NOTES

DISTRIBUTION OF THE 'GENERALISED' MULTIPLE CORRELATION MATRIX IN THE DUAL CASE

By C. G. Khatri[1]

Gujarat University

1. Introduction. Let there be two variable vectors $\mathbf{x}' = (x_1, x_2, \cdots, x_p)$ and $\mathbf{y}' = (y_1, y_2, \cdots, y_q)$ with zero means. Let $\mathbf{X} = (x_{ir}): p \times n$ and $\mathbf{Y} = (y_{jq}): q \times n$ be n independent observations on these variables. Considering the variables of $\mathbf{x}$ as fixed, we have the usual multivariate linear regression analysis of $\mathbf{y}$ set on $\mathbf{x}$-set in the following table:

Source	d. f.	Matrix (of order $q \times q$) of s. s. and s. p.
Linear regression coefficients of $\mathbf{y}$-set on $\mathbf{x}$-set	p	$\mathbf{YX}' (\mathbf{XX}')^{-1} \mathbf{XY}' = \mathbf{B}$
Residual	$n - p$	$\mathbf{YY}' - \mathbf{YX}' (\mathbf{XX}')^{-1} \mathbf{XY}' = \mathbf{A} - \mathbf{B}$
Total due to $\mathbf{y}$-set	n	$\mathbf{YY}' = \mathbf{A}$

Now let us define a matrix $\mathbf{L}$ by the relations

$$\mathbf{B} = \mathbf{TLT}' \text{ and } \mathbf{A} = \mathbf{TT}', \tag{1}$$

where $\mathbf{T}: q \times q$ is a non-singular matrix and as a special case, if need be, $\mathbf{T}$ can be looked on as a triangular one. Since $\mathbf{T}$ is the matrix square root of $\mathbf{A}$ in a certain sense, $\mathbf{L}$ is called the 'generalised' multiple correlation matrix of the $\mathbf{y}$-set on the $\mathbf{x}$-set, for if $q = 1$, it is the square of the multiple correlation of y on the $\mathbf{x}$-set. It may be noted that we may consider any other type of the matrix square root of $\mathbf{A}$, namely, $\mathbf{A} = \mathbf{T}^2$, where $\mathbf{T}: q \times q$ is a symmetric matrix and denoted by $\mathbf{T} = \mathbf{A}^{\frac{1}{2}}$. Also, it may be noted that a nonzero characteristic root of $\mathbf{L}$ gives us the square of a canonical correlation coefficient between the $\mathbf{y}$-set and the $\mathbf{x}$-set.

When $q = 1$, Bartlett [3] pointed out *the duality of relationship* between the distributions of the multiple correlations between y and the $\mathbf{x}$-set (i) when the $\mathbf{x}$-set is fixed, and y is normal and (ii) when y is fixed and the $\mathbf{x}$-set is normal.

The purpose of this paper is to point out that such a duality exists for the distribution of $\mathbf{L}$ under appropriate null hypotheses only. This is done by obtaining the null and non-null distribution of $\mathbf{L}$ (for a linear case only) under the assumption of a linear regression of the $\mathbf{x}$-set (being normal) on the $\mathbf{y}$-set con-

Received January 17, 1963; revised April 6, 1964.

[1] Now at University of North Carolina, Chapel Hill.

sidered as fixed, and under the assumption of a linear regression of the **y**-set (being normal) on the **x**-set considered as fixed. Application of this distribution in discriminant analysis is also given.

2. Non-null distribution of L (in the linear case). Under the assumptions given above, the distribution of **X** when **Y** is fixed is written as

(2) $$\text{constant} \exp[-\tfrac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\beta}\mathbf{Y})(\mathbf{X} - \boldsymbol{\beta}\mathbf{Y})'] \, d\mathbf{X}, \quad (-\infty \leqq \mathbf{X} \leqq \infty),$$

where $\mathbf{\Sigma}: p \times p$ is positive definite and $\boldsymbol{\beta}: p \times q$ is a regression matrix.

