SOME CONTRIBUTIONS TO THE THEORY OF UNIVARIATE AND MULTIVARIATE STATISTICAL ANALYSIS

Govind Shrikrishna Mudholkar

Institute of Statistics
Mimeo Series No. 360
May 1963

UNIVERSITY OF NORTH CAROLINA
Department of Statistics
Chapel Hill, N. C.

SOME CONTRIBUTIONS TO THE THEORY OF UNIVARIATE AND MULTIVARIATE STATISTICAL ANALYSIS

by

Govind Shrikrishna Mudholkar

May 1963

Grant No. AF AFOSR-84-63

A number of problems in univariate and multivariate statistical analysis are considered. Extensive classes of solutions with desirable properties are obtained for some of them. Properties from the classical and modern viewpoint are investigated.

This research was supported by the Mathematics Division of the of the Air Force Office of Scientific Research.

Institute of Statistics
Mimeo Series No. 360

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge my gratitude to all those who have been of help, direct and indirect, during the course of this work. I can mention here but a few of them.

I an highly indebted to Professor S. N. Roy for getting me started on this work and providing encouragement and valuable guidance at the early stages.

To Dr. W. J. Hall I extend very sincere thanks for agreeing to be the chairman of my examination committee and for reading the manuscript very carefully and suggesting many valuable improvements.

To Dr. W. Hoeffding and Dr. Harold Hotelling I extend warm appreciation for their valuable criticism.

I am very thankful to Dr. George E. Nicholson, Jr. for his help and advice during my study here. Through him I wish to thank the faculty of the Department for excellent teaching of a balanced and ambitious graduate program in statistics without which this work would have been impossible.

To the members of the office staff of the Department I am grateful; in particular to Miss Martha Jordan for her kindness and to Mrs. Doris Gardner for excellent typing of this dissertation form a messy manuscript.

I am very thankful to the Mathematics Division of the Air Force Office of Scientific Research and the H. E. Fund, Bombay, for the financial assistance which made this work possible.

In conclusion I must record, although very inadequately, my sentiments towards my family. I owe them most for their unfailing faith in me.

TABLE OF CONTENTS

CHAPTER		PAGE
	ACKNOWLEDGEMENTS	ii
	INTRODUCTION AND SUMMARY	vi
I	FOUR PROBLEMS IN MULTIVARIATE ANALYSIS	1
	1.0 Introduction and Summary	1
	1.1 The Multivariate Analysis of Variance	3
	1.1.1 The multivariate linear model and the general multivariate linear hypothesis	3
	1.1.2 Canonical form and notation	8
	1.1.3 Preliminaries and development	9
	1.2 The Problem of Testing Independence Between Two Sets of Variates	18
	1.2.1 The model, the hypothesis and reduction	18
	1.2.2 Canonical form	20
	1.2.3 Further reduction	21
	1.2.4 Preliminaries and development	23
	1.3 Testing the Equality of Two Dispersion Matrices	27
	1.3.1 The model and the hypothesis	27
	1.3.2 Canonical form	28
	1.3.3 The development	29
	1.4 Testing Whether a Dispersion Matrix Equals a Given Matrix	34
	1.4.1 The model and the hypothesis	34
	1.4.2 Canonical form	35
	1.4.3 The development	35

CHAPTER PAGE

II SOME PROPERTIES OF PERCENTAGE POINTS AND PROBABILITY INTEGRALS OF SOME RANDOM VARIABLES - - - - - - - - - - - - 39

2.0 Introduction and Summary - - - - - - - - - - - - - - 39

2.1 Some Properties of Percentage Points and Integrals Connected with Chi-Square Tests - - - - - - - - - - - 40

2.2 Some Properties of Percentage Points and Integrals Connected with Variance-ratio Tests - - - - - - - - 45

2.3 Some Properties of Percentage Points and Integrals Associated with Some Multivariate Tests - - - - - - 52

2.3.1 Power function of T^2-tests - - - - - - - - 52

2.3.2 Percentage points of the distribution of a class of statistics - - - - - - - - - - 54

III A STUDY OF SOME UNION-INTERSECTION PROCEDURES - - - - - 58

3.0 Introduction and Summary - - - - - - - - - - - - - - 58

3.1 A Problem due to Neyman and Pearson - - - - - - - - 59

3.2 The Union-Intersection Principle - - - - - - - - - 63

3.3 Simultaneous Analysis of Variance - - - - - - - - - 64

3.3.1 Model - 64

3.3.2 Canonical form - - - - - - - - - - - - - - - - 67

3.3.3 Comparison of the analysis of variance test and the simultaneous analysis of variance test - - - - - - - - - - - - - - - - 68

3.4 Modified Simultaneous Analysis of Variance, a Multiple Decision Procedure - - - - - - - - - - - 71

3.5 A Property of the Power Functions of Some Intersection Tests - - - - - - - - - - - - - - - - - - 74

CHAPTER PAGE

IV SIMULTANEOUS INTERVAL ESTIMATION - - - - - - - - - - - - - - 78

4.0 Introduction and Summary - - - - - - - - - - - - - 78

4.1 Simultaneous Interval Estimation - - - - - - - - - - 79

4.2 Simultaneous Interval Estimation in the Univariate Analysis of Variance - - - - - - - - - - 81

4.2.1 Univariate model - - - - - - - - - - - - - - - 81

4.2.2 Univariate estimation - - - - - - - - - - - - 81

4.2.3 Applications - - - - - - - - - - - - - - - - - 83

4.3 Simultaneous Interval Estimation in Multivariate Analysis of Variance - - - - - - - - - 84

4.3.1 Multivariate model - - - - - - - - - - - - - - 84

4.3.2 Multivariate estimation - - - - - - - - - - - 85

4.3.3 Applications - - - - - - - - - - - - - - - - - 87

4.4 Comparison of Various Bounds - - - - - - - - - - - - 88

BIBLIOGRAPHY - 91

INTRODUCTION AND SUMMARY

Statistical problems which do not admit 'known optimum' solutions are not rare; but nonexistence of 'known optimum' procedures is almost a characteristic property of multivariate problems, that is, problems associated with statistical analysis of multiresponse data. It has been generally recognized that these problems can be reduced to an extent by invariance consideration. One of the earliest explicit uses of the invariance method is due to Hotelling [21] in his treatment of a multivariate problem of two sets of variables. The invariance reduction, however, does not lead to a unique solution. Under these circumstances, reasonable heuristic principles like the likelihood ratio principle and the union-intersection principle, which are known to have some good properties, have been used to obtain valid and workable procedures for many of these problems. All these procedures have the desirable invariance properties. There is, thus, no available way of deciding under what conditions one of these procedures is preferable to the others. Therefore the current interest is in studying intrinsic and comparative properties of these procedures (See, for instance Ito [27]).

Roy and Mikhail [53] and Mikhail [39] have shown that, in case of many of the multivariate testing problems, power functions of the union-intersection tests have certain monotonicity properties. In the first chapter of this dissertation we shall consider four multivariate problems of testing hypotheses. For each of these problems we shall obtain an infinite class of invariant procedures with the power functions having the desirable monotonicity properties.

In the second chapter of this dissertation we shall obtain a number of properties of percentage points and probability integrals associated with chi-square tests and variance-ratio tests in univariate statistical analysis and with some tests in multivariate statistical analysis. Some of these properties have been generally known from tables and charts; but proofs are not known to exist in related literature. Our proofs shall be based on some results in the classical theory of testing statistical hypotheses.

The studies in chapter two have some bearing on the study of some union-intersection procedures, which we shall undertake in the third chapter. The problems discussed here have the characteristic property mentioned in the first paragraph of this introduction - Mikhail [39] has used a method due to Stein [57], with slight modification, to prove admissibility of some union-intersection tests in multivariate analysis. We shall use a similar argument for comparing the power functions of some union-intersection procedures. More specifically, we shall compare a simultaneous analysis of variance procedure with the 'corresponding' analysis of variance procedure for distant restricted alternatives. Also, we shall suggest a modified simultaneous analysis of variance procedure and study it as a multiple decision procedure.

In the fourth chapter, we shall obtain a result with a view to comparing sharpness of various multiple comparison methods in analysis of variance.

CHAPTER I

FOUR PROBLEMS IN MULTIVARIATE ANALYSIS[1]

1.0 Introduction and Summary

In this chapter we shall consider the following four problems in multivariate analysis:

(i) testing the general multivariate linear hypothesis, i.e, multivariate analysis of variance, abbreviated MANOVA,

(ii) testing independence between two sets of variates, to be called the 'independence' problem,

(iii) testing equality of two dispersion matrices, to be called the 'equidispersion' problem, and

(iv) testing that a dispersion matrix is equal to a given matrix.

Each of these problems can be reduced to an extent by the method of invariance. Each of the tests which has been suggested for these problems is based on the characteristic roots of some matrices, and these roots are the maximal invariants under the invariance reduction. None of these tests, however, is known to be optimum in a sense which would give it precedence over the others. The current interest is, therefore, in investigation of good properties of these tests and their comparisons (Ito [27][2]).

---

[1]This research was supported by the Mathematics Division of the Air Force Office of Scientific Research.

[2]Numbers in square brackets refer to the bibliography.

The main competitors for each of the above problems are tests given by the likelihood ratio principle and the union-intersection principle, except that for the MANOVA problem Hotelling's trace criterion is also very popular.

It has been shown by Roy and Mikhail [53] and Mikhail [39] that the power functions of the union-intersection tests for the above problems increase monotonically as each of a number of non-centrality parameters, which can be regarded as measures of departure from the hypotheses, increase separately. This property implies that these tests are unbiased. Somewhat weaker results in this direction were, previously, obtained by Anderson [3] and Narain [40]. We shall show that for each of these problems there exists an infinite class of test procedures, characterized by ordered characteristic roots of certain matrices and their elementary symmetric functions, which have the same or similar properties.

More specifically, for the first two of the above problems, we shall prove that any test with acceptance region of either of two special forms given below has the monotonicity property. The two forms are

$$\underline{a}'\underline{e} \leqq \text{constant}$$

and

$$\underline{a}'\underline{\lambda} \leqq \text{constant} \quad ,$$

where $\underline{\lambda}$ and $\underline{e}$ are vectors of ordered characteristic roots of certain matrices and of their elementary symmetric functions respectively and $\underline{a}$ is a vector of arbitrary nonnegative constant a_i's (the a_i corresponding to the largest characteristic root being strictly positive). The largest root test, the likelihood ratio test and the trace criterion belong to this class and, thus, share the monotonicity property. The third and the fourth of the above problems are very similar from the point of view of this investigation. For the third problem we have obtained a class of test procedures, characterized by the form of their acceptance regions as above, which have a weaker monotonicity property. After observing that the class of procedures suggested for the third problem has as its counterpart a class of test procedures for the fourth problem, members of which have analogous weaker monotonicity properties, we shall show that a subclass of the procedures in this class have a stronger monotonicity property.

1.1 The Multivariate Analysis of Variance

1.1.1 The Multivariate Linear Model and the General Multivariate Linear Hypothesis

Let $\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_N$ be a sample of N independent observations from N p-variate normal population with a common covariance matrix Σ, and means $E(\underline{z}_i)$ given by

$$\underset{N\times p}{E(Z')} = \underset{N\times m}{D} \; \underset{m\times p}{\zeta}$$

$$= N[\underset{r}{D_1} \;\vdots\; \underset{m-r}{D_2}] \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}_{p} \begin{matrix} r \\ m-r \end{matrix} ,$$

where

(i) $\underset{p\times N}{Z}$ denotes the matrix formed by $\underline{z}_1, \ldots, \underline{z}_N$ as columns,

(ii) $\underset{N\times m}{D}$ is the design matrix of known constants determined by the design of the experiment, with rank$(D) = r \leq m < N$, and a basis D_1.

(iii) ζ_1, ζ_2 is the partition of ζ, a matrix of unknown parameters, associated with the partition D_1, D_2 of D.

Under this set-up the general multivariate linear hypothesis is

$$H_o: \underset{s\times m}{C} \; \underset{m\times p}{\zeta} \; \underset{p\times u}{U} = \underset{s\times u}{0}$$

i.e. $$s[\underset{r}{C_1} \vdots \underset{m-r}{C_2}] \begin{bmatrix} \zeta_1 \\ \cdots \\ \zeta_2 \end{bmatrix} U = 0$$

against

$$H_1: \quad C \zeta U = \eta \neq 0 ,$$

where C and U are matrices given by the hypothesis with $\text{rank}(C) = s \leqq r \leqq m < n$, C_1, C_2 is the partition of C associated with the partition ζ_1, ζ_2 of ζ, and rank $(U) = u \leqq p$.

It is well known (Roy [4]) that this set-up can be reduced to the following:

Let $\underline{z}_1^*, \underline{z}_2^*, \ldots, \underline{z}_s^*, \underline{z}_{s+1}^*, \ldots, \underline{z}_r^*, \underline{z}_{r+1}^*, \ldots, \underline{z}_N^*$ be a sample of N independent observations from a u-variate normal population with a common covariance matrix and means

$$\begin{aligned} E(\underline{z}_i^*) &= \underline{\zeta}_i^* , \qquad i = 1,2,\ldots,r \\ &= \underline{0} , \qquad i = r+1,\ldots,N . \end{aligned}$$

In this reduced set-up the general multivariate linear hypothesis H_o is reduced to

$$H_o^*: \quad \underline{\zeta}_1^* = \underline{\zeta}_2^* = \cdots = \underline{\zeta}_s^* = \underline{0}$$

against $H_1^*: \quad \underline{\zeta}_i^* = \underline{\eta}_i^* , \qquad i = 1,2,\ldots,s ,$

where not all $\underline{\eta}_i^*$ are $\underline{0}$ simultaneously. Let

$$\underset{u\times s}{Z_1^*} = [\underline{z}_1^*, \ldots, \underline{z}_s^*]$$

and

$$\underset{u\times(N-r)}{Z_3^*} = [\underline{z}_{r+1}^*, \ldots, \underline{z}_N^*] .$$

Then all the suggested tests for the problem use symmetric functions of the characteristic roots of

$$(Z_1^* Z_1^{*\prime})(Z_3^* Z_3^{*\prime})^{-1}$$

as the test statistics.

Roy [49] has shown that the power function of the maximum root test involves, aside from the degrees of freedom, only $t = \min(u,s)$ noncentrality parameters, which are characteristic roots of a matrix of parameters. His argument can, however, be simplified as follows:

The problem of testing the general multivariate linear hypothesis is known to be invariant under the following groups of transformations (Anderson [4], Lehmann [35], Roy [49], James [28]):

G_1: Addition of arbitrary constants to the components of the variables $\underline{z}_{s+1}^*, \ldots, \underline{z}_r^*$.

G_2: Orthogonal transformations of the structure

$$\underset{u\times s}{Z_1^{**}} = \underset{u\times u}{A_1} \underset{u\times s}{Z_1^*} \quad \text{and} \quad \underset{u\times(N-r)}{Z_3^{**}} = \underset{u\times u}{A_3} \underset{u\times(N-r)}{Z_3^*}$$

on the matrices Z_1^* and Z_3^*, where A_1 and A_3 are real orthogonal matrices.

G_3: Transformations of the structure $Z_1^{**} = B Z_1^*$ and $Z_3^{**} = B Z_3^*$, where B is any $u\times u$ nonsingular matrix.

Under these groups of transformations, the $t = \min(u,s)$ roots of the determinantal equation

$$| Z_1^* Z_1^{*\prime} - \theta \; Z_3^* Z_3^{*\prime} | = 0 ,$$

or, in terms of the original variates, the roots of the determinantal equation

$$| S_{H_o} - \theta \; S_E | = 0 ,$$

where S_{H_o} and S_E are two $u \times u$ matrices, called the hypothesis matrix and the error matrix respectively, and

$$S_{H_o} = U'Z \, D_1(D_1'D_1)^{-1} C_1' \left[C_1(D_1'D_1)^{-1} C_1' \right]^{-1} C_1(D_1'D_1)^{-1} D_1' \, Z' \, U$$

$$S_E = U' \left[Z Z' - Z \, D_1(D_1'D_1)^{-1} D_1' \, Z' \right] U ,$$

form a set of maximal invariants. A test for H_o will be invariant under G_1, G_2 and G_3 if, and only if, it depends on the N observations through the t roots of the above determinantal equation.

Each of the above groups of transformations induces a corresponding group of transformations in the space of parameters Σ and $\underline{\xi}_i$, $i = 1,2,\ldots,r$. A set of maximal invariants under these groups is the set of t noncentrality parameters, which are the roots of the determinantal equation

$$| \Gamma - \theta I | = 0 ,$$

where $\Gamma = (U'\Sigma U)^{-1} \, \eta' \, C_1' \left[C_1(D_1'D_1)^{-1} C_1' \right]^{-1} C_1 \, \eta$.

Then a reference to the following theorem (Lehmann [35]) shows that the power functions of all the tests based on the characteristic roots of $(S_{H_o} S_E^{-1})$ will depend only on the noncentrality parameters mentioned above.

Theorem: Let $X \in \mathcal{X}$ be distributed according to a probability density P_θ, $\theta \in \Omega$. Let G be a group of transformations of the sample space $\mathcal{X}$ onto itself; and let $\bar{G}$ be the group of transformations of the parameter space induced by G. Let $T(X)$ be invariant under G. Then, if $V(\theta)$ is maximal invariant under the induced group $\bar{G}$, the distribution of $T(X)$ depends only on $V(\theta)$.

When $t = 1$ there exists a UMP invariant test which coincides with the T^2-test due to Hotelling [20]. But for $t > 1$, the problem of MANOVA, even though, considerably reduced, does not have an optimum test. There are three well-known tests which are all invariant under G_1, G_2 and G_3. These are, in terms of critical regions:

(i) Roy's maximum characteristic root test (Roy [47, 49]):

$$\underset{max}{Ch}\,(S_{H_o} S_E^{-1}) \geq \text{constant} \quad ,$$

(ii) Likelihood ratio test (Wilks [62], Rao [46]):

$$\Lambda = \frac{|S_E|}{|S_{H_o} + S_E|} \leq \text{constant} \quad ,$$

(iii) Hotelling's T_o^2 test (Hotelling [22], Lawley [30, 31])

$$\text{trace}\,(S_{H_o} S_E^{-1}) \geq \text{constant} \quad .$$

Pillai [43] has suggested three other functions of the characteristic roots as the test statistics; but, because of their arbitrary nature, we do not consider them here.

If we wish to restrict to tests, which are invariant under the groups G_1, G_2, G_3, we can use the following canonical form due to Roy and Mikhail [53], which can be obtained by transformations in G_1, G_2 and G_3.

1.1.2 Canonical Form and Notation

Let $\underset{u\times s}{X} = (x_{ij})$ and $\underset{u\times n}{Y} = (y_{ij})$ be two matrices of random variables x_{ij} $(i = 1,\ldots,u;\ j = 1,2,\ldots,s)$ and y_{ij} $(i = 1,2,\ldots,u;\ j = 1,2,\ldots,n)$ with joint probability density

$$\text{constant}\ \exp\left[-\frac{1}{2}\left\{\sum_{i=1}^{u}\sum_{j=1}^{n} y_{ij}^2 + \sum_{i=1}^{t}(x_{ii}-\theta_i)^2 + \sum_{i=t+1}^{u} x_{ii}^2 + \sum_{i=1}^{u}\sum_{j\neq i}^{s} x_{ij}^2\right\}\right] \prod_{i,j} dx_{ij} \prod_{i,j} dy_{ij} ,$$

where u and s are the same as in the original model, namely, the effective number of variates and the number of degrees of freedom (d.f.) for the hypothesis; $n = N - r$ is the number of d.f. for error; and the quantities θ_i^2 are the $t = \min(u,s)$ noncentrality parameters mentioned above. The general multivariate linear hypothesis in canonical form is

H_0: $\theta_1 = \theta_2 = \ldots = \theta_t = 0$ against

H_1: At least one $\theta_i \neq 0$ $i = 1,2,\ldots,t$.

Under this canonical form we shall restrict to tests which involve x_{ij}'s and y_{ij}'s only through the characteristic roots of

$$(XX')(YY')^{-1} .$$

We shall order these roots and denote them by

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_u .$$

Let $e_1 = \sum_i \lambda_i$, $e_2 = \sum_{i\neq j} \lambda_i\lambda_j$, ..., $e_u = \lambda_1\ldots\lambda_u$ be the u elementary symmetric functions of the u roots. Obviously only t λ_i's and t e_j's will be nonzero. Let

$$\underline{\lambda} = \begin{bmatrix}\lambda_1\\ \vdots \\ \lambda_u\end{bmatrix} \quad \text{and} \quad \underline{e} = \begin{bmatrix}e_1\\ \vdots \\ e_u\end{bmatrix} .$$

We shall denote a test by $\phi = \phi(X, Y)$, its acceptance region by A_ϕ in the space of (X, Y) and the boundary of the acceptance region by B_ϕ.

