THE NON-CENTRAL MULTIVARIATE BETA DISTRIBUTION

By A. M. Kshirsagar[1]

Manchester University

1. Introduction. Let $\mathbf{A}$ and $\mathbf{B}$ be two symmetric matrices of order p, having independent Wishart distributions with degrees of freedom f_1 and f_2 respectively. The density function of $\mathbf{A}$ is

$$W(\mathbf{A} \mid \mathbf{\Sigma} \mid f_1) = \frac{\mid \mathbf{A} \mid^{(f_1-p-1)/2} e^{-\frac{1}{2}\mathrm{tr}\mathbf{A}\mathbf{\Sigma}^{-1}}}{2^{f_1p/2} \pi^{p(p-1)/4} \mid \mathbf{\Sigma} \mid^{f_1/2} \prod_{i=1}^{p} \Gamma[\frac{1}{2}(f_1 + 1 - i)]} \tag{1.1}$$

for $\mathbf{A}$ positive definite and 0 otherwise. The density function of $\mathbf{B}$ is $W(\mathbf{B} \mid \mathbf{\Sigma} \mid f_2)$. There exists a lower triangular matrix $\mathbf{C}$ such that

$$\mathbf{A} + \mathbf{B} = \mathbf{C}\mathbf{C}'. \tag{1.2}$$

Let $\mathbf{L}$ be defined by

$$\mathbf{A} = \mathbf{C}\mathbf{L}\mathbf{C}'. \tag{1.3}$$

Then it has been proved (Hsu [8], Anderson [3], Khatri [6]) that $\mathbf{L}$ and $\mathbf{C}$ are independently distributed, the density function of $\mathbf{L}$ being

$$\pi^{-p(p-1)/4} \prod_{i=1}^{p} \Gamma[\tfrac{1}{2}(f_1 + f_2 + 1 - i)]/\{\Gamma[\tfrac{1}{2}(f_1 + 1 - i)]\Gamma[\tfrac{1}{2}(f_2 + 1 - i)]\} \\ |\mathbf{L}|^{(f_1-p-1)/2} \mid \mathbf{I} - \mathbf{L} \mid^{(f_2-p-1)/2} \tag{1.4}$$

for both $\mathbf{L}$ and $\mathbf{I} - \mathbf{L}$ positive definite and 0 otherwise. This distribution may be called the multivariate beta distribution, on account of its similarity in form with the univariate beta distribution. Since the distribution of $\mathbf{L}$ does not involve $\mathbf{\Sigma}$, there is no loss of generality in assuming it to be $\mathbf{I}$. When $\mathbf{B}$ has a non-central Wishart distribution, the corresponding distribution of $\mathbf{L}$ can be called non-central multivariate beta distribution. In this paper, this distribution is derived in the special case when the non-central Wishart distribution of $\mathbf{B}$ belongs to the "linear case" (Anderson [1], [2]). In the last section of this paper, it is shown how the distribution of $\mathbf{L}$ becomes untractable in the "planar case". It is well-known that Wilks's $\mathbf{\Lambda}$ criterion, in the null-case, is distributed as $t_{11}^2 t_{22}^2 \cdots t_{pp}^2$, where the t_{ii}^2 are independent beta variables. In this paper, $\mathbf{\Lambda}$ is expressed explicitly in terms of certain multiple correlation coefficients and hence, by also using the non-central multivariate beta distribution, the distribution of t_{ii}^2 in the non-null but linear case is obtained. The same method can be used in the planar case but the stubborn nature of the non-central Wishart distribution (planar case) is an obstacle. The conditional distribution of the canonical corre-

---

Received March 21, 1960; revised August 29, 1960.

[1] Present address: University College, London.

lations, first obtained by Williams [10] can be easily derived from the non-central multivariate beta distribution (linear case). This is done in Section 4, while Section 5 deals with the role of $|\mathbf{L}|$ in discriminant analysis.

