ON CERTAIN TESTS AND THE MONOTONICITY OF THEIR POWERS

by

K. V. Ramachandran

Special report to the Office of Naval Research of work at Chapel Hill under Contract NR 042 031, Project N7-onr-284(02), for research in probability and Statistics.

Institute of Statistics
Mimeograph Series No. 120
November 16, 1954

ACKNOWLEDGMENT

I wish to express my sincere and heartfelt thanks to Professor S. N. Roy, who suggested the problem, and without whose constant advice and help the work would not have been completed.

To the Institute of Statistics, and particularly to Professor Harold Hotelling, I am obliged for the financial and other facilities which were provided me during the progress of this work.

My thanks are due to the United States Educational Foundation in India and the Institute of International Education for their financial assistance.

Finally, I wish to thank Mrs. Mary Ann Taylor for her careful typing of the manuscript.

TABLE OF CONTENTS

Table of Contents (Continued)

Table of Contents (Continued)

INTRODUCTION

The subject of this dissertation is the monotonicity property of a class of tests which are derived by the union-intersection principle [18][1], the tests being all within the Neyman-Pearson set-up of two-decision problems.

It is well known that most non-sequential parametric tests (within the Neyman-Pearson set-up), some of them in current use and some of them more or less recently proposed by various workers for different situations, have power functions with the following property. In each case the power happens to be a function of certain parameters, which are functions of the original parameters, being much less in number than the original parameters. Furthermore, this set of parametric functions is such that the set as a whole can be appropriately regarded as a measure of deviation from the hypothesis to be tested. For example, in most cases it turns out that every member of the set is zero, and in some cases every member is unity if and only if the hypothesis tested (to be called the null hypothesis) is true. For many of the current non-sequential parametric tests (to be called the classical tests), it is further known that this power function has the additional property of unbiasedness in the sense of Neyman; later designated

---

1. Numbers in square brackets refer to bibliography.

as complete or uniform unbiasedness in contrast with local unbiasedness, whose meaning is also obvious in the context of the Neyman Pearson theory. Many of these uniformly or completely unbiased tests happen to possess what is indeed a stronger property and one that automatically implies complete unbiasedness. This property is that the power function of the test monotonically increases, as each of the parameters involved in the power function (to be called the deviation parameters), tends away from its value on the null hypothesis; the value being usually zero or unity. This property has a considerable significance in terms of the loss functions of the general Wald theory, but this point need not be elaborated here. We shall call this property, when it exists, the monotonicity property of the power function and note that it implies the weaker property of complete unbiasedness.

In this dissertation a number of tests recently proposed in relation to means and variances of univariate and multivariate normal populations will be studied from this point of view, and it will be shown that many of these tests have the monotonicity property, while some of them are completely unbiased but do not have the monotonicity property. It is also shown, incidentally, that each of these tests can be derived from a uniform principle of test construction called the union intersection principle which yields in most cases a lower bound to the power function which turns out in many cases to be pretty good [19].

The monotonicity property of the multivariate analysis of variance test, and of the test for the independence of two sets of variables from a (p + q)-variate normal population, each test being based on the union-intersection principle, has been proved by Roy [19].

In Chapter I, we have proved, that under a certain partition of the tail areas, the two-sided F test for the equality of two variances from two univariate normal populations, has the monotonicity property. The case of one population follows as a corollary to the above result. Percentage points connected with the modified test are given for different values of the parameters.

In Chapter II, we have shown, that if we impose a certain restriction on the tail areas, we get a test for the equality of two covariance matrices from two multinormal populations, which has the monotonicity property; when all the characteristic roots of a certain matrix are equal. The case of one population follows as a corollary to the above result. The distribution problem connected with the test procedure is also given. The results given in this Chapter, generalize those given in Chapter I for multinormal populations.

In Chapter III, we have investigated certain power properties of the Tukey Studentized range q and the Hartley F_{max} ratio tests. We have shown that these tests are uniformly unbiased, but that the power functions do not have the monotonicity property. Also we have obtained certain useful lower bounds to the power functions of these tests. Multivariate extensions of these tests are also considered.

In Chapter IV, power properties of the sim. anova test are investigated. We have shown that the sim. anova test has the monotonicity property. We have also obtained lower bounds to certain probability statements connected with the sim. anova test. The distribution problem connected with the test has been solved and upper 5 per cent points of the Studentized largest chi-square are given for different values of the parameters.

In Chapter V, certain optimum power properties of the Studentized maximum modulus test have been proved. The distribution problem connected with the test was solved by the author jointly with Pillai [16]. It is shown that the Studentized maximum modulus test has the monotonicity property. Using a certain result due to Kimball [11], we have obtained useful lower bounds to certain probability statements connected with the test.

NOTATION

All vectors are column vectors and primes indicate their transposes.

"$X(p \times q)$" denotes "a matrix with p rows and q columns."

"$c(X)$" means "the characteristic roots of $X(p \times p)$."

"Sup $c(X)$" means "the largest $c(X)$."

"u_{max}" stands for "the maximum in the set $(u_1, \ldots, u_k)$."

"u_{min}" stands for "the minimum in the set $(u_1, \ldots, u_k)$."

"$|x|$" stands for "the absolute value of x."

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

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

"n.d." stands for "negative definite."

"p.d.f." stands for "probability density function."

"c.d.f." stands for "cumulative distribution function."

"d.f." stands for "degrees of freedom."

"$N(\xi, \sigma^2)$" stands for "a random variable x having a normal p.d.f. with mean ξ and variance σ^2."

"$N(\underline{\xi}, \Sigma)$" stands for "a random vector $\underline{x}(p \times 1)$ having a multinormal p.d.f. with mean vector $\underline{\xi}(p \times 1)$ and covariance matrix $\Sigma(p \times p)$."

CHAPTER I

UNIVARIATE TESTS ON VARIANCES FROM NORMAL POPULATIONS.

1.1 Case of one population. Let x_i ($i = 1, 2, \ldots, n$) be a random sample of size n from a $N(\xi, \sigma^2)$ population. To test the hypothesis $\frac{\sigma^2}{\sigma_0^2} = 1$ against all $\frac{\sigma^2}{\sigma_0^2} = \delta^2 \neq 1$, we have the test procedure:

(1.1.1) accept $\sigma^2 = \sigma_0^2$ if $\chi_1^2 \leq \chi^2 \leq \chi_2^2$, where

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \sum_1^n (x_i - \bar{x})^2 \Big/ \sigma_0^2 \; , \; \bar{x} = \sum_1^n x_i \Big/ n, \text{ and}$$

χ^2 is a chi-square variable with n-1 d.f, and reject the hypothesis otherwise.

The usual procedure is to choose χ_1^2 and χ_2^2 such that

$$(1.1.2) \qquad \int_0^{\chi_1^2} p(\chi^2/1; n-1)\, d\chi^2 = \int_{\chi_2^2}^{\infty} p(\chi^2/1; n-1)\, d\chi^2 = \frac{\alpha}{2},$$

where α is the given level of significance of the test.

$p(\chi^2/\delta^2; n-1)$ is defined to be equal to

$$(1.1.3) \qquad \text{Const.} \left(\frac{\chi^2}{\delta^2}\right)^{\frac{n-3}{2}} e^{-\frac{\chi^2}{2\delta^2}} .$$

If we do not impose the restriction (1.1.2), then the test procedure given above depends on two quantities χ_1^2 and χ_2^2 which we can choose in an infinite number of ways such that the level of significance of the test is α. Equation (1.1.2) gives one partition of the tail areas.

We shall choose the tail areas in such a way that the resulting test has the monotonicity property. This presumably can be accomplished in various ways, one being as follows:

Under the alternative hypothesis let P denote the power of the test procedure, i.e., let

$$(1.1.4) \qquad P = 1 - \int_{\chi_1'^2}^{\chi_2'^2} p(\chi^2/\delta^2;n-1)\, d\chi^2 ,$$

denote the power of the test procedure based on $\chi_1'^2$ and $\chi_2'^2$, where

$$(1.1.5) \qquad 1 - \alpha = \int_{\chi_1'^2}^{\chi_2'^2} p(\chi^2/1;n-1)\, d\chi^2 ,$$

and $\delta^2 = \sigma^2/\sigma_0^2 \neq 1$ is the deviation parameter. We shall choose $\chi_1'^2$ and $\chi_2'^2$ such that, in addition to (1.1.5), we also have $\partial P/\partial\delta^2 = 0$, when $\delta^2 = 1$, i.e., when $\sigma^2 = \sigma_0^2$.

This condition imposes a further restriction on $\chi_1'^2$ and $\chi_2'^2$, which, together with (1.1.5), fixes $\chi_1'^2$ and $\chi_2'^2$.

We shall presently show that under this condition the test procedure has the monotonicity property.

Now,

$$(1.1.6) \qquad P = 1 - c \int_{\chi_1'^2}^{\chi_2'^2} (\chi^2/\delta^2)^{\frac{n-3}{2}} e^{-\chi^2/2\delta^2} \, d(\chi^2/\delta^2)$$

is the power function of the test procedure. Notice that $c > 0$ is a pure constant independent of δ^2.

Hence

$$(1.1.7) \qquad \frac{\partial P}{\partial \delta^2} = \frac{-c}{\delta^2} \left[\left(\frac{\chi_1'^2}{\delta^2}\right)^{\frac{n-1}{2}} e^{-\frac{\chi_1'^2}{2\delta^2}} - \left(\frac{\chi_2'^2}{\delta^2}\right)^{\frac{n-1}{2}} e^{-\frac{\chi_2'^2}{2\delta^2}} \right].$$

The condition $\partial P/\partial\delta^2 = 0$ when $\delta^2 = 1$ is equivalent to

$$(1.1.8) \qquad (\chi_1'^2)^{\frac{n-1}{2}} e^{-\frac{\chi_1'^2}{2}} = (\chi_2'^2)^{\frac{n-1}{2}} e^{-\frac{\chi_2'^2}{2}},$$

that is,

$$\left(\frac{\chi_1'^2}{\chi_2'^2}\right)^{\frac{n-1}{2}} = e^{-\frac{1}{2}(\chi_2'^2 - \chi_1'^2)}.$$

Now (1.1.7) reduces to

$$(1.1.9) \qquad \frac{-c}{\delta^2}\left(\frac{X_2'^2}{\delta^2}\right)^{\frac{n-1}{2}} e^{-\frac{X_1'^2}{2\delta^2}} \left[\left(\frac{X_1'^2}{X_2'^2}\right)^{\frac{n-1}{2}} - e^{-\frac{1}{2\delta^2}(X_2'^2 - X_1'^2)}\right]$$

$$= \frac{-c}{\delta^2}\left(\frac{X_2'^2}{\delta^2}\right)^{\frac{n-1}{2}} e^{-\frac{X_1'^2}{2\delta^2}} \left[e^{-\frac{1}{2}(X_2'^2 - X_1'^2)} - e^{-\frac{1}{2\delta^2}(X_2'^2 - X_1'^2)}\right],$$

using (1.1.8).

It is now easy to check that

$$\partial P/\partial \delta^2 \quad > 0 \qquad \text{if } \delta^2 > 1, \text{ that is, if } \sigma^2 > \sigma_0^2,$$

and

$$< 0 \qquad \text{if } \delta^2 < 1.$$

Thus P, the power of the modified test is a monotonic increasing function of δ^2 if $\delta^2 > 1$, and a monotonic decreasing function of δ^2 if $\delta^2 < 1$. Hence the modified test has the monotonicity property. The values of $X_1'^2$ and $X_2'^2$ for $\alpha = .05$ are given in Table 1 (see appendix) for different values of n.

1.2 Case of two populations. Let x_{1i} $(i = 1, 2, \ldots n_1)$ and $x_{2i}(i = 1, 2, \ldots n_2)$ be independent samples of sizes n_1 and n_2 from $N(\xi_1, \sigma_1^2)$ and $N(\xi_2, \sigma_2^2)$ respectively. To test the hypothesis

$\sigma_1^2 = \sigma_2^2$ against ~~the~~ all alternatives $\sigma_1^2 \neq \sigma_2^2$ we have the test procedure:

accept $\sigma_1^2 = \sigma_2^2$, that is $\sigma_1^2/\sigma_2^2 = 1$, if $F_1 \leq F \leq F_2$, where

(1.2.1) $$F = s_1^2/s_2^2 = \frac{\sum_1^{n_1} (x_{1i}-\bar{x}_1)^2}{(n_1-1)} \frac{(n_2-1)}{\sum_1^{n_2}(x_{2i}-\bar{x}_2)^2},$$

$$n_j\bar{x}_j = \sum_{i=1}^{n_j} x_{ji}\ ,(j = 1,2),$$ and F has an ordinary F distribution with $n_1 - 1$, $n_2 - 1$ d.f., and reject the hypothesis otherwise.

The usual procedure is to choose F_1 and F_2 such that

(1.2.2) $$\int_0^{F_1} p(F/1;n_1 - 1, n_2 - 1)\, dF = \int_{F_2}^{\infty} p(F/1;n_1-1,n_2-1)dF = \frac{\alpha}{2},$$

where α is the given level of significance of the test.

$p(F/\delta^2; n_1 - 1, n_2 - 1)$ is defined to be equal to, $(\delta^2 = \sigma_1^2/\sigma_2^2 \neq 1)$,

(1.2.3) $$\text{Const.}\ (F/\delta^2)^{\frac{n_1-3}{2}} \Big/ \left[1 + (n_1-1)F/(n_2-1)\delta^2\right]^{\frac{n_1+n_2-2}{2}}.$$

If we do not impose the restriction (1.2.2), then the test procedure given above depends on two quantities F_1 and F_2 which we can choose in an infinite number of ways such that the level of significance of the test is α. Equation (1.2.2) gives one partition of the tail areas.

We shall choose the tail areas in such a way that the resulting test has the monotonicity property. This presumably can be accomplished in various ways, one being as follows:

Under the alternative hypothesis let P denote the power of the test procedure, i.e., let

$$(1.2.4) \qquad P = 1 - \int_{F_1'}^{F_2'} p(F/\delta^2; n_1 - 1, n_2 - 1)\, dF,$$

denote the power of the test procedure based on F_1' and F_2', where

$$(1.2.5) \qquad 1 - \alpha = \int_{F_1'}^{F_2'} p(F/1; n_1 - 1, n_2 - 1)\, dF; \text{and}$$

$\delta^2 = \sigma_1^2/\sigma_2^2 \neq 1$ is the deviation parameter. We shall choose F_1' and F_2' such that, in addition to (1.2.5), we also have $\partial P/\partial\delta^2 = 0$ when $\delta^2 = 1$, i.e., when $\sigma_1^2 = \sigma_2^2$. This condition imposes a further restriction on F_1' and F_2' which, together with

(1.2.5), fixes F_1' and F_2'.

We shall presently show that under this condition the test procedure has the monotonicity property.

Now,

$$(1.2.6) \quad P = 1-c \int_{F_1'}^{F_2'} (F/\delta^2)^{\frac{n_1-3}{2}} d(F/\delta^2) \Big/ \left[1+ {(n_1-1)F}/{(n_2-1)\delta^2}\right]^{\frac{n_1+n_2-2}{2}}$$

is the power function of the test procedure. Notice that $c > 0$ is a pure constant independent of δ^2.

Hence

$$(1.2.7) \qquad \frac{\partial P}{\partial \delta^2} = \frac{-c}{\delta^2} \left[(F_1'/\delta^2)^{\frac{n_1-1}{2}} \Big/ \left\{ 1+ \frac{(n_1-1)F_1'}{(n_2-1)\delta^2} \right\}^{\frac{n_1+n_2-2}{2}} \right.$$

$$\left. - (F_2'/\delta^2)^{\frac{n_2-1}{2}} \Big/ \left\{ 1 + \frac{(n_1-1)F_2'}{(n_2-1)\delta^2} \right\}^{\frac{n_1+n_2-2}{2}} \right] .$$

The condition $\partial P/\partial\delta^2 = 0$ when $\delta^2 = 1$ is equivalent to

$$(1.2.8) \qquad ({F'_1}/{F'_2})^{\frac{n_1-1}{2}} = \left[\frac{1 + \frac{n_1-1}{n_2-1} F'_1}{1 + \frac{n_1-1}{n_2-1} F'_2} \right]^{\frac{n_1+n_2-2}{2}} .$$

Now (1.2.7) reduces to

$$\frac{-c}{\delta^2}(F'_2/\delta^2)^{\frac{n_1-1}{2}} \left(1 + \frac{(n_1-1)F'_1}{(n_2-1)\delta^2}\right)^{-\frac{(n_1+n_2-2)}{2}} \left[(F'_1/F'_2)^{\frac{n_1-1}{2}} - \left\{ \frac{1 + \frac{(n_1-1)F'_1}{(n_2-1)\delta^2}}{1 + \frac{(n_1-1)F'_2}{(n_2-1)\delta^2}} \right\}^{\frac{n_1+n_2-2}{2}} \right]$$

$$= \frac{-c}{\delta^2}(F'_2/\delta^2)^{\frac{n_1-1}{2}} \left(1 + \frac{(n_1-1)F'_1}{(n_2-1)\delta^2}\right)^{-\frac{(n_1+n_2-2)}{2}} \left[\left\{ \frac{1 + \frac{(n_1-1)F'_1}{(n_2-1)}}{1 + \frac{(n_1-1)F'_2}{(n_2-1)}} \right\}^{\frac{n_1+n_2-2}{2}} - \left\{ \frac{1 + \frac{(n_1-1)F'_1}{(n_2-1)\delta^2}}{1 + \frac{(n_1-1)F'_2}{(n_2-1)\delta^2}} \right\}^{\frac{n_1+n_2-2}{2}} \right] ,$$

using (1.2.8).

It is now easy to check that

$$\partial P / \partial \delta^2 \quad > 0 \quad \text{if } \delta^2 > 1, \text{ that is, } \sigma_1^2 > \sigma_2^2 ,$$

and

$$< 0 \quad \text{if } \delta^2 < 1.$$

Thus P, the power of the modified test is a monotonic increasing function of δ^2 if $\delta^2 > 1$ and a monotonic decreasing function of δ^2 if $\delta^2 < 1$. Hence the modified test has the monotonicity property. The values of F_1' and F_2' for $\alpha = .05$ are given in Table 2 (see appendix) for different values of n_1 and n_2.

It is worth noting that the results for the case of one population can be derived from those for the case of two populations by making n_2 large. Also it is worth noting that the likelihood ratio criterion in either of these situations does not give a test which has the monotonicity property. As a matter of fact the tests based on the likelihood ratio criterion in the two situations are each biased.

CHAPTER II

MULTIVARIATE TESTS ON COVARIANCE MATRICES

2.1 Case of one population. Let $X(p \times n)$ be a random sample of size n from a $N(\xi, \Sigma)$ population. To test the hypothesis H_0: $\Sigma = \Sigma_0$, i.e., $\Sigma\Sigma_0^{-1} = I(p)$ against ~~the~~ all alternatives $\Sigma\Sigma_0^{-1} = \Gamma(p \times p) \neq I(p)$ we have the following test procedure [18], accept H_0 if

$$(2.1.1) \qquad \theta_0 \leq \theta_1 < \ldots < \theta_p \leq \theta_0'$$

and reject H_0 otherwise, where $\theta_1, \ldots, \theta_p$ are the p characteristic roots of $(n-1)S\Sigma_0^{-1}$, where $(n-1)S = XX' - n\,\overline{x}\,\overline{x}'$, $\overline{x}$ being the sample mean vector. θ_0 and θ_0' are so chosen that

$$(2.1.2) \qquad \int\limits_{\theta_0 \leq \theta's \leq \theta_0'} p(\theta_1, \ldots, \theta_p/\Sigma\Sigma_0^{-1} = I(p))\, d\theta_1 \ldots d\theta_p = 1 - \alpha,$$

α being the given level of significance of the test.

It is well known that under the null hypothesis

$$(2.1.3) \quad p(\theta_1, \ldots, \theta_p/\Sigma\Sigma_0^{-1} = I(p)) = \text{Const.} \prod_{i=1}^{p} \theta_i^{\frac{n-p-2}{2}} e^{-\frac{1}{2}\sum_{1}^{p}\theta_i} \prod_{i>j}(\theta_i - \theta_j)$$

$$0 < \theta_1 < \ldots < \theta_p < \infty .$$

The test procedure given above depends on two quantities θ_0 and θ_0' which we can choose in an infinite number of ways such that the level of significance of the test is α. We shall choose θ_0 and θ_0' such that in addition to (2.1.2) we have

$$(2.1.4) \qquad \left.\frac{\partial P}{\partial \gamma_i}\right]_{\underline{\gamma} = \underline{1}} = 0 \quad (i = 1, 2, \ldots, p),$$

where $\gamma_1, \ldots, \gamma_p (= \underline{\gamma}')$ are the characteristic roots of Γ, and P is the power of the test based on θ_0 and θ_0'. It is easy to check and it will be shown in (2.6) that the p equations (2.1.4) are all equivalent each to the other, each being also equivalent to $\left.\frac{\partial}{\partial \gamma}(P)_{\gamma_1, \ldots, \gamma_p = \gamma}\right]_{\gamma=1} = 0$. Thus (2.1.4) imposes just one further restriction on θ_0 and θ_0', which, together with (2.1.2), fixes θ_0 and θ_0'. We have shown that under this restriction the resulting test will have the monotonicity property if all the γ's are assumed to be equal and to stay equal.

It is conjectured that when the γ's are not equal, the test based on θ_0 and θ_0' will have the monotonicity property if either all γ's are > 1 or if all are < 1.

The distribution problem connected with the test will be solved in (2.5).

2.2 Case of two populations. Let $X_1(p \times n_1)$ and $X_2(p \times n_2)$ be random samples of sizes n_1 and n_2 from $N(\xi_1, \Sigma_1)$ and $N(\xi_2, \Sigma_2)$ respectively.