We shall be interested in certain sections of A_ϕ. To describe these sections let us express the matrix X as

$$X = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_s]$$

$$= [\underline{x}_1, X_1] .$$

Then the sections we shall be interested in are those of A_ϕ for fixed X_1 and Y. We denote these and their boundaries, respectively by $A_{\phi,X_1,Y}$ and $B_{\phi,X_1,Y}$, both in the space of $\underline{x}_1' = (x_{11}, x_{21}, \ldots, x_{u1})$.

1.1.3 Preliminaries and Development

Roy and Mikhail [53] have shown that the power function of ϕ with

$$A_\phi: \lambda_u \leqq \text{constant}$$

is a monotonically increasing function of $(\theta_1, \ldots, \theta_t)$ separately. In this section we shall obtain two infinite classes of tests for which this property holds. For this we shall need investigation of some properties of ϕ. We shall define these:

Definition 1: A test ϕ for the MANOVA problem is said to have the 'monotonicity' property if its power function increases montonically as each of the noncentrality parameters $|\theta_i|$, $i = 1, 2, \ldots, t$ increases separately.

Definition 2: A test ϕ for the MANOVA problem is said to have the 'property $\mathcal{Q}$' if $B_{\phi,X_1,Y}$ is a homogeneous quadric in the space of the components $x_{11}, \ldots, x_{u1}$ of $\underline{x}_1$.

Definition 3: A test ϕ for the MANOVA problem is said to have the 'property $\mathcal{S}$' if $\underline{x}_1 \in A_{\phi,X_1,Y}$ implies $c\,\underline{x}_1 \in A_{\phi,X_1,Y}$, $0 \leqq c \leqq 1$.

Definition 4: A test ϕ for the MANOVA problem is said to have the 'property $\mathcal{F}$' if the intersection of $A_{\phi,X_1,Y}$ and any line L, through the origin in the space of $\underline{X}_1$, is a finite segment of the line.

Now we shall prove a theorem, which is implicit in Roy and Mikhail [53], connecting these properties:

Theorem 1; If a test ϕ for the MANOVA problem has all the three properties $\mathcal{Q}$, $\mathcal{S}$ and $\mathcal{F}$ then it has the monotonicity property.

Proof:- We have to show that

$$\int_{A_\phi} \text{const.} \exp\left[-\frac{1}{2}\left\{\sum_{i,j} y_{ij}^2 + \sum_{i=1}^{t}(x_{ii} - \theta_i)^2 + \sum_{i=t+1}^{u} x_{ii}^2 + \sum\sum x_{ij}^2\right\}\right] \prod_{i,j} dx_{ij} \prod_{i,j} dy_{ij}$$

decreases as each $|\theta_i|$, $i = 1,2,\ldots,t$ increases separately; or equivalently, the integral, with $\theta_i = 0$ in the integrand, decreases monotonically as A_ϕ is given a translation composed of translations θ_i, $i = 1,2,\ldots,t$ in the directions of x_{ii}, $i = 1,2,\ldots,t$. We assert that, becuase of the form of the integrand, we may give these translations successively. Moreover, as A_ϕ involves x_{ij}'s only through the characteristic roots of $(XX')(YY')^{-1}$, whatever is true of $\underline{x}_1$ will be true of all the columns of X. It will be, thus, sufficient to show that the integral, with $\theta_i = 0$, $i = 1,2,\ldots,t$ in the integrand, decreases as A_ϕ is given a translation θ_1 in the direction x_{11}.

To do this observe that $A_{\phi, X_1, Y}$ because of the properties $\mathcal{Q}$, $\mathcal{G}$ and $\mathcal{J}$ must be an ellipsoid in the space of $(x_{11}, \ldots, x_{u1})$. This ellipsoid can always be referred to with respect to its principal axes. In other words, there always exists an orthogonal transformation

$$\underset{(u\times 1)}{\underline{z}} = \underset{(u\times u)(u\times 1)}{M\, \underline{x}_1} \quad , \quad M \text{ orthogonal} \quad ,$$

$$\underline{z}' = (z_1, \ldots, z_u) \, ,$$

such that $\underline{z}_1, \ldots, \underline{z}_u$ are the principal axes of the ellipsoid. If the rows of M are $(m_{i1}, \ldots, m_{iu})$ $i = 1,2,\ldots,u$ then a translation θ_1 in the direction X_{11} is equivalent to $m_{11}\theta_1$ along $\underline{z}_1$, $m_{21}\theta_1$ along $\underline{z}_2, \ldots, m_{u1}\theta_1$ along $\underline{z}_u$. Now the equation of an ellipsoid referred to principal axes is free from the product terms, and $\frac{1}{\sqrt{2\pi}} \int_{-\mu+\theta}^{\mu+\theta} e^{-\frac{1}{2}x^2} dx$ decreases monotonically as $|\theta|$ increases.

Hence we have the theorem.

We shall now proceed to develop the relationships among these properties in the following lemmas and theorems.

<u>Lemma 1:</u> <u>A test ϕ for the MANOVA problem with</u>

$$B_\phi: \; \underline{a}'\,\underline{\lambda} = \text{constant, or}$$

$$B_\phi: \; \underline{a}'\,\underline{e} = \text{constant, where}$$

<u>$\underline{a}' = (a_1, a_2, \ldots, a_u)$ is a fixed real vector, has the property $\mathcal{Q}$.</u>

<u>Proof:</u>- (YY') is a $(u\times u)$ symmetric and p.d. (a.e.) matrix. Therefore, there exists a nonsingular $(u\times u)$ triangular matrix $\tilde{V}$ with zeros above the diagonal such that

$$YY' = (\tilde{V}'\,\tilde{V})^{-1} \, .$$

Then the characteristic roots of

$$(XX')(YY')^{-1} = (XX')(\tilde{V}'\tilde{V})$$

are the same as those of

$$S^* = X^* X^{*\prime} = (\tilde{V} X)(\tilde{V} X)' \ .$$

Now consider determinantal equation

$$|S^* - \nu I| = 0 \, .$$

We can write this as

$$\nu^u - \not{z}_1 \nu^{u-1} + \not{z}_2 \nu^{u-2} - \ldots\ldots + (-1)^u \not{z}_u = 0 \ ,$$

where

$$\not{z}_j = \text{sum of all } \binom{u}{j} \text{ principal minors of } |S^*| \ ,$$
$$j = 1,2,\ldots, u \ .$$

Thus $\not{z}_1 = \text{trace}(S^*)$ and $\not{z}_u = |S^*|$. It is clear from the theory of equations that $\not{z}_j = e_j$ the j-th elementary symmetric function of the roots of the equation.

Now let us fix X_1 and $\underset{\sim}{Y}$, i.e. $\tilde{V}$, and investigate the nature of e_j as a function of $\underline{x}_1$. When X_1 and $\tilde{V}$ are fixed, the elements of $\underline{x}_1^*$, the first column of X^*, are fixed linear functions of the elements of $\underline{x}_1$; and X_1^*, the submatrix of X^* corresponding to X_1 of X, is fixed. Then any j-rowed principal minor of

$$|S^*| = \left| \begin{bmatrix} x_{11}^* & \cdots\cdots & x_{1s}^* \\ \vdots & & \vdots \\ x_{u1}^* & \cdots\cdots & x_{us}^* \end{bmatrix} \cdot \begin{bmatrix} x_{11}^* & \cdots\cdots & x_{u1}^* \\ \vdots & & \vdots \\ x_{1s}^* & \cdots\cdots & x_{us}^* \end{bmatrix} \right|$$

is a homogeneous quadratic function of $(x^*_{11},\ldots,x^*_{u1})$ with coefficients which are functions of the other x^*_{ij}'s, plus a constant which is also a function of the other x^*_{ij}'s. But $(x^*_{11},\ldots,x^*_{u1})$ are linear functions of $(x_{11},\ldots,x_{u1})$. Therefore, given Y and X_1 any e_j, and hence any linear combination $\underline{a}'\underline{e}$ of e_j, is a homogeneous quadratic function of $(x_{11},\ldots,x_{u1})$ with coefficients which are functions of $\underline{Y}$ and X_1, plus a constant depending upon Y and X_1. Thus $\underline{a}'\,\underline{e}$ = constant is a homogeneous quadric in the space of $(\underline{x}_1)$ when $(X_1, \dot{Y})$ is fixed. This proves the first half of the lemma.

For the second part we observe that if λ_j is a root of $(XX')(YY')^{-1}$ it satisfies the equation

$$|S^* - \nu I| = 0 .$$

Thus if we have

$$\lambda_j = \text{constant}$$

and X_1 and Y are fixed

$$|S^* - \lambda_j I| = 0$$

will be a homogeneous quadric in the space of $\underline{x}_1$. The same is true of any fixed linear combination $\underline{a}'\underline{\lambda}$ of λ_j's.

An immediate consequence of this lemma is the following theorem:

Theorem 2: Any test ϕ for the MANOVA problem with the properties $\mathcal{S}$ and $\mathcal{F}$ and with

$$B_\phi:\ \underline{a}'\,\underline{e} = \text{constant, or}$$

$$B_\phi:\ \underline{a}'\,\underline{\lambda} = \text{constant} ,$$

where $\underline{a}' = (a_1, \ldots, a_u)$ is a real vector, has the monotonicity property.

Examples:- The maximum characteristic root test has B_ϕ of the form B_ϕ: $\underline{a}'\underline{\lambda}$ = const. with $\underline{a}' = (0,0,\ldots,1)$. It has been shown that it has the properties $\mathfrak{I}$ and $\mathfrak{F}$ (Roy and Mikhail [53]).

The trace test has B_ϕ with either of the following two forms

$$B_\phi: \underline{a}'\underline{\lambda} = \text{constant}, \quad \underline{a}' = (1,1,\ldots,1)$$

or

$$B_\phi: \underline{a}'\underline{e} = \text{constant}, \quad \underline{a}' = (1,0,\ldots,0) .$$

The likelihood ratio criterion has B_ϕ of the form:

$$B_\phi: \underline{a}'\underline{e} = \text{constant}, \quad \underline{a}' = (1,1,\ldots,1) .$$

In what follows we shall show that both these tests have the properties $\mathfrak{I}$ and $\mathfrak{F}$. It will then follow that all three tests have the monotonicity property.

Now we shall state a lemma which will be very useful in the sequel.

Lemma 2: If $\lambda_1 \leqq \ldots \leqq \lambda_u$ are the u ordered characteristic roots of a symmetric $(u \times u)$ matrix A and $\Lambda_1 \leqq \ldots \leqq \Lambda_u$ are the u ordered characteristic roots of a matrix $(A+B)$, where B is a symmetric at least p.s.d. matrix, then

$$\lambda_j \leqq \Lambda_j, \qquad j = 1,2,\ldots,u \text{ with strict inequality}$$

if B is p.d.

This result, though known, is not very widely available. Bellman [6] has given a proof which involves properties of continuous functions. A matrix proof of this result has been given by Everitt [12].

Lemma 3: *Let $\lambda_1(c) \leqq \dots \leqq \lambda_u(c)$ denote u characteristic roots of the matrix*

$$\left[(c\underline{x}_1)(c\underline{x}_1)' + X_1X_1'\right]\left[YY'\right]^{-1} .$$

Then $\lambda_j(c)$, $j = 1,2,\dots,u$ are nondecreasing functions of c.

Proof:- We can write

$$\left[c_2^2\ \underline{x}_1\underline{x}_1' + X_1X_1'\right](YY')^{-1}$$

$$= \left[c_1^2\, \underline{x}_1\underline{x}_1' + X_1X_1'\right](YY')^{-1} + \left[(c_2^2 - c_1^2)\underline{x}_1\underline{x}_1' + X_1X_1'\right](YY')^{-1} .$$

Application of the previous lemma proves that the ordered characteristic roots of $\left[c_2^2\, \underline{x}_1\underline{x}_1' + X_1X_1'\right](YY')^{-1}$ are greater than or equal to the ordered characteristic roots of $\left[c_1^2\, \underline{x}_1\underline{x}_1' + X_1X_1'\right](YY')^{-1}$ if $c_2 > c_1$.

We shall, now, prove the following:

Theorem 3: *Any test ϕ, for the MANOVA problem with the property $\mathcal{F}$ and*

$$A_\phi:\quad \underline{a}'\,\underline{\lambda} \leqq \text{constant} \quad\text{or}$$

$$A_\phi:\quad \underline{a}'\,\underline{e} \leqq \text{constant} ,$$

$$\underline{a}' = (a_1,\dots, a_u), \quad a_j \geqq 0, \quad j = 1,2,\dots,u$$

has the monotonicity property.

Proof:- The tests have the property $\mathcal{Q}$. To see that they also have the property $\mathcal{J}$ we observe that,

(i) All the ordered characteristic roots are non-negative,

(ii) The ordered characteristic roots of $(c^2\underline{x}_1\underline{x}_1' + X_1X_1')(YY')^{-1}$ $0 \leq c \leq 1$ are not greater than the ordered characteristic roots of $(\underline{x}_1\underline{x}_1' + X_1X_1')(YY')^{-1}$; and the same is true of a linear function

$\underline{a}'\underline{\lambda}$ or $\underline{a}'\underline{e}$ if $a_j \geqq 0$, $j = 1,2,\ldots,u$. Therefore, if

$$\underline{x}_1 \in A_{\phi,X_1,Y} \text{ then } c\,\underline{x}_1 \in A_{\phi,X_1,Y} \quad , \quad 0 \leqq c \leqq 1 .$$

By hypothesis of the theorem the tests have the property $\mathfrak{A}$. Therefore by Theorem 1 they have the monotonicity property.

Examples:- All the three tests mentioned in this section, namely, the largest root test, the likelihood ratio test and the trace test have the properties mentioned in the theorem.

Theorem 4: All tests ϕ for the MANOVA problem with

$$A_\phi : \quad \underline{a}'\underline{e} \leqq \text{constant} \quad ,$$

$$\underline{a}' = (a_1,\ldots,a_u), \quad a_j \geqq 0, \ j = 1,2,\ldots,u \quad ,$$

have the monotonicity property.

Proof:- We know that all the tests ϕ of the theorem have the property $\mathfrak{A}$ and the property $\mathfrak{I}$. We shall show that they have the property $\mathfrak{F}$.

Consider a line L in the space of $\underline{x}_1$, with direction cosines $\underline{\ell}' = (\ell_1,\ldots,\ell_u)$. The equations of L may be written as

$$\frac{x_{11}}{\ell_1} = \frac{x_{21}}{\ell_2} = \cdots = \frac{x_{u1}}{\ell_u} \quad ,$$

and any point on L may be written as $(\ell_1 r,\ldots,\ell_u r)$, where r is the distance of the point from the origin.

Now if $(\ell_1 r,\ldots,\ell_u r) \in A_{\phi,X_1,Y}$ then at least one elementary symmetric function of the characteristic roots of the matrix

$$(r^2 \underline{\ell}\,\underline{\ell}' + X_1 X_1')(YY')^{-1}$$

is nonzero and finite. This implies that the maximum root of the matrix is finite. But

$$\underset{max}{ch}\,(r^2 \ell \ell')(YY')^{-1} \leq \underset{max}{ch}\,(r^2 \ell \ell' + X_1X_1')(YY')^{-1}$$

Thus

$$r^2 \text{ (constant)} \leq \underset{max}{ch}\,(r^2 \ell \ell' + X_1X_1')(YY')^{-1} < \infty .$$

This proves that the tests ϕ have the property $\mathfrak{F}$. Hence the theorem is proved.

Theorem 5: All tests ϕ for the MANOVA problem with

$$A_\phi: \underline{a}'\underline{\lambda} \leq \text{constant},$$

$$\underline{a}' = (a_1, \ldots, a_u),\quad a_j \geq 0,\quad j = 1,2,\ldots,u-1,\quad a_u > 0,$$

have the montonicity property:

Proof: Since $a_u > 0$ and λ's are all nonnegative

$$\underline{a}'\underline{\lambda} \leq \text{constant}$$

implies that $\lambda_u \leq \text{constant} < \infty$. The proof of Theorem 4. indicates that this implies that the tests ϕ have the property $\mathfrak{F}$. It has been shown before that the tests have the property $\mathfrak{A}$ and the property $\mathfrak{J}$. Hence the theorem is proved.

Examples:- Theorems 4 and 5 show that the three tests for the MANOVA problem have the monotonicity property.

Remark: We have shown that all the three tests for the MANOVA problem have the monotonicity property. For this we have, essentially, verified that these tests have the properties $\mathfrak{A}$, $\mathfrak{J}$ and $\mathfrak{F}$. It may be remarked that for the individual cases the verification of the properties is much easier.

1.2 The Problem of Testing Independence Between Two Sets Of Variates.

1.2.1 The Model, The Hypothesis and Reduction

Let $\underset{(p+q)x1}{\underline{z}} = \begin{bmatrix} \underline{z}_1 \\ \underline{z}_2 \end{bmatrix}_1 \begin{matrix} p \\ q \end{matrix}$, $p \leqq q$ be distributed as a (p+q)-variate normal with a $(p+q)x(p+q)$ symmetric p.d. matrix Σ as the covariance matrix. Let Σ be partitioned as

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{12}' & \Sigma_{22} \end{array}\right] \begin{matrix} p \\ q \end{matrix} ,$$
$$\qquad\qquad p \qquad q$$

where p and q rows and columns of Σ correspond, respectively, to the p-set $\underline{z}_1$ and q-set $\underline{z}_2$. Let $\underset{(p+q)xn^*}{Z} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}_{n^*} \begin{matrix} p \\ q \end{matrix}$ be a sample of n^* from the above population. Then the problem is that of testing independence between the p-set and the q-set, i.e.

$$H_o: \underset{pxq}{\Sigma_{12}} = \underset{pxq}{0}$$

against

$$H_1: \Sigma_{12} \neq 0 .$$

As in the case of the general multivariate linear hypothesis we can reduce the problem by invariance. The problem of testing H_o against H_1 remains invariant under the following groups of transformations:

G_1: Addition of arbitrary constants to the elements of Z i.e.

$$\underset{(pxq)n^*}{Z^*} = Z + B$$

G_2: Nonsingular linear transformations of the variates within the sets, i.e.

$$\underset{(p+q)\cdot n^*}{Z^*} = C \underset{(p+q)\cdot n^*}{Z}$$

where C is a $(p+q)x(p+q)$ nonsingular matrix with structure

$$C = \left[\begin{array}{c|c} C_{11} & 0 \\ \hline 0 & C_{22} \end{array}\right] \begin{array}{l} p \\ q \end{array}$$
$$\quad\quad p \quad\quad q$$

A set of maximal invariants, under the groups of transformations induced in the space of sufficient statistics, is the set of p characteristic roots of the matrix

$$S_{11}^{-1} \; S_{12} \; S_{22}^{-1} \; S_{12}' \quad ,$$

where

$$\underset{(p+q)x(p+q)}{S} = Z\,Z' - n^* \underline{\bar{z}}\,\underline{\bar{z}}' \quad ,$$

$$\underset{(p+q)x1}{\bar{z}} = \frac{1}{n^*} \sum_{i=1}^{n^*} \underline{z}_i \; , \quad \underline{z}_i \text{ are the columns of } Z, \text{ and}$$

S is partitioned as

$$S = \left[\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{12}' & S_{22} \end{array}\right] \begin{array}{l} p \\ q \end{array} \quad .$$
$$\quad\quad p \quad\quad q$$

Also a set of maximal invariants under groups $\bar{G}_1, \bar{G}_2$ of transformation induced by the groups G_1, G_2 in the space of parameters is the set of p characteristic roots of matrix

$$\Sigma_{11}^{-1} \; \Sigma_{12} \; \Sigma_{22}^{-1} \; \Sigma_{12}' \quad .$$

These are the p population canonical correlations (Hotelling [21]). All the invariant tests of the hypothesis H_o will involve observations only through the characteristic roots of $(S_{11}^{-1} S_{12} S_{22}^{-1} S'_{12})$; and their power function will depend, aside from the degrees of freedom, only on the characteristic roots of $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma'_{12}$.

Thus, as for the case of the MANOVA problem, for the independence problem no UMP invariant test exists. The two well known tests for the problem are:

(i) The largest root test with critical region

$$\underset{max}{ch} (S_{11}^{-1} S_{12} S_{22}^{-1} S'_{12}) \geqq \text{constant.}$$

(ii) The likelihood ratio test with critical region

$$\frac{|S|}{|S_{11}| \cdot |S_{22}|} \leqq \text{constant.}$$

Roy and Mikhail [53] have shown that the power function of the largest characteristic root test is a monotonically increasing function of the p characteristic roots of $\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma'_{12}$ separately. We shall show that there exist at least two infinite classes of invariant tests which have this property. We shall work in the following canonical form due to Roy [49]

1.2.2 Canonical Form

Let X and Y be random matrices of the form

$$\underset{p \times n}{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{bmatrix}, \quad \underset{q \times n}{Y} = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & & \vdots \\ y_{q1} & \cdots & y_{qn} \end{bmatrix}$$

with probability density:

$$\left\{ (2\pi)^{\frac{(p+q)n}{2}} \prod_{i=1}^{p} (1-\rho_i^2)^{n/2} \right\}^{-1} \exp\left\{ -\frac{1}{2}\left[\sum_{i=1}^{p} \frac{1}{1-\rho_i^2} \cdot \right.\right.$$

$$\left.\left. \sum_{j=1}^{n} (x_{ij}^2 + y_{ij}^2 - 2\rho_i x_{ij} y_{ij}) + \sum_{i=p+1}^{q} \sum_{j=1}^{n} y_{ij}^2 \right]\right\} \cdot$$

$$\prod_{i,j} dx_{ij} \prod_{i,j} dy_{ij} \quad .$$

The independence hypothesis is then

H_o: $\rho_1 = \rho_2 = \cdots = \rho_p = 0$ against

H_1: at least one $\rho_i \neq 0$, $i = 1,2,\ldots,p$.

