0 \le R^2 \le 1 \\ 0 \le \theta \le \frac{\pi}{2} \\ 0 \le \chi^2 \le \infty \\ -\infty \le z \le \infty \end{array} \right.$$

which is given as

$$\text{Const. } e^{-\frac{h^2\Delta^2}{2}} (R^2)^{\frac{p}{2}-1} (1-R^2)^{\frac{N-p-2}{2}} dR^2$$

$$\times \prod_{r=1}^{p} \left\{ e^{-\frac{1}{2}\chi_r^2} (\chi_r^2)^{\frac{N-r-2}{2}} \sum_{l_{rr}=0}^{\infty} \frac{(h^2R^2\Delta^2)^{l_{rr}}}{(2l_{rr})!} (c_r^2 \chi_r^2)^{l_{rr}} (\cos\theta_1 \cdots \cos\theta_{r-1} \sin\theta_r)^{2l_{rr}} d\chi_r^2 \right.$$

$$e^{-\frac{1}{2} z_{rr-1}^2} \sum_{l_{rr-1}=0}^{\infty} \frac{(h^2R^2\Delta^2)^{l_{rr-1}}}{(2l_{rr-1})!} (c_r^2)^{l_{rr-1}} (z_{rr-1}^2)^{l_{rr-1}-\frac{1}{2}} (\cos\theta_1 \cdots \cos\theta_{r-2} \sin\theta_{r-1})^{2l} dz_{rr-1}^2$$

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

$$e^{-\frac{1}{2} z_{ru}^2} \sum_{l_{ru}=0}^{\infty} \frac{(h^2R^2\Delta^2)^{l_{ru}}}{(2l_{ru})!} (c_r^2)^{l_{ru}} (z_{ru}^2)^{l_{ru}-\frac{1}{2}} (\cos\theta_1 \cdots \cos\theta_{u-1} \sin\theta_u)^{2l_{ru}} dz_{ru}^2$$

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

$$\left. e^{-\frac{1}{2} z_{r1}^2} \sum_{l_{r1}=0}^{\infty} \frac{(h^2R^2\Delta^2)^{l_{r1}}}{(2l_{r1})!} (c_r^2)^{l_{r1}} (z_{r1}^2)^{l_{r1}-\frac{1}{2}} (\sin\theta_1)^{2l_{r1}} dz_{r1}^2 \right\}$$

$$d\theta_1 \, d\theta_2 \cdots d\theta_{p-1} \qquad (2.14)$$

When we integrate out for

$$\{x_r^2, z_{rr-1}^2, \ldots z_{r1}^2\} \quad (r = 1, 2, \ldots p)$$

and making use of the relation

$$\frac{1}{\sqrt{\pi}} \frac{2^{2l_{ru}} \Gamma(l_{ru} + \frac{1}{2})}{\Gamma(2l_{ru}+1)} = \frac{1}{\Gamma(l_{ru}+1)}$$

we get at the joint distribution of

$$(R^2, \cos\theta_1, \ldots \cos\theta_{p-1}) \quad \text{as}$$

$$\text{Const. } e^{-\frac{h^2\Delta^2}{2}} (R^2)^{\frac{p}{2}-1} (1-R^2)^{\frac{N-p-2}{2}} dR^2 \cos^{p-2}\theta_1 \cdots \cos^2\theta_{p-3} \cos\theta_{p-2}$$

$$\times \prod_{r=1}^{p} \left\{ \sum_{l_{rr}=0}^{\infty} \left(\frac{h^2R^2\Delta^2}{2}\right)^{l_{rr}} \frac{1}{l_{rr}!} (c_r^2)^{l_{rr}} (\cos\theta_1 \cdots \cos\theta_{r-1} \sin\theta_r)^{2l_{rr}} \frac{\Gamma(\frac{N-r}{2} + l_{rr})}{\Gamma(l_{rr} + \frac{1}{2})} \right.$$

$$\left[\prod_{u=1}^{r-1} \left(\sum_{l_{ru}=0}^{\infty} \left(\frac{h^2R^2\Delta^2}{2}\right)^{l_{ru}} \frac{1}{l_{ru}!} (c_r^2)^{l_{ru}} (\cos\theta_1 \cdots \cos\theta_{u-1} \sin\theta_u)^{2l_{ru}} \right. \right.$$

(2.15)

Writing $\sum_{u=1}^{r} l_{ru} = l_{r.}$, $\sum_{r=u}^{p} l_{ru} = l_{.u}$ and $l = \sum_{r=1}^{p} l_{r.} = \sum_{u} l_{.u}$ (2.15) can be written as

$$\text{Const. } e^{-\frac{h^2\Delta^2}{2}} (R^2)^{\frac{p}{2}-1} (1-R^2)^{\frac{N-p-2}{2}} dR^2$$

$$\sum_{l=0}^{\infty}\Bigg[\left(\frac{h^2R^2\Delta^2}{2}\right)^{l}\frac{1}{l!}\sum_{(\sum_{r,u} l_{ru}=l)}\Bigg\{\frac{l!}{l_{11}!\,l_{21}!\,l_{22}!\cdots l_{p1}!\cdots l_{pp}!}(c_1^2)^{l_{1.}}(c_2^2)^{l_{2.}}\cdots(c_p^2)^{l_{p.}}$$

$$\prod_{r}\frac{\Gamma\left(\frac{N-r}{2}+l_{rr}\right)}{\Gamma\left(l_{rr}+\frac{1}{2}\right)}$$

$$(\sin\theta_1)^{2l_{.1}}(\cos\theta_1)^{2(l-l_{.1}-l_{1.}+l_{11})+p-2}$$

$$\cdots\cdots$$

$$(\sin\theta_u)^{2l_{.u}}(\cos\theta_u)^{2(l-l_{.1}\cdots-l_{.u}-l_{1.}+l_{11})+p-u}$$

$$\cdots\cdots$$

$$(\sin\theta_{p-1})^{2l_{.p-1}}(\cos\theta_{p-1})^{2(l_{pp}-l_{1.}+l_{11})}\Bigg\}\Bigg]$$

$$d\theta_1\, d\theta_2 \cdots d\theta_{p-1} \tag{2.16}$$

The general term of the subsummation (under the curly brackets inside the square bracket) in (2.16) when integrated for $\theta_1, \theta_2, \cdots \theta_{p-1}$ $(0 \leq \theta \leq \frac{\pi}{2})$ gives

$$\frac{l!}{l_{11}!\,l_{21}!\,l_{22}!\cdots l_{p1}!\cdots l_{pp}!}(c_1^2)^{l_{1.}}(c_2^2)^{l_{2.}}\cdots(c_p^2)^{l_{p.}}$$

$$\prod_{r=1}^{p}\left\{\frac{\Gamma\left(\frac{N-r}{2}+l_{rr}\right)}{\Gamma\left(l_{rr}+\frac{1}{2}\right)}\right\}\prod_{u=1}^{p-1}\frac{\Gamma\left(\frac{2l_{.u}+1}{2}\right)\Gamma\left(\frac{2\,\overline{l-l_{.1}\cdots-l_{.u}-l_{1.}+l_{11}}+p-u}{2}\right)}{\Gamma\left(\frac{2\,\overline{l-l_{.1}\cdots-l_{.u-1}-l_{1.}+l_{11}}+p-u+1}{2}\right)} \tag{2.17}$$

Noting that $l_{1.} = l_{11}$, and $l_{.p} = l_{pp}$, (2.17) will be written as

$$\frac{l!}{l_{1.}!\,l_{2.}!\cdots l_{p.}!}\,(c_1^2)^{l_{1.}}\cdots(c_p^2)^{l_{p.}}\,\frac{1}{\Gamma\left(\frac{p}{2}+l\right)}$$

$$\cdot\frac{l_{1.}!}{l_{11}!}\cdot\frac{l_{2.}!}{l_{21}!\,l_{22}!}\cdots\frac{l_{u.}!}{l_{u1}!\cdots l_{uu}!}\cdots\frac{l_{p.}!}{l_{p1}!\cdots l_{pp}!}$$

$$\Gamma\left(\frac{2l_{.1}+1}{2}\right)\Gamma\left(\frac{2l_{.2}+1}{2}\right)\cdots\cdots\Gamma\left(\frac{2l_{.p-1}+1}{2}\right)$$

$$\frac{\Gamma\left(\frac{N-1}{2}+l_{11}\right)\Gamma\left(\frac{N-2}{2}+l_{22}\right)\cdots\cdots\Gamma\left(\frac{N+1-p}{2}+l_{p-1\,p-1}\right)\Gamma\left(\frac{N-p}{2}+l_{pp}\right)}{\Gamma\left(\frac{1}{2}+l_{11}\right)\Gamma\left(\frac{1}{2}+l_{22}\right)\cdots\cdots\Gamma\left(\frac{1}{2}+l_{p-1\,p-1}\right)}\Bigg\} \tag{2.18}$$

$$\Gamma\left(\frac{N-2}{2}\right)\cdots\cdots\Gamma\left(\frac{N-p}{2}\right).$$

$$\times\Bigg\{\frac{\Gamma\left(l+\frac{N-1}{2}\right)}{\Gamma\left(l+\frac{p}{2}\right)}\;\frac{l!}{l_{1.}!\,l_{2.}!\cdots l_{p.}!}\,(c_1^2)^{l_{1.}}(c_2^2)^{l_{2.}}\cdots(c_p^2)^{l_{p.}}$$

$$\frac{l_{1.}!}{l_{11}!}\;\frac{l_{2.}!}{l_{21}!\,l_{22}!}\cdots\frac{l_{u.}!}{l_{u1}!\cdots l_{uu}!}\cdots\frac{l_{p.}!}{l_{p1}!\cdots l_{pp}!}$$

$$\frac{B\left(\frac{N-p}{2}+l_{.p},\,l_{.p-1}+\frac{1}{2}\right)B\left(\frac{N-1-p}{2}+l_{.p}+l_{.p-1},\,l_{.p-2}+\frac{1}{2}\right)\cdots B\left(\frac{N-2}{2}+l_{.p}+\cdots+l_{.2},\,l_{.1}+\frac{1}{2}\right)}{B\left(\frac{N-2}{2},\,l_{11}+\frac{1}{2}\right)B\left(\frac{N-3}{2},\,l_{22}+\frac{1}{2}\right)\cdots\cdots B\left(\frac{N-p}{2},\,l_{p-1\,p-1}+\frac{1}{2}\right)}\Bigg\} \tag{2.19}$$

This expression is to be summed over all values of l_{ru} satisfying $\sum_{u\leq r} l_{ru} = l$ for a fixed value of l. The summation will be done in two stages. As a first step we will sum the latter portion of (2.19) (viz)

$$\frac{l_{1.}!}{l_{11}!}\;\frac{l_{2.}!}{l_{21}!\,l_{22}!}\cdots\frac{l_{u.}!}{l_{u1}!\cdots l_{uu}!}\cdots\frac{l_{p.}!}{l_{p1}!\cdots l_{pp}!}$$

$$\frac{B\left(\frac{N-p}{2}+l_{.p},\,l_{.p-1}+\frac{1}{2}\right)\cdots\cdots B\left(\frac{N-2}{2}+l_{.p}+\cdots+l_{.2},\,l_{.1}+\frac{1}{2}\right)}{B\left(\frac{N-2}{2},\,l_{11}+\frac{1}{2}\right)\cdots\cdots B\left(\frac{N-p}{2},\,l_{p-1\,p-1}+\frac{1}{2}\right)} \tag{2.20}$$