2. Non-central multivariate beta distribution (linear case). Let us assume that the density function of $\mathbf{B}$ is

$$W(\mathbf{B} \mid \mathbf{I} \mid f_2) \times e^{-\lambda^2/2} \sum_{r=0}^{\infty} \frac{(\lambda^2/2)^r}{r!} \cdot \frac{b_{11}^r \Gamma(f_2/2)}{2^r \Gamma[\frac{1}{2} f_2 + r]}. \tag{2.1}$$

This is the density function of a non-central Wishart distribution of f_2 degrees of freedom, corresponding to what Anderson [1], [2] calls the "linear case". By using the transformation from $\mathbf{A}$ and $\mathbf{B}$ to $\mathbf{L}$ and $\mathbf{C}$ given by (1.2) and (1.3), it can be shown that the density function of $\mathbf{L} = [l_{ij}]$ and $\mathbf{C} = [c_{ij}]$ is

$$\begin{aligned} &\text{const} \prod_{i=1}^{p} c_{ii}^{f_1+f_2-i} e^{-\frac{1}{2}\operatorname{tr}\mathbf{cc}'} \mid \mathbf{L} \mid^{(f_1-p-1)/2} \mid \mathbf{I} - \mathbf{L} \mid^{(f_2-p-1)/2} \\ &\qquad \times e^{-\lambda^2/2} \sum_{r=0}^{\infty} \frac{(\lambda^2/2)^r}{r!} \frac{c_{11}^{2r}(1 - l_{11})^r \Gamma(\frac{1}{2} f_2)}{2^r \Gamma(\frac{1}{2} f_2 + r)} \end{aligned} \tag{2.2}$$

Integrating out c_{11}, the density function of $\mathbf{L}$ comes out as

$$\begin{aligned} &\pi^{-p(p-1)/4} \prod_{i=1}^{p} \Gamma[\tfrac{1}{2}(f_1 + f_2 + 1 - i)] / \{\Gamma[\tfrac{1}{2}(f_1 + 1 - i)\Gamma[\tfrac{1}{2}(f_2 + 1 - i)]\} \\ &\qquad \times e^{-\lambda^2/2} {}_1F_1(\tfrac{1}{2}(f_1 + f_2), \tfrac{1}{2} f_2, \tfrac{1}{2}\lambda^2(1 - l_{11})) \\ &\qquad \times |\mathbf{L}|^{(f_1-p-1)/2} |\mathbf{I} - \mathbf{L}|^{(f_2-p-1)/2}. \end{aligned} \tag{2.3}$$

3. Application to Wilks's Λ criterion. Let $\mathbf{x}' = (x_1, \cdots, x_p)$ and $\mathbf{y}' = (y_1, \cdots, y_q)$ be $p + q$ variables ($p \leqq q$) having a multivariate normal distribution with zero means and variance-covariance matrix

$$E \begin{bmatrix} \mathbf{x} \\ \hline \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \hline \mathbf{y} \end{bmatrix}' = \left[\begin{array}{c|c} \mathbf{I}_p & \mathbf{P} \\ \hline \mathbf{P}' & \mathbf{I}_q \end{array}\right] \tag{3.1}$$

where $\mathbf{P}$ is a diagonal matrix of order $p \times q$, the diagonal elements being ρ_1, $\rho_2, \cdots, \rho_p, 0, \cdots, 0$. The ρ's are population canonical correlations of $\mathbf{x}$ with $\mathbf{y}$. Let there be n independent observations x_{it} ($i = 1, \cdots, p$; $t = 1, \cdots, n$), y_{jt} ($j = 1, \cdots, q$; $t = 1, \cdots, n$) forming the matrix

$$\begin{bmatrix} \mathbf{X} \\ \hline \mathbf{Y} \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \hdashline x_{p1} & \cdots & x_{pn} \\ \hline y_{11} & \cdots & y_{1n} \\ \hdashline y_{q1} & \cdots & y_{qn} \end{bmatrix} \tag{3.2}$$

and let

$$\mathbf{S} = \begin{bmatrix} \mathbf{X} \\ \hline \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \hline \mathbf{Y} \end{bmatrix}' = \begin{array}{c} \left[\begin{array}{c|c} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \hline \mathbf{S}_{21} & \mathbf{S}_{22} \end{array}\right] \begin{array}{l} p \\ q \end{array} \\ \begin{array}{cc} p & q \end{array} \end{array}. \tag{3.3}$$

Let

$$\mathbf{A} = \mathbf{S}_{11} - \mathbf{S}_{12}\,\mathbf{S}_{22}^{-1}\,\mathbf{S}_{21} \tag{3.4}$$

$$\mathbf{B} = \mathbf{S}_{12}\,\mathbf{S}_{22}^{-1}\,\mathbf{S}_{21}\,. \tag{3.5}$$