The invariant tests will involve x_{ij}'s and y_{ij}'s only through the characteristic roots of

$$(XX')^{-1} \; (XY') \; (YY')^{-1} (XY')'$$

Let us denote these characteristic roots by $(\nu_1 \leq \nu_2 \leq \cdots \leq \nu_p)$.

1.2.3 Further Reduction

This canonical form can be further reduced for our purpose (Roy and Mikhail [53]). Towards this end we put

$$\underset{q\times n}{Y} = \underset{q\times q}{\tilde{T}} \; \underset{q\times n}{L} \quad ,$$

where $\tilde{T}$ is a traingular matrix with zeros above the diagonal and L is an orthonormal matrix. L can be completed into an orthogonal matrix $\begin{matrix} q \\ n-q \end{matrix}\begin{bmatrix} L \\ M \end{bmatrix}$, so that $\begin{bmatrix} L \\ M \end{bmatrix} [L' \; M'] = [L' \; M'] \begin{bmatrix} L \\ M \end{bmatrix} = I(n)$.

Let $\underset{p\times n}{X^*} = \underset{p\times n}{X} [\underset{n\times n}{L' \; M'}]$

$= [\underset{q}{U^*} \quad \underset{n-q}{V^*}] \, p$, say .

Then the probability density of U^*, V^* and $\tilde{T}$ can be written as (Roy and Mikhail [53]):

$$\text{Constant} \prod_{i=1}^{p} (1-\rho_i^2)^{-n/2} \exp\Big[-\frac{1}{2}\Big\{\sum_{i=1}^{p} \frac{1}{1-\rho_i^2} \sum_{j=1}^{q} (u^*_{ij} - \rho^*_{ij}\, t_{ij})^2$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{i} t_{ij}^2 + \sum_{i=1}^{p} \frac{1}{1-\rho_i^2} \sum_{j=1+q}^{n} v^{*2}_{ij} \Big\}\Big] \times$$

$$\prod_i \prod_j d\, u^*_{ij} \quad \prod_i \prod_j d\, v^*_{ij} \quad \prod_{i=1}^{q} t_{ii}^{n-i} \prod_{i=1}^{q} \prod_{j=1}^{i} dt_{ij},$$

where $\rho^*_{ij} = \rho_i$, $j = 1,2,\ldots,i;\ i = 1,2,\ldots,p$

$= 0$, otherwise .

Next put

$$\underset{p\times q}{U} = \underset{p\times p}{D_{1/(1-\rho_i^2)^{1/2}}} \quad \underset{p\times q}{U^*} ,$$

$$\underset{p\times(n-q)}{V} = \underset{p\times p}{D_{1/\sqrt{1-\rho_i^2}}} \quad \underset{p\times(n-q)}{V^*} \quad \text{and}$$

$$\gamma_{ij} = \frac{\rho_{ij}}{\sqrt{1-\rho_i^2}} = \gamma_i , \quad j = 1,2,\ldots,i$$
$$i = 1,2,\ldots,p$$

$$= 0 \qquad \text{otherwise} ,$$

where $D_{1/\sqrt{1-\rho_i^2}}$ is a diagonal matrix with elements $1/\sqrt{1-\rho_i^2}$, $i = 1,\ldots,p$. Then the distribution of U, V and $\tilde{T}$ is

$$\text{Constant} \quad \exp \Big[-\frac{1}{2} \Big\{ \sum_{i=1}^{p} \sum_{j=1}^{q} (u_{ij} - \gamma_{ij}\, t_{ij})^2 + \sum_{i=1}^{q} \sum_{j=1}^{i} t_{ij}^2$$

$$+ \sum_{i=1}^{p} \sum_{j=q+1}^{n} v_{ij}^2 \Big\} \Big] dU \cdot dV \prod_{i=1}^{q} t_{ii}^{n-i}\, d\tilde{T} \ .$$

The independence hypothesis now reduces to

$$H_o: \quad \gamma_{ij} = 0$$

against H_1: at least one $\gamma_{ij} \neq 0$.

Also, the characteristic roots $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_p$ of $(UU')(VV')^{-1}$ are related to the characteristic roots $(\nu_1, \nu_2, \ldots, \nu_p)$ by (Roy and Mikhail [53]):

$$\nu = \frac{\lambda}{1 + \lambda}$$

or

$$\lambda = \frac{\nu}{1 - \nu} \ .$$

Thus λ's are the characteristic roots of

$$(S_{11}^{-1} S_{12} S_{22}^{-1} S_{12}') (I - S_{11}^{-1} S_{12} S_{22}^{-1} S_{12}')^{-1} \ .$$

1.2.4 Preliminaries and Development

Let us denote

$$\underline{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix} \quad \text{and} \quad \underline{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_p \end{bmatrix} ,$$

where, $e_1, e_2, \ldots, e_p$ are the p elementary symmetric functions of $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_p$. Then all tests based on the roots λ_i, $i = 1, \ldots, p$ or their symmetric functions will be invariant, under the groups G_1 and G_2.

Now let $\phi = \phi(U,V)$ be a test for H_0 with the acceptance region A_ϕ with boundary B_ϕ, both in the space of (U, V). We shall consider tests ϕ based on the characteristic roots of $(UU')(VV')^{-1}$ i.e. of $(S_{11}^{-1} S_{12} S_{22}^{-1} S_{12}')(I - S_{11}^{-1} S_{12} S_{22}^{-1} S_{12}')^{-1}$. The power function of test will involve, aside from the degrees of freedom, only p canonical correlations in the population viz. ρ_i, $i = 1,2,\ldots,p$ or their functions $\gamma_i = \dfrac{\rho_i}{\sqrt{1-\rho_i^2}}$, $i = 1,2,\ldots,p$.

It is easy to see that there exists an easy tie-up between the distribution problem of the MANOVA problem and the independence problem. This was exploited by Roy and Mikhail [53] to show the monotonicity property for the greatest characteristic root test for the independence problem. We shall define:

<u>Definition 5:- A test ϕ for the independence problem is said to have the monotonicity property if its power function increases monotonically as each of the p population canonical correlations increase separately.</u>

Let $U = [\underline{u}_1 \ U_1]$ and let $A_{\phi,U_1,V}$ and $B_{\phi,U_1,V}$, both in the space of $\underline{u}_1' = (u_{11}, u_{21}, \ldots, u_{p1})$ denote the sections of A_ϕ and their boundaries for different values of U_1 and V. Then we can define the properties $\mathcal{Q}$, $\mathcal{J}$ and $\mathcal{F}$ for the tests $\phi(U,V)$ as in the case of the MANOVA problem. And we have the following theorem which is implicit in Roy and Mikhail [53].

Theorem 6: If a test $\phi(U, V)$ for the independence problem has the properties $\mathcal{Q}$, $\mathcal{G}$ and $\mathcal{F}$ then it has the monotonicity property.

Proof:- We have to show that

$$\int_{A_\phi} \text{const. } \exp\left[-\frac{1}{2} \left\{ \sum_i \sum_j (u_{ij} - \gamma_{ij} t_{ij})^2 + \sum_i \sum_j t_{ij}^2 + \right.\right.$$
$$\left.\left. + \sum_i \sum_j v_{ij}^2 \right\} \right] dU\, dV \prod_{i=1}^{q} t_{ii}^{n-i}\, d\widetilde{T}$$

decreases as each of γ_{ij}'s increase separately. Or, equivalently, we have to show that

$$\int_{A_\phi^*} \text{const. } \exp\left[-\frac{1}{2} \left\{ \sum_i \sum_j u_{ij}^2 + \sum_i \sum_j t_{ij}^2 + \sum_i \sum_j v_{ij}^2 \right\} \right]$$
$$dU\, dV \prod_{i=1}^{q} t_{ii}^{n-i} \quad d\widetilde{T} ,$$

where A_ϕ^* is A_ϕ translated by $\gamma_{ij} t_{ij}$ along u_{ij} i.e. $\gamma_i t_{ii}$ along u_{ii}, $i = 1,2,\ldots,p$, decreases as each of γ_i increases separately.

By an argument similar to one used for proving Theorem 1, it can be easily shown that, for each fixed $\widetilde{T}$, the integral decreases as A_ϕ is translated by $\gamma_1 t_{11}$ along u_{11}. We can then introduce the density function of $\widetilde{T}$ and see that the integral decreases as A_ϕ is moved by $\gamma_1 t_{11}$ along u_{11}. We can then repeat this for the other γ_i's $i = 2, \ldots, p$ and we shall get the theorem.

From the method of proof of the Theorem 6 in this section the following theorem is immediate.

Theorem 7:- Any test ϕ for the independence problem with

$$A_\phi: \quad \underline{a}'\,\underline{e} \leqq \text{constant}, \quad \underline{a}' \geqq (0, 0, \ldots, 0)$$

or

$$A_\phi: \quad \underline{a}'\,\underline{\lambda} \leqq \text{constant}, \quad \underline{a}' = (a_1, a_2, \ldots, a_p) ,$$

$$a_1, a_2, \ldots, a_{p-1} \geqq 0, \quad a_p > 0 ,$$

has the monotonicity property.

Examples:- The greatest characteristic root test has

$$A_\phi : \nu_p \leqq \text{constant} ,$$

which is equivalent to

$$\lambda_p \leqq \text{constant} .$$

The likelihood ratio test has

$$A_\phi : \frac{|S|}{|S_{11}| \cdot |S_{22}|} \geqq \text{constant}$$

This can be reduced as:

$$|I - S_{11}^{-1} S_{12} S_{22}^{-1} S_{12}' | \geqq \text{constant}$$

i.e. $$\prod_{i=1}^{p} (1 - \nu_i) \geqq \text{constant}$$

i.e. $$\prod_{i=1}^{p} \left(1 - \frac{\lambda_i}{1+\lambda_i}\right) \geqq \text{constant}$$

i.e. $$\prod_{i=1}^{p} (1 + \lambda_i) \leqq \text{constant}$$

i.e. $$1 + e_1 + \ldots + e_p \leqq \text{constant} .$$

This proves that the likelihood ratio test has the monotonicity property.

1.3 Testing the Equality of Two Dispersion Matrices.

1.3.1 The Model and the Hypothesis

Let $\underset{p\times(n_1+1)}{Z_1}$ and $\underset{p\times(n_2+1)}{Z_2}$ be two samples of sizes $(n_1 + 1)$ and $(n_2 + 1)$ from $N(\underset{p\times1}{\zeta_1}, \underset{p\times p}{\Sigma_1})$ and $N(\underset{p\times1}{\zeta_2}, \underset{p\times p}{\Sigma_2})$ respectively, where Σ_1 and Σ_2 are both symmetric and p.d. We are interested in testing the hypothesis

$$H_o: \quad \Sigma_1 = \Sigma_2$$

against various alternatives.

This problem is invariant under the group of nonsingular transformations of structure

$$Z_1^* = \underset{p\times p}{C} Z_1 , \qquad Z_2^* = \underset{p\times p}{C} Z_2 , \quad C \text{ nonsingular},$$

and under the group of translations of type

$$Z_1^* = Z_1 + B_1 , \qquad Z_2^* = Z_2 + B_2 ,$$

B_1 and B_2 real matrices. The characteristic roots of

$$S_1S_2^{-1} = [Z_1Z_1' - n_1 \bar{z}_1 \bar{z}_1'] [Z_2Z_2' - n_2 \bar{z}_2 \bar{z}_2']^{-1}$$

where $\underset{p\times1}{\bar{z}_i} = \begin{bmatrix} \bar{z}_{i1} \\ \vdots \\ \bar{z}_{ip} \end{bmatrix}$ and $\bar{z}_{ij} = \sum_{j=1}^{n_i+1} z_{ij}/(n_i+1)$

$$i = 1,2; \quad j = 1,2,\ldots,p ,$$

form a set of maximal invariants under these groups of transformations. Also the maximal invariants under the groups of transformations induced in the parametric space are the characteristic roots of $(\Sigma_1\Sigma_2^{-1})$, say,

$\gamma_1 \leqq \cdots \leqq \gamma_p$. The hypothesis of equal dispersion matrices then becomes

$$H_o: \gamma_1 = \gamma_2 = \cdots\cdots = \gamma_p .$$

Roy and Gnanadesikan [52] have considered this hypothesis against several alternatives. Out of these alternatives Mikhail [39] has considered the following four alternatives:

H_1 : All γ_i's > 1 .

H_2 : All γ_i's < 1.

H_3 : The largest $\gamma_i = \gamma_p > 1$.

H_4 : The smallest $\gamma_i = \gamma_1 < 1$.

For each of these four alternatives Roy and Gnanadesikan [52] have given procedures, some of which are three-decision procedures. Mikhail [39] has obtained different kinds of monotonicity properties for these procedures. We shall show that in some cases we can obtain very broad classes of procedures having properties he has obtained and in other cases classes of procedures having somehwat weaker monotonicity properties. We shall consider only invariant tests and can restrict consideration to the following canonical form.

1.3.2 <u>Canonical Form</u>

Let $\underset{p \times n_1}{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n_1} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn_1} \end{bmatrix}$ and $\underset{p \times n_2}{Y} = \begin{bmatrix} y_{11} & \cdots & y_{1n_2} \\ \vdots & & \vdots \\ y_{p1} & \cdots & y_{pn_2} \end{bmatrix}$

be two matrices of random variables with probability density

$$\text{constant} \prod_{i=1}^{p} (\gamma_i)^{-n_1/2} \exp\left[-\frac{1}{2} \operatorname{tr}\left(D_{1/\gamma_i} XX' + YY'\right)\right] dX\, dY .$$

Under this canonical form the hypothesis is $\gamma_1 = \gamma_2 = \cdots = \gamma_p = 1$, and all the tests will involve X and Y only through the characteristic roots of $(XX')(YY')^{-1}$ say $\lambda_1 \leqq \cdots \leqq \lambda_p$. Let

$$\underline{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{bmatrix} \quad \text{and} \quad \underline{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_p \end{bmatrix},$$

where e_j, $j = 1,2,\ldots,p$ are the p elementary symmetric functions of $\lambda_1, \ldots, \lambda_p$.

1.3.3 The Development

Now we shall consider the problems of testing H_o against H_1 and H_2 as case 1 and those of testing H_o against H_3 and H_4 as case 2.

Case 1.

To test H_o: $\gamma_1 = \gamma_2 = \cdots = \gamma_p = 1$

against H_1: all γ's > 1

and against H_2: all γ's < 1.

For these problems Roy and Gnanadesikan [52] have given the following three-decision procedures:

i) Accept H_o against H_1 if $\underset{max}{ch}\,(XX')(YY')^{-1} \leqq \mu_1$, a constant ;
   Accept H_1 against H_o if $\underset{min}{ch}\,(XX')(YY')^{-1} > \mu_1$
   Make no decision otherwise

ii) Accept H_o against H_2 if $\underset{min}{ch}\,(XX')(YY')^{-1} > \mu_2$, a constant
   Accept H_2 against H_o if $\underset{max}{ch}\,(XX')(YY')^{-1} \leqq \mu_2$
   Make no decision otherwise.

Constants μ_1 and μ_2 are determined by the condition that the probability of accepting H_o when it is true equals a given number $1-\alpha$.

Mikhail [39] has shown that the probability of accepting H_o against H_1, when H_1 is true, decreases monotonically as each of the noncentrality parameters increase separately; and the probability of accepting H_2 against H_o, when H_2 is true, increases montonically as γ's decrease separately. We shall call these the '<u>strong monotonicity properties</u>'. We shall, however, concern ourselves with a weaker property, namely, monotonicity of the above mentioned probabilities when the γ_i's are all equal to say γ and γ increases or decreases. We shall call this the <u>weak monotonicity property</u>.

We shall, in this section, give procedures ϕ characterized only by regions over which H_o or H_2 may be accepted; these regions will be denoted by A_ϕ. The procedures themselves may be two-decision procedures or three-decision procedures. More specifically, we shall prove the following:

<u>Theorem 8</u>:- <u>In the canonical form of section 1 replace s_1 by n_1, n by n_2 and u by p. Let A_ϕ be the acceptance region of any test ϕ for this MANOVA problem which has the properties $\mathcal{Q}$, $\mathcal{J}$ and $\mathcal{F}$. Then all the procedures for testing H_o against H_1 and H_o against H_2, which</u>

<u>(i) accept H_o against H_1 if $(X, Y) \in A_\phi$, ~~reject it otherwise,~~</u>

<u>(ii) accept H_2 against H_o if $(X, Y) \in A_\phi$, ~~reject it otherwise,~~</u>

<u>have the weak monotonicity property.</u>

<u>Proof</u>:- We are concerned with

$$\text{Constant} \int_{A_\phi} \prod_{i=1}^{p} \gamma_i^{-n/2} \exp \left[- \frac{1}{2} \operatorname{tr} \left(D_{1/\gamma_i} XX' + YY' \right) \right] dX\, dY$$

This is the same as

$$\text{Const.} \int_{A^*_\emptyset} \exp\left[- \frac{1}{2}\left(\sum_{i=1}^{p} \sum_{j=1}^{n_1} x_{ij}^2 + \sum_{i=1}^{p} \sum_{j=1}^{n_2} y_{ij}^2 \right) \right] dX \cdot dY ,$$

where $A^*_\emptyset$ is the same as the domain $A_\emptyset$ expanded by quantity $\sqrt{\gamma_i}$ in the directions $(x_{i1}, x_{i2}, \ldots, x_{i,n_1})$, $i = 1,2,\ldots,p$.

But we are concerned with alternatives $\gamma_1 = \gamma_2 = \cdots = \gamma_p = \gamma$, which is > 1 in testing H_o against H_1 and < 1 in testing H_o against H_2. Therefore we may regard $A^*_\emptyset$ as the domain $A_\emptyset$ expanded by the quantity $\sqrt{\gamma}$ in directions $(x_{1j}, x_{2j}, \ldots, x_{pj})$, $j = 1, 2, \ldots, n_1$.

Now these expansions are cumulative. It is, therefore, sufficient to see what happens when $A_\emptyset$ is expanded by quantity $\sqrt{\gamma}$ along $(x_{11}, x_{21}, \ldots, x_{p1})$.

In the notation of section 1 with proper changes, the region $A_{\emptyset,X_1,Y}$ in the space of $(x_{11}, x_{21}, \ldots, x_{p1})$ is an ellipsoid in the space of $(x_{11},\ldots,x_{p1})$. This ellipsoid can be referred to its principal axes. The expansion of $A_{\emptyset,X_1,Y}$ along $(x_{11}, x_{21},\ldots,x_{p1})$ by quantity $\sqrt{\gamma}$ is then equivalent to expansion of this ellipsoid by quantity $\sqrt{\gamma}$ along its principal axes.

Thus

$$A^*_\emptyset \subset A_\emptyset \qquad \text{if } \gamma > 1$$

and

$$A^*_\emptyset \supset A_\emptyset \qquad \text{if } \gamma < 1 .$$

Thus, the probability of accepting H_o against H_1 and the probability of accepting H_2 against H_o ~~both~~ resp. decreases and increases monotonically as $\gamma(= \gamma_1 = \gamma_2 = \cdots = \gamma_p)$, respectively, increases or decreases.

Examples:- Procedures, which accept H_o against H_1 (and reject it otherwise) and which accept H_2 against H_o (and reject it otherwise) over

$$A_\phi : \underline{a}' \underline{e} \leqq \text{constant}, \quad \underline{a}' = (a_1, a_2, \ldots, a_p),$$
$$a_j \geqq 0, \quad j = 1, 2, \ldots p$$

or

$$A_\phi : \underline{a}' \underline{\lambda} \leqq \text{constant}, \quad \underline{a}' = (a_1, \ldots, a_p)$$
$$a_j \geqq 0, \quad j = 1, 2, \ldots, p-1; \quad a_p > 0$$

have the weak monotonicity property.

Case 2

To test $H_o : \gamma_1 = \gamma_2 = \ldots = \gamma_p = 1$

against H_3 : the largest $\gamma_i = \gamma_p > 1$

or against H_4 : the smallest $\gamma_i = \gamma_1 < 1$.

For these problems the following test procedures are known (Roy and Gnanadesikan [52]):

(i) Accept H_o against H_3 if $\underset{max}{ch} (XX')(YY')^{-1} \leqq$ const. and reject it otherwise.

