Osaka Mathematical Journal
Vol. 5, No. 1, June, 1953

On the Sampling Distributions of Classical Statistics in Multivariate Analysis

Junjiro OGAWA

CONTENTS

Introduction

The most fundamental distribution of the exact sampling theory in normal multivariate analysis is the so-called "Wishart's distribution", namely the joint distribution of the $\frac{1}{2}k(k+1)$ central sample moments of the second order formed by a random sample of size n drawn from a k-variate normal population. This distribution was obtained by R. A. Fisher(1) in 1915 for the special case when $k=2$, and the derivation for the general case was first given, in 1928, by Prof. John Wishart(2), and later various methods of derivation were given by various authors(3).

The important statistics in normal multivariate theory are the classical inter-class correlation coefficient, multiple correlation coefficient, partial correlation coefficient, and the Hotelling's generalised Student's ratio T. The exact sampling distributions of these statistics were derived on the basis of Wishart distribution, and they are well-known and now are classical.

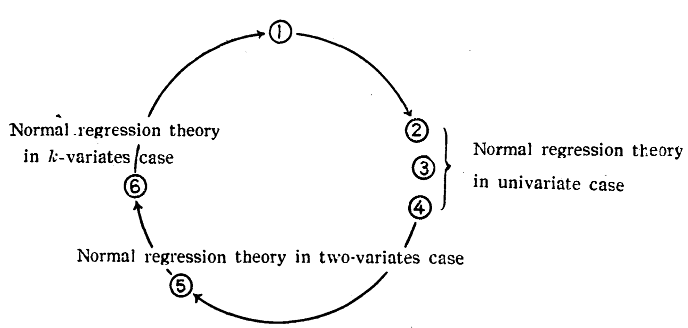
In 1933, Prof. M. S. Bartlett(4) considered in detail the processes of derivations of the sampling distributions of some multivariate statistics from the Wishart distribution and established the "Decomposition

Theorem " of the Wishart distribution.

Later, in 1947, Prof. G. Elfving[5] considered the problem of deducing the sampling distributions of the classical statistics in multivariate analysis in a systematic way, and by the use of the geometrical interpretations of the results of the normal regression theory, he succeeded in the systematic derivations of the sampling distributions of the multiple correlation coefficient, Hotelling's T^2, Bartlett's decomposition theorem and some others, but only in the special cases when the null hypotheses are true.

Prof. G. Elfving's fundamental idea was as follows: if we consider the conditional distribution of some variables $x_1, \dots, x_s$, for instance, of the k-dimensional normal distribution, fixing the other variables $x_{s+1}, \dots, x_k$, then the regressions of $x_1, \dots, x_s$ on the fixed variables $x_{s+1}, \dots, x_k$ are all linear, and therefore the results of the theory of normal regression are available effectively.

The purpose of this paper is to deiive the sampling distributions of the above mentioned classical statistics appearing in the multivariate analysis in general cases, following the idea of the conditional distributions. The diagram shown below may help to understand the rough idea on the order of derivations and the methods used.



① Wishart distribution
② Multiple correlation coefficient
③ Inter-class correlation coefficient
④ Hotelling's T^2
⑤ Partial correlation coefficient
⑥ Bartlett's "Decomposition Theorem"

The results which will be presented in the following are all well-known, but there seems to be some interest in the method used from the methodological point of view.

§ 1. Wishart Distribution

The arguments which will be presented in the following are essentially due to Prof. D. Fog's idea[6].

We shall consider the k-dimensional row vector variate

$$\mathfrak{x} = (x_1, x_2, \cdots, x_k)$$

which follows the non-singular k-dimensional normal distribution with the mean vector

$$\mathfrak{m} = (m_1, m_2, \cdots, m_k)$$

and the variance-covariance matrix

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots\cdots & \lambda_{1k} \\ \lambda_{21} & \lambda_{22} & \cdots\cdots & \lambda_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_{k1} & \lambda_{k2} & \cdots\cdots & \lambda_{kk} \end{pmatrix}, \quad \text{where} \quad \lambda_{ij} = \lambda_{ji}.$$

The population probability distribution has the probability element

$$p(\mathfrak{x})d\mathfrak{x} = (2\pi)^{-k/2}\Lambda^{-1/2}\exp\left[-\frac{1}{2}(\mathfrak{x}-\mathfrak{m})\mathbf{\Lambda}^{-1}(\mathfrak{x}-\mathfrak{m})'\right]d\mathfrak{x}, \qquad (1.1)$$

where the symbol $d\mathfrak{x}$ is the abbreviation of the differential $dx_1dx_2 \cdots dx_k$, and $\mathfrak{x}'$ denotes the transposed vector of $\mathfrak{x}$, and the capital Greek letter, as for example Λ, corresponding to the bold face letter, $\mathbf{\Lambda}$, denotes the determinant of the corresponding matrix.

If we denote the conditional frequency function of x_1, given $x_2, \cdots, x_k$, by

$$p(x_1 | x_2, x_3, \cdots, x_k)$$

and the conditional frequency function of x_2 given $x_3, \ldots, x_k$, by

$$p(x_2 | x_3, x_4, \cdots, x_k)$$

and so on, then we have

$$\begin{aligned} p(\mathfrak{x})\,d\mathfrak{x} = p(x_1 | x_2, \cdots, x_k)\,dx_1 \cdot p(x_2 | x_3, \cdots, x_k)\,dx_2 \cdot \cdots \cdot \\ p(x_{k-1} | x_k)\,dx_{k-1} \cdot p(x_k)\,dx_k, \end{aligned} \qquad (1.2)$$

where, of course, $p(x_k)$ is the marginal frequency function of x_k.

Our purpose of this section is to obtain the joint distribution of the following $\frac{1}{2}k(k+1)$ statistics:

$$m_{11}=\sum_{\alpha=1}^{n}(x_{1\alpha}-\bar{x}_1)^2,\ m_{12}=\sum_{\alpha=1}^{n}(x_{1\alpha}-\bar{x}_1)(x_{2\alpha}-\bar{x}_2),\ \cdots,\ m_{1k}=\sum_{\alpha=1}^{n}(x_{1\alpha}-\bar{x}_1)(x_{k\alpha}-\bar{x}_k),$$
$$m_{22}=\sum_{\alpha=1}^{n}(x_{2\alpha}-\bar{x}_2)^2,\ \cdots\cdots\cdots,\ m_{2k}=\sum_{\alpha=1}^{n}(x_{2\alpha}-\bar{x}_2)(x_{k\alpha}-\bar{x}_k),$$
$$\ddots \qquad \vdots \qquad (1.3)$$
$$m_{kk}=\sum_{\alpha=1}^{n}(x_{k\alpha}-\bar{x}_k)^2,$$

where

$$\mathfrak{x}_\alpha=(x_{1\alpha},x_{2\alpha},\cdots,x_{k\alpha}),\quad \alpha=1,2,\cdots,n$$

is a random sample of size n from the population being considered.

Put

$$\begin{pmatrix}\sqrt{n}\,\bar{x}_i\\ y_{i1}\\ \vdots\\ y_{i,n-1}\end{pmatrix}=\boldsymbol{H}\cdot\begin{pmatrix}x_{i1}\\ x_{i2}\\ \vdots\\ x_{in}\end{pmatrix},\quad i=1,2,\cdots,k, \qquad (1.4)$$

where the matrix $\boldsymbol{H}$ is the well-known "Helmert's orthogonal matrix", i.e.

$$\boldsymbol{H}=\begin{pmatrix}\frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots\cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}}\\ \sqrt{\frac{n-1}{n}} & \frac{-1}{\sqrt{n(n-1)}} & \frac{-1}{\sqrt{n(n-1)}} & \cdots\cdots & \frac{-1}{\sqrt{n(n-1)}} & \frac{-1}{\sqrt{n(n-1)}}\\ 0 & \sqrt{\frac{n-2}{n-1}} & \frac{-1}{\sqrt{(n-1)(n-2)}} & \cdots\cdots & \frac{-1}{\sqrt{(n-1)(n-2)}} & \frac{-1}{\sqrt{(n-1)(n-2)}}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots\cdots & \sqrt{\frac{1}{2}} & \frac{-1}{\sqrt{2\cdot 1}}\end{pmatrix}, \qquad (1.5)$$

then it is clear that

$$m_{ij}=\sum_{\alpha=1}^{n-1}y_{i\alpha}y_{j\alpha},\quad i\leqq j\quad i,j=1,2,\cdots,k. \qquad (1.6)$$

If we consider the conditional probability distribution of

$$\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots\cdots,\boldsymbol{y}_{k-1}$$

given $\boldsymbol{y}_k$, where the column vector $\boldsymbol{y}_i$ denotes the $(n-1)$-dimensional column vector

$$\boldsymbol{y}_i=\begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{i,n-1} \end{pmatrix}, \quad i=1,2,\dots,k,$$

and apply the following orthogonal transformation to $\boldsymbol{y}_1,\cdots,\boldsymbol{y}_{k-1}$, then it follows the result, i.e.,

$$\boldsymbol{y}_1^{(1)}=\begin{pmatrix} \xi_{11} \\ y_{12}^{(1)} \\ \vdots \\ y_{1,n-1}^{(1)} \end{pmatrix} \quad \boldsymbol{y}_2^{(1)}=\begin{pmatrix} \xi_{21} \\ y_{22}^{(1)} \\ \vdots \\ y_{2,n-1}^{(1)} \end{pmatrix}, \cdots\cdots, \quad \boldsymbol{y}_{k-1}^{(1)}=\begin{pmatrix} \xi_{k-1,1} \\ y_{k-1,2}^{(1)} \\ \vdots \\ y_{k-1,n-1}^{(1)} \end{pmatrix}, \tag{1.7}$$

where

$$\xi_{k1}^2=\sum_{\alpha=1}^{n-1} y_{k\alpha}^2, \quad \xi_{i1}=\frac{1}{\xi_{k1}}\sum_{\alpha=1}^{n-1} y_{i\alpha}y_{k\alpha}, \quad i=1,2,\cdots,k-1, \tag{1.8}$$

and the matrix of tranformation is

$$\boldsymbol{H}_k=\begin{pmatrix} \frac{y_{k1}}{\sqrt{y_{k1}^2+\cdots+y_{k,n-1}^2}} & \frac{y_{k2}}{\sqrt{y_{k1}^2+\cdots+y_{k,n-1}^2}} & \cdots\cdots & \frac{y_{k,n-2}}{\sqrt{y_{k1}^2+\cdots+y_{k,n-1}^2}} & \frac{y_{k,n-1}}{\sqrt{y_{k1}^2+\cdots+y_{k,n-1}^2}} \\ \sqrt{\frac{y_{k2}^2+\cdots+y_{k,n-1}^2}{y_{k1}^2+y_{k2}^2+\cdots+y_{k,n-1}^2}} & \frac{-y_{k1}y_{k2}}{\sqrt{(y_{k1}^2+\cdots+y_{k,n-1}^2)(y_{k2}^2+\cdots+y_{k,n-1}^2)}} & \cdots\cdots & \frac{-y_{k1}y_{k,n-2}}{\sqrt{(y_{k1}^2+\cdots+y_{k,n-1}^2)(y_{k2}^2+\cdots+y_{k,n-1}^2)}} & \frac{-y_{k1}y_{k,n-1}}{\sqrt{(y_{k1}^2+\cdots+y_{k,n-1}^2)(y_{k2}^2+\cdots+y_{k,n-1}^2)}} \\ 0 & \sqrt{\frac{y_{k3}^2+\cdots+y_{k,n-1}^2}{y_{k2}^2+y_{k3}^2+\cdots+y_{k,n-1}^2}} & \cdots\cdots & \frac{-y_{k2}y_{k,n-2}}{\sqrt{(y_{k2}^2+\cdots+y_{k,n-1}^2)(y_{k3}^2+\cdots+y_{k,n-1}^2)}} & \frac{-y_{k2}y_{k,n-1}}{\sqrt{(y_{k2}^2+\cdots+y_{k,n-1}^2)(y_{k3}^2+\cdots+y_{k,n-1}^2)}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots\cdots & \sqrt{\frac{y_{k,n-1}^2}{y_{k,n-2}^2+y_{k,n-1}^2}} & \frac{-y_{k,n-2}y_{k,n-1}}{\sqrt{(y_{k,n-2}^2+y_{k,n-1}^2)y_{k,n-1}^2}} \end{pmatrix} \tag{1.9}$$

which we call the "generalized Helmert's orthogonal matrix" for the time being.

Next, consider the conditional distribution of

$$\boldsymbol{y}_1^{(1)},\ \boldsymbol{y}_2^{(1)},\ \cdots\cdots,\ \boldsymbol{y}_{k-2}^{(1)}$$

given $\boldsymbol{y}_{k-1}^{(1)}$ and $\boldsymbol{y}_k$, and applying the orthogonal transformation of which the matrix of transformation is

$$\boldsymbol{H}_{k-1}=\left(\begin{array}{cccccc}
1 & 0 & 0 & \cdots\cdots & 0 & 0\\
0 & \dfrac{y^{(1)}_{k-1,2}}{\sqrt{y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1}}} & \dfrac{y^{(1)}_{k-1,3}}{\sqrt{y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1}}} & \cdots\cdots & \dfrac{y^{(1)}_{k-1,n-2}}{\sqrt{y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1}}} & \dfrac{y^{(1)}_{k-1,n-1}}{\sqrt{y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1}}}\\
0 & \sqrt{\dfrac{y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1}}{y^{(1)2}_{k-1,2}+y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1}}} & \dfrac{-y^{(1)}_{k-1,2}y^{(1)}_{k-1,3}}{\sqrt{(y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1})(y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1})}} & \cdots & \dfrac{-y^{(1)}_{k-1,2}y^{(1)}_{k-1,n-2}}{\sqrt{(y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1})(y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1})}} & \dfrac{-y^{(1)}_{k-1,2}y^{(1)}_{k-1,n-1}}{\sqrt{(y^{(1)2}_{k-1,2}+\cdots+y^{(1)2}_{k-1,n-1})(y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1})}}\\
0 & 0 & \sqrt{\dfrac{y^{(1)2}_{k-1,4}+\cdots+y^{(1)2}_{k-1,n-1}}{y^{(1)2}_{k-1,3}+y^{(1)2}_{k-1,4}+\cdots+y^{(1)2}_{k-1,n-1}}} & \cdots\cdots & \dfrac{-y^{(1)}_{k-1,3}y^{(1)}_{k-1,n-2}}{\sqrt{(y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1})(y^{(1)2}_{k-1,4}+\cdots+y^{(1)2}_{k-1,n-1})}} & \dfrac{-y^{(1)}_{k-1,3}y^{(1)}_{k-1,n-1}}{\sqrt{(y^{(1)2}_{k-1,3}+\cdots+y^{(1)2}_{k-1,n-1})(y^{(1)2}_{k-1,4}+\cdots+y^{(1)2}_{k-1,n-1})}}\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
0 & 0 & 0 & \cdots\cdots & \sqrt{\dfrac{y^{(1)2}_{k-1,n-1}}{y^{(1)2}_{k-1,n-2}+y^{(1)2}_{k-1,n-1}}} & \dfrac{-y^{(1)}_{k-1,n-2}y^{(1)}_{k-1,n-1}}{\sqrt{(y^{(1)2}_{k-1,n-2}+y^{(1)2}_{k-1,n-1})y^{(1)2}_{k-1,n-1}}}
\end{array}\right) \tag{1.10}$$

and we obtain the result

$$\boldsymbol{y}_1^{(2)} = \begin{pmatrix} \xi_{11} \\ \xi_{12} \\ y_{13}^{(2)} \\ \vdots \\ y_{1,n-1}^{(2)} \end{pmatrix}, \quad \boldsymbol{y}_2^{(2)} = \begin{pmatrix} \xi_{21} \\ \xi_{22} \\ y_{23}^{(2)} \\ \vdots \\ y_{2,n-1}^{(2)} \end{pmatrix} \quad \cdots\cdots, \quad \boldsymbol{y}_{k-2}^{(2)} = \begin{pmatrix} \xi_{k-2,1} \\ \xi_{k-2,2} \\ y_{k-2,2}^{(2)} \\ \vdots \\ y_{k-1,n-1}^{(2)} \end{pmatrix} \tag{1.11}$$

where

$$\xi_{k-1,2}^2 = \sum_{\alpha=2}^{n-1} y_{k-1,\alpha}^{(1)^2},$$

$$\xi_{i2} = \frac{1}{\xi_{k-1,2}} \sum_{\alpha=2}^{n-1} y_{i\alpha}^{(1)} y_{k-1,\alpha}^{(1)} \qquad i = 1, 2, \cdots, k-2 \tag{1.12}$$

We continue the similar processes until we attain the vector

$$\boldsymbol{y}_1^{(k-1)} = \begin{pmatrix} \xi_{11} \\ \xi_{12} \\ \vdots \\ \xi_{1,k-1} \\ y_{1k}^{(k-1)} \\ \vdots \\ y_{1,n-1}^{(k-1)} \end{pmatrix}, \tag{1.13}$$

The notations here used will be easily understood by the continuation of (1.8) and (1.12).