... values of $l_{1.}, l_{2.}, \ldots l_{p.}$

The product of Beta functions in the numerator can be written as a constant multiplied by

$$\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\cdots\int_0^{\frac{\pi}{2}} \Big\{ (\sin^2\lambda_1)^{l_{.1}} (\cos^2\lambda_1)^{\frac{N-3}{2}+l_{.p}+\cdots+l_{.2}} d\lambda_1$$
$$(\sin^2\lambda_2)^{l_{.2}} (\cos^2\lambda_2)^{\frac{N-4}{2}+l_{.p}+\cdots+l_{.3}} d\lambda_2$$
$$(\sin^2\lambda_u)^{l_{.u}} (\cos^2\lambda_u)^{\frac{N-2-u}{2}+l_{.p}+\cdots+l_{.u+1}} d\lambda_u$$
$$\cdots$$
$$(\sin^2\lambda_{p-1})^{l_{.p-1}} (\cos^2\lambda_{p-1})^{\frac{N-1-p}{2}+l_{.p}} d\lambda_{p-1}\Big\} \qquad (2.21)$$

The integrand above can be split up as

$$\Big[\{ (\sin^2\lambda_1)^{l_{p1}} (\cos^2\lambda_1 \sin^2\lambda_2)^{l_{p2}} \cdots (\cos^2\lambda_1\cdots\cos^2\lambda_{p-1})^{l_{pp}} \} (\sin^2\lambda_{p-1})^{l_{p-1\,p-1}} (\cos^2\lambda_{p-1})^{\frac{N-1+p}{2}} \Big]$$

$$\Big[\{ (\sin^2\lambda_1)^{l_{p-1\,1}} (\cos^2\lambda_1 \sin^2\lambda_2)^{l_{p-1\,2}} \cdots (\cos^2\lambda_1\cdots\cos^2\lambda_{p-2})^{l_{p-1,p-1}} \} (\sin^2\lambda_{p-2})^{l_{p-2\,p-2}} (\cos^2\lambda_{p-2})^{\frac{N-p}{2}} \Big]$$

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

$$\Big[\{ (\sin^2\lambda_1)^{l_{u1}} (\cos^2\lambda_1 \sin^2\lambda_2)^{l_{u2}} \cdots (\cos^2\lambda_1\cdots\cos^2\lambda_{u-1})^{l_{uu}} \} (\sin^2\lambda_{u-1})^{l_{u-1\,u-1}} (\cos^2\lambda_{u-1})^{\frac{N-1+u}{2}} \Big]$$

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

$$\Big[\{ (\sin^2\lambda_1)^{l_{21}} (\cos^2\lambda_1)^{l_{22}} \} (\sin^2\lambda_1)^{l_{11}} (\cos^2\lambda_1)^{\frac{N-1}{2}} \Big]$$

$$(2.22)$$

We then consider a rearrangement of (2.20) where the expressions in the curly brackets of the successive components in (2.22) are to be combined with

$$\frac{l_{p.}!}{l_{p1}!\,l_{p2}!\cdots l_{pp}!},\quad \frac{l_{p-1.}!}{l_{p-1\,1}!\cdots l_{p-1\,p-1}!},\ \cdots\ \frac{l_{u.}!}{l_{u1}!\cdots l_{uu}!},\ \cdots\ \frac{l_{2.}!}{l_{21}!\,l_{22}!}$$

respectively (2.23)

The expression in the curly bracket of the first component in (2.22) when combined with the appropriate factor in (2.23), will be summed up over all values of $l_{p1}, l_{p2}, \cdots l_{pp}$ for a fixed value of l_p. (ie)

$$\sum_{\substack{l_{p1}+\cdots+l_{pp} \\ = l_p}} \left\{ \frac{l_p!}{l_{p1}! \cdots l_{pp}!} (\sin^2\lambda_1)^{l_{p1}} \cdots (\cos^2\lambda_1 \cdots \cos^2\lambda_{p-1})^{l_{pp}} \right\}$$

$$= (\sin^2\lambda_1 + \cos^2\lambda_1 \sin^2\lambda_2 + \cdots + \cos^2\lambda_1 \cdots \cos^2\lambda_{p-1})^{l_p}$$

$$= 1$$

The remaining expression of the first component (viz) $(\sin^2\lambda_{p-1})^{l_{p-1\,p-1}} (\cos^2\lambda_{p-1})^{\frac{N-1+p}{2}}$ which and which alone now contains λ_{p-1} will be integrated w.r.t. λ_{p-1} so that the net result of the first summation followed by the first integration is $B(l_{p-1\,p-1} + \frac{1}{2}, \frac{N-p}{2})$, and this Beta function cancels out with the last Beta function of the denominator of (2.20). Similarly carrying on similar summations each time followed by a similar integration with each of the successive components of (2.22) we arrive at the last component of (2.22)

(viz)

$$\left[\left\{ (\sin^2\lambda_1)^{l_{21}} (\cos^2\lambda_1)^{l_{22}} \right\} (\sin^2\lambda_1)^{l_{11}} (\cos^2\lambda_1)^{\frac{N-3}{2}} \right]$$

the expression in the curly brackets of which will be combined with the respective factor of (2.23)

(viz). $\dfrac{l_2.!}{l_{21}!\ l_{22}!}$ and summed over all values of l_{21} and l_{22} for fixed values of $l_{21}+l_{22}=l_2$. This summation resulting unity, will then be followed by the integration of the remaining expression in the component (viz) $(\sin^2\lambda_1)^{l_{11}} (\cos^2\lambda_1)^{\frac{N-3}{2}}$ w.r.t λ_1

The result of this integration is a Beta function which will again cancel out with the ~~last~~ first Beta function left in the denominator of (2.20).

Thus the subsummation of (2.16) finally reduces to the summation

$$\sum_{l_1+l_2+\cdots+l_p=l} \frac{l!}{l_1!\cdots l_p!}(c_1^2)^{l_1}\cdots(c_p^2)^{l_p}\frac{\Gamma(\frac{N-1}{2}+l)}{\Gamma(\frac{p}{2}+l)}$$

$$= \frac{\Gamma(\frac{N-1}{2}+l)}{\Gamma(\frac{p}{2}+l)} \quad \text{because of (2.8).}$$

Thus the distribution of R^2 in (2.16) finally comes out as

$$\text{Const.}\; e^{-\frac{h^2\Delta^2}{2}} (R^2)^{\frac{p}{2}-1} (1-R^2)^{\frac{N-p-2}{2}} dR^2$$

$$\sum_{l=0}^{\infty} \left\{\frac{h^2\Delta^2 R^2}{2}\right\}^l \frac{1}{l!}\cdot\frac{\Gamma(\frac{N-1}{2}+l)}{\Gamma(\frac{p}{2}+l)}.$$

where the constant is given by $\frac{1}{B(\frac{p}{2},\frac{N-p}{2})}\frac{\Gamma(\frac{p}{2})}{\Gamma(\frac{N-1}{2})}$

So the distribution written in the standard form is

$$\frac{1}{B(\frac{p}{2},\frac{N-p}{2})}\cdot e^{-\frac{h^2\Delta^2}{2}} (R^2)^{\frac{p}{2}-1} (1-R^2)^{\frac{N-p-2}{2}} {}_1F_1\left(\frac{N-1}{2},\frac{p}{2};\frac{h^2\Delta^2R^2}{2}\right) \quad (2.24)$$

Derivation of the T^2 distribution.

Let $\{x_{ir}\}$ $\left(\begin{smallmatrix} i=1,2,\cdots N \\ r=1,2,\cdots p\end{smallmatrix}\right)$ be a sample of size N from a p-variate normal population. If $\{\bar{x}_r\}$ denotes the set of sample means and

$$b_{rs} = \sum_{i=1}^{N}(x_{ir}-\bar{x}_r)(x_{is}-\bar{x}_s), \text{ then}$$

Hotelling's T^2 is defined as

$$T^2 = N(N-1)\sum b^{rs}\,\bar{x}_r\,\bar{x}_s \quad (2.25)$$

Now in our problem setting

$x_{i0} = 1$ for any i, we observe that

$$h^2 = \sum_{i=1}^{N} x_{i0}^2 = N$$

$$h^2 w_r = \sum_{i=1}^{N} x_{i0}\,x_{ir} = N\bar{x}_r \text{ or } w_r = \bar{x}_r$$

and

$$a_{rs} = \sum_{i=1}^{N}(x_{ir}-w_r x_{i0})(x_{is}-w_s x_{i0})$$

$$= b_{rs}. \quad (2.26)$$

$\therefore$ From (2.4), (2.25) and (2.26) we have

$$1-R^2 = \frac{|a_{rs}|}{|v_{rs}|}$$

$$= \frac{|a_{rs}|}{|a_{rs}+h^2 w_r w_s|} = \frac{1}{1+h^2\Sigma a^{rs} w_r w_s}$$

$$= \frac{1}{1+\frac{T^2}{N-1}} \quad (2.27)$$

Thus for the above transformation of R^2 to T^2, we have

$$dR^2 = \frac{1}{N-1}\left(\frac{1}{1+\frac{T^2}{N-1}}\right)^2 dT^2$$

and $R^2 = \dfrac{T^2}{(N-1)+T^2}$

and (2.24) leads to the distribution of T^2 given by

$$\frac{1}{B(\frac{p}{2},\frac{N-p}{2})} e^{-\frac{N\Delta^2}{2}} \left\{\frac{T^2}{(N-1)+T^2}\right\}^{\frac{p}{2}-1} \cdot \left\{\frac{N-1}{N-1+T^2}\right\}^{\frac{N-p}{2}-1} \frac{(N-1)}{(N-1+T^2)^2}$$

$$\,_1F_1\left(\frac{N-1}{2},\frac{p}{2};\frac{N\Delta^2 T^2}{N-1+T^2}\right) dT^2$$

$$= \frac{1}{B(\frac{p}{2},\frac{N-p}{2})} e^{-\frac{N\Delta^2}{2}} \frac{(N-1)^{\frac{N-p}{2}}(T^2)^{\frac{p}{2}-1}}{\{(N-1)+T^2\}^{\frac{N}{2}}} dT^2$$

$$\,_1F_1\left(\frac{N-1}{2},\frac{p}{2};\frac{N\Delta^2 T^2}{N-1+T^2}\right) \quad (2.28)$$

Derivation of the Distribution of the D^2 statistic.

