CONTRIBUTIONS TO MULTIVARIATE ANALYSIS INCLUDING UNIVARIATE
AND MULTIVARIATE VARIANCE COMPONENTS ANALYSIS
AND FACTOR ANALYSIS

by

Ramanathan Gnanadesikan
University of North Carolina

This research was supported by the United States Air Force through the Air Force
Office of Scientific Research of the Air Research and Development Command under
contract No. AF 18(600) - 83.

Institute of Statistics
Mimeograph Series No. 158
June, 1956
ACKNOWLEDGMENT

I wish to express my deep gratitude and heartfelt thanks to Professor S. N. Roy for his constant and inspiring guidance without which this work would not have been completed. I am thankful to Professors R. C. Bose and G. E. Nicholson for going through sections of the manuscript and for their criticisms and suggestions.

My thanks are due to the Institute of Statistics for the financial and other facilities which were provided me during the progress of this work.

Finally, I wish to thank Mrs. H. Kattsoff and Mrs. F. Spencer for their careful typing of the manuscript, and Miss M. Jordan for help received from her at various times.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>ii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>vii</td>
</tr>
<tr>
<td>NOTATION AND PRELIMINARIES</td>
<td>xiii</td>
</tr>
<tr>
<td>I. FURTHER IMPLICATIONS OF CERTAIN PREVIOUS RESULTS ON</td>
<td></td>
</tr>
<tr>
<td>MULTIVARIATE CONFIDENCE BOUNDS</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Certain lemmas</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Confidence bounds on roots connected with $\xi$ of $N(\xi, \Sigma)$</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Confidence bounds on roots connected with $\xi_1$ and $\xi_2$ of $N(\xi_1, \Sigma_1)$ and $N(\xi_2, \Sigma_2)$</td>
<td>8</td>
</tr>
<tr>
<td>1.5 Confidence bounds connected with the regression matrix $\beta(p \times q)$ of a $p$-set on a $q$-set in a $(p + q)$-variate normal</td>
<td>14</td>
</tr>
<tr>
<td>1.6 Confidence bounds on roots connected with multivariate linear hypotheses on means</td>
<td>17</td>
</tr>
<tr>
<td>II. ON EQUALITY OF VARIANCES AND OF DISPERSION MATRICES</td>
<td>27</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>27</td>
</tr>
<tr>
<td>2.2 The simultaneous variance ratios test</td>
<td>27</td>
</tr>
<tr>
<td>2.3 Choice of $F_{11}, F_{12}$ for $i = 1, 2, \ldots, k$</td>
<td>29</td>
</tr>
<tr>
<td>2.4 Evaluation of a probability integral in connection with the test of section 2.2</td>
<td>29</td>
</tr>
<tr>
<td>2.5 Properties of the power of the test proposed in 2.2</td>
<td>34</td>
</tr>
<tr>
<td>2.6 The associated simultaneous confidence bounds on $\sigma_1^2/\sigma^2, \ldots, \sigma_k^2/\sigma^2$</td>
<td>39</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>PAGE</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>IV. MULTIVARIATE VARIANCE COMPONENTS</td>
<td>92</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>92</td>
</tr>
<tr>
<td>4.2 The Model</td>
<td>92</td>
</tr>
<tr>
<td>4.3 Certain observations on, and deductions from, the model</td>
<td>93</td>
</tr>
<tr>
<td>4.4 Certain results on linear estimability</td>
<td>96</td>
</tr>
<tr>
<td>4.5 Multivariate variance components</td>
<td>97</td>
</tr>
<tr>
<td>4.6 Further multivariate analogues of results in section 3.5</td>
<td>110</td>
</tr>
<tr>
<td>4.7 Tie-up between the analyses of the multivariate models, I and II, for a k-way classification</td>
<td>120</td>
</tr>
<tr>
<td>4.8 Tests of hypotheses connected with the multivariate variance components</td>
<td>124</td>
</tr>
<tr>
<td>4.9 Simultaneous confidence bounds on ( \sigma_1^2, \ldots, \sigma_k^2 ) and ( \sigma(\Sigma) )</td>
<td>126</td>
</tr>
<tr>
<td>4.10 Alternative set of confidence bounds for the individual ( \sigma_1^2, \ldots, \sigma_k^2 )</td>
<td>130</td>
</tr>
<tr>
<td>4.11 Concluding remarks</td>
<td>135</td>
</tr>
</tbody>
</table>

<p>| V. FACTOR ANALYSIS | 137 |
| 5.1 Introduction | 137 |
| 5.2 The Factor Analysis &quot;Model&quot; | 138 |
| 5.3 Clarification of the relationship between p and q | 141 |
| 5.4 The problems in factor analysis | 144 |
| 5.5 Confidence bounds on the parameters in the factor analysis &quot;model&quot; | 145 |
| 5.6 Distribution problem involved in sections 1.2, 2.9, 4.9, and 5.5 | 153 |</p>
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7 Remarks on the results of section 5.6</td>
<td>170</td>
</tr>
<tr>
<td>5.8 Concluding remarks</td>
<td>171</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>173</td>
</tr>
</tbody>
</table>
INTRODUCTION

The main emphasis throughout this dissertation is on confidence bounds, simultaneous and/or individual, on parameters or parametric functions which are natural measures of departure from certain usual null hypotheses. Even where certain tests are proposed, as in chapter II, they are developed as a means to an end, namely that of obtaining confidence bounds on such parameters or parametric functions. A distinctive feature of the following investigations on simultaneous or separate confidence bounds consists in an attempt to obtain these, if possible, without running into difficult or intractable distribution problems.

All parent populations, from which our random samples are drawn, are assumed to be univariate or multivariate normal, as the case may be. On this assumption of the underlying distributions being normal, first certain general results are obtained and, then, a more detailed discussion is made of various problems in the sectors of variance components (both univariate and multivariate) and factor analysis.

Wherever the confidence bounds are obtained by inverting certain tests of hypotheses, it will be noted that the desirable power properties, if any, of the test will be carried over into similar desirable properties in terms of shortness of the confidence intervals, i.e., the probability of the confidence intervals cover-
ing wrong values of the parameters or parametric functions. Thus, for example, if the power of a test has the monotonicity property with respect to a certain parametric function (in the sense of\textsuperscript{1} \cite{34, 43}), then the associated confidence bounds on the parametric function will have the good property of \textit{monotonically increasing shortness}, i.e., the greater the deviation of a wrong value of the parametric function from its true value, the smaller is the probability (under the true value) that our interval includes the wrong value.

It will be seen during the course of this dissertation that fewer restrictions are required for obtaining separate confidence bounds than for obtaining simultaneous confidence bounds on the parameters or parametric functions, if we want, in either case, to avoid running into intractable distribution problems. This will be repeatedly exemplified in chapters III and IV. It may be noted that while simultaneous confidence bounds, which are physically very meaningful, have not received as much attention, they do not, however, involve anything new in principle. The general theory of simultaneous confidence interval estimation is subsumed under the well-known theory, now classic, due to Neyman \cite{30, 31, 32}.

In chapter I, the implications of certain results due to Roy and Bose \cite{38}, and Roy \cite{43, 44, 45}, are worked out fully.

\textsuperscript{1}The numbers in square brackets refer to the bibliography listed at the end.
These hitherto unnoticed implications, once they are brought out in explicit forms, provide, in a certain sense, tools for a complete analysis of that class of statistical problems which can be handled meaningfully in terms of confidence bounds on (a single or many) parametric functions which are natural measures of departure from the customary null hypotheses on the general multivariate normal distributions.

In chapter II, two tests are proposed. The first, derived by the heuristic union-intersection principle $\sum_{i=1}^{k^2}$, is for testing the composite hypothesis of equality of variances of $(k + 1)$ univariate normal populations in terms of $(k + 1)$ independent random samples from them. The proposed test turns out to be equivalent to the well-known Hartley $F_{max}$ test $\sum_{i=1}^{k^2}$ when all the samples are of the same size. For a certain choice of the acceptance (or, critical) region of the test, we show that the proposed test has the monotonicity property and is completely unbiased in the sense of Neyman. The second test, which is a multivariate extension of the first, is for the composite hypothesis of equality of $(k + 1)$ dispersion matrices of $(k + 1)$ non-singular $p$-variate normal populations. For both tests, the associated simultaneous confidence bounds on parametric functions which are natural measures of departure from the respective null hypotheses are also worked out.

In chapter III, we consider univariate variance components, or the Model II and the Mixed Model of anova (Analysis of Variance) in the terminology of Eisenhart $\sum_{i=13}^{12}$ and Crump $\sum_{i=12}^{12}$, which, in re-
cent years, have received a great deal of attention. A unified
general treatment for Model II and Mixed Model of anova, does not,
so far, appear to have been given in an explicit form. Such a treat-
mant is, of course, available for Model I of anova, and, among
others, Bose \cite{47} using vectors space methods, Bose \cite{57} using
matrix methods, and Roy \cite{43a} using matrix methods, have given a
general treatment of the problems here. On the other hand, much of
the published work on Model II of anova is concerned with analogues
of certain simple designs under Model I. In the first part of chap-
ter III of this dissertation, an attempt is made to give a general
treatment for Model II of anova along the lines of the general treat-
ment for Model I of anova. Since this inquiry was completed, it has
been brought to the attention of the author that a similar treatment
of the Mixed Model of anova was given by Bose \cite{57}, with the differ-
ence that, while this dissertation discusses the case of several
sets of random components, Bose discusses, in an explicit form, the
case of only one set of random components, although the case of
several sets of random components would also easily lend itself to
his general treatment. Little or no work has been done in this field
in the direction of simultaneous confidence interval estimation. The
results of Wald \cite{50} and Thompson \cite{48}, which were for special
cases and for individual parametric functions, have not been followed
up. The second part of chapter III is concerned with simultaneous
confidence interval estimation and gives some results in that area.
Chapter IV is concerned with multivariate variance components.

In the sector of multivariate anova, a development due to Roy [43] is available, which is essentially a multivariate generalization of the univariate Model I of anova. In chapter IV, a general multivariate Model II of anova is defined, but, for reasons discussed in section 4.6, we have to content ourselves, in the main, with a more restricted model. In terms of this relatively restricted model, results in point estimation, tests of hypotheses and the associated simultaneous confidence bounds are given, which are all believed to be new. For this restricted model, using certain results of Rao [35], which are extensions of results due to Bartlett [2] and Wald and Broockner [49], certain alternative confidence bounds are also given.

In the last chapter, factor analysis is viewed from the standpoint of analytic statistics. Much of this chapter consists in an attempt, from this standpoint, at a clarification and, in some sense, a partial reformulation of the problems, so as to make them feasible of statistical treatment and yet useful and meaningful to the factor analysts. Aside from properly posing the problems, so far as the actual statistical solution is concerned, it has not been possible to develop it in full for the general case in this dissertation. Most of the previous statistical work in this field has been on models which are, by no means, free from obscurity and reproach. Under such models, the statistical analysis so far has been in terms of point
estimation based, largely, on the maximum likelihood method, and in terms of testing of hypotheses based on the use of the likelihood ratio criterion. In this dissertation, however, starting from a reformulation of the problem, solutions in terms of confidence interval estimation of the relevant parameters are given in explicit form for the cases of three and four variates, and are indicated for the general case. A fuller treatment of the latter will be given in subsequent papers. Finally, in chapter V, an independent solution of a distribution problem which occurs repeatedly in this dissertation is given, using methods that are similar to those of $\mathcal{L}^{32}$, $\mathcal{L}^{34}$ and $\mathcal{L}^{40}$. 
NOTATION AND PRELIMINARIES

"\( \mathbf{x}(m \times 1) \)" denotes a column vector with \( m \) elements (all real unless otherwise stated), and its prime denotes its transpose, a row vector.

"\( \mathbf{X}(p \times q) \)" denotes a matrix (with all real elements unless otherwise stated) with \( p \) rows and \( q \) columns, and its prime denotes its transpose. In addition, we shall use the following notation to denote certain special vectors and matrices:

"\( \mathbf{l}(m \times 1) \)" denotes a column vector of \( m \) unities.

"\( \mathbf{0}(m \times 1) \)" denotes a column vector of \( m \) zeros.

"\( \mathbf{I}(m) \)" denotes the identity matrix of order \( m \).

"\( \mathbf{J}(p \times q) \)" denotes a matrix whose elements are all unity.

"\( \mathbf{D}_{\mathbf{a}}(m \times m) \)" denotes a diagonal matrix whose non-zero elements are \( a_1, a_2, \ldots, a_m \).

"\( \tilde{T}(m \times m) \)" denotes a triangular matrix whose non-zero elements are along and below the diagonal.

"\( \mathbf{O}(p \times q) \)" denotes a matrix whose elements are all zero.

"\( |a| \)" denotes the absolute value of the scalar, \( a \).

"\( |M(p \times p)| \)" denotes the determinant of the matrix \( M(p \times p) \).

"\( M^{-1}(p \times p) \)" denotes the inverse of the matrix \( M(p \times p) \).

"\( \text{tr}(X) \)" denotes the rank of the matrix \( X(p \times q) \).

"\( \text{tr} M \)" denotes the trace of the matrix \( M(p \times p) \).

"\( \text{tr}_s M \)" denotes the sum of all \( s \)th order principal minors of the matrix \( M(p \times p) \), \( s \leq p \).
"c(M)" stands for "all the characteristic roots of the matrix M", and \( c_{\text{max}}(M), c_{\text{min}}(M) \) stand respectively for the largest and smallest characteristic roots of \( M \).

"\((p \times p) \times (q \times q)\)" denotes the Kronecker product of the matrices \( A \) and \( B \), viz., the matrix

\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1q} \\
  a_{21} & a_{22} & \cdots & a_{2q} \\
  \cdots & \cdots & \cdots \\
  a_{q1} & a_{q2} & \cdots & a_{qq}
\end{bmatrix}
\]

with \( pq \) rows and \( pq \) columns.

"\( J(X;Y) \)" denotes the Jacobian of a transformation from the matrix \( X \) to the matrix \( Y \).

"a.e." stands for "almost everywhere".

"anova" stands for "analysis of variance".

"\( \text{min}(p,q) \)" stands for "the smaller of \( p \) and \( q \)".

"NSC" stands for "necessary and sufficient condition".

"p.d." stands for "positive definite".

"p.s.d." stands for "positive semi-definite".

\( \Rightarrow \) stands for "implies" and \( \iff \) stands for "is equivalent to".

\( \cap \) and \( \cup \) stand respectively for "intersection" and "union".

\( A \subset B \) denotes \( A \) is a subset of \( B \).

\( \text{det}(M(p \times p)) \) is \( \bot \) means that the matrix \( M \) is orthogonal.

All other special notations which may be used will be explained clearly where they are used.
CHAPTER I

FURTHER IMPLICATIONS OF CERTAIN PREVIOUS RESULTS
IN MULTIVARIATE CONFIDENCE BOUNDS

1.1 Introduction.

In this chapter results, which are latent in solutions given in 43, 44, 45 for simultaneous confidence bounds on parametric functions connected with multivariate normal populations, are fully worked out. The earlier results were obtained by inverting tests of certain customary null hypotheses, which tests were obtained by the union-intersection principle 42. The tests themselves are verified to have some desirable properties, (unbiasedness, monotonicity or near monotonicity, admissibility, etc.), which are, therefore, reflected in certain good properties, (in terms of shortness, i.e. probability of covering wrong values, etc.), of the confidence bounds. The results obtained in the present chapter (as the earlier results) are concerned with parametric functions, which are natural measures of departure (distance functions) from the null hypotheses in the respective situations. The hitherto unnoticed confidence bounds, which are presented here, are on characteristic roots connected with

(i) one population dispersion matrix;
(ii) two population dispersion matrices;
(iii) the regression matrix of a p-set on a q-set in a (p+q)-
     variate normal;
and (iv) the multivariate linear hypothesis on means, including, in particular, the univariate case and the problem of discriminant analysis.

Before proceeding to obtain these results a few lemmas, which are repeatedly used, are stated, while their proofs are either explicitly or implicitly contained in \(43\).

1.2 Certain Lemmas.

**Lemma 1.2a:** The non-zero characteristic roots of \( A(p \times q) \times B(q \times p) \) are the same as those of \( B(q \times p) A(p \times q) \).

For a proof see \(43\).

**Lemma 1.2b:** If \( A, B \) and \( C \) are real symmetric p.d. matrices of order \( p \) each, then

\[
(1.2.1) \quad c_{\min}(AB^{-1})c_{\min}(BC) \leq \forall c(AC) \leq c_{\max}(AB^{-1})c_{\max}(BC)
\]

For a proof see \(43\).

**Lemma 1.2c:** If \( B(p \times p) \) is symmetric p.d. and \( A(p \times p) \) is symmetric and at least p.s.d. of rank \( r(\leq p) \), then, for all nonnull \( a(p \times 1) \)'s, the stationary values of \( a^t A_a / a^t B_a \) (under variation of \( a \)) are the characteristic roots \( c(AB^{-1}) \).

For a proof see \(43\).

**Lemma 1.2d:** The statement \( \mu^t (1 \times p) \times (p \times l) \leq h, (h > 0) \) is exactly equivalent to the statement \( \| a^t x \| + 1/h \) for all unit vectors \( a(p \times 1) \).

The proof may be obtained by using Cauchy's inequality as follows:
If $x^1x \leq h$, $(h > 0)$, then, by Cauchy's inequality
\[(a^1x)^2 \leq (a^1a) (x^1x)\]
i.e., $\leq x^1x$ for all unit vectors $a(p \times 1)$
\[
\leq h
\]
\[
\because |a^1x| \leq +\sqrt{h} \text{ for all unit vectors } a(p \times 1).
\]
Next, suppose $|a^1x| \leq +\sqrt{h}$ for all unit vectors $a(p \times 1)$, and consider the particular unit vector $a(p \times 1)$ whose ith element
\[= x_1/\sqrt{\sum_{i=1}^{p} x_i^2}, (i = 1, 2, \ldots, p).\]
Then we have
\[
\frac{x^1x}{\sqrt{x^1x}} \leq +\sqrt{h}
\]
i.e., $x^1x \leq h$, $(h > 0)$.

**Lemma 1.2**: If $a(p \times q)$ is a matrix of rank $M \leq \min(p, q)$, then, for all unit vectors $a_1(p \times 1)$ and $a_2(q \times 1)$, the stationary values of $\int a_1^1 A a_2 - J^2$ (under variation of $a_1$ and $a_2$) are the characteristic roots $\sigma(\mathbb{A}^1a^1)$. It follows, of course, that the stationary values of $a_1^1 A a_2$ are $\pm \sigma^{1/2}(\mathbb{A}^1a^1)$.

**Proof**: This result can be proved in several ways using other known results, or independently. An independent proof is given here.

Let $\sigma = \int a_1^1 A a_2 - J^2$

We want to find the stationary values of $\sigma$ under variation of $a_1(p \times 1)$ and $a_2(q \times 1)$ subject to $a_{1-1}^1 = 1$ and $a_{2-2}^1 = 1$, i.e., we want to find the stationary values of $\sigma - \lambda(a_{1-1}^1 - 1) - \mu(a_{2-2}^1 - 1)$, where $\lambda, \mu$ are Lagrangian multipliers.
Differentiating with respect to the vectors $a_1$ and $a_2$ and equating to zero we have,

\[(1.2.2) \quad 2\sqrt{a_1'A_2} \cdot A_2 - 2\lambda a_1 = 0(p \times 1)\]
\[2\sqrt{a_1'A_2} \cdot a_1 - 2\mu a_2 = 0'(1 \times q)\]

i.e., $\lambda = \mu = \sqrt{a_1'A_2} \cdot a_1$, so that (1.2.2) may be rewritten as

\[A_2 - \sqrt{a_1'A_2} \cdot a_1 = 0(p \times 1)\]
\[\sqrt{a_1'A_2} \cdot a_2 - A_1a_1 = 0(q \times 1)\]

For a non-trivial solution of $a_1, a_2$ we, therefore, must have,

\[
\begin{vmatrix}
A & -(a_1'A_2)I(p) \\
(a_1'A_2)I(q) & -A'
\end{vmatrix}
= 0
\]

or,

\[
\begin{vmatrix}
I(p) & \frac{1}{c}A \\
\frac{1}{c}A' & I(q)
\end{vmatrix}
= 0, \text{ remembering}
\]

\[c = \sqrt{a_1'A_2} \cdot a_2 = 0, \text{ i.e., } |I(p) - \frac{1}{c}AA'| = 0, \text{ since } \frac{p}{q} \begin{vmatrix}
A & B \\
C & D
\end{vmatrix} = |D| |A-BD^{-1}C|,
\]

if $D(q \times q)$ is non-singular $\sqrt{a_1'A_2}$.

Hence the lemma.

\underline{Lemma 1.2f:} We carry over from set-theoretic logic that the
statement "if $E_1$ then $E_2$" is equivalent to $E_1 \subset E_2$, and, hence, $P(E_2) \leq P(E_1)$.

1.3 Confidence bounds on roots connected with $N(\mu(p \times 1), \Sigma(p \times p))$.

We start from an earlier statement,

(1.3.1) $nc_{1\alpha}^{-1}(p, n) \geq \alpha'ZS^{-1}a \geq nc_{2\alpha}^{-1}(p, n)$

for all nonnull $a(p \times 1)$, where $(n + 1)$ is the sample size, $S(p \times p)$ is the sample dispersion matrix (a.e.p.d.), and $c_{1\alpha}(p, n)$, $c_{2\alpha}(p, n)$ are obtained from the joint distribution of $c_1 = c_{\text{min}}(nS)$ and $c_p = c_{\text{max}}(nS)$ such that $P(c_{1\alpha} \leq c_1 \leq c_p \leq c_{2\alpha} \mid \Sigma = \Sigma_0 = I(p)) = (1 - \alpha)$. If $c_{1\alpha}$ and $c_{2\alpha}$ are chosen subject only to this last condition, then the confidence bounds obtained below will still be valid, the only difference being that they will not, necessarily, have the property of monotonically increasing shortness possessed by those that are based on $c_{1\alpha}$, $c_{2\alpha}$ determined such that the further condition of local unbiasedness, (in the sense of Neyman), of the test for the null hypothesis, $H_0: \Sigma = I(p)$, is satisfied. The distribution of $c_1$ and $c_p$, which is needed for the evaluation of $P(c_0 \leq c_1 \leq c_p \leq c_0' \mid \Sigma = I(p))$, may be obtained as the limiting case of a distribution explicitly obtained in $\sim \chi^2, \chi^2, \chi^2$ for the case of two dispersion matrices. The problem may, however, be attacked independently, and this is done in the last chapter of the present inquiry, where we run into the same distribution problem again.

It is easily seen that the statement (1.3.1) \(\Longleftrightarrow\)
\begin{equation}
\lambda_1 \frac{a'Sa}{a'a} \geq \lambda_2 \frac{a'Sa}{a'a}
\end{equation}

for all nonnull \(a(p \times 1)\), where \(\lambda_1, \lambda_2\) stand for \(\text{nc}_{1\alpha}(p,n)\) and \(\text{nc}_{2\alpha}(p,n)\) respectively. Choosing \(\mathbf{a}\) so as to minimize \(\frac{a'Sa}{a'a}\) we observe, using lemma 1.2c, that the first part of the inequality (1.3.2) implies that \(\lambda_1 c_{\min}(S) \geq c_{\min}(\Sigma)\); and choosing \(\mathbf{a}\) so as to minimize \(\frac{a'Sa}{a'a}\) we see that the second part of the inequality implies that \(c_{\min}(\Sigma) \geq \lambda_2 c_{\min}(S)\). Likewise, choosing \(\mathbf{a}\) so as to maximize \(\frac{a'Sa}{a'a}\) we observe that the first part of the inequality (1.3.2) implies that \(c_{\max}(\Sigma) \leq \lambda_1 c_{\max}(S)\); and choosing \(\mathbf{a}\) so as to maximize \(\frac{a'Sa}{a'a}\), that the second part of the inequality implies that \(\lambda_2 c_{\max}(S) \leq c_{\max}(\Sigma)\). Thus (1.3.2) \(\implies\)

\begin{equation}
\lambda_1 c_{\min}(S) \geq c_{\min}(\Sigma) \geq \lambda_2 c_{\min}(S)
\end{equation}

and

\begin{equation}
\lambda_1 c_{\max}(S) \geq c_{\max}(\Sigma) \geq \lambda_2 c_{\max}(S)
\end{equation}

Since (1.3.1), or (1.3.2), has a confidence coefficient \(= (1 - \alpha)\), we use lemma 1.2f and notice that (1.3.3) has a confidence coefficient \(\geq (1 - \alpha)\). Also it might be noted that (1.3.3) implies an earlier \(\text{W}\) statement,

\begin{equation}
\lambda_1 c_{\max}(S) \geq c_{\max}(\Sigma) \geq c_{\min}(\Sigma) \geq \lambda_2 c_{\min}(S)
\end{equation}
Going back to (1.3.2) let us take $\mathbf{s}(p \times 1)$ such that the $i$th component ($i = 1, 2, \ldots, p$) is zero. Then arguing exactly as above, we observe that (1.3.2) also implies

$$\lambda_1^{c_{\min}}(\mathbf{s}(i)) \geq c_{\min}(\mathbf{S}(i)) \geq \lambda_2^{c_{\min}}(\mathbf{s}(i))$$

(1.3.5)

and

$$\lambda_1^{c_{\max}}(\mathbf{s}(i)) \geq c_{\max}(\mathbf{S}(i)) \geq \lambda_2^{c_{\max}}(\mathbf{s}(i))$$

for $i = 1, 2, \ldots, p$, where $\mathbf{s}(i)$ and $\mathbf{S}(i)$ stand respectively for the truncated sample and population dispersion matrices obtained by cutting out the $i$th variate.

In the same way, if we take $\mathbf{s}(p \times 1)$ such that the $i$th and $j$th components ($i \neq j$) are zero, and then argue in a similar manner, we see that (1.3.2) also implies.

$$\lambda_1^{c_{\min}}(\mathbf{s}(i, j)) \geq c_{\min}(\mathbf{S}(i, j)) \geq \lambda_2^{c_{\min}}(\mathbf{s}(i, j))$$

(1.3.6)

$$\lambda_1^{c_{\max}}(\mathbf{s}(i, j)) \geq c_{\max}(\mathbf{S}(i, j)) \geq \lambda_2^{c_{\max}}(\mathbf{s}(i, j))$$

for $i \neq j = 1, 2, \ldots, p$, where $\mathbf{s}(i, j)$ and $\mathbf{S}(i, j)$ stand respectively for the truncated sample and population dispersion matrices obtained by cutting out the $i$th and $j$th variates.

We can continue this process on to the stage of cutting out any $(p - 1)$ variates, that is, where any one variate is retained, which gives us just the confidence statements on variances in the univariate case. Thus it is seen that (1.3.2) implies a pair of statements (1.3.3), $p$ pairs of statements like (1.3.5), $\binom{p}{2}$ pairs of statements like (1.3.6), and so on down to $\binom{p}{p-1}$, i.e., $p$ state-
ments involving only one variate. All such statements will, therefore, by lemma 1.2f, have a joint confidence coefficient \( \geq (1 - \alpha) \). Also, at each stage of truncation, we have statements like (1.3.4), which are implied by the corresponding pair of statements, and all such statements will also have a joint confidence coefficient \( \geq (1 - \alpha) \). It may be noted that the statements obtained above provide us, from a certain standpoint, with a complete analysis of what the psychologists call the problem of principal components \( \mathcal{L}19\mathcal{L} \).

1.4  **Confidence bounds on roots connected with** \( \Sigma_1(p \times p) \) and \( \Sigma_2(p \times p) \) of \( N(\varepsilon_1(p \times 1), \Sigma_1) \) and \( N(\varepsilon_2(p \times 1), \Sigma_2) \).

Let \( S_1(p \times p) \) and \( S_2(p \times p) \) be the sample dispersion matrices based on samples of sizes \((n_1+1)\) and \((n_2+1)\) from the two populations. We start, without any loss of generality for our purpose, from an earlier \( \mathcal{L}14\mathcal{L} \) statement in the canonical form

\[
(l.4.1) \quad \frac{n_1}{n_2} c_{1a}(p,n_1,n_2) \geq \text{all } c(S_2^{-1} \Sigma_1^{-1} S_2^{-1}) \geq \frac{n_1}{n_2} c_{2a}(p,n_1,n_2)
\]

where the \( p \) elements of the diagonal matrix \( D(p \times p) \) are the characteristic roots \( c(\Sigma_2^{-1}) \), and \( c_{1a}(p,n_1,n_2) \), \( c_{2a}(p,n_1,n_2) \) are such that

\[
\frac{n_1}{n_2} c_{1a}(p,n_1,n_2) \leq c_{\text{min}}(\Sigma_2^{-1} S_1 S_2^{-1}) \leq c_{\text{max}}(\Sigma_2^{-1} S_1 S_2^{-1}) \leq c_{2a}(p,n_1,n_2) \quad \Sigma_1 = \Sigma_2 = (1 - \alpha). \]

Here, as in 1.3, the confidence bounds that will be obtained are valid under just this one condition on \( c_{1a} \) and \( c_{2a} \), except that, if we want to insure the good property of monotonically increasing shortness of the statements, then
we might impose the further condition of local unbiasedness, (in the
sense of Neyman), on the region defined by $c_{1\alpha}$ and $c_{2\alpha}$ for the
test of the associated null hypothesis $H_0: \Sigma_1 = \Sigma_2$. The distribu-
tion problem involved here is completely solved $\gamma_1, \gamma_0, \mu_3 \gamma$ and
the tabulation is being carried on at present in England at the
London School of Economics. These tables, which will make these
results on confidence bounds, and the earlier results for tests of
multivariate hypotheses, applicable in practical situations, are ex-
pected to be available for general use in the next year or two.

Writing $\lambda_1 = \frac{n_1}{n_2}c_{1\alpha}(p, n_1, n_2)$ and $\lambda_2 = \frac{n_1}{n_2}c_{2\alpha}(p, n_1, n_2)$, we can
rewrite (1.4.1) as,

$$\lambda_1 \geq \text{all } c(S_2D/\gamma_1S_1^{-1}D/\gamma) \geq \lambda_2$$  

(1.4.2)

or,

$$\lambda_1 \geq \text{all } c(S_2S_1^{-1}D/\gamma_1S_1^{-1}D/\gamma) \geq \lambda_2$$

or, using lemma 1.2c, as

$$\lambda_1 \geq \frac{a's_1s_2^{-1}a}{a'a} \geq \frac{a's_1D/\gamma_1S_1^{-1}D/\gamma a}{a'a} \geq \lambda_2 \frac{a's_1S_2^{-1}a}{a'a}$$  

(1.4.3)

for all nonnull $a(p \times 1)$, which, therefore, is a set of simultaneous
statements with a joint probability $= (1 - \alpha)$. Now choosing $a$ so
as to maximize the middle term of (1.4.3), we observe that the first
part of (1.4.3) implies that $\lambda_1c_{\max}(S_1S_2^{-1}) \geq c_{\max}(S_1D/\gamma_1S_1^{-1}D/\gamma)$; and
choosing $a$ so as to minimize the middle term of (1.4.3), we observe
that the second part of (1.4.3) implies that
\[ c_{\text{min}}(S_{1}/\gamma S_{1}^{\text{D}}/\gamma) \geq \lambda_{2}c_{\text{min}}(S_{1}S_{2}^{-1}). \] Thus (1.4.3) \[ \implies \]

(1.4.4)

\[ \lambda_{1}c_{\text{max}}(S_{1}S_{2}^{-1}) \geq c_{\text{max}}(S_{1}/\gamma S_{1}^{\text{D}}/\gamma) \geq c_{\text{min}}(S_{1}/\gamma S_{1}^{\text{D}}/\gamma) \geq \lambda_{2}c_{\text{min}}(S_{1}S_{2}^{-1}). \]

Using lemma 1.2b it can be seen that

(1.4.5) \[ c_{\text{max}}(S_{1}/\gamma S_{1}^{\text{D}}/\gamma) \geq \text{all } c(\gamma) \geq c_{\text{min}}(S_{1}/\gamma S_{1}^{\text{D}}/\gamma), \]

so that, we finally have that (1.4.3) \[ \implies \]

(1.4.6)

\[ \lambda_{1}c_{\text{max}}(S_{1}S_{2}^{-1}) \geq \text{all } c(\gamma_{1}^{\text{D}}) \geq \lambda_{2}c_{\text{min}}(S_{1}S_{2}^{-1}), \]

which, therefore, is a confidence statement with a confidence coefficient \[ \geq (1 - \alpha). \]

It may be observed that, since

\[ c_{\text{min}}(S_{1})c_{\text{min}}(S_{2}^{-1}) \leq \text{all } c(S_{1}S_{2}^{-1}) \leq c_{\text{max}}(S_{1})c_{\text{max}}(S_{2}^{-1}), \]

therefore, (1.4.6) \[ \implies \]

\[ \lambda_{1}c_{\text{max}}(S_{1})c_{\text{max}}(S_{2}^{-1}) \geq \text{all } c(\gamma_{1}^{\text{D}}) \geq \lambda_{2}c_{\text{min}}(S_{1})c_{\text{min}}(S_{2}^{-1}), \]

(1.4.7) i.e., \[ \lambda_{1} \frac{c_{\text{max}}(S_{1})}{c_{\text{min}}(S_{2}^{-1})} \geq \text{all } c(\gamma_{1}^{\text{D}}) \geq \lambda_{2} \frac{c_{\text{min}}(S_{1})}{c_{\text{max}}(S_{2}^{-1})}, \]

which is another set of simultaneous confidence bounds, (obviously wider than (1.4.6)), with a confidence coefficient \[ \geq (1 - \alpha) . \]

Another derivation of (1.4.7) is given later on in this section.

Going back to (1.4.3) and taking, as in section 1.3, \( \gamma(p \times 1) \)
such that the ith component is zero, we may reason in the same way as from (1.4.3) - (1.4.6) and observe that (1.4.3) also implies,

\[(1.4.8) \lambda_1 \max\left( S_1^{(1)} S_2^{(1)} \right) \geq \alpha \left( \sum_1^{(1)} \sum_2^{(1)} \right) \geq \lambda_2 \min\left( S_1^{(1)} S_2^{(1)} \right), \]

where $S_1^{(i)}$, $S_2^{(i)}$, $\Sigma_1^{(i)}$ and $\Sigma_2^{(i)}$ stand for the truncated sample and population dispersion matrices obtained by cutting out any ith variate ($i = 1, 2, \ldots, p$). Likewise, just as in section 1.3, in an obvious notation, we also see that (1.4.3) implies,

\[(1.4.9) \lambda_1 \max\left( S_1^{(i,j)} S_2^{(i,j)} \right) \geq \alpha \left( \sum_1^{(i,j)} \sum_2^{(i,j)} \right) \geq \lambda_2 \min\left( S_1^{(i,j)} S_2^{(i,j)} \right) \]

for $i \neq j = 1, 2, \ldots, p$, and so on. The process of truncation can be carried on till we reach the stage where any ($p - 1$) variates have been cut out, i.e., any one variate is retained, at which stage we shall have just the confidence bounds on variance ratios in the univariate case. We thus have, with a joint confidence coefficient $\geq (1 - \alpha)$, the confidence statement (1.4.6), $p$ confidence statements like (1.4.8), ($\binom{p}{2}$) confidence statements like (1.4.9), and so on.

We next proceed to obtain a slightly different kind of set of confidence bounds, which will imply confidence statements like (1.4.7). We rewrite (1.4.2) in the form,

\[(1.4.10) \quad \lambda_1 \frac{a' S_2^{-1} a}{a'a} \geq \frac{a'D S_1^{-1} D a}{a'D a} \geq \lambda_2 \frac{a'S_2^{-1} a}{a'a} \]

for all nonnull $a(p \times 1)$. 

\[\]
Since \( c_{\text{max}}(S_{2}^{-1}) \geq \forall c(S_{2}^{-1}) \geq c_{\text{min}}(S_{2}^{-1}) \), therefore, (1.4.10) obviously implies

\[ (1.4.11) \quad \frac{\lambda_{1}c_{\text{max}}(S_{2}^{-1})}{c_{\text{min}}(S_{2}^{-1})} \geq \frac{a^{1}D_{y}a}{a^{1}a} \geq \frac{\lambda_{2}c_{\text{min}}(S_{2}^{-1})}{c_{\text{max}}(S_{2}^{-1})} \]

for all nonnull \( a(p \times 1) \), or, equivalently

\[ (1.4.12) \quad \frac{\lambda_{1}}{c_{\text{min}}(S_{2}^{-1})} \geq \frac{a^{1}D_{y}a}{a^{1}a} \geq \frac{\lambda_{2}}{c_{\text{max}}(S_{2}^{-1})} \]

for all nonnull \( a(p \times 1) \). Now choosing \( a \) so as to maximize \( \frac{a^{1}D_{y}a}{a^{1}a} \), we observe that the first part of (1.4.12) implies that,

\[ \frac{\lambda_{1}c_{\text{max}}(S_{1})}{c_{\text{min}}(S_{2}^{-1})} \geq c_{\text{max}}(D_{y}) = c_{\text{max}}(\Sigma_{1}^{\Sigma_{2}^{-1}}) \]

and choosing \( a \) so as to maximize \( \frac{a^{1}S_{1}a}{a^{1}a} \), we observe that the second part of (1.4.12) implies that \( c_{\text{max}}(D_{y}) = c_{\text{max}}(\Sigma_{1}^{\Sigma_{2}^{-1}}) \geq \lambda_{2} \frac{c_{\text{max}}(S_{1})}{c_{\text{max}}(S_{2}^{-1})} \). Similarly, choosing \( a \) to minimize \( \frac{a^{1}S_{1}a}{a^{1}a} \), we observe that the first part of (1.4.12) implies that

\[ \frac{\lambda_{1}c_{\text{min}}(S_{1})}{c_{\text{min}}(S_{2}^{-1})} \geq c_{\text{min}}(\Sigma_{1}^{\Sigma_{2}^{-1}}) \]

and choosing \( a \) so as to minimize \( \frac{a^{1}D_{y}a}{a^{1}a} \), the second part of (1.4.12) implies that

\[ c_{\text{min}}(\Sigma_{1}^{\Sigma_{2}^{-1}}) \geq \lambda_{2} \frac{c_{\text{min}}(S_{1})}{c_{\text{max}}(S_{2}^{-1})} \]. Hence, we see that (1.4.12), or (1.4.10) implies the pair of statements,
\[
\lambda_1 \frac{c_{\text{min}}(S_1)}{c_{\text{min}}(S_2)} \geq c_{\text{max}}(\Sigma_1 \Sigma_2^{-1}) \geq \lambda_2 \frac{c_{\text{max}}(S_1)}{c_{\text{max}}(S_2)}
\]  
(1.4.13)

and

\[
\lambda_1 \frac{c_{\text{min}}(S_1)}{c_{\text{min}}(S_2)} \geq c_{\text{min}}(\Sigma_1 \Sigma_2^{-1}) \geq \lambda_2 \frac{c_{\text{min}}(S_1)}{c_{\text{max}}(S_2)}
\]

which, therefore, is a pair of statements with joint confidence coefficient \( \geq (1 - \alpha) \). It can now be seen that (1.4.13) implies (1.4.7).

Going back to (1.4.10) and carrying out the process of truncation just as before, we observe that (1.4.10) implies, in addition to (1.4.13), \( p \) pairs of statements like,

\[
\lambda_1 \frac{c_{\text{max}}(S_1^{(i)})}{c_{\text{min}}(S_2^{(i)})} \geq c_{\text{max}}(\Sigma_1^{(i)} \Sigma_2^{(i)-1}) \geq \lambda_2 \frac{c_{\text{max}}(S_1^{(i)})}{c_{\text{max}}(S_2^{(i)})}
\]  
(1.4.13)

\[
\lambda_1 \frac{c_{\text{min}}(S_1^{(i)})}{c_{\text{min}}(S_2^{(i)})} \geq c_{\text{min}}(\Sigma_1^{(i)} \Sigma_2^{(i)-1}) \geq \lambda_2 \frac{c_{\text{min}}(S_1^{(i)})}{c_{\text{max}}(S_2^{(i)})}
\]

for \( i = 1, 2, \ldots, p; \) \( \binom{p}{2} \) pairs of statements like

\[
\lambda_1 \frac{c_{\text{max}}(S_1^{(i,j)})}{c_{\text{min}}(S_2^{(i,j)})} \geq c_{\text{max}}(\Sigma_1^{(i,j)} \Sigma_2^{(i,j)-1}) \geq \lambda_2 \frac{c_{\text{max}}(S_1^{(i,j)})}{c_{\text{max}}(S_2^{(i,j)})}
\]  
(1.4.14)

\[
\lambda_1 \frac{c_{\text{min}}(S_1^{(i,j)})}{c_{\text{min}}(S_2^{(i,j)})} \geq c_{\text{min}}(\Sigma_1^{(i,j)} \Sigma_2^{(i,j)-1}) \geq \lambda_2 \frac{c_{\text{min}}(S_1^{(i,j)})}{c_{\text{max}}(S_2^{(i,j)})}
\]

for \( i \neq j = 1, 2, \ldots, p; \) and so on. The joint confidence coefficient of all these statements is again \( \geq (1 - \alpha) \).
The analysis given above, from a certain standpoint, provides a partial analysis of problems which occur in the multivariate generalization of the customary variance components analysis in univariate analysis of variance and covariance. This generalization is considered in greater length in chapter IV of this inquiry.