To test the hypothesis H_0: $\Sigma_1 = \Sigma_2$, i.e., $\Sigma_1\Sigma_2^{-1} = I(p)$ against ~~the~~ all alternatives $\Sigma_1\Sigma_2^{-1} = \Gamma(p \times p) \neq I(p)$ we have the following test procedure [18]:

accept H_0 if

$$(2.2.1) \qquad \theta_0 \leq \theta's \leq \theta_0' ,$$

and reject H_0 otherwise, where $\theta_1, \ldots, \theta_p$ are the characteristic roots of $S_1(S_1 + \frac{n_2-1}{n_1-1} S_2)^{-1}$, and $(n_i - 1) S_i = X_iX_i' - n_i \bar{x}_i\bar{x}_i'$, $\bar{x}_i$ being the sample mean vector of the i th population (i = 1, 2). θ_0 and θ_0' are so chosen that

$$(2.2.2) \qquad \int\limits_{\theta_0 \leq \theta's \leq \theta_0'} P(\theta_1, \ldots, \theta_p/\Sigma_1\Sigma_2^{-1} = I(p))d\theta_1\ldots d\theta_p = 1 - \alpha,$$

α being the given level of significance of the test.

It is well known that

$$(2.2.3) \quad p(\theta_1, \ldots, \theta_p/\Sigma_1\Sigma_2^{-1} = I(p)) = \text{Const.} \prod_1^p \theta_i^m(1-\theta_i)^n \prod_{i>j} (\theta_i - \theta_j) ,$$

where $m = \frac{n_1-p-2}{2}$ and $n = \frac{n_2-p-2}{2}$ and $0 < \theta_1 < \ldots < \theta_p < 1$.

The test procedure given above depends on two quantities θ_0 and θ_0' which we can choose in an infinite number of ways such that the level of significance of the test is α. We shall choose θ_0 and θ_0' such that in addition to (2.2.2), we have

$$\left.\frac{\partial P}{\partial \gamma_i}\right]_{\underline{\gamma} = \underline{1}} = 0 \quad (i = 1, 2, \ldots, p)\,, \tag{2.2.4}$$

where $\gamma_1, \ldots, \gamma_p (= \underline{\gamma}')$ are the characteristic roots of Γ, and P is the power of the test based on θ_0 and θ_0'. It is easy to check and it will be shown in (2.6) that the p equations (2.2.4) are all equivalent to $\frac{\partial}{\partial\gamma}(P)\Big]_{\gamma_1, \ldots, \gamma_p = \gamma,\ \gamma=1} = 0$. Thus (2.2.4) imposes just one further restriction on θ_0 and θ_0' which, together with (2.2.2) fixes θ_0 and θ_0'. We have shown that under this restriction the resulting test will have the monotonicity property if all the γ's are assumed to be equal and to stay equal.

It is conjectured that when the γ's are not equal, the test based on θ_0 and θ_0' will have the monotonicity property if either all γ's are > 1 or if all are < 1.

To solve the distribution problem connected with the test we need the joint distribution of the largest and smallest roots of certain determinental equations. We shall derive in the next few sections certain mathematical results which we shall use to obtain the joint c.d.f. of the largest and smallest roots.

2.3 Properties of a certain special function. Let

$$(2.3.1)\qquad \beta\left[x, y; (m_s, n_s; \ldots; m_1, n_1)\right]$$

$$= \int_x^y \int_x^{x_s} \cdots \int_x^{x_2} \prod_{i=1}^{s} dx_i \begin{vmatrix} x_s^{m_s} (1-x_s)^{n_s} & \cdots & x_1^{m_s} (1-x_1)^{n_s} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ x_s^{m_1} (1-x_s)^{n_1} & \cdots & x_1^{m_1} (1-x_1)^{n_1} \end{vmatrix}$$

$$= \beta\left[x, y; \begin{pmatrix} m_s, n_s & \cdots & m_1, n_1 \\ \cdot & \cdots & \cdot \\ m_s, n_s & \cdots & m_1, n_1 \end{pmatrix}\right] .$$

The last expression is in the form of a pseudo-determinant whose meaning is made clear by considering, for illustration, the case of $s = 3$ for which

$$(2.3.2)\qquad \beta\left[x, y; \begin{pmatrix} m_3, n_3 & m_2, n_2 & m_1, n_1 \\ \cdot & \cdot & \cdot \\ m_3, n_3 & m_2, n_2 & m_1, n_1 \end{pmatrix}\right]$$

$$= \int_x^y x_3^{m_3} (1-x_3)^{n_3} dx_3 \left[\int_x^{x_3} x_2^{m_2} (1-x_2)^{n_2} dx_2 \int_x^{x_2} x_1^{m_1} (1-x_1)^{n_1} dx_1\right.$$

$$- \int_x^{x_3} x_2^{m_1} (1-x_2)^{n_1} dx_2 \int_x^{x_2} x_1^{m_2} (1-x_1)^{n_2} dx_1 \Big]$$

$$- \int_x^{y} x_3^{m_2} (1-x_3)^{n_2} dx_3 \Big[\int_x^{x_3} x_2^{m_3} (1-x_2)^{n_3} dx_2 \int_x^{x_2} x_1^{m_1} (1-x_1)^{n_1} dx_1$$

$$- \int_x^{x_3} x_2^{m_1} (1-x_2)^{n_1} dx_2 \int_x^{x_2} x_1^{m_3} (1-x_1)^{n_3} dx_1 \Big]$$

$$+ \int_x^{y} x_3^{m_1} (1-x_3)^{n_1} dx_3 \Big[\int_x^{x_3} x_2^{m_3} (1-x_2)^{n_3} dx_2 \int_x^{x_2} x_1^{m_2} (1-x_1)^{n_2} dx_1$$

$$- \int_x^{x_3} x_2^{m_2} (1-x_2)^{n_2} dx_2 \int_x^{x_2} x_1^{m_3} (1-x_1)^{n_3} dx_1 \Big] .$$

In opening out the pseudo-determinant it is very important to stick to the order of the factors, indicated in the expansion on the right side of (2.3.2), and to keep in mind that the factors are non-commutative. It is also clear that the whole expression will be zero if any two columns become equal in

$$\begin{pmatrix} m_s, n_s & m_{s-1}, n_{s-1} & \cdots & m_1, n_1 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ m_s, n_s & m_{s-1}, n_{s-1} & \cdots & m_1, n_1 \end{pmatrix} .$$

Again let

(2.3.3) $\quad \beta(x,y;\, m_s,\, n_s;\, \ldots;\, m_1 n_1)$

$$= \int_x^y x_s^{m_s}(1-x_s)^{n_s}\, dx_s \int_x^{x_s} x_{s-1}^{m_{s-1}}\, (1-x_{s-1})^{n_{s-1}} dx_{s-1} \cdots$$

$$\cdots \int_x^{x_2} x_1^{m_1}\, (1-x_1)^{n_1}\, dx_1 \; ,$$

so that $\beta\, (x,y;\, m,n) = \int_x^y x_1^{m}\, (1-x_1)^n \; dx_1$.

Also let

(2.3.4) $\quad x^m(1-x)^n = \beta_0(x;\, m,n).$

Using (2.3.3), we can write (2.3.2) as

(2.3.5) $\Sigma(-1)^r \beta(x,y;\ m_3', n_3';\ m_2', n_2';\ m_1', n_1')$,

where $(m_3', n_3';\ m_2', n_2';\ m_1', n_1')$ is any permutation of $(m_3, n_3;\ m_2, n_2;\ m_1, n_1)$, the summation is taken over all such permutations; r being the total number of inversions of the order of the subscripts in $(m_3, n_3;\ m_2, n_2;\ m_1, n_1)$.

Care is to be taken to preserve on the right side of (2.3.2) the order of the operations with regard to the x's from x_3 through x_2 to x_1.

Similarly,(2.3.1) can be rewritten as

(2.3.6) $\Sigma(-1)^r \beta(x, y;\ m_s', n_s';\ m_{s-1}', n_{s-1}';\ \ldots;\ m_1', n_1')$,

where $(m_s', n_s';\ \ldots;\ m_1', n_1')$ is any permutation of $(m_s, n_s;\ \ldots;\ m_1, n_1)$, the summation is taken over all such permutations; r being the total number of inversions of the order of the subscripts in $(m_s, n_s;\ \ldots;\ m_1, n_1)$. Care is to be taken to preserve the order of the operations with regard to the x's from x_s through x_{s-1}, ... to x_1.

Lemma 1.

(2.3.7) $$\int_{x_0}^{y_0} x^m(1-x)^n f(x,x_0)\,dx = \frac{1}{m+n+1}\Big[-y_0^m(1-y_0)^{n+1} f(y_0,x_0)$$

$$+ x_0^m (1-x_0)^{n+1} f(x_0,x_0) + \int_{x_0}^{y_0} x^m (1-x)^{n+1} f'(x,x_0)\,dx$$

$$+ m \int_{x_0}^{y_0} x^{m-1} (1-x)^n f(x,x_0)\,dx \Big] ,$$

where $f'(x,x_0) = \frac{d}{dx} f(x,x_0)$.

Proof: The result follows immediately by integration by parts, integrating $(1-x)^{m+n}$, and differentiating $(\frac{x}{1-x})^m$ and $f(x,x_0)$. We shall assume here that $m, n > -1$, $0 \leq x_0 < y_0 \leq 1$, and $f(x,x_0)$ is such that $f'(x,x_0)$ and the two integrals on the right side of (2.3.7) all exist.

Lemma 2.

(2.3.8) $\Sigma\beta(x,y;\, m_s',n_s';\, m_{s-1}',\, n_{s-1}';\, \ldots;\, m_1',n_1') = \prod_{i=1}^{s} \beta(x,y;\, m_i,n_i)$,

where on the left side, $(m_s',n_s';\, \ldots;\, m_1',\, n_1')$ is any permutation of $(m_s,\, n_s;\, \ldots;\, m_1,\, n_1)$, the summation being taken over all such permutations.

Proof: The mechanism of the proof will be evident if we consider, for simplicity of algebra, the case of $s = 2$.

We have

$$(2.3.9) \quad \int_x^y x_2^{m_2} (1-x_2)^{n_2} \, dx_2 \int_x^{x_2} x_1^{m_1} (1-x_1)^{n_1} \, dx_1$$

$$+ \int_x^y x_2^{m_1} (1-x_2)^{n_1} \, dx_2 \int_x^{x_2} x_1^{m_2} (1-x_1)^{n_2} \, dx_1$$

$$= \int_x^y x_2^{m_2} (1-x_2)^{n_2} \, dx_2 \int_x^{x_2} x_1^{m_1} (1-x_1)^{n_1} \, dx_1$$

$$+ \int_x^y x_2^{m_2} (1-x_2)^{n_2} \, dx_2 \int_{x_2}^{y} x_1^{m_1} (1-x_1)^{n_1} \, dx_1$$

(obtained by interchanging, in the second term on the left side of (2.3.9), the variables x_2 and x_1, and rewriting the domain of integration in the appropriate manner)

$$= \int_x^y x_2^{m_2} (1-x_2)^{n_2} \, dx_2 \int_x^{y} x_1^{m_1} (1-x_1)^{n_1} \, dx_1$$

$$= \beta(x,y;\ m_2,n_2)\ \beta\ (x,y;\ m_1,n_1).$$

Lemma 3.

$$(2.3.10)\quad \sum_r \beta_r(x,y;\ m_{s-1},\ n_{s-1};\ \ldots;\ m_{r+1},n_{r+1};\ m,n;\ m_{r-1},n_{r-1};\ \ldots\ m_1,n_1)$$

$$= \beta(x,y;\ m,n)\ \beta\ (x,y;\ m_{s-1},n_{s-1};\ \ldots;\ m_{r+1},n_{r+1};\ m_{r-1},n_{r-1};\ \ldots;\ m_1,\ n_1),$$

where β_r ( ) is the result of putting (m,n) in the r th place and filling up the other positions with (m_{s-1},n_{s-1}), ..., (m_1,n_1); r going from 1 to s. Note that each β_r ( ) is an s-fold integral, while $\beta(x,y;\ m_{s-1},n_{s-1};\ \ldots;\ m_{r+1},n_{r+1};\ m_{r-1},n_{r-1};\ \ldots;\ m_1,n_1)$ is an (s-1) fold integral.

Proof: The mechanism of the proof is brought out by considering, in particular, the case s=3, where we have

$$(2.3.11)\quad \beta_1(x,y;\ m_2,n_2;m_1,n_1;\ m,n) + \beta_2(x,y;\ m_2,n_2;\ m,n;\ m_1,n_1) + \beta_3(x,y;\ m,n;\ m_2,n_2;\ m_1,n_1)$$

$$= \int_x^y x_3^{m_2}(1-x_3)^{n_2}\,dx_3 \int_x^{x_3} x_2^{m_1}(1-x_2)^{n_1}\,dx_2 \int_x^{x_2} x_1^{m}(1-x_1)^{n}\,dx_1$$

$$+ \int_x^y x_3^{m_2}(1-x_3)^{n_2}dx_3 \int_x^{x_3} x_2^{m_1}(1-x_2)^{n_1}dx_2 \int_{x_2}^{x_3} x_1^{m}(1-x_1)^{n}dx_1$$

$$+ \int_x^y x_3^{m_2}(1-x_3)^{n_2}dx_3 \int_x^{x_3} x_2^{m_1}(1-x_2)^{n_1}dx_2 \int_{x_3}^{y} x_1^{m}(1-x_1)^{n}dx_1$$

(by interchanging the variables and adjusting the domain of integration)

$$= \beta\ (x,y;\ m,n)\ \beta\ (x,y;\ m_2,n_2;\ m_1,\ n_1).$$

Lemma 4.

(2.3.12) $$\sum_r (-1)^{r-1}\ \beta_r \left[x,y; \begin{pmatrix} m_s,n_s & \cdots & m_1,n_1 \\ \cdot & \cdots & \cdot \\ m_s,n_s & \cdots & m_1,n_1 \end{pmatrix} \right]$$

$$= \sum_r (-1)^{r-1}\ \beta(x,y;\ m'_{s-r+1},n'_{s-r+1})\ \beta_{rr} \left[x,y; \begin{pmatrix} m_s,n_s & \cdots & m_1,n_1 \\ \cdot & \cdots & \cdot \\ m_s,n_s & \cdots & m_1,n_1 \end{pmatrix} \right],$$

where $\beta_r [\quad]$ on the left side is the result of replacing the r th row of $\beta [\quad]$ by $(m'_s,n'_s;\ \ldots;\ m'_1,n'_1)$, and $\beta_{rr} [\quad]$ on the right side is the result of suppressing the r th row and r th column of $\beta[\quad]$. Notice that $\beta_r[\quad]$ is an s x s, and $\beta_{rr}[\quad]$ is an (s -1) x (s - 1) pseudo-determinant.

Proof: The proof will be obvious if we consider, for simplicity of algebra, the case of s = 3, and pick out from the expansion of each pseudo-determinant on the left side of (2.3.12), (for the case s = 3), the term involving the index, say, (m_3', n_3') and then put together all such terms (with index m_3', n_3'). We have thus the following contribution from such terms

(2.3.13)
$$\beta\,(x,y;\ m_3',n_3';\ m_2,n_2;\ m_1,n_1)$$
$$-\ \beta\,(x,y;\ m_3',n_3';\ m_1,n_1;\ m_2,n_2)$$
$$+\ \beta\,(x,y;\ m_2,n_2;\ m_3',n_3';\ m_1,n_1)$$
$$-\ \beta\,(x,y;\ m_1,n_1;\ m_3',n_3';\ m_2,n_2)$$
$$+\ \beta\,(x,y;\ m_2,n_2;\ m_1,n_1;\ m_3',n_3')$$
$$-\ \beta\,(x,y;\ m_1,n_1;\ m_2,n_2;\ m_3',n_3')$$
$$=\ \beta\,(x,y;\ m_3',n_3')\ [\beta\,(x,y;\ m_2,n_2;\ m_1,n_1)$$
$$-\ \beta\,(x,y;\ m_1,n_1;\ m_2,n_2)]$$

(using lemma 3)

$$= \beta\ (x,y;\ m_3',n_3')\ \beta\ [\,x,y;\ \begin{pmatrix} m_2,n_2 & m_1,n_1 \\ m_2,n_2 & m_1,n_1 \end{pmatrix}\,]$$

$$= \beta\ (x,y;m_3',n_3')\ \beta_{11}\ [\,x,y;\ \begin{pmatrix} m_3,n_3 & m_2,n_2 & m_1,n_1 \\ m_3,n_3 & m_2,n_2 & m_1,n_1 \\ m_3,n_3 & m_2,n_2 & m_1,n_1 \end{pmatrix}\,]$$

(by definition).

This shows that if, in the general case, from the expansion of each pseudo-determinant (with the proper sign) on the left side of (2.3.12) we pick out the term with the index (m_s',n_s') and add together such terms (with the same index (m_s',n_s')), we shall have the following contribution

$$(2.3.14)\quad \beta\ (x,y;\ m_s',n_s')\ \beta_{11}\ [\,x,y;\begin{pmatrix} m_s,n_s & \cdots & m_1,n_1 \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_s,n_s & \cdots & m_1,n_1 \end{pmatrix}\,],$$

whence the proof is obvious by combining together different expressions like (2 .3.14) involving the different indices (m_r',n_r') $(r = 1, \ldots, s)$.

2.4 Reduction and evaluation of the integral:

$$\beta \left[x,y; \begin{pmatrix} m_s,n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ m_s,n & \cdots & m_1,n \end{pmatrix} \right] , \tag{2.4.1}$$

where $m_s > \ldots > m_1 > -1$, $n > -1$; and the m's differ by integers. We have already seen from (2.3.6) that the pseudo-determinant can be expanded into $\Sigma\, (-1)^r\, \beta\, (x,y;\, m_s',n;\, \ldots;\, m_1',n)$, where $(m_s', \ldots, m_1')$ is any permutation of $(m_s, \ldots, m_1)$, the summation is taken over all such permutations, s ! in number; and r is the total number of inversions of the order of the subscripts in $(m_s', \ldots, m_1')$. Recalling from (2.3.2) that β will be zero if any two columns of the pseudo determinant become equal, let us try to reduce m_s to m_{s-1} by successive integration by parts. To this end consider the typical term in the expansion. The largest exponent in this will of course be m_s. To reduce this by 1 we proceed as follows:

We have

$$\beta\, (x,y;\, m_s',n\,;\, \ldots;\, m_{r+1}',n;\, m_s,n;\, m_{r-1}',n;\, \ldots;\, m_1',n) \tag{2.4.2}$$

$$= \Big[\int_x^y x_s^{m_s'} (1-x_s)^n \, dx_s \cdots \int_x^{x_{r+2}} x_{r+1}^{m_{r+1}'} (1-x_{r+1})^n \, dx_{r+1}$$

$$\cdot \int_x^{x_{r+1}} x_r^{m_s} (1-x_r)^n dx_r \int_x^{x_r} x_{r-1}^{m_{r-1}'} (1-x_{r-1})^n dx_{r-1} \cdots \int_x^{x_2} x_1^{m_1'} (1-x_1)^n dx_1 \Big].$$

Using (2.3.7), we get

$$(2.4.3) \quad \Big[\int_x^{x_{r+1}} x_r^{m_s} (1-x_r)^n dx_r \int_x^{x_r} x_{r-1}^{m_{r-1}'} (1-x_{r-1})^n \, dx_{r-1} \cdots$$

$$\int_x^{x_2} x_1^{m_1'} (1-x_1)^n \, dx_1 \Big]$$

$$= \int_x^{x_{r+1}} x_r^{m_s} (1-x_r)^n dx_r \, \beta(x, x_r; m_{r-1}', n; \ldots; m_1', n)$$

$$= \frac{1}{m_s+n+1} \Big[- x_{r+1}^{m_s} (1-x_{r+1})^{n+1} \beta(x, x_{r+1}; m_{r-1}', n; \ldots; m_1', n)$$

$$+ x^{m_s} (1-x)^n \beta (x,x;\ m'_{r-1},n;\ \ldots;\ m'_1,n)$$

$$+ \int_x^{x_{r+1}} x_r^{m_s} (1-x_r)^{n+1}\, dx_r \beta' (x,x_r;\ m'_{r-1},n;\ \ldots;\ m'_1,n)$$

$$+ m_s \int_x^{x_{r+1}} x_r^{m_s-1} (1-x_r)^n\, dx_r\ \beta\ (x,x_r;\ m'_{r-1},\ n;\ \ldots;\ m'_1,n) \Big]$$

$$= \frac{1}{m_s+n+1} \Big[-x_{r+1}^{m_s} (1-x_{r+1})^{n+1} \beta(x,x_{r+1};\ m'_{r-1},n;\ \ldots;\ m'_1,n)$$

$$+ x^{m_s} (1-x)^{n+1} \beta(x,x;\ m'_{r-1},n;\ \ldots;\ m'_1,n)$$

$$+ \beta\ (x,x_{r+1};\ m_s + m'_{r-1},\ 2n+1;\ m'_{r-2},\ n;\ \ldots;\ m'_1,\ n)$$

$$+ m_s\ \beta\ (x,\ x_{r+1};\ m_s-1,\ n;\ m'_{r-1},n;\ \ldots;\ m'_1,\ n) \Big] .$$