Since $\mathbf{T}^{-1}\mathbf{Y} = \mathbf{M}_1 : q \times n$ is an orthonormal matrix (i.e. $\mathbf{M}_1\mathbf{M}_1' = \mathbf{I}_q$), we can find a matrix $\mathbf{M}_2 : (n - q) \times n$ of rank $(n - q)$ such that $\mathbf{M}' = (\mathbf{M}_1' \, \mathbf{M}_2')$ is an orthogonal matrix. Using the transformation $\mathbf{X}\mathbf{M}' = (\mathbf{Z}_1 \, \mathbf{Z}_2)$, $\mathbf{Z}_1 : p \times q$ and $\mathbf{Z}_2 : p \times (n - q)$, the joint distribution of $\mathbf{Z}_1$ and $\mathbf{Z}_2$ comes out as

(3) $$\text{constant} \exp[-\tfrac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1}(\mathbf{Z}_1 - \boldsymbol{\beta}\mathbf{T})(\mathbf{Z}_1 - \boldsymbol{\beta}\mathbf{T})' - \tfrac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1}\mathbf{Z}_2\mathbf{Z}_2'] \, d\mathbf{Z}_1 \, d\mathbf{Z}_2,$$

and the matrix **L**, given by (1), can be written as

(4) $$\mathbf{L} = \mathbf{Z}_1'(\mathbf{Z}_1\mathbf{Z}_1' + \mathbf{Z}_2\mathbf{Z}_2')^{-1}\mathbf{Z}_1.$$

In (3), we first use Hsu's theorem (Anderson, Lemma (13.3.1.), ([1], p. 319)) for deriving the distribution of $\mathbf{S}_2 = \mathbf{Z}_2\mathbf{Z}_2'$ and then use the transformation

(5) $$\mathbf{S} = \mathbf{S}_2 + \mathbf{Z}_1\mathbf{Z}_1' \text{ and } \mathbf{W} = \mathbf{S}^{-\frac{1}{2}}\mathbf{Z}_1.$$

The jacobian of the transformation is $J(\mathbf{S}_2, \mathbf{Z}_1; \mathbf{S}, \mathbf{W}) = J(\mathbf{S}_2; \mathbf{S})J(\mathbf{Z}_1; \mathbf{W}) = |\mathbf{S}|^{\frac{1}{2}q}$ (refer Khatri [5]). Hence, we derive the joint distribution for **S** and **W** as

(6) $$\text{constant } |\mathbf{S}|^{\frac{1}{2}(n-p-1)}|\mathbf{I}_p - \mathbf{W}\mathbf{W}'|^{\frac{1}{2}(n-q-p-1)} \exp[-\tfrac{1}{2} \operatorname{tr}\mathbf{\Sigma}^{-1}\mathbf{S} - \tfrac{1}{2}\lambda + \operatorname{tr}(\mathbf{T}'\boldsymbol{\beta}'\mathbf{\Sigma}^{-1}\mathbf{S}^{\frac{1}{2}}\mathbf{W})] \, d\mathbf{S} \, d\mathbf{W},$$

where $d\mathbf{S} = \prod_{i \geqq j} ds_{ij}$, $\mathbf{S}$ and $\mathbf{I}_p - \mathbf{W}\mathbf{W}'$ are positive definite $\lambda = \operatorname{tr}(\mathbf{\Sigma}^{-1}\boldsymbol{\beta}\mathbf{A}\boldsymbol{\beta}')$, and $\mathbf{L} = \mathbf{W}'\mathbf{W}$.

We shall consider the rank of $\boldsymbol{\beta}$ as one only. In this case, we can write the matrix $\mathbf{T}'\boldsymbol{\beta}'\mathbf{\Sigma}^{-1}\mathbf{S}^{\frac{1}{2}}$ as

(7) $$\mathbf{T}'\boldsymbol{\beta}'\mathbf{\Sigma}^{-1}\mathbf{S}^{\frac{1}{2}} = \boldsymbol{\delta}_1(s \; \mathbf{O})\mathbf{\Delta}_2,$$

where $\mathbf{\Delta}_2 : p \times p$ is an orthogonal matrix, $\boldsymbol{\delta}_1$ is a column vector of q elements such that $\boldsymbol{\delta}_1'\boldsymbol{\delta}_1 = 1$ and **O** on the right is a null row vector with $(p - 1)$ elements. Using the transformation $\mathbf{V} = \mathbf{\Delta}_2\mathbf{W}$ and denoting $\boldsymbol{v}$ the first column vector of $\mathbf{V}'$, the jacobian of the transformation is unity and the joint distribution of **V** and **S** with the help of $|\mathbf{I}_p - \mathbf{V}\mathbf{V}'| = |\mathbf{I}_q - \mathbf{V}'\mathbf{V}|$ can be written as