Since the transformations used were all orthogonal, it is clear that

$$\begin{array}{llll} m_{11} = \sum_{\alpha=1}^{k} \xi_{1\alpha}^2, & m_{12} = \sum_{\alpha=1}^{k-1} \xi_{1\alpha}\xi_{2\alpha}, & \cdots\cdots\cdots, & m_{1k} = \xi_{11}\xi_{k1} \\ & m_{22} = \sum_{\alpha=1}^{k-1} \xi_{2\alpha}^2, & \cdots\cdots\cdots, & m_{2k} = \xi_{21}\xi_{k1} \\ & & \ddots & \vdots \\ & & & m_{kk} = \xi_{k1}^2 \end{array} \tag{1.14}$$

where, of course,

$$\xi_{1k}^2 = y_{1k}^{(k-1)^2} + y_{1,k+1}^{(k-1)^2} + \cdots + y_{1,n-1}^{(k-1)^2}$$

The conditional probability element of $\boldsymbol{y}_1$, given $\boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$ is

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_{1\cdot 23\cdots k}}\right)^{n-1} \exp\left[-\frac{1}{2\sigma_{1\cdot 23\cdots k}^2} \sum_{\alpha=1}^{n-1} (y_{1\alpha} - \beta_{12}y_{2\alpha} \cdots - \beta_{1k}y_{k\alpha})^2\right] dy_{11}dy_{12}\cdots dy_{1,n-1},$$

where $\sigma^2_{1\cdot 23\cdots k}$, and β_{1i}, $i=2,\cdots,k$ are the residual variance of y_1, given $y_2,\cdots,y_k$ and the partial regression coefficients of y_1 on y_i, $i=2,\cdots,k$ in the population respectively. Therefore, the conditional distribution of $\boldsymbol{y}_1^{(k-1)}$ has the probability element

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_{1\cdot 23\cdots k}}\right)^{n-1}\exp\Big[-\frac{1}{2\sigma^2_{1\cdot 23\cdots k}}\Big\{\xi_{11}^2+\cdots+\xi_{1,k-1}^2+y_{1,k}^{(k-1)^2}+\cdots+y_{1,n-1}^{(k-1)^2}$$
$$-2\beta_{12}(\xi_{11}\xi_{21}+\cdots+\xi_{1,k-1}\xi_{2,k-1})-2\beta_{13}(\xi_{11}\xi_{31}+\cdots+\xi_{1,k-2}\xi_{3,k-2})\cdots-2\beta_{1k}\xi_{11}\xi_{k1}$$
$$+\sum_{\alpha=1}^{n-1}(\beta_{12}y_{2\alpha}+\cdots+\beta_{1k}y_{k\alpha})^2\Big\}\Big]\,d\xi_{11}\cdots d\xi_{1,k-1}dy_{1,k}^{(k-1)}\cdots dy_{1,n-1}^{(k-1)}, \tag{1.15}$$

Hence, we obtain the conditional joint distribution of $\xi_{11},\cdots,\xi_{1k-1},\xi_{1k}$, given $\boldsymbol{y}_2,\cdots,\boldsymbol{y}_{k-1}$, i.e.,

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_{1\cdot 23\cdots k}}\right)^{k-1}\frac{1}{2^{\frac{n-k}{2}}\Gamma\left(\frac{n-k}{2}\right)\sigma^{n-k}_{1\cdot 23\cdots k}}\exp\Big[-\frac{1}{2\sigma^2_{1\cdot 23\cdots k}}\Big\{\xi_{11}^2+\cdots+\xi_{1,k-1}^2+\xi_{1k}^2$$
$$-2\beta_{12}(\xi_{11}\xi_{21}+\cdots+\xi_{1,k-1}\xi_{2,k-1})\cdots-2\beta_{1k}\xi_{11}\xi_{k1}+\sum_{\alpha=1}^{n-1}(\beta_{12}y_{2\alpha}+\cdots+\beta_{1k}y_{k\alpha})^2\Big\}\Big]\times$$
$$(\xi_{1k}^2)^{\frac{n-k}{2}-1}d\xi_{11}\cdots d\xi_{1,k-1}d(\xi_{1k}^2), \tag{1.16}$$

In a similar manner, we obtain the conditional distribution of $\xi_{21},\cdots,\xi_{2,k-1}$ given $\boldsymbol{y}_3,\cdots,\boldsymbol{y}_k$, successively ......, and finally the marginal distribution ξ_{k1}; i.e.,

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_{2\cdot 34\cdots k}}\right)^{k-2}\frac{1}{2^{\frac{n-k+1}{2}}\Gamma\left(\frac{n-k+1}{2}\right)\sigma^{n-k+1}_{2\cdot 34\cdots k}}\exp\Big[-\frac{1}{2\sigma^2_{2\cdot 34\cdots k}}\Big\{\xi_{21}^2+\cdots+\xi_{2,k-1}^2$$
$$-2\beta_{23}(\xi_{21}\xi_{31}+\cdots+\xi_{2,k-2}\xi_{3,k-2})\cdots-2\beta_{2k}\xi_{21}\xi_{k1}+\sum_{\alpha=1}^{n-1}(\beta_{23}y_{3\alpha}+\cdots+\beta_{2k}y_{k\alpha})^2\Big\}\Big]\times$$
$$(\xi_{2,k-1}^2)^{\frac{n-k+1}{2}-1}d\xi_{21}\cdots d\xi_{2,k-2}d(\xi_{2,k-1}^2), \tag{1.17}$$

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_{3\cdot 45\cdots k}}\right)^{k-3}\frac{1}{2^{\frac{n-k+2}{2}}\Gamma\left(\frac{n-k+2}{2}\right)\sigma^{n-k+2}_{3\cdot 45\cdots k}}\exp\Big[-\frac{1}{2\sigma^2_{3\cdot 45\cdots k}}\Big\{\xi_{31}^2+\cdots+\xi_{3,k-2}^2$$
$$-2\beta_{34}(\xi_{31}\xi_{41}+\cdots+\xi_{3,k-3}\xi_{4,k-3})\cdots-2\beta_{3k}\xi_{31}\xi_{k1}+\sum_{\alpha=1}^{n-1}(\beta_{34}y_{4\alpha}+\cdots+\beta_{3k}y_{k\alpha})^2\Big\}\Big]\times$$
$$(\xi_{3,k-2}^2)^{\frac{n-k+2}{2}-1}d\xi_{31}\cdots d\xi_{3,k-3}d(\xi_{3,k-2}^2), \tag{1.18}$$

....

$$\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)\sigma_k^{n-1}}\,e^{-\frac{\xi_{k1}^2}{2\sigma_k^2}}(\xi_{k1}^2)^{\frac{n-1}{2}-1}d(\xi_{k1}^2)\,. \tag{1.19}$$

Hence the absolute joint probability distribution of the $\frac{1}{2}k(k+1)$ variables

$$\boldsymbol{\Xi} = \begin{pmatrix} \xi_{11} & \xi_{12} & \cdots\cdots & \xi_{1,k-2} & \xi_{1,k-1} & \xi_{1k} \\ \xi_{21} & \xi_{22} & \cdots\cdots & \xi_{2,k-2} & \xi_{2,k-1} & \\ \xi_{31} & \xi_{32} & \cdots\cdots & \xi_{3,k-2} & & \\ \vdots & & \cdot^{\cdot^{\cdot}} & & & \\ \xi_{k1} & & & & & \end{pmatrix}$$

is given by the product of (1.16), (1.17), (1.18) and ······ (1.19).

If we denote the cofactor of the element λ_{ij} in the determinant Λ of the moment matrix $\boldsymbol{\Lambda}$ by Λ_{ij}, and denote the cofactor of the element λ_{pq} of Λ_{11} by $\Lambda_{11\cdot pq}$, and so on, then it is well-known that

$$\sigma^2_{1\cdot 23\cdots k} = \frac{\Lambda}{\Lambda_{11}}\ , \quad \beta_{1i} = -\frac{\Lambda_{1i}}{\Lambda_{11}}\ ; \quad i = 2, \cdots, k\,,$$

$$\sigma^2_{2\cdot 34\cdots k} = \frac{\Lambda_{11}}{\Lambda_{11\cdot 22}}\,, \quad \beta_{2i} = -\frac{\Lambda_{11\cdot 2i}}{\Lambda_{11\cdot 22}}\,, \quad i = 3, \cdots, k\,.$$

and so on. Whence it follows by easy calculations that

$$\begin{aligned} &\frac{1}{\sigma^2_{1\cdot 23\cdots k}} \sum_{\alpha=1}^{n-1} (\beta_{12}y_{2\alpha} + \cdots + \beta_{1k}y_{k\alpha})^2 \\ &+\frac{1}{\sigma^2_{2\cdot 34\cdots k}} \Big\{\xi^2_{21} + \cdots + \xi^2_{2,k-1} - 2\beta_{23}(\xi_{21}\xi_{11} + \cdots + \xi_{2,k-2}\xi_{3,k-2}) \cdots\cdots - 2\beta_{2k}\xi_{21}\xi_{k1} \\ &\qquad + \sum_{\alpha=1}^{n-1} (\beta_{23}y_{3\alpha} + \cdots + \beta_{2k}y_{k\alpha})^2\Big\} \\ &+\frac{1}{\sigma^2_{3\cdot 45\cdots k}} \Big\{\xi^2_{31} + \cdots + \xi^2_{3,k-2} - 2\beta_{34}(\xi_{31}\xi_{41} + \cdots + \xi_{3,k-3}\xi_{4,k-3}) \cdots\cdots - 2\beta_{3k}\xi_{31}\xi_{k1} \\ &\qquad + \sum_{\alpha=1}^{n-1} (\beta_{34}y_{4\alpha} + \cdots + \beta_{3k}y_{k\alpha})^2\Big\} \qquad (1.20) \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &+\frac{1}{\sigma^2_k}\cdot \xi^2_{k1} = \sum_{i,j=2}^{k} \frac{\Lambda_{ij}}{\Lambda} m_{ij} \end{aligned}$$

Consequently, we have

$$\frac{2^k}{\pi^{\frac{k(k-1)}{4}} 2^{\frac{k(n-1)}{2}} \Gamma\left(\frac{n-1}{2}\right) \cdots \Gamma\left(\frac{n-k}{2}\right)} \Lambda^{-\frac{n-1}{2}} \exp\left(-\frac{1}{2}\sum_{i,j=1}^{k} \frac{\Lambda_{ij}}{\Lambda} m_{ij}\right) (\xi_{1k}\xi_{2,k-1}\cdots\xi_{k1})^{n-k-2}$$

$$(\xi_{1k}\xi^2_{2,k-1} \cdots \xi^{k-1}_{k-1,2}\xi^k_{k1})\, d\xi_{11} \cdots d\xi_{1k} d\xi_{21} \cdots d\xi_{2,k-1} \cdots d\xi_{k1} \qquad (1.21)$$

for the joint distribution of $\boldsymbol{\Xi}$.

Since

$$\mathbf{M} \equiv (m_{ij}) = \Xi'\Xi\,,$$
$$M = \Xi^2 = (\xi_{1k}\xi_{2,k-1} \cdots\cdots \xi_{k1})^2\,,$$

and the Jacobian is

$$\frac{\partial(\mathbf{M})}{\partial(\Xi)} = \frac{\partial m_{kk}}{\partial \xi_{k1}} \cdot \frac{\partial(m_{k-1,k-1}, m_{k-1,k})}{\partial(\xi_{k-1,1}, \xi_{k-1,2})} \cdots\cdots \frac{\partial(m_{11}, m_{12}, \cdots, m_{1k})}{\partial(\xi_{11}, \xi_{12}, \cdots, \xi_{1k})}$$
$$= 2^k \cdot \xi_{1k}\xi_{2,k-1}^2\xi_{3,k-2}^3 \cdots\cdots \xi_{k-1,2}^{k-1}\xi_{k1}^k\,,$$

the joint distribution of M has the probability element

$$\frac{1}{2^{\frac{k(n-1)}{2}}\pi^{\frac{k(k-1)}{4}}\prod_{i=1}^{k}\Gamma\left(\frac{n-i}{2}\right)}\Lambda^{-\frac{n-1}{2}}\mathrm{M}^{\frac{n-k-2}{2}}\exp\left(-\frac{1}{2}\Lambda^{-1}\cdot\cdot\,\mathbf{M}\right)d\mathbf{M}\,, \qquad (1.22)$$

where the symbol $\Lambda^{-1}\cdot\cdot\,\mathbf{M}$ means

$$\sum_{ij=1}^{k}\frac{\Lambda_{ij}}{\Lambda}m_{ij}$$

From (1.22) we can easily derive the required Wishart distribution.

§ 2. Normal Regression Theory In Univariate Case.

In this section we shall summarize the main results in the normal regression theory in univariate case somewhat more precise than those which are seen in the literature(7).

Let

$$\boldsymbol{Z} = \begin{pmatrix} z_{11} & z_{21} & \cdots\cdots & z_{s1} \\ z_{12} & z_{22} & \cdots\cdots & z_{s2} \\ \vdots & \vdots & & \vdots \\ z_{1n} & z_{2n} & \cdots\cdots & z_{sn} \end{pmatrix}$$

be a $n \times s$ matrix of fixed variables of rank s, then the $s \times s$ matrix $\boldsymbol{Z}'\boldsymbol{Z}$ is symmetric and positive definite, and consequently $\boldsymbol{Z}'\boldsymbol{Z}$ has a symmetric and positive-definite square root(8), which we denote by $(\boldsymbol{Z}'\boldsymbol{Z})^{1/2}$. Further let the column vector

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{pmatrix}$$

be a vector of s unknown parameters, and let the vector variate

$$\boldsymbol{\eta} = \boldsymbol{x} - \boldsymbol{Z}\cdot\boldsymbol{\beta} \tag{2.1}$$

be a random sample of size n from a normal population $N(0, \sigma^2)$, where

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

is a system of independent observations and η is an error.

The least-square estimate $\boldsymbol{b}$ of $\boldsymbol{\beta}$ is the value of $\boldsymbol{\beta}$ which minimizes

$$S = \eta'\eta = (\boldsymbol{x} - \boldsymbol{Z}\cdot\boldsymbol{\beta})'(\boldsymbol{x} - \boldsymbol{Z}\cdot\boldsymbol{\beta}) \tag{2.2}$$

considered as a function of $\boldsymbol{x}$, and it is determined by the so-called "normal equation"

$$\boldsymbol{Z}'\cdot(\boldsymbol{x} - \boldsymbol{Z}\cdot\boldsymbol{b}) = 0\,, \tag{2.3}$$

therefore we have

$$\boldsymbol{b} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\cdot\boldsymbol{Z}'\boldsymbol{x}\,. \tag{2.4}$$

Representing $\boldsymbol{b}$ in terms of the original variate $\boldsymbol{\eta}$, we have

$$\boldsymbol{b} - \boldsymbol{\beta} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\cdot\boldsymbol{Z}'\cdot\boldsymbol{\eta}\,. \tag{2.5}$$

In (2.1), substituting $\boldsymbol{\beta}$ by its least-square estimate $\boldsymbol{b}$, we have the residual variate

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{Z}\boldsymbol{b} = (\boldsymbol{I} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}')\cdot\boldsymbol{\eta} \tag{2.6}$$

where $\boldsymbol{I}$ denotes the unit matrix of degree n.