(a) *Null case*: If it is assumed that all the ρ's are zero, it is well-known (Anderson [3]) that **A** and **B** have independent Wishart distributions with $f_1 = n - q$ and $f_2 = q$ degrees of freedom respectively. If **L** is defined as in (1.3), its density function will be given by (1.4). It can be seen that Wilks's Λ criterion for testing the independence of **x** and **y** is

$$\Lambda = |\mathbf{A}| \;/\; |\mathbf{A} + \mathbf{B}| = |\mathbf{L}|. \tag{3.6}$$

If we make the transformation

$$\mathbf{L} = \mathbf{T}\mathbf{T}', \tag{3.7}$$

where **T** is a lower triangular matrix $[t_{ij}]$, in the distribution of **L**, the density function of **T** comes out as

$$\text{Const.} \prod_{i=1}^{p} t_{ii}^{f_1 - i} |\mathbf{I} - \mathbf{T}\mathbf{T}'|^{(f_2 - p - 1)/2}. \tag{3.8}$$

Integrating out the non-diagonal elements of **T**, it can be observed that the diagonal elements t_{ii} are independently distributed, the density function of t_{ii}^2 being

$$\text{Const.}\ (t_{ii}^2)^{\frac{1}{2}(f_1 + 1 - i) - 1}(1 - t_{ii}^2)^{\frac{1}{2}f_2 - 1}. \tag{3.9}$$

Since

$$\Lambda = |\mathbf{L}| = \prod_{i=1}^{p} t_{ii}^2$$

it follows that Λ is distributed as the product of p independent beta variables t_{ii}^2. This result is usually established in the literature ([3], [4], [9]) by identifying the moments of Λ with those of a variable distributed as the product of p independent beta variables like (3.9). That method, however, does not explicitly bring out the relation between Λ and the beta variables. These t_{ii}^2 have another interpretation. If $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{C}_i$, $\mathbf{T}_i$ and $\mathbf{L}_i$ are matrices obtained from **A, B, C, T** and **L** respectively by considering the first i rows and columns only, it can be seen from (1.2), (1.3) and (3.7) that

$$\mathbf{A}_i = \mathbf{C}_i\mathbf{T}_i\mathbf{T}_i'\mathbf{C}_i' = \mathbf{C}_i\mathbf{L}_i\mathbf{C}_i', \tag{3.10}$$

$$\mathbf{A}_i + \mathbf{B}_i = \mathbf{C}_i\mathbf{C}_i'\,, \tag{3.11}$$

because **C** and **T** are lower triangular matrices. Hence

$$t_{ii}^2 = \frac{|\mathbf{L}_i|}{|\mathbf{L}_{i-1}|} = \frac{|\mathbf{A}_i|/|\mathbf{A}_{i-1}|}{|\mathbf{A}_i + \mathbf{B}_i|/|\mathbf{A}_{i-1} + \mathbf{B}_{i-1}|} \tag{3.12}$$

for $i = 1, 2, \cdots, p$, with the definitions $|\mathbf{A}_0| = 1$, $|\mathbf{A}_0 + \mathbf{B}_0| = 1$. By a direct substitution in the formula for the sample multiple correlation coefficient, it can be proved with a little algebra, that

$$|\mathbf{A}_i|/|\mathbf{A}_{i-1}| = \left(\sum_{t=1}^{n} x_{it}^2\right)\left(1 - R^2_{x_i(x_1,\cdots x_{i-1},y_1,\cdots y_q)}\right) \tag{3.13}$$

and that

$$\frac{|\mathbf{A}_i + \mathbf{B}_i|}{|\mathbf{A}_{i-1} + \mathbf{B}_{i-1}|} = \left(\sum_{t=1}^{n} x_{it}^2\right)\left(1 - R^2_{x_i(x_1,\cdots x_{i-1})}\right) \tag{3.14}$$

where $R_{x_i(x_1,\cdots,x_{i-1},y_1,\cdots,y_q)}$ denotes the multiple correlation coefficient between x_i and $x_1, x_2, \cdots, x_{i-1}, y_1, \cdots, y_q$. Therefore,

$$t_{ii}^2 = \left(1 - R^2_{x_i(x_1,\cdots x_{i-1},y_1,\cdots y_q)}\right) / \left(1 - R^2_{x_i(x_1,\cdots x_{i-1})}\right). \tag{3.15}$$

This, incidentally, proves that the density function of t_{ii}^2 is given by (3.9), if the regression coefficients attached to $y_1, y_2, \cdots, y_q$ in the regression of x_i on $x_1, \cdots, x_{i-1}, y_1, \cdots, y_q$ are all zero. If $\mathbf{x}$ and $\mathbf{y}$ are independent, this is true for every $i = 1, 2, \cdots, p$.