1.5 Confidence bounds connected with the regression matrix $\beta(p \times q)$ of a p-set on a q-set in a $(p + q)$-variate normal,

$$\mathbf{N} \bigg( \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \bigg) \bigg( \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \bigg)^T \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{12} & \mathbf{Z}_{22} \end{bmatrix} \bigg( \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{12} & \mathbf{Z}_{22} \end{bmatrix} \bigg)^{-1} \mathbf{J}.$$ 

We have a random sample of size $(n + 1)$ and the sample dispersion matrix is $\mathbf{S}(p+q \times p+q) = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{12} & \mathbf{S}_{22} \end{bmatrix}$. 

Let $\mathbf{B}(p \times q) = \mathbf{S}_{12}^{-1} \mathbf{S}_{22}^{-1}$ be the sample regression matrix of the p-set on the q-set, just as $\beta(p \times q) = \mathbf{S}_{12}^{-1} \mathbf{S}_{22}^{-1}$ is the population regression matrix. Let us write $\mathbf{S}_{1.2}(p \times p) = \mathbf{S}_{11}^{-1} \mathbf{S}_{12}^{-1} \mathbf{S}_{22}^{-1} \mathbf{S}_{12}$. 

Starting from the statement,

$$(1.5.1) \quad \sqrt{-\mathbf{S}_{1.2}^{-1} (\mathbf{B} - \beta) \mathbf{S}_{22} (\mathbf{B} - \beta)^T} \leq \mathbf{J} = \mathbf{J}_{\mathbf{\beta}}$$

with probability $= (1 - \alpha)$, where $\mathbf{\beta} = \mathbf{\beta}_n(p, q, n)$ is such that $\mathbf{P} \sqrt{-\mathbf{J}} \geq \mathbf{\beta}_n = \mathbf{\beta}_n(p, q, n)$ is such that $\mathbf{P} \sqrt{-\mathbf{J}} \geq \mathbf{\beta}_n = \mathbf{\beta}_n(p, q, n)$ is such that $\mathbf{P} \sqrt{-\mathbf{J}} = \mathbf{\beta}_n$, it has been shown $\sqrt{h_3, h_4} \mathbf{J}$ that the statement,
(1.5.2) \[ \alpha c \gamma (B - \beta)(B' - \beta') \leq \lambda^2 \omega_{max}(S_{1,2}) c_{max}(S_{2,2}^{-1}) \]

with probability \( \geq (1 - \alpha) \), where \( \lambda^2 = \theta_a / \theta_a' \), can be obtained.
Also, using lemma 1.2e, we can obtain the equivalent set of statements,

(1.5.3) \[ a_1 B a_2 - \lambda c_{max}(S_{1,2}) c_{max}(S_{2,2}^{-1}) \leq a_1^1 \beta a_2 \leq a_1 B a_2 \]

+ \[ \lambda c_{max}(S_{1,2}) c_{max}(S_{2,2}^{-1}) \]

for all unit vectors \( a_1(p \times 1) \) and \( a_2(q \times 1) \).

Going back to lemma 1.2e again, we notice that, with respect to variation over \( a_1 \) and \( a_2 \), the maximum values of \( a_1^1 B a_2 \) and \( a_1^1 \beta a_2 \) are respectively \( c_{max}(BB') \) and \( c_{max}(\beta \beta') \). First choosing \( a_1 \) and \( a_2 \) so as to maximize \( a_1^1 B a_2 \), and then choosing them so as to maximize \( a_1^1 \beta a_2 \), and arguing in the same way as in the two preceding sections 1.3 and 1.4, we note that (1.5.3) \( \Rightarrow \)

(1.5.4) \[ c_{max}(BB') - \lambda c_{max}(S_{1,2}) c_{max}(S_{2,2}^{-1}) \leq c_{max}(\beta \beta') \leq c_{max}(BB') \]

+ \[ \lambda c_{max}(S_{1,2}) c_{max}(S_{2,2}^{-1}) \]

which, therefore, is a confidence statement with a confidence coefficient \( \geq (1 - \alpha) \). It may be noted that the middle term of (1.5.4), i.e., \( c_{max}(\beta \beta') \), is zero if, and only if, the null hypothesis \( H_0: \Sigma_{12} = 0 \) is true. Thus \( c_{max}(\beta \beta') \) is a measure of departure
(distance function) from the null hypothesis.

Now going back to (1.5.1), we may rewrite it in the equivalent form,

\[(1.5.5) \quad \frac{s' (B' - B) S_{22} (B' - B') s}{s' S_{11} s} \leq \lambda^2\]

for all nonnull \(s(p \times 1)\), which, therefore, is a simultaneous confidence statement with confidence coefficient \(= (1 - \alpha)\). As before we take \(s(p \times 1)\) such that the ith component is zero, define \(s^{(i)}\), \(B^{(i)}\) and \(\beta^{(i)}\) as the truncated matrices obtained by cutting out the ith variate of the \(p\)-set, and observe that (1.5.5) also implies,

\[(1.5.6) \quad c_{\max}^{1/2}(B^{(i)}B^{(i)\prime}) - \lambda c_{\max}^{1/2}(S^{(i)}_1S^{-1}_2) \leq c_{\max}^{1/2}(\beta^{(i)}\beta^{(i)\prime}) \leq c_{\max}^{1/2}(B^{(i)}B^{(i)\prime}) + \lambda c_{\max}^{1/2}(S^{(i)}_1S^{-1}_2)\]

for \(i = 1, 2, \ldots, p\), which, again, is a confidence interval statement with confidence coefficient \(\geq (1 - \alpha)\). Likewise, just as before we observe that (1.5.5) also implies,

\[(1.5.7) \quad c_{\max}^{1/2}(B^{(i,j)}B^{(i,j)\prime}) - \lambda c_{\max}^{1/2}(S^{(i,j)}_1S^{-1}_2) \leq c_{\max}^{1/2}(\beta^{(i,j)}\beta^{(i,j)\prime}) \leq c_{\max}^{1/2}(B^{(i,j)}B^{(i,j)\prime}) + \lambda c_{\max}^{1/2}(S^{(i,j)}_1S^{-1}_2)\]

for \(i \neq j = 1, 2, \ldots, p\); and so on. We can carry this process on until we cut out any \((p - 1)\)-variates, i.e., we retain any one variate, at which stage our confidence bounds will be on the square-root of
the sum of squares of the partial regressions of one variate on the
q others. We thus have, with a joint confidence coefficient
> (1 - α), the statement (1.5.4), p statements like (1.5.6), (p
2)
statements like (1.5.7), and so on. Obviously, the process of
truncation described above could be carried out with respect to the
q-set as well. The technique involves nothing new and hence the re-
sults obtained thus are not explicitly given here.

A logical question that might be asked is, "Why not start the
truncation process with the $\mathbf{a}_1(p \times 1)$ of (1.5.3) rather than with
the $\mathbf{a}(p \times 1)$ of (1.5.5)?". Theoretically, this can be done and we
should obtain statements like (1.5.6) and (1.5.7), except for the
difference that the second term in the bounds will not involve the
truncated $S_{12}^{(1)}$, $S_{12}^{(i,j)}$, but the same $S_{12}$ as in (1.5.4). Hence,
the bounds obtained by the first method of truncating using
$\mathbf{a}(p \times 1)$ may be expected to be closer than those obtained by the
other method, at each stage.

1.6 Confidence bounds on roots connected with multivariate linear
hypotheses on means.

(1) On $\mathbf{\xi}(p \times 1)$ of $N(\mathbf{\xi}(p \times 1); \mathbf{Z}(p \times p))$, where $\mathbf{Z}(p \times p)$ is un-
known.

We have a random sample of size $(n + 1)$ and the sample mean
vector $\mathbf{\bar{X}}(p \times 1)$ and dispersion matrix $S(p \times p)$.

We start from an earlier statement,
(1.6.1) \[ \frac{a^T X - \frac{T^2}{n+1} a^T S a}{(a^T a)^{1/2}} \leq \lambda \leq \frac{a^T \xi}{(a^T a)^{1/2}} \leq \frac{a^T X + \frac{T^2}{n+1} a^T S a}{(a^T a)^{1/2}} \]

for all nonnull \( a(p \times 1) \), where \( T^2 = T^2_a(p, n+1-p) \) is the upper \( \alpha \) point of Hotelling's \( T^2 \)-distribution with \( p \) and \( (n+1-p) \) degrees of freedom.

Writing \( \chi^2 = T^2_{a/(n+1)} \), we may rewrite (1.6.1) as

(1.6.2) \[ \frac{(a^T S a)^{1/2}}{(a^T a)^{1/2}} \leq \frac{(a^T S a)^{1/2}}{(a^T a)^{1/2}} + \frac{(a^T \xi^T)}{(a^T a)^{1/2}} \]

for all nonnull \( a(p \times 1) \), which, therefore, is a set of simultaneous confidence statements with a confidence coefficient \( = (1 - \alpha) \). We next observe that the maximum values of \( \frac{a^T X}{(a^T a)^{1/2}} \), \( \frac{a^T \xi}{(a^T a)^{1/2}} \) and \( \frac{(a^T S a)^{1/2}}{(a^T a)^{1/2}} \), with respect to variation over \( a \)'s, are, respectively, \( (\xi^T \xi)^{1/2} \), \( (\xi^T \xi)^{1/2} \) and \( \lambda_{\max}^1/z(s) \). Reasoning in exactly the same way as in the previous sections of this chapter, we deduce that (1.6.2) implies

(1.6.3) \[ (\xi^T \xi)^{1/2} \leq \lambda_{\max}^1/z(s) \leq (\xi^T \xi)^{1/2} + (\xi^T \xi)^{1/2} + \lambda_{\max}^1/z(s) \]

which, therefore, is a confidence statement with a confidence coefficient \( \geq (1 - \alpha) \). It may be noted here again that the middle term in (1.6.3), i.e., \( (\xi^T \xi)^{1/2} \), is a measure of departure from the null hypothesis \( H_0 : \xi = 0 \), (which can be taken, without any loss of gen-
erality, as the general null hypothesis on $\xi$, and it is equal to zero if, and only if, $H_0$ is true. Next, arguing exactly as before and using a similar notation for the truncated $\bar{X}$, $\bar{\xi}$ and $S$ obtained by cutting out the $i$th variate ($i = 1, 2, \ldots, p$), the $i$th and $j$th variates ($i \neq j = 1, 2, \ldots, p$), and so on, we have, with a joint confidence coefficient $\geq (1 - a)$, in addition to (1.6.3), $p$ statements like

(1.6.4)

$$
\left(\bar{X}(1) - \bar{X}(1)\right)^{1/2} - \lambda_{\max}^{1/2}(S(1)) \leq (\bar{\xi}(1) - \bar{\xi}(1))^{1/2} - \lambda_{\max}^{1/2}(S(1)),
$$

($p$) statements like,

(1.6.5)

$$
\left(\bar{X}(1, j) - \bar{X}(1, j)\right)^{1/2} - \lambda_{\max}^{1/2}(S(1, j)) \leq (\bar{\xi}(1, j) - \bar{\xi}(1, j))^{1/2} - \lambda_{\max}^{1/2}(S(1, j)),
$$

and so on down to the stage of cutting out any $(p - 1)$ variates, i.e., retaining any one variate, which will thus be bounds on univariate means.

(ii) Some observations on multivariate linear hypotheses on means.

Certain confidence bounds connected with univariate and with multivariate linear hypotheses on means have already been obtained and are discussed in chapters 15 and 16 of [7]. In this section we shall first set up a physically more general hypothesis, although it is, in fact, subsumed under the one discussed in [7].
Let $X(p \times n)$, where $p < n$, consist of $n$ independently distributed column vectors, $X_i(p \times 1), \quad (i = 1, 2, \ldots, n)$, each being $N(\mu X_i), \Sigma_i$, and let $E(X')(n \times p) = A(n \times m)\xi(m \times p)$, where $m < n$ and rank of $A = r \leq m$. If $A_1(n \times r)$ is a basis of $A$, then we may, without any loss of generality, write

$$A(n \times m) = n \sum_{r}^A A_1 \quad A_2 \quad \xi$$

and rewrite the expectation condition as,

$$(1.6.6) \quad E(X')(n \times p) = n \sum_{r}^A A_1 \quad A_2 \quad \xi$$

$$(\xi_1 \quad \xi_2 \quad \xi_p)$$

$$(r \quad m-r \quad m-r)$$

Here $X$ is a matrix of observable stochastic variables, $\xi$ is a matrix of unknown parameters and $A$ is the model matrix, whose elements can be obtained from a knowledge of the design. These elements of $A$ might be numbers like, say, 0, 1 etc., and/or a set of observed non-stochastic quantities, as in the case of regression problems with concomitant variates. The population dispersion matrix, $\Sigma$, is also unknown.

Under this model a test of the set of hypotheses,

$$(1.6.7) \quad H_0: C(q \times m) \quad \xi(m \times p) = 0(q \times p),$$

where $C(q \times m) = C_{11} C_{12} C_{21} C_{22}$ such that $r(C) = s$, $s$ and $q - s$.  


has been obtained \( \sum_{i=1}^{k} y_3 y_i \) by using the union-intersection principle.

The critical region of the test is obtained as,

\[
(1.6.8) \quad c_t \geq c_a(p, s, n-r) = c_a \text{ (say)}
\]

where \( t = \min(p, s) \), \( c_t \) is the largest characteristic root of

\[
S_1 S^{-1},
\]

where

\[
sS_1(p \times p) = A_1^t (A_1^t A_1)^{-1} c_{11} c_{11}^t (A_1^t A_1)^{-1} c_{11}^t X
\]

\[
(1.6.9) \quad (n-r)S(p \times p) = XX^t - A_1^t (A_1^t A_1)^{-1} A_1^t X
\]

\( S_1 \) and \( S \) may be called, respectively, the sample dispersion matrices due to the hypothesis and due to error. The associated simultaneous confidence bounds,

\[
(1.6.10) \quad \eta^t X_1 (A_1^t A_1)^{-1} c_{11}^t b + (s^t S a_1)^{1/2} \sqrt{sc_a} \eta^t b \leq \eta^t b
\]

\[
\leq \eta^t X_1 (A_1^t A_1)^{-1} c_{11}^t b + (s^t S a_1)^{1/2} \sqrt{sc_a} \eta^t b
\]

for all nonnull \( a(p \times 1) \) and all \( b(s \times 1) \) subject to

\[
b^t \sqrt{c_{11} (A_1^t A_1)^{-1} c_{11}^t} b = 1
\]

and where \( \eta(s \times p) \) is given by \( s(C_{11} C_{12}) \xi(m \times p) = \eta(s \times p) \),

\[
\begin{bmatrix} C_{11} & C_{12} \\ C_{12}^t & C_{22} \end{bmatrix}
\]

have also been obtained, \( \sum_{i=1}^{k} y_3 y_i \). The joint confidence coefficient of (1.6.10) is \( 1 - \alpha \).

We may normalize \( b(s \times 1) \) into \( \bar{b}(s \times s) \bar{b}_1(s \times 1) \), where \( \bar{b}_1 \)

is a unit vector and \( \bar{U} \bar{U}^t = \sqrt{c_{11} (A_1^t A_1)^{-1} c_{11}^t} \eta^t b \). Then it is seen that (1.6.10) \( \iff \)
\[(1.6.11)\quad \alpha X_{A_1} (A_1^t A_1)^{-1} \mathbf{c}_{11} \mathbf{b}_1 \triangleq (\mathbf{e}^t \mathbf{a})^{1/2} \sum_{\mathbf{sc}_{a}}^{1/2} \leq \alpha \eta^t \mathbf{b}_1 \]

\[\leq \alpha^t X_{A_1} (A_1^t A_1)^{-1} \mathbf{c}_{11} \mathbf{b}_1 + (a^t s)^{1/2} \sum_{\mathbf{sc}_{a}}^{1/2} \mathbf{b}_1 \]

for all non-null \(\alpha(p \times 1)\) and unit vectors \(\mathbf{b}_1(s \times 1)\).

Now under the model defined in (1.6.6) we consider, in place of (1.6.7), the physically more general hypothesis

\[(1.6.12)\quad H_0: \quad \mathbf{s} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \cdot \mathbf{r} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = 0(p \times u) \quad M(p \times u) = 0(q \times u)\]

where \(C(q \times m)\) and \(M(p \times u)\) are given hypothesis matrices. It is assumed that \(r(M) = u \leq p\), and, as before, that \(r(C) = s \leq r\) such that, row wise \(\sum_{\mathbf{C}}^{1/2} C_{12} = \mathbf{J}\) is a basis of \(C(q \times m)\). Using the "testability" condition of Roy \(\sum_{\mathbf{J}}^{1/2} \mathbf{J}\), we see that we should have

\[(1.6.13)\quad C_{12} = C_{11} (A_1^t A_1)^{-1} A_1^t A_2 \quad (i = 1, 2) \]

We now go back to \(\mathbf{X}(p \times n)\) and observe that \(\mathbf{X}^*(u \times n) = M^t(u \times p) \mathbf{X}(p \times n)\) will be \(n\) independently distributed column vectors, \(\mathbf{X}_i^*(u \times 1), (i = 1, 2, \ldots, n)\), each being \(\sum_{\mathbf{E}}^{1/2} M_i^t \mathbf{E}\), \(M_i^t \mathbf{E} \mathbf{J}\), \(i.e.,\ \sum_{\mathbf{E}}^{1/2} \mathbf{E}\), \(\mathbf{E}^* \mathbf{J}\), (say), and that

\[(1.6.14)\quad E(\mathbf{X}_i^*)(n \times u) = \sum_{\mathbf{A}}^{1/2} A_1 A_2 \mathbf{J} \quad \mathbf{r} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \sum_{\mathbf{m} \mathbf{r}}^{1/2} \mathbf{M}_i \sum_{\mathbf{A}}^{1/2} A_1 A_2 \mathbf{J} \begin{bmatrix} \xi_1^* \\ \xi_2^* \end{bmatrix}, \quad \text{say,}\]

where
\[
\begin{align*}
& \begin{bmatrix}
\xi_1^* \\
\xi_2^*
\end{bmatrix}^r = \begin{bmatrix}
\xi_1 \\
\xi_2
\end{bmatrix}^r \cdot M(p \times u) \\
& u = \begin{bmatrix}
\xi_1^* \\
\xi_2^*
\end{bmatrix}^m-r \begin{bmatrix}
\xi_1 \\
\xi_2
\end{bmatrix}^m-r \\
\end{align*}
\]

The \( H_0 \) of (1.6.12) can now be rewritten as

\[
(1.6.15) \quad H_0: s(C_{11} \ C_{12}) \begin{bmatrix}
\xi_1^* \\
\xi_2^*
\end{bmatrix}^r = 0(s \times u) \begin{bmatrix}
\xi_1 \\
\xi_2
\end{bmatrix}^m-r \\
\]

and the alternative, \( H \), to \( H_0 \) can be expressed as,

\[
(1.6.16) \quad H_1: s(C_{11} \ C_{12}) \begin{bmatrix}
\xi_1^* \\
\xi_2^*
\end{bmatrix}^r = \eta^*(s \times u) \begin{bmatrix}
\xi_1 \\
\xi_2
\end{bmatrix}^m-r \\
\]

It can now be seen that (1.6.14) and (1.6.15) are of exactly the same form as (1.6.6) and (1.6.7), so that the physically more general case considered can be formally subsumed under the case already considered.

Going back to (1.6.10) let us substitute, \( X^* = M'X \) for \( X \), \( a^*(u \times 1) \) for \( a(p \times 1) \), \( u \) for \( p \) and \( \eta^*(s \times u) \) for \( \eta(s \times p) \). Then we have the statement,

\[
(1.6.17) \quad a^* x^* A_1(A_1^* A_1)^{-1} C_{11} \tilde{U}_{p_1} (a^* S^* a) \frac{1}{2} \sqrt{sc_a(u, s, n-r)} \frac{1}{2} \\
\quad \leq a^* X^* A_1(A_1^* A_1)^{-1} C_{11} \tilde{U}_{p_1}^* (a^* S^* a) \frac{1}{2} \sqrt{sc_a(u, s, n-r)} \frac{1}{2} \\
\quad \text{for all nonnull } a^*(u \times 1) \text{ and unit vectors } b_1(s \times 1), \text{ where}
\]
\[(n-r)S^*(u \times u) = M'(u \times p) X(p \times n)\eta^*-1A_1'A_1^{-1}A_1'A_1^{-1}X'M.\]

(1.6.17) is a set of simultaneous confidence statements with a joint confidence coefficient \((1 - a)\), on bilinear compounds of \(\eta^*\), where \(\eta^*\), defined in (1.6.16), may be regarded as measuring the departure from the null hypothesis \(H_0\).

(iii) Further implications of (1.6.10) and (1.6.17).

From the above discussion it follows that we can always switch over from \(H_0: \xi = 0\) to \(H_0: \xi M = 0\), and back and forth, as far as mathematical properties are concerned. Obvious substitutions in the results for one case yield the corresponding results for the other case. Hence, in the following considerations of further implications of the confidence bounds obtained earlier, we shall consider only one of the set-ups, say, (1.6.17).

In (1.6.17) set \(\lambda^2 = s_0(u, s, n-r)\) and argue exactly as in the previous sections of this chapter to see that (1.6.17) \(\implies\)

\[c_{\text{max}}^{1/2}X^*A_1'(A_1'A_1)^{-1}C_{11}^{1/2}\tilde{U}U'C_{11}^{-1}A_1'A_1^{-1}A_1'A_1^{-1}X'^*J - \lambda c_{\text{max}}^{1/2}(S^*)\]

\[
\leq c_{\text{max}}^{1/2}\eta^*'(\tilde{u}U')\eta^*J - c_{\text{max}}^{1/2}X^*A_1'(A_1'A_1)^{-1}C_{11}^{1/2}\tilde{U}U'C_{11}^{-1}A_1'A_1^{-1}A_1'A_1^{-1}X'^*J + \lambda c_{\text{max}}^{1/2}(S^*) ,
\]

or, substituting \(\tilde{U}\tilde{U}' = C_{11}^{-1}(A_1'A_1)^{-1}C_{11}'^{-1}\),

\[c_{\text{max}}^{1/2}(sS^*) - \lambda c_{\text{max}}^{1/2}(S^*) \leq c_{\text{max}}^{1/2}\eta^*'(C_{11}(A_1'A_1)^{-1}C_{11})^{-1}\eta^*J - c_{\text{max}}^{1/2}(sS^*) + \lambda c_{\text{max}}^{1/2}(S^*) ,
\]
where the dispersion matrix due to the hypothesis, i.e., $S_1^*$, is given by

$$s_1^*(u \times u) = x^* A_1 (A_1^* A_1)^{-1} C_1' \sigma_{11} (A_1^* A_1)^{-1} C_1' - C_1 (A_1^* A_1)^{-1} A_1 x'^*$$

$$= M X A_1 (A_1^* A_1)^{-1} C_1' \sigma_{11} (A_1^* A_1)^{-1} C_1' - C_1 (A_1^* A_1)^{-1} A_1 x'M$$

and $(n - r) S^*$ is given under (1.6.17).

The statement (1.6.19) will, of course, have confidence coefficient $\geq (1 - \alpha)$, and it may be noted that the middle term of (1.6.19) is zero if, and only if, $H_0$ is true.

For $p = 1$ the problem reduces to the univariate general linear hypothesis case, which may thus be studied as a special case. When $p = 1$, $M(p \times u)$ will drop out, (except for a trivial scalar factor, since $u \leq p$), and we also have, $c_{\text{max}}(ss^*_1)$, $c_{\text{max}}(S^*)$ reduced, respectively, to the sum of squares due to the hypothesis and the error mean sum of squares. Also, $c_{\text{max}}(\eta^* (C_1 (A_1^* A_1)^{-1} C_1')^{-1} \eta^*)$ will reduce to the scalar $\eta^*(1 \times s)'C_1 (A_1^* A_1)^{-1} C_1' - C_1 (A_1^* A_1)^{-1} A_1 x* (s \times 1)$.

Starting from (1.6.17) and reasoning in exactly the same way as in the previous sections, we see that (1.6.17) implies, in addition to (1.6.19), $p$ statements like

$$c_{\text{max}}^{1/2}(ss_1^*(i)^*) - \alpha c_{\text{max}}^{1/2}(S(i)^*) \leq c_{\text{max}}^{1/2}(\eta(i)^* (C_1 (A_1^* A_1)^{-1} C_1')^{-1} \eta(i)^*)$$

$$\leq c_{\text{max}}^{1/2}(ss_1^*(i)^*) + \alpha c_{\text{max}}^{1/2}(S(i)^*)$$

involving truncated $S(i)^*$, $s_1^*(i)^*$ and $\eta(i)^*$ obtained by cutting out
any $i$th variate ($i = 1, 2, \ldots, p$), $(\frac{p}{2})$ statements involving $s(i,j)^*$, $s(i,j)^*$ and $\eta(i,j)^*$ obtained by cutting out any $i$th and $j$th variates ($i \neq j = 1, 2, \ldots, p$), and so on. All these confidence statements will have a joint confidence coefficient $\geq (1 - \alpha)$.

It may be mentioned that the problem discussed in (i) under section 1.6 is, in fact, a very special case of the one discussed in (iii) under section 1.6.
CHAPTER II

ON EQUALITY OF VARIANCES AND OF DISPERSION MATRICES

2.1 Introduction:
In this chapter, two tests are proposed and the associated confidence bounds are obtained. The first is for the equality of 
(k + 1) variances of (k + 1) univariate normal populations, where- 
in we choose one of the variances as a standard, and compare the other k variances with it. The proposed test may be called the simultaneous variance ratios test. The well-known Hartley's \( F_{\text{max}} \) test \( \chi^2 \) for the case of equal sample sizes is equivalent to the present test, when all samples are of the same size. The second test is a multivariate extension of the first. The present chapter also considers the distribution problems that arise in connection with both the tests. The non-availability of tables makes the immediate practical application of the tests not possible.

2.2 The simultaneous variance ratios test.
For (k + 1) normal populations we want to test \( H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2 \). Suppose we have (k + 1) independent random samples of sizes \( (n_1 + 1), \ldots, (n_k + 1) \) and \( (n_0 + 1) \) respectively from the (k + 1) populations, and let \( s_1^2, \ldots, s_k^2 \) and \( s^2 \) be the estimates of \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( \sigma^2 \) based on \( n_1, \ldots, n_k \) and \( n_0 \) degrees of freedom respectively. Let us choose \( \sigma^2 \) as standard and compare \( \sigma_1^2, \ldots, \sigma_k^2 \).
with \( \sigma^2 \), so that \( H_0 \) is equivalent to \( H_0': \sigma_1^2 / \sigma^2 = \ldots = \sigma_k^2 / \sigma^2 = 1 \).

The alternative hypothesis is \( H_1': \) Not \( H_0' \), i.e., at least one

\( \sigma_i^2 / \sigma^2 \neq 1 \). For each hypothesis like \( H_{0i'}: \sigma_i^2 / \sigma^2 = 1 \) against

\( H_{1i'}: \sigma_i^2 / \sigma^2 \neq 1 \), it is well known that we have the usual variance ratio test with the acceptance region,

\[
(2.2.1) \quad F_{11} \leq F_{i1} (n_1, n_0) \leq F_{12},
\]

where \( F_{i1} (n_1, n_0) = s_i^2 / s^2 \) has the central F-distribution with \( n_1 \) and \( n_0 \) degrees of freedom under \( H_{0i'} \). It is also easily seen that

\( H_0' = \bigcap_{i=1, \ldots, k} H_{0i'} \) and \( H_1' = \bigcup_{i=1, \ldots, k} H_{1i'} \). Therefore, by the heuristic union-intersection principle, we shall take for our test of \( H_0' \), i.e., of \( H_0' \), the acceptance region,

\[
(2.2.2) \quad F_{11} \leq F_{i1} (n_1, n_0) \leq F_{12}, \quad F_{21} \leq F_{21} (n_2, n_0) \leq F_{22}, \ldots,
\]

\[
F_{k1} \leq F_{k1} (n_k, n_0) \leq F_{k2},
\]

which is the intersection (over \( i \)) of the regions \((2.2.1)\). For the critical region, therefore, we take the union (over \( i \)) of the complements of the regions \((2.2.1)\).
2.3 **Choice of** \( F_{i1} \) **and** \( F_{i2} \) **for** \( i = 1, 2, \ldots, k \).

The optimum choice of \( F_{i1}, F_{i2} \), for \( i = 1, 2, \ldots, k \), is not known, and, in the absence of this knowledge, the following choice is suggested as one possible way:

Following the usual procedure for obtaining a Type I union-intersection region, choose \( F_{i1} \) and \( F_{i2} \) such that all the individual regions \((2.2.1)\) will have the same size \((1 - \beta)\), where \( \beta \) is such that the size of the intersection \((2.2.2)\) is \((1 - \alpha)\), for a pre-assigned \( \alpha \). In general, of course, \((1 - \alpha) \neq (1 - \beta)^k\), but assuming non-triviality, given \( \alpha \) we can determine \( \beta \) and vice versa. This condition, however, still does not determine the region \((2.2.2)\) completely. In order to do so, we impose the further condition that, for each \( i \), the test with acceptance region \((2.2.1)\) be locally unbiased (in the sense of Neyman). This latter condition as will be shown in section 2.5, ensures a desirable property of the simultaneous variance ratios test with acceptance region \((2.2.2)\).

2.4 **Evaluation of a probability integral in connection with the test of section 2.2.**

Notice that the numerators of \( F_1(n_1, n_0), \ldots, F_k(n_k, n_0) \) are independently distributed as \( \sigma^2 X^2_1(n_1), \ldots, \sigma^2 X^2_k(n_k) \) respectively, and they are distributed independently of the common denominator, which is distributed as \( \sigma^2 X^2(n_0) \), where \( X^2(m) \) denotes
the central $\chi^2$ variate with $m$ degrees of freedom. Under $H_0$, therefore, $F_i$'s are quasi-independent variance ratios and, writing

$$F_i^* = \frac{n_i}{n_0} F_i,$$

the joint distribution of $F_i^*$'s is known $\mathcal{F}_{\chi^2}$ to be

$$(2.4.1) \quad p(F_1^*, F_2^*, \ldots, F_k^*) = \frac{\prod_{i=0}^{n_0} \frac{n_1}{n_0}}{\Gamma \left( \frac{n_1}{2} \right) \Gamma \left( \frac{n_0}{2} \right)} \frac{\Gamma \left( \frac{n_1}{2} \right) \Gamma \left( \frac{n_0}{2} \right)}{\Gamma \left( \frac{n_1 + n_0}{2} \right)} \prod_{i=1}^{k} F_i^* \frac{n_i^2 - 2}{2},$$

which can be easily obtained from the joint distribution of

$s_1^2, s_2^2, \ldots, s_k^2$ and $s^2$, under $H_0$, given by

$$(2.4.2) \quad p(s_1^2, s_2^2, \ldots, s_k^2, s^2) = \text{const.} \ \frac{\prod_{i=1}^{k} (s_i^2)^{n_i - 2}}{(s^2)^{n_0 - 2}} \exp \left[ -\frac{1}{2\sigma^2} \left( \sum_{i=1}^{k} n_i s_i^2 + n_0 s^2 \right) \right].$$

Now,

$$(2.4.3) \mathbb{P}(F_{11} \leq F_1(n_1, n_0) \leq F_{12}, \ldots, F_{k1} \leq F_k(n_k, n_0) \leq F_{k2} / H_0)$$

$$= \mathbb{P}(F_{11}^* \leq F_1^*(n_1, n_0) \leq F_{12}^*, \ldots, F_{k1}^* \leq F_k^*(n_k, n_0) \leq F_{k2}^* / H_0),$$

where $F_{11}^* = \frac{n_1}{n_0} F_{11}, F_{12}^* = \frac{n_1}{n_0} F_{12} (i = 1, 2, \ldots, k)$.
\[
\frac{\Gamma \left( \frac{\sum_{j=0}^{n_j} 2}{2} \right)}{k \sum_{j=0}^{n_j} \Gamma \left( \frac{n_j}{2} \right)} F_{11}^{F_{12}^*} \cdots F_{kl}^{F_{k2}^*} \frac{k^{n_i-2}}{\prod_{i=1}^{k} F_i^{F_i^{F_i^{F_i^*}/2}}} dF_i^* \\
\int 1 + \sum_{i=1}^{k} F_i F_i^{-1}
\]

where \( c(n_0, \ldots, n_k) = \frac{\Gamma \left( \frac{k \sum_{j=0}^{n_j} n_j/2}{2} \right)}{k \prod_{j=0}^{n_j} \Gamma \left( \frac{n_j}{2} \right)} \)

\[
= I(F_{11}^{F_{12}^*}; \ldots; F_{kl}^{F_{k2}^*}; n_0, n_1, \ldots, n_k)
\]

We next obtain a recurrence relation for \( I(\ldots) \).

Integrating by parts we see that,

\[
(2.4.11) \quad I(F_{11}^{F_{12}^*}; \ldots; F_{k1}^{F_{k2}^*}; n_0, n_1, \ldots, n_k)
\]

\[
= -c(n_0, n_1, \ldots, n_k) \int \cdots \int \left\{ \sum_{j=0}^{\sum n_j/2} \prod_{i=1}^{k-1} F_i^{F_i^{F_i^*}/2} dF_i^* F_k^{F_k^*/2} \right\} F_{kl}^{F_{k2}^*}
\]

\[
\int 1 + \sum_{i=1}^{k} F_i F_i^{-1}
\]
\[ + \frac{(n_k - 2)}{2} \frac{c(n_0, \ldots, n_k)}{\sum_{j=0}^{\frac{n_k - 2}{2}}} \int \ldots \int \frac{F_{12}^*}{F_{11}^*} \frac{F_{k2}^*}{F_{k1}^*} \frac{\prod_{i=1}^{k-l} \frac{F_i^*}{2} dF_i^*}{\prod_{i=1}^{k} dF_i^*} \]

\[ \frac{k}{\sum_{j=0}^{\frac{n_k - 2}{2}}} \frac{\prod_{i=1}^{k} \frac{F_i^*}{2} dF_i^*}{\prod_{i=1}^{k} dF_i^*} \]

i.e.,

\[ \frac{(F_{k1}^*)}{(1 + F_{k1}^*)} \frac{n_k - 2}{\sum_{j=0}^{\frac{n_k - 2}{2}}} \frac{c(n_0, \ldots, n_k)}{\sum_{j=0}^{\frac{n_k - 2}{2}}} x \]

\[ \frac{F_{12}^*}{1 + F_{k1}^*} \int \ldots \int \frac{F_{(k-1)2}^*}{l + F_{k1}^*} \frac{\prod_{i=1}^{k-l} \frac{F_i^*}{2} dF_i^*}{\prod_{i=1}^{k} dF_i^*} \]

\[ \frac{F_{11}^*}{l + F_{k1}^*} \frac{F_{(k-1)1}^*}{l + F_{k1}^*} \]

\[ \frac{(F_{k2}^*)}{(1 + F_{k2}^*)} \frac{n_k - 2}{\sum_{j=0}^{\frac{n_k - 2}{2}}} \frac{c(n_0 n_1, \ldots, n_k)}{\sum_{j=0}^{\frac{n_k - 2}{2}}} x \]
\[
\begin{align*}
\frac{F_{12}^*}{1+F_{k2}^*} & \quad \frac{F^{*(k-1)2}_{11}}{l+F_{k2}^*} \\
\int \frac{F_{11}^*}{1+F_{k2}^*} & \quad \int \frac{F^{*(k-1)1}_1}{l+F_{k2}^*} \\
\frac{n_1-2}{k} & \quad \left[ \frac{k-1 \Sigma F_i^*}{l+\Sigma F_i^*} \right]_{j=0}^{j=2} n_j-2
\end{align*}
\]

\[
\begin{align*}
+ \frac{(n_k-2)}{j=0} & \quad \frac{c(n_0',n_1',\ldots,n_k')}{c(n_0',n_1',\ldots,n_{k-1}',n_k')}
\end{align*}
\]

Notice that the multiple integral on the left side of

\((2.4.4)\) is of order \(k\), while the first and second integrals on the

right side of \((2.4.4)\) are of order \((k-1)\), (being of the same type

as that on the left side), and the last integral on the right side

is of order \(k\) with \(n_k\) reduced by 2. Proceeding along the chain

of reductions, by repeated application of the reduction formula

\((2.4.4)\), it can be seen that we shall eventually have to evaluate

integrals of the type,

\[(2.4.5)\]

\[
\int_{a_1}^{b_1} \cdots \int_{a_t}^{b_t} \frac{\Gamma_{i=1}^{t} F_i^* \cdot l/2}{l+\Sigma F_i^*} \frac{dF_i^*}{m+t} 
\]
where \( m \) is a function of \( n_i \)’s, and \( t \) can take any value between 1 and \( k \).

It can be seen that (2.4.5) is equivalent to the integral,

\[
(2.4.6) \quad \int_{0}^{\infty} e^{-\frac{m-2}{2 \nu}} x \left[ \prod_{i=1}^{t} \int_{a_i^v}^{b_i^v} u_i^{-\frac{1}{2}} e^{-u_i} du_i \right] dv.
\]

For evaluating this we may notice that the integral,

\[
(2.4.7) \quad \int_{x}^{y} u^{-\frac{1}{2}} e^{-u} du
\]

is essentially a normal probability integral, and we may then apply anyone of the several well-known quadrature methods to evaluate the integral.

Thus it is seen that the evaluation of the integral (2.4.3), for given values of \( F_{11}^*, F_{12}^* \) (\( i = 1, 2, \ldots, k \)), can be carried out, at least in theory, by using the reduction formula (2.4.4). However, it must also be noted that for practical purposes tables of (2.4.3) are essential, and the tabulation, in general, does not seem to be very easy. This is a limitation on the proposed test from the standpoint of immediate practical applications.

2.5 Properties of the power of the test proposed in 2.2

We shall first note that the power, or, equivalently, the
second kind of error, $\beta$, of the test could involve as parameters only the $k$ ratios $\delta_i = \sigma_i / \sigma$, $(i = 1, 2, \ldots, k)$.

$$\beta = \Pr \left( F_{11} \leq F_1(n_1, n_0) \leq F_{12}, \ldots, F_{k1} \leq F_k(n_k', n_0) \leq F_{k2} \mid H_1 \right)$$

$$= \Pr \left( F_{11} \leq \frac{F_1(n_1, n_0)}{\delta_1} \leq F_{12}, \ldots, F_{k1} \leq \frac{F_k(n_k', n_0)}{\delta_k} \leq F_{k2} \mid H_0 \right)$$

$$= \Pr \left( F_{11} \delta_1 \leq F_1(n_1, n_0) \leq F_{12} \delta_1 \delta_2 \ldots, F_{k1} \delta_k \leq F_k(n_k, n_0) \right) \leq F_{k2} \delta_k \mid H_0.$$

It now follows that $\beta$ could involve as parameters only $\delta_1, \delta_2, \ldots, \delta_k$.

We shall next show that, for the choice of $F_{11}, F_{12}$ mentioned in section 2.3, the power of the test has the monotonicity property, i.e., that as each $\delta_i$, $(i = 1, 2, \ldots, k)$, tends away from unity the power has the monotonicity property.

