Open access • Book Chapter • DOI:10.1016/S0169-7161(80)01008-5

6 Monotonicity and unbiasedness properties of ANOVA and MANOVA tests

Somesh Das Gupta

Published on: 01 Jan 1980 - Handbook of Statistics (Elsevier)

Topics: Multivariate statistics and Multivariate analysis of variance

Monotonicity and Unbiasedness Properties of ANOVA and MANOVA Tests.[1]

Somesh Das Gupta*
University of Minnesota

Technical Report No. 318

(1) To be published in "Handbook of Statistics, Vol. 1: Analysis of Variance" by North-Holland Publishing Co.

*Partially supported by a grant from the Mathematics Division, U.S. Army Research Office, Durham, N.C. Grant no. DAAG-29-76-G-0038.

1. Introduction

The multivariate analysis of variance problem for the normal case may be posed as follows: Let $X: p \times n$ be a random matrix such that its column vectors are independently distributed as $\mathfrak{n}_p(\cdot,\Sigma)$, where Σ is an unknown positive-definite matrix; moreover,

(1.1) $$\mathcal{E}(X') = A\Theta ,$$

where $A: n \times m$ is a known matrix of rank r and $\Theta: m \times p$ is a matrix of unknown parameters. The problem is to test $H_0: G'\Theta = 0$ against $H_1: G'\Theta \neq 0$, where G' is a known $s \times m$ matrix of rank s such that $G = A'B$ for some $B: n \times s$. This problem can easily be reduced to the following canonical form: Let $Y_1, \ldots, Y_n$ be n independently distributed $p \times 1$ random vectors such that $Y_\alpha \sim \mathfrak{n}_p(\mu_\alpha,\Sigma)$, where $\mu_{r+1} = \ldots = \mu_n = 0$, and Σ along with $\mu_1, \ldots, \mu_r$ are unknown, Σ being positive-definite. The problem is to test

(1.2) $$H_0: \mu_1 = \ldots = \mu_s = 0$$

against H_1: "not H_0", where $s \leq r$. In this set-up $S_0 \equiv \sum_{\alpha=1}^{s} Y_\alpha Y_\alpha'$ and $S_e = \sum_{\alpha=r+1}^{n} Y_\alpha Y_\alpha'$ are called the sums of products (s.p.) matrices due to the hypothesis H_0 and error, respectively; the corresponding degrees of freedom are s and $n_e \equiv n-r$.

The following tests (represented by their acceptance regions) are most-often considered in the literature:

(a) Likelihood-ratio test:

(1.3) $$\det(S_e)/\det(S_e + S_0) \geq k_1 \quad (0 < k_1 < 1)$$

(b) Roy's maximum-root test:

(1.4) $\max[\text{characteristic root of } S_0 S_e^{-1}] \le k_2 \qquad (0 < k_2)$

(c) Lawley-Hotelling's trace test:

(1.5) $\operatorname{tr}(S_0 S_e^{-1}) \le k_3 \qquad (0 < k_3)$

(d) Pillai's trace test:

(1.6) $\operatorname{tr}[S_0(S_0 + S_e)^{-1}] \le k_4 \qquad (0 < k_4 < \min(p,s)).$

Note that the first three tests are defined only when $n_e \ge p$ in which case S_e is non-singular with probability 1. The last test is defined when $n_e + s > p$. All these four tests are members of a class of invariant tests which is defined as follows.

Let

(1.7)
$$Y_{(1)} = (Y_1, \ldots, Y_s),\ Y_{(2)} = (Y_{s+1}, \ldots, Y_r),$$
$$Y_{(3)} = (Y_{r+1}, \ldots, Y_n)\ .$$

A set of sufficient statistics is given by

(1.8) $$(Y_{(1)}, Y_{(2)}, S_t \equiv Y_{(1)}Y'_{(1)} + Y_{(3)}Y'_{(3)})\ .$$

S_t is positive-definite when $n_e + s \ge p$. Consider the following transformation:

(1.9) $$(L, A, B)\ (Y_{(1)}, Y_{(2)}, S_t) = (AY_{(1)}L,\ AY_{(2)}+B,\ AS_tA')\ ,$$

where

$L \in \mathcal{O}_s$ = the class of all $s \times s$ orthogonal matrices,

$A \in \mathcal{L}_p$ = the class of all $p \times p$ nonsingular matrices,

$B \in \mathfrak{M}_{p,r-s}$ = the class of all $p \times (r-s)$ matrices.