(note that on the right side of (2.4.3), the first, second and third β's are each an $(r-1)$ fold integral, while the fourth β is an r fold integral). Now using (2.4.3), we have (2.4.2) reducing to

$$(2.4.4)\quad \left[-\beta\,(x,y;\,m_s',n;\,\ldots;\,m_{r+1}'+m_s,\,2n+1;\,m_{r-1}',\,n;\,\ldots;\,m_1',n) \right.$$

$$+\,\beta\,(x,y;m_s',n;\,\ldots;\,m_{r+1}',\,n;\,m_{r-1}'+m_s,\,2n+1;\,\ldots;\,m_1',n)$$

$$\left. +\,m_s\,\beta\,(x,y;\,m_s',n;\,\ldots;\,m_{r+1}',n;\,m_s-1,n;\,\ldots;\,m_1',n) \right]$$

$$\cdot \left[\frac{1}{(m_s+n+1)} \right],$$

where the first and second β's are each an (s-1) fold integral, while the third β is an s fold integral with index reduced to $m_s - 1$. It is easy to check that this reduction holds for $r = s-1, \ldots, 2$. If $r = s$, it is easy to check that (2.4.4) will be replaced by

$$(2.4.5)\quad \left[-\beta_0\,(y;\,m_s,n+1\,)\,\beta\,(x,y;\,m_{s-1}',n;\,\ldots;\,m_1',n) \right.$$

$$+\,\beta\,(x,y;\,m_{s-1}'+m_s,\,2n+1;\,m_{s-2}',n;\,\ldots;\,m_1',\,n)$$

$$\left. +\,m_s\,\beta\,(x,y;\,m_s-1,n;\,m_{s-1}',n;\,\ldots;\,m_1',n) \right] \Big/ (m_s+n+1),$$

and, if $r = 1$, (2.4.4) will be replaced by

$$(2.4.6) \quad \left[- \beta\,(x,y;\, m_s',n;\, \ldots;\, m_3',n;\, m_2' + m_s,\, 2n+1) \right.$$

$$+ \beta_0\,(x;\, m_s,\, n+1)\,\beta\,(x,y;\, m_s',\, n;\, \ldots;\, m_3',n;\, m_2',n)$$

$$\left. + m_s\,\beta\,(x,y;\, m_s',\, n;\, \ldots;\, m_2',n;\, m_s-1,\, n) \right] \Big/ (m_s+n+1) \quad .$$

We shall now introduce the rather convenient notations

$$(2.4.7) \quad \beta\,(x,y;\, m_s',n;\, \ldots;\, m_{r+1}' + m_s,\, 2n+1;\, m_{r-1}',n;\, \ldots;\, m_1',n)$$

$$= \beta\,(x,y;\, m_s',n;\, \ldots;\, m_{r+1}',n;\, \overleftarrow{m_s},\overleftarrow{n+1};\, \ldots;\, m_1',n) \ ,$$

where $(\overleftarrow{m_s},\overleftarrow{n+1})$ is supposed to be added to the (m_{r+1}',n) on the left so as to reduce the integral by one dimension;

$$(2.4.8) \quad \beta_0\,(y;\, m_s,n+1)\,\beta\,(x,y;\, m_{s-1}',n;\, \ldots;\, m_1',n)$$

$$= \beta\,(x,y;\, \overleftarrow{m_s},\, \overleftarrow{n+1}\,;\, m_{s-1}',n;\, \ldots;\, m_1',n);$$

$$(2.4.9) \quad \beta\,(x,y;\, m_s',n;\, \ldots;\, m_{r-1}' + m_s,\, 2n+1;\, m_{r-2}',n;\, \ldots;\, m_2',n)$$

$= \beta\ (x,y;\ m'_s,n;\ \ldots;\ \overrightarrow{m_s},\ \overrightarrow{n+1};\ m'_{r-1},n;\ m'_{r-2},n;\ \ldots;\ m'_1,n),$

where $(\overrightarrow{m_s},\ \overrightarrow{n+1})$ is supposed to be added to the $(m'_{r-1},\ n)$ on the right so as to reduce the integral by one dimension; and

(2.4.10) $\beta_0\ (x;\ m_s,\ n+1)\ \beta\ (x,y;\ m'_s,n;\ \ldots;\ m'_3,n;\ m'_2,n)$

$$= \beta\ (x,y;\ m'_s,n;\ \ldots;\ m'_2,n;\ \overrightarrow{m_s},\ \overrightarrow{n+1}).$$

Hence using (2.4.2) - (2.4.10), we have

(2.4.11) $$\beta \left[x,y; \begin{pmatrix} m_s,n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ m_s,n & \cdots & m_1,n \end{pmatrix} \right]$$

$$= \frac{-1}{m_s+n+1}\ \beta \left[x,y; \begin{pmatrix} \overleftarrow{m_s},\overleftarrow{n+1} & m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \overleftarrow{m_s},\overleftarrow{n+1} & m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right]$$

$$+ \frac{1}{m_s+n+1}\ \beta \left[x,y; \begin{pmatrix} \overrightarrow{m_s},\overrightarrow{n+1} & m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \overrightarrow{m_s},\overrightarrow{n+1} & m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right]$$

$$+ \frac{m_s}{m_s+n+1} \; \beta \,[\, x,y; \begin{pmatrix} m_s-1,n & m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ m_s-1,n & m_{s-1},n & \cdots & m_1,n \end{pmatrix}] \, .$$

Recalling the notations introduced earlier, and using lemma 4, it is easy to see that

$$(2.4.12) \quad \beta \,[\, x,y; \begin{pmatrix} \overleftarrow{m_s,n+1} & m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \overleftarrow{m_s,n+1} & m_{s-1},n & \cdots & m_1,n \end{pmatrix}]$$

$$= \beta_0 \,(y;\, m_s,n+1)\; \beta \,[\, x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix}]$$

$$+ \sum_{r=1}^{s-1} (-1)^r \; \beta_r \,[\, x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ m_s+m_{s-1},2n+1 & \cdots & m_s+m_1,2n+1 \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix}] \, ,$$

where $\beta_r \,[\quad]$ is an (s-1) fold pseudo-determinant obtained by substituting $(m_s + m_{s-1}, 2n+1), \ldots, (m_s + m_1, 2n+1)$ in the r the row of the (s-1) fold pseudo-determinant

$$\beta \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right] .$$

Thus (2.4.12) is equal to

$$\beta_0\,(y;\, m_s, n+1)\beta \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right]$$

$$+ \sum_{r=1}^{s-1} (-1)^r \beta\,(x,y;\, m_s+m_{s-r};\, 2n+1)\;\; \beta_{rr} \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right],$$

where $\beta_{rr} [\quad]$ is the (s-2) fold integral obtained by suppressing the r th row and r th column of the (s-1) fold pseudo determinant $\beta [\quad]$.

Similarly

$$(2.4.13) \quad \beta \left[x,y; \begin{pmatrix} \overrightarrow{m_s}, \overrightarrow{n+1} & m_{s-1}, n & \cdots & m_1, n \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \overrightarrow{m_s}, \overrightarrow{n+1} & m_{s-1}, n & \cdots & m_1, n \end{pmatrix} \right]$$

$$= (-1)^{s-1} \beta_0 (x; m_s, n+1) \beta \left[x,y; \begin{pmatrix} m_{s-1}, n & \cdots & m_1, n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1}, n & \cdots & m_1, n \end{pmatrix} \right]$$

$$+ \sum_{r=1}^{s-1} (-1)^{r-1} \beta (x,y; m_s + m_{s-r}, 2n+1) \beta_{rr} \left[x,y; \begin{pmatrix} m_{s-1}, n & \cdots & m_1, n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1}, n & \cdots & m_1, n \end{pmatrix} \right].$$

Using (2.4.12) and (2.4.13), we find that (2.4.11) becomes equal to

(2.4.14)

$$\frac{2}{m_s+n+1} \sum_{r=1}^{s-1} (-1)^{r-1} \beta(x,y; m_s+m_{s-r}, 2n+1) \beta_{rr} \left[x,y; \begin{pmatrix} m_{s-1}, n & \cdots & m_1, n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1}, n & \cdots & m_1, n \end{pmatrix} \right]$$

$$- \left[\beta_0 (y;m_s,n+1) + (-1)^s \beta_0(x;m_s,n+1) \right] \beta \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right]$$

$$+ \frac{m_s}{m_s+n+1} \beta \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right] .$$

Notice that the expression on the left side of (2.4.11) is an s th order pseudo-determinant, while on the right side of (2.4.14), the second $\beta[\ \ ]$ is an (s-1) th order pseudo-determinant, the **first** group of terms involves $\beta_{rr}[\ \ ]$, each such $\beta_{rr}[\ \ ]$ being an (s-2)th order pseudo-determinant, and the last term has a $\beta[\ \ ]$ which is an s th order pseudo-determinant. It may also be noticed that $\beta_{rr}[\ \ ]$ can be written as

$$\beta \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_{r+1},n & m_{r-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_{r+1},n & m_{r-1},n & \cdots & m_1,n \end{pmatrix} \right] .$$

(2.4.14) thus gives a recurrence relation, whereby proceding along the chain and reducing m_s to m_{s-1} (in which case the pseudo-determinant will vanish) we have the following reduction of the integral by one dimension.

$$(2.4.15) \qquad \beta \left[x,y; \begin{pmatrix} m_s,n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_s,n & \cdots & m_1,n \end{pmatrix} \right]$$

$$= - \beta \left[x,y; \begin{pmatrix} m_{s-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_1,n \end{pmatrix} \right]$$

$$\cdot \sum_{r'=1}^{m_s-m_{s-1}} \left[\beta_0(y;m_s-r'+1,n+1) - (-1)^s \beta_0(x;m_s-r'+1,n+1) \right]$$

$$\cdot \left[(m_s)_{r'-1}/(m_s+n+1)_{r'} \right]$$

$$+\left\{ 2 \sum_{r=1}^{s-1} \sum_{r'=1}^{m_s-m_{s-1}} (-1)^{r-1} \beta\, (x,y;\, m_s + m_{s-r}-r'+ 1,\, 2n+1) \right\}$$

$$\cdot\ \beta\ [\ x,y;\ \begin{pmatrix} m_{s-1},n & \cdots & m_{r+1},n & m_{r-1},n & \cdots & m_1,n \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ m_{s-1},n & \cdots & m_{r+1},n & m_{r-1},n & \cdots & m_1,n \end{pmatrix}\ ]\ \frac{(m_s)'_{r-1}}{(m_s+n+1)'_r},$$

where $(m)_p = m(m-1)\ \ldots\ (m-p+1)$. The s th order pseudo-determinant is thus thrown back on (s-1)th and (s-2)th order pseudo-determinants, and so on till we get to first order pseudo-determinants which are easily evaluated from the incomplete beta function tables [13].

2.5 On the joint c.d.f. of the largest and smallest roots.

We have noticed in (2.2) that in testing for the equality of covariance matrices from two multivariate normal populations we run into the joint distribution of the largest and smallest roots of a determinental equation: $\left| S_1 - \theta(S_1 + cS_2) \right| = 0$, where $c = \frac{n_2-1}{n_1-1}$ and $(n_i-1)S_i = X_iX_i' - n_i\ \bar{x}_i\ \bar{x}_i'$ $(i = 1,2)$. Note that S_1 and S_2 will be a.e.p.d.

Now

(2.5.1) $\quad P(\sigma < \theta_0 \leq \theta's \leq \theta_0' < 1\ )$

$$= c(p,m,n) \begin{vmatrix} \int_{\theta_0}^{\theta_0'} \theta_p^{m+p-1} (1-\theta_p)^n \, d\theta_p & \cdots & \int_{\theta_0}^{\theta_2} \theta_1^{m+p-1} (1-\theta_1)^n d\theta_1 \\ \vdots & \cdots & \vdots \\ \int_{\theta_0}^{\theta_0'} \theta_p^{m} (1-\theta_p)^n \, d\theta_p & \cdots & \int_{\theta_0}^{\theta_2} \theta_1^{m} (1-\theta_1)^n \, d\theta_1 \end{vmatrix} ,$$

where $m = \frac{n_1-p-2}{2}$ and $n = \frac{n_2-p-2}{2}$, and $c(p,m,n) =$

$$\left[\pi^{\frac{p}{2}} \prod_{1}^{p} \Gamma\left(\frac{2m+2n+p+i+2}{2}\right) \right] \Big/ \left[\prod_{1}^{p} \Gamma\left(\frac{2n+i+1}{2}\right) \Gamma\left(\frac{2m+i+1}{2}\right) \Gamma\left(\frac{i}{2}\right) \right] .$$

Now, using the results given in (2.3) and (2.4), it is easy to obtain the final reduction for the exact c.d.f. of the largest and smallest roots.

For $p = 2$,

(2.5.2) $$P_2 = P(\theta_0 \leq \theta\text{'s} \leq \theta_0') = \frac{c(2,m,n)}{(m+n+2)} \left[2\beta(\theta_0,\theta_0';\ 2m+1,2n+1) \right.$$

$$\left. - \beta\,(\theta_0,\theta_0';\ m,n) \left\{ \beta_0(\theta_0';\ m+1,n+1) + \beta_0(\theta_0;\ m+1,n+1) \right\} \right] .$$

For $p = 3$,

$$(2.5.3) \quad P_3 = P(\theta_0 \leq \theta's \leq \theta_0') = \frac{c(3,m,n)}{(m+n+3)} \Big[2\beta(\theta_0,\theta_0';\ m,n)$$

$$\cdot\ \beta\ (\theta_0,\theta_0';\ 2m+3,\ 2n+1) - 2\beta\ (\theta_0,\theta_0';\ m+1,n)$$

$$\cdot\ \beta\ (\theta_0,\theta_0';\ 2m+2,2n+1) \quad \frac{-P_2}{c(2,m,n)} \Big\{ \ \beta_0\ (\theta_0';\ m+2,\ n+1)$$

$$-\ \beta_0\ (\theta_0;\ m+2,\ n+1) \Big\} \Big].$$

For $p = 4$,

$$(2.5.4) \quad P_4 = P\ (\theta_0 \leq \theta's \leq \theta_0')$$

$$= \frac{c(4,m,n)}{(m+n+4)} \Big[\ 2\beta\ (\theta_0,\ \theta_0'\ ;\ 2m+5,\ 2n+1) \frac{P_2}{c(2,m,n)}$$

$$\frac{-\ P_3}{c(3,m,n)} \Big\{ \beta_0(\theta_0';\ m+3,n+1) + \beta_0(\theta_0;\ m+3,n+1) \Big\}$$

$$+2\beta\ \frac{(\theta_0,\theta_0';\ 2m+3,2n+1)}{(m+n+3)} \Big\{ -\beta_0(\theta_0';\ m+2,n+1) \quad \beta(\theta_0,\theta_0';m+1,n)$$

$$-\beta_0(\theta_0;\ m+2,n+1)\ \beta(\theta_0,\theta_0';\ m+1,n) + 2\beta(\theta_0,\ \theta_0';\ 2m+3,2n+1) \Big\}$$

$$- 2\beta \frac{(\theta_0,\theta_0';2m+4,2n+1)}{(m+n+3)} \Big\{ -\beta_0 \ (\theta_0'; \ m+2,n+1) \ \beta \ (\theta_0,\theta_0';m,n)$$

$$- \beta_0 \ (\theta_0; \ m+2,n+1) \ \beta \ (\theta_0,\theta_0'; \ m,n) + 2\beta \ (\theta_0,\theta_0'; \ 2m+2,2n+1)$$

$$+ \frac{(m+2)}{c(2,m,n)} \ P_2 \Big\} \Big] .$$

The joint c.d.f. of the largest and smallest roots in the case of one population can be easily obtained from the results given in (2.5) by making n_2 large.

It is worth noting at this point that the c.d.f. of the largest or smallest root can easily be obtained from the expressions (2.5.2) - (2.5.4). The joint c.d.f. of the largest and smallest roots for $p > 4$ can easily be obtained from (2.4.14). But since the expressions are lengthy, they are not given here.

2.6 Power function of the test procedure given in (2.2).

It has been shown in [19] that the second kind of error of the test procedure given in (2.2.1) is

(2.6.1) $$\beta = \text{Const.} \int_D \prod_1^p \gamma_i^{\frac{-n_1+1}{2}} \text{Exp} \Big[-\frac{1}{2} \text{ tr} \Big\{ D_{\frac{1}{\gamma}} \ Z_1Z_1' + Z_2Z_2' \Big\} \Big] dZ_1 dZ_2 ,$$

where D is the domain: $\Big\{ \theta_0 \leq c(Z_1Z_1' \ (Z_2Z_2')^{-1}) \leq \theta_0' \Big\}$.

We shall show that when $\gamma_1 = \gamma_2 = \ldots = \gamma_p = \gamma$, i.e., $\Sigma_1 \Sigma_2^{-1} = \gamma\ I\ (p)$, i.e., $\Sigma_1 = \gamma\ \Sigma_2$,

(2.6.2) $\frac{\partial\beta}{\partial\gamma} < 0$ if $\gamma > 1$

and > 0 if $\gamma < 1$

provided that we choose θ_0 and θ_0' in such a way that in addition to (2.2.2), we also have

(2.6.3) $\frac{\partial\beta}{\partial\gamma} = 0$ when $\gamma = 1$.

In the case where the γ's are not equal it is conjectured that

(2.6.4) $\frac{\partial\beta}{\partial\gamma_i} < 0$ if γ min > 1

and > 0 if γ max < 1

provided that we choose θ_0 and θ_0' in such a way that in addition to (2.2.2), we also have

(2.6.5) $\frac{\partial\beta}{\partial\gamma_i} = 0$ if $\underline{\gamma} = \underline{1}$ $(i = 1, \ldots, p)$.

There are p equations in (2.6.5). We shall show that these p equations are all equivalent each to the other, each being equivalent to $\frac{\partial\beta}{\partial\gamma}\Big\}_{\gamma_1,\gamma_2,\ldots,\gamma_p=\gamma}\Big]_{\gamma=1} = 0$. Thus (2.6.5) imposes just one further restriction on θ_0 and θ_0' which, together with (2.2.2), fixes θ_0 and θ_0'. To prove that the p equations (2.6.5) are all equivalent to just one condition $\frac{\partial\beta}{\partial\gamma}\Big]_{\gamma=1} = 0$, we proceed as follows:

Differentiating β given in (2.6.1) with respect to γ_i, we get,

(2.6.6) $$\frac{\partial\beta}{\partial\gamma_i} = \text{Const.} \int_D \prod_1^p \gamma_k^{\frac{-n_1+1}{2}} \operatorname{Exp}\Big[-\frac{1}{2}\operatorname{tr}\Big\{D_{\frac{1}{\gamma}} Z_1 Z_1' + Z_2 Z_2'\Big\}\Big]$$

$$\cdot\Big[-\frac{n_1-1}{2}\,\gamma_i^{-1} + \frac{1}{2}(Z_1 Z_1')_{ii}\,\gamma_i^{-2}\Big]\,dZ_1 dZ_2.$$

Now make use of the transformation

(2.6.7) $$Z_1(p \times n_1-1) = U(p \times p)\, D_{\sqrt{\theta}}(p)\; L_1(p \times n_1-1), \text{ and}$$

$$Z_2(p \times n_2-1) = U(p \times p)\, L_2\,(p \times n_2-1),$$ where θ's are the $c(Z_1 Z_1' (Z_2 Z_2')^{-1})$ and L_1 and L_2 are such that $L_1 L_1' = L_2 L_2' = I(p)$.

Under this transformation, (after integrating out L_1 and L_2), we get [19], (2.6.8) $$\frac{\partial\beta}{\partial\gamma_i} = \text{Const.} \int_{D'} \prod_1^p \gamma_k^{\frac{-n_1+1}{2}} \text{Exp} \left[-\frac{1}{2}\text{tr} \left\{ D_{\frac{1}{\gamma}} UD_\theta U' + UU' \right\} \right]$$

$$\cdot \left[-\frac{n_1-1}{2} \gamma_i^{-1} + \frac{1}{2} (UD_\theta U')_{ii} \gamma_i^{-2} \right] \left[|U|^{n_1+n_2-p} \prod_1^p \theta_i^m \, d\theta_i \prod_{i>j} (\theta_i - \theta_j) dU \right],$$

where D' is the domain: $\left\{ \begin{array}{c} \theta_0 \leq \theta's \leq \theta_0' \\ -\infty \leq u_{ij} \leq \infty \end{array} \right\}$, and $\theta_1 < \theta_2 < \ldots < \theta_p$.

Now if $\underline{\gamma} = \underline{1}$, i.e., if $\Sigma_1 = \Sigma_2$, we get

(2.6.9) $$\left. \frac{\partial\beta}{\partial\gamma_i} \right]_{\underline{\gamma}=\underline{1}} = \text{Const.} \int_{D'} \text{Exp} - \frac{1}{2} \left[\text{tr } UD_{1+\theta} U' \right]$$

$$\cdot \left[-\frac{n_1-1}{2} + \frac{1}{2} (UD_\theta U')_{ii} \right] \left[|U|^{n_1+n_2-p} dU \prod_1^p \theta_i^m d\theta_i \prod_{i>j} (\theta_i - \theta_j) \right]$$

$$= \text{Const.} \int_{D'} \text{Exp} - \frac{1}{2} \left[\text{tr } UD_{1+\theta} U' \right] \left\{ -\frac{n_1-1}{2} + \frac{1}{2} \sum_{k=1}^p u_{ik}^2 \theta_k \right\}$$

$$\cdot \left[|U|^{n_1+n_2-p} \prod_1^p \theta_i^m \, d\theta_i \prod_{i>j} (\theta_i - \theta_j) \, dU \right].$$

Similarly

$$(2.6.10) \quad \frac{\partial\beta}{\partial\gamma_j}\Big]_{\gamma=1} = \text{Const.} \int_{D'} \text{Exp}\left[-\tfrac{1}{2}\,\text{tr}\,U D_{1+\theta} U'\right]$$

$$\cdot\left[|U|^{n_1+n_2-p} \prod_1^p \theta_i^m \, d\theta_i \prod_{i>j} (\theta_i-\theta_j) dU\right] \left\{-\frac{n_1-1}{2} + \frac{1}{2}\sum_{k=1}^{p} u_{jk}^2\theta_k\right\},$$

$$(\, j \neq i \,).$$

Now since the u_{ij}'s are to be integrated out in the region $-\infty \leq u_{ij} \leq \infty$, and since the expressions in curly brackets on the right side of (2.6.9) and (2.6.10) are symmetric in θ's, it is evident that

$$(2.6.11) \quad \frac{\partial\beta}{\partial\gamma_i}\Big]_{\gamma=1} = \frac{\partial\beta}{\partial\gamma_j}\Big]_{\gamma=1} .$$

Hence the p equations $\frac{\partial\beta}{\partial\gamma_i}\Big]_{\gamma=1} = 0$ $(i = 1,\ldots,p)$ are all equivalent.