(ii) Accept H_4 against H_o if $\underset{max}{ch} (XX')(YY')^{-1} \leqq$ const. and reject it otherwise.

Mikhail [39] has shown that the power functions of both these tests are monotonically increasing functions of the noncentrality parameters, i.e. the power functions increase monotonically as γ_p increases or γ_1 decreases respectively. We shall prove the following:

Theorem 9: With the changes in the notation of the canonical form of section 1, as suggested in Theorem 8, if we accept H_o against H_3 over A_ϕ of Theorem 8 or accept H_4 against H_o over A_ϕ, and reject them otherwise then the power function of ϕ for

(i) testing H_o against H_3 increases monotonically as γ_p increases,

(ii) testing H_o against H_4 increases monotonically as γ_1 decreases.

Proof:- Again we are concerned with

$$\text{const.} \int_{A_\phi} \prod_{i=1}^{p} \gamma_i^{-n_1/2} \exp \left[-\frac{1}{2} \operatorname{tr}\left(D_{1/\gamma_i} XX' + YY' \right) \right] dX\, dY .$$

For given values $(\gamma_1^*, \ldots, \gamma_{p-1}^*)$ of $(\gamma_1, \ldots, \gamma_{p-1})$ we may write this as

$$\text{const.} \int_{A_\phi^*} \exp \left[-\frac{1}{2} \left(\sum_{i=1}^{n_1} x_{pi}^2 + \sum_{i=1}^{p-1} 1/\gamma_i \sum_{j=1}^{n_1} x_{ij}^2 + \sum_{i=1}^{p} \sum_{j=1}^{n_2} y_{ij}^2 \right) \right] dX\, dY .$$

Where A_ϕ^* is A_ϕ expanded by $\sqrt{\gamma_p} > 1$ along $(X_{p1}, X_{p2}, \ldots, X_{p,n_1})$. The expansions being cumulative we may give them successively. It is enough to show that the integral decreases by an expansion $\sqrt{\gamma_p}$ along X_{p1}.

Now $A_{\phi, X_1, Y}$ is interior of an ellipsoid in the space of $(x_{11}, x_{21}, \ldots, x_{p1})$ for any given value $(\gamma_1^*, \gamma_2^*, \ldots, \gamma_{p-1}^*)$ of $(\gamma_1, \gamma_2, \ldots, \gamma_p)$. An expansion of this region by $\sqrt{\gamma_p}$ along x_{p1} can be considered as a result of a number of expansions $\ell_1\sqrt{\gamma_p}, \ell_2\sqrt{\gamma_p}, \ldots, \ell_p\sqrt{\gamma_p}$ along the principal axes of the ellipsoid, where $(\ell_1, \ell_2, \ldots, \ell_p)$ are the direction cosines of x_{p1} with respect to the principal axes. Since $\sqrt{\gamma_p} > 1$ it follows that

$$A^*_{\phi,X_1,Y} \subset A_{\phi,X_1,Y} .$$

Thus the power function of each test ϕ, for testing H_o against H_3, given by the previous theorem increases as γ_p increases.

The proof for the monotonicity of the power functions of the tests ϕ for testing H_o against H_4 is similar.

Examples: All tests ϕ which accept H_o against H_3 over A_ϕ and reject it otherwise or accept H_4 against H_o over A_ϕ and reject it otherwise, where

$$A_\phi : \underline{a}' \underline{e} \leqq \text{constant} , \quad \underline{a}' = (a_1,\ldots,a_p) ,$$

$$a_j \geqq 0, \quad j = 1,2,\ldots, p$$

or

$$A_\phi : \underline{a}' \lambda \leqq \text{constant}, \quad \underline{a}' = (a_1, a_2, \ldots, a_p) ,$$

$$a_j \geqq 0, \quad j = 1, 2, \ldots, p-1, \quad a_p > 0$$

have power functions which increase monotonically as, respectively, γ_p increases or γ_1 decreases.

1.4 Testing Whether a Dispersion Matrix Equals a Given Matrix.

1.4.1 The Model and the Hypothesis.

Let $\underset{p\times(n+1)}{Z}$ be a sample of $(n+1)$ observations from a $N(\underset{p\times 1}{\underline{\zeta}}, \underset{p\times p}{\Sigma})$ where Σ is a symmetric p.d. matrix. We are interested in testing

$$H_o : \quad \Sigma = \Sigma_o \quad ,$$

Σ_o being a given symmetric p.d. matrix.

In many respects this problem is similar to the "equidispersion" problem; and especially so in case of the alternatives to be considered and the monotonicity properties of the power functions.

It is well known and easy to see that the problem can be reduced by invariance and we may restrict our consideration to the following canonical form:

1.4.2 Canonical Form

Let $X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{bmatrix}$ be a matrix of random variables with probability density.

$$\text{Constant} \prod_{i=1}^{p} \gamma_i^{-n/2} \exp \left[-\frac{1}{2} \operatorname{tr} D_{1/\gamma_i} XX' \right] dX .$$

We are interested in testing

$$H_o : \gamma_1 = \gamma_2 = \cdots = \gamma_p = 1$$

against various alternatives:

H_1 : all γ_i's > 1

H_2 : all γ_i's < 1

H_3 : the largest $\gamma_i = \gamma_p > 1$

H_4 : the smallest $\gamma_i = \gamma_1 < 1$

The tests $\phi(X)$ are restricted to involve X only through the characteristic roots $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_p$ of (XX').

1.4.3 The Development

All the results in the previous section on the tests of equality of two dispersion matrices can be extended to this section, that is, results on the weak monotonicity when testing against H_1 or H_2 and the monotonicity in γ_p and γ_1, respectively, when testing against

H_3 or H_4. We shall not consider these results further, but instead we shall derive some strong monotonicity properties for a single class of tests of H_o against H_1.

Theorem 10: All procedures ϕ with

$$A_\phi: \underline{a}'\underline{e} \leq \text{constant}, \quad \underline{a}' = (a_1, \ldots, a_p), \quad a_j \geq 0,$$

$j = 1,2,\ldots, p$, for testing H_o against H_1 have power functions which increase monotonically as $\gamma_1, \gamma_2, \ldots, \gamma_p$ increase separately.

Proof:- We are interested in showing that

$$\int_{A_\phi} \text{constant} \prod_{i=1}^{p} \gamma_i^{-n/2} \exp\left[-\frac{1}{2} \text{ tr } D_{1/\gamma_i} XX'\right] dX$$

which is the same as

$$\int_{A_\phi^*} \text{constant} \exp\left[-\frac{1}{2} \text{ tr } XX'\right] dX \quad ,$$

where A_ϕ^* is A_ϕ expanded by quantity $\sqrt{\gamma_i}$ along $(x_{i1}, x_{i2},\ldots,x_{in})$, $i = 1, 2, \ldots, p$. It is sufficient to show that the integral decreases when A_ϕ is expanded by $\sqrt{\gamma_1}$ along $(x_{11}, x_{12}, \ldots, x_{1n})$. Let

$$X = p\begin{bmatrix} \underline{x}_1' \\ \vdots \\ \underline{x}_p' \end{bmatrix} = \begin{bmatrix} \underline{x}_1' \\ X_1' \end{bmatrix} \begin{matrix} 1 \\ p-1 \end{matrix}$$

and consider the sections $A_{\phi,\underline{x}_1}$ of ϕ. We shall show that these are ellipsoids in the n-dimensional space of $(x_{11}, \ldots, x_{1n})$. This will prove the previous contention.

Consider the equation

$$| XX' - \lambda I | = 0 .$$

This can be written in two ways, as before, as

$$\lambda^p - \not{e}_1 \lambda^{p-1} + \not{e}_2 \lambda^{p-2} - \dots + (-1)^p \not{e}_p = 0$$

and

$$\lambda^p - e_1 \lambda^{p-1} + e_2 \lambda^{p-2} - \dots + (-1)^p e_p = 0 ,$$

where $\not{e}_j$ is the sum of all $\binom{p}{j}$ j-rowed principal minors of

$$X X' = \begin{bmatrix} \underline{x}_1' \underline{x}_1 & \underline{x}_1' X_1 \\ X_1' \underline{x}_1 & X_1' X_1 \end{bmatrix} ,$$

$j = 1, 2, \dots, p$. It is easy to see that given X_1, e_j will be a homogeneous quadratic function of x_{1j}, plus a constant. Thus $B_{\emptyset, X_1}$ will be a homogeneous quadric in the n-dimensional space of $(x_{11}, \dots, x_{1n})$. Next consider matrix

$$\begin{bmatrix} c^2 \underline{x}_1' \underline{x}_1 & c\, \underline{x}_1' X_1 \\ c\, X_1' \underline{x}_1 & X_1' X_1 \end{bmatrix} , \qquad 0 < c < 1 .$$

Any $\not{e}_j$, $j = 1, 2, \dots, p$, of this will contain $\binom{p-1}{j-1}$ principal minors which contain the first row and column, whereas, the remaining will not contain these. Now each of the $\binom{p-1}{j-1}$ principal minors which contains the first row and column is c^2 times the corresponding minor when $c = 1$, and each of the remaining $\binom{p}{j} - \binom{p-1}{j-1}$ is the same as the corresponding minor when $c = 1$. Thus when $0 < c < 1$, any e_j, $j = 1, 2, \dots, p$, of $\begin{pmatrix} c\, \underline{x}_1' \\ X_1' \end{pmatrix} (c \cdot \underline{x}_1 \;\; X_1)$ is less than the corresponding e_j of XX'. Therefore, for $a_j \geqq 0$, $j = 1, 2, \dots, p$, and

$$A_\emptyset : \underline{a}' \underline{e} \leqq \text{constant}$$

ϕ has the property , i.e. if $\underline{x}_1 \in A_{\phi,X_1}$ then $c\,\underline{x}_1 \in A_{\phi,X_1}$, $0 \leq c \leq 1$.

Next consider a line L in the space of $(x_{11},\ldots,x_{1n})$ with equations

$$\frac{x_{11}}{\ell_1} = \frac{x_{12}}{\ell_2} = \ldots = \frac{x_{1n}}{\ell_n} \; (= r) \; .$$

Consider the intersection of L and $A_{\phi\, X_1}$. We notice that this is a segment of the line itself. Now if a point $(\ell_1 r, \ell_2 r, \ldots, \ell_n r)$ on this line belongs to A_{ϕ,X_1}, the section of the acceptance region, then for given X_1, some e_j, $j = 1, 2, \ldots, p$ of

$$\begin{bmatrix} r \cdot \ell' \\ \hdashline X_1' \end{bmatrix} \left[r \cdot \ell \;\vdots\; X_1 \right] = \begin{bmatrix} r^2 \ell' \ell & r\,\ell' X_1 \\ r\, X_1' \ell & X_1 X_1' \end{bmatrix}$$

is finite; also it is positive. Now this e_j will be the sum of $\binom{p-1}{j-1}$ j-rowed principal minors containing the first row and column and the remaining j-rowed principal minors not containing the first row and column. Thus $0 \leqq e_j < \infty$ implies that $0 \leqq r^2 K_1 + K_2 < \infty$, where K_1, K_2 are positive and finite. This implies that r is finite. Therefore, the section A_{ϕ,X_1} is interior of an ellipsoid in the space of $(x_{11},\ldots,x_{1n})$. This completes the proof of Theorem 10.

It is easy to prove on parallel lines that all procedures ϕ for testing H_0 against H_2, which accept over A_ϕ of Theorem 10, have analogous strong monotonicity property.

CHAPTER II

SOME PROPERTIES OF PERCENTAGE POINTS AND PROBABILITY INTEGRALS OF SOME RANDOM VARIABLES

2.0 Introduction and Summary

In this chapter we shall obtain a number of properties of percentage points and integrals associated with the distributions of some random variables. Some of these properties are generally known from charts and tables of these quantities; but proofs of these are not known to have appeared in the related literature. In the first section we shall study the monotonicity properties of the percentage points and some integrals associated with a noncentral chi-square distribution with respect to various parameters and obtain some order relations for the integrals. In the second section we shall do the same for noncentral F-distribution and pose some unsolved problems in this area. In the third section we shall extend some of the results of the second section to the distribution of Hotelling's T^2-statistic and investigate some monotonicity properties for the percentage points of a class of statistics suggested in Chapter I. The methods of investigation and the proofs are based on various results in the theory of hypotheses testing, for which we may refer to Lehmann [32, 35].

2.1 Some properties of percentage points and integrals connected with chi-square tests.

Many problems of testing hypotheses in univariate and multivariate inference use statistics with central or noncentral chi-square distributions (Kendall and Stuart [29], Patnaik [42]). The percentage points, as well as incomplete integrals, are available in extensive tables for the chi-square distribution (Bancroft [5], Greenwood [16]). Tabulation for the noncentral chi-square distribution, however, is not extensive. In this section we shall list and prove some properties of the percentage points and the integrals connected with the noncentral chi-square distribution. These properties have been generally recognized from the tables and charts, but proofs have not heretofore appeared.

Definition 1: A stochastic variable $X \geqq 0$ with probability density

$$\chi^2(x;\ n,\ \delta) = e^{-\delta^2/2} \sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\delta^2}{2}\right)^r \chi^2(x;\ n+2r,\ 0)\ ,$$

where

$$\chi^2(x;\ \nu,\ 0) = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})}\, e^{-\frac{1}{2}x}\, x^{\frac{\nu}{2}-1}\ ,\quad x \geqq 0$$

$$= 0\ ,\quad x < 0$$

is said to have the $\chi^2(n, \delta)$-distribution; i.e., the noncentral chi-square distribution with n degrees of freedom and the noncentrality parameter δ, or equivalently, X is said to be distributed as a $\chi^2(n, \delta)$-variate. When $\delta = 0$ the distribution is said to be the (central) chi-square distribution with n degrees of freedom. We shall denote

$$I_{\chi^2}(\mu;\ n,\ \delta) = \int_0^{\mu} \chi^2(x;\ n,\ \delta)\, d\,x\ .$$

It is well known that if X_i are independently normally distributed random variables, with common variance σ^2 and means $E(X_i) = a_i$, $i = 1, 2, \ldots, n$, then $\sum_1^n X_i^2/\sigma^2$ is distributed as a $\chi^2(n, \delta)$-variate, where $\delta^2 = \sum_1^n a_i^2/\sigma^2$. Conversely, a random variable X, distributed as a $\chi^2(n, \delta)$-variate may be expressed as a sum of squares of n independently normally distributed random variables with unit variance.

Theorem 1: If α is a fixed number, $0 < \alpha < 1$, and $I_{\chi^2}(\mu_\alpha, n, \delta) = \int_0^{\mu_\alpha} \chi^2(x; n, \delta)\, dx = 1 - \alpha$ then μ_α is an increasing function of n for each fixed δ and an increasing function of δ for each fixed n .

Proof: Consider a random variable $X = X_1^2 + \ldots + X_n^2$, where X_i, $i = 1, 2, \ldots, n$, are independently normally distributed random variables with variance unity and means $E\, X_1 = \delta$, $E\, X_i = 0$, $i = 2, 3, \ldots, n$. Let Y be a $N(0,1)$ variable independent of X. Then $X + Y^2$ is a $\chi^2\ (n+1, \delta)$-variate. Now the event $X + Y^2 \leqq b$ implies the event $X \leqq b$, where $b > 0$ is constant. From this the first half of the theorem easily follows.

For the second half of the theorem we observe that

$$I_{\chi^2}(b; n, \delta) = \int_0^b \chi^2(t; n-1, 0)\, dt \int_{-\sqrt{b-t}+\delta}^{\sqrt{b-t}+\delta} \frac{e^{-\frac{1}{2}x^2}\, dx}{\sqrt{2\pi}} .$$

The integral on the r.h.s. decreases as δ increases. The proof of the theorem is complete.

In the next theorem we shall consider the behaviour of $I_{\chi^2}(\mu;n,\delta)$ for varying values of n and δ when μ is selected by some boundary condition.

<u>Theorem 2:</u> <u>If n and n^* are positive integers, $n^* > n$, then</u>

$$I(\mu; n, \Delta^*) < I_{\chi^2}(\mu^*; n^*, \Delta^*) \quad , \quad \Delta^* > \Delta$$

$$> I_{\chi^2}(\mu^*; n^*, \Delta^*) \quad , \quad \Delta^* < \Delta \ ,$$

<u>where Δ , Δ^* are positive numbers and μ, μ^* satisfy</u>

$$I_{\chi^2}(\mu; n, \Delta) = I_{\chi^2}(\mu^*; n^*, \Delta) \ .$$

<u>Also, for $\Delta = O(1)$ and $\Delta^* = O(1)$, as $n \rightarrow \infty$</u>

$$I_{\chi^2}(\mu; n, \Delta^*) - I_{\chi^2}(\mu; n, \Delta) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty \ .$$

<u>Proof:-</u> Let X be a $\chi^2(n, \delta)$-variate and Y be a $\chi^2(n^*-n, 0)$-variate independent of X. Consider the problem of testing

$$H_0 : \delta = \Delta$$

against $H_1 : \delta = \Delta^* > \Delta$.

For this problem there exists a U.M.P. test given by a critical region

$$X \geqq \text{constant} \ .$$

To see this let $p_\delta(x,y)$ denote the joint probability density of X and Y and notice that

$$p_\delta(x,y) = \chi^2(x; n, \delta) \cdot \chi^2(y; n^*-n, 0)$$

By the Neyman-Pearson lemma the U.M.P. test rejects when

$$\frac{p_{\Delta^*}(x,y)}{p_\Delta(x,y)} \geqq \text{constant}$$

i.e.

$$\frac{\chi^2(x; n, \Delta^*)}{\chi^2(x; n, \Delta)} \geqq \text{constant}$$

But we have a

<u>Lemma 1: The ratio</u> $\dfrac{\chi^2(x;\ n,\ \Delta^*)}{\chi^2(x;\ n,\ \Delta\ )}$ <u>is a monotonically increasing function of x provided $\Delta^* > \Delta$. (Lehmann [35])</u> .

Thus the rejection region of the U.M.P. test is

$$X \geqq \text{constant} \quad ,$$

the constant being chosen to satisfy the level condition. The test with critical region $X \geqq$ constant is, therefore, more powerful than the test with critical region $X + Y \geqq$ constant. Now, $X + Y$ is distributed as a $\chi^2(n^*,\ \delta)$-variate. Thus if

$$I_{\chi^2}(\mu;\ n,\ \Delta) \quad = \quad I_{\chi^2}(\mu^*;\ n^*,\ \Delta)$$

then $$I_{\chi^2}(\mu;\ n,\ \Delta^*) \leqq I_{\chi^2}(\mu^*;\ n^*,\ \Delta^*) \ , \quad \Delta^* \geqq \Delta$$

$$\geqq I_{\chi^2}(\mu^*;\ n^*,\ \Delta^*) \ , \quad \Delta^* < \Delta \ .$$

To prove the second part we shall use an approximation of a $\chi^2(n,\ \delta)$-variate by a $\chi^2(\nu,\ 0)$-variate, due to Patnaik [42] . According to this approximation, the incomplete integrals of a $\chi^2(n,\ \delta)$-variate can be approximated by those of a $\rho \cdot \chi^2(\nu,\ 0)$-variate, where

$$\rho \ = \ 1 + \frac{\delta}{n + \delta} \ , \qquad \nu = n + \frac{\delta^2}{n + 2\delta}$$

and the error of approximation will tend to zero as n tends to infinity, provided that $\delta = O(1)$. Thus for large n and $\Delta,\ \Delta^* = O(1)$, we can approximate both $I_{\chi^2}(\mu;\ n,\ \Delta)$ and $I_{\chi^2}(\mu;\ n,\ \Delta^*)$ by $I_{\chi^2}(\mu;\ n,\ 0)$.

This proves the result.

This theorem explains the nature of the power curves of a chi-square test, depending upon noncentral chi-square distribution, e.g. analysis of variance with known population variance. For fixed n, a curve of this family starts at $\delta = 0$ with power equal to the level of significance, and increases monotonically to 1 as $\delta \to \infty$. The curve for $n^* > n$ lies completely below the curve for n. As n increases the curves gradually become parallel to the abscissa.

The next theorem will prove a property of incomplete chi-square integrals.

Theorem 3: Let n, n^* be positive integers, $n^* > n$, and let μ and μ^* satisfy

$$I_{\chi^2}(\mu;\, n,\, 0) = I_{\chi^2}(\mu^*;\, n^*,\, 0) .$$

Then

$$I_{\chi^2}(c\mu;\, n,\, 0) > I_{\chi^2}(c\mu^*,\, n^*,\, 0) , \qquad c < 1$$

$$< I_{\chi^2}(c\mu^*,\, n^*,\, 0) , \qquad c > 1 .$$

Proof:- Let $X_1, X_2, \ldots, X_n;\ Y_1, Y_2, \ldots, Y_{n^*-n}$ be independent normally distributed random variables with zero means and common variance σ^2. Consider the problem of testing

$$H_o : \sigma^2 \leqq 1$$

against $\quad H_1 : \sigma^2 > 1 .$

For this problem there exists a U.M.P. test which rejects H_o when

$$\sum_{i=1}^{n} X_i^2 + \sum_{j=1}^{n^*-n} Y_j^2 \geqq \text{constant} ,$$

the constant being chosen so as to satisfy the level condition. The U.M.P test is certainly more powerful than the test with the same level of significance and with critical region

$$\sum_{i=1}^{n} X_i^2 \geqq \text{constant}$$

when the alternative is $\sigma^2 = \frac{1}{c}$, $c < 1$. This proves the first inequality; the second may be proved analogously by considering the uniformly least powerful test of H_1 against H_o.