(8) $$\text{constant } |\mathbf{S}|^{\frac{1}{2}(n-p-1)}|\mathbf{I}_q - \mathbf{V}'\mathbf{V}|^{\frac{1}{2}(n-q-p-1)}\exp(-\tfrac{1}{2} \operatorname{tr}\mathbf{\Sigma}^{-1}\mathbf{S} - \tfrac{1}{2}\lambda + s\boldsymbol{v}'\boldsymbol{\delta}_1) \, d\mathbf{V} \, d\mathbf{S},$$

where **S** and $\mathbf{I}_q - \mathbf{V}'\mathbf{V}$ are positive definite, and $\mathbf{L} = \mathbf{V}'\mathbf{V}$.

Now, we may note that when $q \leqq p$, the distribution of **L** is simple and derivable from $q > p$ with necessary changes in the parameter. Hence, we shall only consider the case for $q > p$ and finally write down the result for $q \leqq p$.

When $q > p$, the rank of $\mathbf{L}$: $(q \times q)$ is p. Let us write $\mathbf{V} = (\mathbf{V}_1 \, \mathbf{V}_2)$, where $\mathbf{V}_1 : p \times p$ and $\mathbf{V}_2 : p \times (q - p)$. Let us use the transformation

$$\mathbf{V}_1' = \mathbf{G}_1 \mathbf{\Delta}_3 \text{ and } \mathbf{V}_2' = \mathbf{G}_2 \mathbf{\Delta}_3 , \tag{9}$$

where $\mathbf{G}_2 : (q - p) \times p$, $\mathbf{G}_1 : p \times p$ is a lower triangular with diagonal elements $g_{ii} > 0$, and $\mathbf{\Delta}_3 : p \times p$ is an orthogonal matrix. Using the jacobian results given by Roy ([10], Appendix 5 and 6) or Olkin [9], the jacobian of the transformation given in (9) is

$$J(\mathbf{V}_1 , \mathbf{V}_2 ; \mathbf{G}_1 , \mathbf{\Delta}_3 , \mathbf{G}_2) = J(\mathbf{V}_1 ; \mathbf{G}_1 , \mathbf{\Delta}_3) J(\mathbf{V}_2 ; \mathbf{G}_2) = 2^p \prod_{i=1}^{p} g_{ii}^{p-i} J(\mathbf{\Delta}_3), \tag{10}$$

where $J(\mathbf{\Delta}_3)$ is a function of the elements of $\mathbf{\Delta}_3$. Let $\boldsymbol{\delta}$ be a first column vector of $\mathbf{\Delta}_3$ and let $\boldsymbol{\nu} = s(\mathbf{G}_1' \, \mathbf{G}_2')\boldsymbol{\delta}_1$. Also, by an indirect method similar to that given by Roy ([10], p. 197), it is easy to verify that

$$\int_{\mathbf{\Delta}_3' \mathbf{\Delta}_3 = \mathbf{I}} \exp(\boldsymbol{\nu}' \boldsymbol{\delta}) \, J(\mathbf{\Delta}_3) \, d\mathbf{\Delta}_3 = \sum_{j=0}^{\infty} c_1(j) \, \alpha^j, \tag{11}$$

where $c_1(j)$ is a positive constant depending on j and p, and $\alpha = \boldsymbol{\nu}'\boldsymbol{\nu} = \boldsymbol{\delta}_1' \mathbf{L} \boldsymbol{\delta}_1 s^2 = \operatorname{tr}(\mathbf{\Sigma}^{-1} \boldsymbol{\beta} \mathbf{T} \mathbf{L} \mathbf{T}' \boldsymbol{\beta}' \mathbf{\Sigma}^{-1} \mathbf{S})$, $\mathbf{L}$ being $\binom{\mathbf{G}_1}{\mathbf{G}_2}(\mathbf{G}_1' \, \mathbf{G}_2')$.