Now adding the p equations, we get,

$$(2.6.12) \quad \frac{\partial\beta}{\partial\gamma_1}\Big]_{\gamma=1} + \ldots + \frac{\partial\beta}{\partial\gamma_p}\Big]_{\gamma=1}$$

$$= \text{Const.} \int_{D'} \text{Exp}\left[-\frac{1}{2} \operatorname{tr} UD_{1+\theta}U' \right] \left[|U|^{n_1+n_2-p} dU \prod_1^p \theta_i d\theta_i \prod^m (\theta_i-\theta_j) \right]$$

$$\cdot \left\{ -\frac{n_1-1}{2} p + \frac{1}{2} \sum_1^p \sum_1^p u_{ij}^2 \theta_j \right\} .$$

Now when $\gamma_1 = \ldots = \gamma_p = \gamma$, we have,

(2.6.13) $$\frac{\partial\beta}{\partial\gamma}\Big]_{\gamma=1} = \text{Const.} \int_{D'} \text{Exp}\left[-\frac{1}{2} \operatorname{tr} U D_{1+\theta} U' \right]$$

$$\cdot \left[|U|^{n_1+n_2-p} dU \prod_1^p \theta_i d\theta_i \prod_{i>j}^m (\theta_i-\theta_j) \right] \left\{ -\frac{n_1-1}{2} p + \frac{1}{2} \sum_1^p \sum_1^p u_{ij}^2 \theta_j \right\}$$

$$= \frac{\partial\beta}{\partial\gamma_1}\Big]_{\underline{\gamma}=\underline{1}} + \ldots + \frac{\partial\beta}{\partial\gamma_p}\Big]_{\underline{\gamma}=\underline{1}} .$$

Hence $\frac{\partial\beta}{\partial\gamma_i}\Big]_{\underline{\gamma}=\underline{1}} = 0$ is equivalent to $\frac{\partial\beta}{\partial\gamma}\Big]_{\gamma=1} = 0$.

We shall now prove the following theorem:

Theorem 1. If in the set up given in (2.2) all the γ's are equal and equal to γ, say, then the power function P of the test procedure given in (2.2) will be a monotonic increasing function of γ if $\gamma > 1$, and a monotonic decreasing function of γ is $\gamma < 1$, provided that we choose θ_0 and θ_0' in such a way that in addition to (2.2.2), we also have $\frac{\partial P}{\partial \gamma}\Big]_{\gamma=1} = 0$.

Proof: From (2.6.1), the second kind of error of the test procedure is

(2.6.14) $$\beta = \text{Const.} \int_D \gamma^{-\frac{n_1-1}{2}p} \operatorname{Exp}\Big[-\frac{1}{2}\operatorname{tr}\Big\{\frac{1}{\gamma} Z_1Z_1' + Z_2Z_2'\Big\}\Big] dZ_1 dZ_2$$

$$= \text{Const.} \int_{D_1} \operatorname{Exp}\Big[-\frac{1}{2}\operatorname{tr}\Big\{Z_1Z_1' + Z_2Z_2'\Big\}\Big] dZ_1 dZ_2 ,$$

where D_1 is the domain: $\Big\{\frac{\theta_0}{\gamma} \leq c(Z_1Z_1'(Z_2Z_2')^{-1}) \leq \frac{\theta_0'}{\gamma}\Big\}$.

Now using the transformation given in (2.6.7) and integrating out L_1, L_2 and U, we get,

(2.6.15) $$\beta = \text{Const.} \int_{D_1} \prod_1^p \theta_i^m (1+\theta_i)^{-n'} \prod_{i>j} (\theta_i-\theta_j)\, d\theta_i$$

(where $n' = m+n+p+1$)

$$= \text{Const.} \int_{D_1} f(\theta_1, \ldots, \theta_p)\, d\theta_1 \ldots d\theta_p \quad , \qquad \text{(say).}$$

Now it is easy to verify that

(2.6.16) $$\frac{\partial \beta}{\partial(\frac{1}{\gamma})} = c \Big[\theta_0' \int_{\frac{\theta_0}{\gamma}}^{\frac{\theta_0'}{\gamma}} \int_{\frac{\theta_0}{\gamma}}^{\theta_{p-1}} \cdots \int_{\frac{\theta_0}{\gamma}}^{\theta_2} f(\theta_1, \ldots, \theta_{p-1}, \frac{\theta_0'}{\gamma}) \prod_1^{p-1} d\theta_i$$

$$- \theta_0 \int_{\frac{\theta_0}{\gamma}}^{\frac{\theta_0'}{\gamma}} \int_{\frac{\theta_0}{\gamma}}^{\theta_p} \cdots \int_{\frac{\theta_0}{\gamma}}^{\theta_3} f(\frac{\theta_0}{\gamma}, \theta_2, \ldots, \theta_p) \prod_2^p d\theta_i \Big]$$

(where c is positive)

$$= \frac{c\theta_0' (\frac{\theta_0'}{\gamma})^{m}}{(1+\frac{\theta_0'}{\gamma})^{n'}} \int_{\frac{\theta_0}{\gamma}}^{\frac{\theta_0'}{\gamma}} \int_{\frac{\theta_0}{\gamma}}^{\theta_{p-1}} \cdots \int_{\frac{\theta_0}{\gamma}}^{\theta_2}$$

$$\left[\prod_{1}^{p-1} \theta_i^{m} (1+\theta_i)^{-n'} d\theta_i \prod_{i>j=1}^{p-2} (\theta_i-\theta_j) \prod_{1}^{p-1} \left(\frac{\theta_0'}{\gamma} - \theta_i\right) \right]$$

$$\frac{-c\theta_0 \left(\frac{\theta_0}{\gamma}\right)^m}{\left(1+\frac{\theta_0}{\gamma}\right)^{n'}} \int_{\frac{\theta_0}{\gamma}}^{\frac{\theta_0'}{\gamma}} \int_{\frac{\theta_0}{\gamma}}^{\theta_p} \cdots \int_{\frac{\theta_0}{\gamma}}^{\theta_3}$$

$$\left[\prod_{2}^{p} \theta_i^{m} (1+\theta_i)^{-n'} d\theta_i \prod_{i>j=2}^{p-1} (\theta_i-\theta_j) \prod_{2}^{p} \left(\theta_i - \frac{\theta_0}{\gamma}\right) \right]$$

$$= K_1(\gamma)\, I_1(\gamma) - K_2(\gamma)\, I_2(\gamma) \,, \qquad \text{(say).}$$

The proof of the theorem will be complete if we show that

$$(2.6.17) \qquad \frac{\partial \beta}{\partial(\frac{1}{\gamma})} \quad > 0 \qquad \text{if } \gamma > 1$$

and $\quad < 0 \qquad \text{if } \gamma < 1$

subject to the condition

$$\frac{\partial \beta}{\partial (\frac{1}{\gamma})} = 0 \quad \text{if } \gamma = 1. \tag{2.6.18}$$

Condition (2.6.18) is equivalent to

$$\frac{c\theta_0'(\theta_0')^m}{(1+\theta_0')^{n'}} \int_{\theta_0}^{\theta_0'} \int_{\theta_0}^{\theta_{p-1}} \cdots \int_{\theta_0}^{\theta_2}$$

$$\left[\prod_1^{p-1} \theta_i^{\,m} (1+\theta_i)^{-n'} d\theta_i \prod_{i>j=1}^{p-2} (\theta_i - \theta_j) \prod_1^{p-1} (\theta_0' - \theta_i) \right]$$

$$\frac{-c\theta_0(\theta_0)^m}{(1+\theta_0)^{n'}} \int_{\theta_0}^{\theta_0'} \int_{\theta_0}^{\theta_p} \cdots \int_{\theta_0}^{\theta_3}$$

$$\left[\prod_2^{p} \theta_i^{\,m} (1+\theta_i)^{-n'} d\theta_i \prod_{i>j=2}^{p-1} (\theta_i - \theta_j) \prod_2^{p} (\theta_i - \theta_0) \right]$$

$$= K_1 I_1 - K_2 I_2 \text{ , (say)}$$

$$= 0.$$

Hence the proof will be complete if we show that if $\gamma > 1$,

$$(2.6.19) \quad \frac{(1+\frac{\theta_0}{\gamma})^{n'}}{(1+\frac{\theta_0'}{\gamma})^{n'}} \quad \frac{I_1(\gamma)}{I_2(\gamma)} > (\frac{\theta_0}{\theta_0'})^{m+1} \quad ,$$

$$\text{i.e.,} > (\frac{1+\theta_0}{1+\theta_0'})^{n'} \quad \frac{I_1}{I_2}$$

and if $\gamma < 1$,

$$\frac{(1+\frac{\theta_0}{\gamma})^{n'}}{(1+\frac{\theta_0'}{\gamma})^{n'}} \quad \frac{I_1(\gamma)}{I_2(\gamma)} < (\frac{1+\theta_0}{1+\theta_0'})^{n'} \quad \frac{I_1}{I_2} \; .$$

Now if $\gamma > 1$, $$\frac{1+\frac{\theta_0}{\gamma}}{1+\frac{\theta_0'}{\gamma}} > \frac{1+\theta_0}{1+\theta_0'}$$

and if $\gamma < 1$, $\quad \dfrac{1+\frac{\theta_0}{\gamma}}{1+\frac{\theta_0'}{\gamma}} < \dfrac{1+\theta_0}{1+\theta_0'}$.

Thus, the proof will be complete if we show that $I_1(\gamma)$ is an increasing function of γ and $I_2(\gamma)$ is a decreasing function of γ.

Now

$$(2.6.20)\quad \frac{\partial}{\partial(\frac{1}{\gamma})} I_1(\gamma) = - \frac{\theta_0(\frac{\theta_0}{\gamma})^m}{(1+\frac{\theta_0}{\gamma})^{n'}} \int_{\frac{\theta_0}{\gamma}}^{\frac{\theta_0'}{\gamma}} \int_{\frac{\theta_0}{\gamma}}^{\theta_{p-1}} \cdots \int_{\frac{\theta_0}{\gamma}}^{\theta_3}$$

$$\left[\prod_2^{p-1} \theta_i^m (1+\theta_i)^{-n'} d\theta_i \prod_{i>j=2}^{p-2} (\theta_i-\theta_j) \prod_2^{p-1} (\theta_i-\frac{\theta_0}{\gamma}) \prod_2^{p-1} (\frac{\theta_0'}{\gamma}-\theta_i)(\frac{\theta_0'}{\gamma}-\frac{\theta_0}{\gamma})\right]$$

$$< 0$$

and

$$\frac{\partial}{\partial(\frac{1}{\gamma})} I_2(\gamma) = \frac{\theta_0'(\frac{\theta_0'}{\gamma})^m}{(1+\frac{\theta_0'}{\gamma})^{n'}} \int_{\frac{\theta_0}{\gamma}}^{\frac{\theta_0'}{\gamma}} \int_{\frac{\theta_0}{\gamma}}^{\theta_{p-1}} \cdots \int_{\frac{\theta_0}{\gamma}}^{\theta_3}$$

$$\Big[\prod_{2}^{p-1} \theta_i^{m} (1+\theta_i)^{-n'} d\theta_i \prod_{i>j=2}^{p-2} (\theta_i-\theta_j) \prod_{2}^{p-1} \left(\frac{\theta_0'}{\gamma}-\theta_i\right) \prod_{2}^{p-1} \left(\theta_i-\frac{\theta_0}{\gamma}\right)\left(\frac{\theta_0'}{\gamma}-\frac{\theta_0}{\gamma}\right) \Big]$$

$$> 0.$$

Hence $I_1(\gamma)$ is a decreasing function of $\frac{1}{\gamma}$ and hence is an increasing function of γ, and $I_2(\gamma)$ is an increasing function of $\frac{1}{\gamma}$ and hence is a decreasing function of γ.

Hence the theorem.

Making n_2 large, we get, as a corollary to theorem 1, the following theorem:

<u>Theorem 2</u>. If in the set up given in (2.1), all the γ's are equal and equal to γ, say, then the power function P of the test procedure given in (2.1) will be a monotonic increasing function of γ if $\gamma > 1$ and a monotonic decreasing function of γ if $\gamma < 1$, provided we choose θ_0 and θ_0' in such a way that in addition to (2.1.2), we also have $\left[\frac{\partial P}{\partial \gamma}\right]_{\gamma=1} = 0$.

CHAPTER III

TUKEY TEST ON THE EQUALITY OF MEANS AND HARTLEY TEST ON THE EQUALITY OF VARIANCES

3.1 Introduction. In testing the equality of means of k univariate normal populations with a common but unknown variance σ^2, Fisher proposed the analysis of variance z or the equivalent F test based on the ratio of two independent mean squares. This test has several optimum properties including that of the monotonicity of its power function. As an alternative procedure Tukey [23] proposed the short cut test based on the Studentized range q. This test has the advantage that it is rather easy to carry out from the arithmetical point of view. We have shown that the q test is completely unbiased, but its power function does not have the monotonicity property. Another feature of this test is that its power depends on k - 1 parameters, whereas, in the case of the anova F test, the power depends only on one simple function of the deviation parameters. Moreover it is well known [17], [21] that the anova F test is an all contrast test for the means where as the q test is a test built around all two by two differences of the means and hence is a test related to a sub set of all contrasts. It is also well known that the anova F test is a likelihood ratio test.

When analysing data the experimenter is frequently faced with the necessity of testing the homogeneity in a set of estimated variances. When it is desired to combine a number of variances to obtain an estimate of the common variance, it is necessary to apply such a test. For general use in such cases Neyman and Pearson [12] have suggested a test, namely, the L_1 test. The L_1 test has been modified by Bartlett [1]. The exact distribution problem connected with the likelihood as well as the modified Bartlett's criterion is rather difficult. Several approximations have been suggested for the distribution of the likelihood ratio criterion [2], [9] and percentage points tabulated [22]. G. W. Brown [3] has shown that the power function of the likelihood ratio criterion is unbiased when the d.f. of all the sample variances are equal. Also he has shown that when the d.f. are not equal the likelihood ratio criterion is biased but is unbiased in the limit. It is easy to check that the likelihood ratio criterion test does not have the monotonicity property. Also it is known that the power of the likelihood ratio criterion depends on k - 1 parameters.

Cochran [4] has suggested for use in these situations a rather simple test based on

$$w_k = s^2_{max} / \sum_1^k s^2_i \text{, where } s^2_1, s^2_2, \ldots, s^2_k \text{ are the k sample}$$

variances with n.d.f. each. The distribution of w_k has been tabulated [7]. No study of the power function of this test is yet available.

Hartley [10] has intuitively suggested for use in these situations a test based on the statistic $F_{max} = s^2_{max} / s^2_{min}$. He recommends the short cut test based on F_{max} when each s_i^2 is based on the same number of d.f. It is easy to check that the Hartley F_{max} ratio acceptance region is the intersection of the acceptance regions based on all the $\binom{k}{2}$ two by two variance ratio F's. The distribution problem connected with the F_{max} ratio test has been solved and percentage points tabulated [5]. We have shown that the F_{max} ratio test is completely unbiased, but its power function does not have the monotonicity property. Also the F_{max} test depends on k - 1 parameters.

An all contrast test in the case of testing for the equality of several variances is unknown.

The Hartley F_{max} ratio test whose power properties we shall study in detail, being a test built around two by two ratios of variances, is a test related to a subset of all contrasts. An all contrast for k variances will be $\prod_1^k \sigma_i^{2c_i}$ subject to $\sum_1^k c_i = 0$.

3.2 <u>The Tukey q test</u>. Let x_{ij} (i=1, ..., k; j=1, ..., n) be k independent samples of size n from $N(\xi_i, \sigma^2)$ (i=1,...,k) population. Also let s^2 be an independent estimate of σ^2 based on m.d.f. (say, the error mean square in anova). To test the hypothesis: $\xi_1 = \dots = \xi_k$, the anova F test of Fisher which is also equivalent to an all contrast test of the form

$\sum_1^k c_i \xi_i = 0$ for all c_i subject to $\sum_1^k c_i = 0$, is well known. In this situation it is also well known that if the null hypothesis is not true, then the power of the test would involve as a deviation parameter only the quantity $\eta^2 = \sum_1^k (\xi_i - \bar{\xi})^2/\sigma^2$ where $\bar{\xi} = \sum_1^k \xi_{i/k}$. Also it is well known that the power of the test is a monotonic increasing function of the absolute value $/\eta/$. We shall show that an alternative test for the equality of several means from normal populations with a common variance σ^2 based on the Studentized range q which is much simpler than the anova F test from the computational point of view has the following properties. The power of the test would involve as parameters the k - 1 differences $\eta_{i-1} = \xi_1 - \xi_i$ (i = 2, ..., k). The test is completely unbiased, but the power function does not have the monotonicity property.

3.3 Power function of the q test. Under the set up given in (3.2), $\bar{x}_i = \sum_{j=1}^n x_{ij}/n$ is $N(\xi_i, \frac{\sigma^2}{n})$. Let s^2 be an independent and unbiased estimate of σ^2 (say, the error mean square in anova) with m.d.f.

The hypothesis H_0: $\xi_1 = \dots = \xi_k$ is equivalent to the hypothesis H_{ij}: $\xi_i = \xi_j$ (all $i \neq j$). Now for any two ξ's, the hypothesis $\xi_i = \xi_j$ can be tested using Fisher's 't' with m.d.f. The hypothesis $\xi_i = \xi_j$ is accepted if $\left|\bar{x}_i - \bar{x}_j\right| \leq t_\alpha s\sqrt{\frac{2}{n}}$, where t_α is the upper α point of Fishers 't' with m.d.f. Now since H_0 is equivalent to H_{ij}, (for all $i \neq j = 1,\dots,k$), we get a test of H_0 as follows: Take the intersection of all the $\binom{k}{2}$ two by two Fishers 't_{ij}' acceptance regions and accept H_0 if

$$\text{largest } |t_{ij}| = \sup_{i \neq j} |\bar{x}_i - \bar{x}_j| \Big/ s\sqrt{\frac{2}{n}} \leq t_\alpha .$$

It is easy to check that this is the same as accepting H_0 if $q = \frac{\bar{x}_{max} - \bar{x}_{min}}{s} \leq Q \sqrt{\frac{2}{n}}$, where Q is the upper α point of the Studentized range with m.d.f. This is the Tukey q test.

If the hypothesis H_0 is true, then

$$(3.3.1) \qquad 1 - \alpha = P \left[\frac{\bar{x}_{max} - \bar{x}_{min}}{s\sqrt{\frac{2}{n}}} \leq Q \right]$$

$$= \sum_{i=1}^{k} \int_0^\infty p(s)ds \int_{-\infty}^{\infty} p(\bar{x}_i)d\bar{x}_i \left[\prod_{j \neq i}^{k} \int_{\bar{x}_i}^{\bar{x}_i + Q's} p(\bar{x}_j)d\bar{x}_j \right] ,$$

where

$$p(\bar{x}) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2}\frac{\bar{x}^2}{\sigma^2}} ,$$

$$p(s) = \frac{2(\frac{m}{2})^{\frac{m}{2}}}{\Gamma(\frac{m}{2})\sigma^m} s^{m-1} e^{-\frac{ms^2}{2\sigma^2}} ,$$

$$Q' = Q\sqrt{\frac{2}{n}} ,$$

and α is the given level of significance of the test.

If H_0 is not true, then the second kind of error of the test is

$$(3.3.2) \qquad \beta = P \left[\frac{\bar{\bar{x}}_{max} - \bar{\bar{x}}_{min}}{s} \leq Q' \middle/ \underline{\xi} \right]$$

$$= \sum_{i=1}^{k} \int_0^{\infty} p(s)ds \int_{-\infty}^{\infty} p(\bar{x}_i;\xi_i)\, d\bar{x}_i \left[\prod_{j\neq i}^{k} \int_{\bar{x}_i}^{\bar{x}_i+Q's} p(\bar{x}_j;\xi_j)d\bar{x}_j \right],$$

where $p(\bar{x};\xi) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2}\frac{(\bar{x}-\xi)^2}{\sigma^2}}$,

and $\underline{\xi}' = (\xi_1, \ldots, \xi_k)$.

It is easily checked that the right side of (3.3.1) is equal to

$$(3.3.3) \quad \sum_{i=1}^{k} \int_0^{\infty} p(s)ds \int_{-\infty}^{\infty} p(y_i)dy_i \left[\prod_{j\neq i}^{k} \int_{y_i}^{y_i+Q''s} p(y_j)dy_j \right],$$

where

$$Q'' = Q\sqrt{2} \; ,$$

and $p(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$, and $p(s) = \text{Const.}\ s^{m-1} e^{-\frac{ms^2}{2}}$.