Now the general nature of the incomplete noncentral chi-square integral is similar to that of the incomplete chi-square integral. It would be interesting to know whether there exists a property for non-central chi-square integral analogous to that given by Theorem 3 for chi-square integral. More specifically we pose the following problem:

Let $n^* > n$ and let

$I_{\chi^2}(\mu; n, \delta) = I_{\chi^2}(\mu^*; n^*, \delta)$, δ fixed then does there exist any order between $I_{\chi^2}(c\mu; n, \delta)$ and $I_{\chi^2}(c\mu^*; n, \delta)$ for varying c?

The property of incomplete chi-square distribution given by Theorem 3 can be readily translated into a similar property of incomplete gamma integrals.

2.2 Some Properties of Percentage Points and Integrals connected with the Variance-ratio Test.

Definition: A stochastic variate $X \geqq 0$, with probability density function

$$\mathfrak{F}(x; n_1, n_2, \delta) = e^{-\delta^2/2} \sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\delta^2}{2}\right)^r \mathfrak{F}(x; n_1+2r, n_2, 0) ,$$

where

$$\mathcal{F}(x;\ \nu_1,\ \nu_2,\ 0) = \frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\frac{\nu_1}{2},\ \frac{\nu_2}{2})} \cdot \frac{x^{\frac{\nu_1}{2}-1}}{(1+\frac{\nu_1}{\nu_2}x)^{\frac{\nu_1+\nu_2}{2}}},$$

is said to be distributed as $\mathcal{F}(n_1,\ n_2,\ \delta)$- variate or, equivalently, is said to have the noncentral F-distribution with n_1, n_2 degrees of freedom and the noncentrality parameter δ.

It is then well known that the distribution of a $\mathcal{F}(n_1,\ n_2,\ \delta)$-variate is the same as that of a stochastic variate $\frac{n_2}{n_1}\ \frac{Y}{Z}$, where Y and Z are independently distributed random variates with probability density functions $\chi^2(y;\ n_1,\ \delta)$ and $\chi^2(z;\ n_2,\ 0)$ respectively (Kendall and Stuart [29], Patnaik [42]).

Now let α be a fixed number $0 < \alpha < 1$, and let

$$I_{\mathcal{F}}(\mu;\ n_1,\ n_2,\ \delta) = \int_0^{\mu} \mathcal{F}(x;\ n_1,\ n_2,\ \delta)\ dx = 1 - \alpha .$$

Then values of $\mu(n_1,\ n_2)$ for $\delta = 0$, $\alpha = 0.05$ or 0.01 and various values of n_1 and n_2, are very widely available. The tables of $\mu(n_1,\ n_2)$ do not show any monotonicity with respect to n_1 and n_2. The following theorem gives a monotonicity property of $\mu(n_1,\ n_2)$.

Theorem 4: $\mu^* = \frac{n_1}{n_2}\mu(n_1,\ n_2)$ is an increasing function of n_1 and and a decreasing function of n_2.

Proof:- Let Y, Y_1, Z, Z_1 be independent random variables with respective probability densities $\chi^2(y;\ n_1,\ \delta)$, $\chi^2(y_1;\ 1,\ 0)$, $\chi^2(z;\ n_2, 0)$ and $\chi^2(z_1;\ 1,\ 0)$. Then the prob. densities of $\frac{n_2}{n_1}\ \frac{Y}{Z}$, $\frac{n_2}{n_1+1}\ \frac{Y+Y_1}{Z}$,

$\frac{Y}{Z + Z_1}$ $\frac{n_2 + 1}{n_1}$ are respectively, $\mathcal{F}(\cdot\,; n_1, n_2, \delta)$ $\mathcal{F}(\cdot\,, n_1+1, n_2, \delta)$ and $\mathcal{F}(\cdot\,, n_1, n_2+1, \delta)$. Now

$$\frac{Y + Y_1}{Z} \leqq k \quad \text{implies that} \quad \frac{Y}{Z} \leqq k$$

and

$$\frac{Y}{Z} \leqq k \quad \text{implies that} \quad \frac{Y}{Z + Z_1} \leqq k \ ,$$

where k is a constant. Therefore

$$\Pr\left\{\frac{Y + Y_1}{Z} \leqq k\right\} < \Pr\left\{\frac{Y}{Z} \leqq k\right\}$$

and

$$\Pr\left\{\frac{Y}{Z} \leqq k\right\} < \Pr\left\{\frac{Y}{Z + Z_1} \leqq k\right\} .$$

The theorem follows from this.

We shall next consider a property of incomplete integrals of $\mathcal{F}(n_1, n_2, 0)$-variates.

Theorem 5:- Let n_1, n_2, n_1^*, n_2^* be positive integers and let

$$I_{\mathcal{F}}(\mu; n_1, n_2, 0) = I_{\mathcal{F}}(\mu^*; n_1^*, n_2^*, 0) ;$$

then

$$I_{\mathcal{F}}(c\mu; n_1, n_2, 0) > I_{\mathcal{F}}(c\mu^*; n_1^*, n_2^*, 0) , \quad c < 1$$

$$< I_{\mathcal{F}}(c\mu^*; n_1^*, n_2^*, 0) , \quad c > 1 ,$$

provided $n_1^* \geqq n_1$, $n_2^* \geqq n_2$, and $n_1^* + n_2^* > n_1 + n_2$.

Proof:- Let $X_1, X_2, \ldots, X_{n_1}, X_{n_1+1}, \ldots, X_{n^*}$; $Y_1, \ldots, Y_{n_2}, Y_{n_2+1}, \ldots, Y_{n^*_2}$ be independently normally distributed random variables with zero means and variances σ_x^2, σ_y^2 respectively. Consider the problem of testing

$$H_o : \frac{\sigma_x^2}{\sigma_y^2} = 1$$

against $$H_1 : \frac{\sigma_x^2}{\sigma_y^2} > 1 .$$

For this problem there exists a U.M.P. unbiased test with critical region

$$\frac{\sum_{i=1}^{n_1^*} X_i^2}{\sum_{j=1}^{n_2^*} Y_j^2} \geqq \text{constant} .$$

The power of this test against the alternative $\frac{\sigma_x^2}{\sigma_y^2} = \frac{1}{c} > 1$ will be greater than that of the unbiased test which rejects when

$$\frac{\sum_{i=1}^{n_1} X_i^2}{\sum_{j=1}^{n_2} Y_j^2} \geqq \text{constant} .$$

This proves the theorem.

Whether there exists any similar property for a noncentral F-distribution is an open question. Also since the F-distribution and the Beta-distribution may be obtained from each other by monotone transformations, the above property can be easily extended to incomplete beta integrals.

We shall next prove a property of the power function of the analysis of variance test.

Theorem 6:- Let

$$I_{\mathcal{F}}(\mu;\, n_1,\, n_2,\, \Delta) \;=\; I_{\mathcal{F}}(\mu^*;\, n_1,\, n_2^*,\, \Delta)\,, \quad n_2^* > n_2. \text{ Then}$$

$$I_{\mathcal{F}}(\mu;\, n_1,\, n_2,\, \Delta) \;>\; I_{\mathcal{F}}(\mu^*;\, n_1,\, n_2^*,\, \Delta^*), \quad \Delta^* > \Delta$$

$$<\; I_{\mathcal{F}}(\mu^*;\, n_1,\, n_2^*,\, \Delta^*), \quad \Delta^* < \Delta$$

Proof:- Let $X_1, X_2, \ldots, X_{n_1};\, Y_1, Y_2, \ldots, Y_{n_2}, Y_{n_2+1}, \ldots, Y_{n_2^*}$ be independent, normally distributed random variables with common variance σ^2 and means

$$E\,X_i \;=\; \eta_i\,, \qquad i = 1, 2, \ldots, n_1$$

$$E\,Y_i \;=\; 0\,, \qquad i = 1, 2, \ldots, n_2^*\,.$$

Consider the problem of testing

$$H_0 :\quad \sum_{i=1}^{n_1} \eta_i^2/\sigma^2 \;=\; \Delta^2$$

against

$$H_1 :\quad \sum_{i=1}^{n_1} \eta_i^2/\sigma^2 \;=\; \Delta^{*2}\; (> \Delta^2)^{n_1}$$

For this problem there exists a U.M.P. invariant test, invariant under the group of orthogonal transformations of the variates X_i, $i = 1, 2, \ldots, n_1$, and the group of scale changes $X_i^* = cX_i$, $i = 1, 2, \ldots, n_1$, $Y_j^* = cY_j$, $j = 1, 2, \ldots, n_2^*$. The critical region of this test is given by

$$\frac{\sum_{i=1}^{n_1} X_i^2}{\sum_{j=1}^{n_2^*} Y_j^2} \;\geqq\; \text{constant}\,.$$

This test is more powerful than the invariant test with critical region

$$\frac{\sum_{i=1}^{n_1} X_i^2}{\sum_{j=1}^{n_2} Y_j^2} \geqq \text{constant} .$$

This proves the result.

Now it is well known from tables and charts that the power of the analysis of variance test increases as the number of degrees of freedom for error increase, and decreases as the number of degrees of freedom for hypothesis increase (Feldt and Moharram [13], Fox [15], Hartley and Pearson [18], Lehmer [37], Tang [58]). The former of these two results is a special case of the theorem. The latter result is , however, not yet proved. As can be seen, most of the properties of integrals, proved in this chapter, are proved, very simply, by statistical methods. The last problem stated above, namely, the variation of the power of the analysis of variance test with respect to hypothesis degrees of freedom, gives rise to an interesting statistical problem which is not yet solved.

Let $X_1, \ldots, X_{n_1}, Y, Z_1, Z_2, \ldots, Z_{n_2}$ be independent, normally distributed random variables with common variance σ^2 and means

$$E X_i = \eta_i , \quad i = 1, 2, \ldots, n_1 ,$$
$$E Y = \xi ,$$
$$E Z_i = 0 , \quad i = 1, 2, \ldots, n_2 .$$

Consider the problem of testing

$$H_o : \eta_1 = \eta_2 = \ldots = \eta_{n_1} = 0 .$$

The problem can be reduced by invariance (Lehmann [35]) and we get the U.M.P. invariate test which rejects H_o when

$$\frac{\sum_{i=1}^{n_1} X_i^2}{\sum_{j=1}^{n_2} Z_j^2} \geqq \text{constant} .$$

This test can also be obtained as the U.M.P. similar region test. It is, therefore, better than any other similar region test.

Now suppose that $E\,Y = \xi = 0$. Then tests with critical regions

$$\frac{\sum_{i=1}^{n_1} X_i^2 + Y^2}{\sum_{j=1}^{n_2} Z_j^2} \geqq \text{constant}$$

and

$$\frac{\sum_{i=1}^{n_1} X_i^2}{Y^2 + \sum_{j=1}^{n_2} Z_j^2} \geqq \text{constant}$$

are both similar region and invariant tests. But now the last of the three tests is U.M.P. invariant, and is more powerful than the first two. We are interested, however, in comparing the first two tests between themselves. We want to show that the first test is more powerful than the second test. It can be shown that the first test is U.M.P. among tests which have Neyman structure with respect to the sufficient statistic $T = (\Sigma X_i^2 + \Sigma Z_j^2, Y^2)$. But this statistic is not complete, and hence the second test need not necessarily have Neyman structure with respect to T. It can be shown that it does not have it. Thus the comparison is not possible by this argument.

It may be observed that as $n_2 \to \infty$ all three tests above are equivalent to the corresponding chi-square tests with power function depending on the corresponding noncentral chi-square distribution. It will, therefore, follow that, for a sufficiently large number of observations, the power function of the analysis of variance test is a decreasing function of the number of hypothesis degrees of freedom for fixed values of the noncentrality parameter.

2.3 Some Properties of Percentage Points and Integrals Associated with Some Multivariate Tests.

In this section we shall consider some properties of percentage points and power functions connected with some multivariate tests.

2.3.1 Power Function of the T^2-test

Let us first consider the MANOVA problem of Chapter I. We have noted that when $t = 1$, there exists a U.M.P. invariant test which is Hotelling's T^2-test. This test depends on the only nonzero characteristic root of the matrix $(XX')(YY')^{-1}$. To see this we observe that in this case X is either a s-row vector $(u = 1)$ or a u-column vector $(s = 1)$. When $u = 1$, the critical region of the test is

$$T^2 = \frac{\sum_1^s X_{1j}^2}{\sum_1^n Y_{1j}^2} > \text{constant}$$

or

$$\frac{\sum_1^s X_{1j}^2/s}{\sum_1^n Y_{1j}^2/n} > \text{constant}$$

When $s = 1$, the critical region is

$$T^2 = \sum_{i=1}^{u} \sum_{j=1}^{u} s^{ij} x_{i1} x_{j1} > \text{constant},$$

where s^{ij} are the elements of $(YY')^{-1}$.

In the former case $\frac{nT^2}{s}$ is distributed as a noncentral F-variate with s and n degrees of freedom and the noncentrality parameter θ_1^2. When the null-hypothesis is true the distribution is the corresponding central distribution.

In the latter case $T^2 . \frac{n-u+1}{u}$ is distributed as a noncentral F-variate with u and $(n-u+1)$ degrees of freedom and noncentrality parameter θ_1^2. Under the null hypothesis the distribution is the corresponding central distribution (Anderson [4]). Thus we get:

Theorem 7:- When $t = 1$, the power function of the U.M.P. invariant test for the multivariate linear hypothesis is a monotonically increasing function of the number of degrees of freedom of error and hence of the number of observations.

In view of the remarks concerning the variation of the power function of the analysis of variance test with respect to the hypothesis degrees of freedom we may remark that the test mentioned in the theorem above is a monotonically decreasing function of s, if u, the effective number of variates, equals unity, and a monotonically decreasing function of u if s, the number of degrees of freedom for the hypothesis, equals unity.

2.3.2 Percentage Points of the Distribution of a Class of Statistics

In Chapter I we have characterized two classes of tests for the MANOVA problem by statistics which are linear functions of the ordered characteristic roots of the matrix $(XX')(YY')^{-1}$, or of the elementary symmetric functions of these characteristic roots. For the other three problems also, in that chapter, we have given such classes. The distribution problems associated with these problems are very closely related. Restricting our attention to the MANOVA problem, the test statistics in the notation of Chapter I are

$$\underline{a}'\,\underline{\lambda} = R(\underline{a}) \text{ say}$$

and

$$\underline{a}'\,\underline{e} = E(\underline{a}) \text{ say .}$$

Let $\text{Prob.}\left\{ R(\underline{a}) \leqq \mu_\alpha(\underline{a}, u, s, n) \right\} = 1 - \alpha$

and $\text{Prob.}\left\{ E(\underline{a}) \leqq \nu_\alpha(\underline{a}, u, s, n) \right\} = 1 - \alpha$.

Then for using a test belonging to one of the above two classes we shall need μ_α's and ν_α's. However these percentage points are not available for all u, s and n for even the three well known test statistics, associated with the three tests belonging to these classes, namely, the maximum characteristic root, the likelihood ratio and the trace. A large number of workers have investigated the distribution problem associated with these statistics (e.g. Anderson [1, 2, 4], Hotelling [22], Hsu [23, 24, 25], Ito [26, 27 and references therein], Rao [46], Roy [49]). We, however, have not been able to use their methods for determining percentage points μ_α and ν_α .

Now we shall show that these percentage points have certain monotonicity properties with respect to s and n. For this let, as in

Chapter I,

$$X = (\underline{x}_1, X_1) .$$

Then we know that the ordered characteristic roots $\lambda_i(X, Y)$, $i = 1, 2, \ldots, u$, of

$$(XX')(YY')^{-1}$$

are greater than or equal to the corresponding ordered characteristic roots $\lambda_i(X_1, Y)$, $i = 1, 2, \ldots, u$, of

$$(X_1X_1') (YY')^{-1} .$$

This implies that for fixed $\underline{a}$

$$\underline{a}' \underline{\lambda} (X, Y) \leqq a$$

and

$$\underline{a}' \underline{e} (X_1, Y) \leqq b$$

will imply, respectively, that

$$\underline{a}' \underline{\lambda} (X_1, Y) \leqq a$$

and

$$\underline{a}' \underline{e} (X_1, Y) \leqq b ,$$

where a and b are constants. This will imply that the percentage points of the distribution of $\underline{a}' \underline{\lambda}$ or $\underline{a}' \underline{e}$ will be monotonically increasing functions of s, the number of degrees of freedom for hypothesis, when u and n are fixed.

Now let us denote

$Y = (\underline{y}_1, Y_1)$, and assume that $s \geqq u$. Then (XX') is p.d. (a.e.) nonsingular matrix. Then roots of $(XX')(YY')^{-1}$ are reciprocals of the roots of $(YY')(XX')^{-1}$. By a method similar to that above it can be shown that the ordered characteristic roots of (YY') $\cdot(XX')^{-1}$ are not less than the corresponding ordered characteristic roots of $(Y_1Y_1')(XX')^{-1}$. From this it will follow that for fixed $\underline{a}$

$$\underline{a}' \underline{\lambda} (X, Y_1) \leqq a$$

and

$$\underline{a}' \underline{e} (X, Y_1) \leqq b$$

will, respectively, imply that

$$\underline{a}' \underline{\lambda} (X, Y) \leqq a$$

and

$$\underline{a}' \underline{e} (X, Y) \leqq b \quad ,$$

where a and b are positive constants. Thus we have proved the following:

Theorem 8: The percentage points of the distributions, of the class of statistics $R(\underline{a})$ and $E(\underline{a})$ are monotonically increasing functions of s, the number of degrees of freedom for hypothesis, and are monotonically decreasing functions of n, the number of degrees of freedom for error, provided for the latter case $s > u$.

We shall now make some remarks concerning the power functions of the tests for the MANOVA problem mentioned above. It is believed that for a given set of noncentrality parameter values, each of these power functions is a monotonically decreasing function of s for fixed n and u; a monotonically increasing function of n for fixed s and u, and a monotonically decreasing function of u for fixed s and n.

Hsu [23, 24] has proved that as N, the number of observations, becomes large the likelihood ratio test and the trace criterion become equivalent. More specifically it has been shown that, for large numbers of observation and $\Theta = \sum_{i=1}^{t} \Theta_i = O(1)$, the statistics $n\,T_o^2$ and $-n \log \Lambda$ can both be approximated by a noncentral chi-square statistic with $u \cdot s$ degrees of freedom and the noncentrality parameter Θ. Thus we can consider both the likelihood ratio test and the trace test as chi-square tests when the number of observations is large. We can then approximate the powers of these tests by the incomplete noncentral chi-square integrals of section 2.1 . It then follows that these approximations to power functions will be monotonically decreasing functions of both the parameters u and s.

CHAPTER III

A STUDY OF SOME UNION-INTERSECTION PROCEDURES

3.0 Introduction and Summary

The object of this chapter is to discuss power properties of some union-intersection tests and also to study them as multiple decision procedures. For this purpose the simultaneous analysis of variance test of Ghosh [15] and Ramachandran [44, 45], and a modification of it, will be taken as a model union-intersection test.

An empirical investigation of the power properties of various intuitively appealing test procedures, for a problem which does not admit a U.M.P. test, was done by Neyman and Pearson [41] with a view to comparing these tests when the alternatives are restricted to 'certain directions'. In section 3.1 we shall present and explain the relevant part of their findings. After observing that the procedures considered are union-intersection procedures we shall state the union-intersection principle in section 3.2. In section 3.3 we shall generalize the problem and the findings of Neyman and Pearson to the simultaneous analysis of variance situation. We shall show that against certain alternatives the simultaneous analysis of variance test is more powerful than the 'corresponding' analysis of variance test. In section 3.4 we shall investigate a modified simultaneous analysis of variance procedure. We shall show that under certain conditions this procedure is unbiased and that it is a Bayes solution. In the final section 3.5 we shall study the variation of the power of some union - intersection test procedures with respect to the number of intersections. We shall show that the power decreases as the number of intersections increase.

3.1 A Problem due to Neyman and Pearson

Neyman and Pearson [41] have considered the following problem.