If all the variates $(x_0; x_1, x_2, \dots x_p)$ were measured from the respective sample means $\hat{x}_0, \{\hat{x}_r\}$, the power of $|a_{rs}|$ in (2.5) will be $\frac{N-p-3}{2}$ and this will be carried on to the end to the power of $(1-R^2)$ in (2.24) and the distribution in this case will be

$$\frac{1}{B(\frac{p}{2},\frac{N-p-1}{2})} e^{-\frac{h^2\Delta^2}{2}} (R^2)^{\frac{p}{2}-1} (1-R^2)^{\frac{N-p-3}{2}} \,_1F_1\left(\frac{N-1}{2},\frac{p}{2};\frac{h^2\Delta^2 R^2}{2}\right) \quad (2.29)$$

Let $\{x_{ir}\}$ $(i = 1, 2, \dots n_1;\ r = 1, 2, \dots p)$ and $\{x_{ir}\}$ $(i = n_1+1, \dots n_1+n_2;\ r = 1, 2, \dots p)$ be two samples of sizes n_1 and n_2 respectively from two p variate normal populations having identical but unspecified covariance matrix $|\alpha_{rs}|$. If $\{\bar{x}_r\}$, and $\{\bar{x}'_r\}$ where $\bar{x}_r = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_{ir})$ and $\bar{x}'_r = \frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2}(x_{ir})$, be the two sets of sample means and if

$$(n_1+n_2)\, c_{rs} = \sum_{i=1}^{n_1}(x_{ir} - \bar{x}_r)(x_{is} - \bar{x}_s) + \sum_{i=n_1+1}^{n_1+n_2}(x_{ir} - \bar{x}'_r)(x_{is} - \bar{x}'_s)$$

then the D^2 statistic is defined by

$$p(D^2) = \sum_{r,s=1}^{p} c^{rs}(\bar{x}_r - \bar{x}'_r)(\bar{x}_s - \bar{x}'_s)$$

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

(2.30)

To consider this as a special case of the problem we have considered, let the $n_1 + n_2 = N$ sets of observations be looked upon as a single sample. All the variates $\{x_r\}$ are to be measured from the respective sample means $\{\hat{x}_r\}$ where

$$\hat{x}_r = \frac{1}{N}\sum_{i=1}^{N} x_{ir} = \frac{n_1\bar{x}_r + n_2\bar{x}'_r}{N}$$

or the origin of measurement for the r th variate is taken at this point. In accordance with this general requirement let us set

$$x_{i0} = \frac{n_2}{N} \quad \text{for } i = 1, 2, \dots n_1$$

$$= -\frac{n_1}{N} \quad \text{for } i = n_1+1, \dots (n_1+n_2).$$

Then

$$h^2 = \sum_{i=1}^{N} x_{i0}^2 = \frac{n_1 n_2}{N}$$

$$h^2 w_r = \sum_{i=1}^{N} x_{i0}\, x_{ir} = \frac{n_1 n_2}{N}(\bar{x}_r - \bar{x}'_r)$$

$$\text{or, } w_r = (\bar{x}_r - \bar{x}'_r)$$

and $a_{rs} = \sum_{i=1}^{N} (x_{ir} - w_r x_{i0})(x_{is} - w_s x_{i0})$

$$= \sum_{i=1}^{n_1} \left\{x_{ir} - \frac{n_2}{N}(\bar{x}_r - \bar{x}'_r)\right\}\left\{x_{is} - \frac{n_2}{N}(\bar{x}_s - \bar{x}'_s)\right\}$$

$$+ \sum_{i=n_1+1}^{n_1+n_2} \left\{x_{ir} + \frac{n_1}{N}(\bar{x}_r - \bar{x}'_r)\right\}\left\{x_{is} + \frac{n_1}{N}(\bar{x}_s - \bar{x}'_s)\right\}$$

$= N C_{rs}$ because of the origin of measurement.

Thus we will have –

$$1 - R^2 = \frac{1}{1 + h^2 \sum a^{rs} w_r w_s}$$

$$= \frac{1}{1 + \frac{\bar{n} p D^2}{2N}} \quad \text{where } \frac{2}{\bar{n}} = \frac{1}{n_1} + \frac{1}{n_2}$$

$$\therefore dR^2 = \left(\frac{1}{1 + \frac{\bar{n} p D^2}{2N}}\right)^2 \frac{\bar{n} p}{2N} \, dD^2$$

$$R^2 = \frac{\bar{n} p D^2}{2N + \bar{n} p D^2} \quad \text{so that (2.29)}$$

leads to the distribution of D^2 given by

$$\frac{1}{B\left(\frac{p}{2}, \frac{N-p-1}{2}\right)} e^{-\frac{\bar{n}\Delta^2}{4}} \left(\frac{\bar{n} p D^2}{2N + \bar{n} p D^2}\right)^{\frac{p}{2}-1} \left(\frac{2N}{2N + \bar{n} p D^2}\right)^{\frac{N-p-3}{2}} \frac{2N\bar{n}p}{(2N + \bar{n}pD^2)^2} dD^2$$

$$ {}_1F_1\left(\frac{N-1}{2}, \frac{p}{2}; \frac{\bar{n}^2 \Delta^2 p D^2}{8N + 4\bar{n} p D^2}\right)$$

which ~~reduces to~~ can be rewritten as

$$\frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{N-p-1}{2}\right)} \left(\frac{\bar{n} p}{2N}\right)^{\frac{p}{2}} e^{-\frac{\bar{n}\Delta^2}{4}} (D^2)^{\frac{p}{2}-1} \left(1 + \frac{p\bar{n}D^2}{2N}\right)^{-\frac{N-1}{2}} dD^2$$

$$ {}_1F_1\left(\frac{N-1}{2}, \frac{p}{2}; \frac{\bar{n}^2 \Delta^2 p D^2}{8N + 4\bar{n} p D^2}\right) \quad (2.31)$$

Here $\Delta^2 = \sum \alpha^{rs}(m_r - m'_r)(m_s - m'_s)$ and thus ~~pD~~ $p\Delta^2/p$ will be the population parameter corresponding to D^2 defined by (2.30); where m_r and m'_r have the meaning

$$E(x_{ir}) = m_r x_{i0} \qquad i = 1 \text{ to } n_1$$

$$= m'_r x_{i0} \qquad i = n_1 + 1 \text{ to } n_1 + n_2.$$

3. Non null and Null distributions of the Wilks statistic for $k=3$ when the set of $k-1\ (=2)$ variates is fixed and the other set of p variates follows a normal law.

Let $\{x_{i\alpha};\ x_{ir}\}$, $(i=1,2,\cdots N;\ \alpha=0,1;\ r=2,3\cdots p+1)$ be a sample of size N, x_0 and x_1 being the two dependent variates and $\{x_r\}$ $(r=2,3,\cdots p+1)$ being the set of p independent variates. The origin of measurement for each of the variate (both on the dependent and independent side) is taken at the respective sample mean. Given that

1) $\{x_r\}$ follows a normal law

2) $E(x_r) = \gamma_r x_{1\cdot 0} + \delta_r x_0$
where $x_{1\cdot 0}$ is the residual $= x_1 - b_{10} x_0$

and 3) $|\alpha_{rs}|$ is the covariance matrix of $\{x_r\}$

then the probability of the sample is given by

$$\text{Const}\cdot \text{Exp} -\frac{1}{2}\Big[\sum_{r,s=2}^{p+1} \alpha^{rs} \Big\{\sum_{i=1}^{N}(x_{ir}-\gamma_r x_{i1\cdot 0}-\delta_r x_{i0})(x_{is}-\gamma_s x_{i1\cdot 0}-\delta_s x_{i0})\Big\}\Big]$$

$$\times \prod_{i=1}^{N}\prod_{r=2}^{p+1}(dx_{ir}) \qquad (3.1)$$

where $|\alpha^{rs}|$ is the matrix reciprocal to $|\alpha_{rs}|$.

Under these conditions, the likelihood ratio λ appropriate to test the hypothesis $H\ (\gamma_2=\gamma_3=\cdots=\gamma_{p+1}=0;\ \delta_2=\delta_3=\cdots=\delta_{p+1}=0)$

is then given by

$$\lambda^{\frac{2}{N}} = 1-W^2 = \frac{\begin{vmatrix} v_{00} & v_{01} & \cdots & v_{0p+1}\\ v_{10} & v_{11} & \cdots & v_{1p+1}\\ \vdots & \vdots & & \vdots\\ v_{p+10} & v_{p+11} & \cdots & v_{p+1p+1}\end{vmatrix}}{\begin{vmatrix} v_{00} & v_{01}\\ v_{10} & v_{11}\end{vmatrix}\begin{vmatrix} v_{22} & v_{23} & \cdots & v_{2p+1}\\ v_{32} & v_{33} & \cdots & v_{3p+1}\\ \vdots & & & \\ v_{p+12} & v_{p+13} & \cdots & v_{p+1p+1}\end{vmatrix}} \qquad (3.2)$$

where $v_{rs} = \sum_{N} x_{ir} x_{is}$.

Writing $\sum_{i=1}^{N} x_{i1\cdot 0}^2 = g^2$, $\sum_{i=1}^{N} x_{i0}^2 = h^2$

$$g^2 c_r = \sum_{i=1}^{N} x_{ir} \cdot x_{i1\cdot 0}, \quad h^2 d_r = \sum_{i=1}^{N} x_{ir} x_{i0} \quad (r=2 .. p+1)$$

and $$a_{rs} = \sum_{i=1}^{N} (x_{ir} - c_r x_{i1\cdot 0} - d_r x_{i0})(x_{is} - c_s x_{i1\cdot 0} - d_s x_{i0})$$
$$(r, s = 2, 3, .. p+1) \qquad (3\cdot 3)$$

it can be seen that

$$\left.\begin{aligned} v_{rs} &= a_{rs} + g^2 c_r c_s + h^2 d_r d_s \\ \text{and } 1 - W^2 &= \frac{|a_{rs}|}{|v_{rs}|} \end{aligned}\right\} \qquad (3\cdot 4)$$

We will now proceed to get the distribution of W^2, known as Wilks statistic, when H is not true, for fixed values of x_0 & $x_{1\cdot 0}$. The distribution, when H is true can be deduced at that stage. The distribution of $[\{a_{rs}\}, \{c_r\}$ and $\{d_r\}]$ $(r = 2, .. p+1)$ derived from (3.1) is given by

$$\text{Const. Exp} - \tfrac{1}{2}\Big[\sum_{r,s=2}^{p+1} \alpha^{rs} \{a_{rs} + g^2 (c_r - \zeta_r)(c_s - \zeta_s) + h^2 (d_r - \delta_r)(d_s - \delta_s)\}\Big]$$
$$|a_{rs}|^{\frac{N-p-4}{2}} \prod_{r,s} (da_{rs}) \cdot \prod_{r} (dc_r\, dd_r) \qquad (3\cdot 5)$$

In virtue of $(3\cdot 4)$, $(3\cdot 5)$ leads to the distribution of $[\{v_{rs}\}, \{c_r\}, \& \{d_r\}]$ given by

$$\text{Const. Exp} - \tfrac{1}{2}\Big[g^2 \sum \alpha^{rs} \zeta_r \zeta_s + h^2 \sum \alpha^{rs} \delta_r \delta_s\Big]$$
$$\times \text{ Exp} - \tfrac{1}{2}\Big[\sum_{r,s=2}^{p+1} \alpha^{rs} \{v_{rs} - 2g^2 c_r \zeta_s - 2h^2 d_r \delta_s\}\Big]$$
$$|v_{rs}|^{\frac{N-p-4}{2}} (1-W^2)^{\frac{N-p-4}{2}} \prod_{r,s} dv_{rs} \prod (dc_r\, dd_r)$$
$$(3\cdot 6)$$

The distribution problem is enormously simplified if we make use of the property of invariance of the Wilks statistic for linear transformations of variates within sets. In fact this property is partly used already when we started the problem with $x_{1\cdot 0}$ and x_0 on the dependent side. Without loss of generality, we can further write

i) $\alpha^{rs} = 0$ if $r \neq s$ and (ii) the variances of
$= 1$ if $r = s$ $\quad x_0$ and $x_{1\cdot 0}$ are unity.