Similarly the right side of (3.3.2) is equal to

$$(3.3.4) \quad \sum_{i=1}^{k} \int_{0}^{\infty} p(s)ds \int_{-\infty}^{\infty} p(y_i;\xi_i)dy_i \left[\prod_{j\neq i}^{k} \int_{y_i}^{y_i+Q''s} p(y_j;\xi_j)dy_j \right],$$

where

$$p(y;\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\xi)^2} \; .$$

3.4 Canonical reduction of the power function. We shall presently show that β given in (3.3.2) could involve as parameters only the $k - 1$ differences $\eta_{i-1} = \xi_1 - \xi_i$ $(i = 2, \ldots, k)$.

Form (3.3.4) we see that

$$(3.4.1)\quad \beta = \sum_{i=1}^{k} I(0, \infty; m; -\infty, \infty; \xi_i; y_1, y_1 + Q''s; \xi_1; \ldots$$

$$y_1, y_1 + Q''s; \xi_{i-1}; y_1, y_1 + Q''s; \xi_{i+1}; \ldots; y_1, y_1 + Q''s; \xi_k),$$

where

$$I(0,\infty; m; -\infty,\infty; \xi_1; y_1, y_1 + Q''s; \xi_2; \ldots; y_1, y_1 + Q''s; \xi_k)$$

$$= \int_0^\infty p(s)ds \int_{-\infty}^{\infty} p(y_1;\xi_1)dy_1 \int_{y_1}^{y_1+Q''s} p(y_2; \xi_2)\, dy_2 \ldots$$

$$\ldots \int_{y_1}^{y_1+Q''s} p(y_k; \xi_k)dy_k.$$

Now putting $y_i - \xi_i = z_i$ and $\xi_1 - \xi_i = \eta_{i-1}$, after a little simplification, we get

$$(3.4.2)\quad \beta = \sum_{i=1}^{k-1} I(0,\infty\,;\, m;\, -\infty\,,\infty\,;\, 0;\, z_1 - \eta_i,\, z_1 - \eta_i + Q''s;$$

$$0;\, z_1 - \eta_i + \eta_1,\, z_1 - \eta_i + \eta_1 + Q''s;\, 0;\, \ldots;\, z_1 - \eta_i + \eta_{k-1},$$

$$z_1 - \eta_i + \eta_{k-1} + Q''s;\, 0)$$

$$+\, I(0,\infty\,;\, m;\, -\infty\,,\infty\,;\, 0;\, z_1 + \eta_1,\, z_1 + \eta_1 + Q''s;\, 0;\, z_1 + \eta_2,$$

$$z_1 + \eta_2 + Q''s;\, 0;\, \ldots;\, z_1 + \eta_{k-1},\, z_1 + \eta_{k-1} + Q''s;\, 0).$$

From (3.4.2) it is evident that β could involve as parameters only the k-1 η's.

Hence the power (= 1 - β) of the q test could involve as parameters only the k - 1 η's. It is worth noting at this point that the right side of (3.4.2) is symmetric in the η's. Hence the power of the q test is also symmetric in the η's.

3.5 Uniform unbiased nature of the q test. To prove the uniform unbiased nature of the q test we need to use certain lemmas which we shall prove now.

Lemma 1. If

(1) in the domain D: $\left\{ \underline{x}\text{: } a_i \leq x_i \leq b_i,\ i = 1, \ldots, k \right\}$,
$f(x_1, \ldots, x_k)$ exists,

all partial derivatives of order one and two exist, all partial derivatives of order one vanish simultaneously at one and only one inner point $P = (x_{10}, x_{20}, \ldots, x_{k0})$ of D,

(2) the matrix of second partials evaluated at P is negative definite (n.d.), and

(3) at every point $(x_1, \ldots, x_k)$ on the boundary of D,
$f(x_1, \ldots, x_k) < f(x_{10}, \ldots, x_{k0})$, that is, $< A$, (say), then

(3.5.1) $f(x_1, \ldots, x_k) < f(x_{10}, \ldots, x_{k0})$ that is, $< A$ for all $x \epsilon D$.

Proof: Condition (1) implies there is one and only one stationary point P inside the domain D. By condition (2) we see that at this point there is actually a local maximum. Hence there cannot be another point inside D where $f(x_1, \ldots, x_k) > A$, for otherwise there will be a contradiction. But there can be points on the boundary of D where $f(x_1, \ldots, x_k) > A$. But by (3) this is impossible.

Hence $f(x_1, \ldots, x_k) < A$ for all $x \epsilon D$.

Lemma 2. If the conditions of lemma 1 are satisfied, as $a_i \longrightarrow -\infty$ or $b_i \longrightarrow \infty$, for any i and for fixed values of a_j, b_j ($j \neq i = 1, \ldots, k$),

$$f(x_1, \ldots, x_k) < A \text{ for all } x \epsilon D' : \left\{ \underline{x} : -\infty < x's < \infty \right\}$$

Proof: The proof follows obviously from lemma 1.

Theorem 1. The Studentized range test of Tukey is completely unbiased, but its power function does not have the monotonicity property.

Proof: The second kind of error of the q test for a given significance level α is

$$(3.5.2) \qquad \beta = P\left[\frac{\bar{x}_{max} - \bar{x}_{min}}{s} \le Q \,/\, \xi \right]$$

$$= \sum_{i=1}^{k-1} I(0,\infty\,; m; -\infty\,,\infty\,; 0; z_1 - \eta_i; z_1 - \eta_i + Qs; 0;$$

$$z_1 - \eta_i + \eta_1, z_1 - \eta_i + \eta_1 + Qs; 0; \ldots; z_1 - \eta_i + \eta_{i-1},$$

$$z_1 - \eta_i + \eta_{i-1} + Qs; 0; z_1 - \eta_i + \eta_{i+1}, z_1 - \eta_i + \eta_{i+1} + Qs;$$

$$0; \ldots; z_1 - \eta_i + \eta_{k-1}, z_1 - \eta_i + \eta_{k-1} + Qs; 0)$$

$$+ I(0,\infty\,; m; -\infty\,,\infty\,; 0; z_1 + \eta_1, z_1 + \eta_1 + Qs; 0; \ldots$$

$$\ldots; z_1 + \eta_{k-1}, z_1 + \eta_{k-1} + Qs; 0)\,,$$

where $Q > 0$ is so chosen that

$$k\, I(0,\infty\,; m; -\infty\,,\infty\,; 0; z_1, z_1 + Qs; 0; z_1, z_1 + Qs; 0; \ldots$$

$$\ldots; z_1, z_1 + Qs; 0) = 1 - \alpha.$$

Now differentiating β with respect to η_1 we get, after some simplification,

$$(3.5.3) \quad (2\pi)^{\frac{k}{2}} \frac{\partial\beta}{\partial\eta_1} = \int_0^\infty p(s)ds \Big[\int_{-\infty}^\infty e^{-\frac{1}{2}z_1^2-\frac{1}{2}(z_1+\eta_1+Qs)^2} dz_1$$

$$- \int_{-\infty}^\infty e^{-\frac{1}{2}(z_1+\eta_1)^2-\frac{1}{2}(z_1+Qs)^2} dz_1 \Big]\Big[\int_{z_1+\eta_2}^{z_1+\eta_2+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \ \ldots$$

$$\ldots \int_{z_1+\eta_{k-1}}^{z_1+\eta_{k-1}+Qs} e^{-\frac{1}{2}z_{k-1}^2} dz_{k-1} \Big]$$

$$+ \int_0^\infty p(s)ds \Big[\int_{-\infty}^\infty e^{-\frac{1}{2}(z_1+\eta_2)^2 -\frac{1}{2}(z_1+\eta_1+Qs)^2} dz_1$$

$$- \int_{-\infty}^\infty e^{-\frac{1}{2}(z_1+\eta_1)^2-\frac{1}{2}(z_1+\eta_2+Qs)^2} dz_1 \Big]\Big[\int_{z_1}^{z_1+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \ \ldots$$

$$\ldots \int_{z_1+\eta_{k-1}}^{z_1+\eta_{k-1}+Qs} e^{-\frac{1}{2} z_{k-1}^2} dz_{k-1} \Big]$$

$$+ \ldots$$

$$+ \int_0^\infty p(s)ds \Big[\int_{-\infty}^{\infty} e^{-\frac{1}{2}(z_1+\eta_{k-1})^2 - \frac{1}{2}(z_1+\eta_1+Qs)^2} dz_1$$

$$- \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z_1+\eta_1)^2 - \frac{1}{2}(z_1+\eta_{k-1}+Qs)^2} dz_1 \Big]\Big[\int_{z_1}^{z_1+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \ldots$$

$$\ldots \int_{z_1+\eta_{k-2}}^{z_1+\eta_{k-2}+Qs} e^{-\frac{1}{2}z_{k-1}^2} dz_{k-1} \Big].$$

It is easy to check that the right side of (3.5.3) will be negative if $\eta_1 > 0$ and $\eta_1 > \eta_i$ $(i = 2, \ldots, k-1)$, and positive if $\eta_1 < 0$ and $\eta_1 < \eta_i$ $(i = 2, \ldots, k - 1)$. By the symmetry in the variables the same is true of $\partial\beta/\partial\eta_i$ $(i = 2, \ldots, k - 1)$, i.e., (3.5.4) $\quad \partial\beta/\partial\eta_i \quad < 0 \quad$ if $\eta_i > 0$ and $\eta_i = \eta_{max}$

and $\quad > 0 \quad$ if $\eta_i \not\leq 0$ and $\eta_i = \eta_{min}$.

Also it is evident that

(3.5.5) $$\frac{\partial\beta}{\partial\eta_1}\Big]_{\underline{\eta} = \underline{0}} = 0.$$

Similarly $\frac{\partial\beta}{\partial\eta_i}\Big]_{\underline{\eta} = \underline{0}} = 0$, $(i = 2, \ldots, k-1)$.

Now suppose $\underline{\eta}^0 \neq \underline{0}$. Then either $\eta^0_{max} > 0$ or $\eta^0_{min} < 0$. Hence the first partials can vanish simultaneously only at (0, 0, ..., 0).

Again it is easily verified that

(3.5.6) $$\frac{\partial^2\beta}{\partial\eta_i^2}\Big]_{\underline{\eta} = \underline{0}} = -2(k-1)Q\,C\,(Q),$$

where

$$C(Q) = \frac{(2\pi)^{-\frac{k}{2}}}{2} \int_0^\infty s\,p(s)ds \int_{-\infty}^{\infty} e^{-\frac{1}{2}z_1^2\, -\frac{1}{2}(z_1+Qs)^2} \Big[\int_{z_1}^{z_1+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \Big]^{k-2}$$

$> 0.$

Hence $\frac{\partial^2\beta}{\partial\eta_i^2}\Big]_{\underline{\eta}=\underline{0}} < 0$, $(i = 1, \ldots, k-1)$.

Also

(3.5.7) $$\frac{\partial^2\beta}{\partial\eta_i\partial\eta_j}\Big]_{\underline{\eta}=\underline{0}} = 2Q\ c\ (Q) > 0\ ,\ (i \neq j = 1, \ldots, k-1).$$

Hence the matrix of second partials when $\underline{\eta} = \underline{0}$

(3.5.8) $$M = \left\| \frac{\partial^2\beta}{\partial\eta_i\partial\eta_j}\Big]_{\underline{\eta}=\underline{0}} \right\|$$

$$= \left\| \begin{matrix} -2(k-1)f(Q) & 2f(Q) & \cdots & 2f(Q) \\ 2f(Q) & -2(k-1)f(Q) & \cdots & 2f(Q) \\ \cdot & \cdot & \cdots & \cdot \\ 2f(Q) & 2f(Q) & \cdots & -2(k-1)f(Q) \end{matrix} \right\| ,$$

where $f(Q) = Q\ C\ (Q)$, is n.d.

It will now suffice to show that $\beta \longrightarrow 0$ on each point of the boundary of the domain D: $\left\{ \underline{\eta}: \epsilon_i \leq \eta_i \leq \lambda_i;\ i=1,\ldots,k-1 \right\}$ as, say, $\epsilon_1 \longrightarrow -\infty$ or $\lambda_1 \longrightarrow \infty$ for fixed values of ϵ_i, λ_i $(i=2,\ldots,k-1)$.

Now

$$(3.5.9) \quad (2\pi)^{\frac{k}{2}}\beta = \int_0^{\infty} p(s)ds \int_{-\infty}^{\infty} e^{-\frac{1}{2}z_1^2} dz_1 \int_{z_1+\eta_1}^{z_1+\eta_1+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \cdots$$

$$\cdots \int_{z_1+\eta_{k-1}}^{z_1+\eta_{k-1}+Qs} e^{-\frac{1}{2}z_k^2} dz_k$$

$$+ \int_0^{\infty} p(s)ds \int_{-\infty}^{\infty} e^{-\frac{1}{2}z_1^2} dz_1 \int_{z_1-\eta_1}^{z_1-\eta_1+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \cdots$$

$$\cdots \int_{z_1-\eta_1+\eta_{k-1}}^{z_1-\eta_1+\eta_{k-1}+Qs} e^{-\frac{1}{2}z_k^2} dz_k$$

$$+ \cdots$$

$$+ \int_{0}^{\infty} p(s)ds \int_{-\infty}^{\infty} e^{-\frac{1}{2}z_1^2} dz_1 \int_{z_1-\eta_{k-1}}^{z_1-\eta_{k-1}+Qs} e^{-\frac{1}{2}z_2^2} dz_2 \cdots$$

$$\cdots \int_{z_1-\eta_{k-1}+\eta_{k-2}}^{z_1-\eta_{k-1}+\eta_{k-2}+Qs} e^{-\frac{1}{2}z_k^2} dz_k \quad .$$

Now consider the case where $\eta_1 = \epsilon_1$ and $\eta_2, \ldots, \eta_{k-1}$ are in the domain D': $\left\{ \epsilon_i \leq \eta_i \leq \lambda_i \text{, } i=2,\ldots,k-1 \right\}$.

(3.5.10) $$\underset{\epsilon_1 \to -\infty}{\text{Lt.}} \int_{z_1+\epsilon_1}^{z_1+\epsilon_1+Qs} e^{-\frac{z^2}{2}} dz = 0, \text{and} \quad \underset{\epsilon_1 \to -\infty}{\text{Lt.}} \int_{z_1-\epsilon_1}^{z_1-\epsilon_1+Qs} e^{-\frac{1}{2}z^2} dz = 0.$$

Similarly when $\eta_1 = \lambda_1$ and $\eta_2, \ldots, \eta_{k-1}$ are in D', we have,

$$(3.5.11)\quad \mathop{Lt}_{\lambda_1 \to \infty} \int_{z_1+\lambda_1}^{z_1+\lambda_1+Qs} e^{-\frac{1}{2}z^2}\, dz = 0 \text{ and } \mathop{Lt}_{\lambda_1 \to \infty} \int_{z_1-\lambda_1}^{z_1-\lambda_1+Qs} e^{-\frac{1}{2}z^2}\, dz = 0.$$

Also $\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx$ exists finitely.

Hence as $\epsilon_1 \to -\infty$ or $\lambda_1 \to \infty$, the value of β at each point on the boundary $\to 0$ while the value of β at the point where η's $= 0$ is $1 - \alpha > 0$. Hence

$$(3.5.12)\quad 1 - \alpha = \beta(\underline{0}) > \beta(D \text{ as } \epsilon \to -\infty \text{ or } \lambda \to \infty) = 0 .$$

Hence all the conditions given in lemma 2 are satisfied by the function $\beta(\underline{\eta})$.

Hence

$$(3.5.13)\quad \beta(\underline{0}) > \beta(\underline{\eta}) \quad \text{for every } \underline{\eta} \neq \underline{0} .$$

Hence the Tukey q test is completely unbiased.

Now by equation (3.5.4) we have

(3.5.14) $\frac{\partial\beta}{\partial\eta_i} < 0 \quad$ if $\eta_i > 0$ and $\eta_i = \eta_{max}$

$> 0 \quad$ if $\eta_i < 0$ and $\eta_i = \eta_{min}$, $(i = 1,2,\ldots,k-1)$.

It can be shown that $\partial\beta/\partial\eta_i > 0$ or < 0 according as (i) $0 < \eta_i << \eta_j$ or (ii) $\eta_j << \eta_i < 0$ (for all $j \neq i = 1, 2, \cdots k-1$).

Hence the power function of the q test does not have the monotonicity property.

3.6 Lower bound to the power of the q test. The second kind of error of the q test is

$$(3.6.1) \qquad \beta(\underline{\eta}) = P\left[\frac{\bar{x}_{max}-\bar{x}_{min}}{s} \leq Q \Big/ \xi\right]$$

$$\Longleftrightarrow P\left[\frac{\bar{x}_i-\bar{x}_j}{s} \leq Q \Big/ \xi, \text{all } i \neq j = 1,2,\ldots,k\right]$$

$$\leq P_i\left[\left|\frac{\bar{x}_i-\bar{x}_j}{s}\right| \leq Q \Big/ \xi, \; j=1,2,\ldots,i-1,i+1,\ldots,k\right].$$

Hence

(3.6.2)

$$\text{Power of q test} \geq 1-P_i\left[\left|\frac{\bar{x}_i-\bar{x}_j}{s}\right| \leq Q\Big/\xi, j=1,2,\ldots,i-1,i+1,\ldots,k\right].$$

There are k such lower bounds and the power of the q test will be greater than the g.l.b. (greatest lower bound).

The lower bound given in (3.6.2) can be evaluated by using the distribution of a multivariate analogue of Students' t considered by Dunnett and Sobel [6]. Entensive tables of the distribution of the bivariate analogue of Student's t are given in [6].

It is to be noticed that the evaluation of the exact value of the power function of the q test is extremely difficult. The evaluation of the lower bounds given in (3.6.2) is difficult, but the tabulation of the expression on the right side of (3.6.2) is useful not only in this situation but also in other situations including the important situation of ranking means of normal populations.

3.7 The Hartley F_{max} ratio test. Let

x_{ij} $(i = 1, \ldots, k;\ j = 1, \ldots, n + 1)$ be independent samples of size $(n + 1)$ from $N(\xi_i, \sigma_i^2), (i = 1, \ldots, k)$.

Let $s_i^2 = \sum_{j=1}^{n+1} (x_{ij} - \bar{x}_i)^2/n$, where $\bar{x}_i = \sum_{j=1}^{n+1} x_{ij}/(n+1)$, be the unbiased estimate of σ_i^2 based on n.d.f. To test the hypothesis $H_0: \sigma_1^2 = \ldots = \sigma_k^2$, we can obtain a test procedure as follows: To test the hypothesis H_{ij}: $\sigma_i^2 = \sigma_j^2$ we have the well known variance ratio F test of Fisher given by $F_{ij} = (s_i^2/s_j^2)$ with d.f. (n,n). The hypothesis $\sigma_i^2 = \sigma_j^2$ is accepted if $1/F_\alpha \leq (s_i^2/s_j^2) \leq F_\alpha$, where F_α is the upper α point of Fisher's F with d.f. (n,n). Since H_0 is equivalent to H_{ij} (all $i \neq j$) we get a test of H_0 as follows: Take the intersection of all the $\binom{k}{2}$ Fisher's F_{ij} acceptance regions and accept H_0 if

$$\text{largest } F_{ij} \leq F_\alpha \ .$$

It is easy to check that this is the same as accepting H_0 if $F_{max} = (s^2_{max} \ / \ s^2_{min}) \leq F_\alpha$, where F_α is the upper α point of the F_{max} distribution with d.f. (n,n). This is the Hartley F_{max} ratio test.

We shall show that the power function of the F_{max} test could involve as parameters only the k-1 ratios $\eta_{i-1} = (\sigma_1^2/\sigma_i^2)$, (i=2,.. ,k). Also we shall show that the F_{max} test is completely unbiased, but the power function does not have the monotonicity property. A set of useful lower bounds is obtained on the power of the F_{max} test, which can be evaluated using tables of the distribution of the Studentized largest chi-square. The distribution of the Studentized largest chi-square is given in Chapter IV. A multivariate generalization of the F_{max} test is given in (3.12).

3.8 Power function of the F_{max} test. Let

x_{ij} (i=1, ..., k; j=1, ..., n+1) be k (n+1) independent $N(\xi_i, \sigma_i^2;\ i=1, \ldots, k)$ variables. It is well known that $s_i^2/\sigma_i^2 = \sum_{j=1}^{n+1} (x_{ij} - \bar{x}_i)^2/n\sigma_i^2$ where $\bar{x}_i = \sum_{j=1}^{n+1} x_{ij}/(n+1)$, has a chi-square distribution with n.d.f.

When the hypothesis $H_0: \sigma_1^2 = \ldots = \sigma_k^2$ is true, then

$$(3.8.1)\quad 1 - \alpha = P\left[(s_i^2/s_j^2) \leq F, \text{for all } i \neq j\right] \Longleftrightarrow P\left[\frac{s_{max}^2}{s_{min}^2} \leq F\right]$$

$$= \sum_{i=1}^{k} \int_0^{\infty} p(u_i^2) \left[\prod_{j \neq i}^{k} \int_{u_i^2}^{Fu_i^2} p(u_j^2)\, du_j^2\right] du_i^2 \quad ,$$

where $p(u^2) = \frac{1}{2\Gamma(\frac{n}{2})} \left(\frac{u^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{u^2}{2}}$

and α is the given level of significance of the test.

If H_0 is not true, then the second kind of error of the test is

$$(3.8.2)\quad \beta = P\left[s_i^2/s_j^2 \leq F, \text{ all } i \neq j \;\middle/\; \underline{\sigma}^2\right] \Longleftrightarrow P\left[\frac{s_{max}^2}{s_{min}^2} \leq F \;\middle/\; \underline{\sigma}^2\right]$$

$$= \sum_{i=1}^{k} \int_{0}^{\infty} p(v_i^2;\ \sigma_i^2) \left[\prod_{j \neq i}^{k} \int_{v_i^2}^{Fv_i^2} p(v_j^2;\ \sigma_j^2)\, dv_j^2 \right] dv_i^2 ,$$

where $p(v_i^2;\ \sigma_i^2) = \frac{1}{2\Gamma(\frac{n}{2})\sigma_i^2} \left(\frac{v_i^2}{2\sigma_i^2}\right)^{\frac{n-2}{2}} e^{-\frac{v_i^2}{2\sigma_i^2}}$

and $\underline{\sigma}^{2'} = (\sigma_1^2, \ldots, \sigma_k^2)$.