The two variates $\boldsymbol{b}$ and $\boldsymbol{y}$ are mutually independent in the stochastic sense, because the row vectors of their coefficient matrices are orthogonal with each other[(9)], i.e.,

$$(\boldsymbol{I} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}')\cdot((\boldsymbol{Z}'\boldsymbol{Z})^{-1}\cdot\boldsymbol{Z}')' = 0$$

Hence, the minumum value S_0 of S and $\boldsymbol{b}$ are mutually independent stochastically, because

$$S_0 = \boldsymbol{y}'\boldsymbol{y} = (\boldsymbol{x} - \boldsymbol{Z}\cdot\boldsymbol{b})'\cdot(\boldsymbol{x} - \boldsymbol{Z}\cdot\boldsymbol{b})\,.$$

Next, we shall consider the case when the following partitions of $\boldsymbol{Z}$ and $\boldsymbol{\beta}$ are given; i.e.,

$$\boldsymbol{Z} = (\boldsymbol{Z}_1\boldsymbol{Z}_2),\ \text{where}\ \boldsymbol{Z}_1 = \begin{pmatrix} z_{11} & \cdots\cdots & z_{r1} \\ z_{12} & \cdots\cdots & z_{r2} \\ \vdots & & \vdots \\ z_{1n} & \cdots\cdots & z_{rn} \end{pmatrix},\ \boldsymbol{Z}_2 = \begin{pmatrix} z_{r+1,1} & \cdots\cdots & z_{s1} \\ z_{r+1,2} & \cdots\cdots & z_{s2} \\ \vdots & & \vdots \\ z_{r+1,n} & \cdots\cdots & z_{sn} \end{pmatrix},$$

and

$$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \\ \boldsymbol{\beta}_2 \end{pmatrix}, \quad \text{where } \boldsymbol{\beta}_1 = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}, \ \boldsymbol{\beta}_2 = \begin{pmatrix} \beta_{r+1} \\ \vdots \\ \beta_s \end{pmatrix}.$$

The least-square estimate $\boldsymbol{b}_1^*$ of $\boldsymbol{\beta}_1$ under the statistical hypothesis

$H_0: \boldsymbol{\beta}_2 = \boldsymbol{\beta}_2^0$, where $\boldsymbol{\beta}_2^0 = \begin{pmatrix} \beta_{r+1}^0 \\ \vdots \\ \beta_s^0 \end{pmatrix}$ is a certain specified vector,

is the value of $\boldsymbol{\beta}_1$ which minimizes

$$S^* = (\boldsymbol{x} - \boldsymbol{Z}_1\boldsymbol{\beta}_1 - \boldsymbol{Z}_2\boldsymbol{\beta}_2^0)'(\boldsymbol{x} - \boldsymbol{Z}_1\boldsymbol{\beta}_1 - \boldsymbol{Z}_2\boldsymbol{\beta}_2^0), \tag{2.7}$$

and $\boldsymbol{b}_1^*$ is determined by the matrix equation

$$\boldsymbol{Z}_1'\boldsymbol{Z}_1 \cdot \boldsymbol{b}_1^* = \boldsymbol{Z}_1' \cdot (\boldsymbol{x} - \boldsymbol{Z}_2\boldsymbol{\beta}_2^0), \tag{2.8}$$

consequently we have

$$\boldsymbol{b}_1^* = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}_1' \cdot (\boldsymbol{x} - \boldsymbol{Z}_2 \cdot \boldsymbol{\beta}_2^0),$$

or represented by the original variate η, we have

$$\boldsymbol{b}_1^* - \boldsymbol{\beta}_1 = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}_1'\eta + (\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1'\boldsymbol{Z}_2 \cdot (\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0). \tag{2.9}$$

The residual variate $\boldsymbol{y}^*$, corresponding to (2.6), is in this case

$$\begin{aligned} \boldsymbol{y}^* &= \boldsymbol{x} - \boldsymbol{Z}_1\boldsymbol{b}_1^* - \boldsymbol{Z}_2 \cdot \boldsymbol{\beta}_2^0 \\ &= (\boldsymbol{I} - \boldsymbol{Z}_1(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1')\eta + (\boldsymbol{I} - \boldsymbol{Z}_1(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1')\boldsymbol{Z}_1 \cdot (\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0). \end{aligned} \tag{2.10}$$

Hence $\boldsymbol{y}^*$ and $\boldsymbol{b}_1^*$ are mutually independent in the sense of the probability, and consequently the minimum value S_0^* of S^* under the statistical hypothesis H_0 and $\boldsymbol{b}_1^*$ are mutually independent in the stochastic sense.

2.1 We shall first consider a special case when the orthogonality condition

$$\boldsymbol{Z}_1'\boldsymbol{Z}_2 = \boldsymbol{0} \tag{2.11}$$

between the two partitioned submatrices of $\boldsymbol{Z}$ holds. In this case, it is easily seen that

$$\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}' = \boldsymbol{Z}_1(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1' + \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-1}\boldsymbol{Z}_2',$$

whence, from (2.6) aud (2.10), we have

$$\begin{aligned} \boldsymbol{y} &= (\boldsymbol{I} - \boldsymbol{Z}_1(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1')\eta - \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-1}\boldsymbol{Z}_2' \cdot \eta \\ &= \boldsymbol{y}^* - \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-1}\boldsymbol{Z}_2' \cdot \eta - \boldsymbol{Z}_2 \cdot (\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0), \end{aligned}$$

so, we obtain the following relation between $\boldsymbol{y}^*$ and $\boldsymbol{y}$; i.e.,

$$\boldsymbol{y}^* = \boldsymbol{y} + \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-1}\boldsymbol{Z}_2'\cdot\boldsymbol{\eta} + \boldsymbol{Z}_2\cdot(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0). \tag{2.12}$$

If we take as a new variate vector

$$\boldsymbol{\zeta} = \begin{pmatrix} \boldsymbol{\zeta}_1 \\ \boldsymbol{\zeta}_2 \\ \boldsymbol{\zeta}_3 \end{pmatrix},$$

where

$$\boldsymbol{\zeta}_1 \equiv \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_r \end{pmatrix} = (\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-\frac{1}{2}}\boldsymbol{Z}_1'\cdot\boldsymbol{\eta},\ \boldsymbol{\zeta}_2 \equiv \begin{pmatrix} \zeta_{r+1} \\ \vdots \\ \zeta_s \end{pmatrix} = (\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{Z}_2'\cdot\boldsymbol{\eta},\ \boldsymbol{\zeta}_3 \equiv \begin{pmatrix} \zeta_{s+1} \\ \vdots \\ \zeta_n \end{pmatrix} = \boldsymbol{U}\cdot\boldsymbol{\eta},$$

and the $(n-s)\times n$ matrix $\boldsymbol{U}$ is chosen such that the square matrix

$$\boldsymbol{C} = \begin{pmatrix} (\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-\frac{1}{2}}\boldsymbol{Z}_1' \\ (\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{Z}_2' \\ \boldsymbol{U} \end{pmatrix}$$

is an orthogonal one, then it is evident that the variance-convariance matrix of $\boldsymbol{\zeta}$ is $\sigma^2\boldsymbol{I}$; i.e.,

$$E(\boldsymbol{\zeta}\boldsymbol{\zeta}') = \sigma^2\boldsymbol{I}.$$

Since

$$\boldsymbol{\zeta} = \boldsymbol{C}\cdot\boldsymbol{\eta},$$

and the matrix $\boldsymbol{C}$ was orthogonal, we have

$$\boldsymbol{\eta} = \boldsymbol{C}'\cdot\boldsymbol{\zeta},$$

i.e.,

$$\boldsymbol{\eta} = \boldsymbol{Z}_1(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-\frac{1}{2}}\boldsymbol{\zeta}_1 + \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{\zeta}_2 + \boldsymbol{U}'\cdot\boldsymbol{\zeta}_3, \tag{2.13}$$

and it is clearly seen that the three summands on the right-hand side of the above equation are mutually independent in the stochastic sense.

From (2.5), (2.6), (2.9), (2.10) and (2.13), we have

$$\left.\begin{aligned} \boldsymbol{b} - \boldsymbol{\beta} &= (\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-\frac{1}{2}}\boldsymbol{\zeta}_1 + (\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{\zeta}_2, \\ \boldsymbol{y} &= \qquad\qquad \boldsymbol{U}'\cdot\boldsymbol{\zeta}_3. \\ \boldsymbol{b}_1^* - \boldsymbol{\beta}_1 &= (\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-\frac{1}{2}}\boldsymbol{\zeta}_1 \\ \boldsymbol{y}^* &= \qquad \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{\zeta}_3 + \boldsymbol{U}'\cdot\boldsymbol{\zeta}_3 + \boldsymbol{Z}_2\cdot(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0). \end{aligned}\right\} \tag{2.14}$$

The variates $\boldsymbol{y}$ and

$$\boldsymbol{y}^* - \boldsymbol{y} = \boldsymbol{Z}_2(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{\zeta}_2 + \boldsymbol{Z}_2\cdot(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0)$$

are mutually independent in the stochastic sence, and consequently $S_0 = \boldsymbol{y}'\boldsymbol{y}$ and $S_0^* - S_0 = \boldsymbol{y}^{*\prime}\boldsymbol{y}^* - \boldsymbol{y}'\boldsymbol{y}$ are mutually independent in the stochastic sense, because

$$S_0^* - S_0 = \boldsymbol{y}^{*\prime}\boldsymbol{y}^* - \boldsymbol{y}'\boldsymbol{y} = (\boldsymbol{y}^* - \boldsymbol{y})'(\boldsymbol{y}^* - \boldsymbol{y}).$$

The variate

$$\begin{aligned}\sigma^{-2}\cdot(S_0^* - S_0) &= \sigma^{-2}(\boldsymbol{y}^* - \boldsymbol{y})'(\boldsymbol{y}^* - \boldsymbol{y}) \\ &= \sigma^{-2}\boldsymbol{\zeta}_2'\boldsymbol{\zeta}_2 + (\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0)'(\boldsymbol{Z}_2'\boldsymbol{Z}_2)^{-\frac{1}{2}}\boldsymbol{\zeta}_2 + (\boldsymbol{\beta}_2 - \boldsymbol{\beta})'\boldsymbol{Z}_2'\boldsymbol{Z}_2\cdot(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0)\end{aligned}$$

is distributed according to the so-called "non-central chi-square distribution" with degrees of freedom $(s-r)$ and with parameter

$$\lambda = \frac{1}{2\sigma^2}(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0)'\boldsymbol{Z}'\boldsymbol{Z}\cdot(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0). \tag{2.16}$$

If the statistical hypothesis H_0 is true, then $\lambda = 0$, and consequently the distribution of the variate (2.15) is the ordinary chi-square distribution of degrees of freedom of $(s-r)$. The variate

$$\sigma^{-2}S_0 = \sigma^{-2}\cdot\boldsymbol{y}'\boldsymbol{y} = \sigma^{-2}\boldsymbol{\zeta}_3'\boldsymbol{\zeta}_3$$

is distributed according to the chi-square distribution of degrees of freedom of $(n-s)$. Consequeutly, the statistic

$$G = \frac{S_0^* - S_0}{S_0} = \frac{(\boldsymbol{y}^* - \boldsymbol{y})'(\boldsymbol{y}^* - \boldsymbol{y})}{\boldsymbol{y}'\boldsymbol{y}} \tag{2.17}$$

is distributed according to the G-distribution[10] of degrees of freedom of $(s-r, n-s)$ and with parameter λ defined in (2.16); i.e., its probability element is given by

$$f(G)dG = \sum_{\nu=0}^{\infty} e^{-\lambda}\frac{\lambda^\nu}{\nu!}\frac{\Gamma\left(\frac{n-r}{2}+\nu\right)}{\Gamma\left(\frac{s-r}{2}+\nu\right)\Gamma\left(\frac{n-s}{2}\right)}\left(\frac{G}{1+G}\right)^{\frac{s-r-2}{2}+\nu}\left(\frac{1}{1+G}\right)^{\frac{n-s}{2}+1} dG. \tag{2.18}$$

If we take the statistic

$$t = \frac{(\boldsymbol{y}^* - \boldsymbol{y})_i}{\sqrt{\boldsymbol{y}'\boldsymbol{y}}}\sqrt{n-s}, \tag{2.19}$$

where the notation $(\boldsymbol{y}^* - \boldsymbol{y})_i$ denotes the i-th component of the vector $\boldsymbol{y}^* - \boldsymbol{y}$, then its distribution is the so-called "non-central t-distribution"; i.e., its probability element is

$$f(t)dt = \frac{\left(\frac{n-s}{2}\right)^{\frac{n-s}{2}} e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}\,\Gamma\left(\frac{n-s}{2}\right)}\sum_{\nu=0}^{\infty}\frac{\Gamma\left(\frac{n-s+\nu-1}{2}\right)}{\nu!}(\delta t)^\nu\left(\frac{2}{n-s+t^2}\right)^{\frac{n-s+\nu+1}{2}} dt, \tag{2.20}$$

where

$$\delta = \frac{1}{\sigma}(\boldsymbol{Z}\cdot(\boldsymbol{\beta}_2 - \boldsymbol{\beta}_2^0))_i \tag{2.21}$$

If the hypothesis H_0 is true, then $\delta = 0$, and consequently the distribution of the statistic t is the ordinary t-distribution of degrees of freedom of $(n-s)$.

2. 2. The general case. When the orthogonality condition (2.10) does not hold, we can transform the fixed variables such that

$$\boldsymbol{Y} = (\boldsymbol{Y}_1\boldsymbol{Y}_2) = (\boldsymbol{Z}_1\boldsymbol{Z}_2)\cdot\begin{bmatrix} \boldsymbol{P} & \boldsymbol{Q} \\ 0 & \boldsymbol{R} \end{bmatrix},$$

where $\boldsymbol{P}$ and $\boldsymbol{R}$ are $r\times r$ and $(s-r)\times(s-r)$ non-singular square matrices respectively and $\boldsymbol{Q}$ is a $r\times(s-r)$ matrix and

$$\boldsymbol{Y}_1'\boldsymbol{Y}_2 = 0^{(11)} \tag{2.22}$$

Since

$$\boldsymbol{Y}_1 = \boldsymbol{Z}_1\boldsymbol{P} \qquad \boldsymbol{Y}_2 = \boldsymbol{Z}_1\boldsymbol{Q} + \boldsymbol{Z}_2\boldsymbol{R},$$

the condition (2.22) implies the relation

$$\boldsymbol{Z}_1'\boldsymbol{Z}_1\boldsymbol{Q} + \boldsymbol{Z}_1'\boldsymbol{Z}_2\boldsymbol{P} = 0\,. \tag{2.23}$$

Let

$$\boldsymbol{\alpha}_1 = \boldsymbol{P}^{-1}\cdot\boldsymbol{\beta}_1 - \boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{\beta}_2, \qquad \boldsymbol{\alpha}_2 = \boldsymbol{R}^{-1}\boldsymbol{\beta}_2, \tag{2.24}$$

and

$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{bmatrix},$$

then it follows that

$$\eta = \boldsymbol{x} - \boldsymbol{Y}_1\boldsymbol{\alpha}_1 - \boldsymbol{Y}_2\boldsymbol{\alpha}_2,$$

and the statistical hypothesis H_0; $\boldsymbol{\beta}_2 = \boldsymbol{\beta}_2^0$ is equivalent to the derived hypothesis $\tilde{H}_0 : \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_2^0$ $(\equiv \boldsymbol{R}^{-1}\cdot\boldsymbol{\beta}_2^0)$.

The least-square estimate $\boldsymbol{a}$ of $\boldsymbol{\alpha}$, and the least-square estimate $\boldsymbol{a}_1^*$ of $\boldsymbol{\alpha}_1$ under the hypothesis $\tilde{H}_0$ satisfy the following relations; i.e.

$$\boldsymbol{a} = \begin{bmatrix} \boldsymbol{P}^{-1} & -\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{R}^{-1} \\ 0 & \boldsymbol{R}^{-1} \end{bmatrix}\cdot\boldsymbol{b}, \tag{2.25}$$

and

$$\boldsymbol{a}_1^* = \boldsymbol{P}^{-1}\cdot\boldsymbol{b}_1^* - \boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{R}^{-1}\cdot\boldsymbol{\beta}_2^0, \tag{2.26}$$

If we put

$$\boldsymbol{K}=\begin{pmatrix}\boldsymbol{P} & \boldsymbol{Q}\\ 0 & \boldsymbol{R}\end{pmatrix},$$

it is easily seen that

$$\boldsymbol{K}^{-1}=\begin{pmatrix}\boldsymbol{P}^{-1} & -\boldsymbol{P}^{-1}\boldsymbol{Q}\boldsymbol{R}^{-1}\\ 0 & \boldsymbol{R}^{-1}\end{pmatrix}$$

In a completely similar manner as in 2.1, we have

$$\left.\begin{aligned}\boldsymbol{a}-\boldsymbol{\alpha}&=(\boldsymbol{Y}_1'\boldsymbol{Y}_1)^{-\frac{1}{2}}\cdot\tilde{\boldsymbol{\zeta}}_1+(\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_2,\\ \tilde{\boldsymbol{y}}&=\tilde{\boldsymbol{U}}'\cdot\tilde{\boldsymbol{\zeta}}_3\\ \boldsymbol{a}_1^*-\boldsymbol{\alpha}_1&=(\boldsymbol{Y}_1'\boldsymbol{Y}_1)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_1,\\ \tilde{\boldsymbol{y}}^*&=\boldsymbol{Y}_2(\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_2+\tilde{\boldsymbol{U}}'\cdot\tilde{\boldsymbol{\zeta}}_3+\boldsymbol{Y}_2\cdot(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_2^0)\end{aligned}\right\}\quad(2.27)$$

where

$$\tilde{\boldsymbol{\zeta}}=\begin{pmatrix}\tilde{\boldsymbol{\zeta}}_1\\ \tilde{\boldsymbol{\zeta}}_2\\ \tilde{\boldsymbol{\zeta}}_3\end{pmatrix}=\tilde{\boldsymbol{C}}\cdot\boldsymbol{\eta},$$

and the matrix $\tilde{\boldsymbol{U}}$ is so chosen that the square matrix

$$\tilde{\boldsymbol{C}}=\begin{pmatrix}(\boldsymbol{Y}_1'\boldsymbol{Y}_1)^{-\frac{1}{2}}\boldsymbol{Y}_1'\\ (\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\boldsymbol{Y}_2'\\ \tilde{\boldsymbol{U}}\end{pmatrix}=\begin{pmatrix}(\boldsymbol{Y}'\boldsymbol{Y})^{-\frac{1}{2}}\boldsymbol{Y}'\\ \tilde{\boldsymbol{U}}\end{pmatrix}$$

is an orthogonal one. Hence also in this case, we have

$$E(\tilde{\boldsymbol{\zeta}}\cdot\tilde{\boldsymbol{\zeta}}')=\sigma^2\boldsymbol{I}.$$