**Lemma 2.5a**

$$\int_0^\infty p(v) f(v) dv \leq 0$$

according as $f(v) \leq 0$ provided $f(v)$ maintains a constant sign in the range of integration, where

$$p(v) = \frac{n/2}{\Gamma(\frac{n}{2})} \exp\left(-\frac{rv}{2}\right) \frac{n-2}{v^{n/2-1}}$$

$$\frac{1}{\Gamma(\frac{n}{2})} \exp\left(-\frac{rv}{2}\right) \frac{n-2}{v^{n/2-1}}$$
We shall need this lemma in what follows. Let us write
\[ v_0 = s^2 / \sigma^2 \text{ and } v_i = s_i^2 / \sigma_i^2, \quad i = 1, 2, \ldots, k. \]

\[(2.5.1) \quad \frac{\partial \theta}{\partial \theta_1} = \frac{\partial}{\partial \theta_1} \int_0^\infty p(v_0) \left( \prod_{i=1}^k \int_{\frac{F_{i2}v_0}{\theta_i}} \frac{p(v_i) \, dv_i}{\theta_i} \right) \, dv_0 \]

\[= \int_0^\infty p(v_0) \, dv_0 \frac{\partial}{\partial \theta_1} \left[ \prod_{i=1}^k \int_{\frac{F_{i2}v_0}{\theta_i}} p(v_i) \, dv_i \right], \text{ this being valid,} \]

\[= \int_0^\infty p(v_0) \left[ \prod_{i=2}^k \int_{\frac{F_{i2}v_0}{\theta_i}} p(v_i) \, dv_i \right] \, dv_0 \]
\[
\begin{align*}
&= \int_0^{\infty} p(v_0) \left[ \left( \frac{n_1}{2} \right)^{n_1/2} \cdot \frac{1}{r(\delta_1)} \cdot \frac{1}{8_1} \cdot \left\{ \frac{F_{11}v_0}{8_1} \right\}^{n_1/2} - \frac{n_1F_{11}v_0}{28_1} \right] dv_0 \\
&+ \left( \frac{F_{12}v_0}{8_1} \right)^{n_1/2} e^{-\frac{n_1F_{12}v_0}{28_1}} \sum_{i=2}^{k} \frac{F_{12}v_0}{8_1} \int_{\delta_i}^{8_1} p(v_1) dv_1 \\
&= \text{const.} \int_0^{\infty} p(v_0) f(v_0) dv_0 , \text{ say,}
\end{align*}
\]

where

\[
f(v_0) = v_0 \left[ \left( \frac{F_{11}}{28_1} \right)^{n_1/2} e^{-\frac{n_1F_{11}v_0}{28_1}} - \left( \frac{F_{12}}{28_1} \right)^{n_1/2} e^{-\frac{n_1F_{12}v_0}{28_1}} \right]
\]

\[
\sum_{i=2}^{k} \frac{F_{12}v_0}{8_1} \int_{\delta_i}^{8_1} p(v_1) dv_1 ,
\]

and the constant factor is non-negative. Noticing that \( v_0, \delta_i, \)

\( F_{12} > F_{11}, (i=1,2,\ldots,k), \) are all essentially positive and that
each of the integrals in the product \( \frac{k}{i^2} \ldots \) is positive, 
(lying between 0 and 1), we may apply lemma 2.5a to obtain,

\[
\frac{\partial \theta}{\partial \theta_1} \geq 0 \quad \text{according as } F_{11}^{n_1/2} e^{-\frac{n_1 F_{11} V_0}{2 \theta_1}} \\
-F_{12}^{n_1/2} e^{-\frac{n_1 F_{12} V_0}{2 \theta_1}} \geq 0
\]

i.e., according as \( \left( \frac{F_{11}}{F_{12}} \right)^{n_1/2} e^{-\frac{n_1 V_0}{2 \theta_1} (F_{12} - F_{11})} \)

It can be shown that the condition of local unbiasedness of the region, \( F_{11} \leq F_1(n_1, n_0) \leq F_{12} \), reduces to,

\[
\left( \frac{F_{11}}{F_{12}} \right)^{n_1/2} = e^{-\frac{n_1 V_0}{2} (F_{12} - F_{11})}
\]

Substituting in (2.5.2), we see, after some simplification, that

\[
\frac{\partial \theta}{\partial \theta_1} \geq 0 \quad \text{according as } \theta_1 \leq 1 \quad \text{(irrespective of } V_0 > 0)\).

Hence, the power is a monotone increasing or decreasing function
of $\delta_1$ according as $\delta_1 \geq 1$. The same property with respect to
$\delta_2, \ldots, \delta_k$ can be proved similarly.

Also, if $\delta' (l \times k) = (\delta_1 \delta_2 \ldots \delta_k)$, then from (2.5.2) it
appears that if $F_{i1}$, $F_{i2}$ are chosen so as to make the region (2.2.1)
locally unbiased, then

$$
\frac{\partial \delta}{\partial \delta_1} \mid \delta' = (1, \ldots, l) = 0.
$$

Therefore the proposed test is locally unbiased, and, as a
consequence of its monotonicity property, it will be completely
unbiased.

2.6 The associated simultaneous confidence bounds on $\sigma_1^2 / \sigma^2$,
($i = 1, 2, \ldots, k$)

Under the alternative hypothesis it is known that $F_i(n_i, n_0)/\delta_i$, 
for $i = 1, 2, \ldots, k$, are distributed as quasi-independent variance
ratios. Hence, we can make the following simultaneous statements,

$$
F_{i1} \leq \frac{F_i(n_i, n_0)}{\delta_i} \leq F_{i2}, \ldots, 
F_{k1} \leq \frac{F_k(n_k, n_0)}{\delta_k} \leq F_{k2},
$$

where $F_{i1}$, $F_{i2}$ ($i = 1, 2, \ldots, k$) are such that

$$
P(F_{i1} \leq F_i(n_i, n_0) \leq F_{i2}, \ldots, 
F_{k1} \leq F_k(n_k, n_0) \leq F_{k2}) = (1-\alpha),
$$
so that the probability associated with (2.6.1) is \((1 - \alpha)\).

By inverting the statements (2.6.1), it is easily seen that we obtain the following simultaneous confidence interval statements,

\[(2.6.3) \quad \frac{s_1^2}{s_{F12}^2} \leq \delta_1 \leq \frac{s_1^2}{s_{F11}^2}, \quad \frac{s_2^2}{s_{F22}^2} \leq \delta_2 \leq \frac{s_2^2}{s_{F21}^2}, \quad \ldots, \]

\[
\frac{s_k^2}{s_{Fk2}^2} \leq \delta_k \leq \frac{s_k^2}{s_{Fkl}^2}
\]

with a joint confidence coefficient \((1 - \alpha)\).

These results are valid for all choices of \(F_{i1}, F_{i2}, \) \((i = 1, 2, \ldots, k)\), satisfying (2.6.2). However, if \(F_{i1}, F_{i2}, \) \((i = 1, 2, \ldots, k)\), are chosen as in section 2.3, then, from the unbiasedness and monotonicity properties of the associated test, (proved in section 2.5), we shall have the desirable property of monotonically increasing shortness (in terms of probability of covering wrong values) for the confidence bounds (2.6.3).

2. The Multivariate Test

In the multivariate situation we need a test for the hypothesis of equality of the dispersion matrices of \((k+1) p\)-variate normal populations, \(N_{p \times 1}(p x 1), \Sigma_1(p x p)\) \(i = 1, 2, \ldots, k\), and \(N_{p \times 1}(p x 1), \Sigma(p x p)\). That is, the null hypothesis is \(H_0: \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k = \Sigma.\)
Suppose that $X_i (p \times n_i + 1), \ldots, X_k (p \times n_k + 1)$ and $X (p \times n_0 + 1)$,
where $p \leq n_i$ for $i = 0, 1, \ldots, k$, are independent random samples
respectively from $N_{\Sigma_1}, \Sigma_i, \ldots, N_{\Sigma_k}, \Sigma_k$ and $N_{\Sigma}, \Sigma$. Let,

$$(2.7.1) \quad Y_i Y'_i = X_i X'_i - (n_i + 1) \bar{X}_i \bar{X}'_i = n_i S_i (p x p), \quad i = 1, 2, \ldots, k$$

$$(2.7.2) \quad p(Y_i) \, dY_i = \frac{1}{pn_i^{n_i/2}} \exp \left( -\frac{1}{2} \text{tr} \Sigma_i^{-1} Y_i Y'_i \right) \, dY_i,$$

$$(2\pi)^{-\frac{k}{2}} |\Sigma_i|^{n_i/2}$$

$$(2.7.2) \quad p(Y) \, dY = \frac{1}{pn_0^{n_0/2}} \exp \left( -\frac{1}{2} \text{tr} \Sigma^{-1} Y Y' \right) \, dY$$

$$(2\pi)^{-\frac{k}{2}} |\Sigma|^{n_0/2}$$

Just as in the univariate case discussed in 2.2, we may,

for the multivariate case, choose $\Sigma (p x p)$ as standard and compare
the \( k \) matrices \( \Sigma_1, \ldots, \Sigma_k \) with \( \Sigma \). Notice that \( S_0, S_1, \ldots, S_k \) are symmetric and a.e.p.d., having independent Wishart distributions with appropriate degrees of freedom, and \( \Sigma, \Sigma_1, \ldots, \Sigma_k \) are symmetric p.d. matrices being dispersion matrices of non-singular \( p \)-variate normal distributions.

Consider, in analogy with (2.2.2), the test for \( H_0: \Sigma_1 = \ldots = \Sigma_k = \Sigma \), whose acceptance region is

\[
(2.7.3) \quad \frac{\mu_{11}}{\frac{c_{\min}(S_0)}{c_{\max}(S_0)}} \leq \frac{\mu_{12}}{\frac{c_{\min}(S_0)}{c_{\max}(S_0)}} \leq \ldots,
\]

\[
\frac{\mu_{k1}}{\frac{c_{\min}(S_0)}{c_{\max}(S_0)}} \leq \frac{\mu_{k2}}{\frac{c_{\min}(S_0)}{c_{\max}(S_0)}}
\]

or, equivalently,

\[
(2.7.4) \quad \frac{\lambda_{11}}{\frac{c_{\min}(YY^t)}{c_{\max}(YY^t)}} \leq \frac{\lambda_{12}}{\frac{c_{\min}(YY^t)}{c_{\max}(YY^t)}} \leq \ldots,
\]

\[
\frac{\lambda_{k1}}{\frac{c_{\min}(YY^t)}{c_{\max}(YY^t)}} \leq \frac{\lambda_{k2}}{\frac{c_{\min}(YY^t)}{c_{\max}(YY^t)}}
\]

where \( \lambda_{i1} = \frac{n_1}{n_0} \mu_{11}, \lambda_{i2} = \frac{n_1}{n_0} \mu_{i2} \) for \( i = 1, 2, \ldots, k \).
2.8 **Choice of \( \lambda_{11}, \lambda_{12} \) for \( i = 1, 2, \ldots, k \).**

Here also the optimum choice of \( \lambda_{11}, \lambda_{12} \), for \( i = 1, 2, \ldots, k \), is not known. We shall, however, consider a choice in analogy with our choice, discussed in section 2.3, for the univariate case. Let us choose \( \lambda_{11}, \lambda_{12} \), for \( i = 1, 2, \ldots, k \), so that all the individual regions,

\[
\lambda_{11} \leq \frac{c_{\min}(Y_i Y_i')}{c_{\max}(YY')} \leq \frac{c_{\max}(Y_i Y_i')}{c_{\min}(YY')} \leq \lambda_{12}
\]

are of the same size \((1 - \beta)\), where \( \beta \) is such that the region of intersection, \((2.7.4)\), is of size \((1 - \alpha)\). Here again, in general, \((1 - \alpha) \neq (1 - \beta)^k\), but we shall assume non-triviality, i.e., given \( \alpha \) we can find \( \beta \) and vice-versa. As a further condition to determine the \( \lambda \)'s completely, let us impose the condition that the individual tests with acceptance regions \((2.8.1)\) are to be locally unbiased.

Investigations, similar to those of section 2.5, for desirable power properties, which might follow from the second condition on the \( \lambda \)'s, have not been made in this inquiry due to the difficulties involved.
2.9 Evaluation of a probability integral in connection with the test of section 2.7.

Note that \(0 < c(Y_i Y_i^t) < \infty\) for all \(i\) and the same is true for \(c(Y Y^t)\). Also, \(Y, Y_1, \ldots, Y_k\) are distributed independently, so that, from (2.7.2), we have the joint distribution of \(Y, Y_1, \ldots, Y_k\),

\[
(2.9.1) \quad p(Y, Y_1, \ldots, Y_k) dY dY_1 \ldots dY_k
\]

\[
= \frac{\exp \left( -\frac{1}{2} \text{tr} \left( \sum_{i=1}^{k} \begin{bmatrix} 0 & Y_i^t Y_i \\ Y_i Y_i^t & \Sigma^{-1} \end{bmatrix} \right) \right)}{p(n_0 + \sum_{i=1}^{k} n_i)} \left( \frac{1}{(2\pi)^{n_0/2}} \right)^{k/2} \left( \frac{1}{|\Sigma|} \right)^{n_1/2}
\]

The required probability can be written as,

\[
(2.9.2) \quad \mathcal{P} \left| \lambda_{11} \leq c_{\text{min}}(Y Y^t) \leq c_{\text{max}}(Y Y^t) \leq \lambda_{12} \leq \lambda_{22} \right| H_0
\]

We can, without loss of generality, take all the \((k+1)\) dispersion matrices to be equal to the identity matrix under \(H_0\). Under such a \(H_0\), the joint distribution of \(c_{\text{min}}(Y_i Y_i^t)\) and \(c_{\text{max}}(Y_i Y_i^t)\), for each \(i\), when \(Y_i(p x n_i)\) has the distribution (2.7.2), is known and
and a general method of obtaining it is given in section 5.6 of this inquiry. In the notation of that section, in fact,

\[(2.9.3) \quad P \bigwedge_{i=1}^{\lambda_{11}} c_{\max}(YY') \leq c_{\min}(Y_i Y_i') \leq c_{\max}(Y_i Y_i') \leq \lambda_{12} c_{\min}(YY') \bigwedge \]

\[= K(p, n_1) \prod_{i=1}^{\lambda_{11}} c_{\max}(YY'), \lambda_{12} c_{\min}(YY'), \frac{m^{(1)}_1}{2}, \ldots, \frac{m^{(1)}_1}{2}, \frac{m^{(1)}_1}{2}, \ldots, \frac{m^{(1)}_1}{2}, \]

where \(m^{(i)}_j = m^{(i)} + j - 1\), \(m^{(i)} = \frac{n_i - p - 1}{2}\) for \(j = 1, 2, \ldots, p\) and \(i = 1, 2, \ldots, k\). Let us denote by \(0 < x_1 \leq \ldots \leq x_p < \infty\) the characteristic roots of \(YY'\). Then from (5.6.3) the joint distribution of \(0 < x_1 \leq \ldots \leq x_p < \infty\) is,

\[(2.9.4) \quad K(p, n_0) \prod_{i=1}^{p} dx_1, \]

\[\begin{array}{ccccccc}
  m_p - x_p, & m_p - 1 - x_p, & \cdots & m_1 - x_1 \\
  x_p e, & x_p e, & \cdots & x_p e \\
  m_p - x_p - 1, & m_p - 1 - x_p - 1, & \cdots & m_1 - x_1 - 1 \\
  x_p e - x_p, & x_p e - 1, & \cdots & x_p e \\
  \cdots & \cdots & \cdots & \cdots \\
  m_p - x_1, & m_p - 1 - x_1, & \cdots & m_1 - x_1 \\
  x_1 e, & x_1 e, & \cdots & x_1 e
\end{array} \]

where \(m_j = m + j - 1\), \((j = 1, 2, \ldots, p)\), and \(m = \frac{n_0 - p - 1}{2}\).
Hence, it can be seen that the probability (2.9.2) is given by

\[(2.9.5) \quad \prod_{j=0}^{k} K(p, n_j) \int_{x_p = 0}^{\infty} \int_{x_{p-1} = 0}^{x_p} \cdots \int_{x_1 = 0}^{x_2} x_p^{m_p-1} x_{p-1}^{m_{p-1}-1} \cdots x_1^{m_1-1} dx_p \cdots dx_1 \]

\[\left\{ \prod_{i=1}^{k} \int_{\lambda_{i,p}}^{\infty} \int_{\lambda_{i,1}}^{\lambda_{i,p}} \cdots \int_{\lambda_{i,1}}^{\lambda_{i,l_i}} \int_{m_{i,1}}^{m_{i,1}} \cdots \int_{m_{i,l_i}}^{m_{i,l_i}} dx_{1} \cdots dx_p \right\}

The reduction and tabulation of (2.9.5) must be made before the practical applications of the proposed test can be made.

2.10 The associated simultaneous confidence bounds on \(c(\Sigma_1 \Sigma^{-1})\)

for \(i = 1, 2, \ldots, k\).

Starting from the joint probability (2.9.1), let us make the following transformations,

\[(2.10.1) \quad \Sigma_1 (p \times p) = \Gamma'_i (p \times p) \Sigma_1 (p \times p) \Gamma_i (p \times p), i = 1, 2, \ldots, k\]

\[\Sigma (p \times p) = \Gamma'_{x}(p \times p) \Sigma (p \times p) \Gamma (p \times p)\]

where \(\Gamma(p \times p)\) and \(\Gamma'_i (p \times p)\) are orthogonal matrices, and the \(p\) diagonal
elements of $D_\gamma(pxp)$ are the $p$ characteristic roots, $\gamma_1, \ldots, \gamma_p$,
of $\Sigma_1(pxp)$, while those of $D_\gamma(pxp)$ are the $p$ characteristic roots,$\gamma_1, \ldots, \gamma_p$, of $\Sigma(pxp)$.

Then the joint distribution of $Y_1, \ldots, Y_k$ and $Y$ may be rewritten as,

$$\frac{p(n_0^+ \cdots + n_k)}{2(2\pi)^{\frac{k}{2}}} \exp\left\{-\frac{1}{2} \text{tr}\left\{D_1 \frac{Y_1 Y_1'}{\gamma_1} \right\}ight.$$  

$$\left. + \gamma_k^{D_1} \frac{Y_k Y_k'}{\gamma_k} \right\} \int dy_1 \cdots dy_k,$$

or, remembering that $\text{tr} A(pxp)B(qxp) = \text{tr} B(qxp)A(pxp)$ as,

$$\text{(2.10.2) const.} \exp\left\{-\frac{1}{2} \text{tr}\left\{D_1 \frac{Y_1 Y_1'}{\gamma_1} \right\}ight.$$  

$$\left. + D_1 \frac{Y_k Y_k'}{\gamma_k} \right\} \int dy_1 \cdots dy_k,$$

Let us next make the transformations,
\[ D_1 / \sqrt{\Gamma_1} \ \gamma_i \ y_i = y_i^* (pxn_i) \]

(2.10.3)

\[ D_1 / \sqrt{\Gamma} \ y (pxn_0) = y^* (pxn_0) \]

so that the jacobian of the transformation is,

(2.10.4)

\[ J(Y_1, \ldots, Y_k, Y; Y_1^*, \ldots, Y_k^*, Y^*) = \text{mod} \left( \frac{\prod_{i=1}^{n_0} \gamma_i}{\prod_{j=1}^{n_1} \gamma_j^{n_0/2}} \frac{p}{\prod_{j=1}^{n_1} \gamma_j^{n_1/2}} \right) = \frac{p}{\prod_{j=1}^{n_1} \gamma_j^{n_0/2}} \frac{p}{\prod_{j=1}^{n_1} \gamma_j^{n_1/2}} \]

Thus the joint distribution of \( Y_1^*, \ldots, Y_k^* \) and \( Y^* \) is,

(2.10.5)

\[ \frac{1}{p(n_0 + \ldots + n_k)} \exp \left( - \frac{1}{2} \text{tr} \left\{ Y_1^* Y_1^* + \ldots + Y_k^* Y_k^* + Y^* Y^* \right\} \right) \]

\[ \frac{2}{(2\pi)^{k/2}} \]

\[ dY_1^* \ldots dY_k^* dY^* , \]

which is of the same form as the joint distribution of \( Y_1, \ldots, Y_k \) and \( Y \) under \( H_0 \).

Therefore, it follows that we can find constants \( \lambda_{11}, \lambda_{12}, \ldots, \lambda_{1k} \) such that the simultaneous statements
(2.10.6) \[
\lambda_{11} \leq \frac{c_{\min}(Y^*_1 Y^*_1)}{c_{\max}(Y^* Y^*)} \leq \frac{c_{\max}(Y^*_1 Y^*_1)}{c_{\min}(Y^* Y^*)} \leq \lambda_{12}, \ldots,
\]

\[
\lambda_{kl} \leq \frac{c_{\min}(Y^*_k Y^*_k)}{c_{\max}(Y^* Y^*)} \leq \frac{c_{\max}(Y^*_k Y^*_k)}{c_{\min}(Y^* Y^*)} \leq \lambda_{k2}
\]

have a joint probability = (1 - \alpha), for a preassigned \alpha.

Let \( n_i S_i^*(p_x p) = Y^*_i Y^*_i = D_i / \gamma_1 \Gamma_i Y_i Y_i \Gamma_i D_i / \gamma_1 \)

\[
= n_i D_i / \gamma_1 \Gamma_i S_i \Gamma_i D_i / \gamma_1, \quad (i = 1, 2, \ldots, p)
\]

and \( n_0 S_0^*(p_x p) = Y^* Y^* = D_0 / \gamma \Gamma Y Y \Gamma D_0 / \gamma = n_0 D_0 / \gamma S_0 \Gamma D_0 / \gamma \).

It is well known that all non-zero \( c(AB) = c(BA) \).

Hence, all \( c(\Gamma_i S_i \Gamma_i') = c(S_i \Gamma_i' \Gamma_i) = c(S_i), \) since

\( \Gamma_i(p_x p) \) is \( \perp \), \( (i = 1, 2, \ldots, k). \)

Similarly all \( c(\Gamma S_0 \Gamma') \)

= all \( c(S_0). \)

Furthermore, \( c(S_i^*) = c(D_i / \gamma_1 \Gamma_i S_i \Gamma_i') \) and \( c(S_0^*) = c(D_0 / \gamma S_0 \Gamma') \), where \( S_j^*(p_x p), \) for \( j = 0, 1, \ldots, k, \) are symmetric and a.e.p.d. matrices.
Consider, for any $i = 1, 2, \ldots, k$,

\[(2.10.7) \quad \lambda_{i1} \leq \frac{c_{\min}(Y_i^*)}{c_{\max}(Y^*)} \leq \frac{c_{\max}(Y_i^*)}{c_{\min}(Y^*)} \leq \lambda_{i2},\]

which is equivalent to,

\[
\frac{1}{\lambda_{i1}} \geq \frac{n_0}{n_1} \frac{c_{\max}(S_{i1}^{*-1})}{c_{\min}(S_{0}^{*-1})} \geq \frac{n_0}{n_1} \frac{c_{\min}(S_{i1}^{*-1})}{c_{\max}(S_{0}^{*-1})} \geq \frac{1}{\lambda_{i2}},
\]

or, to

\[(2.10.8) \quad \frac{n_i}{n_0 \lambda_{i1}} \geq \frac{c_{\max}(D_{\gamma_i} \Gamma_i S_{i1}^{-1} \Gamma_i')} {c_{\min}(D_{\gamma_i} \Gamma S_{0}^{-1} \Gamma')} \geq \frac{c_{\min}(D_{\gamma_i} \Gamma_i S_{i1}^{-1} \Gamma_i')} {c_{\max}(D_{\gamma_i} \Gamma S_{0}^{-1} \Gamma')} \geq \frac{n_i}{n_0 \lambda_{i2}}.
\]

Since, by lemma 1.2b, we have

\[
c_{\max}(D_{\gamma_i} \Gamma_i S_{i1}^{-1} \Gamma_i') c_{\max}(\Gamma_i S_{i1} \Gamma_i') \geq c_{\max}(D_{\gamma_i} \Gamma_i S_{i1}^{-1} \Gamma_i' \Gamma S_i \Gamma_i')
\]

\[
= c_{\max}(D_{\gamma_i});^p
\]

i.e.,

\[
c_{\max}(D_{\gamma_i} \Gamma_i S_{i1}^{-1} \Gamma_i') \geq c_{\max}(D_{\gamma_i}) / c_{\max}(\Gamma_i S_{i1} \Gamma_i')
\]

\[
= c_{\max}(D_{\gamma_i}) / c_{\max}(S_i),
\]
and, similarly, \( c_{\min(D_{\gamma_{i}}S_{0}^{-1}R_{i}')} \leq \frac{c_{\min(D_{\gamma_{i}})}}{c_{\min(S_{0}R_{i}')}} = \frac{c_{\min(D_{\gamma_{i}})}}{c_{\min(S_{0})}} \),

we see that the first part of (2.10.8) implies that \( \frac{n_{0}}{n_{i}^{\lambda_{i1}}} \)

\[
\geq \frac{c_{\max(D_{\gamma_{i}})}}{c_{\min(S_{0})}} \frac{c_{\min(S_{0})}}{c_{\max(S_{i})}}.
\]

Again, using lemma 1.2b,

we note that \( c_{\min(D_{\gamma_{i}})} \leq \frac{c_{\min(D_{\gamma_{i}})}}{c_{\min(S_{i})}} \) and \( c_{\max(D_{\gamma}S_{0}^{-1}R_{i}')} \)

\[
\geq \frac{c_{\max(D_{\gamma})}}{c_{\max(S_{0})}},
\]

so that the second part of (2.10.8) implies that,

\[
\frac{c_{\min(D_{\gamma_{i}})}}{c_{\max(D_{\gamma_{i}})}} \frac{c_{\max(S_{0})}}{c_{\min(S_{i})}} \geq \frac{n_{i}}{n_{0}^{\lambda_{i2}}}.
\]

Therefore, we observe that (2.10.8) implies the pair of statements,

(2.10.9)

\[
\frac{n_{i}}{n_{0}^{\lambda_{i1}}} \geq \frac{c_{\max(D_{\gamma_{i}})}}{c_{\min(D_{\gamma_{i}})}} \frac{c_{\min(S_{0})}}{c_{\max(S_{i})}}
\]

and \( \frac{c_{\min(D_{\gamma_{i}})}}{c_{\max(D_{\gamma_{i}})}} \frac{c_{\max(S_{0})}}{c_{\min(S_{i})}} \geq \frac{n_{i}}{n_{0}^{\lambda_{i2}}} \),
or, taken together,

\[(2.10.10) \quad \frac{n_1}{n_0^{\lambda_{11}}} \frac{c_{\max}(S_i)}{c_{\min}(S_0)} > \frac{c_{\max}(D_{Y_i})}{c_{\min}(D_{\gamma})} > \frac{c_{\min}(D_{Y_i})}{c_{\max}(D_{\gamma})} = \frac{n_1}{n_0^{\lambda_{12}}} \frac{c_{\min}(S_i)}{c_{\max}(S_0)}\]

Again, using lemma 1.2b,
\[\frac{c_{\max}(D_{Y_i})}{c_{\min}(D_{\gamma})} = c_{\max}(D_{Y_i})c_{\max}(D_{\lambda/\gamma})\]

\[= c_{\max}(Z_i)c_{\max}(Z^{-1}) \geq c_{\max}(Z_i Z^{-1}), \text{ and,} \quad \frac{c_{\min}(D_{Y_i})}{c_{\max}(D_{\gamma})} = c_{\min}(D_{Y_i})c_{\min}(D_{\lambda/\gamma})\]

\[= c_{\min}(Z_i)c_{\min}(Z^{-1}) \leq c_{\min}(Z_i Z^{-1}). \text{ Therefore, we observe that (2.10.10)}\]

\[\Rightarrow\]

\[(2.10.11) \quad \frac{n_1}{n_0^{\lambda_{11}}} \frac{c_{\max}(S_i)}{c_{\min}(S_0)} > c_{\max}(Z_i Z^{-1}) \geq c_{\min}(Z_i Z^{-1}) \geq \frac{n_1}{n_0^{\lambda_{12}}} \frac{c_{\min}(S_i)}{c_{\max}(S_0)}\]

Combining all statements like (2.10.11) for \(i = 1, 2, \ldots, k\), we see that the statements (2.10.6) imply the simultaneous confidence statements,

\[(2.10.12) \quad \frac{n_1}{n_0^{\lambda_{11}}} \frac{c_{\max}(S_1)}{c_{\min}(S_0)} \geq \text{all } c(Z_1 Z^{-1}) \geq \frac{n_1}{n_0^{\lambda_{12}}} \frac{c_{\min}(S_1)}{c_{\max}(S_0)}\]

\[\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\]

\[\frac{n_k}{n_0^{\lambda_{11}}} \frac{c_{\max}(S_k)}{c_{\min}(S_0)} \geq \text{all } c(Z_k Z^{-1}) \geq \frac{n_k}{n_0^{\lambda_{12}}} \frac{c_{\min}(S_k)}{c_{\max}(S_0)}\]

with a joint confidence coefficient \(\geq (1 - \alpha)\).
CHAPTER III

UNIVARIATE VARIANCE COMPONENTS

3.1 Introduction.

Several authors (\cite{1, 43}) have given the most general treatment of the usual problems of analysis of variance, using the general linear model and the theory of least squares, under Model I, in the terminology of Eisenhart \cite{13}. In this model it is well known that, for estimation of the fixed parameters, we assume (i) that the expectation of each of a set of uncorrelated observations is a linear function of some unknown fixed parameters (usually less in number than the observations); (ii) that the observations have the same unknown variance; and, for testing of hypotheses on the fixed parameters, we assume, in addition, (iii) that the observations are distributed normally. The general treatment includes, as special cases, all the usual designs of experiment, and studies have been made to investigate the effects of departures from the assumptions made in the model \cite{7}. For the components of variance case, or Model II and Mixed Model of analysis of variance, in the terminology of Eisenhart and Crump \cite{12}, however, no such general treatment, in an explicit form, exists, to the knowledge of the author. Except for the case of a Mixed Model with one set of random components considered explicitly by Bose \cite{5}, inquiries, thus far, have been concerned with analogues of particular designs under Model I, \cite{11, 16, 17, 46}. In this chapter an attempt has been made to give a general treatment of the
Model II situation, and, following the usual procedure, a tie-up between the Model I and Model II analyses is presented. The treatment is general enough to include the usually considered cases as simple special cases, but is not the most general that may be possible, nor is it, probably, the most realistic. This latter limitation will be brought out later. From the point of view of simultaneous confidence statements, again, a lot more has been done under Model I than under Model II. Results in this direction are also given in this chapter. Finally, a "mixed" model is also presented, and the problems associated with this case are studied. While neither of the models I and II will be a particular case of the other, the "mixed" model will include both the models as special cases.

3.2 The Model.

(i) \( x(n \times 1) = A(n \times m) \xi(m \times 1) + \eta(n \times 1), \quad m < n \)

\[ (3.2.1) \]

\[ = n \sum_{i=1}^{k} A_i \begin{bmatrix} \xi_1 \cr \xi_2 \cr \vdots \cr \xi_k \end{bmatrix} m_1 m_2 \ldots m_k + \eta(n \times 1), \]

where \( \sum_{i=1}^{k} m_i = m \) and \( r(A) = r \leq m \) \(<n\). The elements of \( A(n \times m) \) are known from a knowledge of the design.

(ii) \( \xi_1(m_1 \times 1), \ldots, \xi_k(m_k \times 1) \) are random samples of sizes \( m_1, m_2, \ldots, m_k \) respectively from \( N(u_i, \sigma_i^2) \), \( i = 1, 2, \ldots, k \), and these vectors are distributed independently.
of one another and independently of $\gamma(n \times 1)$.

(iii) $\gamma(n \times 1)$ is a random sample of size $n$, whose elements are physically of the nature of errors, from $N(0, \sigma^2)$.

3.3 Certain observations and deductions from the model.

Following the terminology due to Eisenhart \[137\], we shall refer to the above model as a complete Model II of anova. In the particular case, where the model matrix $A(n \times m)$, (whose elements are constants given by the design, or obtained by suitable transformations from a knowledge of the design), is such that the conditional expectation of each observation (element of $x(n \times 1)$), for a given $\xi$, involves only one element of each of $\xi_1, \xi_2, \ldots, \xi_k$, then we shall call the model as the Model II of anova for the general $k$-way classification. We shall consider this latter case in detail, and observe that the usual applications of variance components analysis will be simple special cases.

It may be noted that the assumptions of normal distributions, in (ii) and (iii) of the model, are made so as to be able to tie up the usual anova techniques under Model I with those under Model II. The normality assumption simplifies the distribution problems, and the analysis, under Model II, will be of the same degree of difficulty as the usual anova, under Model I. It must be admitted, however, that a more realistic assumption to make, regarding the distributions of $\xi_1(1 = 1, 2, \ldots, k)$, would, probably, be that they are samples from certain finite populations, This assump-
tion will, of course, lead to more complex distribution and analysis problems.

Under the model it is seen that \( \mathbf{x}^{'(n \times 1)} \) is a vector of observations from an \( n \)-variate normal, \( \mathcal{N}(\mathbf{\mu}, \mathbf{\Sigma}) \), where

\[
(3.3.1) \quad \mathbf{E}(\mathbf{x})^{(n \times 1)} = A(n \times m) \begin{bmatrix}
\mathbf{E}(\mathbf{\xi}_1) \\
\mathbf{E}(\mathbf{\xi}_2) \\
\vdots \\
\mathbf{E}(\mathbf{\xi}_k)
\end{bmatrix}
\begin{bmatrix}
m_1 \\
m_2 \\
\vdots \\
m_k
\end{bmatrix}
= A(n \times m) \begin{bmatrix}
\mu_1^{\perp} \\
\mu_2^{\perp} \\
\vdots \\
\mu_k^{\perp}
\end{bmatrix}
\begin{bmatrix}
m_1 \\
m_2 \\
\vdots \\
m_k
\end{bmatrix}
\]

where \( \mathbf{1}^{(m_1 \times 1)} \) stands for a column vector of \( m_1 \) unities, and,

\[
(3.3.2) \quad \mathbf{E}(\mathbf{x})^{(n \times n)} = \mathbf{E}(\mathbf{x}^{(n \times n)}) = E(\mathbf{x}^{\prime} \mathbf{x})^{-1} E(\mathbf{x}) E(\mathbf{x}^{\prime})
\]

\[
= A(n \times m) \begin{bmatrix}
\sigma_1^2 \mathbf{I}(m_1) & 0 & \cdots & 0 \\
0 & \sigma_2^2 \mathbf{I}(m_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & \sigma_k^2 \mathbf{I}(m_k)
\end{bmatrix}
A^{\prime}(m \times n)
\]

\[
= \sigma^2 \mathbf{I}(n)
\]

\[
= \sigma_1^2 \mathbf{A}_1 A_1^{\prime} + \cdots + \sigma_k^2 \mathbf{A}_k A_k^{\prime} + \sigma^2 \mathbf{I}(n)
\]

Our objectives will be, (i) to estimate any "estimable" linear function of \( \mu_1, \mu_2, \ldots, \mu_k \), and to test "testable" linear hypotheses on \( \mu_1, \ldots, \mu_k \); (ii) to obtain estimates of, and test hypotheses on, "estimable" linear functions of the variance components, \( \sigma_1^2, \ldots, \sigma_k^2, \sigma^2 \), and, in particular, the individual \( \sigma^2 \)'s themselves; (iii)
to obtain simultaneous confidence bounds on the parameters, or certain parametric functions, e.g., ratios like \( \frac{c_1^2}{c_2^2} \), if possible, without running into intractable distribution problems.

3.4 Certain lemmas in linear estimation.

Following the usual terminology, we shall call a parametric function estimable if there exists a function of the observations, of the same degree as the parametric function, which is an unbiased estimate of the parametric function. Also, we shall call that estimate, if it exists, with the least variance among the class of unbiased estimates, as the best unbiased estimate.

**Lemma 3.4:** If \( x(n \times 1) \) is a random sample from \( \mathcal{N}(A(n \times m)x(m \times 1), \sigma^2 I(n)) \), where \( A(n \times m) \) is a known matrix of rank \( r(A) = r \leq m < n \), and the elements of \( x(m \times 1) \) and \( \sigma^2 \) are unknown parameters, then the NSC for estimability of \( \sigma^1(1 \times m)x(m \times 1) = \sigma^1_I(1 \times r)x_I(1 \times 1) + \sigma^1_D(1 \times m-r)x_D(m-r \times 1) \), where \( x(m \times 1) = \begin{bmatrix} x_I \\ x_D \end{bmatrix} \), is that \( \sigma^1_I \) be related to \( \sigma^1_I \) by the same post-factor \( \begin{bmatrix} x_I \\ x_D \end{bmatrix} \) by which \( A_D \) is related to \( A_I \), where \( A(n \times m) = n^{-1} A_I A_D \) such that \( A_I(n \times r) \) is a basis of \( A(n \times m) \), and the columns of \( A_D(n \times m-r) \) are functions of the columns of \( A_I(n \times r) \).

For a proof see \( \text{[Ref 3]} \).
Lemma 3.4b: If \( \mathbf{x}(n \times 1) \) is a random sample from \( N_{\Sigma}(n \times m) \)
\( \mathbf{z}(m \times 1), \sigma^2 \Sigma(n \times n) \), where \( A, \Sigma \) and \( \sigma^2 \) are as in lemma 3.4a
and \( \Sigma(n \times n) \) is a known symmetric p.d. matrix, then the NSC for
\( \mathbf{z}(1 \times m) \mathbf{x}(m \times 1) = \mathbf{a}_I^t(1 \times r) \mathbf{z}(r \times 1) + \mathbf{a}_D^t(1 \times m-r) \mathbf{z}(m-r \times 1) \)
to be estimable is that \( \mathbf{a}_D^t \) should be related to \( \mathbf{a}_I^t \) by the same post-factor by which \( A_D(n \times m-r) \) is related to \( A_I(n \times r) \).

Proof: Since \( \Sigma(n \times n) \) is symmetric p.d., there exists \( \Sigma^{-1} \) a transformation,
\[ \Sigma(n \times n) = \mathbf{T}(n \times n) \mathbf{T}^t(n \times n) \]
Let
\[ \mathbf{y}(n \times 1) = \mathbf{T}^{-1} \mathbf{x}(n \times 1) \]
Then \( \mathbf{y}(n \times 1) \) is distributed as \( N_{\Sigma} \mathbf{T}^{-1} \Sigma \mathbf{I}(n) \).

Considering \( \mathbf{T}^{-1} \Sigma \) in place of \( A \) in the set-up in lemma 3.4a, and proceeding exactly along the lines of its proof in \( \Sigma \),
we get the NSC for estimability of \( \mathbf{a}_I^t \) to be

\[ (3.4.1) \quad \mathbf{a}_D^t(1 \times m-r) = \mathbf{a}_I^t(1 \times r) (A_I^t \Sigma^{-1} A_I)^{-1} A_I^t \Sigma^{-1} A_D(n \times m-r) \]

But we have,
\[ m-r \begin{bmatrix} A_I^t \Sigma^{-1} A_I \\ A_D^t \Sigma^{-1} A_D \end{bmatrix} = m-r \begin{bmatrix} \mathbf{T}^t_1 \\ \mathbf{T}^t_2 \end{bmatrix} L(r \times n), \]

where \( LL' = I(r), \Sigma \), so that \( \mathbf{T}^{-1} \Sigma A_D(n \times m-r) = L^t \mathbf{T}^t_2 = \mathbf{T}^{-1} A_I(\mathbf{T}^t_1)^{-1} \mathbf{T}^t_2 \), i.e., \( A_D = A_I(\mathbf{T}^t_1)^{-1} \mathbf{T}^t_2 \). Substituting in (3.4.1) we have,
\[ \mathbf{a}_D^t(1 \times m-r) = \mathbf{a}_I^t(1 \times r) = \mathbf{a}_I^t(1 \times r) (\mathbf{T}^t_1)^{-1} \mathbf{T}^t_2 \]
and, since \( A_D(n \times m) = A_I(n_1')^{-1}T_2' \), the lemma is established.

The best unbiased estimate for the present set-up, as obtained by Roy [437], is

\[
(3.4.2) \quad \hat{c}'(A'^{-1}A'X)'X^{-1}A'I^{-1}X
\]

Thus we observe that, while the best unbiased estimate, (3.4.2), involves the matrix \( \Sigma(n \times n) \), the NSC for estimability is independent of \( \Sigma(n \times n) \) and, of course, \( \sigma^2 \). The lemma holds, therefore, even when \( \Sigma \) is unknown.

**Corollary:** In the special case when the model matrix \( A(n \times m) \) is such that each row of the sub-matrices \( A_1(n \times m_1), \ldots, A_k(n \times m_k) \) has only one non-zero element, which is equal to unity, i.e., for the general k-way classification, the NSC for the estimability, (i) under Model I, of \( \hat{c}' = \sqrt{a_1'}c_1' \xi_1 + \cdots + c_k' \xi_k \), is

\[
\begin{bmatrix}
\xi_1 \\
\vdots \\
\xi_k
\end{bmatrix}_{m_1 \ m_2 \ m_3} = \begin{bmatrix}
c_1' \\
\vdots \\
c_k'
\end{bmatrix}_{m_1 \ m_2 \ m_3}
\]

that, (sum of the elements of \( c_1' \)) = (sum of the elements of \( c_2' \)) = \ldots = (sum of the elements of \( c_k' \)); and, (ii) under Model II, of \( \hat{c}'E(\xi) \) is that, (coefficient of \( \mu_1 \)) = (coefficient of \( \mu_2 \)) = \ldots = (coefficient of \( \mu_k \)).