3.9 <u>Canonical reduction of the power function</u>. We shall presently show that β given in (3.8.2) could involve as parameters only the k-1 ratios $\eta_{i-1} = \sigma_1^2/\sigma_i^2$ $(i = 2, \ldots, k)$.

From (3.8.2) we see that

(3.9.1) $\beta = \sum_{i=1}^{k} I(0, \infty,\ \sigma_i^2;\ s_1^2,\ Fs_1^2,\ \sigma_1^2;\ s_1^2,\ Fs_1^2,\ \sigma_2^2;\ \ldots$

$s_1^2,\ Fs_1^2,\ \sigma_{i-1}^2;\ s_1^2,\ Fs_1^2,\ \sigma_{i+1}^2;\ \ldots;\ s_1^2,\ Fs_1^2,\ \sigma_k^2)$,

where

$I(0, \infty,\ \sigma_1^2;\ s_1^2,\ Fs_1^2,\ \sigma_2^2;\ \ldots;\ s_1^2,\ Fs_1^2,\ \sigma_k^2)$

$$= \int_{0}^{\infty} p(s_1^2;\sigma_1^2)ds_1^2 \int_{s_1^2}^{Fs_1^2} p(s_2^2;\sigma_2^2)ds_2^2 \cdots \int_{s_1^2}^{Fs_1^2} p(s_k^2;\sigma_k^2)ds_k^2.$$

Now putting $s_i^2/\sigma_i^2 = v_i^2$ and $\eta_{i-1} = \sigma_1^2/\sigma_i^2$ we get, after a little simplification,

$$(3.9.2)\quad \beta = \sum_{i=1}^{k-1} I(0,\infty,1;\ \frac{v_1^2}{\eta_i},\ \frac{Fv_1^2}{\eta_i},\ 1;\ \frac{v_1^2\eta_1}{\eta_i},\ \frac{Fv_1^2\eta_1}{\eta_i},\ 1;\ldots$$

$$\ldots;\ \frac{v_1^2}{\eta_i}\eta_{k-1},\ F\,\frac{v_1^2}{\eta_i}\eta_{k-1},\ 1)$$

$$+\ I(0,\infty,1;v_1^2\eta_1,Fv_1^2\eta_1,1;\ v_1^2\eta_2,Fv_1^2\eta_2,1;\ldots;\ v_1^2\eta_{k-1},Fv_1^2\eta_{k-1},1).$$

From (3.9.2) it is evident that β could involve as parameters only the k-1 η's. Hence the power (= 1 - β) of the F_{max} test could involve as parameters only the k-1 η's. It is worth noting at this point that the right side of (3.9.2) is symmetric in the η's. Hence the power of the F_{max} test is also symmetric in the η's.

3.10 Uniform unbiased nature of the F_{max} test. To prove the uniform unbiased nature of the F_{max} test we need to use a lemma which is

Lemma 3. If the conditions of lemma 1 are satisfied, as $a_i \longrightarrow 0$ or $b_i \longrightarrow \infty$, for any i and for fixed values of a_j, $b_j (j \neq i = 1, \ldots, k)$, $f(x_1, \ldots, x_k) < A$ for all $x \in D'$: $\{\underline{x}: 0 < x's < \infty\}$.

Proof: The proof follows obviously from lemma 1.

Theorem 2. The F_{max} ratio test of Hartley is completely unbiased, but its power function does not have the monotonicity property.

Proof: The second kind of error of the F_{max} test for a given significance level α is

$$(3.10.1)\quad \beta = P\left[\frac{s^2_{max}}{s^2_{min}} \leq F \,\middle/\, \underline{\sigma}^2\right]$$

$$= \sum_{i=1}^{k-1} I\left(0,\infty,1;\ \frac{v_1^2}{\eta_i}, F\frac{v_1^2}{\eta_i}, 1;\ \frac{v_1^2\eta_1}{\eta_i}, F\frac{v_1^2\eta_1}{\eta_i}, 1;\ \ldots;\ v_1^2\frac{\eta_{i-1}}{\eta_i}, Fv_1^2\frac{\eta_{i-1}}{\eta_i}, 1;\right.$$

$$\left. v_1^2\frac{\eta_{i+1}}{\eta_i}, Fv_1^2\frac{\eta_{i+1}}{\eta_i}, 1;\ \ldots;\ v_1^2\frac{\eta_{k-1}}{\eta_i}, Fv_1^2\frac{\eta_{k-1}}{\eta_i}, 1\right)$$

$$+\ I(0,\infty,1;\ v_1^2\eta_1, Fv_1^2\eta_1, 1; \ldots; v_1^2\eta_{k-1}, Fv_1^2\eta_{k-1}, 1),$$

where $F > 1$ is so chosen that

$$(3.10.2)\quad k\, I(0,\infty,1;\ v_1^2, Fv_1^2, 1; \ldots; v_1^2, Fv_1^2, 1) = 1 - \alpha.$$

Now differentiating β with respect to η_1 we get, after some simplification,

$$(3.10.3) \quad \Gamma^k\left(\frac{n}{2}\right) \frac{\partial\beta}{\partial\eta_1} = \Big[\int_0^\infty \left(\frac{u_1^2}{2}\right)^{n-1} e^{-\frac{u_1^2}{2}(1+F\eta_1)} \, d\left(\frac{u_1^2}{2}\right)$$

$$- \int_0^\infty \left(\frac{u_1^2}{2}\right)^{n-1} e^{-\frac{u_1^2}{2}(F+\eta_1)} \, d\left(\frac{u_1^2}{2}\right) \Big] \Big[\int_{u_1^2\eta_2}^{Fu_1^2\eta_2} \left(\frac{u_2^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{u_2^2}{2}} \, d\left(\frac{u_2^2}{2}\right)$$

$$\cdots \int_{u_1^2\eta_{k-1}}^{Fu_1^2\eta_{k-1}} \left(\frac{u_{k-1}^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{u_{k-1}^2}{2}} \, d\left(\frac{u_{k-1}^2}{2}\right) \Big]$$

$$+ \Big[\int_0^\infty \left(\frac{u_1^2}{2}\right)^{n-1} e^{-\frac{u_1^2}{2}(\eta_2+F\eta_1)} \, d\left(\frac{u_1^2}{2}\right)$$

$$- \int_0^\infty \left(\frac{u_1^2}{2}\right)^{n-1} e^{-\frac{u_1^2}{2}(\eta_1+F\eta_2)} \, d\left(\frac{u_1^2}{2}\right) \Big]$$

$$\cdot\left[\int_{u_1^2}^{Fu_1^2} \left(\frac{u_2^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_2^2}{2}} d\left(\frac{u_2^2}{2}\right) \cdots \int_{u_1^2\eta_{k-1}}^{Fu_1^2\eta_{k-1}} \left(\frac{u_{k-1}^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_{k-1}^2}{2}} d\left(\frac{u_{k-1}^2}{2}\right) \right]$$

$+ \cdots$

$$+ \left[\int_0^{\infty} \left(\frac{u_1^2}{2}\right)^{n-1} e^{-\frac{u_1^2}{2}(\eta_{k-1}+F\eta_1)} d\left(\frac{u_1^2}{2}\right) - \int_0^{\infty} \left(\frac{u_1^2}{2}\right)^{n-1} e^{\frac{-u_1^2}{2}(\eta_1+F\eta_{k-1})} d\left(\frac{u_1^2}{2}\right) \right]$$

$$\cdot\left[\int_{u_1^2}^{Fu_1^2} \left(\frac{u_2^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_2^2}{2}} d\left(\frac{u_2^2}{2}\right) \cdots \int_{u_1^2\eta_{k-2}}^{Fu_1^2\eta_{k-2}} \left(\frac{u_{k-1}^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_{k-1}^2}{2}} d\left(\frac{u_{k-1}^2}{2}\right) \right].$$

It is easy to check that the right side of (3.10.3) will be negative if $\eta_1 > 1$ and $\eta_1 > \eta_i$ $(i=2,\ldots,k-1)$ and positive if $\eta_1 < 1$ and $\eta_1 < \eta_i$ $(i = 2, \ldots, k-1)$. By the symmetry in the variables, the same is true of $\frac{\partial\beta}{\partial\eta_i}$ $(i=1,\ldots,k-1)$, i.e.,

(3.10.4) $\frac{\partial\beta}{\partial\eta_i} < 0$ if $\eta_i > 1$ and $\eta_i = \eta_{max}$

and > 0 if $\eta_i < 1$ and $\eta_i = \eta_{min}$.

Also it is evident that

$$(3.10.5) \qquad \frac{\partial\beta}{\partial\eta_1}\Big]_{\underline{\eta} = \underline{1}} = 0 \; .$$

Similarly $\quad \frac{\partial\beta}{\partial\eta_i}\Big]_{\underline{\eta} = \underline{1}} = 0 \qquad (i=2,\ldots,k-1).$

Now suppose $\underline{\eta}^0 \neq \underline{1}$. Then either $\eta^0_{max} > 1$ or $\eta^0_{min} < 1$. Hence the first partials can vanish simultaneously only at $(1,\ldots,1)$.

Again it is easily verified that

$$(3.10.6) \qquad \frac{\partial^2\beta}{\partial\eta_i^2}\Big]_{\underline{\eta} = \underline{1}} = (k-1)\,(1-F)\;C\;(F),$$

where

$$C(F) = \frac{1}{\Gamma^k(\frac{n}{2})} \int_0^\infty (\tfrac{u^2}{2})^n \; e^{\frac{-u^2}{2}} \, d(\tfrac{u^2}{2}) \Big[\int_{u^2}^{Fu^2} (\tfrac{v^2}{2})^{\frac{n-2}{2}} e^{-\frac{v^2}{2}} \, d(\tfrac{v^2}{2}) \Big]^{k-2}$$

$> 0.$

Hence $\frac{\partial^2\beta}{\partial\eta_i^2}\Big]_{\underline{\eta}=\underline{1}} < 0 , \qquad (i=1,\ldots,k-1).$

Also $\frac{\partial^2\beta}{\partial\eta_i\partial\eta_j}\Big]_{\underline{\eta} = \underline{1}} = (F-1)C(F) \quad > 0 , \quad (i\neq j=1,\ldots,k-1).$

Hence the matrix of second partials when $\underline{\eta} = \underline{1}$

$$M = \left\| \frac{\partial^2 \beta}{\partial \eta_i \partial \eta_j} \right]_{\underline{\eta}=\underline{1}} \right\|$$

$$= \left\| \begin{matrix} -(k-1)g(F) & g(F) & \cdots & g(F) \\ g(F) & -(k-1)g(F) & \cdots & g(F) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ g(F) & g(F) & \cdots & -(k-1)g(F) \end{matrix} \right\| ,$$

where $g(F) = (F-1)\, C\, (F)$, is n.d.

It will now suffice to show that $\beta \longrightarrow 0$ on each point of the boundary of the domain D: $\left\{ \underline{\eta}:\ \epsilon_i \leq \eta_i \leq \lambda_i;\ i=1,\ \ldots,\ k-1 \right\}$ as, say, $\epsilon_1 \longrightarrow 0$ or $\lambda_1 \longrightarrow \infty$ and for fixed values of ϵ_i and λ_i $(i=2,\ \ldots,\ k-1)$.

Now

$$(3.10.7)\quad \Gamma^k\left(\frac{n}{2}\right) \beta = \int_0^\infty \left(\frac{u_1^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_1^2}{2}} d\left(\frac{u_1^2}{2}\right) \int_{u_1^2\eta_1}^{Fu_1^2\eta_1} \left(\frac{u_2^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_2^2}{2}} d\left(\frac{u_2^2}{2}\right) \cdots$$

$$\int_{u_1^2\eta_{k-1}}^{Fu_1^2\eta_{k-1}} \left(\frac{u_k^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}u_k^2} d\left(\frac{u_k^2}{2}\right)$$

$$+ \ldots$$

$$+ \int_{0}^{\infty} \left(\frac{u_1^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{-u_1^2}{2}} d\left(\frac{u_1^2}{2}\right) \int_{\frac{u_1^2}{\eta_{k-1}}}^{\frac{Fu_1^2}{\eta_{k-1}}} \left(\frac{u_2^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}u_2^2} d\left(\frac{u_2^2}{2}\right)$$

$$\ldots \int_{u_1^2 \frac{\eta_{k-2}}{\eta_{k-1}}}^{\frac{Fu_1^2}{\eta_{k-1}}\eta_{k-2}} \left(\frac{u_k^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}u_k^2} d\left(\frac{u_k^2}{2}\right).$$

Now consider the case where $\eta_1 = \epsilon_1$ and $\eta_2, \ldots, \eta_{k-1}$ are in the domain D': $\left\{ \epsilon_i \leq \eta_i \leq \lambda_i ;\ i=2, \ldots, k-1 \right\}$.

$$(3.10.8) \qquad \underset{\epsilon_1 \to 0}{\mathrm{Lt.}} \int_{u_1^2 \epsilon_1}^{Fu_1^2 \epsilon_1} \left(\frac{u^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{u^2}{2}} d\left(\frac{u^2}{2}\right) = 0 ,$$

and

$$\underset{\epsilon_1 \to 0}{\mathrm{Lt.}} \int_{u_1^2/\epsilon_1}^{Fu_1^2/\epsilon_1} \left(\frac{u^2}{2}\right)^{\frac{n-2}{2}} e^{\frac{u^2}{2}} d\left(\frac{u^2}{2}\right) = 0.$$

Similarly when $\eta_1 = \lambda_1$ and $\eta_2, \ldots, \eta_{k-1}$ are in D', we have,

(3.10.9) $$\underset{\lambda_1 \to \infty}{\text{Lt.}} \int_{u_1^2\lambda_1}^{Fu_1^2\lambda_1} \left(\frac{u^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{u^2}{2}} d\left(\frac{u^2}{2}\right) = 0$$

and $$\underset{\lambda_1 \to \infty}{\text{Lt.}} \int_{u_1^2/\lambda_1}^{Fu_1^2/\lambda_1} \left(\frac{u^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{u^2}{2}} d\left(\frac{u^2}{2}\right) = 0.$$

Also $\int_0^{\infty} \left(\frac{u^2}{2}\right)^{\frac{n-2}{2}} e^{-\frac{u^2}{2}} d\left(\frac{u^2}{2}\right)$ exists finitely.

Hence as $\epsilon_1 \to 0$ or $\lambda_1 \to \infty$, the value of β at each point on the boundary $\to 0$ while the value of β at the point where all η's = 1 is $1 - \alpha > 0$. Hence

(3.10.10) $$1 - \alpha = \beta(\underline{1}) > \beta\,(D \text{ as } \epsilon \to 0 \text{ or } \lambda \to \infty) = 0.$$

Hence all the conditions given in lemma 3 are satisfied by the function $\beta(\underline{\eta})$.

Hence $\beta(\underline{1}) > \beta(\underline{\eta})$ for every $\underline{\eta} \neq \underline{1}$.

Hence the Hartley F_{max} ratio test is completely unbiased.

Now by equation (3.10.4), we have,

$$(3.10.11) \qquad \frac{\partial\beta}{\partial\eta_i} \quad < 0 \qquad \text{if } \eta_i > 1 \text{ and } \eta_i = \eta_{max}$$

$$\text{and} \qquad\qquad > 0 \qquad \text{if } \eta_i < 1 \text{ and } \eta_i = \eta_{min} \quad (i=1,\ldots,k-1).$$

It can be shown that $\partial\beta/\partial\eta_i > 0$ or < 0 according as (i) $1 < \eta_i << \eta_j$ or (ii) $\eta_j << \eta_i < 1$ (for all $j \neq i = 1, 2, \cdots k-1$).

Hence the power function of the F_{max} test does not have the monotonicity property.

3.11 Lower bound to the power of the F_{max} test. The second kind of error of the F_{max} test is

$$(3.11.1) \qquad \beta(\underline{\eta}) = P\left[\frac{s^2_{max}}{s^2_{min}} \leq F \,\Big/\, \underline{\sigma}^2\right]$$

$$\Longleftrightarrow P\left[\frac{s^2_i}{s^2_j} \leq F \,\Big/\, \underline{\sigma}^2;\ \text{all } i \neq j = 1, \ldots, k\right]$$

$$\leq P_i \left[\frac{1}{F} \leq \frac{s^2_i}{s^2_j} \leq F \,\Big/\, \underline{\sigma}^2;\ j=1, \ldots, i-1, i+1, \ldots, k\right].$$

Hence

(3.11.2) Power of F_{max} test

$$\geq 1 - P_i \left[\frac{1}{F} \leq \frac{s_i^2}{s_j^2} \leq F \,\Big/\, \underline{\sigma}^2 ;\ j = 1, \ldots, i-1, i+1, \ldots, k \right].$$

There are k such lower bounds and the power of the F_{max} test will be greater than the g.l.b.

The lower bound given in (3.11.2) can be evaluated by using tables of the distribution of Studentized largest chi-square. The distribution of the Studentized largest chi-square is derived in Chapter IV.

It is to be noted that the evaluation of the exact value of the power function of the F_{max} test is extremely difficult. The evaluation of the lower bound given in (3.11.2) is rather easy.

3.12 Multivariate analogue of Tukey and Hartley tests. Given random samples $X_i(p \times n_i + 1)$ of sizes $(n_i + 1)$ $(i = 1, \ldots, k)$ from k independent $N(\underline{\xi}_i, \Sigma), (i=1,\ldots,k)$; the hypothesis H_0: $\underline{\xi}_1 = \ldots = \underline{\xi}_k$ can be tested by using a multivariate analogue of the q test. This test based on the largest of a set of Hotelling's T^2 is due to Roy and Bose [20].

Let $X_i(p \times n + 1), (i = 1, \ldots, k)$ be k independent samples of sizes $(n + 1)$ from $N(\underline{\xi}_i, \Sigma_i), (i = 1, \ldots, k)$. To test the hypothesis H_0: $\Sigma_1 = \ldots = \Sigma_k$ against the alternative $H \neq H_0$, we can get a multivariate extension of the Hartley test as follows: The test of the hypothesis $H^0_{ij} : \Sigma_i = \Sigma_j$ can be carried out, as given in Chapter II, using the joint distribution of the largest and smallest roots of $S_i(S_i + cS_j)^{-1}$ or equivalently of $S_iS_j^{-1}$ where $S_i = X_iX_i' - (n+1)\,\bar{\underline{x}}_i\bar{\underline{x}}_i'$. Hence the test of H_0 will be based on $\operatorname{Sup}_{i \neq j} c(S_iS_j^{-1})$, where the supremum is to be taken over all $i \neq j = 1, \ldots, k$. The distribution problem connected with this test is quite difficult. Power properties of this test will not be discussed here.

So far in this chapter we considered the F_{max} test when all the s_i^2's are based on the same number of d.f.n. Investigation is proceding on the behaviour of the F_{max} test when the d.f. are unequal. Power properties of similar generalizations of the q test are also being investigated.

CHAPTER IV

THE SIMULTANEOUS ANALYSIS OF VARIANCE TEST.

4.1 Introduction. It is well known that in situations involving the testing of the significance of k mean squares, the usual method of anova gives tests which are not independent. To test all the k hypotheses together we can either use the usual analysis of variance test (which we shall call the joint test) or else we can use a simultaneous test (which we shall call the sim. anova test). This sim. anova test is due to M.N. Ghosh [8]. It will be presently seen that the sim. anova test has a very close tie up with the individual tests of hypotheses, unlike the joint test.

The essential difference between the sim. anova test and the usual joint test can be best illustrated if we consider, for simplicity of exposition, the case of two hypotheses H_1 and H_2. Individual tests on H_1 and H_2 are well known. They are based each on the F statistic. The sim. anova test has an acceptance region which is given by the intersection of the acceptance regions of the individual F acceptance regions of the two hypotheses. Thus if the significance levels of the tests of H_1 and H_2 are α_1 and α_2, that of the sim. anova test will be $\geq (\alpha_1, \alpha_2)$ but $\leq \alpha_1 + \alpha_2$. The extension to the case of several hypotheses is immediate. The joint test is an F test obtained by considering the hypotheses H_1 and H_2 together.

We shall in this chapter study certain distribution problems connected with the test and shall also investigate certain power properties of the test. Suppose in a field experiment there are k hypotheses $H_1, \ldots, H_k$ based on, say, $n_1, \ldots, n_k$ d.f. and suppose $s_1^2, \ldots, s_k^2$ are the k mean squares corresponding to these hypotheses. Let s^2 be an independent estimate of the common unknown variance σ^2 based on m.d.f. (s^2 will be the usual error mean square in the anova). In the usual anova situations we test each hypothesis H_i individually by the F ratio

$$F_i = \frac{s_i^2}{s^2} \frac{m}{n_i} , \qquad (i = 1, \ldots, k) \tag{4.1.1}$$

with (n_i, m) d.f. These k tests are not independent because we are using the same estimate of error variance for all the k tests, and because also of the possible non-orthogonality of the estimates. We shall consider the problem of simultaneous tests of hypotheses by the method of anova. We introduce below the notion of quasi-independent tests of multiple hypotheses which proved useful in Ghosh's development of simultaneous tests.