This transformation keeps the above model and the testing problem invariant. The composition of two such transformations is given by

$$(1.10) \qquad (L_1, A_1, B_1)(L_2, A_2, B_2) = (L_2L_1, A_1A_2, A_1B_2 + B_1) .$$

Then the collection G of (L, A, B) with the above binary operation is a group of transformations acting on $\mathfrak{M}_{p,s} \times \mathfrak{M}_{p,r-s} \times \mathcal{S}_p^+$, where

$\mathcal{S}_p^+$ = the collection of all $p \times p$ positive-definite matrices.

Let Φ_G be the class of all non-randomized tests invariant under G.

Lemma. When $n_e + s > p$, a set of maximal invariants under G in the space of sufficient statistics $(Y_{(1)}, Y_{(2)}, S_t)$ is given by the ordered non-zero characteristic roots of $S_0S_t^{-1}$, denoted by $d_1 < \ldots < d_\ell$ where $\ell = \min(s,p)$. When $n_e + s \le p$ there is no non-trivial invariant test.

Suppose $n_e \ge p$ and let $c_1 \ge \ldots \ge c_\ell$ be the ordered non-zero characteristic roots of $S_0S_e^{-1}$. Then $d_i = c_i/(1+c_i)$.

Next we shall consider two important special cases.

(I) $s = 1, n_e \ge p$. The acceptance regions of all the above four tests reduce to

$$(1.11) \qquad c_1 = Y_1'(Y_{(3)}Y_{(3)}')^{-1}Y_1 \le k .$$

This is the UMP invariant test for its size.

(II) $p = 1, n_e \ge 1$. The acceptance regions of all the above four tests reduce to

$$(1.12) \qquad \sum_{\alpha=1}^{s} Y_\alpha^2 \Big/ \sum_{\alpha=r+1}^{n} Y_\alpha^2 \le k .$$

This is also the UMP invariant test for its size.

Except for these two special cases UMP invariant test does not exist. All the above four tests are known to be admissible. Instead of comparing the power functions of different tests we shall be concerned in this paper with the behavior of the power function of a given test with respect to the non-centrality parameters involved; in particular we shall study whether the unbiasedness property is satisfied by a given test.

Let $\tau_1^2, \ldots, \tau_\ell^2$ be the possible non-zero characteristic roots of $\Sigma^{-1}MM'$, where $M = (\mu_1, \ldots, \mu_s)$. Then the power function of any test in Φ_G involves M and Σ only through $\tau_1^2, \ldots, \tau_\ell^2$. We shall study conditions under which the power function of an invariant test increases monotonically in each τ_i^2. Under some additional conditions we shall get a more refined property of this monotonicity.

2. Monotonicity of the power functions of the UMP invariant tests in the two special cases.

The monotonicity property in the above two special cases can be easily proved using the following elementary result.

Theorem 2.1: Let Z be a random variable distributed as $N(0,1)$. Then

$$(2.1) \qquad \pi(\tau) \equiv P\{|Z + \tau| \le k\}$$

for $k > 0$ is a symmetric function of τ and decreases monotonically as τ^2 increases.

The theorem is proved easily by studying the first derivative of π with respect to τ. Later we shall show that this result also holds when the density of Z is symmetric about the origin and unimodal (with the mode at the origin). It will also be extended to the multivariate case.

Corollary 2.1. Let Z_1 and Z_2 be independently distributed according to the non-central chi-square $\chi^2_{n_1}(\tau^2)$ and the central chi-square $\chi^2_{n_2}$ distributions, respectively; n_1 and n_2 are positive integers and τ^2 is the non-centrality parameter of Z_1. Then

(2.2) $\Pr[Z_1/Z_2 \le k] \qquad (0 < k)$

is a monotonically decreasing function of τ^2 .

Proof. Write

$$Z_1 = Z_{11}^2 + \dots + Z_{1n_1}^2 ,$$

where Z_{1i}'s are independently distributed as $N(\cdot,1)$ with $\mathcal{E}Z_{11} = \tau$ and $\mathcal{E}Z_{1\alpha} = 0$ for $\alpha > 1$; moreover Z_{1i}'s are distributed independently of Z_2 . Such a decomposition of Z_1 is clearly possible. Now apply Theorem 2.1 for Z_{11} holding Z_2 and $Z_{1\alpha}$'s for $\alpha > 1$ fixed.

The above corollary is true also for non-integral positive n_1 and n_2 . One may use the monotone likelihood-ratio property of the non-central F-distribution.

Let us now consider the two special cases given by $s = 1$ and $p = 1$.

Case I. $s = 1$, $n_e \ge p$. The critical region of the Hotelling's T^2-test can be expressed as

(2.3) $Y_1'(Y_{(3)}Y_{(3)}')^{-1}Y_{(1)} > \{p/(n_e-p+1)\}F^{\alpha}_{p,n_e-p+1}$

where $F^{\alpha}_{a,b}$ is the upper α-fractile of the F-distribution with a and b degrees of freedom. The power of this test is

(2.4) $\Pr[F_{p,n_e-p+1}(\tau^2) > F^{\alpha}_{p,n_e-p+1}]$,

where $\tau^2 = \mu_1'\Sigma^{-1}\mu_1$. It follows from Corollary 2.1 that the power of this test increases monotonically with τ^2 .