$\zeta_r = 0$ for $r \neq 2$

$\delta_r = 0$ for $r \neq 3$ $\qquad (3\cdot 7)$

x_0 and x_1 being fixed the second assumption leads to
$g^2 = h^2 = N-1$ $\qquad (3\cdot 7a)$

using this invariance property (3.7.), (3.6) reduces to

$$\text{Const}.\ \operatorname{Exp} -\tfrac{1}{2}\left(g^2 \zeta_2^2 + h^2 \delta_3^2\right)$$

$$\times\ \operatorname{Exp} -\tfrac{1}{2}\left[\sum_{r=2}^{p+1} v_{rr}\right](1-W^2)^{\frac{N-p-4}{2}} |v_{rs}|^{\frac{N-p-4}{2}} \prod_{r,s} dv_{rs}$$

$$\times\ \operatorname{Exp}\left(g^2 \zeta_2 c_2 + h^2 \delta_3 d_3\right) \prod_{r=2}^{p+1} \left(dc_r\, dd_r\right) \qquad (3.8)$$

We now change from

$$(c_2, c_3, \ldots c_{p+1};\ d_2, d_3, \ldots d_{p+1})$$

to a new set of variables

$$(B_2, B_3, \ldots B_{p+1};\ b_2, b_3, \ldots b_{p+1})$$

by the following transformations.

$$B_r = b_{1r.(23\ldots r-1;\,0)}$$

$$= \frac{g\begin{vmatrix} c_r & c_{r-1} & \cdots & c_2 \\ v_{rr-1} - h^2 d_r d_{r-1} & \cdots & \cdots & \cdots \\ \vdots & & |v_{ij} - h^2 d_i d_j|_2^{r-1} & \\ v_{r2} - h^2 d_r d_2 & & & \end{vmatrix}}{|v_{ij} - h^2 d_i d_j|_2^{r}}, \qquad (3.9)$$

and

$$b_r = b_{0r.(23\ldots r-1)}$$

$$= \frac{h\begin{vmatrix} d_r & d_{r-1} & \cdots & d_2 \\ v_{r-1\,r} & v_{r-1\,r-1} & \cdots & v_{r-1\,2} \\ \vdots & & & \\ v_{2r} & v_{2\,r-1} & \cdots & v_{22} \end{vmatrix}}{|v_{ij}|_2^{r}} \qquad (3.10)$$

The above transformations in particular lead to

$$B_2 = \frac{g c_2}{v_{22} - h^2 d_2^2}$$

$$b_2 = \frac{h d_2}{v_{22}}$$

and $$b_3 = \frac{h(v_{22} d_3 - v_{32} d_2)}{(v_{33} v_{22} - v_{32}^2)} \qquad (3.11)$$

and the Jacobian of the transformation

$$J = \frac{\partial(c_2, c_3, \ldots c_{p+1}, d_2, d_3, \ldots d_{p+1})}{\partial(B_2, B_3, \ldots B_{p+1}, b_2, b_3, \ldots b_{p+1})}$$

is given by

$$J = \left\{ |v_{ij}|_2^{p+1} \right\} \left\{ |v_{ij} - h^2 d_i d_j|_2^{p+1} \right\} \qquad (3.12)$$

As usual defining

$$x_r^2 = \frac{|v_{ij}|_2^r}{|v_{ij}|_2^{r-1}} \quad \text{and} \quad x_{r.0}^2 = \frac{|v_{ij} - h^2 d_i d_j|_2^r}{|v_{ij} - h^2 d_i d_j|_2^{r-1}}$$

for $r > 2$

and $$x_2^2 = v_{22} \quad \text{and} \quad x_{2.0}^2 = v_{22} - h^2 d_2^2 \qquad (3.13)$$

we observe that

$$R^2_{0(23\ldots p+1)} = \sum_{r=2}^{p+1} b_r^2 x_r^2$$

and $$R^2_{1(23\ldots p+1);0} = \sum_{r=2}^{p+1} B_r^2 x_{r.0}^2 \qquad \ldots (3.14)$$

where $R_{0(23\ldots p+1)}$ is the multiple correlation coefficient of x_0 with the set $\{x_r\}$ and $R_{1(23,\ldots p+1);0}$ is the multiple correlation coefficient of the residual $x_{1.0}$ with the set of residuals $\{x_{r.0}\}$. For simplicity we will write these two multiple correlation coefficients by R_0 and R_1 respectively. We also observe that

$$1 - R_0^2 = \frac{|v_{ij} - h^2 d_i d_j|_2^{p+1}}{|v_{ij}|_2^{p+1}} = \frac{x_{2.0}^2 \cdot x_{3.0}^2 \cdots x_{p+1.0}^2}{x_2^2 \cdot x_3^2 \cdots x_{p+1}^2}$$

using the new quantities defined in (3.13), (3.11) can be rewritten as

$$g c_2 = (B_2 \chi_{2.0})\, \chi_2\, (1 - b_2^2 \chi_2^2)^{\frac{1}{2}}$$

$$h d_2 = b_2 \chi_2^2$$

$$\text{and } h d_3 = b_3 \chi_3^2 + b_2 \chi_2 z_{32} \qquad (3.16)$$

where $z_{32} = \dfrac{\sum_{i=1}^{N} x_{i3}\, x_{i2}}{\chi_2}$

The Jacobian J in (3.12) also can be rewritten with the use of (3.13), (3.14) and (3.15) as

$$J = (1 - R_0^2)^{\frac{1}{2}} (\chi_2^2 \cdots \chi_{p+1}^2)^{\frac{3}{2}} (\chi_{2\cdot 0}^2 \cdots \chi_{p+1\cdot 0}^2)^{\frac{1}{2}}$$

$$= \left(1 - \sum_{r=2}^{p+1} b_r^2 \chi_r^2\right)^{\frac{1}{2}} (\chi_2^2 \cdots \chi_{p+1}^2)^{\frac{3}{2}} (\chi_{2\cdot 0}^2 \cdots \chi_{p+1\cdot 0}^2) \qquad (3.17)$$

using the well known result

$$1 - W^2 = (1 - R_0^2)(1 - R_1^2) \qquad (3.18)$$

$$= \left\{ 1 - \sum_{r=2}^{p+1} b_r^2 \chi_r^2 \right\} \left\{ 1 - \sum_{r=2}^{p+1} B_r^2 \chi_{r\cdot 0}^2 \right\}$$

we see that the portion in (3.8) relevant to the problem of the distribution of W^2 is one which leads to the distribution of

$$\chi_2^2,\ \chi_3^2,\ z_{32},\ \{ b_r \chi_r \} \text{ and } \{ B_r \chi_{r\cdot 0} \},$$

which with the use of (3.16), (3.17) and (3.18) can be written as

$$\text{Const. Exp} - \tfrac{1}{2} (g^2 \zeta_2^2 + h^2 \delta_3^2)$$

$$\times \text{Exp} - \tfrac{1}{2} (\chi_3^2 + z_{32}^2 + \chi_2^2)$$

$$\times \text{Exp} \left\{ g \zeta_2 (B_2 \chi_{2\cdot 0})(1 - b_2^2 \chi_2^2)^{\frac{1}{2}} \chi_2 + h \delta_3 (\chi_3 \cdot b_3 \chi_3 + z_{32} b_2 \right.$$

$$\left(1 - \sum_{r=2}^{p+1} b_r^2 \chi_r^2\right)^{\frac{N-p-3}{2}} \left(1 - \sum B_r^2 \chi_{r\cdot 0}^2\right)^{\frac{N-p-4}{2}} (\chi_2^2)^{\frac{N-3}{2}} (\chi_3^2)^{\frac{N-4}{2}}$$

$$d(\chi_2^2)\, d(\chi_3^2)\, d z_{32} \prod_{r=2}^{p+1} d(b_r \chi_r) \prod_{r=2}^{p+1} d(B_r \chi_{r\cdot 0}) \qquad (3.19)$$

Writing $H_0^2 = \sum_{r=4}^{p+1} b_r^2 \chi_r^2$ and $H_1^2 = \sum_{r=3}^{p+1} B_r^2 \chi_{r\cdot 0}^2$ (3.20)

the distribution of

$$\{ H_0^2,\ H_1^2,\ b_2\chi_2,\ b_3\chi_3,\ B_2\chi_{2\cdot 0},\ \chi_2^2,\ \chi_3^2\ \&\ z_{32})$$

can be immediately derived from (3.19) as

$$\text{Const.}\ \text{Exp} -\tfrac{1}{2}\left(g^2 \zeta_2^2 + h^2 \delta_3^2 \right)$$
$$\times\ \text{Exp} -\tfrac{1}{2}\left(\chi_3^2 + z_{32}^2 + \chi_2^2 \right)$$
$$\times\ \text{Exp}\left[g\zeta_2 (B_2\chi_{2\cdot 0})(1 - b_2^2\chi_2^2)^{\frac{1}{2}} \chi_2 + h\delta_3 (\chi_3 b_3\chi_3 + z_{32} b_2\chi_2 \right.$$
$$(1 - H_0^2 - b_2^2\chi_2^2 - b_3^2\chi_3^2)^{\frac{N-p-3}{2}} (1 - H_1^2 - B_2^2\chi_{2\cdot 0}^2)^{\frac{N-p-4}{2}}$$
$$(H_0^2)^{\frac{p-2}{2}-1} (H_1^2)^{\frac{p-1}{2}-1} (\chi_2^2)^{\frac{N-3}{2}} (\chi_3^2)^{\frac{N-4}{2}}$$
$$d(\chi_3^2)\, d(\chi_2^2)\, dz_{32}\, dH_0^2\, dH_1^2\, d(b_2\chi_2)\, d(b_3\chi_3)\, d(B_2\chi_{2\cdot 0})$$

(3.21)