Let U and V be independent, normally distributed random variables with unit variance and means

$$E\ U = \xi \quad , \quad E\ V = \eta \ .$$

Consider the problem of testing

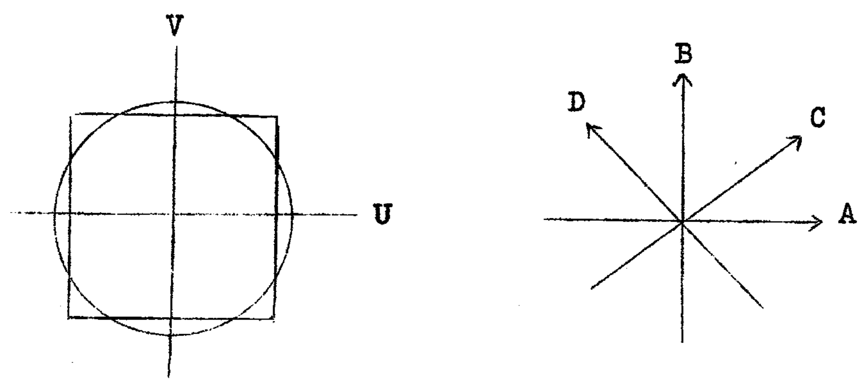
$$H_o : \xi = \eta = 0 \ .$$

The critical regions of the two tests

ϕ_1 : Reject when $u^2 + v^2 \geqq$ constant $= \mu$, say

and ϕ_2 : Reject when $\max [u^2, v^2] \geqq$ constant $= \nu$, say,

where the constants are determined by the significance level of the test, are exteriors of the circle and the square in the following diagram



Consider the problem of comparing the powers of these two tests. Neyman and Pearson obtained the values for the powers of these two tests when the noncentrality is in the directions A, B or C, D. In directions A and B only one of the parameters ξ and η is nonzero, whereas, in the directions C and D $\xi = \pm\eta$. Their computations are given in the following table:

Power functions of the two tests ϕ_1 and ϕ_2

Direction	A or B		C or D	
test	ϕ_1	ϕ_2	ϕ_1	ϕ_2
0.0	0.050	0.050	0.050	0.050
0.2	.053	.053	.053	
0.4	.062	.062	.062	
1.0	.133	.131	.133	0.125
1.5	.249	.250	.249	
2.0	.415	.422	.415	0.369
2.5	.602	.614	.602	0.
3.0	.770	.783	.770	0.702
3.5	.889	.899	.889	
4.0	.956	.962	.956	0.923
5.0	.996	.997	.996	0.991

$$\gamma = \sqrt{\xi^2 + \eta^2} .$$

Now ϕ_1 is a chi-square test and its power does not depend on direction but only on γ as can be seen in the table. From this computation Neyman and Pearson concluded that in the directions A and B the test ϕ_2 is 'slightly more powerful' than the test ϕ_1 when γ is large, but in the directions C and D the test ϕ_2 is 'considerably less powerful' than the test ϕ_1. In this section we shall explain this behavior of the power functions only partially, and pose some problems. Generalization of this problem will be considered in section 3.3.

Lemma 1: $\nu < \mu$

Proof:- The event $u^2 + v^2 \leqq a$ implies the event $\max [u^2, v^2] \leqq a$. Therefore $\text{Prob}\left(u^2 + v^2 \leqq a\right) < \text{Prob.}\left(\max [u^2, v^2] \leqq a\right)$.

Thus if $\text{Prob}\left(u^2 + v^2 \leqq \mu\right) = \text{Prob}\left(\max [u^2, v^2] \leqq \nu\right)$

then $\nu < \mu$.

Lemma 2: If $a > b > 0$ then

$$\underset{\gamma \to \infty}{\text{Lt}} \frac{\int_{-a+\gamma}^{a+\gamma} e^{-\frac{1}{2}x^2}\,dx}{\int_{-b+\gamma}^{b+\gamma} e^{-\frac{1}{2}x^2}\,dx} = \infty .$$

Proof:- As $\gamma \to \infty$ the ratio becomes indeterminate. We, therefore, by using L'Hospital's rule twice get

$$\underset{\gamma \to \infty}{\text{Lt}} \frac{\int_{-a+\gamma}^{a+\gamma} e^{-\frac{1}{2}x^2}\,dx}{\int_{-b+\gamma}^{b+\gamma} e^{-\frac{1}{2}x^2}\,dx} = e^{-\frac{1}{2}(a^2-b^2)} \underset{\gamma \to \infty}{\text{Lt}} \frac{e^{-a\gamma} - e^{a\gamma}}{e^{-b\gamma} - e^{b\gamma}}$$

$$= e^{-\frac{1}{2}(a^2-b^2)} \frac{a}{b} \underset{\gamma \to \infty}{Lt} \frac{e^{-2a\gamma}+1}{e^{-2b\gamma}+1} \cdot \frac{e^{a\gamma}}{e^{b\gamma}}$$

$$= \infty \ .$$

Now, when the noncentrality γ is in the direction A or B, we can write the probabilities of the errors of the second kind for the two tests ϕ_1 and ϕ_2 as:

$$\beta_1(\gamma) = \int_0^{\mu} \chi^2(x;\ 1,\ 0)\ dx \int_{-\sqrt{\mu-x}+\gamma}^{\sqrt{\mu-x}+\gamma} \frac{e^{-\frac{1}{2}t^2}\ dt}{\sqrt{2\pi}}$$

$$\beta_2(\gamma) = \int_{-\sqrt{\nu}}^{\sqrt{\nu}} \frac{e^{-\frac{1}{2}x^2}\ dx}{\sqrt{2\pi}} \cdot \int_{-\sqrt{\nu}+\gamma}^{\sqrt{\nu}+\gamma} \frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}\ dt$$

Now for $0 < X < \mu-\nu$, we have $\sqrt{\mu - x} > \sqrt{\nu}$. Therefore for this range of X where exists a number γ^{**} such that, given a number N however large, for $\gamma \geqq \gamma^{**}$,

$$\int_{-\sqrt{\mu-x}+\gamma}^{\sqrt{\mu-x}+\gamma} \frac{e^{-\frac{1}{2}t^2}\ dt}{\sqrt{2\pi}} > N \cdot \int_{-\sqrt{\nu}}^{\sqrt{\nu}} \frac{e^{-\frac{1}{2}t^2}\ dt}{\sqrt{2\pi}} \cdot \int_{\sqrt{\nu}+\gamma}^{\sqrt{\nu}+\gamma} \frac{e^{-\frac{1}{2}t^2}\ dt}{\sqrt{2\pi}}$$

It is then easy to deduce that there exists a number γ^{*} such that for $\gamma > \gamma^*$

$$\beta_1(\gamma) > \beta_2(\gamma) \ ,$$

that is for $\gamma > \gamma^*$ in directions A or B, the test ϕ_2 is more powerful than the test ϕ_1. At this stage we can not say that γ^*

belongs to the range of γ considered in the computation. We can see from the table that for $5 \geqq \gamma \geqq 1.5$ ϕ_2 is more powerful than ϕ_1. If it is true that, if ϕ_2 is better than ϕ_1 for $\gamma = \gamma_o$ then it is better than ϕ_1 for $\gamma \geqq \gamma_o$, then $\gamma^* = \gamma_o$.

In the table we see that for $\gamma = 1$, ϕ_1 is better than ϕ_2, the corresponding values of power being 0.133 and 0.131. Intuitively one would not expect this. A local comparison of ϕ_1 and ϕ_2 should be interesting. Though ϕ_1 is an L.M.P. test for the problem, it would not be expected that ϕ_1 would be better than ϕ_2 locally but only as good as ϕ_2 when the departure from the null hypothesis is in the directions A or B.

The test ϕ_1 is the likelihood ratio test for the problem of testing $\xi = 0 = \eta$ against the violation of it; it, also, is a union-intersection test for the problem of testing the hypothesis $\xi = 0 = \eta$ against alternatives $c_1\xi + c_2\eta \neq 0$, c_1, c_2 not vanishing simultaneously. The test ϕ_2 is the union-intersection principle test for testing the hypothesis $\xi = 0 = \eta$ against the alternative that only one of ξ and η is nonzero.

3.2 The Union-Intersection Principle.

A heuristic method of test construction was suggested by Roy [47]. This method, later called 'the union-intersection' method, is, in outline, as follows:

Suppose that X is a stochastic variate with distribution depending on a parameter $\xi \in \Xi$. Suppose that we are interested in testing a hypothesis

$$H_o : \xi \in \Xi_o \subset \Xi$$

against

$$H_1 : \xi \in \Xi_1 \subset \Xi .$$

Suppose that H_o and H_1 are expressible as

$$H_o = \bigcap_{\gamma \in \Gamma} H_{o\gamma} \quad \text{and} \quad H_1 = \bigcup_{\gamma \in \Gamma} H_{1\gamma} \quad ,$$

where $H_{o\gamma}$ and $H_{1\gamma}$ are given by

$$H_{o\gamma} ; \xi \in \Xi_{o\gamma} \quad , \quad H_1 = \xi \in \Xi_{1\gamma} \quad ,$$

and Γ is a finite or infinite index set. Suppose that, for each component problem of testing $H_{o\gamma}$ against $H_{1\gamma}$, we have some good test procedures with acceptance regions A_γ , $\gamma \in \Gamma$, say. Then the union-intersection principle dictates that the hypothesis H_o be accepted over $\bigcap_{\gamma \in \Gamma} A_\gamma$ and be rejected otherwise.

In this chapter we intend to study various properties of intersection tests, from both the classical power point of view and the modern loss-function-risk point of view. For this purpose we shall consider the problem of simultaneous analysis of variance, which may be considered as a generalization of the above problem due to Neyman and Pearson.

3.3 Simultaneous Analysis of Variance

3.3.1 Model

We shall borrow the following model for the simultaneous analysis of variance problem from Ramachandran [44] .

Let $\underline{x}$ (nxl) denote a set of observations on n independent, normally distributed random variables with common variance σ^2 and means given by

$$\underset{n\times 1}{E(\underline{x})} = \underset{n\times m}{D} \; \underset{m\times 1}{\underline{\xi}}$$

where D is a matrix of known constants, such that rank $(D) = r \leq m < n$; and ξ is a vector of m unknown parameters.

We wish to test, simultaneously, the following k hypotheses on ζ:

$$\underset{q \times m}{C} \; \underset{m \times 1}{\zeta} = \begin{matrix} q_1 \\ q_2 \\ \vdots \\ q_k \end{matrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{bmatrix}_{m} \underset{m \times 1}{\zeta} = \underset{q \times 1}{\underline{0}} ,$$

where $q = \sum_{i=1}^{k} q_i$ and C is a given hypothesis-matrix of rank $s \leq \min(r,q)$.

Now the analysis of variance test for testing

$$\underset{q_i \times m}{C_i} \; \underset{m \times 1}{\zeta} = \underline{0}$$

is given by the critical region ,

$$F_i \geq \text{constant} ,$$

where

$$F_i = \frac{\underline{x}' D_1 (D_1' D_1)^{-1} C_{i1} \left(C_{i1} (D_1' D_1)^{-1} C_{i1}' \right)^{-1} C_{i1}' (D_1' D_1)^{-1} D_1' \underline{x} / t_i}{\underline{x}' \left(I(n) - D_1 (D_1' D_1)^{-1} D_1' \right) \underline{x} / (n-r)} ,$$

where as in section 1.1.1 D_1, D_2 and C_{i1}, C_{i2} are the partitions of D and C_i, and t_i is the rank of C_i .

F_i defined above is distributed as a variance-ratio with t_i and $(n-r)$ degrees of freedom, when the hypothesis is true.

We shall assume that the k hypotheses are quasi-independent, i.e.

$$D_1 (D_1' D_1)^{-1} C_{i1}' \left[C_{i1} (D_1' D_1)^{-1} C_{i1}' \right]^{-1} C_{i1} (D_1' D_1)^{-1} C_{j1}' \cdot \left[C_{j1} (D_1' D_1)^{-1} C_{j1}' \right]^{-1} C_{j1} (D_1' D_1)^{-1} D_1' = 0$$

for all combinations (i,j), $i \neq j = 1, 2, \ldots, k$.

Under these conditions the numerators of F_i, $i = 1, 2, \ldots, k$, are distributed as k independent $\chi^2(t_i, 0)$-variates, $i = 1, 2, \ldots, k$, under the null-hypothesis, and as k independent $\chi^2(t_i, \delta_i)$-variates, $i = 1, 2, \ldots, k$, under the alternative

$$C \underline{\zeta} = \begin{bmatrix} \mathfrak{D}_1 \\ \vdots \\ \mathfrak{D}_k \end{bmatrix} \begin{matrix} q_1 \\ \vdots \\ q_k \end{matrix} \neq \underline{0} \quad ,$$

$$\delta_i^2 = \mathfrak{D}_i' [C_{i1}(D_1'D_1)^{-1} C_{i1}'] \mathfrak{D}_i / \sigma^2 \quad .$$

The simultaneous analysis of variance test accepts the hypothesis

$$C \underline{\zeta} = \underline{0}$$

if, and only if,

$$F_i \leq \nu_i \qquad i = 1, 2, \ldots, k \;.$$

The optimum choice for ν_i is not known. Ramachandran [44] suggests that ν_i be chosen proportional to t_i, and so as to satisfy the level condition

$$\Pr [F_i \leq \nu_i \;, \quad i = 1, 2, \ldots, k] = 1 - \alpha \;,$$

$0 < \alpha < 1$. We shall comment on this aspect in section 3.4, where we shall consider the problem of simultaneous analysis of variance as a multiple-decision problem. The distribution problem of the simultaneous analysis of variance has been solved by Ramachandran [45] .

In the following section we shall compare the simultaneous analysis of variance test with the corresponding analysis of variance test for a restricted class of alternatives. For this we shall need the following:

3.3.2 Canonical Form

Let Y_{ij} $(i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, t_i;\ \sum_{i=1}^{k} t_i = s)$, Y_i $(i = 1, 2, \ldots, r-s)$, Z_i $(i = 1, 2, \ldots, n-r)$ be n independent normally distributed random variables with common variance σ^2 and means

$$
\begin{aligned}
E(Y_{ij}) &= \xi_{ij} \quad , \quad i = 1, 2, \ldots, k \\
&\qquad\qquad\quad j = 1, 2, \ldots, t_i \\
E(Y_i) &= \xi_i \quad , \quad i = 1, 2, \ldots, r-s \\
E(Z_i) &= 0 \quad , \quad i = 1, 2, \ldots, n-r .
\end{aligned}
$$

We wish to test the hypothesis

$$H_o : \quad \xi_{ij} = 0 \ , \ i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, t_i$$

against the alternative

$$H_1 : \quad \xi_{ij} = \lambda_{ij} \quad i = 1, 2, \ldots, k \quad \text{not all zero.}$$
$$j = 1, 2, \ldots, t_i$$

The usual analysis of variance test rejects when

$$\frac{\sum_{i=1}^{1} \sum_{j=1}^{t_i} Y_{ij}^2}{\sum_{i=1}^{n-r} Z_i^2} \geqq \text{constant} \quad (\mu^* \text{ , say}) .$$

The simultaneous analysis of variance test rejects the hypothesis when

$$\frac{\sum_{j=1}^{t} Y_{ij}^2}{\sum_{i=1}^{n-r} Z_i^2} \geqq \text{constant} \quad (\mu_i \text{ say })$$

for at least one $i = 1, 2, \ldots, k$.

3.3.3 Comparison of the analysis of variance test and the simultaneous analysis of variance test

For this comparison let us assume that $t_1 = t_2 = \ldots = t_k = t$. The situations when all hypothesis degrees of freedom are equal are very common in the analysis of factorial designs, where we simultaneously test the significance of a number of effects and interactions.

We shall assume that we are testing

$$H_0 : \quad \xi_{ij} = 0 \qquad i = 1, 2, \ldots, k$$

$$j = 1, 2, \ldots, t \qquad \text{against}$$

H_1 : Only one of the k t-vectors

$$(\xi_{i1}, \xi_{i2}, \ldots, \xi_{it}) , \quad i = 1, 2, \ldots, k$$

is nonnull.

This alternative corresponds to what Neyman and Pearson [41] call a shift in a particular direction. An interest in restricted alternatives of this type has also been shown by Roy [50], where some advantages of using the union-intersection principle in certain experimental situations are considered. It has been conjectured there that the simultaneous analysis of variance test with $t = 1$ will be more powerful than the corresponding analysis of variance test if the departure from the hypothesis is in particular directions. We shall show that this is true, at least for sufficiently large deviations from the hypothesis in the particular directions.

Let $\sum_{j=1}^{t} \frac{\xi_{ij}^2}{\sigma^2} = \eta_i^2$, $i = 1, 2, \ldots, k$.

Then we can express our hypothesis as all $\eta_i = 0$, $i = 1, 2, \ldots, k$ and alternative as only one of these η_i's is nonzero, say η . The simultaneous analysis of variance test accepts the hypothesis when all

$$\frac{\sum_{j=1}^{t} Y_{ij}^2}{\sum_{i=1}^{n-r} Z_i^2} \leqq \text{constant, say, } \mu \quad ,$$

$i = 1, 2, \ldots, k$, and rejects it otherwise. The analysis of variance test accepts the hypothesis when

$$\frac{\sum_{i=1}^{k} \sum_{j=1}^{t} Y_{ij}^2}{\sum_{i=1}^{k} Z_i^2} \leqq \text{constant, say, } \mu^* \ ,$$

and rejects it otherwise. We can, therefore, write the probabilities of errors of the second kind for the two tests, respectively, as

$$\beta_1(\eta) = \int_0^\infty \chi^2(v;\ n-r,\ 0) \prod_{i=2}^{k} \int_0^{\mu v} \chi^2(u_i;\ t,\ 0)\ d\,u_i \int_0^{\mu v} \chi^2(u_1;\ t,\ \eta)\ d\,u_1 \quad d\,v \quad ,$$

$$\beta_2(\eta) = \int_0^\infty \chi^2(v;\ n-r,\ 0) \int_0^{\mu v} \chi^2(u;\ kt,\ \eta)\ du \quad dv \quad ,$$

where μ and μ^* are the constants mentioned above, being determined by

$$\int_0^\infty \chi^2(v;\ n-r,\ 0) \prod_{i=1}^{k} \int_0^{\mu v} \chi^2(u_i,\ t,\ 0)\ du_i\ dv = 1 - \alpha \quad ,$$

$$\int_0^\infty \chi^2(v;\ n-r,\ 0) \int_0^{\mu^* v} \chi^2(u,\ kt,\ 0)\ du \quad dv = 1 - \alpha \quad ,$$

α being the level of significance. Or

$$\beta_1(\eta) = \int_0^\infty \chi^2(v;\ n-r,\ 0) \prod_{i=2}^{k} \int_0^{\mu v} \chi^2(u_i;t,0)\ du_i$$

$$\int_0^{\mu v} \chi^2(u_1^*;\ t-1,0)\ du_1^* \int_{-\sqrt{\mu v-u_1^*}+\eta}^{\sqrt{\mu v-u_1^*}+\eta} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\ dx\ dv$$

$$\beta_2(\eta) = \int_0^\infty \chi^2(v;n-r,0) \int_0^{\mu v} \chi^2(u^*;kt-1,0)\ du^* \int_{-\sqrt{\mu^* v-u^*}+\eta}^{\sqrt{\mu^* v-u^*}+\eta} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\ dx\ dv$$

Now it can, as before, be easily seen that $\mu < \mu^*$. Then application of the lemma will give us existence of an η^* such that for values of η greater than η^* the simultaneous analysis of variance test will be more powerful than the analysis of variance test. What happens for $\eta < \eta^*$ is not known. However, it is believed that the simultaneous analysis of variance test would be at least as powerful as the analysis of variance test for small values of the noncentrality parameter η when it is in the restricted directions. We may summarize this argument in the following:

Theorem 1: For the restricted simultaneous analysis of variance problem, as stated above, there exists an $\eta^* > 0$ such that for $\eta \geq \eta^*$, the simultaneous analysis of variance test is more powerful than the analysis of variance test.

In the following section we shall consider the simultaneous analysis of variance procedure as a multiple decision procedure.

3.4 Modified Simultaneous Analysis of Variance, a Multiple-decision Procedure.

The problems in analysis of variance, for which the simultaneous analysis of variance procedure was developed, are more in the nature of multiple-decision problems than in the nature of two decision problems. To illustrate this, we may consider analysis of variance of classical factorial experiments. In this the experimenter is more interested in knowing which of the effects and the interactions are significant and which are not than in testing some total hypothesis.

In this section we shall discuss properties of the following modification of the procedure in section 1. In the notation of section 1 suppose that we have tested the k hypotheses

$$H_{oi} : \quad C_i \; \underline{\zeta} = \underline{0}$$

against $\quad H_{1i} : \quad C_i \; \underline{\zeta} = \underline{\eta}_i \neq \underline{0} \,, \quad i = 1, 2, \ldots, k \,,$

by the tests of that section which depend upon the test-statistics F_i, $i = 1, 2, \ldots, k$, at levels of significance α_i, $0 < \alpha_i < 1$, $i = 1, 2, \ldots, k$. Then the modified simultaneous analysis of variance procedure dictates that among these k hypotheses only those, accepted by the component F-tests, be accepted and the remaining be rejected.