Case II. $p = 1$, $n_e \ge 1$. The critical region of the ANOVA F-test can be expressed as

(2.5) $\sum_{\alpha=1}^{s} Y_\alpha^2 \Big/ \sum_{\alpha=r+1}^{n} Y_\alpha^2 > \{s/n_e\}\, F^{\alpha}_{s,n_e}$.

The power of this test is

(2.6) $$\Pr[F_{s,n_e}(\tau^2) > F^{\alpha}_{s,n_e}] ,$$

where $\tau^2 = \sum_{i=1}^{s} \mu_i^2/\Sigma$. Again, Corollary 2.1 shows that the power of this test increases monotonically with τ^2 .

3. Mathematical preliminaries.

The key to all the results in this paper is the following well-known inequality due to Brunn-Minkowski.

Theorem 3.1. Let A_1 and A_2 be two non-empty convex sets in R^n . Then

(3.1) $$V_n^{1/n} (A_1 + A_2) \geq V_n^{1/n} (A_1) + V_n^{1/n} (A_2) ,$$

where V_n stands for the n-dimensional volume, and

$$A_1 + A_2 = \{x_1 + x_2 : x_1 \in A_1, x_2 \in A_2\} .$$

This inequality was first proved by Brunn [5] in 1887 and the conditions for equality to hold were derived by Minkowski [26] in 1910. Later in 1935 Lusternik [25] generalized this result for non-empty arbitrary measurable sets A_1 and A_2 and derived conditions for equality to hold.

This inequality led Anderson [1] to generalize Theorem 2.1 to the multivariate case. We shall present here a minor extension of Anderson's result. Following Anderson we shall call a non-negative function f on R^n unimodal, if

(3.2) $$K_{f,u} \equiv \{x \in R^n : f(x) \geq u\}$$

is convex for all u , $0 \leq u < \infty$. We shall call a (real-valued) function f on R^n centrally symmetric if $f(x) = f(-x)$ for all $x \in R^n$.

Theorem 3.2. Let G be a group of linear Lebesgue measure preserving transformations of R^n onto R^n. Let f be a non-negative (Borel-measurable) function on R^n such that f is unimodal, integrable with respect to the Lebesgue measure μ_n on R^n, and $f(x) = f(gx)$ for all $g \in G$, $x \in R^n$. Let E be a convex set in R^n such that $E = gE$ for all g in G. Then for any fixed $\tau \in R^n$ and any τ^* in the convex-hull of the G-orbit of τ defined by $G(\tau) \equiv \{g\tau: g \in G\}$

$$(3.3) \qquad \int_{E+\tau^*} f(x)dx \geq \int_{E+\tau} f(x)dx .$$

Proof. First note that

$$(3.4) \qquad \int_{E+\tau} f(x)dx = \int_0^\infty \mu_n[K_{f,u} \cap (E+\tau)]du ,$$

where $K_{f,u}$ is defined in (3.2). Then for $g \in G$, $K_{f,u} = gK_{f,u}$, and

$$(3.5) \qquad \mu_n[K_{f,u} \cap (E+\tau)] = \mu_n[gK_{f,u} \cap g(E+\tau)]$$

$$= \mu_n[K_{f,u} \cap (E+g\tau)] .$$

Note that $K_{f,u} \cap (E+\tau)$ and $K_{f,u} \cap (E+g\tau)$ are both either empty or non-empty. Let $g_1, \ldots, g_m$ be in G and $\tau^* = \sum_{i=1}^m \lambda_i g_i \tau$, where $0 \leq \lambda_i \leq 1$, $\sum_{i=1}^m \lambda_i = 1$. Then

$$(3.6) \qquad K_{f,u} \cap (E+\tau^*) \supset \sum_{i=1}^m \lambda_i[K_{f,u} \cap (E+g\tau)] ,$$

whenever $K_{f,u} \cap (E+\tau)$ is non-empty. Theorem 3.1 now yields

$$(3.7) \qquad \mu_n[K_{f,u} \cap (E+\tau^*)] \geq [\sum_{i=1}^m \lambda_i \mu_n^{1/n}\{K_{f,u} \cap (E+g\tau)\}]^n$$

$$= \mu_n[K_{f,u} \cap (E+\tau)] .$$

Integrating with respect to u yields the theorem.