**Proof:** (i) and (ii) of the corollary follow respectively from lemmas 3.4a and 3.4b. For the general k-way classification model matrix \( A(n \times m) \), it can be seen that, the column vector obtained by adding over the columns of the submatrix \( A_1(n \times m_1) = \) the column
vector obtained by adding over the columns of the submatrix
$A_j(n \times m_j) = a column vector of n unities, \mathbf{1}(n \times 1)$, for $i \neq j = 1, 2, \ldots, k$. The rank of the matrix $A(n \times m)$ is then, $r = m-k+1$, and we can, without any loss of generality, take as a basis of $A$, the matrix $A_1(n \times \bar{m-k+1}) = m_1 A_1 \ A_2 \ldots A_k \mathbf{1}$, where $A_{\bar{j}}(n \times \bar{m_j-1})$

is the matrix obtained by taking the first $(m_j-1)$ columns of the submatrix $A_j(n \times m_j)$, for $j = 2, \ldots, k$. It can now be seen, from the properties of the $A$ matrix stated above, that the $(k-1)$ columns of $A_D(n \times \bar{k-1})$ can be obtained from those of $A_1(n \times \bar{m-k+1})$ by post-multiplying $A_1$ by the matrix factor,

$$R(\bar{m-k+1} \times \bar{k-1}) = m_1 \begin{bmatrix} J(m_1 \times k-1) \\ \hline -1 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 0 \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline 0 & 0 & \cdots & -1 \\ 0 & 0 & \cdots & -1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & -1 \end{bmatrix}$$

(3.4.3)

Let

$$c' \xi = (c_1' c_2' \cdots c_k') \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \\ 1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix} = \sum_{i=1}^{k} c_i' \xi_i$$
and
\[ c_i'(1 \times m_i) = (c_{i1} c_{i2} \cdots c_{im_i}), \xi_i'(m_i \times 1) = \begin{bmatrix} \xi_{i1} \\ \xi_{i2} \\ \vdots \\ \xi_{im_i} \end{bmatrix}, \]
for \( i = 1, \ldots, k. \)

Then,
\[ c_i'(1 \times m) \xi(m \times 1) = c_i'(1 \times m-k+1) \xi_i(1 \times m-k+1) + c_i'(1 \times k-1) \xi_d(k-1 \times 1) \]
\[ = \begin{bmatrix} \xi_{11} \\ \xi_{21} \\ \vdots \\ \xi_{k1} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2-1 \\ m_k-1 \end{bmatrix} + \begin{bmatrix} c_{D1} c_{D2} \cdots c_{Dk} \end{bmatrix} \begin{bmatrix} m_1 + c_{D1} \xi_D \\ m_2-1 \\ \vdots \\ m_k-1 \end{bmatrix}, \]
where
\[ c_{ji}'(1 \times m_j-1) = (c_{j1}, \ldots, c_{j(m_j-1)}), \xi_{ji}(1 \times m_j-1) = \begin{bmatrix} \xi_{j1} \\ \vdots \\ \xi_{j(m_j-1)} \end{bmatrix}, \]
for \( j = 2, 3, \ldots, k, \) so that,
\[ c_i'(1 \times k-1) = (c_{2m_2}, c_{3m_3}, \ldots, c_{km_k}) \]
and \[ \xi_d(k-1 \times 1) = \begin{bmatrix} \xi_{2m_2} \\ \xi_{3m_3} \\ \vdots \\ \xi_{km_k} \end{bmatrix}. \]

Applying lemma 3.4a, we have, under Model I, the NSC for estimability.
\( (3,4,4) \quad \mathbf{\xi}^\prime (1 \times k-1) = \mathbf{\xi}^\prime (1 \times m-k+1) \mathbf{P}(m-k+1 \times k-1) \)

i.e.,

\[
\begin{pmatrix} c_{2m_2}, \ldots, c_{km_k} \end{pmatrix} = \begin{pmatrix} c_1^\prime \quad c_2^\prime \quad \ldots \quad c_k^\prime \end{pmatrix} \begin{bmatrix}
\mathbf{J}(m \times k-1) \\
m_1 m_2^{-1} \cdots m_k^{-1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
-1 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
-1 & 0 & \cdots & 0 \\
0 & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & -1 \end{bmatrix}
\]

\[\begin{array}{c}
m_1 \\
m_2^{-1} \\
m_k^{-1}
\end{array}\]

i.e., Sum of the elements of \( c_1^\prime = \) Sum of the elements of \( c_2^\prime = \cdots = \) Sum of the elements of \( c_k^\prime \)

This proves (1) of the corollary.

Next, under Model II, although the dispersion matrix \( \Sigma(n \times n) \) is unknown, the estimability condition is the same by lemma 3.4b, and we shall have

\[
\mathbb{E}(\mathbf{\xi})(m \times 1) = \begin{bmatrix}
\mathbb{E}(\xi_1) \\
\vdots \\
\mathbb{E}(\xi_k)
\end{bmatrix} = \begin{bmatrix}
\mu_{1-1} \\
\mu_{2-1} \\
\vdots \\
\mu_{k-1}
\end{bmatrix}
\]

in place of

\[
\bar{\mathbb{E}}(\mathbf{\xi})(m \times 1) \quad \text{in the preceding steps. Thus,}
\]

\[
\begin{bmatrix}
\mathbf{c}^\prime \bar{\mathbb{E}}(\mathbf{\xi}) = \sum_{i=1}^{k} c_{i1} \mathbb{E}(\xi_i) = \sum_{i=1}^{k} \mu_{i-1} (\text{sum of the elements of } c_1^\prime)
\end{bmatrix}
\]
Hence, we have, under Model II, the NSC for estimability of $\xi' E(\xi)$, which is a linear function of $\mu_1, \ldots, \mu_k$, is that

$$(\text{coefficient of } \mu_1) = (\text{coefficient of } \mu_2) = \ldots = (\text{coefficient of } \mu_k).$$

Hence, (ii) in the corollary.

The following implications of the corollary may be noted:

(i) Under Model I, for the k-way classification, the linearly estimable linear functions of the elements of $\xi(m \times 1)$ are, (a)

\[
\frac{1}{m_1} \sum_{i=1}^{m_1} \xi_{i1} + \frac{1}{m_2} \sum_{i=1}^{m_2} \xi_{i2} + \ldots + \frac{1}{m_k} \sum_{i=1}^{m_k} \xi_{ik}, \quad (b) \text{the } (m_1 - 1) \text{ contrasts between the elements of } \xi_i(1 \times m_1), \text{ for } i = 1, 2, \ldots, k. \quad \text{Neither the individual elements of } \xi(m \times 1), \text{ nor the separate means like}
\]

\[
\frac{1}{m_j} \sum_{i=1}^{m_j} \xi_{ij}, \quad j = 1, \ldots, k, \text{ are estimable.}
\]

(ii) Under Model II, for the k-way classification, there is only one independent linearly estimable function of $\mu_1, \mu_2, \ldots, \mu_k$, viz., $\mu = \mu_1 + \mu_2 + \ldots + \mu_k$, all others being merely multiples of $\mu$.

None of the individual $\mu_i$'s, $(i = 1, \ldots, k)$, is estimable.

Lemma 3.40: If $\xi(n \times 1)$ is a random sample from $N^A(n \times m) \xi(m \times 1)$, $\Sigma(n \times n) - \Sigma$, where $A$ and $\xi$ are as in the preceding lemmas, and, further, $\Sigma(n \times n)$ is an unknown symmetric n.d. matrix, of which there exists an independent estimate $\hat{\Sigma}$, then
(i) the best unbiased estimate of $\xi' \xi = \xi' I + \xi' D$ is

$$c_I' (A_I^{\Sigma^{-1} A_I})^{-1} A_I^{\Sigma^{-1} x}$$

and (ii) the NSC for estimability of $\xi' \xi$ is that $c_D'$ should be related to $c_I'$ by the same post-factor by which $A_D$ is related to $A_I$.

Proof: The proof of (ii) is included in the proof of Lemma 3.4b, because we saw there that the NSC for estimability of $\xi' \xi$ was independent of both $\sigma^2$ and $\Sigma(n \times n)$.

Proceeding exactly as Roy 137, for the problem in lemma 3.4b, we shall obtain the best unbiased estimate of $\xi' \xi = \xi' I + \xi' D$ as

$$c_I' (A_I^{\Sigma^{-1} A_I})^{-1} A_I^{\Sigma^{-1} x}$$

(3.4.5)

Here $\Sigma(n \times n)$ is unknown but we have an independent efficient estimate of $\Sigma$, $\hat{\Sigma}$. So we take as the best unbiased estimate of $\xi' \xi$ the linear function,

$$c_I' (A_I^{\hat{\Sigma}^{-1} A_I})^{-1} A_I^{\hat{\Sigma}^{-1} x}$$

(3.4.6)

The process here is somewhat similar to the maximum likelihood estimation of $\mu$ and $\sigma^2$ for $N(\mu, \sigma^2)$. In this problem we estimate $\mu$ by $\hat{\mu} = \bar{x}$ and $\sigma^2$ by $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$.

For the Model II two-way classification, it has been verified
that \((A_1' \Sigma^{-1} A_1)^{-1} A_1' \Sigma^{-1} = (A_1' A_1)^{-1} A_1'\). Hence, in particular, the best unbiased estimate of \(\mu = \mu_1 + \mu_2 = (\frac{1}{m_1}, \ldots, \frac{1}{m_1}, \frac{1}{m_2}, \ldots, \frac{1}{m_2}) \left[ E(\xi_1) \right]_{m_1}^{m_1}, \left[ E(\xi_2) \right]_{m_2}^{m_2}\) is given by \((\frac{1}{m_1}, \ldots, \frac{1}{m_1}, \frac{1}{m_2}, \ldots, \frac{1}{m_2})^{-1} A_1' A_1^{-1} A_1' x = \bar{x},\) the general mean of the observations.

It may be noted that the above lemmas on estimability and linear estimation will be valid even without the assumption of the normality of the distributions involved.

3.5 **Estimation of the variance components** \(\sigma^2_1, \ldots, \sigma^2_k\) and \(\sigma^2\), and of functions of them.

We shall seek \((k+1)\) quadratic forms, \(q_i = x'_i Q_i x, \) \((i=0,1,\ldots,k),\) of the observations, to be utilized in the estimation of the variance components \(\sigma^2, \sigma^2_1, \ldots, \sigma^2_k\). We shall impose the following restrictions on the \(q_i\)'s:

(i) \(q_i,\) of rank \(n_i,\) is distributed as \(\chi^2(x^2_{n_i}),\) where \(x^2_{n_i}\) denotes the central \(\chi^2\) variate with \(n_i\) degrees of freedom and \(\lambda_{i} = E(q_i/n_i),\) \((i = 0, 1, \ldots, k).\)

(ii) \(q_i\) is distributed independently of \(q_j,\) \((i \neq j = 0, 1, \ldots, k).\)

**Lemma 3.5a:** If \(x(n \times 1)\) is distributed as \(N(\mu, \Sigma(n \times n))\) and \(q_i = x'_i Q_i x, \) \((i = 0, 1, \ldots, k),\) is a quadratic form of rank \(n_i,\) then the NSC for the set \(q_0, q_1, \ldots, q_k\) to satisfy the above restric-
tions (i) and (ii) are,

\[(a) \quad Q_i^TQ_i = \lambda_i Q_i, \quad \sum_i = 0, 1, \ldots, k; \]

\[(b) \quad q_i \{ E(x) \} = E(x')Q_i E(x) = 0 \quad ; \]

\[(c) \quad Q_i^TQ_j = 0(n \times n), \quad \sum_i \neq j = 0, 1, \ldots, k, \]

where (a) insures \( x^2 \) (in general, non-central) distributions, (b) insures the centrality of the \( x^2 \) distributions, and (c) insures independent distributions for \( q_i \) and \( q_j \) \((i \neq j)\).

For a proof of these NSC see \( \sum \).

**Lemma 3.5b:** If \( E \sum x(p \times l) y'(1 \times p) \sim \varepsilon (p \times p) \) then \( E \sum x'Qy \sim \varepsilon \),

where \( Q(p \times p) \) is symmetric, is \( \text{tr } \varepsilon \).

**Proof:** \( E(x'Qy) = E \sum \sum q_{ij}x_i y_j \sim \sum q_{ij} \)

\[= \sum q_{ij}, \quad \text{since } q_{ij} = q_{ji} \]

\[= \text{tr } \varepsilon Q. \]

**Corollary:** Since \( E \sum x(n \times l) x'(1 \times n) \sim \varepsilon(n \times n) + E(x)E(x') \), we have, for the variance components model,

\[E(x'Qx) = \text{tr} \varepsilon Q + \text{tr } E(x)E(x')Q \]

\[= \text{tr} \varepsilon Q + E(x')QE(x), \quad \text{since } \text{tr}(AB) = \text{tr } (BA) \]

and \( \text{tr } (\text{scalar}) = \text{scalar} \).

For the Model II of anova, therefore,

\[\lambda_i = E(q_i/n_1) = \frac{1}{n_1} E(x'Q_i x) \]
\[
= \frac{1}{n_1} \text{tr} \Sigma Q_1 + E(x^t)Q_1E(x)
\]
\[
= \frac{1}{n_1} \text{tr} \Sigma Q_1, \text{ if } q_1 \text{ satisfies (b) of lemma } 3.5a
\]
\[
= \frac{1}{n_1} \sum_{j=1}^{k} \sigma_j^2 \text{tr} A_j A_j^t Q_1 + \sigma^2 \text{tr}(Q_1) \quad \text{using (3.3.2)}.
\]

This holds for \( i = 0, 1, \ldots, k \).

We next observe that, in lemma 3.5a, if \( \Sigma(n \times n) \) is unknown, as it is in Model II of anova, then we may require that the conditions (a) and (γ) of the lemma should be satisfied by \( q_0, q_1, \ldots, q_k \) for all symmetric p.d. matrices \( \Xi(n \times n) \). Under Model II, where \( \Sigma \) is given by (3.3.2), this means that for all \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( \sigma^2 \) we should have (a) and (γ) satisfied by \( q_0, \ldots, q_k \), in addition to (b).

Consider (a) \( Q_i \Sigma Q_1 = \lambda_i Q_1 \) for \( i = 0, 1, \ldots, k \).

This is equivalent to,
\[
Q_1 \sum \sigma_i^2 A_i A_i^t + \sum \sigma_k^2 A_k A_k^t + \sigma^2 I(n) \rightleftharpoons Q_1
\]
\[
= \frac{1}{n_1} \sum \sigma_i^2 \text{tr} A_i A_i^t Q_1 + \sum \sigma_k^2 \text{tr} A_k A_k Q_1 + \sigma^2 \text{tr} Q_1 \rightleftharpoons Q_1 ,
\]
using (3.3.2) and (3.5.1).

If this is to hold for all \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( \sigma^2 \), then we have, by equating coefficients of the \( \sigma^2 \)'s,
\[
(3.5.2) \quad Q_i A_j A_j^t Q_1 \rightleftharpoons \frac{1}{n_1} \text{tr} A_j A_j^t Q_1 = Q_1, j = 1, 2, \ldots, k; Q_i^2 \rightleftharpoons \frac{1}{n_1} \text{tr} Q_1 \rightleftharpoons Q_1 .
\]

Next, consider (γ) \( Q_i \Sigma Q_j = 0(n \times n) \) for \( i \neq j = 0, 1, \ldots, k \). This
is equivalent to,

\[ q_j \sum_{i=1}^{j} \sigma_i^2 A_i^j + \ldots + \sigma_k^2 A_k^j + \sigma^2 I(n) Q_j = 0(n \times n) , \]

and if this is to hold for all \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( \sigma^2 \), we have

\[(3.5.3) \quad q_j A_j^j Q_j = 0(n \times n), j=1,\ldots,k; q_j Q_j = 0(n \times n) \]

We shall now justify the restrictions (i) and (ii) on the quadratic forms \( q_0, q_1, \ldots, q_k \), from the standpoints of (a) point estimation of the variance components; (b) tests of hypotheses on them; and (c) simultaneous confidence interval estimation of the variance components and certain functions of them.

(a) If \( q_i \)'s satisfy (a), (\( \beta \)) and (\( \gamma \)), then, as seen in (3.5.1), \( \lambda_i \)'s are linear functions, with positive coefficients, of \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( \sigma^2 \), and they are independent of \( \mu_1, \ldots, \mu_k \). In fact, (3.5.1) holds even when (\( \beta \)) alone is satisfied. Furthermore, when these conditions are satisfied, \( q_0, q_1, \ldots, q_k \), form a set of sufficient statistics for \( \lambda_0, \lambda_1, \ldots, \lambda_k \) (also for \( \sigma_1^2, \ldots, \sigma_k^2, \sigma^2 \)), and they can be shown to satisfy the completeness condition of Lehmann and Scheffé \( \sum_{27} \), as noted by the authors of \( \sum_{17} \). Then, from a result due to Lehmann and Scheffé \( \sum_{27} \) on completeness and sufficiency, it follows that the unbiased estimate with uniformly least variance, of an estimable linear function of \( \lambda_0, \lambda_1, \ldots, \lambda_k \) is given by the corresponding linear function of \( q_0/n_0, \ldots, q_k/n_k \), and this estimate is a unique (s.e.) function of \( q_i \)'s. This result is essen-
tially the same as that contained in \[\text{17}\]. We may summarize the result in the form of a theorem.

**Theorem 3.5c:** If \(q_0, q_1, \ldots, q_k\) are \((k + 1)\) quadratic forms of ranks \(n_0, n_1, \ldots, n_k\) respectively, satisfying (a), (b) and (\(\gamma\)) of lemma 3.5a, then the unbiased estimate with uniformly least variance of the estimable function \(\frac{1}{\sum_{i=0}^{k} \lambda_i}, \ldots, \frac{1}{\sum_{i=0}^{k} \lambda_i}\), is given by \(\frac{1}{\sum_{i=0}^{k} \lambda_i q_i / n_i}\), and this estimate is a unique (a.e.) function of \(q_0, q_1, \ldots, q_k\).

**Corollary:** In particular, under Model II, if the \(\lambda_i\)'s are linearly independent, i.e., the matrix of coefficients of the equations \(\lambda_i = q_i / n_i, i = 0, 1, \ldots, k,\) is non-singular, so that unique unbiased estimates of \(\sigma^2, \sigma_1^2, \ldots, \sigma_k^2\) exist, then the estimates so obtained have uniformly least variances.

It may be noted that the Theorem 3.5c holds, not only for a linear function of \(\lambda_i\)'s, but also, in general, for any real valued estimable function, \(f(\lambda_0, \lambda_1, \ldots, \lambda_k)\).

(b) Hypotheses on the variance components are composite, the free parameters being \(\mu_1, \mu_2, \ldots, \mu_k\), or, \(\mu = \sum_{i=1}^{k} \mu_i\). A legitimate quest might be to obtain similar region tests for these hypotheses.

From the properties of sufficiency and completeness of \(q_0, q_1, \ldots, q_k\), mentioned above in (a), it follows, from another theorem due to Lehmann and Scheffé \[\text{27}\], that the class of all similar tests of hypotheses on the \(\sigma^2\)'s will be of Neyman structure.
with respect to \( q_0, \ldots, q_k \). In the terminology of Roy \( \sqrt{\Phi} \),
the class of all similar regions will be Neyman mechanism regions.

(c) Finally, if the quadratic forms \( q_0, q_1, \ldots, q_k \) satisfy
(a), (b) and (c), then, as will be seen later in this chapter, we
can obtain simultaneous confidence intervals for \( \sigma_1^2, \ldots, \sigma_k^2, \sigma^2 \),
and for ratios like \( \sigma_i^2/\sigma^2 \), without running into difficult or in-
tractable distribution problems.

3.6 Tie-up between Model I analysis and Model II analysis for the
general k-way classification.

Under the Model I set-up we can obtain the \( k \) sums of squares
due to the \( k \) sets of hypotheses of equality of the elements of
\( \xi_i(m_1 \times 1), i = 1, 2, \ldots, k \). Further, we can obtain a sum of
squares due to error. The hypotheses of equality of the elements
of \( \xi_i(m_1 \times 1) \) can be written as,

\[
(3.6.1) \quad H_{0i} : C_{12} m_1 \xi_1(m \times 1) - C_{11} c_{12} - \xi(m \times 1), \text{ say,}
\]

where \( r \) and \( m \) are, as defined in the model,

\[
r(A) = m - k + 1 \quad \text{and} \quad \sum_{i=1}^{k} m_i, \text{ respectively.}
\]

\[
= (m_1 - 1) \left[ \begin{array}{cccc}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & 1 \\
\end{array} \right]_{m_1 \times k}
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]

\[
= \xi_1(m_1 - 1 \times 1),
\]
so that \( r(C_i) = (m_i - 1) \). This holds for \( i = 1, 2, \ldots, k \). It is known, \( \int \left[ \text{testability} \right] \), that for "testability" we must have \( C_{12}(\overline{m_1} - 1 \times \overline{m-r}) \) related to \( C_{11}(\overline{m_1} - 1 \times r) \) by the same post-factor as \( \Lambda_D(n \times \overline{m-r}) \) to \( \Lambda_I(n \times r) \), and that, if this condition is satisfied, the usual sum of squares due to the hypotheses (3.6.1) is,

\[
(3.6.2) \quad x' A_I (A_I' A_I)^{-1} C_{11}^{-1} C_{11} (A_I' A_I)^{-1} C_{11}^{-1} f^{-1} C_{11} (A_I' A_I)^{-1} A_I' x
\]

with \((m_i - 1)\) degrees of freedom, for \( i = 1, 2, \ldots, k \).

Also, we have the sum of squares due to error,

\[
(3.6.3) \quad x' / \overline{I}(n) - A_I (A_I' A_I)^{-1} A_I' x
\]

with \( n-r = (n-m+k+1) \) degrees of freedom.

We shall refer to these sums of squares, (3.6.2) and (3.6.3), as the anova sums of squares under Model I. The sums of squares (3.6.2), for \( i = 1, 2, \ldots, k \), will not always be (for example in incomplete block designs) the same as the anova sums of squares as they are usually defined. In the customary parlance these sums of squares, (3.6.2), would be called the sums of squares due to the hypotheses (3.6.1).

Now, under Model II, in the notation of section 3.5, suppose we take \( q_0 \) as the error sum of squares, (3.6.3), with \( n_0 = (n-m+k+1) \), and \( q_i \) as the sum of squares (3.6.2) with \( n_i = (m_i - 1) \), \( i = 1, 2, \ldots, k \), so that, \( \sum_{i=0}^{k} n_i = n-1 \), where \( n \) is the total number of observations. Notice that for these quadratic forms,
$q_i (i = 0, 1, \ldots, k)$, we have the following property:

(3.6.4) \hspace{1cm} Q_i^2 = Q_i \hspace{1cm} (i = 0, 1, \ldots, k)$

It can be shown that (3.6.4) is equivalent to,

(3.6.5) \hspace{1cm} Q_i = L_i^1 L_i$

where \hspace{1cm} \begin{align*} & L_i (n_1 \times n) L_i^1 (n \times n_1) = I(n_i), \\ & r(Q_i) = n_i,
\end{align*}$

for $i = 0, 1, \ldots, k$, and that, from (3.6.4) or (3.6.5), we have,

(3.6.6) \hspace{1cm} \text{all non-zero} \hspace{0.5mm} c(Q_i) = 1, \hspace{1mm} \text{so that,} \hspace{1mm} tr(Q_i) = r(Q_i) = n_i$

Keeping these properties, (3.6.4) - (3.6.6), in mind, we can apply the conditions, (a), (b) and (γ), of lemma 3.5a to these $q_i$'s. On doing so, we shall find that (b) is satisfied by all the $q_i$'s, so that centrality of the distribution, (if it is $\chi^2$ at all) is assured for each $q_i, i = 0, 1, \ldots, k$. It would, therefore, follow that,

\[
\lambda_i = \frac{1}{n_i} \text{tr} \Sigma Q_i = \frac{2}{(m_i - 1)} \left\{ \text{sum of the independent elements of } \Gamma_{11} (A_{11}^{-1})^{-1} C_{11}^{-1} \right\} \sigma_i^2
\]

(3.6.7) \hspace{1cm} + \sigma^2$, for $i = 1, 2, \ldots, k$

\[
= \nu_i \sigma_i^2 + \sigma^2, \hspace{1mm} \text{say, where} \hspace{1mm} \nu_i = \frac{2}{(m_i - 1)} \left\{ \text{sum of the independent elements of } \Gamma_{11} (A_{11}^{-1})^{-1} C_{11}^{-1} \right\} > 0
\]

and \hspace{1cm} $\lambda_0 = \frac{1}{n_0} \text{tr} \Sigma Q_0 = \sigma^2$

It is easily seen that, for these $q_i$'s, the $\lambda_i$'s are linearly in-
dependent, so that, the optimum property, mentioned in (a) of section 3.5, will hold for estimates of \( \sigma_1^2, \ldots, \sigma_k^2, \sigma^2 \) obtained by using the \( q_i \)'s, if they satisfy (a), (\( \beta \)) and (\( \gamma \)) of lemma 3.5a.

Again, it is easily verified that \( q_0 \) always satisfies (a) and (\( \gamma \)), i.e., \( q_0 \), the sum of squares due to error, under Model I, is always, under Models I and II, distributed as \( \sigma^2 x_{n_0}^2 \), where \( x_{n_0}^2 \) is the central \( x^2 \) variate with \( n_0 = (n-m+k-1) \) degrees of freedom; and \( q_0 \) is distributed independently of \( q_1, q_2, \ldots, q_k \). Next, applying (a) and (\( \gamma \)) to \( q_1, q_2, \ldots, q_k \), and simplifying the conditions, we obtain them respectively in the forms,

\[
(3.6.8) \quad \sum_{i=1}^{k} C_{i1} C_{i2} \begin{bmatrix}
\sigma_1^2 I(m_1) & 0 & \cdots & 0 \\
0 & \sigma_2^2 I(m_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sigma_k^2 I(m_k)
\end{bmatrix} - \\
(\lambda_1 - \sigma^2) \begin{bmatrix}
C_{11} \quad \sigma_1^2 I(m_1) \\
C_{12} \quad \sigma_2^2 I(m_2) \\
\vdots \quad \vdots \\
C_{k1} \quad \sigma_k^2 I(m_k)
\end{bmatrix} = \begin{bmatrix}
\begin{bmatrix} C_{i1} \\ C_{i2} \end{bmatrix} \\
\begin{bmatrix} C_{i1} \\ C_{i2} \end{bmatrix} \\
\begin{bmatrix} C_{i1} \\ C_{i2} \end{bmatrix} \\
\begin{bmatrix} C_{i1} \\ C_{i2} \end{bmatrix}
\end{bmatrix}
\]

\( (\lambda_1 - \sigma^2) \sum_{i=1}^{k} C_{i1} (A_i' A_i)^{-1} C_{i1} = \lambda_i - 1, \quad \sum_{i=1}^{k} C_{i1} (A_i' A_i)^{-1} C_{i1} \quad \sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \lambda_i \)

and,

\[
(3.6.9) \quad C_{i1} (A_i' A_i)^{-1} C_{j1} = 0 \quad \text{for} \quad i \neq j, \quad 1 \leq i, j \leq k
\]

The condition, (3.6.9), for independent distribution of the sums of squares (3.6.2), which are the sums of squares due to composite hypotheses like (3.6.1), is a relatively trivial result if one approaches anova from the standpoint of linear estimation and breakdown of the total sum of squares into suitable components. It is more important if one approaches anova from the standpoint of
simultaneous testing of different composite hypotheses and simultaneous confidence interval estimation associated with this problem. Finally, this condition is indispensable if one tries to generalize anova to the non-parametric set-up, and specially to the case of categorical data, where the breakdown of the total sum of squares into suitable components, which is an integral part of the customary anova procedures, is the first thing that one has to abandon, while most of the other features of anova carry over in a nice manner.

Using the knowledge of \( \sum C_{i1}^2 \), from (3.6.1), we can simplify (3.6.8) further into,

\[
(3.6.10) \quad C_{i1}(A_i'A_i)^{-1}C_{i1} = \frac{\sigma^2}{\lambda_1 - \sigma^2} \sum I(m_1-1) + J(m_1-1 x m_1-1) \]

\[
= \frac{1}{v_i} \sum I(m_1-1) + J(m_1-1 x m_1-1) ,
\]

using (3.6.7). It may be noted that (3.6.9) and (3.6.10) are independent of the unknown \( \sigma^2 \)’s.

It has been verified that for the usual complete block designs, like the Randomized Block, Latin Square and Factorial designs, both the conditions (3.6.9) and (3.6.10) are satisfied. However in general, both these conditions are not simultaneously satisfied by the incomplete designs, like the Balanced Incomplete Block designs. Thus the restrictions on the \( q_i \)’s, mentioned at the beginning of section 3.5, are not too restrictive in that the usual
complete designs have anova sums of squares under Model I satisfying them, but they are restrictive in that the incomplete block designs have anova sums of squares not satisfying them. More about this will be said in the concluding section of this chapter.

3.7 Simultaneous confidence statements.

In this section we shall be concerned with the k-way classification designs, which have anova sums of squares, under Model I, which satisfy the conditions on quadratic forms, mentioned at the beginning of section 3.5, under Model II, i.e., we shall be concerned with the complete class, k-way classification designs.

We now proceed to obtain certain simultaneous confidence interval statements.

(1) If the \( q_1 \)'s satisfy the conditions of 3.5, then we can find constants, \( \chi^2_{1a_j}(n_j) = \chi^2_{1a_j} \) (say) and \( \chi^2_{2a_j}(n_j) = \chi^2_{2a_j} \) (say), for \( j = 0, 1, \ldots, k \), such that, the simultaneous statements,

\[
(3.7.1) \quad \chi^2_{1a_0} \leq \frac{q_0}{\lambda_0} \leq \chi^2_{2a_0}, \quad \chi^2_{1a_1} \leq \frac{q_1}{\lambda_1} \leq \chi^2_{2a_1}, \ldots, \quad \chi^2_{1a_k} \leq \frac{q_k}{\lambda_k} \leq \chi^2_{2a_k}
\]

have a joint confidence coefficient, \( (1 - a) = \prod_{j=0}^{k}(1-a_j) \), where

\[ P( \chi^2_{1a_j}(n_j) \leq \chi^2_{2a_j}(n_j) = (1-a_j), \quad \chi^2_{(n_j)} \text{ stands for the central} \]

\( \chi^2 \) variate with \( n_j \) degrees of freedom, \( (j = 0, 1, \ldots, k) \). By inverting the statements (3.7.1), we obtain the simultaneous confidence statements,
\[(3.7.2) \quad \frac{\sigma_0}{x_{2\alpha_0}} \leq \frac{\sigma_0}{x_{2\alpha_1}} \leq \frac{\sigma_1}{x_{2\alpha_1}} \leq \frac{\sigma_1}{x_{2\alpha_1}} \leq \cdots \leq \frac{\sigma_k}{x_{2\alpha_k}} \leq \frac{\sigma_k}{x_{2\alpha_k}} \leq \frac{\sigma_k}{x_{2\alpha_k}} \]

or

\[c_{1\alpha_0} \leq c_{2\alpha_0}, \quad c_{1\alpha_1} \leq c_{2\alpha_1}, \quad \cdots c_{1\alpha_k} \leq c_{2\alpha_k}, \]

where \(c_{1\alpha_j} = \frac{q_j}{x_{2\alpha_j}}\) and \(c_{2\alpha_j} = \frac{q_j}{x_{2\alpha_j}}\), with a joint confidence coefficient \(= (1 - \alpha)\).

Recalling (3.6.7), we see that (3.7.2) is equivalent to the statement that the set of points \((\sigma_1, \ldots, \sigma_k, \sigma^2)\) lie in the \((k+1)\) dimensional space bounded by the \((k+1)\) pairs of \(k\)-dimensional flats,

\[(3.7.3) \quad \lambda_0 = \frac{\sigma_0}{x_{2\alpha_0}} = c_{1\alpha_0}, \quad \lambda_0 = \frac{\sigma_0}{x_{2\alpha_0}} = c_{2\alpha_0}, \quad \lambda_1 = \frac{\sigma_1}{x_{2\alpha_1}} = c_{1\alpha_1}, \quad \lambda_1 = \frac{\sigma_1}{x_{2\alpha_1}} = c_{2\alpha_1}, \quad \lambda_k = \frac{\sigma_k}{x_{2\alpha_k}} = c_{1\alpha_k}, \quad \lambda_k = \frac{\sigma_k}{x_{2\alpha_k}} = c_{2\alpha_k}

The associated probability, viz., the probability that the \((k+1)\) dimensional region, defined above, will contain the point \((\sigma_1^2, \ldots, \sigma_k^2, \sigma^2)\), is, of course, \((1 - \alpha)\). To explain the meaning of the region defined by statements like (3.7.3), let us consider the case \(k=1\), i.e., the one-way classification. The statements (3.7.2), in this case, will reduce to,

\[(3.7.4) \quad c_{1\alpha_0} \leq \sigma^2 \leq c_{2\alpha_0}, \quad c_{1\alpha_1} \leq \sigma^2 \leq c_{2\alpha_1} \]
It is easily seen that the region defined by (3.7.4) can be geometrically pictured as the shaded area in the figure below.

The statements (3.7.3), in this case, are seen to define this same region. The confidence region, in the space of $\sigma^2$'s, is visualized easily in the cases when $k = 1$ and $k = 2$. For larger values of $k$, however, the visualization is not so easy. However, remembering that $\lambda_0 = \sigma^2$ and $\lambda_i = \nu_i \sigma_1^2 + \sigma^2$, $(i = 1, 2, \ldots, k)$, we can obtain the following set of simultaneous confidence interval statements on the $\sigma^2$'s, which intervals are implied by (3.7.2):

$$(3.7.5) \quad c_1 a_0 \leq \sigma^2 \leq c_2 a_0, \quad \frac{c_1 a_1 - c_2 a_0}{\nu_1} \leq \sigma_i \leq \frac{c_2 a_1 - c_1 a_0}{\nu_1}, \ldots, \frac{c_1 a_k - c_2 a_0}{\nu_k},$$

and since (3.7.2) $\implies$ (3.7.5), therefore, by lemma 1.2f, we have that the joint confidence coefficient of (3.7.5) is $\geq (1 - \alpha) = \frac{k}{J} (1 - \alpha_j)$. In order to be non-trivial, of course, the $c_1 a_j$ and $j = 0$
\( c_{2i} \), which are defined in (3.7.2), should be such that all the bounds in (3.7.5) are non-negative.

(ii) As a simple extension, we can obtain simultaneous confidence interval statements on \( \mu = \mu_1 + \mu_2 + \cdots + \mu_k \) and \( \sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2, \sigma^2 \).

It is well known that if \( X^{(n \times 1)} \) has the \( n \)-variate non-singular normal distribution \( N^{\mathbb{R}^n(n \times 1), \Sigma(n \times n)} \), then, for nonnull \( \mathcal{F}^t(l \times n) \), \( \mathcal{F}^t X \) has the univariate normal distribution \( N^{\mathbb{R}^t, \mathcal{F}^t \Sigma \mathcal{F}} \). Under this set-up, we also have the lemma,

**Lemma 3.7a:** \( \mathcal{F}^t X \) is distributed independently of \( X^t Q X \), which has a \( \chi^2 \) distribution, if, and only if, \( \mathcal{F}^t Q = \mathcal{F}'(1 \times n) \).

For a proof see \( \mathcal{F}^t \).

Using this lemma, it can be verified that the general mean, \( \bar{x} = \left( \frac{1}{n}, \ldots, \frac{1}{n} \right) X^{(n \times 1)} \), is distributed independently of \( q_0, q_1, \ldots, q_k \). Also, it can be seen that \( E(\bar{x}) = \mu \), so that, \( \bar{x} \) is an unbiased estimator of \( \mu \), although it may not always have the least variance in this class of estimators (for the two-way classification it also has the least variance). Due to the simplicity of calculating \( \bar{x} \), and its being distributed independently of \( q_0, \ldots, q_k \), we shall take \( \bar{x} \) as the estimate of \( \mu \) in general. Now \( \bar{x} \), of course, is distributed as \( N_{\hat{\mu}, \sigma_0^2 + \Sigma_{i=1}^k a_i \sigma_i^2} \), where \( a_0, a_1, \ldots, a_k \) are known functions of \( n, m_1, \ldots, \) and \( m_k \). Thus, we can make the simultaneous statements,
\[ -\frac{t}{c_{k+1}} \sqrt{\frac{\sigma^2 + \sum_{i=1}^{k} a_i \sigma^2}{\sum_{i=1}^{k+1} a_i \sigma^2}} \leq \bar{x} - \mu \leq \frac{t}{c_{k+1}} \sqrt{\frac{\sigma^2 + \sum_{i=1}^{k} a_i \sigma^2}{\sum_{i=1}^{k+1} a_i \sigma^2}} \]

\[ c_{1a_0} \leq \lambda_0 \leq c_{2a_0} \]

(3.7.6)

\[ c_{1a_1} \leq \lambda_1 \leq c_{2a_1} \]

\[ \ldots \quad \ldots \quad \ldots \]

\[ c_{1a_k} \leq \lambda_k \leq c_{2a_k} \]

With a joint probability \( (1 - \alpha) = \prod_{j=0}^{k+1} (1 - a_j) \), where \( t_{a_{k+1}} \) is the \( a_{k+1} \) th percentile of \( N(0,1) \), and \( c_{1a_j}, c_{2a_j} \) are as defined in (3.7.2).

Just as we obtained (3.7.5) from (3.7.2), we can obtain, from (3.7.6), the implied simultaneous confidence interval statements,

\[ \bar{x} - \frac{t_{a_{k+1}}}{\sqrt{b_0 c_{2a_0} + b_1 (c_{2a_1} - c_{1a_0}) + \ldots + b_k (c_{2a_k} - c_{1a_0})}} \leq \mu \leq \bar{x} + \frac{t_{a_{k+1}}}{\sqrt{b_0 c_{2a_0} + \ldots + b_k (c_{2a_k} - c_{1a_0})}} \]

(3.7.7)

\[ c_{1a_0} \leq \sigma^2 \leq c_{2a_0} \]

\[ \frac{c_{1a_0} - c_{2a_0}}{\nu_1} \leq \sigma^2 \leq \frac{c_{2a_1} - c_{1a_0}}{\nu_1} \]

\[ \ldots \quad \ldots \quad \ldots \]

\[ \frac{c_{1a_k} - c_{2a_0}}{\nu_k} \leq \sigma^2 \leq \frac{c_{2a_k} - c_{1a_0}}{\nu_k} \]
with a joint confidence coefficient $\geq (1 - \alpha) = \prod_{j=0}^{k+1} (1-\alpha_j)$, where

$b_0 = a_0$, $b_i = a_i / v_i$ for $i = 1, 2, \ldots, k$, which are, therefore, known functions of $n$, $m_1, \ldots$, and $m_k$, depending only on a knowledge of the model matrix $A(n \times m)$.

(iii) Next we shall obtain simultaneous confidence bounds on the ratios $\sigma_1^2 / \sigma^2$, $\sigma_2^2 / \sigma^2$, $\ldots$, $\sigma_k^2 / \sigma^2$.

When the above sums of squares under Model I, $q_0$, $q_1$, $\ldots$, $q_k$, satisfy the conditions of 3.5, then it is seen that

$F_1(n_1, n_0) = \frac{q_1 / n_1 \lambda_1}{q_0 / n_0 \lambda_0}$ has the central $F$-distribution with $n_1$ and $n_0$ degrees of freedom, $(i = 1, 2, \ldots, k)$. By (3.6.9), the numerators of the $F_1(n_1, n_0)$’s will all be distributed independently, and independently of their common denominator. The $F_1$’s will then be quasi-independent in the sense of Ghosh $\chi_{15}$, Ramachandran $\chi_{34}$, and Roy $\chi_{43}$. The joint distribution of these quasi-independent $F_1$ is known $\chi_{34}$, and, as in section 2.4, constants $F_{11}, F_{12}, \ldots, F_{12}, F_{12}, \ldots, k$, can be determined, such that the simultaneous statements,

$$ F_{11} \leq \frac{q_1 / n_1 \lambda_1}{q_0 / n_0 \lambda_0} \leq F_{12}, \ldots, F_{k1} \leq \frac{q_k / n_k \lambda_k}{q_0 / n_0 \lambda_0} \leq F_{k2} $$

have a joint probability $= (1 - \alpha)$, for a preassigned $\alpha$.

Recalling that $\lambda_1 = v_1 \sigma_1^2 + \sigma^2$, $n_1 = (m_1 - 1)$, for $i = 1, 2, \ldots, k$, and $\lambda_0 = \sigma^2$, $n_0 = (n - m + k - 1)$, we can invent (3.7.8)
to obtain the simultaneous confidence statements,

\[(3.7.9) \frac{n_0}{n_1^{12}} \frac{q_1}{q_0} \leq \frac{\lambda_1}{\lambda_0} \leq \frac{n_0}{n_1^{11}} \frac{q_1}{q_0}, \ldots, \frac{n_0}{n_k^{k2}} \frac{q_k}{q_0} \leq \frac{\lambda_k}{\lambda_0} \leq \frac{n_0}{n_k^{kl}} \frac{q_k}{q_0}\]

or

\[(3.7.10) \frac{1}{\sqrt{n_1^{12}}} \frac{n_0}{n_1^{12}} \frac{q_1}{q_0} - 1 \leq \frac{\sigma_1^2}{\sigma^2} \leq \frac{1}{\sqrt{n_1^{11}}} \frac{n_0}{n_1^{11}} \frac{q_1}{q_0} - 1, \ldots\]

\[\frac{1}{\sqrt{n_k^{k2}}} \frac{n_0}{n_k^{k2}} \frac{q_k}{q_0} - 1 \leq \frac{\sigma_k^2}{\sigma^2} \leq \frac{1}{\sqrt{n_k^{kl}}} \frac{n_0}{n_k^{kl}} \frac{q_k}{q_0} - 1, \]

with a joint confidence coefficient \(= (1 - \alpha)\). Here, again, for non-triviality the bounds should all be positive.

The results of (iii) are essentially the same as those in section 2.6, the difference being that we work here in terms of \(\lambda_0, \lambda_1, \ldots, \lambda_k\), in place of \(\sigma_1^2, \ldots, \sigma_k^2\) in 2.6, and \(q_0, q_1, \ldots, q_k\) in place of \(s^2, s_1^2, \ldots, s_k^2\) in 2.6.

3.8 Tests of hypotheses on \(\sigma_1^2, \ldots, \sigma_k^2\) and \(\sigma^2\).

The usual hypotheses tested are \(H_0^i; \sigma_i^2 = 0\), against \(H_i^i; \sigma_i^2 > 0\), for \(i = 1, 2, \ldots, k\). Working in terms of the anova sums of squares under Model I, it is seen that these hypotheses are equivalent to \(H_0^i; \lambda_i = \lambda_0\) against \(H_i^i; \lambda_i > \lambda_0\), for \(i = 1, 2, \ldots, k\). For each \(i\), therefore, under Model II, we can test \(H_0^i\) against \(H_i^i\) by taking as the critical region the region defined by,

\[(3.8.1) \quad F_i(n_1, n_0) > F_{\lambda_i}(n_1, n_0),\]
where \( F_1(n_1, n_0) = \frac{q_1/n_1}{q_0/n_0} \) has the central F-distribution with \( n_1 \) and \( n_0 \) degrees of freedom, under \( H'_0 \), and, where \( F_{a_1}(n_1, n_0) \) is the upper \( a_1 \) % point of \( F \) with \( n_1 \) and \( n_0 \) degrees of freedom.