Moreover, we may note that

$$\begin{aligned} &\int_{\mathbf{S}} (\operatorname{tr} \mathbf{PS})^j \, |\mathbf{S}|^{\frac{1}{2}(n-p-1)} \exp\left(-\tfrac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1} \mathbf{S}\right) d\mathbf{S} \\ &\quad = \text{coef. of } (\theta^j/j!) \text{ in } \int |\mathbf{S}|^{\frac{1}{2}(n-p-1)} \exp\left[-\tfrac{1}{2} \operatorname{tr} (\mathbf{\Sigma}^{-1} - 2\theta \mathbf{P})\mathbf{S}\right] d\mathbf{S} \\ &\quad = (\text{constant depending on } \mathbf{\Sigma}, n, p, j) \times (\operatorname{tr} \mathbf{P\Sigma})^j, \end{aligned} \tag{12}$$

if the rank of $\mathbf{P}$ is one and $\mathbf{P}$ is symmetric.

Now, we use the transformation (9) in (8). Then, integrating first with $\mathbf{\Delta}_3$ and then with $\mathbf{S}$ with the help of (11) and (12), we obtain the joint distribution of $\mathbf{G}_1$ and $\mathbf{G}_2$ as

$$2^p \exp(-\tfrac{1}{2}\lambda) \sum_{j=0}^{\infty} c_2(j) [\operatorname{tr}(\mathbf{T}\mathbf{L}\mathbf{T}' \boldsymbol{\beta}' \mathbf{\Sigma}^{-1} \boldsymbol{\beta})]^j |\mathbf{I} - \mathbf{L}|^{\frac{1}{2}(n-p-q-1)} \prod_{i=1}^{p} g_{ii}^{p-i} \, d\mathbf{G}_1 \, d\mathbf{G}_2 , \tag{13}$$

where $(\mathbf{I} - \mathbf{L})$ is positive definite, and $c_2(j)$ is a constant possibly depending on $\mathbf{\Sigma}$, p, n, q and j.

Let us finally consider the transformations

$$\mathbf{L}_{11} = \mathbf{G}_1 \mathbf{G}_1' \text{ and } \mathbf{L}_{12} = \mathbf{G}_1 \mathbf{G}_2' . \tag{14}$$

The jacobian of the transformation is

$$J(\mathbf{G}_1 , \mathbf{G}_2 ; \mathbf{L}_{11} , \mathbf{L}_{12}) = J(\mathbf{G}_2 ; \mathbf{L}_{12}) J(\mathbf{G}_1 ; \mathbf{L}_{11}) = 2^{-p} \prod_{i=1}^{p} g_{ii}^{-(p-i+1)-(q-p)},$$

and

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ \mathbf{L}_{12}' & \mathbf{L}_{12}'\mathbf{L}_{11}^{-1}\mathbf{L}_{12} \end{pmatrix},$$

$\mathbf{L}_{11} : p \times p$ and $\mathbf{L}_{12} : p \times (q - p)$, (see Roy [10] or Olkin [9]).

Using these results in (13), we get the distribution of $\mathbf{L}$ as

$$(15) \quad \exp(-\tfrac{1}{2}\lambda) \sum_{j=0}^{\infty} c_2(j)[\mathrm{tr}(\mathbf{TLT}'\beta'\mathbf{\Sigma}^{-1}\beta)]^j |\mathbf{I} - \mathbf{L}|^{\frac{1}{2}(n-p-q-1)} |\mathbf{L}_{11}|^{\frac{1}{2}(p-q-1)} \, d\mathbf{L}_{11} \, d\mathbf{L}_{12},$$

where $\mathbf{I} - \mathbf{L}$ and $\mathbf{L}_{11}$ are positive definite, rank of β is one, $q > p$, $\lambda = \mathrm{tr}(\mathbf{\Sigma}^{-1}\beta\mathbf{A}\beta')$ and $c_2(j)$ is a positive constant depending on n, p, q, j and possibly on $\mathbf{\Sigma}$. If we integrate over $\mathbf{L}_{11}$ and $\mathbf{L}_{12}$, we can get $c_2(j)$ or if we carry out the explicit expressions of the constants from the beginning, we arrive at