We shall improve Theorem 3.2 by using a condition on f which is stronger than unimodality. Following Das Gupta [11] we shall call a non-negative function f on R^n 0-unimodal (or, strongly unimodal) if for any x_0, x_1 in R^n and any $0 < \theta < 1$

(3.8) $$f[(1-\theta)x_0 + \theta x_1] \geq f^{1-\theta}(x_0)f^{\theta}(x_1) .$$

(Borel-measurable)

<u>Theorem 3.3</u>. Let f be a non-negative 0-unimodal function on R^n such that f is integrable with respect to μ_n. Then for any two Borel-measurable non-empty sets E_0 and E_1

(3.9) $$\int_{(1-\theta)E_0+\theta E_1} f(x)dx \geq [\int_{E_0} f(x)dx]^{1-\theta}[\int_{E_1} f(x)dx]^{\theta} .$$

<u>Proof</u>. For $u \in R^1$ define

(3.10) $$C = \{(x,u) \in R^n \times R^1 : f(x) \geq \exp(-u)\} .$$

Let C_u be the u-section of C. Then for any measurable set $E \subset R^n$

(3.11) $$\int_E f(x)dx = \int_{-\infty}^{\infty} \mu_n[C_u \cap E] \exp(-u)du .$$

We assume that the integrals in the left-hand side of (3.9) are positive (excluding the trival cases). Define

(3.12) $$h_{\theta}(u) = \mu_n[C_u \cap \{(1-\theta)E_0 + \theta E_1\}] .$$

Let S_i be the support of h_i $(i = 0,1)$. Then for $u_0 \in S_0$, $u_1 \in S_1$, $u = (1-\theta)u_0 + \theta u_1$

(3.13) $$h_{\theta}(u) \geq [h_0(u_0)]^{1-\theta}[h_1(u_1)]^{\theta} .$$

To see this, note that

(3.14) $$C_u \cap \{(1-\theta)E_0 + \theta E_1\} \supset (1-\theta)(C_{u_0} \cap E_0) + \theta(C_{u_1} \cap E_1) .$$

From Brunn-Minkowski-Lusternik inequality we get

(3.15) $$\mu_n^{1/n}[C_u \cap \{(1-\theta)E_0 + \theta E_1\}] \geq (1-\theta)\mu_n^{1/n}(C_{u_0} \cap E_0) + \theta\mu_n^{1/n}(C_{u_1} \cap E_1) .$$

Applying the arithmetic-mean geometric-mean inequality we finally get

$$(3.16)\qquad \mu_n[C_u \cap \{(1-\theta)E_0 + \theta E_1\}] \geq [\mu_n(C_{u_0} \cap E_0)]^{1-\theta}[\mu_n(C_{u_1} \cap E_1)]^{\theta} .$$

Multiplying both the sides by

$$(3.17)\qquad \exp(-u) = \exp[(1-\theta)u_0]\exp[\theta u_1]$$

we get (3.13). The following lemma will now yield the theorem.

Lemma 3.3.1. Let g_0 and g_1 be non-negative (Borel-measurable) integrable functions on R^1 with non-empty supports given by S_0 and S_1, respectively. Let g be a non-negative Borel-measurable integrable function on R^1 such that for $0 < \theta < 1$ $x = (1-\theta)x_0 + \theta x_1$, $x_i \in S_i$

$$(3.18)\qquad g(x) \geq g_0^{1-\theta}(x_0)g_1^{\theta}(x_1) .$$

Then

$$(3.19)\qquad \int_{(1-\theta)S_0+\theta S_1} g(x)dx \geq [\int_{S_0} g_0(x)dx]^{1-\theta}[\int_{S_1} g_1(x)dx]^{\theta} .$$

Proof. First we shall assume that g_i's are bounded. Let c_i be the supremum of g_i. C_i's are assumed to be positive (excluding the trival case). Define

$$(3.20)\qquad A_i = \{x^* = (x,z) \in R^2 : g_i(x) > c_i z,\ z > 0\ ,\ x \in S_i\} ,$$

$i = 0,1$, and

$$(3.21)\qquad A = \{x^* = (x,z) \in R^2 : g(x) > z\, c_0^{1-\theta}c_1^{\theta},\ z > 0,\ x \in (1-\theta)S_0 + \theta S_1\} .$$

Let $A_i(z)$ and $A(z)$ be the z-sections of A_i and A, respectively. For $0 < z < 1$ both $A_0(z)$ and $A_1(z)$ are non-empty, and

$$(3.22)\qquad A(z) \supset (1-\theta)A_0(z) + \theta A_1(z) .$$

Moreover,

$$(3.23)\qquad \int_{-\infty}^{\infty} g_i(x)dx = c_i \int_0^1 \mu_1(A_i(z))dz$$

We may assume that the integrals in the left-hand side of (3.19) are positive, the result is trivial otherwise.