This modified procedure of the simultaneous analysis of variance belongs to the following formulation of multiple-decision problems (Lehmann [36]).

Let X be a stochastic variable with its distribution depending upon a parameter $\theta \in \text{Ⓗ}$, where Ⓗ is the parameter space. Suppose that we are interested in testing a number of hypotheses ,

$$H_{oi}: \quad \theta \in \Theta_i \subset \Theta$$

against $\quad H_{1i}: \quad \theta \in \Theta_i^{-1} = \Theta - \Theta_i \;, \quad i = 1, 2, \ldots, k \;,$

concerning the parameter θ and determining, simultaneously, which of these hypotheses are true and which are false.

Suppose that we use procedures π_i $(i = 1, 2, \ldots, k)$ for testing H_{oi}, and as a simultaneous decision π accept only those which are accepted by the individual procedures. Let us also suppose that the total procedure π based on the individual procedure is logically consistent, that is, there are no contradictions in the possible decisions arrived at by this procedure based on the component procedures.

For determining some properties of such a total procedure in relation to those of component procedures, let us assume the following

Loss function: Let a_i and b_i $(i = 1, 2, \ldots, k)$ be the losses due to falsely rejecting and accepting H_{oi} by procedure π_i, where a_i and b_i are not necessarily constant with respect to i. Notice that such a loss function is appropriate for the kind of problems in classical factorial experiments which we have mentioned. Losses due to errors in misjudgment concerning main effects and first order interactions, in these experiments, may be more serious than those due to higher order interactions. For this loss function, it may be noted, the risk function is simply a weighted sum of the probabilities of the errors of the two kinds.

Let us further assume that the losses are additive, that is, the loss due to the total procedure is the sum of the losses for the individual procedures.

With this set-up in mind, suppose that we test each of the hypotheses H_{oi} $(i = 1, 2, \ldots, k)$ at levels of significance

$$\alpha_i = \frac{b_i}{a_i+b_i} , \quad i = 1, 2, \ldots, k .$$

Now the analysis of variance test is strictly unbiased in the classical, i.e. Neyman and Pearson [41], sense. It is, therefore, unbiased in the Lehmann [33] sense, i.e. the average loss using the analysis of variance test is minimum when the decision is correct.

Because of the nature of the loss function, the total risk of the total procedure is the sum of the component risks of the component procedures. It, therefore, follows that when the individual F-tests are used at the levels of significance α_i , $i = 1, 2, \ldots, k$, the modified simultaneous analysis of variance procedure is unbiased in the Lehmann sense. This gives us a rule concerning the choice of the levels of significance for the individual F-tests of the modified simultaneous analysis of variance procedure. The distribution problem connected with the test has been solved by Ramachandran [44] .

We shall next prove that the modified simultaneous analysis of variance procedure is a Bayes solution. Towards this end we shall state a result due to Duncan [7, 8] and Lehmann [36].

Lemma 3:- If the component procedures π_i, $i = 1, 2, \ldots, k$, are Bayes solutions with respect to a common a priori distribution on the parameter space, then the total procedure is also a Bayes solution with respect to the same a priori distribution.

Now it is well known (Hsu [25], Wald [61] and Wolfowitz [64]) that the analysis of variance test is a Bayes solution with respect to an a prioro distribution, in the space of parameters, which is uniform on the surfaces where the noncentrality parameter is constant. In the case of the modified simultaneous analysis of variance procedure we can consider an a priori distribution Λ, say, which is uniform on each of the k spheres

$$\sum_{j=1}^{t_i} \lambda_{ij}^2 = c^2 \sigma^2 , \quad i = 1, 2, \ldots, k$$

for all real numbers c. Then each of the component procedures is a Bayes solution with respect to this a priori distribution. Therefore, by virtue of the above lemma, the modified simultaneous analysis of variance is also a Bayes solution with respect to this a priori distribution.

We may summarize the results of this section in the following:

Theorem 2: The modified simultaneous analysis of variance procedure is unbiased if the levels of significance of the component tests are $\alpha_i = b_i/(a_i + b_i)$, $i = 1, 2, \ldots, k$. Furthermore, the modified simultaneous analysis of variance procedure is a Bayes solution for the problem when Λ is the a priori distribution in the parametric space.

3.5 A Property of the Power Functions of Some Intersection Tests.

The simultaneous analysis of variance test considered in the previous section is not the union-intersection test for the problem stated there. The union-intersection test accepts the hypothesis when all k quantities

$$\frac{\sum_{j=1}^{t} Y_{ij}^2}{\sum_{\ell \neq i}^{k} \sum_{j=1}^{t} Y_{\ell j}^2 + \sum_{i=1}^{n-r} Z_i^2} \leqq \text{constant}, \; i = 1,2,\ldots,k$$

and rejects it otherwise. When σ^2, the population variance, is known this test coincides with the corresponding test given by the simultaneous analysis of variance procedure as well as that given by the likelihood ratio principle. This test has an intersection of k chi-square test acceptance regions for its acceptance region. In this section we shall examine the variation of the power function of this test, with the intersection of chi-square test acceptance regions for its acceptance region, with respect to the number of intersections. This problem arises in the following context also.

Let $X_1, X_2, \ldots, X_k$ be independent normally distributed random variables with variance unity and means $E(X_1) = \theta$, $E(X_i) = 0$, $i = 2, 3, \ldots, k$. Consider the problem of testing H_0: $\theta = 0$ against H_1: $\theta \neq 0$. We know that the U.M.P. test accepts H_0 only when $|X_1| \leqq$ constant. Other valid acceptance regions for the hypothesis are

A_j : $\bigcap_{i=1}^{j} \; |x_i| \leqq \text{constant} = \mu_j$; say , $j = 2, 3, \ldots, k$ and

B_j : $\sum_{i=1}^{j} X_i^2 \leqq$ constant, $j = 2, 3, \ldots, k$. We are interested in comparing powers of these tests. We know that, for sufficiently large alternatives θ, A_j will be more powerful than B_j. It is felt that A_j will be at least as good as B_j for all $|\theta| > 0$. Whether A_{j-1}, the acceptance region with j-1 intersections, will be more powerful than B_j, the chi-square with j d.f., for all $|\theta| > 0$ is a problem.

In what follows we shall show that for the intersection of chi-squares tests, the power decreases as the number of intersections increases.

Theorem 3: Let k and k^* be positive integers, $k^* > k$. Let

$$\prod_{i=1}^{k} \int_0^{\mu} \chi^2(x_i;\, t,\, 0)\, dx_i = \prod_{i=1}^{k^*} \int_0^{\mu^*} \chi^2(x_i;\, t,\, 0)\, dx_i \;;$$

then
$$\prod_{i=2}^{k} \int_0^{\mu} \chi^2(x_i;\, t,\, 0)\, dx_i \int_0^{\mu} \chi^2(x_1;\, t,\, \eta)\, dx_1$$

$$< \prod_{i=2}^{k^*} \int_0^{\mu^*} \chi^2(x_i,\, t,\, 0)\, dx_i \int_0^{\mu^*} \chi^2(x_1,\, t,\, \eta)\, dx_1 \;.$$

Proof: It is sufficient to prove that

$$\frac{\int_0^{\mu} \chi^2(x;\, t,\, \eta)\, dx}{\int_0^{\mu} \chi^2(x;\, t,\, 0)\, dx} < \frac{\int_0^{\mu^*} \chi^2(x;\, t,\, \eta)\, dx}{\int_0^{\mu^*} \chi^2(x;\, t,\, 0)\, dx}$$

Towards this end we shall need the following:

Lemma 4: $$R(\mu, \delta) = \frac{\int_0^{\mu} \chi^2(x;\, n,\, \delta)\, dx}{\int_0^{\mu} \chi^2(x;\, n,\, 0)\, dx}$$

is a monotonically increasing function of μ for each fixed $\delta > 0$.

This lemma can be proved by straightforward differentiation of $R(\mu, \delta)$ w.r.t. μ. However, as the distribution of a noncentral chi-square variate has monotone likelihood ratio (M.L.R.), (Lehmann [34, 35]) it is interesting to note that the above lemma is a particular case of the following lemma due to Hall [17].

<u>Lemma 5: Let a continuous random variable X have a M.L.R. family of densities $f_\theta(x)$ with distribution functions $F_\theta(x)$. Then for $\theta' < \theta$, the ratio $F_\theta(x)/F_{\theta'}(x)$ is a monotonically increasing function of x .</u>

<u>Proof:</u> Let us write $F_\theta(x) = F(x)$, $f_\theta(x) = f(x)$, $F_{\theta'}(x) = G(x)$, $f_{\theta'}(x) = g(x)$. Then

$$\operatorname{sgn}\left[\frac{d}{dx}\frac{F(x)}{G(x)}\right] = \operatorname{sgn}\left[f(x)\,G(x) - F(x)\,g(x)\right]$$

$$= \operatorname{sgn}\left[\int_{-\infty}^{x}\left[\frac{f(x)}{g(x)} - \frac{f(y)}{g(y)}\right]g(x)\cdot g(y)\,dy\right].$$

Now f_θ is a M.L.R. family of densities. Therefore, for $y < x$, $\frac{f(x)}{g(x)} > \frac{f(y)}{g(y)}$. Thus, the integrand in the last step being nonnegative, the integral is positive. It follows, therefore, that $F(x)/G(x)$ is an increasing function of x. The proof of theorem is thus complete.

An observation that $\mu^* > \mu$ completes the proof of Theorem 3.

CHAPTER IV

SIMULTANEOUS INTERVAL ESTIMATION

4.0 Introduction and Summary

Even though the concept of simultaneous interval estimation is old (Working and Hotelling [63]) its application to the problem of multiple comparisons in analysis of variance is comparatively recent (Scheffé [54, 55], Tukey [59, 60]). The method of simultaneous interval estimation was first systematized by Roy and Bose [51] and Roy [48]. They obtained both the Scheffé-bounds and the Tukey-bounds as particular cases of their argument. It is well known that pairwise comparisons by Scheffe's method compare unfavorably with those by Tukey's method.

In this chapter we shall discuss the problem of simultaneous interval estimation in univariate and multivariate analysis of variance situations. We shall obtain various solutions to the problem in the univariate analysis of variance situation and extend them to the multivariate situation. In section 4.4 we shall obtain a theorem which will explain the sharpness of Tukey bounds as compared with Scheffé bounds, and of Dunnet bounds as compared with both these bounds. Using the theorem we can extend these sharpness properties from the univariate to the multivariate situation.

4.1 Simultaneous Interval Estimation

Let $y = (y_1, y_2, \ldots, y_n)$ be an observed set of random variables with distribution depending on a set of unknown parameters

$$\xi = (\xi_1, \xi_2, \ldots, \xi_k), \qquad \xi \in \Xi \quad ,$$

where Ξ is a set .

Let

$$\left\{ \omega_\gamma = \omega_\gamma(\xi) \quad , \qquad \gamma \in \Gamma \right\} \quad ,$$

where Γ is a finite or infinite index set, be a set of parametric functions. The problem is that of making simultaneous confidence statements

$$\ell_\gamma \leqq \omega_\gamma \leqq u_\gamma \quad , \qquad \gamma \in \Gamma \quad ,$$

with a known simultaneous confidence coefficient $(1-\alpha)$, $0 < \alpha < 1$.

This problem occurs in several branches of applied statistics, especially, in the analysis of variance and covariance of experiments and regression analysis. The earliest nontrivial example of the simultaneous interval estimation of the kind to be discussed in this chapter, occurs in the paper of Working and Hotelling [63], where they have obtained a confidence bound for a regression line, which may be regarded as a set of simultaneous confidence statements on all linear functions $c_1\beta_1 + c_2\beta_2$ $(c_1 \neq 0)$ of the regression parameters β_1 and β_2. Tukey [59, 60] and Scheffe [54, 55] have given methods for the simultaneous interval estimation of, respectively, $\binom{k}{2}$ pairwise contrasts and all linear contrasts of the parameters $\xi_1, \xi_2, \ldots, \xi_k$.

The earliest systematic investigation of the problem, however, appears to be due to Roy and Bose [51]. Examples of confidence interval estimation can be found in, Roy and Bose [51], Roy [48], Dwass [11], Dunnet [9], Durand [10], Hoel [19], Mandel [38], Scheffe [56]. Roy and Bose [51] have shown that the problem can be ~~found~~ solved under the following circumstances:

<u>Lemma:</u> <u>Roy and Bose</u> [51]

Suppose that it is possible to find a set of functions

$$\phi_\gamma(y, \omega_\gamma) \quad , \quad \gamma \in \Gamma$$

such that

$$\mu_1 \leqq \phi_\gamma \leqq \mu_2 \quad , \quad \gamma \in \Gamma$$

is equivalent to

$$\ell_\gamma \leqq \omega_\gamma \leqq u_\gamma \quad , \quad \gamma \in \Gamma \qquad (\cdot\dot{\underset{\cdot}{x}}\cdot)$$

where μ_1 and μ_2 are constants independent of γ. For a given ξ let

$$E_{\gamma,\xi} = \left\{ y : \mu_1 \leqq \phi_\gamma \leqq \mu_2 \mid \xi \right\}$$

be the set of points y in the sample space E_n, for which $(\cdot\dot{\underset{\cdot}{x}}\cdot)$ is true. Let

$$E_\xi = \bigcap_{\gamma \in \Gamma} E_{\gamma,\xi} \quad .$$

If E_ξ is a Borel set for each ξ, and

$$\Pr \left\{ y \in E_\xi \mid \xi \right\} = 1 - \alpha, \quad 0 < \alpha < 1 \quad ,$$

$\xi \in \Xi$, then $(\cdot\dot{x}\cdot)$ are true with the same probability.

4.2 Simultaneous Interval Estimation in the Univariate Analysis of Variance.

Without any loss of generality we may confine ourselves to two-factor design terminology.

4.2.1 Univariate Model

Let $\underline{y}' = (y_1, y_2, \ldots, y_n)$, where $y_1, y_2, \ldots, y_n$ are independent, normally distributed random observables with common variance σ^2 and means $E(y_i)$, $i = 1, 2, \ldots, n$, given by

$$\underset{n \times 1}{E(\underline{y})} = \underset{n \times v}{A} \; \underset{v \times 1}{\underline{\tau}} + \underset{n \times b}{X} \; \underset{b \times 1}{\underline{\beta}} \quad ,$$

where

(i) A and X are matrices of known constants depending upon the design of the experiment with rank $(A) = r$, rank $(A \vdots X) = r + q$, $r + q < n$

(ii) $\underline{\tau}$ and $\underline{\beta}$ are vectors of unknown parameters. In the design terminology the elements of $\underline{\tau}$ may be regarded as the v treatment effects and the elements of $\underline{\beta}$, the b block effects.

4.2.2 Univariate Estimation

The experimenter is, usually, interested in testing significance of the treatment effects i.e. testing the hypothesis

$$H_0 : \tau_1 = \tau_2 = \ldots = \tau_v = 0$$

and estimating some linear contrasts of the treatment effects.

Suppose that the experimenter is, specially, interested in a particular set of contrasts

$$G_\nu = \left\{ \underline{c}' \underline{\tau} \ : \ \underline{c} \in \Delta_\nu \right\} ,$$

where

$\Delta_\nu \subset \Delta$, Δ being the space of all real vectors

$$\underline{c}' = (c_1, c_2, \ldots, c_v) \text{ with}$$

$$\sum_{j=1}^{v} c_j^2 = 1 \quad , \quad \sum_{j=1}^{v} c_j = 0$$

and the rank of $\Delta_\nu \geqq (v-1)$.

Under these circumstances we can express the null hypothesis as

$$H_o = \bigcap_{\underline{c} \in \Delta_\nu} H_{o,\underline{c}}$$

where $H_{o\underline{c}} : \underline{c}' \underline{\tau} = 0$. Let us consider the problem of testing H_o against

$$H_\nu : \bigcup_{\underline{c} \in \Delta_\nu} [H_{o,\underline{c}}^{-1}]$$

where H^{-1} denotes 'not H'. Suppose that the τ's are estimable. Let

$$\underline{t}' = (t_1, t_2, \ldots, t_v)$$

be the best unbiased linear estimate of the treatment effects

$$\underline{\tau}' = (\tau_1, \ldots\ldots, \tau_v) .$$

Then $\underline{c}' \underline{t}$ is the best unbiased linear estimate of $\underline{c}' \underline{\tau}$. Let $s_{\underline{c}}^2$ be the estimate of the variance of $\underline{c}' \underline{t}$, with n_e degrees of freedom. Then $[\underline{c}' (\underline{t} - \underline{\tau})]^2 / s_{\underline{c}}^2$ is distributed as the square of Student's t with n_e degrees of freedom. We have, therefore, a test for $H_{o\underline{c}}$ as:

Accept if, and only if,

$$\frac{(\underline{c}'\,\underline{t})^2}{s_{\underline{c}}^2} \leqq \text{constant} = \mu^2_{(\nu),\alpha} \quad \text{say}$$

i.e. $-\mu_{(\nu),\alpha}\, s_{\underline{c}} \leqq \underline{c}'\,\underline{t} \leqq \mu_{(\nu),\alpha}\, s_{\underline{c}}$.

Thus by the union-intersection principle we accept H_o over

$$A_{(\nu)} = \bigcap_{\underline{c}\,\epsilon\,\Delta_\nu} [-\mu_{(\nu),\alpha}\, s_{\underline{c}} \leqq \underline{c}'\,\underline{t} \leqq \mu_{(\nu),\alpha}\, s_{\underline{c}}] ,$$

where $\mu_{(\nu),\alpha}$ is given by

$$\Pr [A_{(\nu)} \mid H_o] = 1 - \alpha .$$

It is easy to see that from this test procedure we get the following set of simultaneous interval estimates for $\underline{c}'\,\tau$, $\underline{c}\,\epsilon\,\Delta_\nu$:

$$\left\{ \underline{c}'\,\underline{t} - \mu_{(\nu),\alpha}\, s_{\underline{c}} \leqq \underline{c}'\,\underline{\tau} \leqq \underline{c}'\,\underline{t} + \mu_{(\nu),\alpha}\, s_{\underline{c}} \;:\; \underline{c}\,\epsilon\,\Delta_\nu \right\}$$

with the simultaneous confidence coefficient $1 - \alpha$.

4.2.3 Applications

By considering different sets $\Delta_\nu \subset \Delta$, we shall indicate how some of the previous results in this field are particular cases of this set-up, essentially, due to Roy and Bose [51] .

(i) $\Delta_\nu = \Delta$, the entire space.

In this case we get the simultaneous interval estimates for all the contrasts $\underline{c}' \underline{\tau}$, with a specified simultaneous confidence coefficient $(1 - \alpha)$. This is due to Scheffe [54, 55]. The constant $\mu_{(\nu),\alpha}$ is determined from the tables of the distribution of the variance ratio.

(ii) $\Delta_\nu = \Delta_o$, a set of (v-1) orthogonal vectors $\underline{c} \in \Delta$.

This gives us the simultaneous interval estimates for a set of (v-1) orthogonal contrasts $\underline{c}' \underline{\tau}$. The distribution problem, in this case, has been solved by K. V. Ramachandran [45].

(iii) $\Delta_\nu = \Delta_p$, the set of $\binom{v}{2}$ vectors $\underline{c} \in \Delta$, such that all elements of $\underline{c}$, except two, are zero, the nonnull being $\pm \frac{1}{\sqrt{2}}$.

In this case we get the simultaneous interval estimates for all the pairwise contrasts $(\tau_i - \tau_j) \frac{1}{\sqrt{2}}$. The determination of constant $\mu_{(\nu),\alpha}$ uses distribution of the studentized range statistics. This solution is due to Tukey [59].

(iv) $\Delta_\nu = \Delta_f$, a set of vectors (v-1) $\underline{c} \in \Delta$, such that a fixed element of $\underline{c}$ is $\frac{1}{\sqrt{2}}$, only one of the remaining is $-\frac{1}{\sqrt{2}}$ and the rest are zero.

This gives us the simultaneous interval estimates of the pairwise comparisons of the treatments with a control. The solution is due to Dunnet [9].

4.3 Simultaneous Interval Estimation in Multivariate Analysis of Variance

4.3.1 Multivariate Model

Again, with no loss of generality, we speak in terms of a two factor design model.

Let Y_{nxp} be a matrix, with n p-dimensional random observables $\underline{y}_1, \ldots, \underline{y}_n$ as the columns. Let $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n$ be independent and distributed in a p-variate normal distribution, with common positive definite variance-covariinace matrix Σ, and means $E(\underline{y}_i)$, $i = 1, 2, \ldots, n$, given by

$$\underset{nxp}{E\,Y'} = \underset{nxv}{A}\;\underset{vxp}{\tau} + \underset{nxb}{X}\;\underset{bxp}{\beta} \qquad ,$$

where i) A and X are two matrices with the same meaning as in the univariate case.

ii) τ and β are matrices whose ν and b rows, respectively, signify the treatment effects and the block effects.