(b) *Linear case*: In the linear case, all ρ's except ρ_1 are zero. Let

$$\lambda^2 = \rho_1^2 \sum_{t=1}^{n} y_{1t}^2 / (1 - \rho_1^2). \tag{3.16}$$

Then the density function of $\mathbf{B}$ defined by (3.5) is given by (2.1) with $f_2 = q$. This refers to the conditional distribution of $\mathbf{B}$ when $\mathbf{y}$ is fixed. Proceeding exactly as in the null case and starting from (2.2), it can be shown that $\Lambda = |\mathbf{L}|$ is distributed as $\prod_{i=1}^{p} t_{ii}^2$ where the density function of t_{ii}^2 $(i = 2, \cdots, p)$ is given by (3.9) but that of t_{11}^2 is

$$\text{Const. } (t_{11}^2)^{\frac{1}{2}f_1 - 1}(1 - t_{11}^2)^{\frac{1}{2}f_2 - 1} e^{-\lambda^2/2} \, {}_1F_1\left(\tfrac{1}{2}(f_1 + f_2), \tfrac{1}{2}f_2, \tfrac{1}{2}\lambda^2(1 - t_{11}^2)\right). \tag{3.17}$$

If, however, $\mathbf{y}$ is not fixed, then $\sum_{t=1}^{n} y_{1t}^2$ in λ^2 of (3.16) is a chi-square with n degrees of freedom $(n = f_1 + f_2)$, and therefore the density function of t_{11}^2 will be

$$\begin{aligned} &\frac{(1 - \rho_1^2)^{\frac{1}{2}(f_1+f_2)}}{B(\frac{1}{2}f_1, \frac{1}{2}f_2)} (t_{11}^2)^{\frac{1}{2}f_1 - 1}(1 - t_{11}^2)^{\frac{1}{2}f_2 - 1} \\ &\qquad {}_2F_1\left(\tfrac{1}{2}(f_1 + f_2), \tfrac{1}{2}(f_1 + f_2), \tfrac{1}{2}f_2, \rho_1^2(1 - t_{11}^2)\right). \end{aligned} \tag{3.18}$$

4. Application in the case of canonical correlations. We start with $\mathbf{x}$ and $\mathbf{y}$ as in Section 3. It is, however, assumed that $p = 2$ and ρ_1 is the only non-zero canonical correlation. The density function of

$$\mathbf{L} = \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}$$

will, therefore, be

$$\begin{aligned} &\Gamma(\tfrac{1}{2}n)\Gamma[\tfrac{1}{2}(n - 1)]|\mathbf{L}|^{(n-q-3)/2}|\mathbf{I} - \mathbf{L}|^{(q-3)/2} e^{-\lambda^2/2} {}_1F_1(\tfrac{1}{2}n, \tfrac{1}{2}q, \tfrac{1}{2}\lambda^2(1 - l_{11})) \\ &\qquad \times \pi^{-\frac{1}{2}}\{\Gamma[\tfrac{1}{2}(n - q)]\Gamma[\tfrac{1}{2}(n - q - 1)]\Gamma(\tfrac{1}{2}q)\Gamma[\tfrac{1}{2}(q - 1)]\}^{-1}. \end{aligned} \tag{4.1}$$

Since

$$|\mathbf{B} - \theta(\mathbf{A} + \mathbf{B})| = |\mathbf{C}(\mathbf{I} - \mathbf{L})\mathbf{C}' - \theta\mathbf{C}\mathbf{C}'|$$
$$= |\mathbf{C}|\,|(\mathbf{I} - \mathbf{L}) - \theta\mathbf{I}|\,|\mathbf{C}'|$$

it is evident that the latent roots r_1^2 and r_2^2 of $\mathbf{I} - \mathbf{L}$ are the squares of the canonical correlations of $\mathbf{x}$ and $\mathbf{y}$. Also from (3.4), (3.5) and (1.3),

$$l_{11} = 1 - R^2_{x_1(y_1,\cdots,y_q)}\,, \tag{4.2}$$

which we shall denote simply by $1 - R^2$. Then

$$\begin{aligned} r_1^2 + r_2^2 &= (1 - l_{11}) + (1 - l_{22}), \\ r_1^2 r_2^2 &= (1 - l_{11})(1 - l_{22}) - l_{12}^2. \end{aligned} \tag{4.3}$$