$$\int_{(1-\theta)S_0+\theta S_1} g(x)dx \geq c_0^{1-\theta}c_1^{\theta}\int_0^1 \mu_1(A(z))dz . \tag{3.24}$$

By the one-dimensional Brunn-Minkowski-Lusternik inequality

$$\mu_1(A(z)) \geq (1-\theta)\mu_1(A_0(z)) + \theta\, \mu_1(A_1(z)) , \tag{3.25}$$

for $0 < z < 1$. Now it follows that

$$\begin{aligned} \int_{(1-\theta)S_0+\theta S_1} g(x)dx &\geq c_0^{1-\theta}c_1^{\theta}[(1-\theta)c_0^{-1}\int_{-\infty}^{\infty} g_0(x)dx+\theta c_1^{-1}\int_{-\infty}^{\infty} g_1(x)dx] \\ &\geq [\int_{-\infty}^{\infty} g_0(x)dx]^{1-\theta}[\int_{-\infty}^{\infty} g_1(x)dx]^{\theta} . \end{aligned} \tag{3.26}$$

In the general case, define

$$g_{ik}(x) = \begin{cases} g_i(x), \text{ if } g_i(x) \leq k \\ k , \text{ if } g_i(x) > k \end{cases} . \tag{3.27}$$

Then $g_{ik}(x) \uparrow g_i(x)$ as $k \to \infty$. Now apply the above result to g_{ik}'s and appeal to the monotone convergence theorem.

__Theorem 3.4.__ Let f be a function on R^n satisfying the conditions in Theorem 3.3. Let E be a convex set in R^n , and for $\tau \in R^n$ define

$$h(\tau) = \int_{E+\tau} f(x)dx \tag{3.28}$$

Then h is a 0-unimodal function on R^n , i.e.

$$h[(1-\theta)\tau_0 + \theta\tau_1] \geq h^{1-\theta}(\tau_0)h^{\theta}(\tau_1) \tag{3.29}$$

for $0 < \theta < 1$, $\tau_i \in R^n$.

__Proof.__ Apply theorem 3.3 with $E_0 = E + \tau_0$, $E_1 = E + \tau_1$, and note that $(1-\theta)E_0 + \theta E_1 = E + [(1-\theta)\tau_0 + \theta\tau_1]$.

Corollary 3.4.1. Define h as in Theorem 3.4. Suppose

$$h(\tau_1) = \dots = h(\tau_m) = h(\tau) \tag{3.30}$$

for τ_i's and τ in R^n. Then

$$h\left(\sum_{i=1}^{m} \lambda_i \tau_i\right) \geq h(\tau) \tag{3.31}$$

for $0 \leq \lambda_i \leq 1$, $\sum_{i=1}^{m} \lambda_i = 1$.

4. Study on monotonicity in the general case.

For studying tests in Φ_G we shall reduce the problem further. Recall that $\tau_1^2, \dots, \tau_\ell^2$ are the ℓ largest characteristic roots of $\Sigma^{-1}MM'$. It is possible to write

$$\Sigma^{-\frac{1}{2}}M = Q\Delta(\tau)L', \tag{4.1}$$

where $Q: p \times p$ and $L: s \times s$ are orthogonal matrices, and

$$\Delta(\tau) = \left[\begin{array}{c|c} \Delta^*(\tau) & 0 \\ \hline 0 & 0 \end{array}\right], \quad \Delta^*(\tau) = \mathrm{diag}(\tau_1, \dots, \tau_\ell), \tag{4.2}$$

$$\tau = (\tau_1, \dots, \tau_\ell)'.$$

Define

$$A = Q'\Sigma^{-\frac{1}{2}}, \ U = AY_{(1)}L, \ V = AY_{(3)}. \tag{4.3}$$

Then the columns of U and V are independently distributed as $\mathfrak{n}_p(\cdot, I_p)$, and $\mathcal{E}U = \Delta(\tau)$, $\mathcal{E}V = 0$. Note that the nonzero characteristic roots of $(UU')(UU'+VV')^{-1}$ are the same as those of $S_0S_t^{-1}$. This shows that the power function of any test in Φ_G depends on Σ, M only through τ. We shall now write $S_0 = UU'$, $S_e = VV'$, $S_t = S_0 + S_e$.

For a non-randomized test φ, let A_φ be its acceptance region. We shall first consider acceptance regions in the space of U and V.