Notice that the critical regions, (3.8.1), for the individual hypotheses \( H_{01} \), under Model II, are the same as the critical regions obtained for the individual hypotheses, (3.6.1), under Model I.

This identical nature of critical regions, under Models I and II, is true even when we consider the simultaneous hypotheses, \( H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = 0 \), against, \( H_1: \) at least one \( \sigma_i^2 > 0 \), which is equivalent to considering, \( H_0: \frac{\lambda_1}{\lambda_0} = \ldots = \frac{\lambda_k}{\lambda_0} = 1 \), against, \( H_1: \) at least one \( \frac{\lambda_i}{\lambda_0} > 1 \). The critical region of the simultaneous test is,

\[
(3.8.2) \quad F_1(n_1, n_0) > a_1, \ldots, F_k(n_k, n_0) > a_k
\]

where \( F_1(n_1, n_0) = \frac{q_1/n_1}{q_0/n_0} \), \( F_1 \)'s are quasi-independent variance ratios, and \( a_1 \)'s are such that the size of the region (3.8.2) is \( a \), a pre-assigned level. It is easily seen that the critical region (3.8.2) is the same as the critical region of the simultaneous anova test of Ghosh under Model I \( \text{cf. 34.7} \).

It must be noted, however, that, while the critical regions for both, the individual and the simultaneous tests, of the custom-
any hypotheses, under Models I and II, are the same, yet the power functions are different under the two models.

3.9 Remarks on confidence bounds of section 3.7.

The results obtained in 3.7, for simultaneous confidence bounds, are valid for all choices of \( \chi^2_{1a_j}, \chi^2_{2a_j}, (j = 0, 1, \ldots, k) \), in (3.7.1), and for all choices of \( F_{i1}, F_{i2}, (i = 1, \ldots, k) \), in (3.7.8), for which we have, respectively,

\[
(3.9.1) \quad P\left( \chi^2_{1a_0} \leq \chi^2_{0} \leq \chi^2_{2a_0}, \ldots, \chi^2_{1a_k} \leq \chi^2_{0} \leq \chi^2_{2a_k} \right) = \prod_{j=0}^{k} (1-a_j) = \prod_{j=0}^{k} (1-a_j)
\]

and,

\[
(3.9.2) \quad P\left( F_{i1} \leq F_{12} \ldots, F_{i2} \leq F_{k2} \right) = (1-a).
\]

For the respective associated hypotheses,

(i) \( H_0: \lambda_0 = \lambda_{00}, \ldots, \lambda_k = \lambda_{0k} \), against, \( H_1: \lambda_1 \neq \lambda_{01} \) : for at least one \( i \).

(ii) \( H_0: \frac{\lambda_i}{\lambda_0} = \tau_i (> 1), \) for \( i = 1, 2, \ldots, k \), against, \( H_1: \frac{\lambda_i}{\lambda_0} \neq \tau_i \), for at least one \( i \), we have the respective acceptance regions,

\[
(3.9.3) \quad \chi^2_{1a_0} \leq q_0/\lambda_{00} \leq \chi^2_{2a_0}, \ldots, \chi^2_{1a_k} \leq q_k/\lambda_{0k} \leq \chi^2_{2a_k},
\]

and,

\[
(3.9.4) \quad F_{i1} \leq \frac{1}{\tau_i} \frac{q_i/n_i}{q_0/n_0} \leq F_{i2}, \ldots, F_{k1} \leq \frac{1}{\tau_k} \frac{q_k/n_k}{q_0/n_0} \leq F_{k2}.
\]

If \( \chi^2_{1a_j}, \chi^2_{2a_j}, (j = 0, 1, \ldots, k) \), and \( F_{i1}, F_{i2}, (i = 1, \ldots, k) \), and
..., k), are so chosen as to make the tests with acceptance regions, (3.9.3) and (3.9.4), have the monotonicity property, (which will imply complete unbiasedness as well), then, of course, the confidence bounds in (3.7.2) and (3.7.9) will have the desirable property of monotonically increasing shortness (in terms of probability of covering wrong values).

3.10  **A Mixed Model.**

Suppose, (i) \[ x(n \times 1) = \Lambda(n \times m) \xi(m \times 1) + \eta(n \times 1), \quad m < n \]

\[
\begin{bmatrix}
\xi_0 \\
m_0 \\
m_1 \\
\vdots \\
m_k
\end{bmatrix}
\]

where \[ \sum_{i=0}^{k} m_i = m \quad \text{and} \quad r(\Lambda) = r \leq m (\leq n); \]

(ii) \[ \xi_0 \ (m_0 \times 1) \] is a vector of fixed parameters, and \[ \xi_i \ (m_i \times 1), \]

for \( i = 1, 2, \ldots, k, \) is a random sample of size \( m_i \) from \[ N(\mu_i, \sigma_i^2); \]

the vectors \[ \xi_i \ (m_i \times 1), \] (\( i = 1, 2, \ldots, k, \)) are distributed independently of one another and independently of \( \eta(n \times 1); \)

(iii) \[ \eta(n \times 1) \] is a random sample of size \( \eta \) (whose elements are, physically, in the nature of errors), from \[ N(0, \sigma^2). \]

3.11  **Remarks on, and deductions from, the Mixed Model.**

The term **Mixed Model** is used for the model defined in 3.10, because it is, in a sense, a combination of Models I and II. Both
these latter models are special cases of the mixed model. When 
\( m_i = 0 \), for \( i = 1, 2, \ldots, k \), i.e., when there is no stochastic 
part in the vector \( \xi(m \times 1) \), then we have a purely parametric 
model, or, Model I. On the other hand, if \( m_0 = 0 \), i.e., there is 
no fixed parametric part in \( \xi(m \times 1) \), then we have the Model II 
of section 3.2. Also, \( \xi_0(m_0 \times 1) \) might itself be considered to 
consist of subvectors of fixed parameters, so that, the model de-
defined in 3.10 is quite general. Here, as before, we shall consider 
in detail only the case where the model matrix, \( A(n \times m) \), is such 
that the sub-matrices, \( A_i(n \times m_i) \), \( (i = 0, 1, \ldots, k) \), have only 
one non zero element (equal to unity) in each row.

Under the Mixed Model, it can be seen that \( X(n \times 1) \) has an 
n-variate normal distribution, with

\[
E(X)(n \times 1) = A(n \times m) \begin{bmatrix}
\xi_0 \\
\mu_1 \\
\mu_2 \\
\vdots \\
\mu_k \\
1
\end{bmatrix}
\begin{bmatrix}
m_0 \\
m_1 \\
m_2 \\
\vdots \\
m_k \\
1
\end{bmatrix}, \text{where } 1(m_1 \times 1)
\]

stands for a column vector of \( m_1 \) unities, and

\[
E(n \times n) = \Sigma_{XX'} - E(X)E(X')
\]

\[
= A(n \times m) m_0 \begin{bmatrix}
0 & 0 & \cdots & 0 \\
m_1 & \sigma_1^2(m_1) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
m_k & 0 & \cdots & \sigma_k^2(m_k) \\
m_0 & m_1 & \cdots & m_k
\end{bmatrix}
\]

\( A' \) \( (n \times m) \), \( \sigma^2 I(n) \)
\[ \sigma_1^2 A_1^2 + \ldots + \sigma_k^2 A_k^2 + \sigma^2 I(n) \]

Our aims will be, (i) to find the best linear estimates of estimable linear functions of the elements of \( \xi_0, \mu_1, \mu_2, \ldots, \) and \( \mu_k, \) and to test testable hypotheses on them; (ii) to obtain estimates of, and test hypotheses on, estimable functions of \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( \sigma^2, \) and, in particular, the individual \( \sigma^2 \)'s themselves; (iii) to obtain simultaneous confidence bounds on parametric functions, if possible, without running into intractable distribution problems.

3.1.2 Certain results on linear estimability.

For the special case, when the model matrix, \( A(n \times m), \) is such that each row of \( A_i(n \times m_i), \) for \( i = 0, 1, \ldots, k, \) has only one non-zero element, which is unity, we have \( r(A) = r = m-k. \) Proceeding exactly as in the corollary of lemma 3.4b, we obtain that the NSC for estimability of, \( \mathbf{c} \mathbf{r}(\xi) = \mathbf{c}_0 \xi_0 + \mathbf{c}_1 \mathbf{E}(\xi_1) + \ldots + \mathbf{c}_k \mathbf{E}(\xi_k) = \mathbf{c}_0 \xi_0 + \sum_{i=1}^{k} \mu_i (\text{sum of the elements of } \mathbf{c}_i'), \) is that

\[(3.12.1) \quad \text{Sum of the elements of } \mathbf{c}_0' = \text{coefficient of } \mu_1 \]
\[= \ldots \]
\[= \text{coefficient of } \mu_k. \]

In place of the matrix post-factor, (3.4.3), we shall have, for the present case, the post-factor
As a consequence of (3.12.1), we note that, if we are interested in linear functions of only the elements of $\xi_0$, then, in order to be estimable, they should be contrasts. Also, the only independent linear function of the elements of $\xi_0$ and $\mu_1, \ldots, \mu_k$, which is estimable is, \[ \frac{1}{\xi_0} \sum_{i=1}^{m_0} \xi_{0i} + \mu_1 + \ldots + \mu_k, \]

where $\xi_0'(1 \times m_0) = (\xi_{01}, \ldots, \xi_{0m_0})$. In particular, neither the individual elements of $\xi_0$, nor their mean, nor the individual $\mu_i$'s are separately estimable.

3.13 **Tie-up between a Model I analysis and the Mixed Model analysis.**

We can present a tie-up between a $(k+1)$-way classification Model I analysis, and the above particular case of the Mixed Model.

Under Model I, for the $(k+1)$-way classification, we have $(k+1)$ sums of squares, each due to a hypothesis like,

\[ H_{Ci}: C_i(\bar{m}_{k-1} \times m) \xi = \Omega(\bar{m}_{k-1} \times 1), \]
where \( C_i^I(\bar{m}_1-1 \times m) \) is of a similar form to that given in (3.6.1), for \( i = 0, 1, \ldots, k \). We shall have \((k + 1)\) sums of squares like (3.6.2), and an error sum of squares like (3.6.3). Let us denote the \((k + 1)\) sums of squares due to the different hypotheses (3.13.1), by \( s_0^2, s_1^2, \ldots, s_k^2 \), and the error sum of squares by \( s_e^2 \).

Under the particular case of the Mixed Model, we may obtain conditions in order that \( s_0^2, s_1^2, \ldots, s_k^2 \) and \( s_e^2 \) be distributed independently as \( \lambda_0 \chi^2_{(m_0-1)}, \ldots, \lambda_k \chi^2_{(m_k-1)} \) and \( \lambda_e \chi^2_{(n-m+k)} \) respectively. It will be seen that \( s_e^2 \) is always distributed as \( \lambda_e \chi^2_{(n-m+k)} \), where \( \lambda_e = \frac{\sigma^2}{s_e^2/(n-m+k)} = \sigma^2 \), and that \( s_e^2 \) is distributed independently of \( s_0^2, \ldots, s_k^2 \). Also, under \( H_0 \), i.e., under (3.13.1) with \( i = 0 \), we have \( \lambda_0 = \frac{1}{\nu_0} \left( \frac{s_0^2}{(m_0-1)} \right) = \sigma^2 \), while \( \lambda_i = \nu_i \sigma_i^2 + \sigma^2 \) for \( i = 1, 2, \ldots, k \), as in (3.6.7). Just as in section 3.6, we can obtain the conditions for independent \( \lambda_i \chi^2_{(m_i-1)} \) distributions of \( s_i^2, i = 0, 1, \ldots, k \), as,

\[
(3.13.2) \quad C_{i1}(A_i^I A_i^{-1}) - C_{i1}' = \frac{1}{\nu_i} \chi^2_{(m_i-1)} + J(\bar{m}_1-1 \times \bar{m}_1-1) \cdot J,
\]

\( (i = 0, 1, \ldots, k) \)

and

\[
(3.13.3) \quad C_{i1}(A_i^I A_i^{-1}) - C_{j1}' = 0(\bar{m}_1-1 \times \bar{m}_1-1), \quad (i \neq j = 0, 1, \ldots, k).
\]

and, for \( i = 0 \), in addition to (3.13.2), we note that the distribution of \( s_0^2 \) will be central only under \( H_0 \).

For the designs for which these conditions are satisfied, we
have, therefore, for the customary hypothesis on the fixed parameters, (equality), the usual all-contrast $F$-test based on $(m_0-1)$ and $(n-m+k)$ degrees of freedom. The statistic itself is, $F_0 = \frac{s_0^2/(m_0-1)}{s_e^2/(n-m+k)}$. Also, this $F$ will be quasi-independent of $F_i = F_i(m_i-1, n-m+k) = \frac{s_i^2/(m_i-1)}{s_e^2/(n-m+k)}$, for $i = 1, 2, \ldots, k$, which we use to test hypotheses, like those discussed in section 3.8, on $\sigma_i^2$ ($i = 1, 2, \ldots, k$).

Again, using $s_0^2$ and $s_e^2$, we can, as in section 1.6, set confidence bounds on the deviation (from $H_{00}$) function, $
abla^T C_{01}(\lambda_1, \lambda_2, \ldots, \lambda_k)^{-1} C_{01}^{-1} \nabla$, where the alternative to $H_{00}$ is $H_{10}: \nabla^T C_{01} C_{02} \nabla = n(m_0-1 \times 1)$. Using the fact of independent distributions of $s_0^2$, $s_1^2$, $\ldots$, $s_k^2$ and $s_e^2$, we may also obtain simultaneous confidence interval statements on this parametric function and $\lambda_1$, $\ldots$, $\lambda_k$, $\lambda_0$ (or, $\sigma_1^2$, $\ldots$, $\sigma_k^2$ and $\sigma^2$). These results involve nothing new in technique, and are, therefore, not worked out explicitly here.

3.14 Concluding remarks.

It was mentioned, at the end of section 3.6, that the conditions, (3.6.9) and (3.6.10), are not both satisfied, in general, by incomplete block designs. Let us take the usual balanced incomplete block designs, as an example.
Here, under Model I, we have two sums of squares like (3.6.2) and an error sum of squares like (3.6.3). It can be shown that, under Model I, the error sum of squares is always distributed as a \( \sigma^2 x^2 \) variate with appropriate degrees of freedom, and that it is so distributed independently of the other two sums of squares. Also, it can be shown that, under the two respective null hypotheses, these two sums of squares will be distributed individually as \( \sigma^2 x^2 \) variates with appropriate degrees of freedom, though not independently so, in general.

A similar situation exists when we consider the same anova sums of squares under Model II. The error sum of squares will be distributed, independently of the other two sums of squares, as a \( \sigma^2 x^2 \) variate with appropriate degrees of freedom. The two sums of squares due to hypotheses will be distributed individually as \( \lambda_1 x^2 \) variates, \( (i = 1, 2) \), though not independently so, in general.

The consequence of this is that, we can carry through the separate tests of the individual hypotheses like (3.6.1), under Model I, and of the individual hypotheses like \( H_{01} \) in 3.8, under Model II. Neither the simultaneous anova test, under Model I, nor the simultaneous test (3.8.2), under Model II, is valid. This is also reflected in confidence interval estimation under both the models. Confidence bounds on the individual parametric functions, (which are measures of deviation from the respective null hypotheses), under Model I, and, on the individual \( \sigma^2 \)'s and the individual
ratios, $\frac{\sigma_i^2}{\sigma^2}$, under Model II, can be obtained rather simply. However, the simultaneous confidence statements cannot be obtained so easily.

These results for the balanced incomplete block designs (and partially balanced incomplete block designs) will be given in greater detail in a later paper.

Also, for the problem of simultaneous tests and simultaneous confidence bounds, we need, essentially, the joint distribution of non-independent $\chi^2$ variates. Investigations along these lines are in progress, and, will be discussed in a later paper.
CHAPTER IV

MULTIVARIATE VARIANCE COMPONENTS

4.1 Introduction.

Multivariate extensions of the general univariate Model I have been made by Roy. So far as multivariate extensions of the univariate Model II of anova are concerned, there seems to have been no work done, to the knowledge of the author. The present chapter gives certain such extensions of results obtained in chapter III for the univariate case, and the results are all believed to be new.

Similar multivariate extensions of the univariate Mixed Model of anova can be carried out using the methods of this chapter, and, hence, they have not been separately discussed.

4.2 The Model.

(i) \( x'(n \times p) = A(n \times m) \xi(m \times p) + \eta(n \times p); m < n, p \leq (n-r) \)

\[ = n^{-1/2} A_1 A_2 \ldots A_k \begin{bmatrix} \xi_1 & m_1 + \eta(n \times p), \\ m_1 & m_2 & m_k \\ \xi_2 & m_2 \\ \vdots \\ \xi_k & m_k \end{bmatrix} \]

where \( \sum_{i=1}^{k} m_i = m \) and \( r(A) = r \leq m(< n) \);

(ii) \( \xi_1(m_1 \times p), \xi_2(m_2 \times p), \ldots, \xi_k(m_k \times p) \) are independent random
samples of sizes \( m_1, m_2, \ldots, m_k \), respectively, from the non-singular \( p \)-variate normal distributions \( \mathcal{N}_{\mu_i}(p \times 1), \Sigma_i(p \times p) \), \( i = 1, 2, \ldots, k \), and \( \xi_i(m_1 \times p) \), for all \( i \), is distributed independently of \( \eta(n \times p) \);

(iii) \( \eta(n \times p) \) is a random sample of size \( n \) from non-singular \( \mathcal{N}_{\mu_0}(p \times 1), \Sigma(p \times p) \).

4.3 Certain observations on, and deductions from, the Model.

We shall refer to the model of 4.2 as the complete Model II of multivariate anova. For the multivariate situation also, we shall consider, in detail, the analogue of the \( k \)-way classification in the univariate case, and we shall call this analogue, the multivariate \( k \)-way classification. It is interesting to note that for the multivariate \( k \)-way classification, we have the same model matrix, \( A(n \times m) \), as for the univariate \( k \)-way classification. This, of course, means that, under our general set-up, we can only have the same kind of design for all the variates, in a multivariate anova situation, as for any single variate. It is not possible to have different types of designs for the different variates, when observations are made on all of them, at the same time, in our experiment.

Let us write \( X'(n \times p) = n \left[ \begin{array}{cccc} x_1 & x_2 & \cdots & x_p \\ 1 & 1 & \cdots & 1 \end{array} \right] \), and

\[ X(pn \times 1) = n \left[ \begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_p \end{array} \right] \cdot \text{Then, under the model of 4.2, the elements} \]
of $X'(n \times p)$, i.e., of $X(p \times 1)$, have a $pn$-variate normal distribution, with,

\[(1.3.1) \quad E(X) (pn \times 1) = n \begin{bmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & A \\ m & m & \cdots & m \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \vdots \\ \mu_{k1} \\ \mu_{12} \\ \vdots \\ \mu_{k2} \\ \vdots \\ \mu_{1p} \\ \vdots \\ \mu_{kp} \end{bmatrix} m_l, \quad \text{where} \]

$\mu_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{ip} \end{bmatrix}$, for $i = 1, 2, \ldots, k$

$= A^k (on \times pm) \begin{bmatrix} \mu_{11} \\ \vdots \\ \mu_{k1} \\ \mu_{12} \\ \vdots \\ \mu_{k2} \\ \vdots \\ \mu_{1p} \\ \vdots \\ \mu_{kp} \end{bmatrix} m_l, \quad \text{where} \quad A^k = \begin{bmatrix} A & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & A \end{bmatrix}$

and,
(4.3.2)  $\Sigma^n(p \times p) = E(x x') - E(x)E(x')$

$$= A_1 A_1' x x_1 + \ldots + A_k A_k' x x_k + I \cdot x x,$$

where we recall the Kronecker product notation $\otimes$.

$$C(p q \times p q) = A(p \times p) \cdot x B(q \times q)$$

$$= \begin{bmatrix}
A_{11} & A_{12} & \ldots & A_{1q} \\
A_{21} & A_{22} & \ldots & A_{2q} \\
\ldots & \ldots & \ldots & \ldots \\
A_{q1} & A_{q2} & \ldots & A_{qq}
\end{bmatrix}$$

We shall be considering, in detail, only the case,

(4.3.3)  $\Sigma_i(p \times p) = \sigma_i^2 \Sigma(p \times p)$, for $i = 1, 2, \ldots, k$,

since, as is shown in section 4.6, the more general set-up does not lend itself to an easy treatment. The author of $\xi - 14$ points out that the situation, where dispersion matrices are constant multiples of one another, or of a standard dispersion matrix, occurs very often in genetical problems. Thus, our restricted set-up will still be meaningful in certain physical situations.

Our objectives will be, (i) to estimate and test hypotheses on estimable linear functions of the elements of $\Sigma_1, \ldots, \Sigma_k$; (ii) to obtain estimates of, and test hypotheses on, the characteristic roots, $c(\Sigma_1), \ldots, c(\Sigma_k)$ and $c(\Sigma)$; (iii) to obtain confidence bounds on $c(\Sigma_1), \ldots, c(\Sigma_k)$ and $c(\Sigma)$. Under the restricted set-up, (4.3.3), of course, (ii) is equivalent to obtaining estimates of, and testing hypotheses on, $\sigma_1^2, \ldots, \sigma_k^2$ and $c(\Sigma)$, while (iii) is equiva-
lent to obtaining confidence bounds, (both separate and simultaneous), on \( \sigma_1^2, \ldots, \sigma_k^2 \) and \( c(\Sigma) \).

4.4 Certain results on linear estimability.

To obtain a NSC for estimability of a linear function of the elements of \( \mu_1, \ldots, \mu_k \), we need only the lemmas already proved in section 3.4, remembering only that we shall be working with a \( p \times n \)-variate normal instead of an \( n \times p \)-variate normal distribution.

For the multivariate \( k \)-way classification Model II, we shall have the following matrix post-factor relating \( A_T^*(pn \times \frac{P}{P(m-r)}) \) to \( A_T^*(pn \times pr) \), where \( A_T^*(pn \times pm) = pn \sum A_T^* A_D^* \) and \( r(A) = r(m-k+1) \):

\[
A_T^*(pn \times pr) = \begin{bmatrix}
P & 0 & \cdots & 0 \\
r & P & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
r & 0 & \cdots & P \\
k-1 & k-1 & k-1 & k-1
\end{bmatrix}
\]

where \( P(r \times k-1) \) is given by (3.4.2). Using this \( P^* \), and proceeding exactly as in the corollary of lemma 3.4b, we can obtain that the NSC for the estimability of

\[
\phi_1^t(1 \times p) \mu_1(p \times l) + \cdots + \phi_k^t(1 \times p) \mu_k(p \times l)
\]

is that

\[
\phi_1^t(1 \times p) = \phi_2^t(1 \times p) = \cdots = \phi_k^t(1 \times p)
\]

This means that the only independent estimable linear function
of the elements of \( \mu_1, ..., \mu_k \) is of the form,
\[
\gamma'(1 \times p) \mu_1 + ... + \mu_k \gamma(1 \times p) \mu(p \times 1),
\] (say),

and all other such estimable linear functions have coefficients that can be obtained by post-multiplying \( \gamma' \) by a diagonal matrix. It can also be seen, from (4.4.3), that linear functions of the elements of any one \( \mu_i \) alone are not estimable, and, in particular, the individual elements of \( \mu_i \) are not separately estimable.

4.5 Multivariate Variance Components.

We now proceed to problems arising in the multivariate situation that are similar to those considered in section 3.5 for the univariate situation.

We shall seek \((k + 1)\) matrices, \( S_i(p \times p) = \frac{1}{n_i} X(p \times n)Q(n \times n) x x'(n \times p) \), for \( i = 0, 1, ..., k \), where \( Q_i(n \times n) \) is a symmetric at least p.s.d. matrix of rank \( n_i (\leq n) \), such that

(i) \( \frac{1}{\lambda_i} S_i \), of rank \( \leq p \), is distributed in the central pseudo-Wishart form (defined below) with \( n_i \) degrees of freedom, where an expression for the positive constant \( \lambda_i \) will be derived below.

\[
\sum_{i = 0, 1, ..., k} \lambda_i S_i = 0, 
\]

(ii) \( \frac{1}{\lambda_i} S_i \) is distributed independently of \( \frac{1}{\lambda_j} S_j \), for \( i \neq 0, 1, ..., k \).

**Definition of a pseudo-Wishart distribution:**

Suppose \( X(p \times n) \) has the distribution,
\[(4.5.1) \quad \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} \text{tr} \Sigma^{-1}(X-\mu)(X-\mu)^\prime \right) d\mathbf{x}; -\infty < x_{ij} < \infty, \]

so that \( E(X) = \mu \) and the symmetric p.d. matrix \( \Sigma(p \times p) \)

is interpreted as, \( n\Sigma(p \times p) = E\Sigma(X-\mu)(X-\mu)^\prime \). Then, we

shall call the distribution of \( S(p \times p) = \frac{1}{n} XX^\prime \), the (non-central)

pseudo-Wishart distribution with \( n \) degrees of freedom. The dis-

tribution of this \( S(p \times p) = \frac{1}{n} XX^\prime \) will be called a central pseudo-

Wishart distribution if, and only if, \( \mu(p \times n) = 0(p \times n) \), i.e.,

if, and only if, \( \mu = 0(p \times p) \). For this central pseudo-Wishart

distribution, we shall, therefore, have \( E(XX^\prime) = n\Sigma \), or, \( E(S) = \Sigma \).

Conversely, we shall say that any symmetric at least p.s.d. matrix,

\( S(p \times p) \), has the pseudo-Wishart distribution (in general, non-central)

with \( n \) degrees of freedom, if we can write \( S(p \times p) = \frac{1}{n} X(p \times n) \times

X^\prime(n \times p) \), where \( X(p \times n) \) has the distribution \((4.5.1)\). Further,

if \( \mu = 0(p \times p) \), then the distribution will be said to be

central.

In particular, if \( p \leq n \) and \( r(S) = p \), then a pseudo-Wishart

distribution for \( S \) is equivalent to the ordinary Wishart distribu-
tion.

We shall next prove certain lemmas establishing NSC for the

restrictions \((i)\) and \((ii)\) mentioned above.

Lemma \( 4.5a \) If \( X(p \times n) \) has the distribution \((4.5.1)\), then the NSC

for \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^*} XQX^\prime \), where \( Q(n \times n) \) is a symmetric p.s.d.
matrix of rank $n^*$, to have a central pseudo-Wishart distribution with $n^*$ degrees of freedom are,

(i) $Q_0^2(n \times n) = \lambda Q$, which is $Q=\lambda L_1(n \times n^*)L(n^* \times n)$, where $LL^t = I(n^*)$

(ii) $E(X)Q E(X^t) = 0(p \times p)$

Furthermore, if $\frac{1}{\lambda} S_1^t(p \times p) = \frac{1}{n_1^*} XQ_1X^t$, where $Q_1(n \times n)$ is a symmetric p.s.d. matrix of rank $n_1^*$, then the NSC that $\frac{1}{\lambda} S^*$ and $\frac{1}{\lambda} S_1^*$ are distributed independently is,

(iii) $Q_1 = 0(n \times n)$, which is $L(n^* \times n) L_1^t(n \times n_1^*) = 0(n^* \times n_1^*)$, where $L(n^* \times n)$ is defined in condition (i) and $L_1(n_1^* \times n)$ is such that, $Q_1(n \times n) = \lambda_1 L_1^t L_1$, where $L_1 L_1^t = I(n_1^*)$.

Proof: We shall first establish the necessity and sufficiency of conditions (i) and (ii), and then do the same for condition (iii).

Sufficiency of (i) and (ii): Let $Q(n \times n) = \lambda L_1^t L_1$, where $L_1^t = I(n^*)$.

Then $XX^t = XL_1^t LX^t + XL_1^t L_1X^t - L_1^t L_1X^t$

$= XL_1^t LX^t + XL_1^t L_0L_0^t X^t$, where $L_0(n-n^* \times n)$ is a completion of $L(n^* \times n)$ such that, $\begin{bmatrix} L \\ L_0 \end{bmatrix}$ is $\underline{\underline{1}}$, so that, $L_0 L_0^t = I(n-n^*)$ and $LL_0^t (n \times n-n^*) = 0$, i.e., $XX^t = Z(p \times n^*)Z_1(n^* \times p) + Z(p \times n-n^*)Z_0(n-n^* \times p)$, where $Z(p \times n^*) = X(p \times n) L_1^t(n \times n^*)$ and $Z_0(p \times n-n^*) = X(p \times n)L_0^t(n \times n-n^*)$.

From (4.5.1) the joint distribution of $Z$ and $Z_0$ is,
\[
\frac{1}{pn} \frac{n}{(2\pi)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} \left( \begin{bmatrix} Z \\ Z_0 \end{bmatrix} - \begin{bmatrix} L \\ L_0 \end{bmatrix} \right) \left( \begin{bmatrix} Z' \\ Z'_0 \end{bmatrix} - \begin{bmatrix} L' \\ L'_0 \end{bmatrix} \right) \right\} 
\]

\[
x \, dZ \, dZ_0; \quad -\infty < (\text{elements of } Z \text{ and } Z_0) < \infty ,
\]

since the Jacobian, \( J(X; Z, Z_0) = 1 \).

That is, the joint distribution of \( Z \) and \( Z_0 \) is,

\[
\frac{1}{pn} \frac{n}{(2\pi)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} \left( \begin{bmatrix} Z \\ Z_0 \end{bmatrix} - \begin{bmatrix} \tau \\ \tau_0 \end{bmatrix} \right) \left( \begin{bmatrix} Z' \\ Z'_0 \end{bmatrix} - \begin{bmatrix} \tau' \\ \tau'_0 \end{bmatrix} \right) \right\} \, dZ \, dZ_0,
\]

where \( p \, \begin{bmatrix} \tau \\ \tau_0 \end{bmatrix} \) = \( p \, \begin{bmatrix} L \\ L_0 \end{bmatrix} \), so that, \( \mathbb{E}(Z) = \tau \)

\[n^* \quad n-n^* \quad n^* \quad n-n^*\]

and \( \mathbb{E}(Z_0) = \tau_0 \). We now notice that the joint distribution of \( Z \) and \( Z_0 \) may be rewritten as,

\[
\frac{1}{pn^*} \frac{n^*}{(2\pi)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} (Z - \tau)(Z' - \tau') \right\} \, dZ \times \frac{1}{p(n-n^*)} \frac{n-n^*}{(2\pi)^{\frac{1}{2}}} \left| \Sigma \right|^{-\frac{1}{2}} \]

\[
x \, \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} (Z_0 - \tau_0)(Z'_0 - \tau'_0) \right\} \, dZ_0,
\]

so that the distribution of \( Z(p \times n^*) \) is,

\[
(4.5.2) \quad \frac{1}{pn^*} \frac{n^*}{(2\pi)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} (Z - \tau)(Z' - \tau') \right\} \, dZ; \quad -\infty < z_{ij} < \infty ,
\]

\[
\quad \left| \Sigma \right|^{-\frac{1}{2}} \]

where, of course, \( \mathbb{E}(Z) = \tau \) and \( \mathbb{E}(Z)(Z - \tau)(Z' - \tau') \) \( \mathbb{E} = n^* \Sigma \). Hence,
by definition, \( \frac{1}{n^*} \mathbf{Z}\mathbf{Z}' = \frac{1}{n^*} \mathbf{X}\mathbf{L}'\mathbf{L}\mathbf{X}' = \frac{1}{n^*} \mathbf{X}\mathbf{Q}\mathbf{X}' = \frac{1}{\lambda} \mathbf{S}^* \), will have a pseudo-Wishart distribution with \( n^* \) degrees of freedom. Also, if \( \mathbf{E}(\mathbf{X}) \mathbf{Q} \mathbf{E}(\mathbf{X}') = 0 \), then, since \( \lambda \) is a finite positive constant, \( \frac{1}{\lambda} \mathbf{E}(\mathbf{X}) \mathbf{Q} \mathbf{E}(\mathbf{X}') = 0 \), i.e., \( \mathbf{E}(\mathbf{X}) \mathbf{L}' \mathbf{L} \mathbf{E}(\mathbf{X}') = 0 \), or, \( \mathbf{E}(\mathbf{Z}) \mathbf{E}(\mathbf{Z}') = 0 \), which means that the distribution of \( \frac{1}{n^*} \mathbf{Z}\mathbf{Z}' = \frac{1}{\lambda} \mathbf{S}^* \), is central.

Hence the sufficiency of (i) and (ii).

Also, since \( \mathbf{E}(\mathbf{Z}\mathbf{Z}') = \frac{1}{n^*} \mathbf{E}(\mathbf{X}\mathbf{Q}\mathbf{X}') = n^* \Sigma \), therefore,

\[
\lambda \Sigma = \frac{1}{n^*} \mathbf{E}(\mathbf{X}\mathbf{Q}\mathbf{X}') = \mathbf{E}(\mathbf{S}^*)
\]

We may take (4.5.3) as the equation defining \( \lambda \).

Necessity of (i) and (ii):

Let \( \mathbf{X}(p \times n) = \left[ \begin{array}{c} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_p' \\ \mathbf{n} \end{array} \right] \) and \( \mathbf{\zeta}(p \times n) = \left[ \begin{array}{c} \mathbf{l} \\ \mathbf{l} \\ \vdots \\ \mathbf{l} \\ \mathbf{n} \end{array} \right] \).

It can be seen that if \( \mathbf{X}(p \times n) \) has the distribution (4.5.1), then, for any \( i \), the distribution of \( \mathbf{x}_i(n \times 1) \), \( i = 1, 2, \ldots, p \), is

\[
\frac{1}{n} \exp \left( \frac{1}{\mathbf{2}\sigma_{ii}} \left( \mathbf{x}_i' - \mathbf{\zeta}_i \right) \left( \mathbf{x}_i' - \mathbf{\zeta}_i \right)' \right) \mathbf{I} dx_i \]

\[
\left( \mathbf{2}\pi \sigma_{ii} \right)^{\frac{1}{2}}
\]

\( -\infty < \text{elements of } \mathbf{x}_i < \infty \),

where \( \sigma_{ii} \) is the \( i \)th diagonal element of \( \mathbf{E}(p \times p) \).
If \( \frac{1}{n} XX' \) has a pseudo-Wishart distribution with \( n \) degrees of freedom, then the distribution of any diagonal element of \( XX' \), i.e., \( x_{i1}'x_{i1} \), \( i = 1, 2, \ldots, p \), will have a distribution, whose moment generating function is,

\begin{equation}
(4.5.5) \quad M(t,n) = \frac{1}{n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ \sum_{i=1}^{\infty} \frac{t_{i1} x_{i1}' x_{i1}}{2\sigma_{i1}} \right] 7 \exp \left[ - \frac{1}{2\sigma_{i1}} (x_{i1}' - t_{i1}) (x_{i1}' - t_{i1}) \right] \prod_{i=1}^{\infty} dx_{i1}.
\end{equation}

Writing \( y_{i1} = x_{i1} - t_{i1} \), and, then, \( y_{i1} = z_{i1} + \epsilon_{i1} \), where we choose \( \epsilon_{i1} \) so as to make the coefficient of \( z_{i1} \) in the exponent, zero, we have

\begin{equation}
(4.5.6) \quad M(t,n) = \frac{1}{n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ \sum_{i=1}^{\infty} \frac{t_{i1} z_{i1}' z_{i1}}{2\sigma_{i1}} \right] 7 \exp \left[ \frac{t_{i1}^2 \sigma_{i1}^2}{2(1-2t\sigma_{i1})} \right] \prod_{i=1}^{\infty} dz_{i1}.
\end{equation}
\[\exp \left( -\frac{t}{1-2t\sigma_{11}} \right) \sum_{i=1}^{p} \left( \frac{t_{i} - T_{i-1}}{1-2t\sigma_{11}} \right)^{n} \]

It may be noted that, if \( \frac{1}{n}XX' \) has a central pseudo-Wishart distribution, then \( (\xi' = 0) \) (p x p), so that \( (1-2t\xi_{i-1}) = 0 \), (i = 1, 2, ..., p), and the above moment generating function reduces to

\[\frac{1}{(1-2t\sigma_{11})^{2}} \]

which is the moment generating function of \( \sigma_{11}^{2}(n) \).

Now suppose that \( \frac{1}{n^{*}}XQX' = \frac{1}{n^{*}}XFX' \), where \( F(n \times n) = \frac{1}{\lambda}Q(n \times n) \), has a pseudo-Wishart distribution with \( n^{*} \) degrees of freedom, where \( r_{Q}(Q) = r^{*} \). Then, necessarily, the diagonal element, \( X_{1}'F_{1} \), (i = 1, 2, ..., p), should have a distribution, whose moment generating function is \( M(t,n^{*}) \).

But the moment generating function of \( X_{1}'F_{1} \) is,

\[M_{F}(t) = \frac{1}{(2\pi\sigma_{11}^{2})^{n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left( -\frac{1}{2\sigma_{11}^{2}} \left( X_{1}' - T_{1}' \right) \left( X_{1} - T_{1} \right) \right) \]

\[\cdot \exp \left( -2t\sigma_{11}X_{1}'F_{1} \right) \right) \]

Let \( |t| < \left| \frac{1}{2\sigma_{11}c_{0}} \right| = \frac{1}{2\sigma_{11}c_{0}} \), where \( c_{0} \) is the dominant latent root of \( F \), so that \( \int_{I(n)-2t\sigma_{11}^{2}} \) is p.d. Then, after some simplification, we obtain
\[ M_p(t) = \frac{1}{\sqrt{2\pi\sigma_{11}^2}} \exp \left\{ \frac{t^t I(n) - 2t\sigma_{11}P - \mathcal{L}_1}{2\sigma_{11}} \right\} \]

\[ x \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2\sigma_{11}} \left\{ \sum \left[ x^t y_{11}^t I(n) - 2t\sigma_{11}P \right]^{-1} P^{-1} I(n) - 2t\sigma_{11}P \right\}^{-1} \right\} \]

\[ x^t (I(n) - 2t\sigma_{11}P) x \sum \left[ y_{11}^t I(n) - 2t\sigma_{11}P \right]^{-1} P^{-1} \int_{\mathcal{L}_1} dy_{11}, \]

Make the transformations, \( I(n) - 2t\sigma_{11}P = T^{-1} \), and

\[ \sum y_{11}^t I(n) - 2t\sigma_{11}P = T^{-1} T^{-1} = \sum (I(n) - 2t\sigma_{11}P)^{-1} \]

Then, the Jacobian, \( J(y_{11}^t x_{11}) = |T^{-1}| = |T^{-1}|^{-1/2} = |I(n) - 2t\sigma_{11}P|^{-1/2} \), and the moment generating function of \( x_{11}^t P x_{11} \) reduces to,

\[ (4.5.7) \quad M_p(t) = \frac{\exp \left\{ \frac{t^t I(n) - 2t\sigma_{11}P - \mathcal{L}_1}{2\sigma_{11}} \right\}}{|I(n) - 2t\sigma_{11}P|^{1/2}} \]

In order that \( M_p(t) = M(t, n^*) \), a necessary condition is,

\[ (4.5.8) \quad |I(n) - 2t\sigma_{11}P|^{-1/2} = (1 - 2t\sigma_{11})^{-n^*/2} \]

Craig has shown that a NSC for this is that \( P^p = P \), which is equivalent to, \( \frac{Q}{\lambda} \cdot \frac{Q}{\lambda} = \frac{Q}{\lambda} \), or, \( Q^2 = \lambda Q \).

Hence the necessity of (i) in lemma 4.5a. Next, since \( Q(n \times n) \) is symmetric p.s.d. of rank \( n^* \), and \( \lambda \) is a positive constant, therefore, there exists a transformation, \( \int_{\mathcal{L}_1} \).
\[
\frac{Q}{\lambda} = n^{*} \begin{bmatrix}
\frac{T_{1}}{n} \\
T_{2}
\end{bmatrix}
\begin{bmatrix}
T_{1} \\
T_{2}
\end{bmatrix}.
\]

If \( Y(p \times n^{*}) = X(p \times n) \begin{bmatrix}
T_{1} \\
T_{2}
\end{bmatrix} \), then
\[
\frac{1}{\lambda} XQX^{'} = \frac{1}{n^{*}} YY^{'}.
\]

Therefore, if \( \frac{1}{\lambda} XQX^{'} \) has a central pseudo-Wishart distribution, then, \( E(Y)E(Y^{'}) = O(p \times p). \) i.e., \( E(X)QE(X^{'}) = O \) since \( \lambda > 0. \)

Hence the necessity of (ii) in lemma 4.5a.

Next, we shall proceed to prove the necessity and sufficiency of condition (iii) of the lemma.