$$(16) \quad \begin{aligned} c_2(j) &= \Gamma(\tfrac{1}{2}n + j) \{\pi^{\frac{1}{2}p(q-p)+\frac{1}{4}p(p-1)} j! \, 2^j \, \Gamma\left(\frac{n-q}{2}\right) \Gamma(\tfrac{1}{2}p + j)\}^{-1} \\ &\cdot \left\{\prod_{i=2}^{p} \Gamma\left(\frac{n-i+1}{2}\right)\right\} \left[\prod_{i=2}^{p} \left\{\Gamma\left(\frac{n-q-i+1}{2}\right) \Gamma\left(\frac{p-i+1}{2}\right)\right\}\right]^{-1}. \end{aligned}$$

We shall call the distribution of $\mathbf{L}$ given by (15) as 'pseudo' non-central multivariate beta distribution, the word 'pseudo' being used in the same sense as that used in 'pseudo' Wishart distribution defined by Roy and Gnanadesikan [9].

Now, when $p \geqq q$, $\mathbf{L}$ is positive definite and so we shall get the non-central multivariate beta distribution (see Kshirsagar [6]) which can be written as

$$(17) \quad \exp(-\tfrac{1}{2}\lambda) \sum_{j=0}^{\infty} c_3(j)[\mathrm{tr}(\mathbf{TLT}'\beta'\mathbf{\Sigma}^{-1}\beta)]^j |\mathbf{I} - \mathbf{L}|^{\frac{1}{2}(n-p-q-1)} |\mathbf{L}|^{\frac{1}{2}(p-q-1)} \, d\mathbf{L}$$

where $\mathbf{L}$ and $\mathbf{I} - \mathbf{L}$ are positive definite and

$$(18) \quad \begin{aligned} c_3(j) &= \Gamma(\tfrac{1}{2}n + j) \left\{\pi^{\frac{1}{4}q(q-1)} j! \, 2^j \, \Gamma\left(\frac{n-q}{2}\right) \Gamma(\tfrac{1}{2}p + j)\right\}^{-1} \\ &\cdot \left\{\prod_{i=2}^{q} \Gamma \frac{n-i+1}{2}\right\} \left[\prod_{i=2}^{q} \left\{\Gamma\left(\frac{n-p-i+1}{2}\right) \Gamma\left(\frac{p-i+1}{2}\right)\right\}\right]^{-1}. \end{aligned}$$

Now, for establishing the duality relationship as mentioned in Section 1, let us assume that the columns of $\mathbf{Y} : q \times n$ are independent normals with covariance matrix $\mathbf{\Sigma}_1$ and $E(\mathbf{Y}) = \beta_1\mathbf{X}$ when $\mathbf{X} : p \times n$ is fixed. Then it is easy to verify that the joint distribution of $\mathbf{R} : q \times q = \mathbf{YY}'$ and $\mathbf{U} : q \times p = \mathbf{T}^{-1}\mathbf{YX}'(\mathbf{XX}')^{-\frac{1}{2}}$, ($\mathbf{T}$ being given by $\mathbf{TT}' = \mathbf{R}$), is

$$(19) \quad \begin{aligned} &\text{constant } |\mathbf{R}|^{\frac{1}{2}(n-q-1)} |\mathbf{I}_q - \mathbf{UU}'|^{\frac{1}{2}(n-p-q-1)} \exp[-\tfrac{1}{2}\mathrm{tr}\mathbf{\Sigma}_1^{-1}\mathbf{R} - \tfrac{1}{2}\lambda_1] \\ &\qquad \times \exp[\mathrm{tr}\mathbf{T}'\mathbf{\Sigma}_1^{-1}\beta_1(\mathbf{XX}')^{\frac{1}{2}}\mathbf{U}'] \, d\mathbf{U} \, d\mathbf{R}, \end{aligned}$$