The power function of a test φ is

(4.4) $$\mathcal{E}_{M,\Sigma}\varphi(U,V) = P_{M,\Sigma}[(U,V) \notin A_\varphi] .$$

For $\varphi \in \Phi_G$ the power function of φ will be denoted by $\Pi(\tau;\varphi)$. Given τ_i^2's and the structure of Δ in (4.2) the diagonal elements of Δ in (4.2) are not uniquely defined. In particular, by choosing Q and L appropriately it is possible to write in (4.2) $\Delta = \Delta(D_e\tau)$, as well as, $\Delta = \Delta(\Gamma\tau)$, where D_e is an $\ell \times \ell$ diagonal matrix with diagonal elements as ± 1, and Γ is an $\ell \times \ell$ orthogonal permutation matrix, i.e. $\Gamma\tau = (\tau_{i_1}, \ldots, \tau_{i_\ell})'$ for some permutation $(i_1, \ldots, i_\ell)$ of $(1, \ldots, \ell)$. Hence for $\varphi \in \Phi_G$

(4.5) $$\Pi(\tau;\varphi) = \Pi(D_e\tau;\varphi) = \Pi(\Gamma\tau;\varphi)$$

for any such matrices D_e and Γ and for all $\tau \in R^\ell$.

Let U_i be the i^{th} column vector of U and $\widetilde{U}^{(i)}$ be the matrix U with U_i deleted. For a region A in (U,V) space, let $A(\widetilde{U}^{(i)},V)$ be the section of A in the U_i-space, i.e.

(4.6) $$A(u^{(i)},v) = \{u_i \in R^p : (u,v) \in A\} .$$

For any test $\varphi \in \Phi_G$ and all $u^{(i)}$ and v

(4.7) $$A_\varphi(u^{(i)},v) = -A_\varphi(u^{(i)},v) ,$$

and for all v

(4.8) $$A_{\bar{\varphi}}(v) = -A_\varphi(v) ,$$

where $A_\varphi(v)$ is the section of A_φ in the u-space.

Later we shall require A_φ to be a region in the space of (U,VV') , or in the space of $(U,UU' + VV')$. For that purpose we denote the acceptance region of φ as $\widetilde{A}_\varphi$ to mean that it is a region in $\mathfrak{M}_{p,s} \times \mathcal{S}_p^+$.

Next we shall introduce four subclasses of Φ_G as follows:

(1) $\Phi_G^{(1)}$ is the set of all $\varphi \in \Phi_G$ such that the acceptance region A_φ (in the space of U and V) is convex in the space of each column vector of U for each set of fixed values of V and of the other column vectors of U, i.e. for every i and all $\tilde{u}^{(i)}$ and v the set $A_\varphi(\tilde{u}^{(i)},v)$ is convex.

(2) $\Phi_G^{(2)}$ is the set of all $\varphi \in \Phi_G$ such that the acceptance region A_φ is convex in the space of U for each set of fixed value of V .

(3) $\Phi_G^{(3)}$ is the set of all $\varphi \in \Phi_G$ such that the acceptance region $\tilde{A}_\varphi$ (in the space of (U,VV')) is convex in U and VV' .

(4) $\Phi_G^{(4)}$ is the set of all $\varphi \in \Phi_G$ such that the acceptance region $\tilde{A}_\varphi$ (in the space of $(U, S_t = UU' + VV')$) is convex in U and S_t . Note that $\Phi_G^{(1)} \supset \Phi_G^{(2)} \supset \Phi_G^{(3)}$.

<u>Theorem 4.1.</u> For $\varphi \in \Phi_G^{(1)}$ the power function of φ given by $\Pi(\tau,\varphi)$ is a symmetric function in each τ_i and monotonically increases as each $|\tau_i|$ increases separately.

<u>Proof.</u> The first part of the theorem follows from (4.5). For $i = 1, \ldots, \ell$

$$(4.9)\qquad \mathcal{E}_\tau[1-\varphi(U,V)|U^{(i)} = u^{(i)}, V = v] = \int_{A_\varphi(u^{(i)},v)+\tau_i e_o} f(u_i)du_i \quad ,$$

where f is the p.d.f. corresponding the $n_p(0,I_p)$ and $\underset{\sim}{e}_i$ is the vector in R^p with 1 at the i^{th} position and the other components being 0.

Now we shall use Theorem 3.2. Note that the density function f is unimodal and centrally symmetric. $A_\varphi(u^{(i)},v)$ is convex and centrally symmetric. Specialize G in Theorem 3.2 to be the group of sign transformations on R^p. Note that the distribution of $U^{(i)}$ and V is free from τ_i. Hence

$$P[U_i \in A_\varphi(\tilde{u}^{(i)},v) + \lambda_i\tau_i e_i \mid \tilde{U}^{(i)} = \tilde{u}^{(i)}, V = \tilde{v}]$$

$$(4.10) \qquad = P[U_i \in A_\varphi(\tilde{u}^{(i)},v) + (1+\lambda_i)\tau_i e_i/2 - (1-\lambda_i)\tau_i e_i/2 \mid \tilde{U}_i = \tilde{u}_i, \quad V = v]$$

$$\geq P[U_i \in A\varphi(\tilde{u}^{(i)},v) + \tau_i e_i \mid \tilde{U}_i = \tilde{u}_i, V = v] ,$$

where $-1 \leq \lambda_i \leq 1$ and the conditional p.d.f. of U_i is taken as f. Taking expectation with respect to $\tilde{U}_i$ and V we find that $\Pi(\tau;\varphi)$ increases if τ_i is replaced by $\lambda_i\tau_i$, where $-1 \leq \lambda_i \leq 1$, holding the other components of τ fixed.