From (2.25), (2.26) and (2.27), we have

$$\boldsymbol{b}-\boldsymbol{\beta}=\begin{pmatrix}\boldsymbol{P} & \boldsymbol{Q}\\ 0 & \boldsymbol{R}\end{pmatrix}(\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\begin{pmatrix}\tilde{\boldsymbol{\zeta}}_1\\ \tilde{\boldsymbol{\zeta}}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}(\boldsymbol{Y}_1\boldsymbol{Y}_1)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_1+\boldsymbol{Q}(\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_2\\ \boldsymbol{R}(\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_2\end{pmatrix}\quad(2.28)$$

$$\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{Z}\cdot\boldsymbol{b}=\boldsymbol{x}-\boldsymbol{Y}\cdot\boldsymbol{a}=\tilde{\boldsymbol{y}}=\tilde{\boldsymbol{U}}'\cdot\tilde{\boldsymbol{\zeta}}_3\quad(2.29)$$

$$\begin{aligned}\boldsymbol{b}_1^*-\boldsymbol{\beta}_1&=\boldsymbol{P}\cdot\boldsymbol{a}_1^*+\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{\beta}_2^0-\boldsymbol{\beta}_1=\boldsymbol{P}\cdot(\boldsymbol{Y}_1'\boldsymbol{Y}_1)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_1-\boldsymbol{Q}\boldsymbol{R}^{-1}\cdot(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)\\ &=\boldsymbol{P}(\boldsymbol{P}'\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{P})^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_1+(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1'\boldsymbol{Z}_2\cdot(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)\end{aligned}\quad(2.30)$$

and

$$\begin{aligned}\boldsymbol{y}^*&=\boldsymbol{x}-\boldsymbol{Z}_1\boldsymbol{b}_1^*-\boldsymbol{Z}_2\boldsymbol{\beta}=\boldsymbol{x}-\boldsymbol{Y}_1\boldsymbol{a}_1^*-(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\boldsymbol{R})^{-1}\cdot\boldsymbol{\beta}_2^0\\ &=\boldsymbol{x}-\boldsymbol{Y}_1\boldsymbol{a}_1^*-\boldsymbol{Y}_2'\cdot\boldsymbol{\alpha}_2^0=\tilde{\boldsymbol{y}}^*\\ &=\boldsymbol{Y}_2(\boldsymbol{Y}_2'\boldsymbol{Y}_2)^{-\frac{1}{2}}\tilde{\boldsymbol{\zeta}}_2+\tilde{\boldsymbol{U}}'\cdot\boldsymbol{\zeta}_3+(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\boldsymbol{R})\boldsymbol{R}^{-1}\cdot(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0).\end{aligned}\quad(2.31)$$

The variate

$$\sigma^{-2}\cdot S_0=\sigma^{-2}\cdot\boldsymbol{y}'\boldsymbol{y}=\sigma^{-2}\cdot\tilde{\boldsymbol{\zeta}}_3'\tilde{\boldsymbol{\zeta}}_3$$

is distributed according to the chi-square distribution of degrees of freedom of $(n-s)$ and the variate

$$\begin{aligned}\sigma^{-2}(S_0^* - S_0) &= \sigma^{-2}(\tilde{\boldsymbol{y}}^{*\prime}\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{y}}'\tilde{\boldsymbol{y}}) = \sigma^{-2}(\tilde{\boldsymbol{y}}^* - \tilde{\boldsymbol{y}})'(\tilde{\boldsymbol{y}}^* - \tilde{\boldsymbol{y}}) \\ &= \sigma^{-2}\cdot[\tilde{\zeta}_2'\cdot\tilde{\zeta}_2 + (\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)'\cdot(\boldsymbol{Y}_2'\cdot\boldsymbol{Y}_2)^{-\frac{1}{2}}\cdot\boldsymbol{Y}_2'(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\boldsymbol{R})\cdot\boldsymbol{R}^{-1}\cdot\tilde{\zeta}_2 \\ &\quad +(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)'\boldsymbol{R}'^{-1}(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\cdot\boldsymbol{R})'(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\cdot\boldsymbol{R})\boldsymbol{R}^{-1}\cdot(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)]\end{aligned}$$

is stochastically independent of $\sigma^{-2}S_0$ and is distributed according to the non-central chi-square distribution of degrees of freedom of $(s-r)$ and with the parameter

$$\begin{aligned}\lambda &= \frac{1}{2\sigma^2}(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)'\boldsymbol{R}'^{-1}(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\boldsymbol{R})'(\boldsymbol{Z}_1\boldsymbol{Q}+\boldsymbol{Z}_2\boldsymbol{R})\boldsymbol{R}^{-1}(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0) \\ &= \frac{1}{2\sigma^2}(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0)'(\boldsymbol{Z}_1'\boldsymbol{Z}_1-\boldsymbol{Z}_2'\boldsymbol{Z}_1(\boldsymbol{Z}_1'\boldsymbol{Z}_1)^{-1}\boldsymbol{Z}_1'\boldsymbol{Z}_2)(\boldsymbol{\beta}_2-\boldsymbol{\beta}_2^0) \qquad (2.32)\end{aligned}$$

Hence the statistic

$$G = \frac{S_0^* - S_0}{S_0}$$

is distributed according to the G-distribution of degrees of freedom of $(s-r,\ n-s)$ and with the parameter λ given in (2.32).

§ 3. Multiple Correlation Coefficient And Inter-class Correlation Coefficient. In this section, we shall first derive the sampling distribution of the sample multiple correlation coefficient. The parent population under consideration is a non-singular k-dimensional normal population with means $m_1, \cdots, m_k$ and the moment matrix $\Lambda = (\lambda_{ij})$. The population multiple correlation coefficient of the first variate on the other is

$$\rho_{1(23\cdots k)} = \sqrt{1-\frac{\Lambda}{\lambda_{11}\Lambda_{11}}},$$

the corresponding sample multiple correlation coefficient, which we denote by the corresponding latin letter

$$r_{1(23\cdots k)}$$

is given as follows:

Let

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{21} & \cdots\cdots & x_{k1} \\ x_{12} & x_{22} & \cdots\cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & \cdots\cdots & x_{kn} \end{pmatrix} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots\cdots, \boldsymbol{x}_k)$$

be a random sample of size n, and further let it be

$$\bar{\boldsymbol{x}}_i = \bar{x}_i \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \bar{x}_i = \frac{1}{n}\sum_{\alpha=1}^{n} x_{i\alpha}, \quad i = 1, 2, \cdots, k,$$

$$\tilde{\boldsymbol{X}} = (\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1, \ \boldsymbol{x}_2 - \bar{\boldsymbol{x}}_2, \cdots\cdots, \boldsymbol{x}_k - \bar{\boldsymbol{x}}_k),$$

and

$$\boldsymbol{A} \equiv (a_{ij}) = \tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}},$$

then

$$r_{1(23\cdots k)} = \sqrt{1 - \frac{A}{a_{11}A_{11}}}$$

We shall use the notation

$$\rho = \rho_{1(23\cdots k)} \quad \text{and} \quad r = r_{1(23\cdots k)}$$

throughout this section, unless otherwise stated.

The sampling distribution of r in the special case when $\rho = 0$ is easily obtained and its probability element is

$$\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k-1}{2}\right)} \cdot (r^2)^{\frac{k-3}{2}}(1-r^2)^{\frac{n-k-2}{2}} dr^2, \text{[12]}$$

In the general case when $\rho \neq 0$, the derivations of the sampling distribution of r were given by R. A. Fisher[13], S. S Wilks[14], S. Nabeya[15], and P. A. P. Moran[16], and its probability element

$$\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k-1}{2}\right)}(1-\rho^2)^{\frac{n-1}{2}}(r^2)^{\frac{k-3}{2}}(1-r^2)^{\frac{n-k-2}{2}} \cdot \sum_{\nu=0}^{\infty}\frac{1}{\nu!} \cdot \frac{\Gamma\left(\frac{k-1}{2}\right)\Gamma^2\left(\frac{n-1}{2}+\nu\right)}{\Gamma^2\left(\frac{n-1}{2}\right)\Gamma\left(\frac{k-1}{2}+\nu\right)}$$
$$\times (\rho^2 r^2)^{\nu} \cdot dr^2, \tag{3.1}$$

or represened by means of the hypergeometric function

$$\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k-1}{2}\right)}(1-\rho^2)^{\frac{n-1}{2}}(r^2)^{\frac{k-3}{2}}(1-r^2)^{\frac{n-k-2}{2}} F\left(\frac{n-1}{2}, \frac{n-1}{2}, \frac{k-1}{2}; \rho^2 r^2\right) \cdot dr^2, \tag{3.1'}$$

where, of course,

$$F(a, b, c; x) = \sum_{\nu=0}^{\infty} \frac{a(a+1)\cdots(a+\nu-1)b(b+1)\cdots(b+\nu-1)}{\nu!\, c(c+1)\cdots(c+\nu-1)} x^{\nu}$$

We shall first mention a necessary lemma.

Lemma 3. 1.[17] The probability element of the variate

$$\lambda^* = \frac{\Lambda_{11}}{2\Lambda} \sum_{p,q=2}^{k} \frac{\Lambda_{1p}\Lambda_{1q}}{\Lambda_{11}^2} a_{pq} \tag{3.2}$$

is given by

$$\frac{1}{\left(\frac{\rho^2}{1-\rho^2}\right)^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} \cdot \exp\left(-\frac{\lambda^*}{\frac{\rho^2}{1-\rho^2}}\right) \cdot (\lambda^*)^{\frac{n-1}{2}-1} \cdot d\lambda^* \tag{3.3}$$

If we consider the conditional distribution of x_1 under the condition that the other variates $x_2, x_3, \cdots, x_k$ are fixed, then the variate

$$\eta_1 = \boldsymbol{x}_1 - \boldsymbol{\alpha}_1 - \beta_{12}\boldsymbol{x}_2 \cdots\cdots - \beta_{1k}\boldsymbol{x}_k,$$

where

$$\boldsymbol{\alpha}_1 = (m_1 - \beta_{12}m_2 \cdots\cdots - \beta_{1k}m_k) \cdot \mathbf{1}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

is a random sample of size n from the population

$$N(0, \sigma_{1\cdot 23\ldots k}^2)$$

The least-square estimate

$$\boldsymbol{b}_1 = \begin{pmatrix} b_{12} \\ b_{13} \\ \vdots \\ b_{1k} \end{pmatrix}, \quad \text{of} \quad \boldsymbol{\beta}_1 = \begin{pmatrix} \beta_{12} \\ \beta_{13} \\ \vdots \\ \beta_{1k} \end{pmatrix}$$

is given by

$$\boldsymbol{b}_1 = (\tilde{\boldsymbol{X}}_1'\tilde{\boldsymbol{X}}_1)^{-1}\tilde{\boldsymbol{X}}_1 \cdot (\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1),$$

were the matrix $\tilde{\boldsymbol{X}}_1$ is generated from $\tilde{\boldsymbol{X}}$ by omitting its first column. The conditional expected value of $\boldsymbol{A}_{11}^{\frac{1}{2}}\boldsymbol{b}_1$ is $\boldsymbol{A}_{11}^{\frac{1}{2}}\beta_1$, and the variate

$$q_1 = \sigma_{1.23\ldots k}^{-2}(\boldsymbol{A}_{11}^{\frac{1}{2}}\boldsymbol{b}_1)'(\boldsymbol{A}_{11}^{\frac{1}{2}}\boldsymbol{b}_1) = \sigma_{1.23\ldots k}^{-2}(\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1)' \cdot \tilde{\boldsymbol{X}}_1'\boldsymbol{A}_{11}\tilde{\boldsymbol{X}}_1 \cdot (\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1) \tag{3.4}$$

is distributed according to the non-central chi-square distribution of degrees of freedom of $(k-1)$ and with the parameter

$$\lambda^* = \frac{1}{2\sigma_{1\cdot 23\cdots k}^2}\boldsymbol{\beta}_1'\boldsymbol{A}_{11}\boldsymbol{\beta}_1 = \frac{\Lambda_{11}}{2\Lambda} \sum_{p,q=2}^{k} \frac{\Lambda_{1p}\Lambda_{1q}}{\Lambda_{11}} a_{pq}. \tag{3.5}$$

The variate

$$q_2 = \sigma_{1\cdot 23\cdots k}^{-2}(\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1 - \tilde{\boldsymbol{X}}_1\boldsymbol{b}_1)'(\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1 - \tilde{\boldsymbol{X}}_1\boldsymbol{b}_1)$$

is distributed according to the chi-square distribution of degrees of freedom of $(n-k)$, and further the two variates q_1 and q_2 are mutually independent in the stochastic sense. Consequently, the conditional distribution of the statistic

$$G = \frac{q_1}{q_2} = \frac{(\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1)' \cdot \tilde{\boldsymbol{X}}_1' \boldsymbol{A}_{11} \tilde{\boldsymbol{X}}_1 (\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1)}{a_{11} + (\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1)' \tilde{\boldsymbol{X}}_1' \tilde{\boldsymbol{X}}_1 \cdot (\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1)} = \frac{r^2}{1-r^2} \tag{3.6}$$

is the G-distribution of degrees of freedom of $(k-1, n-k)$, and with the parameter λ^* given by (3.5). Hence its conditional probability element is

$$\sum_{\nu=0}^{\infty} e^{-\lambda^*} \frac{\lambda^{*\nu}}{\nu!} (r^2)^{\frac{k-1}{2}} (1-r^2)^{\frac{n-k-2}{2}} (r^2)^{\nu} \frac{\Gamma\left(\frac{n-1}{2}+\nu\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k-1}{2}+\nu\right)} dr^2 . \tag{3.7}$$

The absolute probability element of r^2 is obtained by integrating out from the product of (3.7) and (3.3); i.e.,

$$\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k-1}{2}\right)} (1-\rho^2)^{\frac{n-1}{2}} (r^2)^{\frac{k-2}{2}} (1-r^2)^{\frac{n-k-2}{2}} \sum_{\nu=0}^{\infty} \frac{\Gamma^2\left(\frac{n-1}{2}+\nu\right)\Gamma\left(\frac{k-1}{2}\right)}{\nu!\,\Gamma^2\left(\frac{n-1}{2}\right)\Gamma\left(\frac{k-1}{2}+\nu\right)} \times$$

$$(\rho^2 r^2)^{\nu} \cdot dr^2 , \tag{3.8}$$

which was to be proved.

If we put $k=2$ in (3.8), we get

$$\frac{2^{n-2}}{\pi \cdot \Gamma(n-2)} (1-\rho^2)^{\frac{n-1}{2}} (1-r^2)^{\frac{n-4}{2}} \cdot \sum_{\nu=0}^{\infty} \frac{\Gamma^2\left(\frac{n+2\nu-1}{2}\right)}{(2\nu)!} (2\rho r)^{2\nu} \cdot dr . \tag{3.9}$$

which is the probability element for the absolute value of the inter-class correlation coefficient. If we wish to get the probability element of the inter-class correlation coefficient itself, we must go as follows:

Let it be

$$a_{11} = ns_1^2 , \ a_{22} = ns_2^2 \quad \text{and} \quad a_{12} = nm_{12}$$

$$u_{2\cdot 1} = s_1(b_{21} - \beta_{21}) , \quad \beta_{21} = \rho \frac{\sigma_2}{\sigma_1} ,$$

then consider the conditional distribution of $s_1 b_{21}/u_{2\cdot 1}$ under the condition that the first variate x_1 is fixed. The conditional distribution of the statistic

$$t=\frac{s_1 b_{21}}{\sqrt{s_2^2(1-r^2)}}\sqrt{n-2}=\frac{r}{\sqrt{1-r^2}}\sqrt{n-2} \tag{3.10}$$

is the non-central t-distribution of degrees of freedom of $(n-2)$ and with the parameter

$$\delta=\frac{s_1}{\sigma_1}\frac{\rho}{\sqrt{1-\rho^2}}, \tag{3.11}$$

i.e.,

$$\frac{\left(\frac{n-1}{2}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}\cdot\frac{\rho^2}{1-\rho^2}\left(\frac{s_1}{\sigma_1}\right)^2}}{\sqrt{2\pi}\,\Gamma\left(\frac{n-2}{2}\right)}\sum_{\nu=0}^{\infty}\frac{\Gamma\left(\frac{n-1+\nu}{2}\right)}{\nu!}\left(\frac{s_1}{\sigma_1}\right)^{\nu}\left(\frac{\rho^2}{1-\rho^2}\right)^{\frac{\nu}{2}}\left(\frac{2}{n-2+t^2}\right)^{\frac{n-1+\nu}{2}} t^{\nu}dt. \tag{3. 12}$$

Since the probability element of s_1^2/σ_1^2 is

$$\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}e^{-\frac{s_1^2}{2\sigma_1^2}}\left(\frac{s_1^2}{\sigma_1^2}\right)^{\frac{n-1}{2}-1}d\left(\frac{s_1^2}{\sigma_1^2}\right).$$

the absolute distribution of r is given by

$$\frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}(1-\rho^2)^{\frac{n-1}{2}}(1-r^2)^{\frac{n-4}{2}}\sum_{\nu=0}^{\infty}\frac{\Gamma^2\left(\frac{n-1+\nu}{2}\right)}{\nu!}(2\rho r)^{\nu}\cdot dr$$

$$=\frac{2^{n-3}}{\pi\cdot\Gamma(n-1)}(1-\rho^2)^{\frac{n-1}{2}}(1-r^2)^{\frac{n-4}{2}}\cdot\sum_{\nu=0}^{\infty}\frac{\Gamma^2\left(\frac{n+\nu-1}{2}\right)}{\nu!}(2\rho r)^{\nu}\cdot dr, \tag{3.13}$$

which was to be derived.