Applying then the transformation

$$\begin{aligned} H_0 &= R_0 \cos\theta_1 \cos\theta_2 \\ b_3\chi_3 &= R_0 \cos\theta_1 \sin\theta_2 \\ b_2\chi_2 &= R_0 \sin\theta_1 \\ H_1 &= R_1 \cos\theta_3 \\ B_2\chi_{2\cdot 0} &= R_2 \sin\theta_3 \end{aligned} \qquad (3.22)$$

the Jacobian of which is given by

$$\frac{\partial(H_0, H_1, B_2\chi_{2\cdot 0}, b_2\chi_2, b_3\chi_3)}{\partial(R_1^2, R_0^2, \theta_1, \theta_2, \theta_3)} = R_0 R_1 \cos\theta_1 \quad \ldots (3.23)$$

the joint distribution of

$(R_0^2, R_1^2, \theta_1, \theta_2, \theta_3, \chi_3^2, \chi_2^2, \& z_{32})$ is given by

$$\left(\begin{array}{l} 0 \le R^2 \le 1 \\ 0 \le \theta \le \frac{\pi}{2} \end{array} \right)$$

$$\text{Const}\cdot \text{Exp} -\tfrac{1}{2}\left(g^2\zeta_2^2 + h^2\delta_3^2 \right) \times \text{Exp} -\tfrac{1}{2}\left(\chi_3^2 + \chi_2^2 + z_{32}^2 \right)$$
$$\sum_{l=0}^{\infty} \frac{(g^2\zeta_2^2)^l}{2l!} (1 - R_0^2 \sin^2\theta_1)^l (\chi_2^2 R_1^2 \sin^2\theta_3)^l$$
$$\sum_{m=0}^{\infty} \frac{(h^2\delta_3^2)^m}{2m!} (\chi_3^2 R_0^2 \cos^2\theta_1 \sin^2\theta_2)^m$$
$$\sum_{n=0}^{\infty} \frac{(h^2\delta_3^2)^n}{2n!} (z_{32}^2 R_0^2 \sin^2\theta_1)^n$$
$$(1 - R_1^2)^{\frac{N-p-4}{2}} (1 - R_0^2)^{\frac{N-p-3}{2}} (R_0 \cos\theta_1 \cos\theta_2)^{p-3} (R_1 \cos\theta_3)^{p-2} R_0 \cos\theta_1$$

N-3 … N-4

Integration of (3.24) for $(\chi_3^2, \chi_2^2, z_{32}, \theta_2, \text{and } \theta_3)$ leads to the joint distribution of $(R_0^2, R_1^2 \text{ and } \theta_1)$ as given by

$$
\begin{aligned}
&\text{Const.}\ \operatorname{Exp} -\tfrac{1}{2}\left(g^2\zeta_2^2 + h^2\delta_3^2\right) \\
&(1-R_0^2)^{\frac{N-p-3}{2}} (1-R_1^2)^{\frac{N-p-4}{2}} (R_0^2 R_1^2)^{\frac{p-2}{2}} (\cos\theta_1)^{p-2} \\
&\sum_{l=0}^{\infty} \left(\frac{g^2\zeta_2^2}{2}\right)^{l} \frac{1}{l!} \left\{R_1^2(1-R_0^2\sin^2\theta_1)\right\}^{l} \frac{\Gamma(\frac{N-1}{2}+l)}{\Gamma(\frac{p}{2}+l)} \\
&\sum_{m=0}^{\infty} \left(\frac{h^2\delta_3^2}{2}\right)^{m} \frac{1}{m!} (R_0^2\cos^2\theta_1)^{m} \frac{\Gamma(\frac{N-2}{2}+m)}{\Gamma(\frac{p-1}{2}+m)} \\
&\sum_{n=0}^{\infty} \left(\frac{h^2\delta_3^2}{2}\right)^{n} \frac{1}{n!} (R_0^2\sin^2\theta_1)^{n} \qquad (3.25) \\
&\qquad d R_0^2 \; d R_1^2 \; d\theta_1
\end{aligned}
$$

In (3.25) if we write the binomial expansion for $(1-R_0^2\sin^2\theta_1)^{l}$ in the first series as

$$(1-R_0^2\sin^2\theta_1)^{l} = \sum_{l_1+l_2=l} \frac{l!}{l_1!\, l_2!} (\cos^2\theta_1)^{l_1} \left\{\sin^2\theta_1(1-R_0^2)\right\}^{l_2}$$

and integrating for θ_1 $(0 \leq \theta_1 \leq \frac{\pi}{2})$ we are led to the joint distribution of R_0^2 and R_1^2 as

$$
\begin{aligned}
&\text{Const.}\ \operatorname{Exp} -\tfrac{1}{2}\left\{\left(g^2\zeta_2^2 + h^2\delta_3^2\right)\right\} (1-R_0^2)^{\frac{N-p-3}{2}} (1-R_1^2)^{\frac{N-p-4}{2}} (R_0^2R_1^2)^{\frac{p-2}{2}} \\
&\sum_{l_1,l_2,m,n} \left\{ \left(\frac{g^2\zeta_2^2}{2}\right)^{l_1+l_2} \left(\frac{h^2\delta_3^2}{2}\right)^{m+n} \frac{\Gamma(\frac{N-1}{2}+l_1+l_2)\,\Gamma(\frac{N-2}{2}+m)\,\Gamma(\frac{p-1}{2}+m+l_1)\,\Gamma(l_2+n+\cdots}{\Gamma(\frac{p}{2}+l_1+l_2)\,\Gamma(\frac{p-1}{2}+m)\,\Gamma(\frac{p}{2}+l_1+l_2+m+n)} \right. \\
&\qquad \left. \frac{(R_0^2)^{m+n}\,(R_1^2)^{l_1+l_2}\,(1-R_0^2)^{l_2}}{l_1!\; l_2!\; m!\; n!} \right\} d R_0^2 \; d R_1^2 \qquad (3.26)
\end{aligned}
$$

We now transform R_0^2 and R_1^2 to a pair of new variables u and W^2 where the latter is given by (3.18) and

$$u = \frac{R_0^2}{W^2}$$

so that the limits of u are 0 and 1 independent of W^2. Then the joint distribution of u & W^2 can be immediately derived from (3.26) as

$$\text{Const. } \operatorname{Exp} - \frac{1}{2} \left(g^2 \zeta_2^2 + h^2 \delta_3^2 \right)$$

$$\times \; (W^2)^{p-1} (1 - W^2)^{\frac{N-p-4}{2}} \{u(1-u)\}^{\frac{p-2}{2}}$$

$$\times \sum_{l_1, l_2, m, n} \left\{ \left(\frac{g^2 \zeta_2^2}{2}\right)^{l_1+l_2} \left(\frac{h^2 \delta_3^2}{2}\right)^{m+n} \frac{\Gamma\left(\frac{N-1}{2} + l_1 + l_2\right) \Gamma\left(\frac{N-2}{2} + m\right) B\left(\frac{p-1}{2} + m + l_1, l_2 + n + \frac{1}{2}\right)}{\Gamma\left(\frac{p}{2} + l_1 + l_2\right) \Gamma\left(\frac{p-1}{2} + m\right) l_1! \, l_2! \, m! \, n!} \right.$$

$$\left. \{1 - u(W^2)\}^{-\left(\frac{p-1}{2} + l_1\right)} u^{m+n} (1-u)^{l_1+l_2} (W^2)^{m+n+l_1+l_2} \right\} dW^2 \, du$$

(3.27)

Integrating (3.27) with respect to u $(0 \leq u \leq 1)$ and while doing so using the well known result

$$B(\beta, \gamma - \beta) \, {}_2F_1(\alpha, \beta; \gamma, x)$$

$$= \int_0^1 t^{\beta-1} (1-t)^{\gamma-\beta-1} (1-xt)^{-\alpha} dt \text{ for } |x| \leq 1$$

we get ~~at~~ the distribution of W^2 as given by

$$\text{Const. } \operatorname{Exp} - \frac{1}{2} \left\{ \left(g^2 \zeta_2^2 + h^2 \delta_3^2 \right) \right\} (W^2)^{p-1} (1 - W^2)^{\frac{N-p+4}{2}}$$

$$\sum_{l_1, l_2, m, n} \left\{ \left(\frac{g^2 \zeta_2^2}{2}\right)^{l_1+l_2} \left(\frac{h^2 \delta_3^2}{2}\right)^{m+n} \frac{(W^2)^{m+n+l_1+l_2}}{l_1! \, l_2! \, m! \, n!} \right.$$

$$\frac{\Gamma\left(\frac{N-1}{2} + l_1 + l_2\right) \Gamma\left(\frac{N-2}{2} + m\right) \Gamma\left(\frac{p}{2} + m + n\right) B\left(\frac{p-1}{2} + m + l_1, l_2 + n + \frac{1}{2}\right)}{\Gamma(p + l_1 + l_2 + m + n) \, \Gamma\left(\frac{p-1}{2} + m\right)}$$

$$\left. F\left(\frac{p-1}{2} + l_1, \frac{p}{2} + m + n; p + m + n + l_1 + l_2, W^2\right) \right\} \quad d(W^2)$$

(3.28)

In view of (3.7), it can be easily seen that all the variates x_0, $x_{1 \cdot 0}$, $\{x_r\}$ correspond to Hotellings canonical variates and that ζ_2^2 and δ_3^2

are the two non vanishing or the two canonical correlations ρ_1^2 and ρ_0^2. Then writing

$$(N-1)\rho_0^2 = E^2, \quad (N-1)\rho_1^2 = F^2,$$ the distribution can be written as

$$\text{Const. } \operatorname{Exp} -\tfrac{1}{2}\{E^2+F^2\}\,(W^2)^{p-1}(1-W^2)^{\frac{N-p-4}{2}}\,dW^2$$

$$\sum_{l_1,l_2,m,n}\Big\{\left(\frac{E^2}{2}\right)^{m+n}\left(\frac{F^2}{2}\right)^{l_1+l_2}\frac{(W^2)^{m+n+l_1+l_2}}{l_1!\,l_2!\,m!\,n!}$$

$$\frac{\Gamma\left(\frac{N-1}{2}+l_1+l_2\right)\Gamma\left(\frac{N-2}{2}+m\right)\Gamma\left(\frac{p}{2}+m+n\right)B\left(\frac{p-1}{2}+m+l_1,\,l_2+n+\frac{1}{2}\right)}{\Gamma(p+l_1+l_2+m+n)\,\Gamma\left(\frac{p-1}{2}+m\right)}$$

$$F\left(\frac{p-1}{2}+l_1,\ \frac{p}{2}+m+n;\ p+m+n+l_1+l_2,\ W^2\right)\Big\} \quad (3.29)$$

The expression for the Constant, say C, in the above distribution ~~is sea~~ can be seen to be independent of E or F and as such it will be same as the constant of the distribution, deduced from (3.29) with $E=F=0$ - (i.e.) constant of the distribution when H is true, which will be deduced from above. Thus the distribution, when H is true is given as

$$~~\text{Const}~~\ C\,(W^2)^{p-1}(1-W^2)^{\frac{N-p-4}{2}}$$

$$\times\ \frac{\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma(p)}\ F\left(\frac{p-1}{2},\ \frac{p}{2};\ p,\ W^2\right)dW^2 \quad (3.30)$$

The hypergeometric series in the above distribution is known to satisfy the differential equation

$$x(1-x)\frac{d^2y}{dx^2}+\left\{p-\left(p+\tfrac{1}{2}\right)x\right\}\frac{dy}{dx}-\frac{p(p-1)}{4}\,y=0$$

with $x = W^2$. By well known methods of solving this type of second order differential equations it will be seen that

$$y = \left(1+\sqrt{1-x}\right)^{-(p-1)}$$ is a principal solution and so it will be seen that

$$F\left(\frac{p-1}{2},\ \frac{p}{2};\ p,\ W^2\right) = \frac{2^{p-1}}{\left(1+\sqrt{1-W^2}\right)^{p-1}} \quad (3.31).$$

Using (3.31), the distribution given in (3.30).