Since f is also 0-unimodal the result would also follow from Corollary 3.4.1.

In the above theorem we need only $n_e + s > p$.

<u>Corollary 4.1.1</u>. If $\varphi \in \Phi_G^{(2)}$ the power function of φ is a symmetric function in each τ_i and increases monotonically in each $|\tau_i|$.

<u>Proof</u>. Simply note that $\Phi_G^{(2)} \subset \Phi_G^{(1)}$.

Let H be the group of transformations acting on R^ℓ defined as follows. For $\tau \in R^\ell$, $h \in H$

$$(4.11) \qquad h\tau = (e_1\tau_{i_1}, \ldots, e_\ell\tau_{i_\ell}) ,$$

where $e_i = \pm 1$ and $(i_1, \ldots, i_\ell)$ is a permutation of $(1, \ldots, \ell)$.

<u>Theorem 4.2</u>. If $\varphi \in \Phi_G^{(3)}$, and $\tau \in R^\ell$

(4.12) $$\Pi(\tau^*;\varphi) \le \Pi(\tau;\varphi) ,$$

where τ^* is any point in the convex-hull of the H-orbit of τ , provided $n_e \ge p+1$.

<u>Proof</u>. The joint density p_0 of U and $S_e = VV'$ under H_0 is O-unimodal when $n_e \ge p+1$. For $h \in H$, $\tau \in R^\ell$

(4.13) $$\Pi(h\tau;\varphi) = \Pi(\tau;\varphi) .$$

For $h_i \in H$ and $0 \le \lambda_i \le 1$, $\sum_1^m \lambda_i = 1$

(4.14) $$\sum_{i=1}^m \lambda_i \Delta(h_i\tau) = \Delta(\sum_{i=1}^m \lambda_i h_i \tau) .$$

Moreover

(4.15) $$\begin{aligned} &P_\tau[(U,S_e) \in \tilde{A}_\varphi | H_1] \\ &= P[(U+\Delta(\tau), S_e) \in \tilde{A}_\varphi | H_0] . \end{aligned}$$

The theorem now follows from Corollary 3.4.1.

Theorem 3.4 also yields the following.

<u>Corollary 4.2.1</u>. If $\varphi \in \Phi_G^{(3)}$ the power function of φ given by $\Pi(\tau;\varphi)$ is a O-unimodal function of τ, provided $n_e \ge p+1$.

<u>Theorem 4.3</u>. If $\varphi \in \Phi_G^{(4)}$ the result in Theorem 4.2 holds provided $n_e \ge p+1$.

<u>Proof</u>. The joint density of U and S_t under H_0 is given by

(4.16) $$\begin{aligned} q(u,s_t) &= C \exp(-\tfrac{1}{2}tr(s_t))[\det(s_t-uu')]^{\frac{n_e-p-1}{2}} , \quad \text{if } s_t-uu' \in \mathcal{S}_p^+ \\ &= 0 \text{ otherwise} . \end{aligned}$$

The following facts show that q is a O-unimodal function when $n_e \ge p+1$

(i) If A_0 and A_1 are $p \times p$ positive-definite marices

(4.17) $$\det((1-\theta)A_0 + \theta A_1) \geq (\det A_0)^{1-\theta}(\det A_1)^{\theta} ,$$

for $0 < \theta < 1$.

(ii) Let $U^{(0)}$, $U^{(1)}$ be elements in $\mathfrak{m}_{p,s}$ and $U \doteq (1-\theta)U^{(0)} + \theta U^{(1)}$ for $0 < \theta < 1$. Then

(4.18) $$(1-\theta)U^{(0)}U^{(0)\prime} + \theta U^{(1)}U^{(1)\prime} = UU' + (1-\theta)\theta(U^{(0)}-U^{(1)})(U^{(0)}-U^{(1)})' .$$

(iii) If A_0 and A_1 are non-negative definite $p \times p$ matrices

(4.19) $$\det(A_0 + A_1) \geq \det(A_0) + \det(A_1) .$$

The rest of the proof is the same as that of Theorem 4.2.

Corollary 4.3.1. If $\varphi \in \Phi_G^{(4)}$ the power function of φ given by $\Pi(\tau;\varphi)$ is a 0-unimodal function of τ , provided $n_e \geq p+1$.