§ 4. **Hotelling's T^2.**(18) Let

$$\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k,1),\quad \boldsymbol{X}^*=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k),$$

and

$$\boldsymbol{A}=\boldsymbol{X}'\boldsymbol{X},\qquad \boldsymbol{A}^*=\boldsymbol{X}^{*\prime}\boldsymbol{X}^*,$$

then the corresponding determinants are

$$A=n\left|\sum_{\alpha=1}^{n}x_{i\alpha}x_{j\alpha}-n\bar{x}_i\bar{x}_j\right|,\qquad A^*=\left|\sum_{\alpha=1}^{n}x_{i\alpha}x_{j\alpha}\right|,$$

As is well-known, Prof. H. Hotelling's square of the generalized Student's ratio T^2 is given by the equation

$$\frac{n}{1+\frac{T^2}{n-1}}=\frac{A}{A^*}, \tag{4.1}$$

hence we have

$$T^2 = n(n-1)\cdot \sum_{i,j=1}^{k} \frac{A_{ij}}{A}\bar{x}_i\bar{x}_j .$$

Decompose A and A^* into the products of residual variances, then we have

$$A = n^{k+1}s^2_{1\cdot 23\cdots k}\cdot s^2_{2\cdot 34\cdots k} \cdots\cdots s^2_{k-1\cdot k}\cdot s^2_k , \tag{4.2}$$

$$A^* = n^k\cdot s^{*2}_{1\cdot 23\cdots k}\cdot s^{*2}_{2\cdot 34\cdots k} \cdots\cdots s^{*2}_{k-1\cdot k}\cdot s^{*2}_k \tag{4.3}$$

where, for instance, $s^2_{1\cdot 23\cdots k}$ is sample residual variance of x_1 when the parameter $\alpha_1 = m_1 - \beta_{12}m_2 \cdots - \beta_{1k}m_k$ and $\beta_{12}, \cdots, \beta_{1k}$ are all estimated, and $s^{*2}_{1\cdot 23\cdots k}$ is the sample residual variance of x_1 under the statistical hypothesis H_1: $\alpha_1 = 0$. Hence we have

$$\frac{A}{A^*} = n\cdot \frac{\dfrac{ns^2_{1.23\cdots k}}{\sigma^2_{1\cdot 23\cdots k}}}{\dfrac{ns^{*2}_{1.23\cdots k}}{\sigma^2_{1\cdot 23\cdots k}}}\cdot \frac{\dfrac{ns^2_{2.34\cdots k}}{\sigma^2_{2.34\cdots k}}}{\dfrac{ns^{*2}_{2.34\cdots k}}{\sigma^2_{2\cdot 34\cdots k}}} \cdots\cdots \frac{\dfrac{ns^2_{k-1\cdot k}}{\sigma^2_{k-1\cdot k}}}{\dfrac{ns^{*2}_{k-1\cdot k}}{\sigma^2_{k-1\cdot k}}}\cdot \frac{\dfrac{ns^2_k}{\sigma^2_k}}{\dfrac{ns^{*2}}{\sigma^2_k}}$$

$$= n\cdot \frac{q_{1\cdot 23\cdots k}}{q^*_{1\cdot 23\cdots k}}\cdot \frac{q_{2\cdot 34\cdots k}}{q^*_{2\cdot 34\cdots k}} \cdots\cdots \frac{q_{k-1\cdot k}}{q^*_{k-1\cdot k}}\cdot \frac{q_k}{q^*_k}, \quad \text{(say)}. \tag{4.4}$$

4.1. We shall first consider the special case when

$$m_1 = m_2 = \cdots\cdots = m_{k-1} = 0 ,$$
$$\lambda_{1k} = \lambda_{2k} = \cdots\cdots = \lambda_{k-1\cdot k} = 0 ,$$

i.e., the k-th variate x_k is uncorrelated with others all of which have zero means.

If we consider the conditional distribution of x_1 under the condition that the other variates x_2, ……, x_k are fixed, then

$$E(x_1) = \beta_{12}x_2 + \cdots + \beta_{1k}x_k + \alpha_1 , \quad \alpha_1 = m_1 - \beta_{12}m_2 \cdots\cdots - \beta_{1k}m_k ,$$

and in this case it is clear that

$$\beta_{1k} = 0, \quad \alpha_1 = 0 ,$$

hence the statistic

$$\frac{q^*_{1\cdot 23k\cdots k} - q_{1\cdot 23\cdots k}}{q_{1\cdot 23\cdots k}}$$

is distributed according to the G-distribution of degrees of freedom of $(1, n-k)$ and with the parameter

$$\lambda_1 = \frac{1}{2\sigma^2_{12\cdot 3\cdots k}}\alpha_1(n - (n\bar{x}_2 \cdots n\bar{x}_k)A^{-1}_{11}\begin{pmatrix} n\bar{x}_2 \\ \vdots \\ n\bar{x}_k \end{pmatrix})\alpha_1 = 0 ,$$

i.e., for any $x > 0$,

$$P\left(\frac{q^*_{1\cdot 23\cdots k}-q_{1\cdot 23\cdots k}}{q_{1\cdot 23\cdots k}}<x\right)=\frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int_0^x\left(\frac{u}{1+u}\right)^{\frac{1}{2}-1}\left(\frac{1}{1+u}\right)^{\frac{n-k+2}{2}}du,$$

therefore, we have

$$P\left(\frac{q_{1\cdot 23\cdots k}}{q^*_{1\cdot 23\cdots k}}<x\right)=\frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int_{\frac{1-x}{x}}^{\infty}\left(\frac{u}{1+u}\right)^{\frac{1}{2}-1}\left(\frac{1}{1+u}\right)^{\frac{n-k+2}{2}}du,$$

and consequently we have the probability element

$$\frac{\Gamma\left(\frac{n-k+1}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{1}{2}\right)}x^{\frac{n-k}{2}-1}(1-x)^{\frac{1}{2}-1}dx \tag{4.5}$$

for the statistic $q_{1\cdot 23\cdots k}/q^*_{1\cdot 23\cdots k}$. It is evident from (4.5) that the distribution of the statistic $q_{12\cdot 3\cdots k}/q^*_{1\cdot 23\cdots k}$ is independent of the conditioning variates $x_2, \ldots\ldots, x_k$, and consequently $q_{1\cdot 23\cdots k}/q^*_{1\cdot 23\cdots k}$ is independent of the other statistics

$$\frac{q_{2\cdot 34\cdots k}}{q^*_{2\cdot 34\cdots k}}, \ldots\ldots, \frac{q_{k-1\cdot k}}{q^*_{k-1\cdot k}}, \frac{q_k}{q^*_k}$$

in the absolute sense.

If we denote the Beta-variate which has the density

$$\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\cdot x^{p-1}(1-x)^{q-1}, \quad \text{for } 0<x<1$$

by B_{pq}, then the following lemma is obtained.

Lemma 4.1. If the two Beta-variates $B_{\mu-k\cdot k}$ and $B_{\mu\cdot\nu}$ are mutually independent in the stochastic sense, then

$$B_{\mu-k\cdot k}B_{\mu\cdot\nu}=B_{\mu-k\cdot k+\nu}. \tag{4.6}$$

(4.5) is equivalent to the equation

$$\frac{q_{1\cdot 23\cdots k}}{q^*_{1\cdot 23\cdots k}}=B_{\frac{n-k}{2}\cdot\frac{1}{2}}. \tag{4.7}$$

Next, if we consider the conditional distribution of x_2 under the condition that $x_3, \ldots\ldots, x_k$ are fixed, then

$$E(x_2)=\beta_{23}x_3+\cdots+\beta_{2k}x_k+\alpha_2, \quad \alpha_2=m_2-\beta_{23}m_3\cdots-\beta_{2k}m_k,$$

and in this special case $\alpha_2=0$, because $m_2=m_3=\cdots=m_{k-1}=0$ and

$\beta_{2k}=0$. Hence, in a completely similar manner as above, we have

$$\frac{q_{2\cdot 34\cdots k}}{q^*_{2\cdot 34\cdots k}}=\mathrm{B}_{\frac{n-k+1}{2}\cdot\frac{1}{2}}\,.$$

and it is independent of all the remaining statistics in the stochastic sense.

We proceed as above and get

$$\left.\begin{aligned}\frac{q_{3\cdot 45\cdots k}}{q^*_{3\cdot 45\cdots k}}&=\mathrm{B}_{\frac{n-k+2}{2}\cdot\frac{1}{2}}\\ &\cdots\cdots\cdots\cdots\\ \frac{q_{k-1\cdot k}}{q^*_{k-1\cdot k}}&=\mathrm{B}_{\frac{n-2}{2}\cdot\frac{1}{2}}\end{aligned}\right\}\tag{4.9}$$

Finally the statistic

$$(q^*_k-q_k)/q_k$$

is distribution according to the G-distribution of degrees of freedom of $(1,\ n-1)$ and with the parameter

$$\lambda=\frac{1}{2\lambda_{kk}}\,nm_k^2\,,\tag{4.10}$$

hence the probability element of q_k/q^*_k is

$$\sum_{\nu=0}^{\infty}e^{-\lambda}\frac{\lambda^\nu}{\nu\,!}\cdot\frac{\Gamma\left(\frac{n}{2}+\nu\right)}{\Gamma\left(\frac{1}{2}+\nu\right)\Gamma\left(\frac{n-1}{2}\right)}\,x^{\frac{n-1}{2}-1}(1-x)^{\frac{1}{2}+\nu-1}\,dx\,,\tag{4.11}$$

we shall denote the statistic q_k/q^*_k symbolically as follows:

$$\frac{q_k}{q^*_k}=\sum_{\nu=0}^{\infty}e^{-\lambda}\frac{\lambda^\nu}{\nu\,!}\mathrm{B}_{\frac{n-1}{2}\cdot\frac{1}{2}+\nu}\,.\tag{4.12}$$

It follows from (4.1), (4.4), (4.7), (4.8), (4.9) and (4.12) that

$$\frac{1}{1+\dfrac{T^2}{n-1}}=\mathrm{B}_{\frac{n-k}{2}\cdot\frac{1}{2}}\cdot\mathrm{B}_{\frac{n-k+1}{2}\cdot\frac{1}{2}}\cdots\cdots\mathrm{B}_{\frac{n-2}{2}\cdot\frac{1}{2}}\cdot\sum_{\nu=0}^{\infty}e^{-\lambda}\frac{\lambda^\nu}{\nu\,!}\mathrm{B}_{\frac{n-1}{2}\cdot\frac{1}{2}+\nu}$$

$$=\sum_{\nu=0}^{\infty}e^{-\lambda}\frac{\lambda^\nu}{\nu\,!}\mathrm{B}_{\frac{n-k}{2}\cdot\frac{1}{2}}\cdots\cdots\mathrm{B}_{\frac{n-k+1}{2}\cdot\frac{1}{2}}\,\mathrm{B}_{\frac{n-2}{2}\cdot\frac{1}{2}}\,\mathrm{B}_{\frac{n-1}{2}\cdot\frac{1}{2}+\nu}\,.\tag{4.13}$$

The $(\nu+1)$-th term of (4.13) will be reduced by means of Lemma 4.1 to $\mathrm{B}_{\frac{n-k}{2}\cdot\frac{k}{2}+\nu}$, whence we get the probability element

$$\sum_{\nu=0}^{\infty}e^{-\lambda}\frac{\lambda^\nu}{\nu\,!}\frac{\Gamma\left(\frac{n}{2}+\nu\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k}{2}+\nu\right)}\cdot\left(\frac{T^2}{n-1}\right)^{\frac{k}{2}+\nu-1}\left(1+\frac{T^2}{n-1}\right)^{-\frac{n}{2}-\nu}d\left(\frac{T^2}{n-1}\right)\tag{4.14}$$

for the statistic $T^2/(n-1)$, where, of course,

$$\lambda = \frac{n}{2\lambda_{kk}} \cdot m_k^2 .$$

4.2. *The general case.* If we transform the original variate

$$\mathfrak{x} = (x_1, x_2, \dots, x_k)$$

into

$$\mathfrak{y} = (y_1, y_2, \dots, y_k)$$

by the relation

$$\mathfrak{y} = \mathfrak{x}\Lambda^{-\frac{1}{2}} , \tag{4.15}$$

then the mean value vector

$$\bar{\mathfrak{m}} = (\bar{m}_1, \bar{m}_2, \dots, \bar{m}_k)$$

of $\mathfrak{y}$ is obtained by

$$\bar{\mathfrak{m}} = \mathfrak{m}\Lambda^{-\frac{1}{2}} . \tag{4.16}$$

and k variates $y_1, y_2, \dots, y_k$ are mutually independent in the stochastic sense.

Let it be

$$W = \begin{pmatrix} \sqrt{\dfrac{\bar{m}_2^2+\cdots+\bar{m}_k^2}{\bar{m}_1^2+\bar{m}_2^2+\cdots+\bar{m}_k^2}} & 0 \cdots\cdots & \cdots\cdots 0 & \dfrac{\bar{m}_1}{\sqrt{\bar{m}_1^2+\cdots+\bar{m}_k^2}} \\ \dfrac{-\bar{m}_1\bar{m}_2}{\sqrt{(\bar{m}_1^2+\cdots+\bar{m}_k^2)(\bar{m}_2^2+\cdots+\bar{m}_k^2)}} & \sqrt{\dfrac{\bar{m}_3^2+\cdots+\bar{m}_k^2}{\sqrt{\bar{m}_2^2+\bar{m}_3^2+\cdots+\bar{m}_k^2}}} \cdots\cdots & \cdots\cdots 0 & \dfrac{\bar{m}_2}{\sqrt{\bar{m}_1^2+\cdots+\bar{m}_k^2}} \\ \dfrac{-\bar{m}_1\bar{m}_3}{\sqrt{\bar{m}_1^2+\cdots+\bar{m}_k^2)(\bar{m}_2^2+\cdots+\bar{m}_k^2)}} & \dfrac{-\bar{m}_2\bar{m}_3}{\sqrt{(\bar{m}_2^2+\cdots+\bar{m}_k^2)(\bar{m}_3^2+\cdots+\bar{m}_k^2)}} \cdots & \cdots\cdots 0 & \dfrac{\bar{m}_3}{\sqrt{\bar{m}_1^2+\cdots+\bar{m}_k^2}} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{-\bar{m}_1\bar{m}_{k-1}}{\sqrt{(\bar{m}_1^2+\cdots+\bar{m}_k^2)(\bar{m}_2^2+\cdots+\bar{m}_k^2)}} & \dfrac{-\bar{m}_2\bar{m}_{k-1}}{\sqrt{\bar{m}_2^2+\cdots+\bar{m}_k^2)(\bar{m}_3^2+\cdots+\bar{m}_k^2)}} \cdots & \sqrt{\dfrac{\bar{m}_k^2}{\bar{m}_{k-1}^2+\bar{m}_k^2}} & \dfrac{\bar{m}_{k-1}}{\sqrt{\bar{m}_1^2+\cdots+\bar{m}_k^2}} \\ \dfrac{-\bar{m}_1\bar{m}_k}{\sqrt{(\bar{m}_1^2+\cdots+\bar{m}_k^2)(\bar{m}_2^2+\cdots+\bar{m}_k^2)}} & \dfrac{-\bar{m}_2\bar{m}_k}{\sqrt{\bar{m}_2^2+\cdots+\bar{m}_k^2)(\bar{m}_3^2+\cdots+\bar{m}_k^2)}} \cdots & \dfrac{-\bar{m}_{k-1}\bar{m}_k}{\sqrt{(\bar{m}_{k-1}^2+\bar{m}_k^2)\bar{m}_k^2}} & \dfrac{\bar{m}_k}{\sqrt{\bar{m}_1^2+\cdots+\bar{m}_k^2}} \end{pmatrix} \tag{4.17}$$

and put

$$\mathfrak{z}=(z_1, z_2, \ldots, z_k)=\mathfrak{y}\cdot \boldsymbol{W}, \tag{4.18}$$

then it is easily seen that

$$E(z_1)=E(z_2)=\cdots=E(z_{k-1})=0.$$

and since $\boldsymbol{W}$ is an orthogonal matrix, we have

$$E\{(\mathfrak{z}-E(\mathfrak{z}))'(\mathfrak{z}-E(\mathfrak{z}))\}=\boldsymbol{W}'\cdot E\{(\mathfrak{y}-\overline{\mathfrak{m}})'\cdot(\mathfrak{y}-\overline{\mathfrak{m}})\}\boldsymbol{W}=\boldsymbol{I},$$

i.e., k variates $z_1, \ldots, z_k$ are uncorrelated with each other.