**Sufficiency of (iii)** We are given that, \( Q = \lambda L^{'}L \), where \( LL^{'} = I(n^{*}) \), and \( r(Q) = n^{*} \leq n \), and, that \( Q_{L} = \lambda L_{L} L_{L}^{'} \), where \( LL_{L}^{'} = I(n_{L}^{*}) \), \( r(Q_{L}) = n_{L}^{*} \leq n \). Condition (iii) is that \( LL_{L}^{'} = O(n^{*} \times n_{L}^{*}) \). Consider a completion \( L_{0}(n-n_{L}^{*}-n^{*} \times n) \) of \( \begin{bmatrix}
L \\
L_{L}
\end{bmatrix} \), such that,
\[
\begin{bmatrix}
n^{*} \\
n_{L}^{*}
\end{bmatrix}
\begin{bmatrix}
L \\
L_{0}
\end{bmatrix}
= \begin{bmatrix}
n-n^{*}-n_{L}^{*} \\
n_{L}^{*}
\end{bmatrix}
\begin{bmatrix}
L \\
L_{0}
\end{bmatrix}
\]

is \( \bot \), so that \( L_{0}L_{0}^{'} = I(n-n^{*}-n_{L}^{*}) \), \( L_{0}L = 0 \) and \( L_{0}L_{L}^{'} = 0 \).

Then,
\[
XX^{'} = XL^{'}LX^{'} + XL^{'}L_{L}X^{'} + XL_{0}L_{0}X^{'}
= XL^{'}LX^{'} + XL^{'}L_{L}X^{'}
+ (p \times n^{*})Z^{'}(n^{*} \times p) + Z_{L}(p \times n_{L}^{*})Z_{L}^{'}(n_{L}^{*} \times p)
+ Z_{0}(p \times n-n_{L}^{*}-n^{*})Z_{0}^{'}(n-n_{L}^{*}-n_{L}^{*} \times p),
\]

where \( Z = XL^{'} \), \( Z_{L} = XL_{L}^{'} \) and \( Z_{0} = XL_{0}^{'} \).

Proceeding exactly as under the proof of the sufficiency of
condition (i), we obtain the joint distribution of \( Z, Z_1 \) and \( Z_0 \) as,

\[
\exp\{-\frac{1}{2} \text{tr} \Sigma^{-1} \left\{ (Z - \mathcal{T})(Z' - \mathcal{T}') + (Z_1 - \mathcal{T}_1)(Z_1' - \mathcal{T}_1') \right\} \}
+ \frac{1}{2} \text{tr} \Sigma^{-1} \left\{ (Z_0 - \mathcal{T}_0)(Z_0' - \mathcal{T}_0') \right\} \text{d}Z \text{d}Z_1 \text{d}Z_0,
\]

where \( E(Z) = \mathcal{T}, E(Z_1) = \mathcal{T}_1, E(Z_0) = \mathcal{T}_0 \), and \( E(Z - \mathcal{T})(Z' - \mathcal{T}') = n*Z \), \( E(Z_1 - \mathcal{T}_1)(Z_1' - \mathcal{T}_1') = n_1^* Z \) and \( E(Z_0 - \mathcal{T}_0)(Z_0' - \mathcal{T}_0') = (n - n^* - n_1^*)Z \).

From (4.5.9), it follows, now, that \( Z(p \times n^*) \) and \( Z_1(p \times n_1^*) \) are distributed independently. Hence \( \frac{1}{n^*} ZZ' = \frac{1}{\lambda} S^*(p \times p) \) and \( \frac{1}{n_1^*} Z_1 Z_1' = \frac{1}{\lambda_1} S_{1*}(p \times p) \) are distributed independently in pseudo-Wishart forms with respective degrees of freedom, \( n^* \) and \( n_1^* \).

Thus, the sufficiency of (iii)

Necessity of (iii): Let \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^*} XQX' = \frac{1}{n_1^*} X_1P_1X_1' \), and \( \frac{1}{n^*} XQX = \frac{1}{n_1^*} X_1P_1X_1' \), where \( P(n \times n) = \frac{1}{\lambda} Q(n \times n) \) and

\[ P_1(n \times n) = \frac{1}{\lambda_1} Q_1(n \times n) \]

be distributed independently in pseudo-Wishart forms with respective degrees of freedom \( n^* \) and \( n_1^* \). Then, necessarily, the diagonal elements, \( X_1^TP_1X_1 \) and \( X_1^TP_1X_1' \), must be distributed independently.

If \( M(t, t_1) \) denotes the moment generating function of the joint distribution of \( X_1^TP_1X_1 \) and \( X_1^TP_1X_1' \), then, as in the proof of the necessity of (i), we can obtain that,
(4.5.10) \[ M(t,t_1) = \mathbb{E}^\gamma \exp \frac{\mathbb{I}^I(t_1 P_1 t_1 P_1)}{2} \sum_{1} \gamma \]

\[ = \exp \int_{\mathbb{I}^I(t_1 P_1 t_1 P_1)(I(n) - 2t_1 \sigma_{11} P_1 - 2t_1 \sigma_{11} P_1)^{-1} \frac{\mathbb{I}^I}{2}} \frac{1}{\sqrt{|I(n) - 2t_1 \sigma_{11} P_1|}} \]

where \( t \) and \( t_1 \) are restricted to values for which \((I(n) - 2t_1 \sigma_{11} P_1 - 2t_1 \sigma_{11} P_1)\) is p.d.

If \( \mathbb{I}^I P \mathbb{I} \) and \( \mathbb{I}^I P_1 \mathbb{I} \) are distributed independently, then we must have,

\[ M(t,t_1) = M_P(t) M_{P_1}(t_1), \text{ where } M_P(t) \text{ is given by (4.5.7)}, \]

\[ = \exp \int_{\mathbb{I}^I(t_1 P_1(I(n) - 2t_1 \sigma_{11} P_1)^{-1} \frac{\mathbb{I}^I}{2}} \frac{1}{\sqrt{|I(n) - 2t_1 \sigma_{11} P_1|}} \]

(4.5.11)

\[ \times \exp \int_{\mathbb{I}^I(t_1 P_1(I(n) - 2t_1 \sigma_{11} P_1)^{-1} \frac{\mathbb{I}^I}{2}} \frac{1}{\sqrt{|I(n) - 2t_1 \sigma_{11} P_1|}} \]

A necessary condition for (4.5.11) is that,

(4.5.12) \[ |I(n) - 2t_1 \sigma_{11} P_1 - 2t_1 \sigma_{11} P_1| = |I(n) - 2t \sigma_{11} P| \times |I(n) - 2t \sigma_{11} P| \]

The authors of (4.5.7) and (4.5.7) have shown that, \( PP_1 = 0(n \times n) \), i.e., \( QQ = 0(n \times n) \), since \( \lambda \) and \( \lambda_1 \) are finite positive constants, is necessary for (4.5.12).

Hence, the necessity of condition (iii) in lemma 4.5a.

This concludes the proof of lemma 4.5a.

It may be noted that, for proving the necessity of conditions
(i) and (iii), we have, essentially, thrown back the multivariate problem on a univariate problem, and the method of solving the univariate problem closely parallels that of (6.7).

**Lemma 4.5b** If \( X(p \times n) \) has the distribution,

\[
(4.5.13) \quad \frac{1}{\text{det} B^{-1/2} \Sigma^{-1} B^{-1/2}} \exp \left( -\frac{1}{2} \text{tr} B^{-1}(X - \mu)(X - \mu') B^{-1} \right) \frac{1}{(2\pi)^{\frac{n(n+1)}{2}}} |B|^{-p/2} |\Sigma|^{-n/2} dX; \infty < x_{i,j} < \infty
\]

where \( B(n \times n) \) is symmetric p.d., then the NSC for \( \frac{1}{\lambda} \Sigma^*(p \times p) = \frac{1}{n^*} XQX' \) to have a central pseudo-Wishart distribution with \( n^* \) degrees of freedom, where \( r(Q) = n^* \) and \( Q(n \times n) \) is symmetric at least p.s.d., are

(a) \( QBQ = \lambda Q \)

(b) \( E(X)QE(X') = 0(p \times p) \).

Further, in order that \( \frac{1}{\lambda_1} \Sigma_1^*(p \times p) = \frac{1}{n_1^*} XQ_1X' \) and \( \frac{1}{\lambda} \Sigma^*(p \times p) = \frac{1}{n\lambda} XQX' \) be distributed independently in pseudo-Wishart forms with respective degrees of freedom, \( n_1^* \) and \( n^* \), we have the NSC,

(c) \( QBQ = 0(n \times n) \).

**Proof:** Since \( B(n \times n) \) is symmetric p.d., therefore, there exists the transformation

\[
B(n \times n) = \tilde{T}(n \times n) \tilde{T}'(n \times n),
\]

so that,
\[ B^{-1} = (\tilde{T}')^{-1} \tilde{T}^{-1} \]

Let \( Y(p \times n) = X(p \times n)(\tilde{T}')^{-1}(n \times n) \), or, \( X(p \times n) = \tilde{Y}' \), and, \( e(p \times n) = t(p \times n)(\tilde{T}')^{-1} \), so that, if \( E(X) = \tilde{t}(p \times n) \), then, \( E(Y) = e(p \times n) \). Also, \( J(X; Y) = |\tilde{T}'| \cdot \frac{p}{|B|^{p/2}} \). Then,

\[ \frac{1}{n^*} XQX' = \frac{1}{n^*} \tilde{Y}' Q \tilde{Y}' \], and the distribution of \( Y(p, n) \) is,

\[ \frac{1}{p n/n} \frac{1}{n^2} \exp \left( -\frac{1}{2} tr \tilde{R}^{-1}(Y - e)(Y' - e') \tilde{S} ; -\infty < y_{ij} < \infty \right) \]

Now, by (i) of lemma 4.5a, the NSC for \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^*} \tilde{T}' Q \tilde{T}' \), to have a pseudo-Wishart distribution with \( n^* \) degrees of freedom, where \( n^* = r(Q) = r(\tilde{T}' Q \tilde{T}) \), is that,

\[ \tilde{T}' Q \tilde{T}' Q \tilde{T}' = \lambda \tilde{T}' Q \tilde{T}' \]

i.e., \( Q \tilde{T}' Q = \lambda Q \), or, \( OBQ = \lambda Q \), which establishes (a).

Again, by (ii) of lemma 4.5a, the NSC for centrality of the distribution of \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^*} XQX' \), is that

\[ E(Y) \tilde{T}' Q \tilde{E}(X') = O(p \times p) \]

i.e., \( E(X)(\tilde{T}')^{-1} \tilde{T}' \tilde{Q} \tilde{T}'^{-1} E(X') = O(p \times p) \), or, \( E(X)Q \tilde{E}(X') = O(p \times p) \), which establishes (b).

Finally, from (iii) of lemma 4.5a, we have that the NSC for independent pseudo-Wishart distributions of, \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^*} \tilde{T}' Q \tilde{T}' \) and \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^*} \tilde{T}' Q \tilde{T}' \), is that, \( \frac{1}{\lambda} S^*(p \times p) = \frac{1}{n^* \lambda} \tilde{T}' Q \tilde{T}' \).
\[ \tilde{T}' \tilde{Q} \tilde{T} \tilde{Q}_1 \tilde{T} = \Theta(n \times n) \]

i.e., \( \tilde{Q} \tilde{T}' \tilde{Q}_1 = O(n \times n) \), or, \( \tilde{Q} \tilde{B} \tilde{Q}_1 = O(n \times n) \), which establishes (\( \gamma \)).

4.6 Further multivariate analogues of results in section 3.5.

Under the general model defined in 4.2, supposing that 
\( \tilde{Q}(n \times n) \) is a symmetric, at least p.s.d. matrix of rank \( n^* \leq n \), 
such that \( \tilde{E}(\tilde{x}) \tilde{Q} \tilde{E}(\tilde{x}') = O(p \times p) \), let us determine \( \frac{1}{n^*} \tilde{E}(\tilde{Q} \tilde{x} \tilde{x}') \).

\[ \Lambda (p \times p) = \frac{1}{n^*} \tilde{E}(\tilde{Q} \tilde{x} \tilde{x}') = \frac{1}{n^*} \tilde{E} \begin{bmatrix} \tilde{x}_1' & \tilde{x}_2' & \cdots & \tilde{x}_p' \end{bmatrix} = \frac{1}{n^*} \begin{bmatrix} \tilde{E}(\tilde{x}_1 \tilde{x}_1), & \cdots & \tilde{E}(\tilde{x}_1 \tilde{x}_p) \\ \tilde{E}(\tilde{x}_2 \tilde{x}_1), & \cdots & \tilde{E}(\tilde{x}_2 \tilde{x}_p) \\ \vdots & \vdots & \vdots \\ \tilde{E}(\tilde{x}_p \tilde{x}_1), & \cdots & \tilde{E}(\tilde{x}_p \tilde{x}_p) \end{bmatrix} \]

From (4.3.2), \( \Sigma^*(pn \times pn) = \tilde{E} \tilde{x} \tilde{x}' \tilde{x} \tilde{x}' - \tilde{E}(\tilde{x}) \tilde{E}(\tilde{x}') \)

\[ = \tilde{E} \begin{bmatrix} \tilde{x}_1' & \tilde{x}_2' & \cdots & \tilde{x}_p' \end{bmatrix} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_p \end{bmatrix} \]
\[
\begin{align*}
- E & \begin{bmatrix}
    x_1' \\
    x_2 \\
    \vdots \\
    x_p 
\end{bmatrix} \begin{bmatrix}
    x_1' x_2' \ldots x_p' 
\end{bmatrix} \quad \text{for } j = \mathcal{E}^* + E \\
\end{align*}
\]

i.e.,
\[
E \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_p 
\end{bmatrix} \begin{bmatrix}
    x_1' x_2' \ldots x_p' 
\end{bmatrix} = E \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_p 
\end{bmatrix} \begin{bmatrix}
    x_1' x_2' \ldots x_p' 
\end{bmatrix} 
\]

\[
= \sum_{i=1}^{k} A_i' A_i x \Sigma_1 + I \cdot x \Sigma + E \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_p 
\end{bmatrix} \begin{bmatrix}
    x_1' x_2' \ldots x_p' 
\end{bmatrix} 
\]

Hence,
\[
E x_j' (n \times 1) \Sigma_k (i \times n) \quad \text{for } j, k = 1, 2, \ldots, p,
\]
where \( \Sigma_1 (p \times p) = \begin{bmatrix}
    \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1p} \\
    \sigma_{12} & \sigma_{22} & \ldots & \sigma_{2p} \\
    \vdots & \vdots & \ddots & \vdots \\
    \sigma_{1p} & \sigma_{2p} & \ldots & \sigma_{pp} 
\end{bmatrix} \)

From lemma 3.5b, it follows, therefore, that
\[
E(x_j' \Sigma_k'^j) = \text{tr} \left\{ \sum_{i=1}^{k} A_i' A_i \sigma_{ij} + \sigma_{jk} I(n) + E(x_j)E(x_k') \right\} 
\]
\[
\begin{align*}
= & \ tr \left\{ \sum_{j=1}^{k} A_j A_j' Q_j Q_j' I(n) \right\} \mathbf{Q} \mathbf{7} + \mathbf{B}(x'_j) \mathbf{Q}(x_j) \\
= & \ tr \left\{ \sum_{j=1}^{k} A_j A_j' Q_j Q_j' I(n) \right\} \mathbf{Q} \mathbf{7},
\end{align*}
\]

since \( \mathbf{E}(x'_j) \mathbf{Q}(x_j) = 0 \), for all \( j, k = 1, 2, \ldots, p \), if \( \mathbf{E}(X) \mathbf{Q}(X') = 0(p \times p) \). Thus,
\[
\mathbf{E}(X \mathbf{Q} X') = \sum_{i=1}^{k} \mathbf{tr}(A_i A_i' Q) \mathbf{Q} + \mathbf{tr}(Q) \mathbf{Q},
\]
so that,
\[
(4.6.1) \quad \Lambda (p \times p) = \frac{1}{n} \mathbf{E}(X \mathbf{Q} X') = \frac{1}{n} \sum_{i=1}^{k} \mathbf{tr}(A_i A_i' Q) \mathbf{Q} + \mathbf{tr}(Q) \mathbf{Q} \mathbf{7}.
\]

For the more restricted set-up, (4.3.3), the result (4.6.1) will reduce to,
\[
(4.6.2) \quad \Lambda (p \times p) = \frac{1}{n} \sum_{i=1}^{k} \sigma_i^2 \mathbf{tr}(A_i A_i' Q) + \mathbf{tr}(Q) \mathbf{7} (p \times p).
\]

Next, for the general multivariate Model II of anova, we shall start from the distribution of \( \mathbf{x}(p n \times 1) \), i.e., all elements of \( X(p n \times n) \), and obtain the distribution of \( X(p n \times n) \) in a form, which is essentially of the type (4.5.13). In doing this, it will be seen that the restricted set-up (4.3.3) arises naturally in the process. The reason for this attempt to obtain the distribution of \( X(p n \times n) \) in, essentially, the same form as (4.5.13), is that, we can then apply the conditions of lemma 4.5b to certain symmetric, at least p.s.d., matrices, obtained under the multivariate Model I of anova, and, as will be shown in section 4.7, a tie-up, analogous to the
univariate one, can be made between the analyses under the multivariate Models I and II.

The distribution of \( \mathbf{x}(p \times 1) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \) is,

\[
\begin{align*}
\frac{1}{(2\pi)^{pn/2} |\Sigma^*|^{1/2}} & \cdot \exp \left[ -\frac{1}{2} \left\{ \sum_{i=1}^{p} (x_1^i - E(x_1^i)) (x_1^i - E(x_1^i))^\top \Sigma^{-1} (x_1^i) \right\} \right] \\
& \cdot \exp \left[ -\frac{1}{2} \left\{ \sum_{i=1}^{p} (x_i^i - E(x_i^i)) (x_i^i - E(x_i^i))^\top \Sigma^{-1} (x_i^i) \right\} \right] \\
& \cdot \exp \left[ -\frac{1}{2} \left\{ \sum_{i=1}^{p} (x_p^i - E(x_p^i)) (x_p^i - E(x_p^i))^\top \Sigma^{-1} (x_p^i) \right\} \right] \\
& \cdot \exp \left[ -\frac{1}{2} \left\{ \sum_{i=1}^{p} (x_i^i - E(x_i^i)) (x_i^i - E(x_i^i))^\top \Sigma^{-1} (x_i^i) \right\} \right]
\end{align*}
\]

where \( E(x_1^i)(n \times 1) \) can be obtained, from (4.3.1), as

\[
A(n \times m) \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{k1} \end{bmatrix} \begin{bmatrix} 1 \\ m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix}
\]

In order to obtain the distribution of \( X(p \times n) \), in, essentially, the same form as (4.5.13), we need to express the exponent in (4.6.3), except for a constant factor \(-\frac{1}{2}\), in the form

\[
\text{tr} M^{-1}_2(X^\top - \xi) M^{-1}_1(X^\top - \xi^\top)
\]
where \( M_1(n \times n) \) and \( M_2(p \times p) \) are symmetric p.d.

**Lemma 4.6a** A NSC for this is that,

\[
\Sigma^*(p n \times p n) = M_1(n \times n) \cdot x M_2(p \times p),
\]

where \( M_1(n \times n) \) and \( M_2(p \times p) \) are symmetric p.d.

**Proof of Sufficiency:** If \( \Sigma^* = M_1 \cdot x M_2 \), then, it is known [28] that, \( \Sigma^*^{-1} = M_1^{-1} \cdot x M_2^{-1} \).

Let

\[
M_1^{-1}(n \times n) = \begin{bmatrix}
m_{11}^{(1)} & m_{12}^{(1)} & \cdots & m_{1n}^{(1)} \\
m_{21}^{(1)} & m_{22}^{(1)} & \cdots & m_{2n}^{(1)} \\
\vdots & \vdots & \ddots & \vdots \\
m_{n1}^{(1)} & m_{n2}^{(1)} & \cdots & m_{nn}^{(1)}
\end{bmatrix}, \quad m_{ij}^{(1)} = m_{ji}^{(1)}
\]

and,

\[
M_2^{-1}(n \times n) = \begin{bmatrix}
m_{11}^{(2)} & m_{12}^{(2)} & \cdots & m_{1p}^{(2)} \\
m_{21}^{(2)} & m_{22}^{(2)} & \cdots & m_{2p}^{(2)} \\
\vdots & \vdots & \ddots & \vdots \\
m_{p1}^{(2)} & m_{p2}^{(2)} & \cdots & m_{pp}^{(2)}
\end{bmatrix}, \quad m_{ij}^{(2)} = m_{ji}^{(2)}
\]

Then,

\[
\sum x_i^{' - E(x_i^{'})}, \ldots, x_p^{' - E(x_p^{'})} \Sigma^*^{-1} \begin{bmatrix}
x_1 - E(x_1) \\
x_2 - E(x_2) \\
\vdots \\
x_p - E(x_p)
\end{bmatrix}
\]
\[ \sum_{i,j=1}^{p} \left( x_1' - E(x_1') \right) M_{1}^{-1} \left( x_j - E(x_j') \right) \]
\[ = \sum_{i,j=1}^{p} m_{ij}^{(2)} \left( x_1' - E(x_1') \right) \left( x_j - E(x_j') \right) \]
\[ = \text{tr} \left( M_{2}^{-1} \begin{bmatrix} x_1' - E(x_1') \\ \vdots \\ x_p' - E(x_p') \end{bmatrix} M_{1}^{-1} \begin{bmatrix} x_1 - E(x_1) \\ \vdots \\ x_p - E(x_p) \end{bmatrix} \right) \]
\[ = \text{tr} \left( M_{2}^{-1} \sum_{1}^{p-1} x - E(x) \right) \left( M_{1}^{-1} \sum_{1}^{p-1} x' - E(x') \right) \]

Hence, the sufficiency.

Proof of necessity: Suppose, next, that

\[ \sum_{1}^{p} x_1' - E(x_1'), \ldots, x_p' - E(x_p') \text{ and } \begin{bmatrix} x_1 - E(x_1) \\ \vdots \\ x_p - E(x_p) \end{bmatrix} \]

Then, writing \( M_{1}^{-1} \) and \( M_{2}^{-1} \) as in (4.6.6), and arguing backwards in the sufficiency part, we obtain that, \( \Sigma^{*} = M_{1}^{-1} \cdot x \cdot M_{2}^{-1} \), so that,

\[ \Sigma^{*} = M_{1} \cdot x \cdot M_{2} \]

Hence, the necessity of the condition (4.6.5).

Next, recalling the form of \( \Sigma^{*}(p \times p) \) from (4.3.2), we can show that, in order that \( \Sigma^{*} \) is of the form \( M_{1} \cdot x \cdot M_{2} \), where \( M_{1}(n \times n) \) and \( M_{2}(p \times p) \) are symmetric p.d., we have the NSC,

\[ \Sigma_{1}(p \times p) = \sigma_{1}^{2} \Sigma(p \times p), \quad (i = 1, 2, \ldots, k) \]
for non-trivial $A^*(p \times p_m)$.

That these conditions on $\Sigma_1(p \times p)$ are sufficient is easily seen. That they are necessary, can be demonstrated as follows, where, for simplicity of argument, we assume $p = 2$.

Suppose $\Sigma^*(p \times p) = \Sigma^*(2n \times 2n)$

$= M_1 \times M_2$,

where $M_1(n \times n)$ and $M_2(2 \times 2)$ are symmetric p.d. Then,

\[ A_1 A_1^t \times \Sigma_1(2 \times 2) + \ldots + A_k A_k^t \times \Sigma_k(2 \times 2) + I(n) \times \Sigma(2 \times 2) = M_1(n \times n) \times M_2(2 \times 2). \]

From this, we obtain the equations,

\[
(16.6.6) \quad A_1 A_1^t \sigma_{11}^{(1)} - c_1 \sigma_{12}^{(1)} - c_2 \sigma_{12}^{(2)} + \ldots + A_k A_k^t \sigma_{11}^{(k)} - c_1 \sigma_{12}^{(k)} - c_2 \sigma_{12}^{(k)} + I(n) \sigma_{11} - c_1 \sigma_{12} = 0
\]

and

\[
A_1 A_1^t \sigma_{11}^{(1)} - c_1 \sigma_{12}^{(1)} - c_2 \sigma_{12}^{(2)} + \ldots + A_k A_k^t \sigma_{11}^{(k)} - c_1 \sigma_{12}^{(k)} - c_2 \sigma_{12}^{(k)} + I(n) \sigma_{11} - c_2 \sigma_{12} = 0,
\]

where $c_1 = (M_2)_{11}/(M_2)_{22}$, $c_2 = (M_2)_{11}/(M_2)_{12}$, $(M_2)_{ij}$ is the $ij$th element $(i,j=1,2)$ of $M_2(2 \times 2)$, and $\Sigma_1(2 \times 2) = \begin{bmatrix} \sigma_{11}^{(1)} & \sigma_{12}^{(1)} \\ \sigma_{12}^{(1)} & \sigma_{12}^{(2)} \end{bmatrix}$, $\Sigma_2(2 \times 2) = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$.
For the equations (1.6.6a) to hold, we must have, either,
\[ \frac{\sigma_{11}}{\sigma_{22}} = \frac{\sigma_{11}}{\sigma_{22}} = \ldots = \frac{\sigma_{11}}{\sigma_{22}} = \sigma_{11}/\sigma_{22} = c_1 \]
\[ \frac{\sigma_{11}}{\sigma_{12}} = \frac{\sigma_{12}}{\sigma_{12}} = \ldots = \frac{\sigma_{11}}{\sigma_{12}} = \sigma_{11}/\sigma_{12} = c_2 \]
or, the matrices \( A_1 A_1', \ldots, A_k A_k' \) must be of the form \( A_1 A_1' = a_1 I(n), \ldots, A_k A_k' = a_k I(n) \). This latter condition on the \( A_i(n \times m_i) \) is too restrictive and unrealistic. Hence, for perfectly general \( A_1(n \times m_1) \), the former conditions hold necessarily, and, it is easily seen that they are equivalent to,
\[ \Sigma_1 = \sigma_1^2 \Sigma \]
for \( i = 1, 2, \ldots, k \), where \( \sigma_1^2 \) are certain (positive) constants.

The proof of the necessity, for general \( p \), follows exactly along the same lines.

It is this NSC that yields our restricted Model II set-up, (4.3.3).

Under (4.3.3), we can identify \( M_1(n \times n) \) and \( M_2(p \times p) \) as follows:
\[ \Sigma^*(pn \times pn) = \sigma_1^2 A_1 A_1' \cdot \Sigma + \ldots + \sigma_k^2 A_k A_k' \cdot \Sigma + I \cdot \Sigma \]
\[ = \sum_1^2 A_1 A_1' + \ldots + \sum_k^2 A_k A_k' + I(n) \cdot \Sigma, \quad \text{cf.} \quad 28.7 \]
\[ = M_1(n \times n) \cdot xM_2(p \times p), \]
where
(4.6.7) \[ M_1(n \times n) = \sigma_1^2 A_1 A_1' + \ldots + \sigma_k^2 A_k A_k' + I(n) \]; \[ M_2(p \times p) = \Sigma(p \times p) \]

Notice that, \( M_2(p \times p) \) is, obviously, symmetric p.d., and \( M_1(n \times n) \),

which is in the form of the sum of a p.d. matrix and a p.s.d. matrix,

is, therefore, symmetric p.d.

Hence, for the restricted set-up (4.3.3), starting from the

distribution (4.6.3), we obtain, for \( X(p \times n) \), the distribution,

(4.6.8) \[
\frac{1}{(2\pi)^{bn/2}} \left( \sum_{i=1}^{k} \sigma_i^2 A_i A_i'^{-1} + I(n) \right)^{-n/2} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1}(X-E(X))(\Sigma^{-1})^{-1} (X'-E(X')) \right\} \\
\int dX,
\]

since \( |\Sigma| = |M_1(n \times n)| \), \( |X M_2(p \times p)| = |M_1|^p |M_2|^n \), by \( \Sigma^2 \).

We notice, of course, that (4.6.8) is, essentially, of the

same form as (4.5.13), with \( B(n \times n) \) being replaced by

\[
\sum_{i=1}^{k} \sigma_i^2 A_i A_i'^{-1} + I(n) \]

Now suppose that we have \( (k + 1) \) symmetric, at least p.s.d.

matrices \( \frac{1}{\lambda_i} S_i(p \times p) = \frac{1}{n_i \lambda_i} X Q_i X' \), where \( Q_i(n \times n) \) is symmetric at

least p.s.d. of rank \( n_i(\leq n) \), for \( i = 0, 1, \ldots, k \), and where

\( X(p \times n) \) has the distribution (4.6.8), under the restricted Model II

of (4.3.3). Then, applying lemma 4.5b, we see that the NSC for the

matrices, \( \frac{1}{\lambda_i} S_i(p \times p), i = 0, 1, \ldots, k \), to have independent central

pseudo-Wishart distributions, with respective degrees of freedom
\( n_0, n_1, \ldots, \text{ and } n_k, \) are

\[
Q_1 \sum_{k=1}^{k} \sigma^2_k A_k^t A_k + I(n) Q_1 = \lambda_1 Q_1, (i = 0, 1, \ldots, k),
\]

\( E(X)Q_1E(X) = 0(p \times p), (i = 0, 1, \ldots, k), \)

and

\[
Q_1 \sum_{k=1}^{k} \sigma^2_k A_k^t A_k + I(n) Q_1 = 0(n \times n), (i \neq j = 0, 1, \ldots, k).
\]

These, of course, give us a set of NSC for the restrictions (i) and (ii), mentioned at the beginning of 4.5, under the restricted Model II of (4.3.3).

Again, if \( X(p \times n) \) has the distribution (4.6.8), then we have,

\[
\lambda_1 \Sigma(p \times p) = \frac{1}{n_1} E \sum_{i} XQ_1X^t = \Lambda_1(p \times p)
\]

\[
= \frac{1}{n_1} \sum_{k=1}^{k} \sigma^2_k \text{tr}(A_k^t A_k Q_1 + trQ_1 J),
\]

using (4.6.2). Hence, we have

\[
\lambda_1 = \frac{1}{n_1} \sum_{k=1}^{k} \sigma^2_k \text{tr}(A_k^t A_k Q_1 + trQ_1 J), \text{ for } i = 0, 1, \ldots, k.
\]

We observe that the conditions (4.6.9)-(4.6.11) are independent of the unknown p.d. matrix, \( \Sigma(p \times p) \), but (4.6.9) and (4.6.11) involve the unknown \( \sigma^2_1, \ldots, \sigma^2_k \). We may require that the \((k + 1)\) matrices

\[
\frac{1}{\lambda_1} S_i(p \times p) \text{ satisfy (4.6.9) and (4.6.11) for all } \sigma^2_1, \ldots, \sigma^2_k.
\]
this case, using (4.6.12), and equating coefficients on the two sides of the equation (4.6.9), we obtain

\[(4.6.13) \quad Q_{1}^1 A^1_k A_k^1 Q_{1}^1 = \frac{1}{n_1} \text{tr}(A^1_k A_k^1 Q_{1}^1) - \mathbf{J}, Q_{1}, Q_{1}^1, k = 1, \ldots, k; Q_{1}^2 = \frac{1}{n_1} \text{tr} Q_{1}^2 - \mathbf{J}, Q_{1} \]

for \( i = 0, 1, \ldots, k \). Similarly, from (4.6.11) we obtain

\[(4.6.11) \quad Q_{1}^1 A^1_k A_k^1 Q_{j}^0 = 0(n \times n), k = 1, 2, \ldots, k; Q_{1} Q_{j}^0 = 0(n \times n) \]

for \( i \neq j = 0, 1, \ldots, k \). It is seen that these conditions, (4.6.13) and (4.6.11), are exactly the same as those obtained for the univariate case, viz., (3.5.2) and (3.5.3). Thus, for a given model matrix \( A(n \times m) \), the same \( Q_{1}^1(n \times n), i = 0, 1, \ldots, k \), which satisfy (3.5.2) and (3.5.3), under the univariate Model II set-up, also satisfy (4.6.13) and (4.6.11), under the multivariate restricted Model II set-up. Also, for given \( Q_{1}^1(n \times n) \), the same model matrix \( A(n \times m) \), which satisfies (3.5.2) and (3.5.3), under the univariate Model II set-up, also satisfies (4.6.13) and (4.6.11), under the multivariate restricted Model II set-up.

4.7 Tie-up between the analyses of the multivariate models, I and II, for a k-way classification.

Under the Model I set-up we can obtain \( k \) matrices due to the \( k \) sets of hypotheses of equality of the row vectors of \( \xi_{i}^1(m_1 \times p) \), \( i = 1, 2, \ldots, k \). Further, we can obtain a matrix due to error. The hypotheses of equality of the row vectors of \( \xi_{i}^1(m_1 \times p) \), may be
written as,

\[ (4.7.1) \quad H_{0i} C_i (\frac{m_i}{m_i-1} \times m) \xi(m \times p) = (m_i-1) \sum C_{11} C_{12} \xi(m \times p), \]

where \( r(A) = r = m - k + 1 \), and \( \sum_{i=1}^{k} m_i = m \),

\[
\begin{bmatrix}
1 & 0 & \ldots & 0 & -1 \\
0 & 1 & \ldots & 0 & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & 1 & -1 \\
m_1 & \ldots & m_i & \ldots & m_k
\end{bmatrix}
\]

\[ = 0(m_i^{-1} \times m) \]

so that \( r(C_i) = (m_i-1) \), for \( i = 1, 2, \ldots, k \), and we notice that the hypotheses matrices \( C_i (\frac{m_i}{m_i-1} \times m) \) here, are the same as those in (3.6.1). The condition of testability will also be same as in the univariate case. If this condition is satisfied, then we can obtain \( H_3 \) the matrices due to the hypotheses (4.7.1) as,

\[ (4.7.2) \quad X A_i (A_i' A_i)^{-1} C_{11} (A_i' A_i)^{-1} C_{11} (A_i' A_i)^{-1} A_i' x' \]

for \( i = 1, 2, \ldots, k \), and these matrices are symmetric, at least p.s.d. of rank, \( \min(p, m_i) \). We also, have the matrix due to error,

\[ (4.7.3) \quad X \sum I(n) - A_i (A_i' A_i)^{-1} A_i' x' \]

which is symmetric p.d.a.e., since we assume in the model that
\[ p \leq n-r, \text{ so that, rank of the matrix due to error} = \min(p, n-r) = p. \]

We shall refer to these matrices as the anova matrices under the multivariate Model I.

Now, under the restricted Model II set-up, in the notation of the preceding sections, suppose we take \( n_0 S_0(p \times p) = XQ_0X' \), as the matrix due to error, \((4.7.3)\), with \( n_0 = (n - r) = (n - m + k - 1) \), and \( n_i S_i(p \times p) = XQ_iX' \), for \( i = 1, 2, \ldots, k \), as the \( k \) matrices \((4.7.2)\), with \( n_i = r(Q_i) = (m_i - 1) \).

It is seen that these \( Q_i(n \times n), i = 0, 1, \ldots, k \), are the same as those of section 3.6, and that, the properties \((3.6.4)\) - \((3.6.6)\) hold here also. Moreover, it can be verified that the condition, \( E(X)Q_iE(X') = 0(p \times p) \), is always satisfied by these \( Q_i \)'s, and, hence, we can obtain that,

\[(4.7.4)\]

\[ \Lambda_i(p \times p) = \frac{1}{n_i} E(XQ_iX') \]

\[ = v_i \Sigma_i + \Sigma, \text{ for the general Model II of 4.2, where } v_i \text{ is defined in (3.6.7)} \]

\[ = (v_i \sigma_i^2 + 1) \Sigma(p \times p), \text{ for the restricted Model II of (4.3.3), } \sum_{i=1}^{k} \]

and,

\[(4.7.5)\]

\[ \Lambda_0(p \times p) = \frac{1}{n_0} E(XQ_0X') = \Sigma(p \times p) \]

Therefore, under the restricted Model II set-up, we have, for the above \( Q_i(n \times n) \), that
(4.7.6) \[ \lambda_i = \nu_i \sigma_i^2 + 1, \quad i = 1, \ldots, k \]
\[ \lambda_0 = 1 \]

For the restricted Model II set-up, using lemma 4.5b, or, equivalently, the conditions (4.6.13) and (4.6.14), it can be verified that \( \frac{1}{\lambda_0} S_0(p \times p) = S_0(p \times p) \), where \( n_0 S_0 \) is the matrix due to error, is always distributed in the central pseudo-Wishart form, (in fact, in the ordinary Wishart form, since \( S_0 \) is symmetric p.d. here), with \( n_0 = (n - m + k - 1) \) degrees of freedom, and that, \( S_0(p \times p) \) is distributed independently of \( \frac{1}{\lambda_i} S_i(p \times p) \), for \( i = 1, 2, \ldots, k \). Also, we obtain that, \( \frac{1}{\lambda_i} S_i(p \times p) \), for \( i = 1, 2, \ldots, k \), are distributed independently in central pseudo-Wishart forms, with respective degrees of freedom \( n_1, \ldots, n_k \), \( (n_i = m_i - 1) \), if, and only if, the same conditions, (3.6.9) and (3.6.10), which were obtained for the univariate case, are satisfied.

The remarks following (3.6.9) in chapter III apply here also.

Recalling our remarks at the end of section 4.6, these conditions for independent central pseudo-Wishart distributions of
\[ \frac{1}{\lambda_i} S_i(p \times p), \quad i = 1, 2, \ldots, k \], are satisfied by the multivariate analogues of the usual univariate complete class designs like, the Randomized Block, the Latin Square and the Factorial designs.

It may be noted, from (4.7.4) and (4.7.5), that we can take
\[ \frac{1}{\nu_i} (S_i - S_0) \] as an unbiased estimate of \( S_i(p \times p) \), for \( i = 1, \ldots, k \).
k, and, \( S_0 \) as an unbiased estimate of \( \Sigma(p \times p) \). We may, therefore, take \( c\sqrt{\frac{1}{\lambda_1}(S_1 - S_0)} \) as estimates of \( c\sqrt{\Sigma_1} \), and \( c(S_0) \) as estimates of \( c(\Sigma) \).

4.8 Tests of hypotheses connected with the multivariate variance components.

The usual hypotheses may be stated as,

\[
(4.8.1) \ H_{01}: \Sigma_1(p \times p) = 0(p \times p) \iff c(\Sigma_1) = 0 \iff \Lambda_1 = \Lambda_0 \iff \lambda_1 = \lambda_0
\]

for the more restricted set-up of (4.3.3), against the alternative,

\[
(4.8.2) \ H_{11}: \lambda_1 > \lambda_0 \text{, again, for the set-up of (4.3.3).}
\]

Suppose that \( \frac{1}{\lambda_0} S_0 = \frac{1}{n_0 \lambda_0} XQ_0 X' \) and \( \frac{1}{\lambda_1} S_1 = \frac{1}{n_1 \lambda_1} XQ_1 X' \), are matrices obtained from the above matrices under Model I, and that, under the restricted Model II, they are distributed independently in central pseudo-Wishart forms with \( n_0 \) and \( n_1 \) degrees of freedom, respectively. Then, \( XQ_0 X' = Y_0(p \times n_0)Y_0'(n_0 \times p) \) and \( XQ_1 X' = Y_1(p \times n_1)Y_1'(n_1 \times p) \), where \( Y_0 \) and \( Y_1 \) have the joint distribution,

\[
\text{const. } \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} \left\{ \frac{1}{\lambda_0} Y_0 Y_0' + \frac{1}{\lambda_1} Y_1 Y_1' \right\} \right\} dy_0 dy_1,
\]

where

\[
E(Y_0 Y_0') = n_0 \lambda_0 \Sigma, \ E(Y_1 Y_1') = n_1 \lambda_1 \Sigma.
\]
Consider \( a'(1 \times p) Y_0(p \times n_0) \) and \( a'(1 \times p) Y_1(p \times n_1) \), for all nonnull \( a(p \times 1) \).

Then,
\[
\frac{1}{n_0} \mathbb{E}(a' Y_0 a) = \lambda_0 a' \Sigma a
\]
and
\[
\frac{1}{n_1} \mathbb{E}(a' Y_1 a) = \lambda_1 a' \Sigma a .
\]

For \( \lambda_1 = \lambda_0 \), against, \( \lambda_1 > \lambda_0 \), we have the test with acceptance region,
\[
\bar{\omega}_{1a}: \frac{a' Y_1 a}{a' Y_0 a} < F_{a}(n_1, n_0) ,
\]
where \( F_{a}(n_1, n_0) \) is the upper \( \alpha \% \) point of the F-distribution with \( n_1 \) and \( n_0 \) degrees of freedom.

Taking, \( \bar{\omega}_1 = \bigcap_{a} \bar{\omega}_{1a} \), as an acceptance region for the hypothesis (4.8.1), we obtain that,
\[
\bar{\omega}_1: c_{\max}(S_1 S_0^{-1}) \leq c_{\beta}(p, n_1, n_0) ,
\]
which is the acceptance region of Roy's test \( \bigcap_{4.3} \) for the hypotheses (4.7.1) under the multivariate Model I.

It must be noted that the above arguments were made, solely, to obtain an acceptance region for the customary hypothesis under (the restricted) Model II, which region, is the same as the acceptance region for the customary hypothesis under Model I. The use of the union-intersection principle to obtain \( \bar{\omega}_1 \) from \( \bar{\omega}_{1a} \) is rather arti-
ficial, since we do not have \( H_{0i} \) itself as an intersection of hypotheses \( H_{0i,a} \).

4.9 **Simultaneous confidence bounds on** \( \sigma^2_1, \ldots, \sigma^2_k \) **and** \( c(\Sigma) \).