Transforming from $\mathbf{L}$ to r_1^2, r_2^2 and R^2 and noting that the jacobian of the transformation is $(r_1^2 - r_2^2)[(r_1^2 - R^2)(R^2 - r_2^2)]^{-\frac{1}{2}}$ it can be seen that the density function of r_1^2, r_2^2 and R^2 is

$$\begin{aligned} &\text{Const. } (r_1^2 r_2^2)^{(q-3)/2}[(1 - r_1^2)(1 - r_2^2)]^{(n-q-3)/2}(r_1^2 - r_2^2) \\ &\qquad [(r_1^2 - R^2)(R^2 - r_2^2)]^{-\frac{1}{2}} e^{-\lambda^2/2} {}_1F_1(n/2, q/2, \lambda^2 R^2/2). \end{aligned} \tag{4.4}$$

From this, the density function of the conditional distribution of r_1^2 and r_2^2 when R^2 is fixed is

$$\frac{\text{Const. } (r_1^2\, r_2^2)^{(q-3)/2}[(1 - r_1^2)(1 - r_2^2)]^{(n-q-3)/2}(r_1^2 - r_2^2)}{[(r_1^2 - R^2)(R^2 - r_2^2)]^{\frac{1}{2}}(R^2)^{q/2-1}(1 - R^2)^{(n-q)/2-1}}\,. \tag{4.5}$$

This distribution was first derived by Williams [10] and, since it is independent of λ, was used by him for testing the adequacy of a hypothetical discriminant function.

5. Role of $|\mathbf{L}|$ in discriminant analysis. Let there be three vector variables,

$$\mathbf{x}'(1) = (x_1, \cdots, x_k), \qquad \mathbf{x}'(2) = (x_{k+1}, \cdots, x_p), \qquad \mathbf{y}' = (y_1, \cdots, y_q) \tag{5.1}$$

and let

$$\mathbf{x}' = [\mathbf{x}'(1) \mid \mathbf{x}'(2)]. \tag{5.2}$$

Let the means of these variables be zero and let the variance-covariance matrix be

$$E\begin{bmatrix} \mathbf{x}(1) \\ \hline \mathbf{x}(2) \\ \hline \mathbf{y} \end{bmatrix}\begin{bmatrix} \mathbf{x}(1) \\ \hline \mathbf{x}(2) \\ \hline \mathbf{y} \end{bmatrix}' = \begin{array}{c} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{23} \\ \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{32} & \boldsymbol{\Sigma}_{33} \end{bmatrix} \begin{matrix} k \\ p-k \\ q \end{matrix} \\ \begin{matrix} k & p-k & q \end{matrix} \end{array} \tag{5.3}$$

Let there be n independent observations on each of these variables, forming the matrix

$$\begin{array}{c} \begin{bmatrix} \mathbf{X}(1) \\ \hline \mathbf{X}(2) \\ \hline \mathbf{Y} \end{bmatrix} \begin{matrix} k \\ p-k \\ q \end{matrix} \\ n \end{array} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \hline \cdots & & \cdots \\ \hline y_{q1} & \cdots & y_{qn} \end{bmatrix} \tag{5.4}$$

and let

$$\begin{bmatrix} \mathbf{X}(1) \\ \hdashline \mathbf{X}(2) \\ \hdashline \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{X}(1) \\ \hdashline \mathbf{X}(2) \\ \hdashline \mathbf{Y} \end{bmatrix}' = \begin{array}{c|c} \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33} \end{bmatrix} & \begin{array}{l} k \\ p-k \\ q \end{array} \\ \begin{array}{ccc} k & p-k & q \end{array} & \end{array} . \tag{5.5}$$