Since

$$\boldsymbol{Z}^*=\begin{pmatrix} z_{11} & z_{21} & \cdots & z_{k1} \\ z_{12} & z_{22} & \cdots & z_{k2} \\ \vdots & \vdots & & \vdots \\ z_{1n} & z_{2n} & \cdots & z_{kn} \end{pmatrix}=\boldsymbol{X}^*\Lambda^{-\frac{1}{2}}\boldsymbol{W},$$

$$\boldsymbol{B}^*=\boldsymbol{Z}^{*\prime}\boldsymbol{Z}^*=\boldsymbol{W}'\Lambda^{-\frac{1}{2}}\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\Lambda^{-\frac{1}{2}}\boldsymbol{W}=\boldsymbol{W}'\Lambda^{-\frac{1}{2}}\boldsymbol{A}^*\Lambda^{-\frac{1}{2}}\boldsymbol{W},$$

it follows that

$$B^*=A^*\Lambda^{-1},$$

and since

$$(\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_k)=(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)\Lambda^{-\frac{1}{2}}\boldsymbol{W},$$

and

$$\boldsymbol{B}^*-n\begin{pmatrix} \bar{z}_1 \\ \bar{z}_2 \\ \vdots \\ \bar{z}_k \end{pmatrix}(\bar{z}_1, \bar{z}_2 \cdots \bar{z}_k)$$

$$=\boldsymbol{W}'\Lambda^{-\frac{1}{2}}\left[\boldsymbol{A}^*-n\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_k \end{pmatrix}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)\right]\Lambda^{-\frac{1}{2}}\boldsymbol{W},$$

we have

$$B=\det(\boldsymbol{Z}'\boldsymbol{Z})=A\Lambda^{-1},$$

where

$$\boldsymbol{Z}=(\boldsymbol{Z}^*\boldsymbol{1})=\begin{pmatrix} z_{11} & z_{21} & \cdots & z_{k1} & 1 \\ z_{12} & z_{22} & \cdots & z_{k2} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ z_{1n} & z_{2n} & \cdots & z_{kn} & 1 \end{pmatrix}$$

Hence we get the relation

$$\frac{n}{1+\dfrac{T^2}{n-1}}=\frac{A}{A^*}=\frac{B}{B^*} \tag{4.19}$$

Applying the arguments of the special case to the transformed variate $\mathfrak{z}$, we have the probability element of $T^2/(n-1)$, which is given by (4.14) with the parameters

$$\lambda_{kk}=(\boldsymbol{W}'\Lambda^{-\frac{1}{2}}\Lambda\Lambda^{-\frac{1}{2}}\boldsymbol{W})_{kk}=1\,,$$

and

$$m_k^2=\mathrm{m}\Lambda^{-\frac{1}{2}}\boldsymbol{W}'\boldsymbol{W}\Lambda^{-\frac{1}{2}}\mathrm{m}'=\mathrm{m}\Lambda^{-1}\mathrm{m}'=\sum_{i,j=1}^{k}\frac{\Lambda_{ij}}{\Lambda}m_i m_j\,.$$

If we use the notation

$$\psi^2=\frac{n}{2}\cdot\sum_{i,j=1}^{k}\frac{\Lambda_{ij}}{\Lambda}m_i m_j\,, \tag{4.20}$$

the probability element of the statistic $T^2/(n-1)$ becomes

$$e^{-\psi^2}\sum_{\nu=0}^{\infty}\frac{(\psi^2)^\nu}{\nu\,!}\;\frac{\Gamma\left(\frac{n}{2}+\nu\right)}{\Gamma\left(\frac{n-k}{2}\right)\Gamma\left(\frac{k}{2}+\nu\right)}\cdot\left(\frac{T^2}{n-1}\right)^{\frac{k}{2}+\nu-1}\left(1+\frac{T^2}{n-1}\right)^{-\frac{n}{2}-\nu}d\left(\frac{T^2}{n-1}\right), \tag{4.12}$$

as was to be proved.

§5. Normal Regression Theory in Multivariate Case. In this section, we shall consider the normal regression theory in the multivariate case.

Let

$$\boldsymbol{Z}=\begin{pmatrix} z_{11} & z_{21} & \cdots & z_{s1} \\ z_{12} & z_{22} & \cdots & z_{s2} \\ \vdots & \vdots & & \vdots \\ z_{1n} & z_{2n} & \cdots & z_{sn} \end{pmatrix} \tag{5.1}$$

be a $n\times s$ matrix of fixed variables of rank s, and

$$\boldsymbol{B}=\begin{pmatrix} \beta_{11} & \beta_{21} & \cdots & \beta_{k1} \\ \beta_{12} & \beta_{22} & \cdots & \beta_{k2} \\ \vdots & \vdots & & \vdots \\ \beta_{1s} & \beta_{2s} & \cdots & \beta_{ks} \end{pmatrix}\equiv(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_2,\ \ldots,\ \boldsymbol{\beta}_k) \tag{5.2}$$

be a $s\times k$ matrix of unknown parameters. Further let it be

$$(E(\boldsymbol{x}_1),\ E(\boldsymbol{x}_2),\ \cdots,\ E(\boldsymbol{x}_k))=\boldsymbol{Z}\mathrm{B}\,, \tag{5.3}$$

and k vector variates

$$\boldsymbol{\eta}_i=\boldsymbol{x}_i-\boldsymbol{Z}\cdot\boldsymbol{\beta}_i,\quad i=1,\ 2,\ \cdots,\ k \tag{5.4}$$

are distributed according to yhe non-singular k-dimensional normal distribution of mean vector $\mathbf{0}$ and with variance-covariance matrix $\boldsymbol{\Lambda}$.

Denote the least-square estimate of β_i, which minimizes

$$S_i = \boldsymbol{\eta}_i'\boldsymbol{\eta}_i = (\boldsymbol{x}_i - \boldsymbol{Z}\boldsymbol{\beta}_i)'(\boldsymbol{x}_i - \boldsymbol{Z}\boldsymbol{\beta}_i), \quad i=1, 2, \dots, k \tag{5.5}$$

by $\boldsymbol{b}_i$, $i=1, 2, \dots, k$, then it is seen that

$$\boldsymbol{b}_i = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\cdot\boldsymbol{x}_i, \quad i=1, 2, \dots, k,$$

or represented in terms of the error variates, we have

$$\boldsymbol{b}_i - \boldsymbol{\beta}_i = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\cdot\boldsymbol{\eta}_i, \quad i=1, 2, \dots, k. \tag{5.6}$$

The residual $\boldsymbol{y}_i$ for $\boldsymbol{\eta}_i$ is

$$\boldsymbol{y}_i = \boldsymbol{x}_i - \boldsymbol{Z}_i\boldsymbol{b}_i = (\boldsymbol{I} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}')\cdot\boldsymbol{x}_i = (\boldsymbol{I} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}')\cdot\boldsymbol{\eta}_i$$
$$i=1,2, \dots, k \tag{5.7}$$

Hence $\boldsymbol{b}_i$ and $\boldsymbol{y}_i$ are mutually independent in the stochastic sense.

Rows of the $s\times n$ matrix $(\boldsymbol{Z}'\boldsymbol{Z})^{-\frac{1}{2}}\boldsymbol{Z}'$ satisfy the orthogonality relations, and therefore we can choose a $(n-s)\times n$ matrix $\boldsymbol{U}$ so that the square matrix

$$\boldsymbol{P} = \begin{pmatrix} (\boldsymbol{Z}'\boldsymbol{Z})^{-\frac{1}{2}}\boldsymbol{Z}' \\ \boldsymbol{U} \end{pmatrix} \tag{5.8}$$

becomes an orthogonal one. Let it be

$$\boldsymbol{\xi}_i = \begin{pmatrix} {}_1\boldsymbol{\xi}_i \\ {}_2\boldsymbol{\xi}_i \end{pmatrix} \equiv \begin{pmatrix} \xi_{i1} \\ \vdots \\ \xi_{is} \\ \xi_{i,s+1} \\ \vdots \\ \xi_{in} \end{pmatrix} = \boldsymbol{P}\,\boldsymbol{\eta}_i, \quad i=1, 2, \dots, k \tag{5.9}$$

then it follows that

$$\boldsymbol{\eta}_i = \boldsymbol{P}'\boldsymbol{\xi}_i = \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-\frac{1}{2}}\cdot{}_1\xi_i + \boldsymbol{U}'\cdot{}_2\xi_i\,.$$

Since $(\boldsymbol{Z}'\boldsymbol{Z})^{-\frac{1}{2}}\boldsymbol{Z}'\cdot\boldsymbol{U}' = \boldsymbol{O}$, it is easily seen that

$$\boldsymbol{y}_i = \boldsymbol{U}'\cdot{}_2\boldsymbol{\xi}_i, \quad i=1,2, \dots, k \tag{5.10}$$

The variance-covariance matrix of vector variates

$$(\xi_{1p}, \xi_{2p}, \dots \xi_{kp}), \quad p=s+1, s+2, \dots, n,$$

can be calculated as follows: Let

$$\boldsymbol{U} = (u_{p\nu}),$$

then we have

$$E(\xi_{ip}\xi_{jp}) = E\{(\sum_{\nu=1}^{n} u_{p\nu}\eta_{i\nu})(\sum_{\mu=1}^{n} u_{p\mu}\eta_{j\mu})\}$$
$$= \sum_{\nu=1}^{n}\sum_{\mu=1}^{n} u_{p\nu}u_{j\mu}\cdot E(\eta_{i\nu}\eta_{j\mu}) = \lambda_{ij}\delta_{\nu\mu}\sum_{\nu,\mu} u_{p\nu}u_{p\mu} = \lambda_{ij},$$

therefore the required variance-covariance matrix is Λ for all p. Hence, if the condition

$$n-s>k$$

is satisfied, then the joint distribution of

$$Y=(y_i'y_j)=({}_2\xi_i'{}_2\xi_j) \tag{5.11}$$

has the probability element

$$\frac{1}{2^{\frac{k(n-s)}{2}}\pi^{\frac{k(k-1)}{4}}\Gamma\left(\frac{n-s}{2}\right)\Gamma\left(\frac{n-s-1}{2}\right)\cdots\Gamma\left(\frac{n-s-k+1}{2}\right)}\Lambda^{-\frac{n-s}{2}}Y^{\frac{n-s-k-1}{2}}$$

$$\exp\left(-\frac{1}{2}\Lambda^{-1}\cdot\cdot Y\right)\cdot dY, \tag{5.12}$$

§ 6. **Partial Correlation Coefficient.** We shall consider a k-dimensional non-singular normal distribution with means $m_1, \dots, m_k$ and variance-covariance matrix Λ, and denote the residuals of x_1 and x_2 when the other variates $x_3, \dots, x_k$ are fixed by $\eta_{1\cdot 34\cdots k}$ and $\eta_{2\cdot 34\cdots k}$ respectively, then

$$\left.\begin{aligned}\eta_{1\cdot 34\cdots k} &= x_1-\alpha_1-\beta_{13}x_3\cdots-\beta_{1k}x_k\,,\ \beta_{1i}=-\frac{\Lambda_{22\cdot 1i}}{\Lambda_{22\ 11}},\ \alpha_1=m_1+\sum_{i=3}^{k}\frac{\Lambda_{22\cdot 1i}}{\Lambda_{22\cdot 11}}m_i\,,\\ \eta_{2\cdot 34\cdots k} &= x_2-\alpha_2-\beta_{23}x_3\cdots-\beta_{2k}x_k\,,\ \beta_{2i}=-\frac{\Lambda_{11\cdot 2i}}{\Lambda_{11\cdot 22}},\ \alpha_2=m_2+\sum_{i=3}^{k}\frac{\Lambda_{11\cdot 2i}}{\Lambda_{11\cdot 22}}m_i\end{aligned}\right\} \tag{6.1}$$

The correlation coefficient $\rho_{12\cdot 34\cdots k}$ is the partial correlation coefficient of x_1 and x_2, and it is given as follows:

$$\rho_{12\cdot 34\cdots k}=\frac{E(\eta_{1\cdot 34\cdots k}\cdot\eta_{2\cdot 34\cdots k})}{\sqrt{E(\eta^2_{1\cdot 34\cdots k})E(\eta^2_{2\cdot 34\cdots k})}}=-\frac{\Lambda_{12}}{\sqrt{\Lambda_{11}\Lambda_{22}}}, \tag{6.2}$$

because

$$E(\eta^2_{1\cdot 34\cdots k})=\sigma^3_{1\cdot 34\cdots k}=\frac{\Lambda_{22}}{\Lambda_{22\cdot 11}},$$

$$E(\eta^2_{2\cdot 34\cdots k})=\sigma^2_{2\cdot 34\cdots k}=\frac{\Lambda_{11}}{\Lambda_{11\cdot 22}}.$$

The conditional distribution of $\eta_{1\cdot 34\cdots k}$ and $\eta_{2\cdot 34\cdots k}$, when $x_3, \dots, x_k$ are fixed, has the probability element

$$\frac{1}{2\pi\sigma_{1\cdot 34\cdots k}\sigma_{2\cdot 34\cdots k}\sqrt{1-\rho^2_{12\cdot 34\cdots k}}}\exp\left(-\frac{1}{2(1-\rho^2_{12\cdot 34\cdots k})}Q\right)\cdot d\eta_{1\cdot 34\cdots k}d\eta_{2\cdot 34\cdots k}, \tag{6.3}$$

where

$$Q=\frac{(x_1-\alpha_1-\beta_{13}x_3\cdots-\beta_{1k}x_k)^2}{\sigma^2_{1\cdot34\cdots k}}$$
$$-2\rho_{12\cdot34\cdots k}\frac{(x_1-\alpha_1-\beta_{13}x_3\cdots-\beta_{1k}x_k)(x_2-\alpha_2-\beta_{23}x_3\cdots-\beta_{1k}x_k)}{\sigma_{1\cdot34\cdots k}\sigma_{2\cdot34\cdots k}}$$
$$+\frac{(x_2-\alpha_2-\beta_{23}x_3\cdots-\beta_{2k}x_k)^2}{\sigma^2_{2\cdot34\cdots k}}. \tag{6.4}$$

If we denote the least-square estimates of

$$\alpha_p,\ \beta_{p3},\ \ldots,\ \beta_{pk},\quad p=1,2,$$

which minimizes

$$S_p=\sum_{\nu=1}^{n}(x_{p\nu}-\alpha_p-\beta_{p3}x_{3\nu}-\cdots-\beta_{pk}x_{k\nu})^2,\quad p=1,2$$

respectively, by

$$a_p,\ b_{p3},\ \ldots,\ b_{pk},\quad p=1,2,$$

then by the arguments given in § 5, the variates

$$y_{p\cdot34\cdots k,\nu}=x_{p\nu}-\bar{x}_p-b_{p3}(x_{3\nu}-\bar{x}_3)\cdots-b_{pk}(x_{k\nu}-\bar{x}_k),$$
$$p=1,2;\ \nu=1,2\ \ldots,\ n$$

are independent of $(b_{p3},\ \ldots,\ b_{pk})$, $p=1,2$, in the stochastic sense.

Let

$$\tilde{\boldsymbol{X}}_{12}=(\boldsymbol{x}_3-\bar{\boldsymbol{x}}_3,\ \boldsymbol{x}_4-\bar{\boldsymbol{x}}_4,\ \ldots,\ \boldsymbol{x}_k-\bar{\boldsymbol{x}}_k),$$

then it follows that

$$\boldsymbol{b}_p-\boldsymbol{\beta}_p=(\tilde{\boldsymbol{X}}'_{12}\tilde{\boldsymbol{X}}_{12})^{-1}\tilde{\boldsymbol{X}}'_{12}\cdot\boldsymbol{\eta}_{p\cdot34\cdots k},\quad p=1,2, \tag{6.5}$$

and

$$\boldsymbol{y}_{p\cdot34\cdots k}=(\boldsymbol{I}-\tilde{\boldsymbol{X}}_{12}(\tilde{\boldsymbol{X}}'_{12}\tilde{\boldsymbol{X}}_{12})^{-1}\tilde{\boldsymbol{X}}'_{12})\cdot\boldsymbol{\eta}_{p\cdot34\cdots k},\quad p=1,2. \tag{6.6}$$

If we choose $(n-k-2)\times n$ matrix $\boldsymbol{U}$ such that the square matrix

$$\boldsymbol{P}=\begin{pmatrix}(\tilde{\boldsymbol{X}}'_{12}\tilde{\boldsymbol{X}}_{12})^{-\frac{1}{2}}\tilde{\boldsymbol{X}}'_{12}\\ \boldsymbol{U}\end{pmatrix}$$

is an orthogonal matrix, and let it be

$$_1\boldsymbol{\zeta}_p=(\tilde{\boldsymbol{X}}'_{12}\tilde{\boldsymbol{X}}'_{12})^{-\frac{1}{2}}\tilde{\boldsymbol{X}}_{12}\cdot\boldsymbol{\eta}_{p\cdot34\cdots k},$$
$$_2\boldsymbol{\zeta}_p=\boldsymbol{U}\cdot\boldsymbol{\eta}_{p\cdot34\cdots k},\qquad p=1,2, \tag{6.7}$$

then the matrix

$$({}_2\boldsymbol{\zeta}_1,\ {}_2\boldsymbol{\zeta}_2)$$

is a random sample of size $(n-k-2)$ from a two-dimensional normal population (6.3), and it is clear that

$$y'_{1\cdot 34\cdots k}\cdot y_{1\cdot 34\cdots k} = {}_2\zeta_1'\cdot {}_2\zeta_1,$$
$$y'_{1\cdot 34\cdots k}\cdot y_{2\cdot 34\cdots k} = {}_2\zeta_1'\cdot {}_2\zeta_2,$$
$$y'_{2\cdot 34\cdots k}\cdot y_{2\cdot 34\cdots k} = {}_2\zeta_2'\cdot {}_2\zeta_2.$$

Now, by easy calculations, it is seen that

$$r_{12\cdot 34\cdots k} = -\frac{A_{12}}{\sqrt{A_{11}A_{22}}} = \frac{y'_{1\cdot 34\cdots k}\cdot y_{2\cdot 33\cdots k}}{\sqrt{(y'_{1\cdot 34\cdots k}y_{1\cdot 34\cdots k})(y'_{2\cdot 34\cdots k}y_{2\cdot 34\cdots k})}}$$
$$= \frac{{}_2\zeta_1'\cdot {}_2\zeta_2}{\sqrt{({}_2\zeta_1'\cdot {}_2\zeta_1)\cdot({}_2\zeta_2'\cdot {}_2\zeta_2)}}. \tag{6.8}$$

Whence we get the probability element of $r_{12\cdot 34\cdots k}$ as that of the inter-class correlation coefficient calculated from the sample of size $(n-k-2)$; i.e.,

$$f_n(r_{12\cdot 34\cdots k})dr_{12\cdot 34\cdots k}$$
$$= \frac{2^{n-k-1}}{\pi\cdot(n-k-1)!}(1-\rho^2_{12\cdot 34\cdots k})^{\frac{n-k+1}{2}}(1-r^2_{2\cdot 234\cdots k})^{\frac{n-k-2}{2}}\sum_{\nu=0}^{\infty}\frac{\Gamma^2\left(\frac{n-k+\nu+1}{2}\right)}{\nu!}$$
$$(2\rho_{12\cdot 34\cdots k}r_{12\cdot 34\cdots k})^\nu\cdot dr_{12\cdot 34\cdots k} \tag{6.9}$$

or

$$= \frac{n-k}{2\pi}(1-\rho^2_{12\cdot 34\cdots k})^{\frac{n-k+1}{2}}(1-r^2_{12\cdot 34\cdots k})^{\frac{n-k-2}{2}}\int_{-\infty}^{\infty}\frac{dz}{(\cosh z-\rho_{12\cdot 34\cdots k}r_{12\cdot 34\cdots k})^{n-k+1}}$$
$$\cdot dr_{12\cdot 34\cdots k}. \tag{6.10}$$

§ 7. **Bartlett's Decomposition Theorem.** In 1933, Prof. M. S. Bartlett[20] gave the so-called "Bartlett's Decomposition Theorem" of the Wishart distribution. In this section, we shall derive this theorem by means of the normal regression theory in the multivariate case.