In this section, we shall consider the multivariate restricted Model II for \( k \)-way classification designs, which have anova matrices under Model I that satisfy the conditions for independent central pseudo-Wishart distributions.

If \( \frac{1}{\lambda_i} S_i (p \times p) = \frac{1}{n_i \lambda_i} XQ_i X' \), for \( i = 0, 1, \ldots, k \), have independent central pseudo-Wishart distributions with respective degrees of freedom \( n_0, n_1, \ldots, n_k \), then, by definition, we have that,

\[ \frac{1}{n_i \lambda_i} XQ_i X' = \frac{1}{n_i} Y_i(p \times n_i) Y_i'(n_i \times p) \]

for \( i = 0, 1, \ldots, k \), where the joint distribution of \( Y_0, \ldots, Y_k \) is,

\[
\frac{1}{p(n-1)} \frac{1}{(n-1)} \exp \left[ - \frac{1}{2} tr \left( \Sigma^{-1} \{ Y_0 Y_0' + Y_1 Y_1' + \cdots + Y_k Y_k' \} \right) \right]
\]

\[
\times dY_0 dY_1 \cdots dY_k ;
\]

\(-\infty < \text{elements of all} \ Y_i < \infty \),

where \( E(Y_i Y_i') = n_i \Sigma (p \times p) \) for \( i = 0, 1, \ldots, k \), and \( \sum_{i=0}^{k} n_i = (n - 1) \).

It is well known that there exists an orthogonal matrix, \( \Gamma (p \times p) \), such that,

\[ \Sigma (p \times p) = \Gamma (p \times p) D (p \times p) \Gamma (p \times p) \]
where the \( p \) elements of \( D_\gamma \) are the \( p \) non-zero (positive) characteristic roots of \( \Sigma (p \times p) \).

Make the transformations,

\[
(4.9.2) \quad D_\gamma \rightarrow (p \times p)\Gamma(p \times p)Y_i(p \times n_i)=Z_i(p \times n_i), i=0,1,\ldots,k,
\]

so that the Jacobian is,

\[
(4.9.3) \quad J(Y_0,\ldots,Y_k;Z_0,\ldots,Z_k) = \prod_{i=0}^{k} \frac{n_i^{1/2}}{\Gamma(n_i/2)} |D_\gamma|^{-n_i/2} \quad = |Z_i|^{(n_i-1)/2}, \text{since } \sum_{i=0}^{k} n_i = (n_i-1) \cdot \]

Therefore, the joint distribution of \( Z_0, Z_1, \ldots, Z_k \) is,

\[
(4.9.4) \quad \frac{1}{p! (n_i-1)!} \exp \left\{-\frac{1}{2} tr \left\{ Z_0 Z_0^T + Z_1 Z_1^T + \ldots + Z_k Z_k^T \right\} \right\} \prod_{i=0}^{k} dZ_i ; \quad -\infty < \text{elements of all } Z_i < \infty .
\]

From (4.9.4), it can be seen that, we can obtain constants \( \mu_{11}(\alpha_i), \mu_{12}(\alpha_i) \) which will depend on \( p \) and \( n_i \), for \( i=0, 1, \ldots, k \), such that the statement,

\[
(4.9.5) \quad \mu_{11}(\alpha_i) \leq c_{\min}(Z_i Z_i^T) \leq c_{\max}(Z_i Z_i^T) \leq \mu_{12}(\alpha_i)
\]

has probability \( (1 - \alpha) \), and, the probability of statements like (4.9.5) holding for all \( i=0, 1, \ldots, k \), simultaneously, is \( (1 - \alpha) \).

\[
\prod_{i=0}^{k} (1 - \alpha_i) . \text{ In (4.9.5), } c_{\min}(Z_i Z_i^T) \text{ and } c_{\max}(Z_i Z_i^T) \text{ stand for the smallest and largest non-zero (positive) characteristic roots, re-
\]
spectively, of $Z_i'Z_i$.

For $i = 0$, we note that $Z_i'Z_i = \frac{n_0}{\lambda_0} D_1/\gamma$ $R S_0' D_1/\gamma$, where $n_0 S_0(p \times p)$, the matrix due to error, is symmetric p.d. and $\lambda_0 = 1$.

Hence, proceeding from (4.9.5) with $i = 0$, exactly as in section 1.3, we shall have,

\( (4.9.6) \quad \frac{n_0}{\mu_{01}} c_{\max}(S_0) \geq c_{\max}(\Sigma) \geq c_{\min}(\Xi) \geq \frac{n_0}{\mu_{02}} c_{\min}(S_0) \),

with probability $\geq (1 - c_0)$.

Next, for any $i = 1, 2, \ldots, k$, we note that (4.9.5) is equivalent to,

\[ \mu_{11}(a_1) \leq c_{\min}(D_1/\gamma_R Y_i'Y_i' \gamma) \leq c_{\max}(D_1/\gamma_R Y_i'Y_i' \gamma) \mu_{12}(a_1) \]

or, to

\( (4.9.7) \quad \mu_{11}(a_1) \leq \frac{n_i}{\lambda_i} c_{\min}(D_1/\gamma R S_i' \gamma) \leq \frac{n_i}{\lambda_i} c_{\max}(D_1/\gamma R S_i' \gamma) \mu_{12}(a_1) \)

But,

\[ c_{\min}(R S_i' D_1/\gamma) \leq \frac{c_{\min}(R S_i' \gamma)}{c_{\min}(D_1/\gamma)} = \frac{c_{\min}(S_i)}{c_{\min}(\Sigma)} \],

and

\[ c_{\max}(R S_i' D_1/\gamma) \geq \frac{c_{\max}(R S_i' \gamma)}{c_{\max}(D_1/\gamma)} = \frac{c_{\max}(S_i)}{c_{\max}(\Sigma)} \],

so that, (4.9.7) implies the pair of statements,

\( (4.9.8) \quad \frac{\mu_{11}(a_1)}{n_i} \leq \frac{1}{\lambda_i} \frac{c_{\min}(S_i)}{c_{\min}(\Sigma)} \), and, $\frac{1}{\lambda_i} \frac{c_{\max}(S_i)}{c_{\max}(\Sigma)} \leq \frac{\mu_{12}(a_1)}{n_i}$.\]
From (4.9.8) we have,

\[ (4.9.9) \quad \frac{n_1}{\mu_{12}(a_1)} \frac{c_{\max}(S_1)}{c_{\max}(\Sigma)} \leq \lambda_1 \leq \frac{n_1}{\mu_{11}(a_1)} \frac{c_{\min}(S_1)}{c_{\min}(\Sigma)}, \]

which, therefore, has a probability \( \geq (1 - a_1) \). We have a similar result for \( i = 1, 2, \ldots, k \).

Taking the statement (4.9.6) together with all statements (4.9.9), for \( i = 1, 2, \ldots, k \), we obtain the simultaneous statements,

\[ (4.9.10) \quad \frac{n_0}{\mu_{02}} \frac{c_{\min}(S_0)}{c_{\min}(\Sigma)} \leq \frac{n_0}{\mu_{01}} \frac{c_{\max}(S_0)}{c_{\max}(\Sigma)} \]

\[ \quad \frac{n_1}{\mu_{12}} \frac{c_{\max}(S_1)}{c_{\max}(\Sigma)} \leq \lambda_1 \leq \frac{n_1}{\mu_{11}} \frac{c_{\min}(S_1)}{c_{\min}(\Sigma)} \]

\[ \quad \ldots \quad \ldots \quad \ldots \quad \ldots \]

\[ \quad \frac{n_k}{\mu_{k2}} \frac{c_{\max}(S_k)}{c_{\max}(\Sigma)} \leq \lambda_k \leq \frac{n_k}{\mu_{k1}} \frac{c_{\min}(S_k)}{c_{\min}(\Sigma)} \]

with a joint probability \( \geq (1 - a) = \prod_{i=0}^{k} (1 - a_i) \).

From (4.9.10), by writing \( \lambda_i = \gamma_i \sigma_i^2 + 1, (i = 1, \ldots, k) \),

and using the first of the statements (4.9.10), we obtain the further statements, implied by (4.9.10),

\[ (4.9.11) \quad \frac{n_0}{\mu_{02}} \frac{c_{\min}(S_0)}{c_{\min}(\Sigma)} \leq \frac{n_0}{\mu_{01}} \frac{c_{\max}(S_0)}{c_{\max}(\Sigma)} \]
\[
\frac{1}{\nu_1} \left[ \frac{n_1}{n_0} \frac{\mu_{01}}{\mu_{12}} \frac{c_{\text{max}}(S_1)}{c_{\text{max}}(S_0)} - 1 \right] \leq \sigma_1^2 \leq \frac{1}{\nu_1} \left[ \frac{n_1}{n_0} \frac{\mu_{02}}{\mu_{11}} \frac{c_{\text{min}}(S_1)}{c_{\text{min}}(S_0)} - 1 \right]
\]

\[
\vdots \quad \vdots \quad \vdots
\]

\[
\frac{1}{\nu_k} \left[ \frac{n_k}{n_0} \frac{\mu_{01}}{\mu_{k2}} \frac{c_{\text{max}}(S_k)}{c_{\text{max}}(S_0)} - 1 \right] \leq \sigma_k^2 \leq \frac{1}{\nu_k} \left[ \frac{n_k}{n_0} \frac{\mu_{02}}{\mu_{k1}} \frac{c_{\text{min}}(S_k)}{c_{\text{min}}(S_0)} - 1 \right]
\]

which, therefore, are a set of simultaneous confidence bounds on all \( \sigma(\zeta) \), \( \sigma_1^2 \), \( \ldots \), and \( \sigma_k^2 \), with a joint confidence coefficient \( \geq (1-\alpha) \), for a preassigned \( \alpha \).

### 4.10

**Alternative set of confidence intervals for the individual**

\( \sigma_1^2 \), \( \sigma_2^2 \), \( \ldots \), and \( \sigma_k^2 \).

In this section we shall use certain results contained in \( \Gamma^{35.7} \), which are extensions of earlier results, contained in \( \Gamma^{2.7} \) and \( \Gamma^{49.7} \).

Suppose \( Y_0(p \times n_0) \) and \( Y_1(p \times n_1) \), where \( p \leq n_0 \) but may be \( \geq n_1 \), such that \( r(Y_0Y_0') = p \) and \( r(Y_1Y_1') = \min(p, n_1) \), have the joint distribution,

\[
(4.10.1) \quad \frac{1}{p(n_0+n_1)} \frac{1}{n_0+n_1} \exp \left\{ -\frac{1}{2} \text{tr} Z^{-1} \left\{ Y_0Y_0' + Y_1Y_1' \right\} Z dY_0 dY_1 \right\}
\]

\[
(2\pi) \frac{1}{2p \times p} \frac{1}{2n_0+n_1} \frac{1}{2}
\]

where \( Z(p \times p) \) is a symmetric p.d. matrix, and \( E(Y_0Y_0') = n_0 Z \),

\( E(Y_1Y_1') = n_1 Z \). Let,

\[
(4.10.2) \quad \Lambda_1 = \frac{|Y_0Y_0'|}{|Y_0Y_0'| + |Y_1Y_1'|}, \quad \text{and}, \quad m_1 = \frac{p+n_1+1}{2}
\]
Then it has been shown, \( L^{\text{35.7}} \), that, for large \( m_1 \), we may take, \(-m_1 \log_e \Lambda_1\), to be distributed as a central \( \chi^2 \)-variate with \( pn_1 \) degrees of freedom. Hence, we can find \( \chi^2_{1a_1} \) and \( \chi^2_{2a_1} \) such that the statement,

\[
(4.10.3) \quad \chi^2_{1a_1} \leq -m_1 \log_e \Lambda_1 \leq \chi^2_{2a_1},
\]

or, equivalently,

\[
(4.10.4) \quad \mu_{1a_1} = \exp(-\frac{1}{m_1} \chi^2_{1a_1}) \leq \Lambda_1 \leq \exp(-\frac{1}{m_1} \chi^2_{2a_1}) = \mu_{2a_1},
\]

will have a probability \( (1 - a_1) \), for a preassigned \( a_1 \).

In the multivariate restricted Model II set-up, when the anova matrices under Model I satisfy the conditions for independent central pseudo-Wishart distributions, we have \( \frac{n_0S_0}{\lambda_0} = Z_0^2Z_0' \), (say), and

\[
\frac{n_1S_1}{\lambda_1} = Z_1^2Z_1', \quad \text{(say)},
\]

where \( \lambda_0 = 1 \) and \( \lambda_1 = \nu_1 \sigma_1^2 + 1 \), and the joint distribution of \( Z_0(p \times n_0) \) and \( Z_1(p \times n_1) \) is,

\[
(4.10.5) \quad \frac{1}{p(n_0+n_1)} \frac{n_0+n_1}{2} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} \left( Z_0^2Z_0' + Z_1^2Z_1' \right) \right\} \frac{1}{2\pi} |\Sigma|^{-\frac{1}{2}} \int dZ_0dZ_1,
\]

where \( E(Z_0^2) = n_0 \Sigma(p \times p) \) and \( E(Z_1^2) = \frac{n_1}{\lambda_1} E(S_1) = n_1 \Sigma \). By analogy with (4.10.1) - (4.10.4), therefore, we can, for large \( m_1 \), find values, \( \mu_{1a_1} \) and \( \mu_{2a_1} \), such that the statement,
\[ \mu_{1a_1} \leq \frac{|z_{0}^{*}z_{0}^{t}|}{|z_{0}^{*}z_{0}^{t} + z_{i}^{*}z_{i}|} \leq \mu_{2a_1} \]

will have probability \(1 - a_1\), for a preassigned \(a_1\). Now, (4.10.6)\

\[ \mu_{1a_1} \leq \frac{|n_{0}^{*}n_{0}|}{|n_{0}^{*}n_{0} + \frac{n_{1}^{*}s_{i}}{\lambda_{1}}|} \leq \mu_{2a_1} , \]

or,

\[ \frac{1}{\mu_{1a_1}} \geq \frac{|n_{0}^{*}n_{0} + \frac{n_{1}^{*}s_{i}}{\lambda_{1}}|}{|n_{0}^{*}n_{0}|} \geq \frac{1}{\mu_{2a_1}} , \]

or,

\[ (4.10.7) \quad \frac{1}{\mu_{1a_1}} \geq |t_{1} s_{i} + I(p)| \geq \frac{1}{\mu_{2a_1}} , \quad \text{(where } t_{1} = \frac{n_{1}}{n_{0}^{*}n_{0} > 0}) , \]

since \(S_{0}(p \times p)\) is symmetric p.d. so that \(S_{0}^{-1}\) exists. Again, (4.10.7) is equivalent to,

\[ (4.10.8) \quad \frac{1}{\mu_{1a_1}} \geq (t_{1})^{s_{i}} + |t_{1} s_{i} + I(p)| \geq \frac{1}{\mu_{2a_1}} \]

where \(s_{i} = r(S_{1}S_{0}^{-1}) = r(S_{1}) = \min(p, n_{1})\), and, \(\text{tr}_s(A)\) denotes the sum of all \(s\)th order principal minors of \(A\). Since, \(S_{0}^{-1}\) is symmetric p.d. and \(S_{1}\) is symmetric at least p.s.d. of rank \(s_{1} = \min(p, n_{1})\),
therefore, we have the transformations,

\((4.10.9)\) \(s_0^{-1}(p \times p) = \tilde{T}(p \times p)\tilde{T}'(p \times p), s_1(p \times p) = s_1 \left[ \begin{array}{c} \tilde{T}_1 \\ \tilde{T}'_1 \end{array} \right] \tilde{T}'_2 \tilde{T}_2,\)

so that,

\[
\text{tr}_{s_2} \left( s_1^{-1} \right) = \text{tr}_{s_2} \left( \begin{array}{c} \tilde{T}_1 \\ \tilde{T}'_1 \end{array} \right) \left( \begin{array}{c} \tilde{T}'_1 \tilde{T}_2 \\ \tilde{T}'_2 \end{array} \right) \tilde{T}'_1 \tilde{T}_1.
\]

\[
= \text{tr}_{s_2} \left( \tilde{T}'_1 \right) \left( \begin{array}{c} \tilde{T}'_1 \\ \tilde{T}_1 \end{array} \right) \left( \begin{array}{c} \tilde{T}'_1 \tilde{T}_2 \\ \tilde{T}'_2 \end{array} \right) \tilde{T}'_1 \tilde{T}_1.
\]

\[
= \text{tr}_{s_2} \left( B(p \times s_1) B'(s_1 \times p) \right),
\]

where \( B(p \times s_1) = \tilde{T}' \left[ \begin{array}{c} \tilde{T}_1 \\ \tilde{T}_2 \end{array} \right] \)

\[
> 0 \text{ for all } s \leq s_1,
\]

since \(BB'\) is symmetric, at least p.s.d. of rank \(s_1\). Hence, all the coefficients of powers of \(\tilde{T}_1\), in the middle part of \((4.10.8)\), are real and positive.

Next, it can be shown that, in order that the bounds in \((4.10.8)\) be non-trivial, we should have \(\frac{1}{\mu_{2\alpha_1}} > 1\). The argument is as follows:

\[(4.10.10)\)

\[
\left| \tilde{T}_1 s_1 s_0^{-1} + I(p) \right| = \prod_{i=1}^{s_1} (1 + c_i),
\]
where \( c_1, \ldots, c_{s_i} \) (some of which may be equal) are the non-zero (in fact, positive) characteristic roots of \( t_1 s_i s_0^{-1} \). That \( c_1, \ldots, c_{s_i} \) are positive, follows from,

\[
c_{1}s_1s_0^{-1} = c(t_1 \begin{bmatrix} t_1' & \sqrt{T_1T_2'} \end{bmatrix}^T ) = c(t_1 \begin{bmatrix} t_1' & \sqrt{T_1T_2'} \end{bmatrix}^T ) = c(t_1BB') .
\]

Hence, from (4.10.10), \( |t_1 s_1 s_0^{-1} + I(p)| > 1 \), so that, \( \frac{1}{\mu_{2\alpha_1}} > 1 \).

Considering now the equality signs in (4.10.7), we obtain the equations,

\[
(4.10.11) \quad \left( t_1 \right)^{s_1} \mathrm{tr}_{s_1} (s_1 s_0^{-1}) + \ldots + t_1 \mathrm{tr}(s_1 s_0^{-1}) - \left( \frac{1}{\mu_{2\alpha_1}} - 1 \right) = 0
\]

\[
(4.10.12) \quad \left( t_1 \right)^{s_1} \mathrm{tr}_{s_1} (s_1 s_0^{-1}) + \ldots + t_1 \mathrm{tr}(s_1 s_0^{-1}) - \left( \frac{1}{\mu_{1\alpha_1}} - 1 \right) = 0
\]

By Descartes' rule of signs, it can be seen that, since the number of variations in sign is 1, both these equations have at the most one positive real root. Also, since the number of variations is odd, therefore, they have at least one positive root. Let these positive roots of (4.10.11) and (4.10.12) be denoted by \( \theta_{2\alpha_1} \) and \( \theta_{1\alpha_1} \), respectively. Then, it is seen that (4.10.7) is equivalent to,
(4.10.13) \[ e_{\lambda_1} > \frac{t_1}{n_0 \lambda_1} > e_{2\lambda_1} \]

with a probability \((1 - \alpha_1)\).

Further, \((4.10.13) \iff \)

\[ e_{\lambda_1} > \frac{n_1}{n_0 \lambda_1} > e_{2\lambda_1} \]

or, to

\((4.10.14) \)

\[ \frac{n_1}{n_0 \lambda_1} \leq \lambda_1 \leq \frac{n_1}{n_0 \lambda_2} \]

or, to

\((4.10.15) \)

\[ \frac{1}{\nu_1} \left[ \frac{n_1}{n_0 \lambda_1} - 1 \right] \leq \sigma_1^2 \leq \frac{1}{\nu_1} \left[ \frac{n_1}{n_0 \lambda_2} - 1 \right] \]

which is a confidence interval statement for \(\sigma_1^2\) with a confidence coefficient = \((1 - \alpha_1)\), for a preassigned \(\alpha_1\).

Thus, starting from \((4.10.6)\), where if we take \(Z_0 Z_0' = \frac{n_0 s_0}{\lambda_0}\)

with \(n_0 s_0\) as the matrix due to error (under Model I), so that, \(Z_0 Z_0'\) is distributed independently of \(Z_1 Z_1'\), the other anova matrices under the multivariate Model I, we can obtain confidence intervals like \((4.10.15)\) on \(\sigma_1^2, \ldots, \sigma_k^2\) separately. However, due to the complexity of the distribution problem involved, we cannot obtain simultaneous confidence bounds on \(\sigma_1^2, \ldots, \sigma_k^2\) by this method.

4.11 **Concluding remarks.**

It may be repeated that, in this chapter, we have, in the main, considered a restricted multivariate Model II. The problems
that arise in the perfectly general model of 4.2, without the restrictions (4.3.3), are still to be solved.

Also, under the restricted Model II, the remarks, on incomplete designs, in the concluding section of chapter III, apply here too, with the change that, here we shall have anova matrices instead of anova sums of squares, and the distributions will be central pseudo-Wishart instead of $\chi^2$. 
CHAPTER V

FACTOR ANALYSIS

5.1 Introduction.

In this chapter, factor analysis is looked at from the point of view of analytic statistics, inasmuch as statistical tools are used by factor analysts. Several authors, 1, 21, 23-26, 35, 46-7, have attacked the statistical problems in factor analysis, using different methods and, usually, arriving at different results. There remains, however, a great deal of ambiguity about certain assumptions that are made, and these need to be clarified. A recent Ph.D. thesis, 21-7, has done a good service by stating some of the problems in factor analysis clearly in statistical terms. Part of the present chapter is concerned with clarifying certain issues, which arise in the use of statistical tools in factor analysis. Again, little, or, no work has been done, in this field, from the standpoint of confidence intervals for the parameters in the model. The interest, so far, has been in point estimation (usually, by the method of maximum likelihood), and tests of hypotheses, using the likelihood ratio criterion. This chapter presents certain results in confidence interval estimation and the associated tests of hypotheses. The results on confidence intervals are based on certain earlier results 38, 43, 44-7. Finally, the relevant distribution problem, which has also occurred already in sections 1.3, 2.9 and 4.9, is also studied in this chapter.
5.2 The Factor Analysis "Model".

The interest in factor analysis, stated in statistical terminology, arises from, what is usually stated as, the possibility of adequately describing a p-dimensional complex of random variables by \( q < p \) other variables. Cf. 217.

Assuming linearity, the usual "model" is postulated as,

\[
\mathbf{x}(p \times 1) = \beta(p \times q) \mathbf{y}(q \times 1) + \mathbf{\eta}(p \times 1),
\]

where, \( \mathbf{x}(p \times 1) \) is a p-dimensional random variable which is observed on the various individuals of the sample, \( \beta(p \times q) \) is the matrix of "factor loadings" of rank \( q \), and \( \mathbf{\eta}(p \times 1) \) is a vector whose elements are in the nature of random errors, so that, they may be assumed (i) to have zero expectations, (ii) to be mutually uncorrelated, and (iii) uncorrelated with the elements of \( \mathbf{y}(q \times 1) \). That is, \( \mathbb{E}(\mathbf{\eta}) = \mathbf{0}(p \times 1) \); \( \mathbb{E}\mathbf{\eta}'\mathbf{\eta} = 0(p \times q) \). The elements of \( \beta(p \times q) \) are assumed to be fixed parameters, which are to be estimated. It is about the vector \( \mathbf{y}(q \times 1) \) that there is a great deal of divergence, among authors, in their assumptions. Some \( \text{247} \) assume that \( \mathbf{y}(q \times 1) \) is a vector of fixed parameters, while others \( \text{357} \) assume that it is a vector of stochastic variates. Statistically, it seems that the latter is the correct assumption to make regarding the "factors" \( \mathbf{y}(q \times 1) \). This is due to
the fact that, if \( \beta(p \times q) \) and \( y(q \times l) \) are both parametric, then the effective (i.e., estimable and testable) parameters are at the most, the \( p \) elements of \( \Theta(p \times l) = \beta(p \times q) y(q \times l) \), and not the individual elements of \( \beta(p \times q) \) and/or \( y(q \times l) \). So it appears that, statistically speaking, the correct assumption to make, regarding the factors \( y(q \times l) \), is that they are stochastic and non-observable. This assumption does not exclude the problem of "estimating" or "predicting" the factor values \( y \), which problem has been considered by several authors. \( L_{17}, L_{217}, L_{61} \). Our assumption means that the "estimation" or, "prediction" of \( y(q \times l) \) will be in the sense of regression, i.e., we use the regression of \( y(q \times l) \) on \( x(p \times l) \), which gives \( E(y|x) \), as a prediction equation, and not in the sense of point estimation of parameters.

Hence, the admissible "model" may be stated as,

\[
(5.2.1) \quad x(p \times l) = \beta(p \times q)y(q \times l) + \eta(p \times l),
\]

where, (i) \( x(p \times l) \) is a vector of stochastic, observable variates; (ii) \( \beta(p \times q) \) is a matrix of rank \( q(\leq p) \) whose elements (factor loadings) are fixed parameters which are to be estimated; (iii) \( y(q \times l) \) is a vector of stochastic, non-observable variates; and, (iv) \( \eta(p \times l) \) is a vector, whose elements are in the nature of random errors, such that, \( E(\eta) = 0(p \times l) \), \( E(\eta') = \sigma^2 (p \times p) \) and \( E(\eta y) = 0(p \times q) \).

A distinction must be made between the factor analysis "model", ...
(5.2.1), and the Models I and II of anova. In both these latter models the basic model (which is not tested for) may be stated in terms of a vector equation like (5.2.1). However, in both the models of anova, the counterpart of the matrix $\mathbf{B}(p \times q)$ of (5.2.1), which, in the terminology of chapters III and IV, is the model matrix, is assumed to be known for any given design. Also, the elements of the counterpart of $\mathbf{Y}(p \times 1)$ of (5.2.1) are assumed to have the same variance in the models of anova.

Under the "model" (5.2.1) we, of course, have

$$(5.2.2) \quad \mathbf{X}(p \times p) = \mathbf{E}(x^2) = \mathbf{B}(p \times q) \mathbf{E}(y'y) \mathbf{B}'(q \times p) + \mathbf{D}(p \times p) \gamma^2(p \times p)$$

$$= \mathbf{B}' + \mathbf{D}, \quad \text{if } \mathbf{E}(y'y) = \mathbf{I(q)},$$

i.e., if the factors are standardized orthogonal factors. (5.2.2) may itself be taken as equivalent to the "model" (5.2.1) for standardized orthogonal factors.

This so-called "model" of factor analysis is not to be taken in the same sense as other models in statistics, which are assumed as basic and not tested for statistically. In fact, in factor analysis, one of the chief interests is to test this hypothesis or postulated model. From this standpoint, (5.2.2) is not an assumed model but a statistical hypothesis to be tested against an alternative.

For maximum likelihood estimation and for tests of hypotheses, it is usually assumed that $\mathbf{X}(p \times 1)$ is $p$-variate normal. It is well
known that $m(<n)$ independent linear functions of $n$ variables which have an $n$-variety normal distribution, have an $m$-variety normal distribution. It is also true that, if $m$ independent linear functions of $n$ independently distributed variables have an $m$-variety normal distribution, then the $n$ variables themselves have normal distributions. \[ \text{Cf. 9, or, 10 p. 112.} \] Using this latter result, we see, from (5.2.1), that if $\mathbf{x}(p \times 1)$ is assumed to be $p$-variety normal, and further that the "factors", $\mathbf{y}(q \times 1)$, are orthogonal and distributed independently of $\mathbf{z}(p \times 1)$, then $\mathbf{y}(q \times 1)$ and $\mathbf{z}(p \times 1)$ will also be normal. From now on we assume this in our discussion.

We note that, since $\Sigma(p \times p)$ is p.d., and $r(\beta) = r(\beta) = q(<p)$, therefore, no element $\gamma_{i}^2 (i = 1, \ldots, p)$ of $D_{\gamma}(p \times p)$ can be zero.

5.3 Clarification of relationship between $p$ and $q$.

As observed by Kendall, the model (5.2.2), i.e., $\Sigma(p \times p) = \beta(p \times q) \beta'(q \times p) + D_{\gamma}(p \times p)$ where $q < p$, is under determined. As it was mentioned earlier, the physical interest of the problem lies in the possibility of $q(<p)$ "factors" adequately describing the correlation structure among the original $p$ variables. For the statistical purposes of unambiguous estimation of the parameters and testing the postulated model, however, we shall see that $q$ should be related to $p$ by a more stringent inequality than $q < p$. This inequality will now be obtained.

Consider, $\Sigma(p \times p) = \beta(p \times q) \beta'(q \times p) + D_{\gamma}(p \times p)$, where $r(\beta) = q < p$. Then, it is known that $\beta(p \times q)$ can be written
in the form \( q \begin{bmatrix} \hat{T}_1 \\ T_2 \end{bmatrix} \), where \( I(q \times q) \) is \( I \), so that, \( (n-q) \begin{bmatrix} \hat{T}_1 \\ T_2 \end{bmatrix} q \)

we have

\[
\beta \beta' = q \begin{bmatrix} \hat{T}_1 \\ T_2 \end{bmatrix} \begin{bmatrix} \sqrt{\hat{T}'_1} & T'_2 \\ & \gamma \end{bmatrix}, \begin{bmatrix} \hat{T}'_1 \\ T_2 \end{bmatrix} q (p-q)
\]

and,

\[
(5.3.2) \quad \Sigma(p \times p) = q \begin{bmatrix} \hat{T}_1 \\ T_2 \end{bmatrix} \begin{bmatrix} \sqrt{\hat{T}'_1} & T'_2 \\ & \gamma \end{bmatrix} D_{\gamma}^2(p \times p) q (p-q)
\]

The number of independent elements on the left side of (5.3.2) is seen to be \( \frac{p(p+1)}{2} \), and the number of independent elements on the right side is seen to be \( \sqrt{\frac{q(q+1)}{2} + (p-q)q} \). It is known that unique estimates of all the elements of \( \Sigma(p \times p) \) exist, and, therefore, for unique estimability, we have

\[
\frac{p(p+1)}{2} \geq \frac{q(q+1)}{2} + (p-q)q + p
\]

or,

\[
(5.3.3) \quad p \leq \frac{(p-q)(p-q+1)}{2}
\]

This inequality, when it is a strict inequality, also means, that the model specifies certain relationships among parameters so that the number of independent parameters to be estimated is reduced. This,
of course, is the nature of a statistical hypothesis.

Now, \((5.3.3)\) is easily seen to be equivalent to,

\[
(5.3.4) \quad 2p > (2q+1) + \sqrt{8q+1}
\]

which, therefore, yields a closer upper-bound on \(q\) in terms of \(p\) than the inequality \(q < p\). Thompson \(\text{[47, p.} 40\text{]}\) mentions \((5.3.4)\), without in any way deriving it, in connection with what he terms "unique estimability of communalities". For a given (integral) value of \(p\), the largest integral value of \(q\) which satisfies \((5.3.4)\) will be denoted as \(q^*\).

The usual likelihood ratio test for the model, as proposed by several authors \(\text{[21, 22]}\), involves a \(\chi^2\) approximation, and the degrees of freedom of the \(\chi^2\) distribution is \(\frac{p(p-1)}{2} - pq + \frac{q(q-1)}{2}\).

For the \(\chi^2\) to be applicable, therefore, we should have,

\[
\frac{p(p-1)}{2} - pq + \frac{q(q-1)}{2} > 0
\]

which, again, is equivalent to \((5.3.3)\), and, hence, to \((5.3.4)\). Thus, both for unambiguous estimation and for using the usual test procedures, we should have the inequality, \((5.3.4)\), between \(p\) and \(q\).

It may be noted that, for purposes of estimation etc., the matrix that is accessible to us is

\[
\begin{bmatrix}
\tilde{T}_1 & q \\
T_2 & (p-q) \\
\end{bmatrix}
\]

where \(L(q \times q)\) is \(\perp\), so that, \(\beta(p \times q)\) is indeterminate to the extent of this
factor, $L(q \times q)$. With this understanding, we shall, from now on, without any loss of effective generality, take $\beta(p \times q) = q \begin{bmatrix} \pi_1 \\ T_1 \\ p-q \begin{bmatrix} \pi_2 \\ T_2 \\ q \end{bmatrix} \end{bmatrix}$, when $r(\beta) = q$.

It may be noted also that, in model (5.2.1), if we write $\bar{y}^*(q \times 1) = L(q \times q)y(q \times 1)$, so that, $\bar{y}^*(q \times 1)$ is an orthogonal transformation of $y(q \times 1)$, then $\bar{y}^*(q \times 1)$, itself will satisfy the same assumptions that $y(q \times 1)$ satisfies.

5.4 The problems in factor analysis.

In factor analysis, the problems may be stated in two stages as follows in statistical terminology:

(i) For a given $p$, when the upper bound on $q$ is given by (5.3.4), we want to test the hypothesis that $r(\beta) = r < q$ against the alternative that $r(\beta) = q$.

When $\beta(p \times q)$ is of rank $r < q(-p)$, then, for our method of approaching the problem through confidence intervals on relevant parameters, the matrix that is accessible to us is $r \begin{bmatrix} \pi^* \\ T_1 \\ (p-r) \begin{bmatrix} \pi^* \\ T_2 \\ q \end{bmatrix} \end{bmatrix}$ and not $q \begin{bmatrix} \pi^* \\ T_1 \\ (p-q) \begin{bmatrix} \pi^* \\ T_2 \\ q \end{bmatrix} \end{bmatrix}$ of section 5.3.

(ii) Having decided on a $q$, for a given $p$, we want to test the hypothesis that there is some simplification in the matrix $\beta(p \times q)$ in terms of some of its elements not being significantly different.
from zero. ("Simple structure")

The problems may also be taken in one stage as follows:

For a given \( p \), take the largest permissible value, \( q^* \), of \( q \) given by (5.3.4), and for this pair of values for \( p \) and \( q^* \), to test the hypotheses that some of the elements of the matrix \( \beta(p \times q^*) \) are zero.

We shall seek a solution of the problem, stated in one stage, by means of putting confidence bounds on the relevant parameters, viz., the elements of \( \beta(p \times q^*) \), or the elements of \( \begin{pmatrix} T_1 \\ (p-q^*)T_2 \\ q^* \end{pmatrix} \) defined in section 5.3.

5.5 Confidence bounds for the parameters in the factor analysis model.

Under the assumption that \( X(p \times l) \) has a \( p \)-variate normal distribution, suppose that \( X(p \times m) \) is a matrix of observations on \( m \) individuals from this \( p \)-variate normal. Let \( \overline{X}(p \times l) \) be the sample mean vector, and \( nS(p \times p) = (m-1)S(p \times p) = XX' - \text{max} \), so that \( S(p \times p) \) is the sample dispersion matrix. We recall from (1.3.2) that we have the statement

\[
\lambda_1 a' S \geq a' \Sigma a \geq \lambda_2 a' S a
\]

for all nonnull vectors \( a(p \times l) \), with a probability \( = (1 - \alpha) \), where \( \Sigma(p \times p) \) is the population dispersion matrix. The distribu-
tion problem connected with the statement, \( \lambda_1 c_{\text{max}}(S) \geq \text{all } c(\Sigma) \geq \lambda_2 c_{\text{min}}(S) \), which is implied by (5.5.1), is considered in the next section.

Under the factor analysis model, we have that
\[
\Sigma(p \times p) = \beta_\beta' + D^2,
\]
so that, \( \sigma_{ii} = (\beta_\beta')_{ii} + \gamma_i^2 \) (i = 1, 2, \ldots, p) and \( \sigma_{ij} = (\beta_\beta')_{ij} = \sigma_{ji} \) (i \neq j = 1, 2, \ldots, p), where \( \Sigma(p \times p) = (\sigma_{ij}) \).

In (5.5.1) take \( g(p \times 1) \) such that all the elements are zero except the ith which is unity. Then we see that (5.5.1) \( \implies \)

\[
\lambda_1 s_{ii} \geq \sigma_{ii} \geq \lambda_2 s_{ii}
\]
with a confidence coefficient \( \geq (1 - \alpha) \), where \( S(p \times p) = (s_{ij}) \).

We have statements like (5.5.2) for i = 1, 2, \ldots, p by choosing the nonnull vector \( g(p \times 1) \) suitably in (5.5.1). Next, take \( g(p \times 1) \) such that all the elements are zero except the ith and jth (i \neq j) which are both equal to 1. Then, we obtain that (5.5.1) \( \implies \)

\[
\lambda_1 (s_{ii} + s_{jj} + 2s_{ij}) \geq \sigma_{ii} + \sigma_{jj} + 2\sigma_{ij} \geq \lambda_2 (s_{ii} + s_{jj} + 2s_{ij})
\]
with a confidence coefficient \( \geq (1 - \alpha) \), or, using (5.5.2), the further implied statement

\[
\frac{(\lambda_1 - \lambda_2)}{2} (s_{ii} + s_{jj}) + \lambda_1 s_{ij} \geq \sigma_{ij} \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{ii} + s_{jj}) + \lambda_2 s_{ij}
\]
for i \neq j = 1, 2, \ldots, p.
Noticing that the number of elements in
\[ q^* \left[ \begin{array}{c} T_1 \\ T_2 \end{array} \right] \]
is \( p \cdot q^* \) the number of independent non-diagonal elements of \( \Sigma(p \times p) \), we can, at least theoretically, obtain confidence bounds on all the elements of the matrix
\[ \left[ \begin{array}{c} T_1 \\ T_2 \end{array} \right] \]
or, equivalently, on all the elements of the matrix \( \beta(p \times q^*) \). We shall, however, consider only the cases of \( p = 3,4 \) explicitly here, and treat the solution of the general problem in a later paper.

For both \( p = 3 \) and \( p = 4 \), from (5.3.4) we have, \( q^* = 1 \), so that the matrix \( \beta(p \times q^*) \), or
\[ q^* \left[ \begin{array}{c} T_1 \\ T_2 \end{array} \right] \]
will reduce, in these cases, to column vectors like \( \beta(3 \times 1) \) and \( \beta(4 \times 1) \).

Let us first consider the case of \( p = 3 \). Let us write,
\[ \beta(3 \times 1) = \left[ \begin{array}{c} \beta_1 \\ \beta_2 \\ \beta_3 \end{array} \right] \]
so that,
\[ \beta\beta' = \left[ \begin{array}{ccc} \beta_1^2, & \beta_1\beta_2, & \beta_1\beta_3 \\ \beta_1\beta_2, & \beta_2^2, & \beta_2\beta_3 \\ \beta_1\beta_3, & \beta_2\beta_3, & \beta_3^2 \end{array} \right] \]
From (5.5.4), recalling that the non-diagonal elements of $\Sigma$ are the non-diagonal elements of $\beta\beta'$, we have the following confidence bounds,

$$
(5.5.5) \quad \frac{(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{22}) + \lambda_1 s_{12} \geq \beta_1 \beta_2 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{22}) + \lambda_2 s_{12},
$$

$$
(5.5.6) \quad \frac{(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{33}) + \lambda_1 s_{13} \geq \beta_1 \beta_3 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{33}) + \lambda_2 s_{13},
$$

$$
(5.5.7) \quad \frac{(\lambda_1 - \lambda_2)}{2} (s_{22} + s_{33}) + \lambda_1 s_{23} \geq \beta_2 \beta_3 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{22} + s_{33}) + \lambda_2 s_{23}.
$$

All these bounds may be positive, or all negative, or some of them positive and some negative. We recall that there is a correspondence between confidence regions on parameters or parametric functions, and the acceptance regions of the associated tests on the parameters or parametric functions. Using this correspondence, we shall define a procedure of testing which of the elements of $\beta(3 \times 1)$ are zero, by making use of the statements (5.5.5) - (5.5.7) simultaneously. We shall use the terminology of rejecting and accepting null hypotheses, where acceptance really means non-rejection.

If none of the confidence intervals, (5.5.5) - (5.5.7), includes zero, then we reject the hypothesis that at least one element of the vector $\beta(3 \times 1)$ is zero. On the other hand, if all three of the intervals include zero, then we accept the hypothesis that at least two elements of $\beta(3 \times 1)$ are zero. We can show that this
conclusion is, effectively, the same as concluding that we accept the hypothesis that all three elements of $\beta(3 \times 1)$ are zero. Suppose, for instance, that $\beta_2$ and $\beta_3$ are inferred to be zero. Then, we have

$$x(3 \times 1) = \beta(3 \times 1) \cdot y_1 + \eta(3 \times 1)$$

i.e.,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \beta_1 y_1 + \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix}$$

so that, $x_1$, $x_2$, and $x_3$ are mutually uncorrelated. So, in order to explain the correlations between $x_1$, $x_2$, and $x_3$, we may, in this case, merely take $x(3 \times 1) = \eta(3 \times 1)$, where the elements of $\eta(3 \times 1)$ are mutually uncorrelated. Therefore, we observe that accepting the hypothesis that at least two elements of $\beta(3 \times 1)$ are zero, is equivalent effectively, to accepting the hypothesis that all three elements of $\beta(3 \times 1)$ are zero. From this it also follows that, if all three intervals, $(5.5.5)$ - $(5.5.7)$, include zero, then we may accept the hypothesis that any two (or, effectively, all three) elements of $\beta(3 \times 1)$ are zero. Next, if two of the intervals, $(5.5.5)$ - $(5.5.7)$, include zero, say, $(5.5.5)$ and $(5.5.6)$, but not $(5.5.7)$, then we accept the hypothesis that $\beta_1$ is zero, and reject the hypothesis that $\beta_2$ and/or $\beta_3$ are/is zero. Depending on which two of the intervals include zero and which one does not, we can make other similar inferences. Finally, if only one of the intervals includes zero, while the other two do not, then we are not able to accept the hy-
thesis that any of the elements of $\beta(3 \times 1)$ is zero. However, in view of the joint confidence coefficient of (5.5.5)-(5.5.7) being $\geq (1 - \alpha)$, this last statement would not lead to any real contradiction.