It can be proved ([3]) that, if the regression coefficients attached to $\mathbf{x}(2)$ in the regression of $\mathbf{y}$ on $\mathbf{x}$ are all zero, then

$$\mathbf{A} = \mathbf{S}_{33} - [\mathbf{S}_{31}\ \mathbf{S}_{32}] \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{S}_{13} \\ \mathbf{S}_{23} \end{bmatrix} \tag{5.6}$$

and

$$\mathbf{B} = [\mathbf{S}_{31}\ \mathbf{S}_{32}] \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{S}_{13} \\ \mathbf{S}_{23} \end{bmatrix} - \mathbf{S}_{31} \mathbf{S}_{11}^{-1} \mathbf{S}_{13} \tag{5.7}$$

have independent Wishart distributions with degrees of freedom

$$f_1 = n - p, \qquad f_2 = p - k \tag{5.8}$$

respectively. Then $\mathbf{L}$ defined by (1.3) has the density function (1.4). There is a close relationship between the problem of discrimination and that of regression. In problems of discrimination between several groups, $\mathbf{y}$ is only implicit and the observations on $\mathbf{y}$ are not random. The canonical variables of the $\mathbf{x}$ set are the discriminant functions and the number of non-zero canonical correlations, in the population, is the number of discriminant functions adequate for discrimination. It can be verified easily that, if the regression coefficients attached to $\mathbf{x}(2)$ in the regression of $\mathbf{y}$ on $\mathbf{x}$ are all zero, the characters $x_{k+1}, \cdots, x_p$ do not further bring out the differences between the populations, when the differences due to $x_1, \cdots, x_k$ are eliminated (Rao [9]). The criterion employed for testing this is $|\mathbf{L}|$; it can be easily seen that it is the ratio of

$$\begin{vmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33} \end{vmatrix} \Bigg/ \begin{vmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{vmatrix} \tag{5.9}$$

and

$$\begin{vmatrix} \mathbf{S}_{11} & \mathbf{S}_{13} \\ \mathbf{S}_{31} & \mathbf{S}_{33} \end{vmatrix} \Bigg/ |\mathbf{S}_{11}| \cdot |\mathbf{S}_{33}| \tag{5.10}$$

and hence it can also be written as

$$\prod_{i=1}^{p} (1 - r_i^2) / \prod_{i=1}^{k} (1 - r_i'^2), \tag{5.11}$$

where r_i ($i = 1, \cdots, p$) are the canonical correlations of $\mathbf{x}$ and $\mathbf{y}$ and r_i' ($i = 1, \cdots, k$) are the canonical correlations of $\mathbf{x}(1)$ with $\mathbf{y}$.

When we are testing the adequacy of a single hypothetical discriminant function, $k = 1$ and x_1 is the assigned discriminant function, and, under the hypothesis that the assigned discriminant function is adequate for discrimination, L will have the density function (1.4), with f_1 and f_2 given by (5.8). The criterion employed in practice, however, is not $\mathbf{L}$ but $|\mathbf{L}|$, and it reduces to

$$\prod_{i=1}^{p} (1 - r_i^2) \,/\, (1 - R^2) \tag{5.12}$$

where R is the multiple correlation of the assigned discriminant function with $\mathbf{y}$. The refinement of using $(1 - r_1^2)/(1 - R^2)$ instead of $|\mathbf{L}|$ and thus removing, at least approximately, the portion corresponding to "non-collinearity", and some other aspects of this problem have been considered by Bartlett [5] and Williams [10], [11]. In particular, for $q = 1$, i.e., when we are discriminating between only two groups, $\mathbf{L}$ reduces to the scalar quantity

$$[1 - R^2_{y(x_1,\cdots,x_p)}]/(1 - R^2) \tag{5.13}$$

and we get the usual test for an assigned discriminant function, due to Fisher [7].

6. Non-central multivariate Beta distribution (planar case). Let $\mathbf{A}$ and $\mathbf{B}$ be, as in Section 1, two symmetric $p \times p$ matrices independently distributed, the density functions of their distributions being $W(\mathbf{A} \mid \mathbf{I} \mid f_1)$ and

$$\begin{aligned} &W(\mathbf{B} \mid \mathbf{I} \mid f_2) \times e^{-(\lambda_1^2+\lambda_2^2)/2}\,\Gamma(f_2/2)\Gamma[\tfrac{1}{2}(f_2 - 1)] \\ &\qquad \times \sum_{\alpha,\beta=0}^{\infty} \frac{(\lambda_1^2\lambda_2^2)^{\alpha}(b_{11}\, b_{22} - b_{12}^2)^{\alpha}(\lambda_1^2 b_{11} + \lambda_2^2 b_{22})^{\beta}}{2^{4\alpha+2\beta}\alpha!\,\beta!\,\Gamma[\tfrac{1}{2}(f_2 - 1) + \alpha]\Gamma[\tfrac{1}{2}f_2 + 2\alpha + \beta]}. \end{aligned} \tag{6.1}$$