leads to the distribution of $L = (1-W^2)$ as given by

$$\text{Const.}\, K \cdot 2^{p-1} (1-\sqrt{L})^{p-1} (\sqrt{L})^{N-p-4}\, dL .$$ where K is the product of the constant, C in (3.30) and the other constant terms of that expression. By integrating the above distribution with respect to $\sqrt{L}$ between 0 & 1 and equating the integral to unity it is seen that

$$K \cdot 2^{p-1} = \frac{1}{2\,B(N-p-2,\, p)}$$ whence in virtue of the relation $\Gamma(x+\frac{1}{2})\,\Gamma(x+1) = \dfrac{\sqrt{\pi}\,\Gamma(2x+1)}{2^{2x}}$

it is seen that
$$C = \frac{2^{N-p-3}}{\pi\, \Gamma(N-p-2)} .$$

Thus the nonnull and null distributions of $1-W^2 = L$ are respectively given by

$$\frac{2^{N-p-3}}{\pi\,\Gamma(N-p-2)}\, e^{-\frac{1}{2}(E^2+F^2)}\, (1-L)^{p-1}\, L^{\frac{N-p-4}{2}}$$

$$\sum_{l_1, l_2, m, n} \Bigg\{ \left(\frac{E^2}{2}\right)^{m+n} \left(\frac{F^2}{2}\right)^{l_1+l_2} \frac{(1-L)^{m+n+l_1+l_2}}{l_1!\, l_2!\, m!\, n!}\; \frac{\Gamma(\frac{N-1}{2}+l_1+l_2)\,\Gamma(\frac{N-2}{2}+m)\,\Gamma(\frac{p}{2}+m+n)}{\Gamma(p+l_1+l_2+m+n)\;\Gamma(\frac{p-1}{2}+m)}$$

$$B\left(\tfrac{p-1}{2}+m+l_1,\; l_2+n+\tfrac{1}{2}\right) F\left[\tfrac{p-1}{2}+l_1,\; \tfrac{p}{2}+m+n;\; p+m+n+l_1+l_2,\; (1-L)\right] \Bigg\}\, dL \qquad (3.32)$$

and

$$\frac{1}{2\,B(N-p-2,\, p)}\, (1-\sqrt{L})^{p-1} (\sqrt{L})^{N-p-4}\, dL . \qquad (3.33)$$

4. Nonnull and Null distributions of the Wilks statistic relevant to the general regression problem with two variates in the dependent set and p variates in the predicting or independent set.

Let $\{x_{i\alpha}\,;\ x_{ir}\}$ $(i = 1, 2, .. N;\ \alpha = 0, 1;\ r = 2, 3, .. p+1)$ be a set of N observations on the dependent set $\{x_\alpha\}$ and the independent set $\{x_r\}$, the latter taking a set of fixed values. The origin of measurement for each of the variates is taken at the respective mean. Given that

i) (x_0, x_1) follow a bivariate normal law with

ii) $E(x_{1\cdot 0}) = \beta_{12}x_{2\cdot 0} + \beta_{13}x_{3\cdot 0} + \cdots + \beta_{1p+1\cdot}x_{p+1\cdot 0}$

and $E(x_0) = \beta_{02}x_2 + \beta_{03}x_3 + \cdots + \beta_{0p+1}x_{p+1}$

where $x_{1\cdot 0}$, $\{x_{r\cdot 0}\}$ are the residuals of x_1, $\{x_r\}$ respectively after eliminating the portion due to regression on x_0.

iii) K_1^2 & K_0^2 are the two population ~~pooled~~ variances defined by

$$\begin{aligned} K_1^2 &= E\,(x_{1\cdot 0} - \Sigma\, \beta_{1r}\, x_{r\cdot 0})^2 \\ K_0^2 &= E\,(x_0 - \Sigma\, \beta_{0r}\, x_r)^2 \end{aligned} \qquad (4\cdot 1)$$

The hypothesis that $\{x_\alpha\}$ is independent of $\{x_r\}$ will mean in terms of the parameters that

$$H: \quad \beta_{1r} = 0\ ;\ \beta_{0r} = 0 \quad \text{for all } r = 2, 3, .. p+1.$$

The sample wilks statistic W^2 is ~~also~~ a function of the appropriate likelihood ratio λ and W^2 is given as

$$\lambda^{\frac{2}{N}} = 1 - W^2 = \left[\frac{\sum_{i=1}^{N} \left\{ x_{i1\cdot 0} - \sum_{r=2}^{p+1} b_{1r\cdot(023\cdots r-1)}\, x_{ir\cdot(023\cdots r-1)} \right\}^2}{\sum_{i=1}^{N} x_{i1\cdot 0}^2} \times \frac{\sum_{i=1}^{N} \left\{ x_{i0} - \sum_{r=2}^{p+1} b_{0r\cdot(23\cdots r-1)}\, x_{ir\cdot(23\cdots r-1)} \right\}^2}{\sum_{i=1}^{N} x_{i0}^2} \right] \qquad (4\cdot 2)$$

$$\text{where} \quad b_{1r\cdot(023\ldots r-1)} = \frac{\sum_{i=1}^{N}\{x_{i1}\, x_{ir\cdot(023\ldots r-1)}\}}{\sum_{i=1}^{N}\{x_{ir\cdot(023\ldots r-1)}\}^2}$$

$$\text{and} \quad b_{0r\cdot(23\ldots r-1)} = \frac{\sum_{i=1}^{N}\{x_{i0}\, x_{ir\cdot(23\ldots r-1)}\}}{\sum_{i=1}^{N}\{x_{ir\cdot(23\ldots r-1)}\}^2} \qquad (4\cdot3)$$

$$\text{with} \quad x_{r\cdot(023\ldots r-1)} = \frac{\begin{vmatrix} x_r & v_{r\,r-1} & \cdots & v_{r2} & v_{02} \\ x_{r-1} & v_{r-1\,r-1} & & v_{r-1\,2} & v_{0\,r-1} \\ \vdots & & & & \\ x_2 & v_{r-1\cdot2} & & v_{22} & v_{02} \\ x_0 & v_{r-1\cdot0} & & v_{20} & v_{00} \end{vmatrix}}{\begin{vmatrix} v_{r-1\,r-1} & v_{r-1\,r-2} & \cdots & v_{r-1\cdot2} & v_{r-1\cdot0} \\ v_{r-2\,r-1} & v_{r-2\,r-2} & & v_{r-2\cdot2} & v_{r-2\,0} \\ \vdots & & & & \\ v_{2\,r-1} & v_{2\,r-2} & \cdots & v_{22} & v_{20} \\ v_{0\,r-1} & v_{0\,r-2} & \cdots & v_{02} & v_{00} \end{vmatrix}}$$

$$\text{and} \quad x_{r\cdot(23\ldots r-1)} = \frac{\begin{vmatrix} x_r & v_{r\,r-1} & \cdots & v_{r2} \\ x_{r-1} & & & \\ \vdots & & |v_{ij}|_2^{r-1} & \\ x_2 & & & \end{vmatrix}}{|v_{ij}|_2^{r-1}} \qquad (4\cdot4)$$

$$\text{and} \quad v_{rs} = \sum_{i=1}^{N} x_{ir}\, x_{is}\,.$$

(4·2) can be rewritten as

$$1-W^2 = \frac{\chi^2_{1\cdot(023\ldots p+1)}}{\chi^2_{1\cdot0}} \times \frac{\chi^2_{0\cdot(23\ldots p+1)}}{\chi^2_0} \qquad \ldots (4\cdot5)$$

where the definition of χ^2's is quite clear by comparing (4·5) with (4·2). We will derive the distribution of $1-W^2$ for fixed values of $\{x_r\}$ when the hypothesis ~~to~~ is being tested is not true -(ie) the non null distribution - from which the null distribution, (ie) the distribution when H is true can be immediately deduced..

$\chi^2_{1\cdot0}$ and χ^2_0 can be split up as

$$\sum_{i=1}^{N} x_{i1\cdot0}^2 = \chi^2_{1\cdot0} = \chi^2_{1\cdot(023\ldots p+1)} + \sum_{r=2}^{p+1} z^2_{1r\cdot(02\ldots r-1)}$$

$$\text{and} \sum_{i=1}^{N} x_{i0}^2 = \chi^2_0 = \chi^2_{0\cdot(23\ldots p+1)} + \sum_{r=2}^{p+1} z^2_{0r\cdot(23\ldots r-1)}$$

where $z_{1r\cdot(02\ldots r-1)}$, for simplicity $z_{1r\cdot0}$ is given by

$$z^2_{1r\cdot0} = b^2_{1r\cdot(023\ldots r-1)} \sum_{i=1}^{N} x^2_{ir\cdot(023\ldots r-1)}$$

$$= b^2_{1r\cdot0}\, \chi^2_{r}, \text{ say.}$$

and $z_{ar.(23\cdots r-1)}$, for simplicity z_{or} is given by $z_{or}^2 = b_{or.(23\cdots r-1)}^2 \sum_{i=1}^{N} x_{ir.(23\cdots r-1)}^2$

$$= b_{or}^2 \, \chi_r^2 \text{ (say)}.$$

Thus introducing these quantities (4.5) can be rewritten as

$$1-W^2 = \frac{\chi^2_{1.(023\cdots p+1)}}{\chi^2_{1.(023\cdots p+1)} + \sum_{r=2}^{p+1} z_{1r.0}^2} \times \frac{\chi^2_{0.(23\cdots p+1)}}{\chi^2_{0(23\cdots p+1)} + \sum_{r=2}^{p+1} z_{or}^2} \qquad (4.6)$$

It can be easily seen that the two factors on the right hand side in (4.6) (viz)

$$\frac{\chi^2_{1.(023\cdots p+1)}}{\chi^2_{1.(023\cdots p+1)} + \sum z_{1r.0}^2}, \quad \frac{\chi^2_{0.(23\cdots p+1)}}{\chi^2_{0.(23\cdots p+1)} + \sum z_{or}^2} \qquad (4.7)$$

are $(1-R_1^2)$ and $(1-R_0^2)$ respectively, where R_1^2 is the sample multiple correlation coefficient of the residual $x_{1.0}$ with the set of residuals $\{x_{r.0}\}$ and R_0^2 is the sample multiple correlation coefficient of x_0 with the set $\{x_r\}$. It can be rightly pointed that the relation between the two sets of variates (x_0, x_1) and $\{x_r\}$ is completely characterised by these two coefficients.