Let $\underline{\tau}_1', \underline{\tau}_2', \ldots, \underline{\tau}_v'$ be the v rows of τ. Then we define a contrast-vector, a multivariate analogue of contrast, as a vector

$$\underset{1xv}{\underline{c}'}\;\underset{vxp}{\tau} \qquad ,$$

where $\underline{c} \in \Delta$, the same space of real vectors $\underline{c}$ as in the previous section.

4.3.2 Multivariate Estimation

The experimenter, as before, is interested in testing the hypothesis of no treatment effects i.e.

$$H_0 : \underline{\tau}_1 = \underline{\tau}_2 = \cdots = \underline{\tau}_v = \underline{0} \quad ,$$

and making simultaneous confidence statements about the $\underline{\tau}$'s .

In this case it is possible to make simultaneous confidence interval estimation of all the linear compounds, like

$$\underline{c}'\tau\,\underline{a} , \qquad \underline{c} \in \Delta, \quad \underline{a} \in \mathcal{Q}$$

where $\mathcal{Q}$ is the space of all real, nonnull p-dimensional vectors $\underline{a}'$, of a contrast-vector $\underline{c}'\tau$. See Roy and Bose [51].

Again suppose that the experimenter is interested in a particular set G_ν of contrast-vectors specified by a set $\Delta_\nu \subset \Delta$, $\underline{c} \in \Delta$, rank $(\Delta_\nu) \geqq \nu-1$. We can decompose the hypothesis of no treatment effects as

$$H_o = \bigcap_{\underline{c} \in \Delta_\nu} \bigcap_{\underline{a} \in \mathcal{Q}} [H_{o,\underline{c},\underline{a}}],$$

where $H_{o,\underline{c},\underline{a}}$: $\underline{c}'\tau\underline{a} = 0$.

Now consider the problem of testing H_o against

$$H_1 = \bigcup_{\underline{c} \in \Delta_\nu} \bigcup_{\underline{a} \in \mathcal{Q}} [H^{-1}_{o,\underline{c},\underline{a}}],$$

where $H^{-1}_{o,\underline{c},\underline{a}}$ denotes 'not $H_{o,\underline{c},\underline{a}}$'.

Suppose that the elements of $\underset{v\times p}{\tau}$ have for the maximum likelihood estimates the corresponding elements of a matrix $\underset{v\times p}{V}$. Also let $S_{\underline{c}}$ denote the covariance matrix of the estimate $\underline{c}V$ of the contrast-vector $\underline{c}'\tau$. Then the variance of the linear compound $\underline{c}'V\underline{a}$ is

$$\underline{a}' S_{\underline{c}} \underline{a} .$$

We have, then, a test for $H_{o,\underline{c},\underline{a}}$ against $H^{-1}_{o,\underline{c},\underline{a}}$ whose acceptance region is

$$A_{\underline{c},\underline{a}} : \frac{(\underline{c}'V\underline{a})^2}{\underline{a}'S_{\underline{c}}\underline{a}} \leqq \text{Constant} = \mu_\alpha^{(\nu)} \quad \text{(say)}$$

Then, as before, by the union intersection principle we accept H_o over

$$A^{(\nu)} = \bigcap_{\underline{c} \in \Delta_\nu} \bigcap_{\underline{a} \in \mathcal{Q}} [A_{\underline{c}, \underline{a}}]$$

$$= \bigcap_{\underline{c} \in \Delta_\nu} \left[\bigcap_{\underline{a} \in \mathcal{Q}} \left\{ \frac{(\underline{c}' V \underline{a})^2}{\underline{a}' S_{\underline{c}} \underline{a}} \leq \mu_\alpha^{(\nu)} \right\} \right] ,$$

which is known to be

$$= \bigcap_{\underline{c} \in \Delta_\nu} [T_{\underline{c}}^2 \leq \mu_\alpha^{(\nu)}] ,$$

where $T_{\underline{c}}^2$ is the Hotelling's T^2-statistic for testing $\underline{c}' \tau = \underline{0}$, and $\mu_\alpha^{(\nu)}$ is determined by

$$\Pr \left[\bigcap_{\underline{c} \in \Delta_\nu} \left\{ T_{\underline{c}}^2 \leq \mu_\alpha^{(\nu)} \right\} \middle| H_0 \right] = 1 - \alpha , \quad 0 < \alpha < 1 .$$

Notice that, as before the acceptance region $A^{(\nu)}$ gives us a set of simultaneous interval estimates, for all linear compounds of the contrast vectors $\underline{c}' \tau$, $\underline{c} \in \Delta_\nu$, namely

$$\left\{ \left[\underline{c}' V \underline{a} - \sqrt{\mu_\alpha^{(\nu)}} \sqrt{\underline{a}' S_{\underline{c}} \underline{a}} \leq \underline{c}' \tau \underline{a} \leq \underline{c}' V \underline{a} + \sqrt{\mu_\alpha^{(\nu)}} \sqrt{\underline{a}' S_{\underline{c}} \underline{a}} \right] : \underline{c} \in \Delta_\nu , \underline{a} \in \mathcal{Q} \right\}$$

with the simultaneous confidence coefficient equal to $1 - \alpha$.

4.3.3 Applications:

As in the univariate case we shall see that different subsets Δ_ν of Δ yield different sets of simultaneous confidence interval estimates.

(i) $\Delta_\nu = \Delta$.

In this we get the simultaneous interval estimates for all the linear compounds of all contrast vectors $\underline{c}' \tau$. The determination of the constant $\mu_\alpha^{(\nu)}$ is done by using the distribution of the maximum characteristic root. This is generalization of Scheffe's procedure to the multivariate set-up.

(ii) $\Delta_\nu = \Delta_o$.

We get the simultaneous confidence interval estimates on all the linear compounds of a set of orthogonal-contrast vectors.

(iii) $\Delta_\nu = \Delta_p$.

We get the estimates on all the linear compounds of the set of pairwise contrast vectors. The distribution problem is unsolved.

(iv) $\Delta_\nu = \Delta_f$.

We get the estimates on all the linear compounds of a set of comparisons with a control-contrast vector .

4.4 Comparison of various Bounds.

Now we shall state and prove a theorem concerning the nature of the constants $\mu_{(\nu),\alpha}$ and $\mu_\alpha^{(\nu)}$ as functions from the class of subsets Δ_ν of Δ to the positive real line.

Theorem: - $\mu_{(\nu),\alpha}$ and $\mu_\alpha^{(\nu)}$ are monotone set functions from the class of subsets Δ_ν of Δ to the positive real line.

Proof: We have to show that if $\Delta_\nu \subset \Delta_{\nu'}$ then

$$\mu_{(\nu),\alpha} \leq \mu_{(\nu'),\alpha} \quad \text{and}$$

$$\mu_\alpha^{(\nu)} \leq \mu_\alpha^{(\nu')} \quad .$$

To see this we recall that $\mu_{(\nu),\alpha}$ and $\mu_{\alpha}^{(\nu)}$ are defined by

$$\Pr.\left\{ \bigcap_{\underline{c} \in \Delta_\nu} \left[\frac{(\underline{c}'\,\underline{t})^2}{s_{\underline{c}}^2} \leqq \mu_{(\nu),\alpha}^2 \right] \mid H_o \right\} = 1 - \alpha$$

and $\Pr.\left\{ \bigcap_{\underline{c} \in \Delta_\nu} \left[T_{\underline{c}}^2 \leqq \mu_{\alpha}^{(\nu)} \right] \mid H_o \right\} = 1 - \alpha$

Now $\left\{ \bigcap_{\underline{c} \in \Delta_{\nu'}} \left[\frac{(\underline{c}'\,\underline{t})^2}{s_{\underline{c}}^2} \leqq \mu_{(\nu'),\alpha}^2 \right] \mid H_o \right\}$

implies

$$\left\{ \bigcap_{\underline{c} \in \Delta_\nu} \left[\frac{(\underline{c}'\,\underline{t})^2}{s_{\underline{c}}^2} \leqq \mu_{(\nu'),\alpha}^2 \right] \mid H_o \right\}$$

if

$$\Delta_\nu \subset \Delta_{\nu'} \quad .$$

$$\therefore \quad \Pr.\left\{ \bigcap_{\underline{c} \in \Delta_{\nu'}} \left[\frac{(\underline{c}'\,\underline{t})^2}{s_{\underline{c}}^2} \leqq \mu_{(\nu'),\alpha}^2 \right] \mid H_o \right\}$$

$$\leqq \quad \Pr.\left\{ \bigcap_{\underline{c} \in \Delta_\nu} \left[\frac{(\underline{c}'\,\underline{t})^2}{s_{\underline{c}}^2} \leqq \mu_{(\nu'),\alpha}^2 \right] \mid H_o \right\}.$$

This proves the result for the univariate case. The argument for $\mu_{\alpha}^{(\nu)}$ is analogous.

This theorem provides us with the explanation of the sharpness of the Dunnet bounds in comparison with the Tukey bounds and of the Tukey bounds as compared with the Scheffe' bounds when all three sets of bounds are obtained from the same data. The reason for the sharpness of the geometrical width of bounds on a set of orthogonal contrasts in comparison with that of the Scheffe bounds is the same. It is easy to see that these comparisons of the geometrical width extend to multivariate extensions of the various bounds. However we still have no means of comparing the sharpness of bounds on two sets of contrasts when one of the sets is not contained in the other set, e.g. the set of pairwise contrasts and the set of orthogonal contrasts.

BIBLIOGRAPHY

[1] Anderson, T. W. (1948). "The asymptotic distribution of the roots of certain determinantal equations," J. Roy. Stat. Soc. Ser. B 10 , 132-139.

[2] Anderson, T. W. (1950). "The asymptotic distribution of certain roots and vectors". Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Los Angeles, pp. 103-130.

[3] Anderson, T. W. (1955). "The integral of a symmetric unimodal function". Proc. Amer. Math. Soc., 6, 170-176.

[4] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis . New York, John Wiley and Sons, Inc.

[5] Bancroft, T. A. (1950). "Probability values for the common tests of significance". J. Amer. Statist. Assoc. 45, 211-217.

[6] Bellman, R. E. (1960). Introduction to Matrix Analysis . New York, McGraw-Hill.

[7] Duncan, D. B. (1955). "Multiple range and multiple F-tests". Biometrics 11, 1-42.

[8] Duncan, D. B. (1961). "Bayes rules for a common multiple comparison problem and related Student-t problems". Ann. Math. Stat. 32, 1013-1033 .

[9] Dunnet, C. W. (1955). "A multiple comparison procedure for comparing several treatments with a control". J. Amer. Stat. Assoc. 50, 1096-1122.

[10] Durand, D. (1954). "Joint confidence region for multiple regression coefficients". J. Amer. Stat. Assoc. 49, 130-146.

[11] Dwass, M. (1955). "A note on simultaneous confidence interval estimation". Ann. Math. Stat. 26 , 146-147.

[12] Everitt, W. N. (1962). "Two theorems in matrix theory". Amer. Math. Monthly 69, 856-859.

[13] Feldt, L. S. and Moharram, W. M. (1958). "Power function and charts for specifying number of observations in analysis of variance of fixed effects". Ann. Math. Stat. 29, 871-877.

[14] Fox, M. (1956). "Charts for the power of F-test". Ann. Math. Stat. 27, 484-497.

[15] Ghosh, M. N. (1953). "Simultaneous test for linear hypothesis by method of analysis of variance". Unpublished manuscript.

[16] Greenwood, J. A. and Hartley, H. O. (1962). Guide to Tables in Mathematical Statistics. Princeton University Press, Princeton, New Jersey.

[17] Hall, W. J. (1963). Personal Communication.

[18] Hartley, H. O. and Pearson, E. S. (1951). "Charts of the power function of analysis of variance tests". Biometrika 38, 112-130.

[19] Hoel, P. G. (1950). "Confidence regions for linear regressions", Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Los Angeles, 75-82.

[20] Hotelling, H. (1931). "The generalization of student's ratio". Ann. Math. Stat. 2 , 360-378.

[21] Hotelling, H. (1936). "Relations between two sets of variables," Biometrika 28, 280-287.

[22] Hotelling, H. (1951). "A generalized T-test and measure of multivariate dispersion." Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Los Angeles, 23-41.

[23] Hsu, P. L. (1940). "On generalized analysis of variance". Biometrika 31, 221-237.

[24] Hsu, P. L. (1941). "On the limiting distribution of roots of a determinantal equation". J. London Math. Soc. 6 , 185-189.

[25] Hsu, P. L. (1945). "On the power function of E^2 and T^2 tests". Ann. Math. Stat. 16 , 278 - 285.

[26] Ito, K. (1956). "Asymptotic formulae for the distribution of Hotelling's generalized T_o^2 statistics". Ann. Math. Stat. 27, 1091-1105.

[27] Ito, K. (1960). "On multivariate analysis of variance test". Bull. Internat. Inst. Stat. 38, 87-98.

[28] James, A. T. (1954). "Normal multivariate analysis and the orthogonal group". Ann. Math. Stat. 25, 40-75.

[29] Kendall, M. G. and Stuart, A. (1958). The Advanced Theory of Statistics. London, C. Griffin.

[30] Lawley, D. N. (1938). "A generalization of Fisher's Z-test." Biometrika 30, 180-187.

[31] Lawley, D. N. (1939). "A correction to 'A generalization of Fisher's Z-test'." Biometrika 30, 467.

[32] Lehmann, E. L. (1950). "Some Principles of the theory of hypothesis testing". Ann. Math. Stat. 21, 1-26.

[33] Lehmann, E. L. (1951). "A general concept of unbiasedness". Ann. Math. Stat. 22 , 587-592.

[34] Lehmann, E. L. (1955). "Ordered families of distributions". Ann. Math. Stat. 26, 399-419.

[35] Lehmann, E. L. (1959). Testing Statistical Hypotheses, New York, John Wiley and Sons, Inc.

[36] Lehmann, E. L. (1957). "A theory of some multiple decision procedures I." Ann. Math. Stat. 28, 1-25.

[37] Lehmer, E. (1944). "Inverse tables of probabilities of errors of the second kind". Ann. Math. Stat. 15, 388-398.

[38] Mandel, J. (1958). "A note on confidence regions in regression problems". Ann. Math. Stat. 29, 903-

[39] Mikhail, W. F. (1960). "On the Monotonicity and Admissibility of Some Tests in Multivariate Analysis". Ph.D. Thesis, University of North Carolina, Chapel Hill, N. C.

[40] Narain, R. D. (1950). "On the completely unbiased character of tests of independence in multivariate analysis". Ann. Math. Stat. 21, 293-298.

[41] Neyman, J. and Pearson, E. S. (1938). "Contributions to the theory of testing statistical hypotheses". Stat. Res. Mem. Vol. II, 25-27.

[42] Patnaik, P. B. (1949). "The noncentral χ^2 and F-distributions and their applications". Biometrika 36, 202-232.

[43] Pillai, K. C. S. (1955). "Some new criteria in multivariate analysis". Ann. Math. Stat. 26, 117-121.

[44] Ramachandran, K. V. (1954). "On certain tests and the monotonicity of their power functions". University of North Carolina Mimeo Series No. 120.

[45] Ramachandran, K. V. (1956). "Simultaneous analysis of variance test". Ann. Math. Stat. 27, 521-528.

[46] Rao, C. R. (1952). Advanced Statistical Methods in Biometric Research. New York, John Wiley and Sons, Inc.

[47] Roy, S. N. (1953). "On a heuristic method of test construction and its use in multivariate analysis". Ann. Math. Stat. 24, 220-238.

[48] Roy, S. N. (1954). "Some further results in simultaneous confidence interval estimation". Ann. Math. Stat. 25, 752-761.

[49] Roy, S. N. (1958). Some Aspects of Multivariate Analysis, New York, John Wiley and Sons, Inc.

[50] Roy, S. N. (1961). "Some remarks on normal multivariate analysis of variance". University of North Carolina Mimeo Series No. 304.

[51] Roy, S. N. and Bose, R. C. (1953). "Simultaneous interval estimation". Ann. Math. Stat. 24, 513-536.

[52] Roy, S. N. and Gnanadesikan, R. (1960). "Equality of two dispersion matrices against alternatives of intermediate specificity". Ann. Math. Stat. 33, 423-438.

[53] Roy, S. N. and Mikhail, W. F. (1961). "On the monotonic character of the power function of two multivariate tests". Ann. Math. Stat. 32, 1145-1151.

[54] Scheffe, H. (1953). "A method for judging all contrasts in the analysis of variance". Biometrika, 40, 87-104.

[55] Scheffe, H. (1959). Analysis of Variance. New York, John Wiley and Sons, Inc.

[56] Scheffe, H. (1961). "Simultaneous interval estimates of linear functions of parameters". Bull. Internat. Inst. Stat. 38, Part IV, 245-253.

[57] Stein, C. (1956). "The admissibility of Hotelling's T^2-test". Ann. Math. Stat. 27, 616-623.

[58] Tang, P. C. (1938). "The power function of the analysis of variance test with tables and illustrations of their use". Stat. Res. Mem. Vol. II, 126-149.

[59] Tukey, J. W. (1951). "Quick and dirty methods in statistics, Part II". Proc. Fifth Annual Convention, Amer. Soc. for Quality Control, 189-197.

[60] Tukey, J. W. (1952). "Allowances for various types of error rates". Unpublished invited address, Blacksburg meeting of the IMS, March 1952.

[61] Wald, A. (1942). "On the power function of the analysis of variance test". Ann. Math. Stat. 13, 434-439.

[62] Wilks, S. S. (1932). "Certain generalizations in the analysis of variance". Biometrika 24, 471-494.

[63] Working, H. and Hotelling, H. (1929). "Application of the theory of errors to interpretation of trends". J. Amer. Stat. Assoc., March Suppl., 73-85.

[64] Wolfowitz, J. (1949). "The power of classical tests associated with the normal distribution". Ann. Math. Stat. 20, 540-551.

INSTITUTE OF STATISTICS
NORTH CAROLINA STATE COLLEGE

(Mimeo Series available for distribution at cost)

328. Schutz, W. M. and C. C. Cockerham. The effect of field blocking on gain from selection. Ph.D. Thesis. August, 1962.
329. Johnston, Bruce O., and A. H. E. Grandage. A sampling study of estimators of the non-linear parameter in the exponential model. M.S. Thesis, June, 1962.
330. Hurst, D. C. and R. J. Hader. Modifications of response surface techniques for biological use. Ph.D. Thesis, June, 1962.
331. Potthoff, Richard F. A test of whether two regression lines are parallel when the variances may be unequal. August, 1962.
332. Searls, Donald. On the "large" observation problem. Ph.D. Thesis. August, 1962.
333. Gupta, Somesh Das. On the optimum properties of some classification rules. August, 1962.
334. Potthoff, Richard F. A test of whether two parallel regression lines are the same when the variances may be unequal. August, 1962.
335. Bhattacharyya, P. K. On an analog of regression analysis. August, 1962.
336. Fish, Frank. On networks of queues. 1962.
337. Warren, William G. Contributions to the study of spatial point processes. 1962.
338. Naor, P, Avi-Itzhak, B., On discretionary priority queueing. 1962.
339. Srivastava, J. N. A note on the best linear unbiased estimates for multivariate populations. 1962.
340. Srivastava, J. N. Incomplete multiresponse designs. 1962.
341. Roy, S. N. and Srivastava, J. N. Hierarchical and p-block multiresponse designs and their analysis. 1962.
342. Srivastava, J. N. On the monotonicity property of the three main tests for multivariate analysis of variance. 1962.
343. Kendell, Peter. Estimation of the mean life of the exponential distribution from grouped data when the sample is censored—with application to life testing. February, 1963.
344. Roberts, Charles D. An asymptotically optimal sequential design for comparing several experimental categories with a standard or control. 1963.
345. Novick, M. R. A Bayesian indifference procedure. 1963.
346. Johnson, N. L. Cumulative sum control charts for the folded normal distribution. 1963.
347. Potthoff, Richard F. On testing for independence of unbiased coin tosses lumped in groups too small to use χ^2.
348. Novick, M. R. A Bayesian approach to the analysis of data for clinical trials. 1963.
349. Sethuraman, J. Some limit distributions connected with fixed interval analysis. 1963.
350. Sethuraman, J. Fixed interval analysis and fractile analysis. 1963.
351. Potthoff, Richard F. On the Johnson-Neyman technique and some extensions thereof. 1963.
352. Smith, Walter L. On the elementary renewal theorem for non-identically distributed variables. 1963.
353. Naor, P. and Yadin, M. Queueing systems with a removable service stations. 1963.
354. Page, E. S. On Monte Carlo methods in congestion problems—I. Searching for an optimum in discrete situations. February, 1963.
355. Page, E. S. On Monte Carlo methods in congestion problems—II. Simulation of queueing systems. February, 1963.
356. Page, E. S. Controlling the standard deviation by cusums and warning lines.
357. Page, E. S. A note on assignment problems. March, 1963.
358. Bose, R. C. Strongly regular graphs, partial geometries and partially balanced designs. March, 1963.