4.2 Quasi-independent test of hypotheses. Consider k hypotheses $H_1, \ldots, H_k$. For any test of hypotheses we consider the first and second kinds of error. The second kind of error depends upon the alternative hypotheses, $H_1(\theta_1), \ldots, H_k(\theta_k)$; where the hypotheses $H_1, \ldots, H_k$ are, say, $H_1\colon \theta_1 = 0, \ldots, H_k\colon \theta_k = 0$. Tests $T_1, \ldots, T_k$ of $H_1, \ldots, H_k$ are defined to be quasi-independent if, for $i = 1, \ldots, k$, (4.2.1) P_{T_i} (accept $H_i / \theta_i \neq 0; \theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k$) is independent of $\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k$,

and

(4.2.2) P_{T_i} (reject $H_i / \theta_i = 0; \theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k$) is independent of $\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k$.

As an example consider the anova of k linear hypotheses $H_1, \ldots, H_k$.

Let $\underline{x}$ (n x 1) be a set of n uncorrelated stochastic variates with the same (unknown) variance σ^2 and let $E(\underline{x})$ be subject to the constraint:

(4.2.3) $$E(\underline{x}) = A(n \times p)\ \underline{\xi}\ (p \times 1)$$

where $p \leq n$ and $\underline{\xi}$ (p x 1) is a set of unknown parameters (to be estimated or about which we are interested in testing certain hypotheses) and A is a matrix of rank $r \leq p \leq n$, whose elements are given by the particular experimental design.

Assuming that each x_i is $N(E(x_i), \sigma^2)$, $(i = 1, \ldots, n)$; to obtain the test for the hypotheses:

$$(4.2.4)\quad \begin{pmatrix} H_1 \\ \cdot \\ \cdot \\ \cdot \\ H_k \end{pmatrix} : \begin{matrix} q_1 \\ \cdot \\ \cdot \\ \cdot \\ q_k \end{matrix} \underset{p}{\begin{pmatrix} C_1 \\ \cdot \\ \cdot \\ \cdot \\ C_k \end{pmatrix}} \underline{\xi}(p \times 1) = \underline{0}(q \times 1),$$

where $q = \sum_{1}^{k} q_i$, and

$$C = \begin{pmatrix} C_1 \\ \cdot \\ \cdot \\ \cdot \\ C_k \end{pmatrix}$$ is a given matrix of rank $t \leq \min(r,q)$.

Taking $(C_{i11} \quad C_{i12})$, $(i=1,\ldots,k)$ to be the set of t_i independent row vectors of C_i, and $A_1(n \times r)$ the matrix formed by the r independent columns of A we get,

$$(4.2.5)\quad t_i s_i^2 = \underline{x}' A_1(A_1'A_1)^{-1} C_{i11}' \left[C_{i11}(A_1'A_1)^{-1}C_{i11}' \right]^{-1} C_{i11}(A_1'A_1)^{-1}A_1'\underline{x}$$

to be a chi-square variable with t_i d.f.

Now if

$$(4.2.6) \qquad A_1(A_1'A_1)^{-1} C_{i11}' \left[C_{i11}(A_1'A_1)^{-1}C_{i11}' \right]^{-1} C_{i11}(A_1'A_1)^{-1}C_{j11}'$$

$$\cdot \left[C_{j11} (A_1'A_1)^{-1} C_{j11}' \right]^{-1} C_{j11}(A_1'A_1)^{-1} \cdot A_1' = 0 ,$$

then $t_i s_i^2$ and $t_j s_j^2$ will be independent.

Let s^2 be an independent and unbiased estimate of σ^2 with m.d.f. (say, the error mean square in the anova). Then we have, on the null hypothesis (4.2.4), estimates $s_1^2, \ldots, s_k^2$ of σ^2 corresponding to $H_1, \ldots, H_k$. We construct $F_i = (s_i^2/s^2)\,\frac{t_i}{m}$, $(i = 1, \ldots, k)$, and obtain, from the joint distributions $s_1^2, \ldots, s_k^2, s_j^2$ given by

$$(4.2.7) \quad p(s_1^2,\ldots,s_k^2;s^2) = \text{Const.} \prod_{i=1}^{k} (s_i^2)^{\frac{t_i-2}{2}} (s^2)^{\frac{m-2}{2}} e^{-\frac{1}{2}\left[\sum_1^k t_i s_i^2 + ms^2\right]},$$

the joint distribution of $F_1, \ldots, F_k$ equal to

$$(4.2.8) \quad p(F_1, \ldots, F_k) =$$

$$= \left[\Gamma\frac{(\sum_1^k t_i+m)}{2} \Big/ \prod_{i=1}^{k} \Gamma\left(\frac{t_i}{2}\right) \Gamma\left(\frac{m}{2}\right) \right] \left[\prod_1^k F_i^{\frac{t_i-2}{2}} \Big/ \left(1+\sum_1^k F_i\right)^{\frac{\sum t_i+m}{2}} \right].$$

The marginal distributions are, of course, the usual distribution of ratios of chi-square variables. Since any deviation of $H_2, \ldots, H_k$ from the null hypothesis does not affect the marginal distribution of F_1, the first and second kinds of errors in the test of H_1 are independent of the parameters under $H_2, \ldots, H_k$. Similarly for the other hypotheses. Thus the usual F tests of multiple hypotheses where these are orthogonal are quasi-independent. Even when the χ_i^2 variables were not independent, if the marginal distribution of F_i did not involve the parameters of the other hypotheses, then the F tests would be quasi-independent by definition.

There may be different points of view for assigning significance limits in the case of simultaneous tests of hypotheses. In certain situations where the decisions regarding the hypotheses $H_1, \ldots, H_k$ are unrelated it is proper to consider the significance level of each hypothesis individually at 5 per cent or 1 per cent (say). But when these decisions have a joint import it is proper to consider the first kind of error of the simultaneous test as the error of rejection of at least one of the hypotheses when all are in fact true. The significance level of a simultaneous test is defined as the probability of rejecting at least one of the hypotheses when all are in fact true.

4.3 Simultaneous analysis of variance model and tests of hypotheses. Let $\underline{x}(n \times 1)$ denote a set of n uncorrelated stochastic variates with the same (unknown) variance σ^2 and let $E(\underline{x})$ be subject to the constraint:

(4.3.1) $$E(\underline{x}) = A(n \times p)\, \underline{\xi}\, (p \times 1),$$

where $p \leq n$, and $\underline{\xi}$ $(p \times 1)$ is a set of unknown parameters (to be estimated or about which we are interested in testing certain hypotheses), and A is a matrix of rank $r \leq p \leq n$, whose elements are given by the particular experimental design. Let us assume that each x_i is $N(E(x_i), \sigma^2)$, $(i = 1, \ldots, n)$. To obtain the simultaneous test for the k hypotheses on ξ:

$$(4.3.2) \qquad \begin{matrix} q_1 \\ \cdot \\ \cdot \\ \cdot \\ q_k \end{matrix} \begin{pmatrix} C_1 \\ \cdot \\ \cdot \\ \cdot \\ C_k \end{pmatrix} \underline{\xi}\ (p \times 1) = \underline{0}\ (q \times 1),$$

where $q = \sum_{1}^{k} q_i$, and $C = \overset{p}{\begin{pmatrix} C_1 \\ \cdot \\ \cdot \\ \cdot \\ C_k \end{pmatrix}}$ is a given matrix of rank $t \leq \min(r,q)$.

Now for testing the hypothesis $C_i\, \underline{\xi} = \underline{0}$ we get, using certain results given in [19], that

$$(4.3.3) \quad F_i = \left[(n-r)\underline{x}' A_1 (A_1' A_1)^{-1} C_{i11}' \left\{ C_{i11} (A_1' A_1)^{-1} C_{i11}' \right\}^{-1} \right.$$

$$\left. \cdot \left\{ C_{i11} (A_1' A_1)^{-1} A_1' \underline{x} \right\} \right] \div \left[t_i \left\{ \underline{x}' \underline{x} - \underline{x}' A_1 (A_1' A_1)^{-1} A_1' \underline{x} \right\} \right],$$

is an F with t_i and $n-r$ d.f. $(i = 1, 2, \ldots k)$.

Notice that A_1 and $(C_{i11}\ C_{i12})$ are as defined in (4.2). We note that each F_i is distributed as an F with t_i and $(n-r)$

d.f. But the F's are not mutually independent. If the tests are quasi-independent, then the numerators of the different F's are independent chi-square variables with t_i d.f. $(i = 1, \ldots, k)$. But if the tests are not quasi-independent, then the different chi-squares are not independent. The hypothesis $C(q \times p)\ \underline{\xi}\ (p \times 1) = \underline{0}\ (q \times 1)$ is accepted if $F_i \leq a_i$ $(i = 1, \ldots, k)$.

The optimum choice of a_i is not known. We shall choose a_i proportional to t_i. If we want the significance level of the simultaneous test to be α, then we choose a_i such that

$$(4.3.4) \qquad P[F_i \leq a_i;\ i = 1, \ldots, k] = 1 - \alpha.$$

A method to evaluate the probability in the left side of (4.3.4) will be presented in (4.4) - (4.9).

If the tests are not quasi-independent, then the numerators of the different F's are not independent. These tests can easily be made quasi-independent and the results given for quasi-independent tests can then be applied.

4.4 Evaluation of the probability statement given in the left side of (4.3.4). From (4.3.4) we see that the sim. anova test depends on the evaluation of expressions of the form, (with $m = n - r$),

$$(4.4.1) \quad c(t_1,\ldots,t_k;m) \int_0^{a_1} \cdots \int_0^{a_k} \frac{\prod_1^k F_i^{\frac{t_i-2}{2}} \, dF_i}{\left[1+\sum_1^k F_i\right]^{\frac{\Sigma t_i+m}{2}}} ,$$

where $c = c(t_1, \ldots, t_k; m)$ is a function of $t_1,\ldots,t_k$ and m.

Usually we will be interested in obtaining $a_1, \ldots, a_k$ such that

$$(4.4.2) \qquad c \int_0^{a_1} \cdots \int_0^{a_k} \frac{\prod_1^k F_i^{\frac{t_i-2}{2}} \, dF_i}{\left[1+\sum_1^k F_i\right]^{\frac{\Sigma t_i+m}{2}}} = 1 - \alpha.$$

This can be evaluated as follows:

Denoting the left side of (4.4.2) by $I(a_1, \ldots, a_k; t_1, \ldots, t_k; m)$, we get, by integration by parts,

$$(4.4.3) \quad I(a_1, \ldots, a_k; t_1, \ldots, t_k; m)$$

$$= \frac{-c(t_1,\ldots,t_k;m)}{\left\{\frac{(\sum_1^k t_i+m-2)}{2}\right\}} \int_0^{a_1} \cdots \int_0^{a_{k-1}} \left[\frac{\prod_1^{k-1} F_i^{\frac{t_i-2}{2}} \, dF_i}{\left[1+\sum_1^k F_i\right]^{\frac{\Sigma t_i+m-2}{2}}} F_k^{\frac{t_k-2}{2}} \right]_0^{a_k}$$

$$+\frac{(t_k-2)}{2}\frac{c(t_1,\ldots,t_k;m)}{\frac{(\Sigma t_i+m-2)}{2}}\int_0^{a_1}\cdots\int_0^{a_k}\frac{\prod_1^{k-1}F_i^{\frac{t_i-2}{2}}F_k^{\frac{t_k-4}{2}}\prod_1^{k}dF_i}{\left[1+\sum_1^k F_i\right]^{\frac{\Sigma t_i+m-2}{2}}},$$

i.e.,

$$=\frac{-a_k^{\frac{t_k-2}{2}}}{(1+a_k)^{\frac{t_k+m-2}{2}}}\frac{c(t_1,\ldots,t_k;m)}{\left\{\frac{(\Sigma t_i+m-2)}{2}\right\}}\int_0^{\frac{a_1}{1+a_k}}\cdots\int_0^{\frac{a_{k-1}}{1+a_k}}\frac{\prod_1^{k-1}F_i^{\frac{t_i-2}{2}}dF_i}{\left[1+\sum_1^{k-1}F_i\right]^{\frac{\Sigma t_i+m-2}{2}}}$$

$$+\frac{(t_k-2)}{(\Sigma t_i+m-2)}\frac{c(t_1,\ldots,t_k;m)}{c(t_1,\ldots,t_{k-1},t_k-2;m)}I(a_1,\ldots,a_k;t_1,\ldots,t_{k-1},t_k-2,m).$$

Successive reduction will leave us with the evaluation of integrals of the form

$$(4.4.4)\qquad \int_0^{b_1}\cdots\int_0^{b_j}\prod_1^j F_i^{-\frac{1}{2}}dF_i\Big/\left[1+\sum_1^j F_i\right]^{\frac{p+j}{2}}.$$

Now it is easy to see that (4.4.4) is equivalent to

$$(4.4.5) \quad \int_0^{\infty} dv \int_0^{b_1 v} \cdots \int_0^{b_j v} \frac{e^{-v} v^{\frac{p-2}{2}}}{\Gamma(\frac{p+j}{2})} \prod_1^j u_i^{-\frac{1}{2}} e^{-u_i} du_i \, dv .$$

Also from [15], we get,

$$\int_0^x u^{-\frac{1}{2}} e^{-u} du = 2x^{\frac{1}{2}} e^{-\frac{x}{3}} \sum_0^{\infty} A_i x^i .$$

The evaluation of (4.4.2) for given values of $a_1, \ldots, a_k$ can be carried out successively for different values of $t_1, \ldots, t_k$ and m by using the reduction formula (4.4.3).

When m is large, (4.4.2) can be evaluated using tables of the incomplete gamma function [14].

It is easy to notice that the tabulation of (4.4.2) is rather tedious because of the large number of parameters involved. In the next section we shall consider the special but important case when $t_i = t$, ($i = 1, \ldots, k$).

4.5 <u>Special case when $t_i = t(i = 1, \ldots, k)$</u>. In this case we have to obtain an a such that

$$(4.5.1) \quad 1 - \alpha = c(k,t;m) \int_0^a \cdots \int_0^a \prod_1^k F_i^{\frac{t-2}{2}} dF_i \Big/ \Big[1+\sum_1^k F_i\Big]^{\frac{kt+m}{2}} ,$$

where $c(k,t;m) = \Gamma(\frac{kt+m}{2}) / \Gamma^k(\frac{t}{2})\Gamma(\frac{m}{2})$.

It is evident that the right side of (4.5.1) is equivalent to the following statement:

$$(4.5.2) \quad P\left[\frac{ts_i^2}{ms^2} \leq a;\ i=1,2,..k\right] \Longleftrightarrow P\left[\frac{ts_{max}^2}{ms^2} \leq a\right] = 1 - \alpha .$$

Let us call the statistic $u_k = \frac{ts_{max}^2}{ms^2}$, the Studentized largest chi-square.

In order to obtain an a such that (4.5.2) is satisfied, we shall study the distribution of the Studentized largest chi-square.

4.6 <u>Studentized largest chi-square</u>. Let $x_1,\ldots,x_k$ be k independent chi-square variables with the common p.d.f. given by

$$(4.6.1) \qquad p(x) = \frac{x^n e^{-x}}{\Gamma(n+1)} \quad .$$

Let y be another independent chi-square variable with the p.d.f.

$$(4.6.2) \qquad p(y) = \frac{y^m e^{-y}}{\Gamma(m+1)} \quad .$$

The Studentized largest chi-square is defined as

(4.6.3) $$u_k = \frac{x_{max}}{y} \quad .$$

We shall derive in the next few sections certain mathematical results which we shall use to obtain the distribution of u_k.

4.7 Power series expansion for the incomplete gamma type integrals.

Let

(4.7.1) $$I(n,k;x) = \left[\int_0^x \frac{u^n e^{-u}}{\Gamma(n+1)} du \right]^k \quad .$$

Using methods similar to those given in [16], we find an appropriate expansion for I(n,k;x) is given by

(4.7.2) $$I(n,k;x) = \frac{x^{k(n+1)}}{\Gamma^k(n+2)} \, e^{-\frac{n+1}{n+2}kx} \sum_0^{\infty} A_i^{(k)} x^i \quad ,$$

where the A's satisfy the recurrence relation

(4.7.3) $$A_i^{(k)} \left[1 + \frac{i}{k(n+1)} \right] = \left[A_i^{(k-1)} - \frac{1}{(n+2)} A_{i-1}^{(k-1)} + \cdots \right.$$

$$\left. + \frac{(-1)^i}{i\,!} \frac{1}{(n+2)^i} A_0^{(k-1)} \right] + \frac{1}{(n+2)} A_{i-1}^{(k)} \quad , \quad (i=0,1,2,\ldots\ldots).$$

Notice that $A_0^{(k)} = 1$ and $A_1^{(k)} = 0$ for all k.

We shall now prove the convergence of the series $\sum_0^{\infty} A_i^{(k)} x^i$.

4.8 Convergence of the series on the right side of (4.7.2).

Consider

$$(4.8.1) \quad I(n;x) = \int_0^x \frac{u^n e^{-u}}{\Gamma(n+1)}\, du = \frac{x^{n+1}}{\Gamma(n+2)}\, e^{-\frac{n+1}{n+2}x} \sum_0^{\infty} A_i^{(1)} x^i ,$$

where

$$(4.8.2) \quad A_i^{(1)}\left(1+\frac{i}{n+1}\right) = \frac{(-1)^i}{i\,!}\,\frac{1}{(n+2)^i} + \frac{1}{n+2}\, A_{i-1}^{(1)} .$$

Since we will be interested in cases where n is of the form $\frac{r}{2}$, $(r = -1, 0, 1, \ldots)$, we shall prove the convergence of the series on the right side of (4.8.1) for the case $n = \frac{r}{2}$, $(r = -1,0,1,\ldots)$. The case when $r = -1$ has been already considered [16].

Case 1. $n = 0$, i.e., $r = 0$.

In this case

$$(4.8.3) \quad A_{2i+1}^{(1)} = 0 \quad (i=0,1,\ldots),\ A_{2i}^{(1)} = \frac{1}{2^{2i}\,2.3\ldots\overline{2i+1}} \quad (i=1,2,\ldots)$$

and $A_0^{(1)} = 1$.

Hence

$$(4.8.4)\qquad \frac{A^{(1)}_{2i}}{A^{(1)}_{2i-2}} = \frac{1}{4\ .\ 2i(2i+1)} < \frac{1}{16i^2}\ .$$

Hence $\sum_0^{\infty} A_i^{(1)}$ is convergent and the value of the ratio of the i th to the (i-1)th term of the power series in (4.8.1) is less than $\frac{x^2}{16i^2}$. Hence the series $\sum_0^{\infty} A_i^{(1)}\ x^i$ is convergent and therefore the powers of the series are also convergent. It may be noticed that the series (4. 7.2) is rather rapidly convergent, so that for a relatively small x, only a few terms of the series will suffice for any degree of accuracy desired in practice.

Case 2. $n > 0$, i.e., $r > 0$.

Now from (4.8.2) after a little simplification, we get,

$$(4.8.5)\qquad 0 \leq A_i^{(1)} = \frac{(n+1)^i\ (n+1)!}{(n+2)^i\ (n+i+1)!}\ \Big[-\frac{(n+2)}{(n+1)}\ \frac{1}{2!} - \frac{(n+2(n+3)}{(n+1)^2\ 3!} + \cdots$$

$$+ \frac{(-1)^i\ (n+i)!}{i!(n+1)!(n+1)^{i-1}} \Big]\ .$$

Hence

$$(4.8.6) \quad \frac{A_i^{(1)}}{A_{i-1}^{(1)}} = \frac{(n+1)}{(n+2)(n+i+1)} \frac{\text{Sum of first } (i+1) \text{ terms in } (1+\frac{1}{n+1})^{-(n+1)}}{\text{Sum of first } i \text{ terms in } (1+\frac{1}{n+1})^{-(n+1)}}.$$

Hence if i is large,

$$(4.8.7) \quad \frac{A_i^{(1)}}{A_{i-1}^{(1)}} < \frac{1}{i} .$$

Hence $\sum_0^{\infty} A_i^{(1)}$ is convergent and the value of the ratio of the i th to the (i-1)th term of the power series in (4.8.1) is less than $\frac{x}{i}$. Hence the series $\sum_0^{\infty} A_i^{(1)} x^i$ is convergent and therefore the powers of the series are also convergent.

4.9 Distribution of the Studentized largest chi-square. Under the set up given in (4.6), the p.d.f. of $x_{max} = v$ is

$$(4.9.1) \quad p(v) = \frac{k}{\Gamma(n+1)} v^n e^{-v} \left[\int_0^v \frac{x^n e^{-x}}{\Gamma(n+1)} dx \right]^{k-1}$$

$$= \frac{k}{\Gamma(n+1)} \frac{v^{k(n+1)-1}}{\Gamma^{(k-1)}(n+2)} e^{-\frac{k(n+1)+1}{n+2}v} \sum_0^{\infty} A_i^{(k-1)} v^i ,$$

using (4.7.2).

Multiplying (4.6.2) and (4.9.1), using the transformation $u = \frac{v}{y}$, and integrating with respect to y in the interval 0 to ∞, we get

$$(4.9.2) \quad p(u) = \frac{k(n+1)}{\Gamma^k(n+2)\Gamma(m+1)} \sum_0^{\infty} A_i^{(k-1)} \frac{\Gamma\{k(n+1)+m+1+i\} \, u^{k(n+1)+i-1}}{\left[1+\frac{k(n+1)+1}{n+2}u\right]^{k(n+1)+m+i+1}} .$$

From (4.9.2) it is evident that the distribution of u can be tabulated using tables of the incomplete beta function [13]. Upper 5 per cent points of u are given in Table 3 (see appendix) for k = 2 and for different values of m and n.