The portion relevant for the distribution problem of W^2, taken out from the total sample probability density, namely, the distribution, when H is not true, of

$$\chi^2_{1.(023\cdots p+1)}, \; \{z_{1r.0}\}, \; \chi^2_{0.(23\cdots p+1)}, \; \{z_{or}\}$$

is given by

$$\text{Const.}\, \text{Exp} - \frac{1}{2K_1^2}\Big[\chi^2_{1.(023\cdots p+1)} + \sum_{r=p+1}^{2} \big\{z_{1r.0} - \chi_{r.0} \sum_{u=p+1}^{r} \beta_{1u} b_{ur.(02r-}\big\}^2$$

$$\times \; \{\chi^2_{1.(023\cdots p+1)}\}^{\frac{N-p-2}{2}-1} d\chi^2_{1.(023\cdot p+1)} \prod \{dz_{1r.0}\}$$

$$\text{Exp} - \frac{1}{2K_0^2}\Big[\chi^2_{0.(23\cdots p+1)} + \sum_{r=p+1}^{2} \big\{z_{or} - \chi_r \sum_{u=p+1}^{r} \beta_{0u} b_{ur.(23\cdots r-1)}\big\}^2$$

$$\times \; \{\chi^2_{0.(23\cdots p+1)}\}^{\frac{N-p-1}{2}-1} d(\chi^2_{0.23\cdots p+1}) \prod \{dz_{or}\} \qquad (4.8)$$

where

$$b'_{ur.(02..r-1)} = \sum_{i=1}^{N} \frac{x_{iu}\, x_{ir.(02..r-1)}}{X_{r.0}^2}$$

$$= 1, \text{ if } u = r$$

$$b_{ur.(23..r-1)} = \sum_{i=1}^{N} \frac{x_{iu}\, x_{ir.(23..r-1)}}{x_r^2}$$

$$= 1 \text{ if } u = r$$

The distribution problem is considerably simplified if we make use of the property of invariance of the Wilks statistic for linear transformations of variates within sets. In fact this property is already partly employed when we started the problem with $x_{1.0}$ and x_0 on the dependent side. Without loss of generality, we can further write

$K_1^2 = K_0^2 = 1$; $\beta_{1r} = 0$ for $r = 3, 4, \dots p+1$, (ie) $\beta_{12} \neq 0$

$\beta_{0r} = 0$ for $r = 2, 4, \dots p+1$. or $\beta_{03} \neq 0$

$\alpha^{rs} = 0$ if $r \neq s$

$= 1$ if $r = s$ where $|\alpha_{rs}|$ is the covariance matrix of the independent set $\{x_r\}$; the latter assumption leads to

$b_{ur.(23..r-1)} = 0$ or $z_{ur} = 0$ for $u > r$ $(r = 2, 3, \dots p+1)$

$X_2^2 = N-1$, $X_3^2 = N-2$ etc for the independent set $\{x_r\}$ is considered to be fixed. (4.9)

Noting that the first part (portion containing $\chi^2_{1.(023..p+1)}$, $\{z_{1r.0}\}$) in (4.8) leads to the distribution of R_1^2 for fixed $\{x_{r.0}\}$ which under the first of assumptions in (4.9) can be written as

$$\text{Const. } e^{-\frac{\beta_{12}^2 X_{2.0}^2}{2}} (1-R_1^2)^{\frac{N-p-4}{2}} (R_1^2)^{\frac{p-2}{2}} \, {}_1F_1\left(\frac{N-2}{2}, \frac{p}{2}; \frac{\beta_{12}^2 X_{2.0}^2}{2} R_1^2\right) dR_1^2 \qquad (4.10)$$

The desired distribution is to be thus derived from the product of (4.10) and the second part (portion containing $\chi^2_{0.(23..p+1)}$, $\{z_{0r}\}$ in (4.8) reduced with the use of (4.9). Since the distribution of $(1-W^2)$ is to be derived for fixed values of $\{x_r\}$

and not $\{x_{r.0}\}$, $\chi^2_{2.0}$ coming in (4.10) is to be suitably expressed in terms of the quantities we will be integrating for in course of the derivation.

Writing,

$$\chi^2_{2.0} = \frac{\chi_2^2}{\chi^2_{0.(23..p+1)} + \sum_{r=2}^{p+1} z_{0r}^2}\left\{\chi^2_{0.(23..p+1)} + \sum_{r=3}^{p+1} z_{0r}^2\right\}$$

$$= \chi_2^2\left\{1 - \frac{z_{02}^2}{\chi^2_{0.(23..p+1)} + \sum_{r=2}^{p+1} z_{0r}^2}\right\}$$

and using the second assumption of (4.9) the above referred product (i.e) the distribution of $[R_1^2, \chi^2_{0.(23..p+1)}$ (χ_0^2 for simplicity) and $\{z_{0r}\}]$ for fixed $\{x_r\}$, restricting the variation of z_{03}, z_{02} to $0 \leq z_{03} \leq \infty$, $0 \leq z_{02} \leq \infty$, can be written as

$$\text{const} \cdot \text{Exp} -\tfrac{1}{2}\left\{\beta_{03}^2 \chi_3^2 + \beta_{12}^2 \chi_2^2\right\}$$

$$\times \text{Exp} -\tfrac{1}{2}\left\{\chi_0^2 + \sum_{r=2}^{p+1} z_{0r}^2\right\}(R_1^2)^{\frac{p-2}{2}}(1-R_1^2)^{\frac{N-p-4}{2}}(\chi_0^2)^{\frac{N-p-3}{2}}$$

$$\sum_{l=0}^{\infty}\frac{\Gamma(\frac{N-2}{2}+l)}{\Gamma(\frac{p}{2}+l)\, l!}\left\{\frac{R_1^2 \beta_{12}^2 \chi_2^2}{2}\left(1 - \frac{z_{02}^2}{\chi_0^2 + \sum_{r=2}^{p+1} z_{0r}^2}\right)\right\}^l$$

$$\sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{\beta_{12}^2 \chi_2^2}{2}\,\frac{z_{02}^2}{\chi_0^2 + \sum z_{0r}^2}\right)^m$$

$$\sum_{n=0}^{\infty}\frac{1}{2n!}\left(\beta_{03}^2 \chi_3^2 \; z_{03}^2\right)^n$$

$$d\chi_0^2 \, d R_1^2 \prod_{r=2}^{p+1}\{dz_{0r}\} \qquad \cdots (4.11)$$

Adopting the transformation

$$\begin{aligned}
z_{02} &= H \sin\theta_1 \\
z_{03} &= H \sin\theta_2 \cos\theta_1 \\
&\cdots\cdots \\
z_{0u} &= H \cos\theta_1 \cdots \cos\theta_{u-1} \sin\theta_u \\
&\cdots\cdots \\
z_{0p+1} &= H \cos\theta_1 \cos\theta_2 \cdots \cos\theta_{p-2} \cos\theta_{p-1}
\end{aligned} \qquad (4.12)$$

the jacobian of which is given by

$$\frac{\partial(z_{02}, z_{03}, \cdots z_{0p+1})}{\partial(H, \theta_1, \theta_2, \cdots \theta_{p-1})} = H^{p-1} \cos^{p-2}\theta_1 \cdots \cos\theta_{p-2} \qquad (4.13)$$

(4.11) leads to the joint distribution of

$\{R_1^2, \chi_0^2, H^2, \theta_1, \theta_2, \cdots \theta_{p-1}\}$ and this joint distribution, when all the immediately possible integr (viz) those with respect to $\theta_2, \theta_3, \cdots \theta_{p-1}$ are carried out, leads to the joint distribution of

$$(R_1^2, \chi_0^2, H^2 \;\&\; \theta_1) \quad \begin{cases} 0 \le R_p^2 \le 1, & 0 \le H^2 \le \infty \\ 0 \le \chi^2 \le \infty, & 0 \le \theta_1 \le \frac{\pi}{2} \end{cases}$$

as

$$\text{Const. Exp} -\tfrac{1}{2}\{\beta_{03}^2 \chi_3^2 + \beta_{12}^2 \chi_2^2\}$$
$$\times \text{Exp} -\tfrac{1}{2}(\chi_0^2 + H^2)(R_1^2)^{\frac{p-2}{2}}(1-R_1^2)^{\frac{N-p-4}{2}}(H^2)^{\frac{p-2}{2}}(\chi_0^2)^{\frac{N-p-}{2}} \cos^{p-2}\theta_1$$
$$\sum_{l=0}^{\infty} \frac{\Gamma(\frac{N-2}{2}+l)}{\Gamma(\frac{p}{2}+l)\, l!} \left\{\frac{R_1^2 \beta_{12}^2 \chi_2^2}{2}\left(1 - \frac{H^2 \sin^2\theta_1}{\chi_0^2 + H^2}\right)\right\}^l$$
$$\sum_{m=0}^{\infty} \frac{1}{m!}\left\{\frac{\beta_{12}^2 \chi_2^2}{2} \; \frac{H^2 \sin^2\theta_1}{\chi_0^2 + H^2}\right\}^m$$
$$\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\beta_{03}^2 \chi_3^2}{2}\right)^n \left(\frac{H^2\cos^2\theta_1}{2}\right)^n \frac{1}{\Gamma(\frac{p-1}{2}+n)}$$
$$dR_1^2\, d\chi_0^2\, dH^2\, d\theta_1 \qquad (4\cdot 14)$$

Transforming (H^2, χ_0^2) into (D^2, R_0^2) where

$$\chi_0^2 = D^2(1-R_0^2)$$
$$H^2 = D^2 R_0^2$$

where R_0 is the multiple correlation coefficient of x_0 with $\{x_r\}$, and integrating the transformed form with respect to D^2 $0 \le D^2 \le \infty$, the joint distribution of $(R_1^2, R_0^2 \;\&\; \theta_1)$ is given by

$$\text{const. exp} -\tfrac{1}{2}\{\beta_{03}^2 \chi_3^2 + \beta_{12}^2 \chi_2^2\}$$
$$(R_0^2 R_1^2)^{\frac{p-2}{2}} (1-R_1^2)^{\frac{N-p-4}{2}} (1-R_0^2)^{\frac{N-p-3}{2}} \cos^{p-2}\theta_1$$
$$\sum_{l=0}^{\infty} \frac{\Gamma(\frac{N-2}{2}+l)}{\Gamma(\frac{p}{2}+l)\, l!}\left\{R_1^2 \frac{\beta_{12}^2 \chi_2^2}{2}\left(1 - R_0^2 \sin^2\theta_1\right)^l\right\}$$
$$\sum_{m=0}^{\infty} \frac{1}{m!}\left(\frac{\beta_{12}^2 \chi_2^2}{2}\right)^m (R_0^2 \sin^2\theta_1)^m$$
$$\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\beta_{03}^2 \chi_3^2}{2}\right)^n \frac{\Gamma(\frac{N-1}{2}+n)}{\Gamma(\frac{p-1}{2}+n)} (R_0^2 \cos^2\theta_1)^n$$
$$dR_0^2\, dR_1^2\, d\theta_1 \qquad (4.15)$$

Setting

$$\frac{1}{l!}(1 - R_0^2 \sin^2\theta_1)^l = \sum_{l_1+l_2=l} \frac{(\cos^2\theta_1)^{l_1}(1-R_0^2 \sin^2\theta_1)^{l_2}}{l_1!\; l_2!}$$

as in (3.25), and integrating for θ_1 $(0 \le \theta_1 \le \frac{\pi}{2})$ we get the joint distribution of (R_0^2, R_1^2) as

$$\text{Const. } \operatorname{Exp} -\tfrac{1}{2}\left\{ \beta_{03}^2 \chi_3^2 + \beta_{12}^2 \chi_2^2 \right\}$$

$$(1-R_1^2)^{\frac{N-p-4}{2}} (1-R_0^2)^{\frac{N-p-3}{2}} (R_0^2 - R_1^2)^{\frac{p-2}{2}}$$

$$\sum_{l_1, l_2} \left[\frac{\Gamma(\frac{N-2}{2} + l_1 + l_2)}{\Gamma(\frac{p}{2} + l_1 + l_2)\, l_1!\, l_2!} \left(\frac{R_1^2 \beta_{12}^2 \chi_2^2}{2} \right)^{l_1 + l_2} (1-R_0^2)^{l_2} \right.$$

$$\left. \sum_{m,n} \left\{ \frac{\Gamma(\frac{N-1}{2} + n)}{\Gamma(n + \frac{p-1}{2})} \frac{B(l_1 + n + \frac{p-1}{2}, l_2 + m + \frac{1}{2})}{m!\, n!} \left(\frac{\beta_{12}^2 \chi_2^2}{2} \right)^m \left(\frac{\beta_{03}^2 \chi_3^2}{2} \right)^n (R_0^2)^{m+n} \right\} \right] dR_0^2\, dR_1^2 \qquad (4.16)$$