where $\mathbf{R}$ is positive definite and $\lambda_1 = \mathrm{tr}(\mathbf{\Sigma}_1^{-1}\beta_1\mathbf{XX}'\beta_1')$. Note that $\mathbf{L} = \mathbf{UU}'$ and (19) is comparable with (6). It is easy to see that when $\beta_1 = \mathbf{0}$ and $\beta = \mathbf{0}$ in (19) and (6) respectively, then the distributions of $\mathbf{L}$ in the two situations are identical. If the **y**-set in (6) is random, and the **x**-set in (19) is random, then $\beta_1 = \mathbf{0}$ or $\beta = \mathbf{0}$ implies the independence of the **x** and **y** sets, but otherwise they have different meanings. Now, for the non-null distribution of **L** (for a linear case) from (19), applying the technique used from (6) onwards, we find that the form of the distribution of **L** is different from (15) or (17) except when $q = 1$, but it is interesting to note that the forms of the distribution of the ch. roots of **L** for the non-null case in two situations are the same except for the non-central parameters—in the first case, they depend on the ch. roots of $(\beta'\Sigma^{-1}\beta\mathbf{YY}'): q \times q.$, while in the latter case, they depend on the ch. roots of the $(\beta_1'\Sigma_1^{-1}\beta_1\mathbf{XX}'): p \times p$. This shows that the duality relationship for the distribution of **L** exists only in the null case (appropriately framed).

3. Application in discriminant analysis. Let there be $(p + 1)$ populations. The multivariate analysis of variance table for q characters $\mathbf{y}' = (y_1, y_2, \cdots, y_q)$ will be of the form:

Source	d. f.	matrix of s. s. and s. p.
Between Populations	p	**B**
Within Populations	$n - p$	**A** − **B**
Total	n	**A**

The p degrees of freedom can be looked upon as corresponding to p dummy variables $\mathbf{x}' = (x_1, x_2, \cdots, x_p)$, (see Bartlett [4]) and the matrix **B** as the matrix of ss. and s.p. due to regression of **y** on **x**. The problems of direction and collinearity factors considered by Kshirsagar [6], [8] and Bartlett [4] are on the **y**-space, but there are situations in discriminant analysis, where the hypothetical discriminant functions are specified on dummy variables **x**. For example, in the discriminator considered by Barnard [2] in the case of Egyptian skull data, 'time' was a variable in the space of dummy variables **x**. The Section 2 shows that we can use the test procedures for testing 'direction' and 'collinearity' factors or a hypothetical discriminant function in dummy variables **x**, exactly in the same way as we do when they are considered in the **y**-space.

4. Acknowledgment. I thank Dr. A. M. Kshirsagar for suggesting the problem and also for the application suggested in Section 3.

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Analysis*. Wiley, New York.

[2] BARNARD, M. M. (1935). The secular variation of skull characters in four series of Egyptian. *Ann. Eugenics* **6** 352–371.

[3] BARTLETT, M. S. (1939). A note on tests of significance in multivariate analysis. *Proc. Cambridge Philos. Soc.* **35** 180–185.

[4] BARTLETT, M. S. (1951). The goodness of fit of a single hypothetical discriminant function in the case of several groups. *Ann. Eugenics* **16** 199–214.

[5] KHATRI, C. G. (1959). On the mutual independence of certain statistics. *Ann. Math. Statist.* **30** 1258–1262.

[6] KSHIRSAGAR, A. M. (1961). Some problems in multivariate analysis. Ph.D. thesis, Manchester University, England.

[7] KSHIRSAGAR, A. M. (1961). Non-central multivariate Beta distribution. *Ann. Math. Stat.* **32** 104–111.

[8] KSHIRSAGAR, A. M. (1962). A note on direction and collinearity factors in canonical analysis. *Biometrika* **33** 255–259.

[9] OLKIN, I. (1951). On distribution problems in Multivariate Analysis. Institute of Statistics, Univ. of North Carolina, mimeograph series No. 43.

[10] ROY, S. N. (1958). *Some Aspects of Multivariate Analysis*. Wiley, New York.

[11] ROY, S. N. and GNANADESIKAN, R. (1959). Some contributions to ANOVA in one or more dimensions: II. *Ann. Math. Stat.* **30** 318–340.