Next, let us consider the case $p = 4$, where

$$
\beta(4 \times 1) = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}, \text{ and, } \beta_0' = \begin{bmatrix} \beta_1^2 & \beta_1 \beta_2 & \beta_1 \beta_3 & \beta_1 \beta_4 \\ \beta_1 \beta_2 & \beta_2^2 & \beta_2 \beta_3 & \beta_2 \beta_4 \\ \beta_1 \beta_3 & \beta_2 \beta_3 & \beta_3^2 & \beta_3 \beta_4 \\ \beta_1 \beta_4 & \beta_2 \beta_4 & \beta_3 \beta_4 & \beta_4^2 \end{bmatrix}
$$

Then, as in the preceding case, we can obtain the following six confidence intervals on the non-diagonal elements of $\beta_0'$, from the corresponding statements (5.5.4):

$$
\frac{(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{22}) + \lambda_1 s_{12} \geq \beta_1 \beta_2 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{22}) + \lambda_2 s_{12}
$$

$$
\frac{(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{33}) + \lambda_1 s_{13} \geq \beta_1 \beta_3 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{33}) + \lambda_2 s_{13}
$$

$$
\frac{(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{44}) + \lambda_1 s_{14} \geq \beta_1 \beta_4 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{11} + s_{44}) + \lambda_2 s_{14}
$$

$$
\frac{(\lambda_1 - \lambda_2)}{2} (s_{22} + s_{33}) + \lambda_2 s_{23} \geq \beta_2 \beta_3 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{22} + s_{33}) + \lambda_2 s_{23}
$$

$$
\frac{(\lambda_1 - \lambda_2)}{2} (s_{22} + s_{44}) + \lambda_2 s_{24} \geq \beta_2 \beta_4 \geq \frac{-(\lambda_1 - \lambda_2)}{2} (s_{22} + s_{44}) + \lambda_2 s_{24}
$$
Recalling from (5.5.1) that \( \lambda_1 \) and \( \lambda_2 \) are functions of \( a, p, n \), we note that the \( \lambda_1 \) and \( \lambda_2 \) occurring in (5.5.5)-(5.5.7) are, in general, different from the \( \lambda_1 \) and \( \lambda_2 \) occurring in (5.5.8)-(5.5.13).

Here again, if none of the intervals, (5.5.8)-(5.5.13), includes zero we reject the hypothesis that at least one of the elements of \( \gamma(4 \times 1) \) is zero. On the other hand, if all the intervals, (5.5.8)-(5.5.13), include zero, then we accept the hypothesis that at least three elements of \( \gamma(4 \times 1) \) are zero. We can reason exactly as in the case of \( p = 3 \) and notice that this last inference is equivalent to accepting the hypothesis that any three (or, effectively, all four) elements of \( \gamma(4 \times 1) \) are zero. Next, suppose that five of the intervals, say, (5.5.8)-(5.5.12), include zero, while (5.5.13) does not. Then, we accept the hypothesis that (a certain pair) \( \beta_1 \) and \( \beta_2 \) are zero, but reject the hypothesis that (the remaining two) \( \beta_3 \) and \( \beta_4 \) are zero. Depending on which of the five intervals include zero and which one does not, we can accept, similarly, the hypothesis that a certain pair of the elements of \( \gamma(4 \times 1) \) are zero and reject the hypotheses that the other two are zero. Next, if three of the intervals, (5.5.8)-(5.5.13), involving the same \( \beta_1(i = 1, 2, 3, \text{ or } 4) \), include zero, while the other three intervals do not, then we accept the hypothesis that this \( \beta_1 \) is zero, but reject the hypothesis that any other \( \beta_j(j \neq 1) \) is zero. It may be noted that, if four of the
intervals, (5.5.8)-(5.5.13), include a set of three intervals involving the same \( \beta_1 \), such that, all four intervals include zero, then we can only accept the hypothesis that \( \beta_1 \) is zero, and go no further. The fact that the fourth interval, which does not involve the \( \beta_1 \) that is common to the other three, includes zero does not provide us with enough additional evidence to be able to accept the hypothesis that any other \( \beta_j \) \((j \neq 1)\) is zero. Finally, if only one or two of the intervals, (5.5.8)-(5.5.13), include zero, while the others do not, then we do not have enough evidence to accept the hypothesis that any of the individual \( \beta_i \) \((i = 1, \ldots, 4)\) are equal to zero. These last two statements would not lead to any real contradiction, in view of the joint confidence coefficient of (5.5.8)-(5.5.13) being \( \geq (1-\alpha) \).

It may be noted that, since the confidence coefficients of the confidence statements, (5.5.5)-(5.5.7), and, (5.5.8)-(5.5.13), are \( \geq (1-\alpha) \), the associated tests of the hypotheses that the \( \beta_i \)'s are zero, are made at a significance level \( \leq \alpha \), for a preassigned \( \alpha \).

The above method of starting from a \( \beta(p \times q^*) \)

\[
\begin{bmatrix}
T_1 \\
T_2 \\
q^* \\
p-q^*
\end{bmatrix}
\]

and then inferring that some elements of this matrix are zero, or, in other words, that \( \beta(p \times q^*) \) is even simpler than the general form

\[
\begin{bmatrix}
T_1 \\
T_2 
\end{bmatrix}
\]

would suggest, will be extended to a general value of \( p \), the number of original variates, in a later paper. It is hoped that this
general treatment will throw more light even on the special cases considered above. Investigations regarding the possibility of obtaining closer bounds than those considered here, are also in progress.

5.6 **Distribution problem involved in sections 1.2, 2.2, 4.9 and 5.5.**

Recall that \( \lambda_1 = n^{-1/2}(p, n), \lambda_2 = n^{-1/2}(p, n) \), where

\[
P \left( \sum_{l=1}^{c_1 \lambda_1} (p, n) \leq c_{\min}^{(nS)} \leq c_{\max}^{(nS)} \leq c_2 \lambda_2 (p, n) \mid \Sigma = I(p) \right) = 1 - \alpha.
\]

We note that \( S(p \times p) \) is a.e. p.d., so that, all \( c(nS) \) are a.e. positive, and they are also a.e. distinct. Let us denote \( c_{\min}^{(nS)} \) and \( c_{\max}^{(nS)} \), respectively, by \( c_1 \) and \( c_p \). Our problem then is the evaluation of the probability,

\[(5.6.1) \quad P \left( \sum_0 \leq c_1 \leq c_p \leq \sum_0' \mid \Sigma = I(p) \right),\]

where \( \sum_0 \) and \( \sum_0' \) are given constants. As mentioned in section 1.3, this problem can be considered as a "limiting" case of the one that is explicitly considered by the authors of \( 3.4 \) and \( 4.3 \). Nanda \( 2.9 \) considers, explicitly, the problems of obtaining the cumulative distribution functions of \( c_1 \) and \( c_p \) separately. The solution given below, for the joint cumulative distribution function of \( c_1 \) and \( c_p \), is, however, based on more general methods which are, essentially, similar to those used in \( 3.3, 3.4 \).

It is well known \( 4.3 \) that the null joint distribution of the characteristic roots, \( 0 < c_1 \leq c_2 \leq \ldots \leq c_p < \infty \), of \( nS(p \times p) \) is
\[(5.6.2) \quad \text{const. } \exp \left( -\frac{1}{2} \sum_{i=1}^{p} c_i \sum_{i=1}^{p} \frac{n-p-1}{2} c_i \prod_{i>j=1}^{p} (c_i - c_j) \right). \]

Making the transformation \( x_i = c_i/2 \) (i = 1, 2, ..., p), we obtain the joint distribution of \( 0 < x_1 \leq x_2 \leq \ldots \leq x_p < \infty \) as,

\[(5.6.3) \quad \text{const. } \exp \left( -\sum_{i=1}^{p} x_i \sum_{i=1}^{p} \frac{n-p-1}{2} x_i \prod_{i>j=1}^{p} (x_i - x_j) \right),\]

where it can be verified that the constant factor is,

\[(5.6.4) \quad K(p,n) = \frac{n^{p/2}}{\prod_{i=1}^{p} (i(n-i+1)r(\frac{p-i+1}{2}))}.\]

Using the fact that \( \prod_{i>j=1}^{p} (x_i - x_j) \) is the Vandermonde's determinant,

\[
\begin{vmatrix}
    x^{p-1}_p & x^{p-1}_p & \ldots & x^{p-1}_1 \\
x^{p-2}_p & x^{p-2}_p & \ldots & x^{p-2}_1 \\
\vdots & \vdots & \ddots & \vdots \\
x_p & x_{p-1} & \ldots & x_1 \\
1 & 1 & \ldots & 1
\end{vmatrix}
\]

we can rewrite (5.6.3) as,
(5.6.5) \( K(p, n) \)  
\[
\begin{array}{c|cccc|c}
\times_{p}^{m+p-1} & x_{p}^{m+p-1} & x_{p-1}^{m+p-1} & \ldots & x_{1}^{m+p-1} & \prod_{i=1}^{p} dx_{i} \\
\times_{p}^{m+p-2} & x_{p}^{m+p-2} & x_{p-1}^{m+p-2} & \ldots & x_{1}^{m+p-2} & \\
\vdots & \vdots & \vdots & \ddots & \vdots & \\
x_{p}^{m} & x_{p}^{m} & x_{p-1}^{m} & \ldots & x_{1}^{m} & \\
\end{array}
\]

where \( m = \frac{n-p-1}{2} \).

We seek to evaluate integrals of the form,

(5.6.6) \( P \int_{x}^{y} x_{1} \leq x_{p} \leq y | E = I(p)_7 \)

\[
= K(p, n) \int_{x}^{y} \int_{x_{p} = x}^{x_{p}} \int_{x_{p-1} = x}^{x_{2}} \int_{x_{1} = x}^{x_{1}} dx_{p} dx_{p-1} dx_{1} dx_{2} \times_{p}^{m+p-1} x_{p}^{m+p-1} x_{p-1}^{m+p-1} x_{1}^{m+p-1} \prod_{i=1}^{p} dx_{i}
\]

\[
= K(p, n) \int_{x}^{y} \begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}^{7}, \text{ say,}
\begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}
\begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}

\begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}
\begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}

\begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}
\begin{pmatrix} m_{p, 1} & m_{p-1, 1} & \ldots & m_{1, 1} \\
\end{pmatrix}

where \( m_{i} = m+1-i, \) for \( i = 1, 2, \ldots, p \).
or,

\[ = \int \cdots \int \frac{m_p, r_p}{x_p} \cdots \frac{m_1, r_1}{x_1} \]  

where

\[ \int \cdots \int \frac{m_p, r_p}{x_p} \cdots \frac{m_1, r_1}{x_1} \]

Some mathematical properties of the function \( \int \cdots \int \frac{m_p, r_p}{x_p} \cdots \frac{m_1, r_1}{x_1} \):

Let us define \( \int \cdots \int \frac{m_p, r_p}{x_p} \cdots \frac{m_1, r_1}{x_1} \)

\[ \int \frac{m_p, r_p}{x_p} dx_p \int \frac{m_p, r_p}{x_p} dx_p \int \frac{m_p, r_p}{x_p} dx_p \int \frac{m_p, r_p}{x_p} dx_p \]

Note that the integral in (5.6.6) is denoted by \( \int \cdots \int \frac{m_p, r_p}{x_p} \), while the repeated integral here is denoted by \( \int \cdots \int \frac{m_p, r_p}{x_p} \).

In particular, \( \int \cdots \int \frac{m_p, r_p}{x_p} = \int \frac{y}{x} e^{-ru} du \).
Also, let us write \( I_0(x;m,r) = x^m e^{-rx} \). We shall now proceed to prove certain lemmas.

**Lemma 5.6a**

\[ I(x; y; m_p, r_p; \ldots; m_1, r_1) = \sum_{s} (-1)^s I(x; y; m'_p, r'_p; \ldots; m'_1, r'_1) \]

where \( (m'_p, r'_p; m'_p, r'_p; \ldots; m'_1, r'_1) \) is any permutation of \( (m_p, r_p; m_p, r_p; \ldots; m_1, r_1) \) and the summation on the right side of (5.6.7) is taken over all such permutations; \( s \) is the number of inversions of the order of the subscripts in \( (m_p, r_p; \ldots; m_1, r_1) \).

**Proof:** The proof is made clear by considering the case of \( p = 3 \).

\[ I(x; y; m_3, r_3; m_2, r_2; m_1, r_1) = \int_{x}^{x_3} \int_{x}^{x_2} \int_{x}^{x_1} \frac{y}{y!} \, dx_1 \]

\[
\begin{bmatrix}
    m_3 & -r_3 x_3 & m_2 & -r_2 x_2 & m_1 & -r_1 x_1 \\
    x_3 e & x_3 e & x_3 e & x_3 e & x_3 e \\
    m_3 & -r_3 x_2 & m_2 & -r_2 x_2 & m_1 & -r_1 x_1 \\
    x_2 e & x_2 e & x_2 e & x_2 e & x_2 e \\
    m_3 & -r_3 x_1 & m_2 & -r_2 x_1 & m_1 & -r_1 x_1 \\
    x_1 e & x_1 e & x_1 e & x_1 e & x_1 e
\end{bmatrix}
\]

\[
= \int_{x}^{x_3} e^{-r_3 x_3} dx_3 \int_{x}^{x_2} e^{-r_2 x_2} dx_2 \int_{x}^{x_1} e^{-r_1 x_1} dx_1
\]

\[
- \int_{x}^{x_3} e^{-r_1 x_2} dx_2 \int_{x}^{x_1} e^{-r_2 x_1} dx_1
\]

\[
- \int_{x}^{x_3} e^{-r_2 x_2} dx_2 \int_{x}^{x_1} e^{-r_1 x_1} dx_1
\]

\[- \int_{x_3}^{x_2} e^{-r_2 x_3} \, dx_3 \int_{x_2}^{x_1} e^{-r_3 x_2} \, dx_2 \int_{x_1}^{x_1} e^{-r_1 x_1} \, dx_1 \]

\[- \int_{x_2}^{x_1} e^{-r_1 x_2} \, dx_2 \int_{x_1}^{x_1} e^{-r_3 x_1} \, dx_1 \]

\[+ \int_{x_3}^{x_1} e^{-r_1 x_3} \, dx_3 \int_{x_2}^{x_2} e^{-r_3 x_2} \, dx_2 \int_{x_1}^{x_1} e^{-r_2 x_1} \, dx_1 \]

\[- \int_{x_2}^{x_2} e^{-r_2 x_2} \, dx_2 \int_{x_1}^{x_1} e^{-r_3 x_1} \, dx_1 \]

\[= x(-1)^8 I(x; y, m_1^{11}, m_2^{11}, m_3^{11}, m_4^{11}, r_1^{11}) .\]

Thus the lemma is true in general and it is seen that
\[
\int_{x}^{y} \left( x, y; m, r, \ldots; m_1, r_1 \right), \text{ of (5.6.6), is in the form of a pseudo-determinant which may be written as,}
\]

\[
\begin{vmatrix}
\int_{x}^{y} e^{-r x} \, dx & \int_{x}^{y} e^{-r x} \, dx & \cdots & \int_{x}^{y} e^{-r x} \, dx \\
\int_{x}^{y} e^{-r x} \, dx & \int_{x}^{y} e^{-r x} \, dx & \cdots & \int_{x}^{y} e^{-r x} \, dx \\
\int_{x}^{y} e^{-r x} \, dx & \int_{x}^{y} e^{-r x} \, dx & \cdots & \int_{x}^{y} e^{-r x} \, dx \\
\int_{x}^{y} e^{-r x} \, dx & \int_{x}^{y} e^{-r x} \, dx & \cdots & \int_{x}^{y} e^{-r x} \, dx
\end{vmatrix}
\]
It must be noted that in opening out the pseudo-determinant it is very important to preserve the order of the factors from $x_p$ through $x_{p-1}$ to $x_{p-2}$, and so on, to $x_1$.

\textbf{Lemma 5.6b}

\[(5.6.8) \sum_{\text{over all permutations}} I(x, y; m_p, r_p; \ldots; m_1, r_1) = \prod_{i=1}^{p} I(x, y; m_i, r_i)\]

\textbf{Proof:} The proof in general will be understood if we consider, for simplicity of algebra, the case of $p = 2$. In this case we have,

\[I(x, y; m_2, r_2; m_1, r_1) + I(x, y; m_1, r_1; m_2, r_2)\]

\[= \int_{x}^{y} x_2^2 e^{-r_2 x_2} dx_2 \int_{x}^{y} x_1^2 e^{-r_1 x_1} dx_1 + \int_{x}^{y} x_1^2 e^{-r_2 x_2} dx_2 \int_{x}^{y} x_2^2 e^{-r_1 x_1} dx_1\]

\[= \int_{x}^{y} x_2^2 e^{-r_2 x_2} dx_2 \int_{x}^{y} x_1^2 e^{-r_1 x_1} dx_1 + \int_{x}^{y} x_2^2 e^{-r_2 x_2} dx_2 \int_{x}^{y} x_1^2 e^{-r_1 x_1} dx_1\]

where the last term is obtained by interchanging, in the last term of the preceding step, $x_2$ and $x_1$, and rewriting the domain of integration.

Thus, we have

\[I(x, y; m_2, r_2; m_1, r_1) + I(x, y; m_1, r_1; m_2, r_2)\]

\[= \int_{x}^{y} x_2^2 e^{-r_2 x_2} dx_2 \int_{x}^{y} x_1^2 e^{-r_1 x_1} dx_1\]

\[= \prod_{i=1}^{2} I(x, y; m_i, r_i).\]
Lemma 5.6c

\[(5.6.9) \sum_{s} I_s(x, y; m_{p-1}, r_{p-1}; \ldots; m_s, r_s, m_{s-1}, r_{s-1}; \ldots; m_1, r_1)\]

\[= I(x, y; m, r) I(x, y; m_{p-1}, r_{p-1}; \ldots; m_s, r_s, m_{s-1}, r_{s-1}; \ldots; m_1, r_1)\]

where \(I_s()\) is the result of writing \((m, r)\) in the \(s\)th place and filling up the remaining \((p-1)\) places of \(I()\) with \((m_{p-1}, r_{p-1}), \ldots, (m_1, r_1)\), for \(s = 1, 2, \ldots, p\). Each \(I_s()\) is, of course, a \(p\)-fold integral, while on the right side of \((5.6.9)\) the second factor is a \((p-1)\)-fold integral.

Proof: The proof in general will follow exactly along the same lines as for the case of \(p=3\), which we shall consider for simplicity of algebra. We have,

\[I_1(x, y; m_2, r_2; m_1, r_1; m, r) + I_2(x, y; m_2, r_2; m, r; m_1, r_1) + I_3(x, y; m, r; m_2, r_2; m_1, r_1)\]

\[= \int_{x}^{y} m_2 \ e^{-r_2 x} \ dx_3 \int_{x}^{m_1} m \ e^{-r_1 x} \ dx_2 \int_{x}^{r_1} \ dx_1\]

\[+ \int_{x}^{y} m_2 \ e^{-r_2 x} \ dx_3 \int_{x}^{m_1} m \ e^{-r_2 x} \ dx_2 \int_{x}^{r_1} \ dx_1\]

\[+ \int_{x}^{y} m \ e^{-r_3 x} \ dx_3 \int_{x}^{m_2} m \ e^{-r_2 x} \ dx_2 \int_{x}^{r_1} \ dx_1\]

\[= \int_{x}^{y} m_2 \ e^{-r_2 x} \ dx_3 \int_{x}^{m_1} m \ e^{-r_2 x} \ dx_2 \int_{x}^{r_1} \ dx_1\]
\[ + \int_{x_3}^{y} \int_{x_2}^{m_2 - r_2 x_3} e^{-r_2 x_3} \, dx_3 \int_{x_2}^{m_1 - r_2 x_2} e^{-r_1 x_2} \, dx_2 \int_{x_1}^{m - r_1 x_1} e^{-r_1 x_1} \, dx_1 \]

which is obtained by interchanging the variables and adjusting the domain of integration suitably,

\[ = I(x, y; m, r) I(x, y; m_2, r_2; m_1, r_1). \]

**Lemma 5.6d**

\[(5.6.10) \sum_{s=1}^{p} (-1)^{s-1} I_s \int x, y; \left( \begin{array}{ccc} m_p, r_p & \cdots & m_1, r_1 \\ \vdots & \ddots & \vdots \\ m_p, r_p & \cdots & m_1, r_1 \\ \vdots & \cdots & \vdots \\ m_p, r_p & \cdots & m_1, r_1 \end{array} \right) \]

\[ = \sum_{s=1}^{p} (-1)^{s-1} \int x, y; \left( \begin{array}{ccc} m_p, r_p & \cdots & m_1, r_1 \\ \vdots & \ddots & \vdots \\ m_p, r_p & \cdots & m_1, r_1 \end{array} \right) \]

where \( I_s \int \) on the left side of (5.6.10) is obtained by replacing the \( s \)th row of \( I \int \) by \( (m_p, r_p; m_{p-1}, r_{p-1}; \cdots; m_1, r_1) \), and \( I_{ss} \int \) on the right side of (5.6.10) is obtained by suppressing the \( s \)th row and \( s \)th column of \( I \int \). It may be noted that each \( I_s \int \) is a \((p \times p)\) pseudo-determinant, and each \( I_{ss} \int \) is a \((p-1 \times p-1)\) pseudo-determinant.
Proof: Again we shall consider the case \( p = 3 \) for simplicity of algebra. For this case, we pick out, from the expansion of each pseudo-determinant on the left side of (5.6.10), the terms involving the index, say, \((m'_3, r'_3)\), and then collect all such terms involving the same index. We shall then have the following contribution from such terms:

\[
I(x, y; m'_3, r'_3; m_2, r_2; m_1, r_1) - I(x, y; m'_3, r'_3; m_1, r_1; m_2, r_2) \\
+ I(x, y; m_2, r_2; m'_3, r'_3; m_1, r_1) - I(x, y; m_1, r_1; m'_3, r'_3; m_2, r_2) \\
+ I(x, y; m_2, r_2; m_1, r_1; m'_3, r'_3) - I(x, y; m_1, r_1; m_2, r_2; m'_3, r'_3) \\
= I(x, y; m'_3, r'_3) \left\{ I(x, y; m_2, r_2; m_1, r_1) - I(x, y; m_1, r_1; m_2, r_2) \right\} ,
\]

using lemma 5.6c,

\[
= I(x, y; m'_3, r'_3) I_\mathcal{L} x, y; \begin{pmatrix} m_2, r_2 & m_1, r_1 \\ m_2, r_2 & m_1, r_1 \end{pmatrix} \mathcal{J} \\
= I(x, y; m'_3, r'_3) I_{11} I_\mathcal{L} x, y; \begin{pmatrix} m_3, r_3 & m_2, r_2 & m_1, r_1 \\ m_3, r_3 & m_2, r_2 & m_1, r_1 \\ m_3, r_3 & m_2, r_2 & m_1, r_1 \end{pmatrix} \mathcal{J}
\]

From this it follows that, in the general case, if we pick out, from the expansion of each pseudo-determinant on the left side of (5.6.10), the term with the index \((m'_p, r'_p)\), and collect together all such terms involving the same index \((m'_p, r'_p)\), then we shall have the following contribution:
By combining different expressions like this, involving the different indices \((m_i', r_i')\) \((i = 1, \ldots, p)\), the result stated in the lemma is obtained.

**Lemma 5.6e**

\[
\int_{x_0}^{y_0} e^{-rx} f(x, x_0) \, dx = \frac{1}{r} \int_{y_0}^{m - ry_0} f(y, x_0) \, dy
\]

\[
+ m \int_{x_0}^{y_0} x^{-r} e^{-rx} f(x, x_0) \, dx
\]

\[
+ m \int_{x_0}^{y_0} x^{m-1} e^{-rx} f(x, x_0) \, dx
\]

where \(m > -1\), \(r > 0\), \(0 < x_0 < y_0 < \infty\) and \(f(x, x_0)\) is such that, \(f'(x, x_0) = \frac{d}{dx} f(x, x_0)\) and the two integrals on the right side exist.

Proof: The proof follows by integration by parts, integrating \(e^{-rx}\) and differentiating \(x^m f(x, x_0)\).

A reduction formula for the evaluation of the integral \(I(\int_{x_0}^{y_0} f(x, y) \, dy)\) \(m_1, \ldots, m_p \neq 0\), where \(m > \ldots > m_1 > -1\) and the \(m_i's\) differ by integers.

Notice that this integral, when \(m_i = m + i - 1\), \((i=1,2,\ldots,p)\) and \(m = \frac{n-p-1}{2}\), is the one that occurs in \((5.6.5)\). It can be verified
that the pseudo-determinant \( I \left( x, y; m_p^1, m_p^2; \ldots ; m_1^1, m_1^2 \right) \) will be zero if any two of its columns are equal. Remembering this, and that \( r_1 = r_2 = \ldots = r_p = 1 \) for the present situation, we shall try to reduce \( m_p \) to \( m_{p-1} \) through successive integration by parts. We shall, for this purpose, consider a typical term in the expansion of the pseudo-determinant, viz., \( I(x, y; m_p^1, m_{p-1}^1; \ldots ; m_1^1, 1) \). Suppose that the largest exponent, \( m_p \), occurs in the \( s \)th place. We then have the typical term,

(5.6.11) \[
I(x, y; m_p^1, m_{s+1}^1; m_p^1, m_{s-1}^1; \ldots ; m_1^1, 1)
\]

\[
= \int x_p \ e^x \ dx \quad \int x_{s+1} \ m_{p-1}^1 \ e^{-x} \ dx_s \quad \int x_{s-1} \ m_1^1 \ e^{-x} \ dx_{s-1}
\]

\[
+ \int x_2 \ m_1^1 \ e^{-x} \ dx_1
\]

\[
= \int x_p \ e^x \ dx \quad \int x_{s+2} \ m_{s+1}^1 \ e^{-x} \ dx_{s+1} \quad \int x_s \ m_p^1 \ e^{-x} \ dx_s
\]

\[
x \ I(x, y; m_{s-1}, 1; \ldots ; m_1, 1) \ dx_s
\]

Using lemma 5.6e, we have

(5.6.12) \[
\int x_{s+1} \ m_p^1 \ e^{-x} \ I(x, y; m_{s-1}, 1; \ldots ; m_1, 1) \ dx_s
\]

\[
= \int x_{s+1} \ e^{-x} \ I(x, y; m_{s-1}, 1; \ldots ; m_1, 1)
\]
\[ + x^{m_p - x} I(x, x; m_{s-1}', 1; \ldots; m_1', 1) \]
\[ + \int_{x}^{x_{s+1}} x^{m_p - x} e^{-x_s} I(x, x_s; m_{s-1}', 1; \ldots; m_1', 1) \, dx_s \]
\[ + m_p \int_{x}^{x_{s+1}} x^{m_p - 1} e^{-x_s} I(x, x_s; m_{s-1}', 1; \ldots; m_1', 1) \, dx_s \]
\[ = -x^{m_p - x_{s+1}} I(x, x_{s+1}; m_{s-1}', 1; \ldots; m_1', 1) \]
\[ + I(x, x_{s+1}; m_{s-1}', 2; m_{s-2}', 1; \ldots; m_1', 1) \]
\[ + m_p I(x, x_{s+1}; m_{s-1}', 1; m_{s-1}', 1; \ldots; m_1', 1) \]

since \( I(x, x; \ldots) = 0 \) and \( I(x, x_s; m_{s-1}', 1; \ldots; m_1', 1) \)
\[ = x^{m_{s-1} - x_s} I(x, x_s; m_{s-2}', 1; \ldots; m_1', 1) \]. Using (5.6.12), we see that (5.6.11) becomes,

\[ (5.6.13) \quad I(x, y; m_p', 1; \ldots; m_{s+1}', 1; m_p, 1; m_{s-1}', 1; \ldots; m_1', 1) \]
\[ = -I(x, y; m_p', 1; \ldots; m_{s+1}', m_p, 2; m_{s-1}', 1; \ldots; m_1', 1) \]
\[ + I(x, y; m_p', 1; \ldots; m_{s+1}', 1; m_{s-1}', m_p, 2; \ldots; m_1', 1) \]
\[ + m_p I(x, y; m_p', 1; \ldots; m_{s+1}', 1; m_{s-1}', 1; m_{s-1}', 1; \ldots; m_1', 1) \]

It may be noted here that the first and second terms on the right side of (5.6.13) are \((p-1)\)-fold integrals, while the last term is a \(p\)-fold integral with the index \(m_p\) reduced to \(m_{p-1}\). It is easy to verify
that the reduction (5.6.13) holds for \( s = p \) if \( s = p \), it can be checked that in place of the right side of (5.6.13) we shall have,

\[
(5.6.14) \quad - \ I_0(y;m_p,1) \ I(x,y;m_{p-1},1;\ldots;m_1,1) \\
+ I(x,y;m_{p-1}+m_p,2;\ldots;m_1,1) \\
+ m_p I(x,y;m_{p-1},1;m_{p-1},1;\ldots;m_1,1) 
\]

and, if \( s = 1 \), we shall have,

\[
(5.6.15) \quad - \ I(x,y;m_p,1;\ldots;m_2,1;m_2+m_p,2) \\
+ I_0(x;m_p,1) \ I(x,y;m_p,1;m_2,1) \\
+ m_p I(x,y;m_p,1;\ldots;m_2,1;m_{p-1},1) 
\]

We shall next introduce a certain simplifying notation. Let

\[
(5.6.16) \quad I(x,y;m_p,1;\ldots;m_{s+1},m_{s-1},1;\ldots;m_1,1) \\
= I(x,y;m_p,1;\ldots;m_{s+1},1;m_p,1;m_{s-1},1;\ldots;m_1,1) 
\]

where \( (\frac{m_p}{m_p}, \frac{1}{1}) \) is supposed to be added to \( (m_{s+1},1) \) on the left, so as to reduce the integral by one dimension;

\[
(5.6.17) \quad I_0(y;m_p,1) \ I(x,y;m_{p-1},1;\ldots;m_1,1) \\
= I(x,y;m_p,1;m_{p-1},1;\ldots;m_1,1) 
\]

\[
(5.6.18) \quad I(x,y;m_p,1;\ldots;m_{s+1},l;m_p,1;m_{s-1},1;\ldots;m_1,1) \\
= I(x,y;m_p,1;\ldots;m_{s+1},1;m_p,1;m_{s-1},1;\ldots;m_1,1) 
\]
where \((\frac{\rightarrow}{m_p}, \frac{\rightarrow}{1})\) is supposed to be added to \(\left(\frac{\rightarrow}{m_{p-1}}, 1\right)\) on the right, so as to reduce the integral by one dimension;

\[
(5.6.1) \quad I_0(x; m_p, 1) I(x, y; m_p, 1; \ldots; m_2, 1) = I(x, y; m_p, 1; \ldots; m_2, 1; \frac{\rightarrow}{m_p}, \frac{\rightarrow}{1}) .
\]

Thus, using (5.6.11)-(5.6.19), we have,

\[
(5.6.20) \quad I \int x, y; m_p, 1; m_p-1, 1; \ldots; m_1, 1 \quad \Rightarrow
\]

\[
\quad = \psi(-1)^s I(x, y; m_p, 1; \ldots; m_1, 1)
\]

\[
\quad = - I \int x, y; \frac{\rightarrow}{m_p}, \frac{\rightarrow}{1}; m_p-1, 1; \ldots; m_1, 1 \quad \Rightarrow
\]

\[
\quad + I \int x, y; \frac{\rightarrow}{m_p}, \frac{\rightarrow}{1}; m_p-1, 1; \ldots; m_1, 1 \quad \Rightarrow
\]

\[
\quad + m_p I \int x, y; m_p-1, 1; m_p-1, 1; \ldots; m_1, 1 \quad \Rightarrow
\]

Remembering the notation introduced in (5.6.16)-(5.6.19), and using lemma 5.6d, we have,

\[
(5.6.21) \quad I \int x, y; \frac{\rightarrow}{m_p}, \frac{\rightarrow}{1}; \ldots; m_1, 1 \quad \Rightarrow
\]

\[
\quad = I_0(y; m_p, 1) I \int x, y; m_p-1, 1; \ldots; m_1, 1 \quad \Rightarrow
\]

\[
\quad + \sum_{s=1}^{p-1} (-1)^s I_s \int x, y; \begin{pmatrix}
\frac{\rightarrow}{m_p-1, 1} & \frac{\rightarrow}{m_p-2, 1} & \ldots & \frac{\rightarrow}{m_1, 1} \\
\frac{\rightarrow}{m_p-2, 1} & \ldots & \ldots & \ldots \\
\frac{\rightarrow}{m_p+m_p-1, 2} & \frac{\rightarrow}{m_p-1, 2} & \ldots & \frac{\rightarrow}{m_1, 2} \\
\frac{\rightarrow}{m_p-1, 1} & \frac{\rightarrow}{m_{p-2}, 1} & \ldots & \frac{\rightarrow}{m_1, 1}
\end{pmatrix} \quad \Rightarrow, where
\]
\[ I_s \backslash \setminus \mathcal{J} \] is obtained by writing \((m_{p,p-1}^2; \ldots; m_p^2, 2)\) for the \(s\)th row of \(I \backslash \setminus x, y; m_{p-1}; \ldots; m_1, 1\).

\[
= I_0(y; m_p, 1) I \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J} \\
+ \sum_{s=1}^{p-1} (-1)^s I(x, y; m_p^2, s, 2) I_{ss} \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J},
\]

where \(I_{ss} \backslash \setminus \mathcal{J}\) is a \((p-2 \times p-2)\) pseudo-determinant obtained by suppressing the \(s\)th row and \(s\)th column of \(I \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1\).

Similarly,

\[
(5.6.22) \quad I \backslash \setminus x, y; \overrightarrow{m_p}, 1; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J} \\
= (-1)^{p-1} I_0(x; m_p, 1) I \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J} \\
+ \sum_{s=1}^{p-1} (-1)^{s-1} I(x, y; m_p^2, m_{p-s}, 2) I_{ss} \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J}.
\]

Substituting from (5.6.21) and (5.6.22) in (5.6.20), we obtain,

\[
(5.6.23) \quad I \backslash \setminus x, y; m_p, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J} \\
= - \left\{ I_0(y; m_p, 1) + (-1)^{p-1} I_0(x; m_p, 1) \right\} \overrightarrow{I \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J}} \\
+ (-1)^{p-1} \sum_{s=1}^{p-1} (-1)^{s-1} I(x, y; m_p^2, m_{p-s}, 2) I_{ss} \overrightarrow{I \backslash \setminus x, y; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J}} \\
+ m_p I \backslash \setminus x, y; m_{p-1}, 1; m_{p-1}, 1; \ldots; m_1, 1 \backslash \setminus \mathcal{J}.
\]

Note that on the left side of (5.6.23) we have a \((p \times p)\) pseudo-determinant, while on the right side of (5.6.23), the first term involves a \((p-1 \times p-1)\) pseudo-determinant, the second group of terms involves
\((p-2 \times p-2)\) determinants like \(I_{ss} {\int} \_j\), and the last term involves the \((p \times p)\) pseudo-determinant \(I_{x,y;m_p-1,1;\ldots;m_1,1} \_j\), in which the highest index is \(m_p-1\). Thus, (5.6.23) gives us a recurrence relation, repeated use of which, reduces \(m_p\) to \(m_p-1\) (in which case the resultant pseudo-determinant will be zero), and we have the following reduction of the pseudo-determinant (or, equivalently, the integral) by one dimension:

\[(5.6.24)\]

\[
I_{x,y;m_p-1,1;\ldots;m_1,1} \_j
\]

\[
= -I_{x,y;m_p-1,1;\ldots;m_1,1} \_j \sum_{s' = 1}^{m_p-m_p-1} \left\{ I_0(y;m_p-s'+1) + (-1)^{p-1} I_0(x;m_p-s'+1) \right\} (m_p)_{s'-1} \\
+ 2 \sum_{s=1}^{p-1} \sum_{s' = 1}^{m_p-m_p-1} (-1)^{s-1} I(x,y;m_p+m_p-s-s'+1,2)
\]

\[
x I_{ss} {\int} x,y;m_p-1,1;\ldots;m_1,1 \_j(m_p)_{s'-1} ,
\]

where \((m)_r = m(m-1) \ldots (m-r+1)\). The \(p\)th order pseudo-determinant (integral) is thus thrown back on \((p-1)\)th and \((p-2)\)th order pseudo-determinants (integrals), and so on, until we get to the first order pseudo-determinants (integrals) which can be evaluated easily by using incomplete gamma function tables.

For the particular problem that we are interested in, we have, \(m_1 = m+1 \cdot 1\) (i = 1, 2, ..., \(p\)) and \(m = \frac{n-p-1}{2}\), so that, \(s'\) in (5.6.24) takes only the value 1. We can, therefore, reduce the probability
integral in (5.6.6),
\[ P \int x \leq x_1 \leq x_p \leq y \mid Z = I(p) \]  
using the reduction formula (5.6.24). The explicit expressions for a few values of \( p \) are given next.

(i) \( p = 2 \):
\[ P_2 = P \int x \leq x_1 \leq x_2 \leq y \mid Z(2 \times 2) = I(2) \]
\[ = K(2, n) \int 2I(x, y; 2m+1, 2) \]
\[ - I(x, y;m, 1) \left\{ I_0(y;m+1,1) + I_0(x;m+1,1) \right\} \]

(ii) \( p = 3 \):
\[ P_3 = P \int x \leq x_1 \leq x_3 \leq y \mid Z = I(3) \]
\[ = K(3, n) \int 2I(x, y;m, 1) I(x, y;2m+3, 2) \]
\[ - 2I(x, y;m+1,1) I(x, y;2m+2,2) \]
\[ - \frac{P_2}{K(2, n)} \left\{ I_0(y;m+2,1) - I_0(x;m+2,1) \right\} \]

5.7 Remarks on the results of section 5.6.
It may be noted that the probability (5.6.1), i.e.,
\[ P(c_0 \leq c_1 \leq c_p \leq c_0' \mid Z = I(p) ) \]
\[ = P(\frac{c_0}{2} \leq x \leq \frac{c_0'}{2} \mid Z = I(p) ) = K(p, n) \int \frac{c_0}{2} \frac{c_0'}{2} \frac{c_0}{4} \frac{c_0'}{4} \frac{c_0}{6} \frac{c_0'}{6} \ldots \frac{c_0}{m} \frac{c_0'}{m} \]
s, so that, given \( c_0 \) and \( c_0' \) we can evaluate the desired probability using the results given in section 5.6. The practical uses of all these results, however, depend on the ready availability of tables.
It is hoped that the work now in progress at the London School of Economics, will eventually include this problem also. Until such tables are available for ready use, the results of section 5.5, and also those of sections 1.3, 2.9 and 4.9, are not immediately applicable to practical situations.

The cumulative distribution function of the largest characteristic root of $nS(p \times p)$ can be obtained as a special case of what has been considered in section 5.6, by taking $x = 0$ and remembering that, for $x = 0, I_0(x;m,r) = 0$. Thus, as a special case of the results obtained in 5.6, one can obtain the asymptotic (as sample size $\rightarrow \infty$) distribution of the largest characteristic root, which figures in Roy's tests $\chi^2_42, \chi^2_3$ for the following hypotheses:

(i) $H_0: \Sigma_{12}^2(p \times q) = 0(p \times q)$, for $\mathbb{N}^{\Sigma_{11}}_{\Sigma_{12}} \left( \begin{array}{cc} \xi_1 & \xi_2 \\ \xi_2 & \end{array} \right)_{q \times q}$, $\mathbb{P}^{\Sigma_{11}}_{\Sigma_{12}} \left( \begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{array} \right)_{p \times p}$

(ii) $H_0: \xi_1(p \times 1) = \xi_2(p \times 1) = \cdots = \xi_k(p \times 1)$, for the $k$ populations, $\mathbb{N}^{\xi_1}_{\xi_1} \left( p \times 1 \right), \mathbb{P}^{\xi_1}_{\xi_1} \left( p \times p \right)$, $i = 1, 2, \ldots, k$.

5.8 Concluding remarks.

After this inquiry had been completed, the work of Danford $\chi^2_{12a}$ and Holzinger and Harman $\chi^2_{18a}$ in the field of factor analysis were brought to the attention of the author. The former gives a statistical treatment of some of the problems in factor analysis. The
latter contains, among its results, certain inequalities between $p$ and $q$ similar to those contained in section 5.3 of this dissertation.


197 Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, XXIV(1933), 417-441 and 498-520.


The Individual Sampling Distributions of the Maximum, the Minimum and Any Intermediate one of the p-statistics on the Null Hypothesis," Sankhya, VII (1945), 133-158.


A Note on 'Some Further Results in Simultaneous Confidence Interval Estimation'," Annals of Mathematical Statistics, XXVII (1956), 856-858.