Transforming R_0^2 and R_1^2 into (u, W^2), where the latter is the Wilks statistic given as

$$1 - W^2 = (1 - R_0^2)(1 - R_1^2)$$

and $u = R_0^2 / W^2$ so that

$$\frac{\partial(R_0^2, R_1^2)}{\partial(W^2, u)} = \frac{-W^2}{1 - uW^2}$$

and the limits of u are 0 and 1 independent of W^2. Then the joint distribution of u & W^2 can ~~thus~~ be immediately derived from (4.16) as

$$\text{const. } \operatorname{Exp} -\tfrac{1}{2}\left\{ \beta_{03}^2 \chi_3^2 + \beta_{12}^2 \chi_2^2 \right\}$$

$$(W^2)^{p-1} (1-W^2)^{\frac{N-p-4}{2}} \{u(1-u)\}^{\frac{p-2}{2}} (1-uW^2)^{-(\frac{p}{2}-1)}$$

$$\times \sum_{l_1, l_2} \left[\frac{\Gamma(\frac{N-2}{2} + l_1 + l_2)}{\Gamma(\frac{p}{2} + l_1 + l_2)\, l_1!\, l_2!} \left(\frac{\beta_{12}^2 \chi_2^2 W^2}{2} \right)^{l_1 + l_2} (1-u)^{l_1 + l_2} (1-uW^2)^{-l_1} \right.$$

$$\left. \sum_{m,n} \left\{ \frac{\Gamma(\frac{N-1}{2} + n)}{\Gamma(\frac{p-1}{2} + n)} \frac{B(l_1 + n + \frac{p-1}{2}, l_2 + m + \frac{1}{2})}{m!\, n!} \left(\frac{\beta_{12}^2 \chi_2^2}{2} \right)^m \left(\frac{\beta_{03}^2 \chi_3^2}{2} \right)^n (W^2 u)^{m+n} \right\} \right] du\, dW^2 \qquad (4.17)$$

Integrating (4.17) with respect u by using the well known result

$$B(\beta, \gamma - \beta)\, F(\alpha, \beta; \gamma, x) = \int_0^1 t^{\beta-1} (1-t)^{\gamma-\beta-1} (1-xt)^{-\alpha}\, dt \quad \text{for } |x| \le 1$$

we get to the distribution of W^2 as given by

$$\text{Const. } \operatorname{Exp} -\tfrac{1}{2}\left(\beta_{03}^2 \chi_3^2 + \beta_{12}^2 \chi_2^2\right)(W^2)^{p-1}(1-W^2)^{\frac{N-p-4}{2}}$$

$$\sum_{l_1, l_2, m, n} \left\{ \left(\frac{\beta_{12}^2 \chi_2^2}{2}\right)^{l_1+l_2+m} \left(\frac{\beta_{03}^2 \chi_3^2}{2}\right)^{n} \frac{(W^2)^{l_1+l_2+m+n}}{l_1!\, l_2!\, m!\, n!} \right.$$

$$\frac{\Gamma\left(\frac{N-2}{2}+l_1+l_2\right)\Gamma\left(\frac{N-1}{2}+n\right)\Gamma\left(\frac{p}{2}+m+n\right)}{\Gamma\left(n+\frac{p-1}{2}\right)\Gamma\left(p+l_1+l_2+m+n\right)} B\left(l_1+n+\frac{p-1}{2}, l_2+m+\frac{1}{2}\right)$$

$$\left. F\left(\frac{p-1}{2}+l_1, \frac{p}{2}+m+n; p+l_1+l_2+m+n, W^2\right)\right\} dW^2 \qquad (4.18)$$

As in the previous section, in view of (4.9) it can be easily seen that all the variates $x_0, x_{1.0}, \{x_r\}$ correspond to Hotellings canonical variates and that β_{12}^2 and β_{03}^2 are the two non-vanishing of the two canonical correlations ρ_1^2 and ρ_0^2. Then writing

$(N-1)\rho_0^2 = E^2$ and $(N-1)\rho_1^2 = F^2$, the distribution can be written as

$$C. \operatorname{Exp} -\tfrac{1}{2}\left\{E^2+F^2\right\}(W^2)^{p-1}(1-W^2)^{\frac{N-p-4}{2}}$$

$$\sum_{l_1, l_2, m, n} \left\{ \left(\frac{E^2}{2}\right)^{n} \left(\frac{F^2}{2}\right)^{l_1+l_2+m} \frac{(W^2)^{m+n+l_1+l_2}}{l_1!\, l_2!\, m!\, n!} \right.$$

$$\frac{\Gamma\left(\frac{N-2}{2}+l_1+l_2\right)\Gamma\left(\frac{N-1}{2}+n\right)\Gamma\left(\frac{p}{2}+m+n\right)}{\Gamma\left(n+\frac{p-1}{2}\right)\Gamma\left(p+l_1+l_2+m+n\right)} B\left(l_1+n+\frac{p-1}{2}, l_2+m+\frac{1}{2}\right)$$

$$\left. F\left(\frac{p-1}{2}+l_1, \frac{p}{2}+m+n; p+l_1+l_2+m+n, W^2\right)\right\} dW^2 \qquad (4.19)$$

Hereonwards proceeding in a way exactly similar to that in the previous section, either while determining the constant C or while deducing and reducing the null distribution, we arrive at the nonnull and null distributions of $1-W^2 = L$, the expressions for which are given by

$$\frac{2^{N-p-3}}{\pi\,\Gamma(N-p-2)}\, e^{-\frac{1}{2}(E^2+F^2)}\,(1-L)^{p-1}\, L^{\frac{N-p-4}{2}}$$

$$\sum_{l_1,l_2,m,n}\left\{\left(\frac{E^2}{2}\right)^{n}\left(\frac{F^2}{2}\right)^{m+l_1+l_2}\frac{(1-L)^{m+n+l_1+l_2}}{m!\,n!\,l_1!\,l_2!}\right.$$

$$\frac{\Gamma\left(\frac{N-2}{2}+l_1+l_2\right)\Gamma\left(\frac{N-1}{2}+n\right)\Gamma\left(\frac{p}{2}+m+n\right)}{\Gamma\left(n+\frac{p-1}{2}\right)\Gamma\left(p+l_1+l_2+m+n\right)}\, B\left(\frac{p-1}{2}+n+l_1,\ l_2+m+\frac{1}{2}\right)$$

$$\left.F\left(\frac{p-1}{2}+l_1,\ \frac{p}{2}+m+n,\ p+l_1+l_2+m+n,\ 1-L\right)\Big]\right\}\, dL \qquad (4.20)$$

and

$$\frac{1}{2B(N-p-2,p)}\,(1-\sqrt{L})^{p-1}\,(\sqrt{L})^{N-p-4}\, dL \qquad (4.21)$$

respectively.

If we put $E = F$ it is seen that the nonnull distributions in both the cases, as given by (4.20) and (3.32) are of the same form. Thus there exists duality when both the nonvanishing correlations are equal and not in the general case.

5. SUMMARY:- The nonnull distributions of Hotelling's T^2 and Mahalanobis D^2 statistics are obtained without assuming the invariance property through a completely analytic approach. The nonnul and null distributions of the Wilk's statistic with two variates in the dependent set and p variates in the independent set, when either of the sets is fixed, are derived. It is interesting indeed to find that there is no duality between the two general nonnul distributions as in the case of only one variate in the dependent set. However, it is shown that there exists duality when the two nonvanishing correlations are equal.

My thanks are due to Dr. R.D. Narain for his help and suggestions during the preparation of this thesis.

REFERENCES

1. Hotelling, H.(1931) ... The generalisation of Student's ratio. Ann.Math.Stat.Vol.2.

2. Bose R.C. & Roy S.N.(1938) .. Distribution of the Studentised D2 Statistics.Sankhya Vol.4.

3. Hsu P.L. & 1938) .. Notes on Hotellings Generalised T. Ann.Math.Stat.Vol.9.

4. Fisher R.A.(1928) .. The general sampling distribution of the multiple correlation coefficient. Proc.Roy.Soc.A.121.

5. Fisher R.A.(1938) .. The statistical utilization of multiple measurements. Ann. Eugenics Vol.VIII.

6. Narain R.D. (1948). .. A new approach to the sampling distributions of the multivariate normal theory (I) Jour. Ind. Soc.Agr.Stat. Vol.I Part 1.

7. Narain R.D. (1948). .. A new approach to the sampling distribution of the multivariate normal theory. (II) Jour.Ind.Soc.Agr.Stat. Vol.I Part 2.

8. Rao C.R.(1946). .. Discriminant functions in Multivariate analysis. Sankhya Vol.7.

9. Anderson, T.W. (1946). .. Non-Central Wishart Distribution and certain problems of Multivariate statistics. Ann.Math.Stat.Vol.17.

10. Hsu P.L.(1939) .. On the roots of a certain determinent equation. Ann.Eugenics. Vol.9.

11. Fisher. R.A.(1939). .. The sampling distribution of some statistics obtained from non-linear equations. Ann.Eugenics. Vol.9.

12. Roy S.N.(1939) .. p.statistics of some generalisations in analysis of variance appropriate to multivariate problems. Sankhya Vol.4,

13. Roy S.N. (1945). .. The individual sampling distribution of the maximum and minimum and any intermediate of the p-statistics on the null hypothesis. Sankhya Vol.7.

14. Nanda, D.N.(1947). .. Distribution of a root of a determinental equation. Ann.Math. Stat. Vol.18.

15. Bartlett M.S.(1941). .. The statistical significance of canonical correlations. Biometrika Vol.32.

16. Bartlett M.S. (1947). .. The general canonical correlation distribution. Ann.Math.Stat.Vol.18.

17. Roy S.N.(1942). .. Analysis of variance for multivariate normal populations. The sampling distribution of the requisite p-statistics on the null and non-nul hypotheses. Sankhya Vol.6.

18-------------

18. Roy S.N.(1946). .. Multivariate analysis of variance. The sampling distribution of the numerically largest of the p-statistics on the non null ~~by~~ hypothesis. Sankhya Vol. 8.

19. Wilks S.S.(1932). .. Certain generalisations in the analysis of variance. Biometrika Vol.24.

20. Bartlett M.S.(1934). .. The vector representation of a sample. Proc.Camb.Phil.Soc.Vol.30.

21. Hotelling.H.(1936). .. Relation between two sets of variates Biometrika Vol.28.

22. Wald.A & Brookner R.J. (1941). .. On the distribution of Wilks' statistic etc. Ann.Math.Stat.Vol.12

23. Hsu P.L.(1941). .. Canonical reduction of the general regression problem. Ann.Eugenics Vol. 11.