Thus $\mathbf{B}$ has the non-central Wishart distribution, belonging to what Anderson [1], [2] calls the planar case. Transforming to $\mathbf{L}$ and $\mathbf{C}$ by (1.2) and (1.3), and noting that the Jacobian is

$$2^p \prod_{i=1}^{p} c_{ii}^{2(p+1)-i} \tag{6.2}$$

the density function of $\mathbf{L}$ and $\mathbf{C}$ comes out as

$$\begin{aligned} &\frac{e^{-(\lambda_1^2+\lambda_2^2)/2} \mid \mathbf{L} \mid^{(f_1-p-1)/2} \mid \mathbf{I} - \mathbf{L} \mid^{(f_2-p-1)/2} e^{-\frac{1}{2}\mathrm{tr}\,\mathbf{c}\mathbf{c}'} \prod_{i=1}^{p} c_{ii}^{f_1+f_2-i}}{2^{(f_1+f_2)p/2}\, 2^{-p}\, \pi^{p(p-1)/2}\{\Gamma[\tfrac{1}{2}f_2]\Gamma[\tfrac{1}{2}(f_2 - 1)]\}^{-1} \cdot \prod_{i=1}^{p} \{\Gamma[\tfrac{1}{2}(f_1 + 1 - i)]\Gamma[\tfrac{1}{2}(f_2 + 1 - i)]\}} \\ &\quad \times \sum_{\alpha,\beta_1,\beta_2=0}^{\infty} \frac{\lambda_1^{2\alpha+2\beta_1}\lambda_2^{2\alpha+2\beta_2}\, c_{11}^{2\alpha+2\beta_1}\, c_{22}^{2\alpha}\, [(1 - l_{11})(1 - l_{22}) - l_{12}^2]^{\alpha}(1 - l_{11})^{\beta_1}}{2^{4\alpha+2\beta_1+2\beta_2}\,\alpha!\,\beta_1!\,\beta_2!\,\Gamma[\tfrac{1}{2}(f_2 - 1) + \alpha]\Gamma[\tfrac{1}{2}f_2 + 2\alpha + \beta_1 + \beta_2]} \\ &\quad \times [c_{21}^2\,(1 - l_{11}) - 2\, c_{21}\, c_{22}\, l_{21} + c_{22}^2\,(1 - l_{22})]^{\beta_2} \end{aligned} \tag{6.3}$$

The multivariate non-central beta distribution of $\mathbf{L}$ in the planar case can be obtained, theoretically, by integrating out c_{11}, c_{21}, c_{22} and other c's from the above distribution. This, however, appears to be possible, explicitly in terms of common functions, only after expanding the last bracket in (6.3). The resulting expression, however, appears to be too complicated; otherwise we could have obtained the distribution of t_{ii}^2 (the factors of Wilks's Λ) in the planar case by using (3.7).

Acknowledgment: The author wishes to thank Professor M. S. Bartlett for helpful discussions during the preparation of this paper.

REFERENCES

[1] T. W. Anderson and M. A. Girschick, "Some extensions of the Wishart distribution," *Ann. Math. Stat.*, Vol. 15 (1944), pp. 345–357.

[2] T. W. Anderson, "The non-central Wishart distribution and certain problems of multivariate statistics," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 409–431.

[3] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley and Sons, New York, 1958.

[4] M. S. Bartlett, "Further aspects of the theory of multiple regression," *Proc. Camb. Phil. Soc.*, Vol. 34 (1938), pp. 33–40.

[5] M. S. Bartlett, "The goodness of fit of a single hypothetical discriminant function in the case of several groups," *Ann. Eugen.*, Vol. 16 (1951), pp. 199–214.

[6] C. G. Khatri, "On the mutual independence of certain statistics," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 1258–1262.

[7] R. A. Fisher, "The statistical utilization of multiple measurements," *Ann. Eugen.*, Vol. 8 (1938), p. 376.

[8] P. L. Hsu, "On the distribution of the roots of certain determinantal equations," *Ann. Eugen.*, Vol. 9 (1939), p. 250.

[9] C. Radhakrishna Rao, *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons, New York, 1952.

[10] E. J. Williams, "Some exact tests in multivariate analysis," *Biometrika*, Vol. 39 (1952), pp. 17–31.

[11] E. J. Williams, "Significance tests for discriminant functions and linear functional relationships", *Biometrika*, Vol. 42 (1955), pp. 360–381.