We shall first explain the theorem in spacial cases.

7.1. (i) *Two variates case.* In this case, the joint distribution of m_{11}, m_{12}, m_{22}, are given by

$$W_2\left(m_{11}m_{12}m_{22};\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix};\ n-1\right)$$

$$\frac{(m_{11}m_{22}-m_{12}^2)^{\frac{n-4}{2}}e^{-\frac{1}{2(1-\rho^2)}\left(\frac{m_{11}}{\sigma_1^2}-2\rho\frac{m_{12}}{\sigma_1\sigma_1}+\frac{m_{22}}{\sigma_2^2}\right)}dm_{11}dm_{12}dm_{22}}{2^{n-1}\pi^{\frac{1}{2}}(\sigma_1\sigma_2\sqrt{1-\rho^2})^{n-1}\Gamma\left(\frac{n-1}{2}\right)\cdot\Gamma\left(\frac{n-2}{2}\right)}. \tag{7.1}$$

If we put

$$\left.\begin{aligned} b_{12} &= \frac{m_{12}}{m_{22}} \\ m_{11\cdot 2} &= \sum_{\alpha=1}^{n}[(x_{1\alpha}-\bar{x}_1)-b_{12}(x_{2\alpha}-\bar{x}_2)]^2 = m_{11}-2b_{12}m_{12}+b_{12}^2\cdot m_{22} \\ &= m_{11}-\frac{m_{12}^2}{m_{22}} = \frac{m_{11}m_{22}-m_{12}^2}{m_{22}} = \frac{M}{m_{22}}, \\ u_{12} &= \sqrt{m_{22}}\cdot(b_{12}-\beta_{12}), \qquad \beta_{12} = \rho\frac{\sigma_1}{\sigma_2} \end{aligned}\right\} \tag{7.2}$$

then we get by easy calculations

$$\frac{\partial(m_{11\cdot 2},\, u_{1\cdot 2},\, m_{22})}{\partial(m_{11},\, m_{12},\, m_{22})} = \frac{1}{\sqrt{m_{22}}}. \tag{7.3}$$

Since

$$\begin{aligned} \frac{1}{1-\rho^2}\left(\frac{m_{11}}{\sigma_1^2}-2\rho\frac{m_{12}}{\sigma_1\sigma_2}+\frac{m_{22}}{\sigma_2^2}\right) &= \frac{1}{1-\rho^2}\left(\frac{m_{11}}{\sigma_1^2}-2\rho\frac{m_{12}}{\sigma_1\sigma_2}+\rho^2\frac{m_{22}}{\sigma_2^2}\right)+\frac{m_{22}}{\sigma_2^2} \\ &= \frac{1}{\sigma_1^2(1-\rho^2)}\sum_{\alpha=1}^{n}[(x_{1\alpha}-\bar{x}_1)-\beta_{12}(x_{2\alpha}-\bar{x}_2)]^2+\frac{m_{22}}{\sigma_2^2} \\ &= \frac{1}{\sigma_{1\cdot 2}^2}\sum_{\alpha=1}^{n}[(x_{1\alpha}-\bar{x}_1)-b_{12}(x_{2\alpha}-\bar{x}_2)+(b_{12}-\beta_{12})(x_{2\alpha}-\bar{x}_2)]^2+\frac{m_{22}}{\sigma_2^2} \\ &= \frac{u_{1\cdot 2}^2}{\sigma_{1\cdot 2}^2}+\frac{m_{11\cdot 2}}{\sigma_{1\cdot 2}^2}+\frac{m_{22}}{\sigma_2^2}, \end{aligned}$$

the joint distribution of $u_{1\cdot 2}$, $m_{11\cdot 2}$, m_{22} has the probability element

$$\begin{aligned} &\frac{e^{-\frac{u_{1\cdot 2}^2}{2\sigma_{1\cdot 2}^2}}du_{1\cdot 2}}{\sqrt{2\pi\sigma_{1\cdot 2}^2}}\cdot\frac{m_{11\cdot 2}^{\frac{1}{2}(n-4)}e^{-\frac{m_{11\cdot 2}}{2\sigma_{1\cdot 2}^2}}dm_{11\cdot 2}}{(2\sigma_{1\cdot 2}^2)^{\frac{n-2}{2}}\Gamma\left(\frac{n-2}{2}\right)}\cdot\frac{m_{22}^{\frac{1}{2}(n-3)}e^{-\frac{m_{22}}{2\sigma_2^2}}dm_{22}}{(2\sigma_2^2)^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} \\ &= N_1(u_{1\cdot 2};\,\sigma_{1\cdot 2}^2)\cdot W_1(m_{11\cdot 2};\,\sigma_{1\cdot 2}^2;\,n-2)\cdot W_1(m_{22};\,\sigma_2^2;\,n-1). \end{aligned} \tag{7.4}$$

That is, that $u_{1\cdot 2}$, $m_{11\cdot 2}$, m_{22} are mutually independent in the stochastic sense, and $u_{1\cdot 2}$ is distributed according to $N(0,\sigma_{1\cdot 2}^2)$, $m_{11\cdot 2}/\sigma_{1\cdot 2}^2$ is distributed rccording to the chi-square distribution of degrees of freedom of $(n-2)$, and m_{22}/σ_2^2 is distributed according to the chi-square distribution of degrees of freedom of $(n-1)$. (7.4) is the Bartlett's decomposition of the 2-dimensional Wishart distribution.

(ii) *Three variates case.* In this case the joint distribution of m_{11}, m_{12}, m_{13}, m_{22}, m_{23}, m_{33} has the probability element

$$\begin{aligned} &W_3(m_{11},\, m_{12},\, m_{13},\, m_{22},\, m_{23},\, m_{33};\,\Lambda;\,n-1) \\ &= \frac{M^{\frac{n-5}{2}}\exp\left(-\frac{1}{2\Lambda}\sum_{\mu\nu=1}^{3}\Lambda_{\mu\nu}m_{\mu\nu}\right)}{2^{\frac{3(n-1)}{2}}\pi^{\frac{3}{2}}\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n-3}{2}\right)}dm_{11}dm_{12}dm_{13}dm_{22}dm_{23}dm_{33} \end{aligned} \tag{7.5}$$

If we put

$$b_{13}=\frac{m_{13}}{m_{33}}, \qquad b_{23}=\frac{m_{23}}{m_{33}},$$

$$\begin{aligned} x_{1\cdot 3\alpha} &= x_{1\alpha}-\bar{x}_1-b_{13}(x_{3\alpha}-\bar{x}_3) \\ x_{2\cdot 3\alpha} &= x_{2\alpha}-x_2-b_{23}(\bar{x}_{3\alpha}-\bar{x}_3) \end{aligned} \qquad \alpha=1,2,\cdots,n$$

and

$$m_{11\cdot 3}=\sum_{\alpha=1}^{n} x_{1\cdot 3\alpha}^2=m_{11}-2b_{13}m_{13}+b_{13}^2\cdot m_{33}=\frac{m_{11}m_{33}-m_{13}^2}{m_{33}}=\frac{M_{22}}{m_{33}},$$

$$m_{12\cdot 3}=\sum_{\alpha=1}^{n} x_{1\cdot 3\alpha}x_{2\cdot 3\alpha}=m_{12}-b_{13}\cdot m_{23}-b_{23}\cdot m_{13}+b_{12}b_{13}\cdot m_{33}$$

$$=\frac{m_{12}m_{33}-m_{13}m_{23}}{m_{33}}=-\frac{M_{12}}{m_{33}},$$

$$m_{22\cdot 3}=\sum_{\alpha=1}^{n} x_{2\cdot 3\alpha}^2=m_{22}-2b_{23}m_{23}+b_{23}^2m_{33}=\frac{M_{33}}{m_{33}},$$

$$u_{1\cdot 3}=\sqrt{m_{33}}\,(b_{13}-\beta_{13}), \qquad u_{2\cdot 3}=\sqrt{m_{33}}\,(b_{23}-\beta_{23}),$$

then it follows that

$$dm_{11}dm_{12}dm_{13}dm_{22}dm_{23}dm_{33}=m_{33}du_{1\cdot 3}du_{2\cdot 3}dm_{11\cdot 3}dm_{12\cdot 3}dm_{22\cdot 3}dm_{33}\,. \tag{7.6}$$

Since

$$\lambda_{13}\Lambda_{11}+\lambda_{23}\Lambda_{21}+\lambda_{33}\Lambda_{31}=0,$$
$$\lambda_{13}\Lambda_{12}+\lambda_{23}\Lambda_{22}+\lambda_{33}\Lambda_{32}=0,$$

multiply the first by λ_{13}, the second by λ_{23}, and add up, then we get

$$\lambda_{13}^2\Lambda_{11}+2\lambda_{13}\lambda_{23}\Lambda_{12}+\lambda_{23}^2\Lambda_{22}+\lambda_{33}(\lambda_{13}\Lambda_{13}+\lambda_{23}\Lambda_{23})=0,$$

or

$$\lambda_{13}^2\Lambda_{11}+2\lambda_{13}\lambda_{23}\Lambda_{12}+\lambda_{22}^2\Lambda_{22}+\lambda_{33}(\Lambda-\lambda_{33}\Lambda_{33})=0,$$

because $\lambda_{13}\Lambda_{13}+\lambda_{23}\Lambda_{23}+\lambda_{33}\Lambda_{33}=0$. Therefore, we have

$$\frac{\Lambda_{33}}{\Lambda}=\left(\frac{\lambda_{13}}{\lambda_{33}}\right)^2\frac{\Lambda_{11}}{\Lambda}+2\frac{\lambda_{13}}{\lambda_{33}}\frac{\lambda_{23}}{\lambda_{33}}\frac{\Lambda_{12}}{\Lambda}+\left(\frac{\lambda_{23}}{\lambda_{33}}\right)^2\frac{\Lambda_{22}}{\Lambda}+\frac{1}{\lambda_{33}}$$

$$=\beta_{13}^2\frac{\Lambda_{11}}{\Lambda}+2\beta_{13}\beta_{23}\frac{\Lambda_{12}}{\Lambda}+\beta_{23}^2\frac{\Lambda_{22}}{\Lambda}+\frac{1}{\lambda_{33}},$$

where it follows that

$$Q=\sum_{\mu\nu=1}^{3}\frac{\Lambda_{\mu\nu}}{\Lambda}m_{\mu\nu}=\frac{\Lambda_{11}}{\Lambda}(m_{11}+\beta_{13}^2m_{33})+2\frac{\Lambda_{12}}{\Lambda}(m_{12}+\beta_{13}\beta_{23}m_{33})$$

$$+\frac{\Lambda_{22}}{\Lambda}(m_{22}+\beta_{23}^2m_3)+2\frac{\Lambda_{13}}{\Lambda}m_{13}+2\frac{\Lambda_{23}}{\Lambda}m_{23}+\frac{m_{33}}{\lambda_{33}}$$

$$=\frac{\Lambda_{11}}{\Lambda}(m_{11}-2\beta_{13}m_{14}+\beta_{13}^2m_{33})+2\frac{\Lambda_{12}}{\Lambda}(m_{12}-\beta_{13}m_{23}-\beta_{23}m_{13}+\beta_{13}\beta_{23}m_{23})$$

$$+\frac{\Lambda_{22}}{\Lambda}(m_{22}-2\beta_{23}m_{23}+\beta_{23}^2m_{33})+2m_{13}\left(\beta_{13}\frac{\Lambda_{11}}{\Lambda}+\beta_{23}\frac{\Lambda_{12}}{\Lambda}+\frac{\Lambda_{13}}{\Lambda}\right)$$

$$+2m_{23}\left(\beta_{13}\frac{\Lambda_{21}}{\Lambda}+\beta_{23}\frac{\Lambda_{22}}{\Lambda}+\frac{\Lambda_{23}}{\Lambda}\right)+\frac{m_{33}}{\lambda_{33}}$$

$$=\frac{\Lambda_{11}}{\Lambda}m_{11\cdot3}+2\frac{\Lambda_{12}}{\Lambda}m_{12\cdot3}+\frac{\Lambda_{22}}{\Lambda}m_{22\cdot2}+\frac{\Lambda_{11}}{\Lambda}u_{1\cdot3}^2+2\frac{\Lambda_{12}}{\Lambda}u_{1\cdot3}\cdot u_{2\cdot3}$$

$$+\frac{\Lambda_{23}}{\Lambda}u_{2\cdot3}^2+\frac{m_{33}}{\lambda_{33}} \tag{7.7}$$

On the other hand, we have

$$M=\begin{vmatrix} m_{11} & m_{12} & m_{13} \\ m_{12} & m_{22} & m_{23} \\ m_{13} & m_{23} & m_{33} \end{vmatrix}=\begin{vmatrix} m_{11}-\frac{m_{13}m_{13}}{m_{33}} & m_{12}-\frac{m_{13}m_{23}}{m_{33}} \\ m_{12}-\frac{m_{13}m_{23}}{m_{33}} & m_{22}-\frac{m_{23}m_{23}}{m_{33}} \end{vmatrix}\cdot m_{33}$$

$$=\begin{vmatrix} m_{11\cdot3} & m_{12\cdot3} \\ m_{12\cdot3} & m_{22\cdot3} \end{vmatrix} m_{33}$$

Consequently, we have the decomposition

$$W_3(\boldsymbol{M};\Lambda;n-1)=\frac{e^{-\frac{1}{2\Lambda}(\Lambda_{11}u_{1\cdot3}^2+2\Lambda_{22}u_{1\cdot3}u_{2\cdot3}+\Lambda_{22}u_{2\cdot3}^2)}du_{1\cdot3}du_{2\cdot3}}{2\pi\cdot\begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{vmatrix}^{\frac{1}{2}}}\times$$

$$\frac{\begin{vmatrix} m_{11\cdot3} & m_{12\cdot3} \\ m_{12\cdot3} & m_{22\cdot3} \end{vmatrix}^{\frac{n-5}{2}} e^{-\frac{1}{2\Lambda}(\Lambda_{11}m_{11\cdot3}+2\Lambda_{12}m_{12\cdot3}+\Lambda_{22}m_{22\cdot3})}dm_{11\cdot3}dm_{12\cdot3}dm_{22\cdot3}}{2^{\frac{2(n-2)}{2}}\pi^{\frac{1}{2}}\begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{vmatrix}^{\frac{n-2}{2}}\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n-3}{2}\right)}\times$$

$$\frac{m_{33}^{\frac{n-3}{2}}e^{-\frac{m_{33}}{2\lambda_{33}}}dm_{33}}{(2\lambda_{33})^{\frac{n-1}{2}}\cdot\Gamma\left(\frac{n-1}{2}\right)} \tag{7.8}$$

where, of course,

$$\begin{pmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{pmatrix}=\begin{pmatrix} \lambda_{11}-\frac{\lambda_{13}\lambda_{13}}{\lambda_{33}} & \lambda_{12}-\frac{\lambda_{13}\lambda_{23}}{\lambda_{33}} \\ \lambda_{12}-\frac{\lambda_{13}\lambda_{23}}{\lambda_{33}} & \lambda_{12}-\frac{\lambda_{13}\lambda_{23}}{\lambda_{33}} \end{pmatrix}$$

The formula (7.8) may be written in the form

$$W_3(M; \Lambda; n-1) = N_2(u_{1\cdot3}, u_{2\cdot3}; \begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{vmatrix})\times$$

$$W_2(m_{11\cdot3}, m_{12\cdot3}, m_{22\cdot3}; \begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{vmatrix}; n-2)\cdot W_1(m_{33}; \lambda_{33}; n-1), \quad (7.9)$$

symbolically.