The methods presented in (4.4) - (4.9) will enable us to evaluate integrals of the form

$$(4.9.3) \quad \int_0^{\infty} du \int_0^{a_1 u} \cdots \int_0^{a_k u} u^m e^{-u} \prod_1^k v_i^{n_i} e^{-v_i} \, dv_i .$$

These integrals are found to be useful in obtaining lower bounds to the power of the Hartley test for equality of several variances from univariate normal populations which is discussed in Chapter III.

Before proceding to the study of the power of the sim. anova test it is interesting to note that a very useful lower bound to the probability statement on the left side of (4.3.4) can be obtained by using a result due to Kimball [11] .

Using the result given in [11], we get

$$(4.9.4) \qquad P\,[\, F_i \le a_i;\ i = 1, \ldots, k \,] > \prod_{1}^{k} P(F_i \le a_i) \ .$$

The expression on the right side of (4.9.4) can be easily obtained from [13].

4.10 Power function of the sim. anova test. When the hypothesis given in (4.3.2) is not true, let the alternative hypothesis be

$$(4.10.1) \qquad \begin{matrix} q_1 \\ \vdots \\ q_k \end{matrix} \underset{p}{\begin{pmatrix} C_1 \\ \vdots \\ C_k \end{pmatrix}} \ \underline{\xi}\ (p \times 1) = \underline{\Delta}\ (q \times 1) = \begin{matrix} q_1 \\ \vdots \\ q_k \end{matrix} \underset{1}{\begin{pmatrix} \underline{\delta}_1 \\ \vdots \\ \underline{\delta}_k \end{pmatrix}} .$$

In the anova situation it is well known that the power function of the test H_1 would involve as a parameter only

$\lambda_1 = \underline{\delta}_{11}'\,[\, C_{111}(A_1'A_1)^{-1}\, C_{111}' \,]^{-1}\ \underline{\delta}_{11}/\sigma^2$; where $\underline{\delta}_{11}\,(t_1 x_1)$ is chosen as to match $C_{111}\,(t_1 \times r)$.

In a similar way it is easy to verify that the power function of the sim. anova test, in the case of quasi-independent tests, would involve as parameters only $\lambda_1, \ldots, \lambda_k$; where λ_j

$= \underline{\delta}_{j1}' \left[C_{j11}(A_1'A_1)^{-1} C_{j11}' \right]^{-1} \underline{\delta}_{j1}/\sigma^2$.

Also it is well known that under this set up the power function would be equal to

$$(4.10.2) \quad P = 1 - c \int_0^{\infty} p(s^2)ds^2 \int_D p(s_1^2,\ldots,s_k^2;\lambda_1,\ldots,\lambda_k;$$

$$t_1,\ldots,t_k;a_1,\ldots,a_k)\, ds_1^2 \ldots ds_k^2,$$

where

$$p(s_1^2,\ldots,s_k^2;\lambda_1,\ldots,\lambda_k;t_1,\ldots,t_k;a_1,\ldots,a_k)$$

$$= \sum_{n_1=0}^{\infty} \frac{(s_1^2)^{\frac{t_1}{2}+n_1-1}}{\Gamma(\frac{t_1}{2}+n_1)} \frac{e^{-t_1\frac{s_1^2}{2}-\lambda_1}}{n_1!} \lambda_1^{n_1} \ldots$$

$$\ldots \sum_{n_k=0}^{\infty} \frac{(s_k^2)^{\frac{t_k}{2}+n_k-1}}{\Gamma(\frac{t_k}{2}+n_k)} \frac{e^{-t_k\frac{s_k^2}{2}-\lambda_k}}{n_k!} \lambda_k^{n_k},$$

$$p(s^2) = \text{Const.}\ (s^2)^{\frac{m-2}{2}}\ e^{-\frac{ms^2}{2}}\ ,$$

D is the domain: $\left\{ \begin{array}{c} 0 \leq s_1^2 \leq \frac{m}{t_1}\, a_1 s^2 \\ \vdots \\ 0 \leq s_k^2 \leq \frac{m}{t_k}\, a_k s^2 \end{array} \right\}$,

and $c > 0$ is a pure constant independent of the λ's .

We shall now prove an optimum property of the sim. anova test.

Theorem 1. The power function of the sim. anova test, in the case of quasi-independent tests, is a monotonic increasing function of the absolute value of the square root of each of the deviation parameters separately.

Proof: The second kind of error (complement of the power) of the sim. anova test is equal to

$$(4.10.3)\quad c \int_0^\infty p(s^2)\, ds^2 \int_D p(s_1^2,\ldots,s_k^2;\lambda_1,\ldots,\lambda_k;t_1,\ldots t_k;a_1,\ldots$$

$$\ldots,a_k)\, ds_1^2 \ldots ds_k^2\ .$$

It is well known that for the purpose of discussing the power properties of the sim. anova test, we can start, without any loss of generality, from the canonical probability law:

$$(4.10.4)\quad \text{Const. Exp}\left[-\frac{1}{2}\left(\sum_{1}^{t_1} x_{1i}^2 + \dots + \sum_{1}^{t_k} x_{ki}^2 + \sum_{1}^{m} y_i^2\right)\right] \prod_{1}^{k}\prod_{1}^{t_i} dx_{ij} \prod_{1}^{m} dy_i ,$$

where $-\infty \leq x\text{'s} \leq \infty$ and $-\infty \leq y\text{'s} \leq \infty$.

Also it is well known that

$$(4.10.5)\quad \left\{\begin{array}{l} t_1 s_1^2 = (x_{11} + \sqrt{\lambda_1})^2 + x_{12}^2 + \dots + x_{1t_1}^2 \\ \quad \cdot \qquad\quad \cdot \qquad\quad \cdot \quad \dots \quad \cdot \\ \quad \cdot \qquad\quad \cdot \qquad\quad \cdot \quad \dots \quad \cdot \\ t_k s_k^2 = (x_{k1} + \sqrt{\lambda_k})^2 + x_{k2}^2 + \dots + x_{kt_k}^2 \\ ms^2 = y_1^2 + \dots + y_m^2. \end{array}\right\} .$$

Under this set up, the second kind of error of the sim. anova test is equal to

$$(4.10.6)\quad \beta = c\int_{D_1} \text{Exp}\left[-\frac{1}{2}\left(\sum_{j=1}^{k}\sum_{i=1}^{t_j} x_{ji}^2 + \sum_{1}^{m} y_i^2\right)\right] \prod_{1}^{k}\prod_{1}^{t_j} dx_{ji} \prod_{1}^{m} dy_i ,$$

where D_1 is the domain:

$$\left\{\begin{array}{l} 0 \leq (x_{11}+\sqrt{\lambda_1})^2+\sum_2^{t_1} x_{1i}^2 \leq a_1 \sum_1^m y_i^2 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ 0 \leq (x_{k1}+\sqrt{\lambda_k})^2+\sum_2^{t_k} x_{ki}^2 \leq a_k \sum_1^m y_i^2 \end{array}\right\}$$

and $c > 0$ is a pure constant independent of the λ's.

It is easy to see that β is symmetric in the λ's. Hence we shall prove the theorem only for λ_1. For any other λ_i the theorem is immediate because of the symmetry in the variables.

Notice that λ_1 occurs only with x_{11}. From (4.10.5) we get the limits of x_{11} to be

$$(4.10.7) \quad -\left(a_1 \sum_1^m y_i^2 - \sum_2^{t_1} x_{1i}^2\right)^{\frac{1}{2}} - \sqrt{\lambda_1} \leq x_{11} \leq \left(a_1 \sum_1^m y_i^2 - \sum_2^{t_1} x_{1i}^2\right)^{\frac{1}{2}} - \sqrt{\lambda_1}.$$

In (4.10.6) perform first the integration over x_{11}. The contribution to the total p.d.f. (4.10.4) made by x_{11} is Const. $e^{-\frac{1}{2} x_{11}^2}$. The upper and lower limits of the x_{11} integration are ℓ_1 and ℓ_2 given by

$$(4.10.8) \qquad \ell_1 = \left(a_1 \sum_1^m y_i^2 - \sum_2^{t_1} x_{1i}^2\right)^{\frac{1}{2}} - \sqrt{\lambda_1}$$

and $\ell_2 = -(a_1 \sum_1^m y_i^2 - \sum_2^{t_1} x_{1i}^2)^{\frac{1}{2}} - \sqrt{\lambda_1}$.

If we now differentiate w.r. to $\sqrt{\lambda_1}$ the integral of (4.10.4) over the domain D_1 we get, through the x_{11} integral an integrand which is

$$(4.10.9) \qquad - e^{-\frac{\ell_1^2}{2}} + e^{-\frac{\ell_2^2}{2}} \quad .$$

For all positive values of $\sqrt{\lambda_1}$, the expression in (4.10.9) will be negative, and for all negative values of $\sqrt{\lambda_1}$ it will be positive. Thus

$$(4.10.10) \qquad \frac{\partial \beta}{\partial \sqrt{\lambda_1}} < 0 \qquad \text{if} \quad \sqrt{\lambda_1} > 0$$

$$\text{and} \qquad > 0 \qquad \text{if} \quad \sqrt{\lambda_1} < 0.$$

By symmetry in the variables, the same is true of any $\sqrt{\lambda_i}$, i.e.,

$$(4.10.11) \qquad \frac{\partial \beta}{\partial \sqrt{\lambda_i}} < 0 \qquad \text{if} \quad \sqrt{\lambda_i} > 0$$

and $\qquad > 0 \qquad$ if $\sqrt{\lambda_i} < 0$, $(i = 1, \ldots, k)$.

Hence the second kind of error of the sim. anova test is a decreasing function of each $\left|\sqrt{\lambda_i}\right|$ separately and hence the power of the test (complement of the second kind of error) is an increasing function of each $\left|\sqrt{\lambda_i}\right|$ separately.

Hence the theorem.

CHAPTER V

THE STUDENTIZED MAXIMUM MODULUS TEST.

5.1 Introduction. In a 2^n factorial experiment suppose we are interested in testing the hypothesis that all linear functions of the treatment effects t_{ij} are simultaneously zero. The estimates of the treatment effects are assumed to be independently and normally distributed with a common variance σ^2 which can be independently estimated by an appropriate multiple of the error mean square in the anova.

The test of this hypothesis can be obtained by taking the intersection of the n Student's 't' acceptance regions [20]. It is easily shown that this test is based on the Studentized maximum modulus given by $u_n = |x| / s$, where $x_1, \ldots, x_n$ are independent $N(0, \sigma^2)$ variates, s^2 is an unbiased and independent estimate of σ^2 based on m.d.f., and $|x|$ is the maximum of $|x_1|, \ldots, |x_n|$.

The distribution problem connected with the test has been solved in [16]. We shall investigate certain optimum power properties of the test in (5.2).

Before proceding with the power properties of the Studentized maximum modulus test we shall prove two lemmas which are useful in demonstrating the optimum properties of the test.

Lemma 1.

(5.1.1) $I(0, \infty; m)\, U(-us, us; \xi_1) \ldots U(-us, us; \xi_n)$

$$< \quad I(0, \infty; m) \left[U(-us, us; 0) \right]^n ,$$

where $I(0, \infty; m) = c \int_0^{\infty} s^{m-1} e^{-\frac{ms^2}{2}} \, ds$

and $U(x,y;\xi) = \int_x^y \frac{e^{-\frac{1}{2}(t-\xi)^2}}{\sqrt{2\pi}} \, dt.$

Proof: It is easy to see that the expression on the left side of (5.1.1) can be put in the form

(5.1.2) $I(0, \infty; m)\, U(-us-\xi_1, us-\xi_1; 0) \ldots U(-us-\xi_n, us-\xi_n; 0).$

Also it is easy to check that

(5.1.3) $U(-a+b, a+b; 0) < U(-a, a; 0)$ for every $b \neq 0$.

Hence the lemma.

Lemma 2. If

$$(5.1.4)\quad \beta(\underline{\xi}) = I(0, oo; m)\ U(-us-\xi_1, us-\xi_1; 0) \ldots U(-us-\xi_n, us-\xi_n; 0),$$

then

$$\frac{\partial\beta}{\partial\xi_i} < 0 \qquad \text{if } \xi_i > 0$$

$$\text{and} \quad > 0 \qquad \text{if } \xi_i < 0 \qquad (i=1,\ldots,n).$$

Proof:

$$(5.1.5)\quad \sqrt{2\pi}\,\frac{\partial\beta}{\partial\xi_i} = I(0,oo;m)\, U(-us-\xi_1, us-\xi_1; 0) \ldots$$

$$U(-us-\xi_{i-1}, us-\xi_{i-1}; 0)\, U(-us-\xi_{i+1}, us-\xi_{i+1}; 0) \ldots$$

$$U(-us-\xi_n, us-\xi_n; 0)\ \left[e^{-\frac{1}{2}(us+\xi_i)^2} - e^{-\frac{1}{2}(us-\xi_i)^2} \right],$$

which is negative if $\xi_i > 0$, and positive if $\xi_i < 0$.

Hence by the symmetry in the variables, we get,

$$(5.1.6)\quad \frac{\partial\beta}{\partial\xi_i} < 0 \qquad \text{if } \xi_i > 0$$

$$\text{and} \quad > 0 \qquad \text{if } \xi_i < 0 \qquad (i=1, \ldots, n).$$

Hence the lemma.

5.2 Power function of the Studentized maximum modulus test. Under the set up given in (5.1), if the hypothesis is not true, let x_1, ..., x_n be independent $N(\xi_i, \sigma^2;\ i=1, \ldots, n)$.

The second kind of error of the Studentized maximum modulus test is

$$(5.2.1)\quad \beta = P\left[\frac{|x_i|}{s} \leq u \Big/ \xi;\ \text{all } i=1,\ldots,n \right] = P\left[\frac{|x|}{s} \leq u \Big/ \xi \right]$$

$$= I(0,\infty;m)U(-us,us;\xi_1) \ldots U(-us,us;\xi_n),$$

where u is so chosen that

$$1 - \alpha = P\left[\frac{|x|}{s} \leq u \right] = I(0,\infty;m) \left[U(-us,us;0) \right]^n ,$$

$|x|$ is the maximum of $|x_1|$, $|x_2|$, ..., $|x_n|$; and α is the given level of significance of the test.

We shall now prove the following properties of the Studentized maximum modulus test.

Property I. The Studentized maximum modulus test is completely unbiased.

Proof: The proof follows from lemma 1 and (5.2.1).

Property II. The power function of the Studentized maximum modulus test is a monotonically increasing function of each of the absolute values of the deviation parameters ξ_1, ..., ξ_n separately.

Proof: The proof follows from lemma 2 and (5.2.1).

Property III.

$$(5.2.2) \qquad P\left[\frac{|x|}{s} \le u\right] > \prod_{i=1}^{n} P\left[\frac{|x_i|}{s} \le u\right] .$$

Proof: We have

$$(5.2.3) \qquad P\left[\frac{|x|}{s} \le u\right] \Longleftrightarrow P\left[\begin{array}{c} |x_1| \le us \\ \vdots \\ |x_n| \le us \end{array}\right]$$

$$> \prod_{i=1}^{n} P\left[\, |x_i| \le us \right]$$

from [11].

Notice that the right side of (5.2.3) is easy to evaluate and hence we can easily obtain an upper bound to the error probability of the first kind.

APPENDIX

TABLE 1

Values of $X_1'^2$ and $X_2'^2$ (see Chapter I) for α = .05 and for different values of n', where $n' = n - 1$ is the d.f. of X^2.

n'	$X_1'^2$	$X_2'^2$	n'	$X_1'^2$	$X_2'^2$
1	$.0^332$	7.82	13	5.32	25.90
2	.08	9.53	14	5.95	27.26
3	.30	11.19	16	7.24	29.95
4	.61	12.80	18	8.58	32.61
5	.99	14.37	20	9.96	35.23
6	1.43	15.90	22	11.36	37.82
7	1.90	17.39	24	12.79	40.39
8	2.41	18.86	26	14.24	42.93
9	2.95	20.31	28	15.71	45.45
10	3.52	21.73	30	17.21	47.96
11	4.10	23.13	40	24.86	60.32
12	4.70	24.52	60	40.93	84.23

TABLE* 2

Values of F_1' and F_2' (see Chapter I) for α = .05 and for different values of n_1' and n_2', where $n_1' = n_1 - 1$ and $n_2' = n_2 - 1$ are the d.f. of F.

n_2' \ n_1'	2	4	6	8	10	12	16	20	24	30
2	39.0	30.5	28.0	26.8	26.1	25.6	25.1	24.8	24.6	24.4
4	12.9	9.60	8.56	8.05	7.75	7.55	7.30	7.16	7.06	6.97
6	9.14	6.64	5.82	5.42	5.17	5.01	4.81	4.69	4.61	4.53
8	7.73	5.53	4.80	4.43	4.21	4.07	3.88	3.77	3.69	3.62
10	7.00	4.95	4.27	3.93	3.72	3.58	3.40	3.29	3.22	3.14
12	6.56	4.61	3.95	3.62	3.41	3.28	3.10	3.00	2.93	2.85
16	6.05	4.21	3.58	3.26	3.06	2.93	2.75	2.65	2.58	2.51
20	5.76	3.98	3.37	3.06	2.87	2.74	2.56	2.46	2.39	2.32
24	5.58	3.84	3.24	2.93	2.74	2.61	2.43	2.34	2.27	2.20
30	5.41	3.70	3.12	2.81	2.62	2.49	2.31	2.22	2.15	2.07

*The values given in the table are F_2'. To obtain the value of F_1' for n_1', n_2'; take the reciprocal of F_2' with n_2', n_1'.

TABLE 3

Upper 5 per cent points of u(see Chapter IV) for different values of m and n when k = 2.

m \ n	0	1	2	3	4	5
2	6.90	5.82	5.38	5.14	4.98	4.89
3	5.86	4.83	4.33	4.19	4.04	3.96
4	5.32	4.34	3.93	3.71	3.56	3.48
5	4.99	4.03	3.63	3.42	3.27	3.19
6	4.77	3.83	3.44	3.21	3.06	2.98
7	4.65	3.69	3.29	3.07	2.92	2.85
8	4.51	3.57	3.19	2.96	2.82	2.75
9	4.41	3.46	3.11	2.88	2.75	2.66
oo	3.69	2.79	2.41	2.19	2.05	1.94

BIBLIOGRAPHY

[1] Bartlett, M. S., "Properties of Sufficiency and Statistical Tests," Procedings of the Royal Society of London, 160 A (1937), 268-282.

[2] Bishop, D. J., and Nair, U. S., "A Note on Certain Methods of Testing for the Homogenity of a Set of Estimated Variances," Supplement to the Journal of the Royal Statistical Society, VI (1939), 89-99.

[3] Brown, G. W., "On the Power of the L_1 Test for Equality of Several Variances," Annals of Mathematical Statistics, X (1939), 119-128.

[4] Cochran, W. G., "The Distribution of the Largest of a Set of Estimated Variances as a Fraction of Their Total," Annals of Eugenics, XI (1941), 47-52.

[5] David, H. A., "Upper 5 and 1 Per Cent Points of the Maximum F-Ratio," Biometrika, XXXIX(1952), 422-424.

[6] Dunnett, C. W., and Sobel, M, "A Bivariate Generalization of Student's Distribution, with Tables for Certain Special Cases," Biometrika, XXXXI (1954), 153-169.

[7] Eisenhart, Hastay and Wallis, Techniques of Statistical Analysis, McGraw-Hill Book Company, Inc., First edition (1947).

[8] Ghosh, M. N., "Simultaneous Test of Linear Hypotheses by the Method of Analysis of Variance," (unpublished manuscript), 1953.

[9] Hartley, H. O., "Testing the Homogeneity of a Set of Variances," Biometrika, XXXI (1939-40), 249-255.

[10] Hartley, H. O., "The Maximum F-Ratio as a Short Cut Test for Heterogeneity of Variance," Biometrika, XXXVII (1950), 308-312.

[11] Kimball, A. W., "On Dependent Tests of Significance in the Analysis of Variance," Annals of Mathematical Statistics, XXII (1951), 600-602.

[12] Neyman, J., and Pearson, E. S., "On the Problem of k Samples," Bulletin de l'academie Polonaise Sciences, Lettres, Serie A (1931), 460-468.

[13] Pearson, K., Tables of the Incomplete Beta Function, London, The "Biometrika" office, 1934.

[14] Pearson, K., Tables of the Incomplete Γ - Function, Cambridge University Press, 1946.

[15] Pillai, K. C. S., "On the Distributions of Midrange and Semirange in Samples from a Normal Population," Annals of Mathematical Statistics, XXI (1950), 100-105.

[16] Pillai, K. C. S., and Ramachandran, K. V., "On the Distribution of the Ratio of the i th Observation in an Ordered Sample from a Normal Population to an Independent Estimate of the Standard Deviation," Annals of Mathematical Statistics, XXV (1954), 565-572.

[17] Roy, S. N. "Notes on Testing Composite Hypotheses - II," Sankhya, IX (1948-49), 19-38.

[18] Roy, S. N., "On a Heuristic Method of Test Construction and Its Use in Multivariate Analysis," Annals of Mathematical Statistics, XXIV (1953), 220-238.

[19] Roy, S. N., "Lecture Notes on Multivariate Analysis," (Unpublished).

[20] Roy, S. N., and Bose, R. C., "Simultaneous Confidence Interval Estimation," Annals of Mathematical Statistics, XXIV (1953), 513-536.

[21] Scheffé, H., "A Method for Judging All Contrasts in the Analysis of Variance," Biometrika, XXXX (1953), 87-104.

[22] Thompson, C. M., and Merrington, M., "Tables for Testing the Homogeneity of a Set of Estimated Variances," Biometrika, XXXIII (1946), 296-304.

[23] Tukey, J. W., "Allowances for Various Types of Error Rates," (unpublished invited address, Blacksburg meeting of the Institute of Mathematical Statistics, March, 1952).