Next, if we put

$$\begin{aligned} m_{11\cdot23} &= \sum_{\alpha=1}^{n} (x_{1\cdot3\alpha} - b_{12\cdot3}x_{2\cdot3\alpha})^2 \\ &= \sum_{\alpha=1}^{n} \left\{x_{1\alpha} - \bar{x}_1 - b_{12\cdot3}(x_{2\alpha} - \bar{x}_2) - (b_{13} - b_{12\cdot3}b_{23})(x_{3\alpha} - \bar{x}_3)\right\}^2 \\ &= \sum_{\alpha=1}^{n} \left\{x_{1\alpha} - \bar{x}_1 - b_{12\cdot3}(x_{3\alpha} - \bar{x}_2) - b_{13\cdot2}(x_{3\alpha} - \bar{x}_3)\right\}^2 = \frac{M}{M_{33}}, \end{aligned}$$

because

$$b_{13} - b_{12\cdot3}b_{23} = b_{13} - \frac{b_{12} - b_{13}b_{32}}{1 - b_{23}b_{32}} b_{23} = \frac{b_{13} - b_{12}b_{23}}{1 - b_{23}b_{32}} = b_{13\cdot2},$$

and

$$u_{1\cdot23} = \sqrt{m_{22\cdot3}}\,(b_{12\cdot3} - \beta_{12\cdot3}),$$

then the probability element

$$W_2(m_{11\cdot3}, m_{12\cdot3}, m_{22\cdot3}; \begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{vmatrix}; n-2)$$

can be decomposed into the product

$$N_1(u_{1\cdot23}; \sigma_{1\cdot23}^2)\cdot W_1(m_{11\cdot23}; \sigma_{1\cdot23}^2; n-2)\cdot W_1(m_{22\cdot3}; \sigma_{2\cdot3}^2; n-2), \quad (7.10)$$

where

$$\sigma_{1\cdot23}^2 = \lambda_{11\cdot23} = \lambda_{11\cdot3} - \frac{\lambda_{12\cdot3}^2}{\lambda_{22\cdot3}},$$

$$\sigma_{2\cdot3}^2 = \lambda_{22\cdot3}$$

Finally we have the decomposition

$$W_3(M; \Lambda; n-1) = N_2(u_{1\cdot3}, u_{2\cdot3}; \begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{12\cdot3} \end{vmatrix})\cdot N_1(u_{1\cdot23}; \sigma_{1\cdot23}^2)\times$$

$$W_1(m_{11\cdot33}, \sigma_{1\cdot23}^2; n-2)W_1(m_{22\cdot3}; \sigma_{2\cdot3}^2; n-2)\cdot W_2(m_{11\cdot3}, m_{12\cdot3}, m_{22\cdot3}; \begin{vmatrix} \lambda_{11\cdot3} & \lambda_{12\cdot3} \\ \lambda_{12\cdot3} & \lambda_{22\cdot3} \end{vmatrix};$$

$$n-2)\cdot W_1(m_{33}; \lambda_{33}; n-1), \quad (7\,11)$$

which is the Bartlett's decomposition of the Wishart distribution in three-varites case.

7.2. *Bartlett's decomposition theorem in the general case.* We shall derive the Bartlett's decomposition theorem for k-dimensional Wishart distribution by means of the normal regression theory in multivariate case described in section 6.

The population considered here has the probability element

$$p(\mathfrak{x})d\mathfrak{x} = (2\pi)^{-\frac{k}{2}}\Lambda^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\mathfrak{x}-\mathfrak{m})\Lambda^{-1}(\mathfrak{x}-\mathfrak{m})'\right]d\mathfrak{x},$$

and the conditional probability distribution of $(x_1, \cdots, x_{k-1})$ under the condition that x_k is fixed is the $(k-1)$-dimensional normal distribution whose probability element is

$$\begin{aligned} &p(x_1, \cdots, x_{k-1} | x_k)dx_1 \cdots dx_{k-1} \\ &= (2\pi)^{-\frac{k-1}{2}}\Lambda^{(1)^{-\frac{1}{2}}}\exp\left[-\frac{1}{2\Lambda^{(1)}}\sum_{i,j=1}^{k-1}\Lambda_{ij}^{(1)}\eta_i^{(1)}\eta_j^{(1)}\right]dx_1 \cdots dx_{k-1}, \end{aligned} \tag{7.12}$$

where

$$\begin{aligned} &\alpha_i^{(1)} = m_i - \beta_{ik}\cdot m_k, \quad \beta_{ik} = \frac{\lambda_{ik}}{\lambda_{kk}}, \\ &\eta_i^{(1)} = x_i - \alpha_i^{(1)} - \beta_{ik}x_k \\ &\qquad i = 1, 2, \cdots, k-1 \end{aligned}$$

The least-square estimates $a_i^{(1)}, b_{ik}, i = 1, 2, \cdots, k-1$ of $\alpha_i^{(1)}, \beta_{ik}, i = 1, 2, \cdots, k-1$ and the residual variates

$$y_{i\cdot k} = \begin{pmatrix} x_{i1} - \bar{x}_i - b_{ik}(x_{k1} - \bar{x}_k) \\ x_{i2} - \bar{x}_i - b_{ik}(x_{k2} - \bar{x}_k) \\ \vdots \\ x_{in} - \bar{x}_i - b_{ik}(x_{kn} - \bar{x}_k) \end{pmatrix}, \quad i = 1, 2, \cdots, k-1$$

are independent in the stochastic sense.

Since

$$b_{ik} - \beta_{ik} = \frac{1}{\sum_{\alpha=1}^{n}(x_{k\alpha} - \bar{x}_k)^2}\sum_{\alpha=1}^{n}(x_{k\alpha} - \bar{x}_k)\cdot\eta_{i\alpha}^{(1)}, \quad i = 1, 2, \cdots, k-1$$

the conditional joint distribution of the variates

$$\begin{aligned} u_{i\cdot k} = \sqrt{m_{kk}}\,(b_{ik} - \beta_{ik}) = \frac{1}{\sqrt{\sum_{\alpha=1}^{n}(x_{k\alpha} - \bar{x}_k)^2}}\sum_{\alpha=1}^{n}(x_{k\alpha} - \bar{x}_k)\eta_{i\alpha}^{(1)} \\ i = 1, 2, \cdots, k-1 \end{aligned} \tag{7.13}$$

is the $(k-1)$-dimensional normal distribution of zero means and variance-covariance matrix

$$\Lambda^{(1)} = \left(\lambda_{ij} - \frac{\lambda_{ik}\lambda_{jk}}{\lambda_{kk}}\right), \quad i, j = 1, 0, \cdots, k-1,$$

The joint distribution of the elements of the matrix

$$\boldsymbol{Y}_k = (\boldsymbol{y}'_{i\cdot k}\boldsymbol{y}_{j\cdot k}), \quad i, j = 1, 2, \ldots, k-1$$

is the Wishart distribution of $(k-1)$-dimensions

$$W_{k-1}(\boldsymbol{Y}_k; \Lambda^{(1)}; n-2) \tag{7.14}$$

All the distributions of the variates above considered are independent of the conditioning variate x_k, so we have the following decomposition; i.e.,

$$W_k(\boldsymbol{M}; \Lambda; n-1) = N_{k-1}(u_{1\cdot k}, \cdots, u_{k-1\cdot k}; \Lambda^{(1)}) \cdot W_{k-1}(\boldsymbol{Y}_k; \Lambda^{(1)}; n-1) \cdot W_1(m_{kk}; \lambda_{kk}; n-1) \tag{7.15}$$

The variates can be written in the form

$$\boldsymbol{y}_{i\cdot k} = \boldsymbol{P}_k \cdot \begin{pmatrix} \xi_{i1} \\ \xi_{i2} \\ \vdots \\ \xi_{i,n-2} \end{pmatrix}, \quad i = 1, 2, \cdots, k-1,$$

where the system of vectors

$$\boldsymbol{\xi}_\alpha = (\xi_{1\alpha}, \xi_{2\alpha}, \cdots, \xi_{k-1,\alpha}), \quad \alpha = 1, 2, \cdots, n-2$$

are the random sample of size $(n-2)$ from the $(k-1)$-dimensional normal distribution whose probability element was given by (7.10), and $\boldsymbol{P}_k$ is a $n \times (n-2)$-matrix depending only on the k-th variate x_k and of which the rows satisfy the orthogonality conditions. Hence it follows that, since

$$\boldsymbol{y}'_{i\cdot k}\boldsymbol{y}_{j\cdot k} = \sum_{\alpha=1}^{n-2} \xi_{i\alpha}\xi_{j\alpha}, \quad i, j = 1, 2, \cdots, k-1,$$

the joint distribution of $\boldsymbol{Y}_k$ is determined completely by the joint distribution of $\xi_1, \xi_2, \cdots, \xi_{n-2}$.

If we apply the above arguments to the joint distribution of $\xi_1, \xi_2, \cdots, \xi_{n-2}$, then the following decomposition can be obtained; i.e.,

$$\begin{aligned} &W_{k-1}(\boldsymbol{Y}_k; \Lambda^{(1)}; n-2) \\ &= N_{k-2}(u_{1\cdot k-1, k}, \cdots, u_{k-2\cdot k-1, k}; \Lambda^{(2)}) \cdot W_{k-2}(\boldsymbol{Y}_{k-1}; \Lambda^{(2)}; n-3) \\ &\quad \cdot W_1(m_{k-1, k-1\cdot k}; \lambda_{k-1, k-1\cdot k}; n-2), \end{aligned} \tag{7.16}$$

where

$$u_{i\cdot k-1,k}=\sqrt{m_{k-1,k-1\cdot k}}\cdot(b_{i\cdot k-1,k}-\beta_{i\cdot k-1,k})$$
$$=\frac{1}{\sqrt{\sum_{\alpha=1}^{n-2}\xi_{k-1\cdot\alpha}^2}}\sum_{\alpha=1}^{n-2}\xi_{k-1,\alpha}\eta_{i\alpha}^{(2)},\quad i=1,2,\cdots,k-2,$$
$$\eta_{i\alpha}^{(2)}=\xi_{i\alpha}-\alpha_i^{(2)}-\beta_{i\cdot k-1,k}\cdot\xi_{k-1,\alpha},\quad i=1,2,\cdots,k-2,$$

and

$$\Lambda^{(1)}=\left(\lambda_{ij\cdot k}-\frac{\lambda_{i,k-1\cdot k}\lambda_{j,k-1\cdot k}}{\lambda_{k-1,k-1\cdot k}}\right),\quad i,j=1,2,\cdots,k-2,$$

Thus we can obtain successively the decomposition of the Wishart distribution into the product of the normal variates, chi-sqnare variate, and Wishart distribution of reduced dimensions by one, and all of them are mutually independent in the stochastic sense absolutely. The Bartlett's decomposition of the Wishart distribution can be easily obtained in this manner.

(Received February 26, 1953)

Notes and References.

(1) R. A. Fisher: Frequency distribution of the value of the correlation coefficient in samples from an infintely large population. Biom. Vol. 10, (1915). pp. 507–521.

(2) J. Wishart: The generalized product moment distribution in sample from a normal multivariate population. Biom. Vol. 20A (1928). pp. 32–52.

J. Wishart and M. S. Bartlett: The generalized product moment distribution in a normal distribution. Proc. Cambridge Phil. Soc., Vol. 29, (1933). pp. 260–270.

(3) H. Cramér: *Mathematical Methods of Statistics,* Princeton Univ. Press, 1946, 29.9, pp. 404–405.

G. Rash: A functional equation for Wishart distribution. The Ann. of Math. Stat. Vol. 19, (1948). pp. 262–266.

David Fog: The geometrical method in the theory of sampling. Biom. Vol. 35, (1948). pp. 46–54.

J. Wishart: Note on "Proofs of the distribution law of the second order moment statistics". Biom. Vol. 35, (1949). p. 422.

A. C. Aitken: On the Wishart distribution in statistics. Biom. Vol. 36, (1949). pp. 59–62.

(4) M. S. Bartlett: On the theory of statistical regression. Proc. Roy. Soc. Edinburgh. Vol. 103, (1932–1933).

(5) G. Elfving: A simple method of deducing certain distributions connected with multivariate sampling. Skandinavisk Aktuarietidskrift, Vol. (1947)

(6) D. Fog: loc. cit.

(7) See, for instance, S. S. Wilks: *Mathematical Statistics*, Princeton, (1943), Chapter VIII Normal Regression Theory, pp. 157–175.

(8) s×s square matrix $\boldsymbol{Z'Z} \equiv (b_{\alpha\beta})$, say, is positive-definite, because, for any set of real numbers $t_1, t_2, \cdots, t_s$, we have

$$\sum_{\alpha\cdot\beta=1}^{s} b_{\alpha\beta} t_\alpha t_\beta = \sum_{\alpha\cdot\beta=1}^{s} \sum_{\mu=1}^{n} z_{\alpha\mu} z_{\beta\mu} t_\alpha t_\beta = \sum_{\mu=1}^{n} (\sum_{\alpha=1}^{s} t_\alpha z_{\alpha\mu})^2 \geqq 0$$

the equality holds when and only when

$$\sum_{\alpha=1}^{t} t_\alpha z_{\alpha\mu} = 0, \qquad \mu = 1, 2, \cdots, n,$$

or

$$\boldsymbol{z}_1 t_1 + \boldsymbol{z}_2 t_2 + \cdots + \boldsymbol{z}_s t_s = 0,$$

where

$$\boldsymbol{z}_\alpha = \begin{pmatrix} z_{\alpha 1} \\ z_{\alpha 2} \\ \vdots \\ z_{\alpha n} \end{pmatrix} \quad \alpha = 1, 2, \cdots, s,$$

whence it follows that $t_1 = t_2 = \cdots = t_s = 0$.

If we take a suitable orthogonal matrix $\boldsymbol{P}$, then

$$\boldsymbol{P'Z'ZP} = \begin{pmatrix} d_1 & 0 \cdots\cdots 0 \\ 0 & d_2 \cdots\cdots 0 \\ \cdots\cdots & \cdots\cdots \\ 0 & 0 \cdots\cdots d_s \end{pmatrix}, \quad d_i > 0, \quad i = 1, 2, \cdots, s,$$

Let

$$\boldsymbol{D} = \begin{pmatrix} \sqrt{d_1} & 0 & \cdots\cdots & 0 \\ 0 & \sqrt{d_2} & \cdots\cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots\cdots & \sqrt{d_s} \end{pmatrix}$$

then it follows that

$$\boldsymbol{Z'Z} = \boldsymbol{P'D^2P} = (\boldsymbol{P'DP})^2$$

Hence we may put

$$(\boldsymbol{Z'Z})^{\frac{1}{2}} = \boldsymbol{P'DP}$$

and it is symmetric and positive-definite.

(9) See, for instance, J. Ogawa: On the independence of bilinear and quadratic forms of a random sample from a normal population. The Ann. Inst. Stat. Math. Vol. 1, No. 1, (1949). pp. 83–108.

10) The statistic G given in (2.17) satisfies the relation

$$\frac{n-s}{s-r} G = F'^{s-r}_{n-s}$$

where F'^{s-r}_{n-s} is the variate which follows the non-central F-distribution of degrees of freedom of $(s-r, n-s)$. For the sake of the brevity of descriptions, we call here the distribution of the statistic G, the "G-distribution". See, for example, H. B. Mann: *Analysis And Designs Of Experiments*, of The Dover Series in Math. and Phys., (1949), p. 69.

(11) This fact is essentially due to the so-called "Schmidt's method of orthogonalization".

(12) See, for example, H. Cramér: *Math. Meth. of Stat.*, 29.13.7, p. 414.

(13) R. A. Fisher: The general sampling distribution of the multiple correlation coefficient. Proc. Roy. Soc. London. Vol. 121, (1928). pp. 654–673.

(14) S. S. Wilks: On the sampling distribution of the multiple correlation coefficient. The Ann. of Math. Stat., Vol. 3, (1932). pp. 196–203.

(15) S. Nabeya: On the sampling distribution of the multiple correlation coefficient. Kohkyuroku of The Inst. of Stat. Math., Vol. 4, No. 9, (1948). pp. 381–385. (Written in Japanese).

(16) P. A. P. Moran: The distribution of the multiple correlation coefficient. Proc. Cambridge Phil. Soc., Vol. 46, (1950). pp. 521–522.

(17) To prove this lemma, calculate the characteristic function of λ^*, and invert it by means of the famous P. Lévy's inversion formula.

(18) Harold Hotelling: The generalizations of Student's ratio. The Ann. of Math. Stat., Vol. 2, (1931). p. 360.

For the sampling distribution of T^2 when the null-hypothesis is true, see, for instance, H. Cramér: *Math. Meth. of Stat.*, p. 409, (29.11.4).

(19) G. Elfving: loc. cit., p. 66, (35).

(20) M. S. Bartlett: loc. cit.

(January 27, 1953).