Next we shall study the four standard invariant tests given in Section 1.

Theorem 4.4. (a) The likelihood-ratio test is in $\Phi_G^{(1)}$.

(b) Roy's maximum root test is in $\Phi_G^{(3)}$.

(c) Lawley-Hotelling's trace test is in $\Phi_G^{(3)}$.

(d) Pillai's trace test is in $\Phi_G^{(4)}$.

(e) Pillai's trace test is in $\Phi_G^{(1)}$ if and only if the cut-off point $R_4 \leq \max(1, p-n_\ell)$.

Proof. (a) Let $W_i = (\bar{U}^{(i)}, V)$ then the acceptance region of the likelihood-ratio test can easily be expressed as

(4.20) $$1 + U_i'(W_iW_i')^{-1}U_i \leq (\det VV')/k \det(W_iW_i')$$

which is clearly convex in U_i for fixed W_i .

(b) Note that

$$\max \operatorname{ch}[(UU')S_e^{-1}] \le k_2$$

$$= \bigcap_{a \in R^p} [(U,S_e): a'UU'a \le k_2 a'S_e a] \ . \tag{4.21}$$

It follows from (4.18) that the region $a'UU'a \le k_2 a'S_e a$ is convex in (U,S_e) .

(c) For a matrix $B \in \mathfrak{M}_{p,s}$

$$\operatorname{tr}(S_e^{\frac{1}{2}}B)'(S_e^{-\frac{1}{2}}U) \le [\operatorname{tr}(B'S_eB)\ \operatorname{tr}(U'S_e^{-1}U)]^{\frac{1}{2}}$$

$$\le (\tfrac{1}{2})\ \operatorname{tr}(B'S_eB + U'S_e^{-1}U) \ . \tag{4.22}$$

Hence

$$\operatorname{tr}(B'U) - \tfrac{1}{2}\operatorname{tr}(B'S_eB) \le \tfrac{1}{2}\ \operatorname{tr}U'S_e^{-1}U \ , \tag{4.23}$$

the equality is attained when $B = S_e^{-1}U$. Hence the region in (U,S_e) given by $\operatorname{tr}(UU')S_e^{-1} \le k_3$ is the intersection of the regions

$$\operatorname{tr}(B'U) - \tfrac{1}{2}\operatorname{tr}(B'S_eB) \le \tfrac{1}{2}k_3 \tag{4.24}$$

for $B \in \mathfrak{M}_{p,s}$. However, each such region (4.24) is convex in (U,S_e) .

(d) The proof is the same as in (c) .

(e) The proof of this result is rather involved and we refer to [29]. Note however that tables for k_4 are partially available and even then they were obtained when $n_e \ge p$.

Examples of other tests in $\Phi_G^{(i)}$ $(i = 1,2,3,4)$ are given in [6], [27], [17], [36], [15]. A step-down test of H_0 vs. H_1 is given in [32]; however, this test is not in Φ_G . This test can easily be shown to be unbiased since it is given in terms of F tests. Only partial results are known for the monotonicity property of this test; see [7] and [10].

For the case $p = 1$ the power function of the F-test increases monotonically in n_e and decreases in s when the other parameters are held fixed. For $s = 1$, the power of the Hotelling's T^2-test increases if n_e increases, or if p decreases when the other parameters are held fixed. The proofs of these two results are given in [8]. Similar results for the general case are only known in very special situations; see [10], [9].

5. General MANOVA models.

The general MANOVA model introduced by Potthoff and Roy [30] may be described as follows: Let $X: p \times n$ be a random matrix such that its column vectors are independently distributed as $N_p(\cdot,\Sigma)$ with an unknown positive-definite matrix Σ; moreover $\mathcal{E}X' = A_1 \Theta A_2$, where $A_1: n \times m$ is a known matrix of rank r, $A_2: q \times p$ is a known matrix of rank q, and $\Theta: m \times q$ is a matrix of unknown parameters. The problem is to test H_0: $A_3 \Theta A_4 = 0$ against H_1: $A_3 \Theta A_4 \neq 0$, where $A_3 \Theta A_4$ is bilinearly estimable, and $A_3: s \times m$ and $A_4: q \times v$ are known matrices of ranks s and v, respectively. This problem can be reduced to the following canonical form: Let

$$(5.1) \qquad Y = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{bmatrix} \begin{matrix} q-v \\ v \\ p-q \end{matrix}$$
$$\qquad\qquad \begin{matrix} s & r-s & n-r \end{matrix}$$

be a random matrix such that its column vectors are independently distributed as $N_p(\cdot,\Sigma)$, and

$$(5.2) \qquad \mathcal{E}Y = \begin{bmatrix} M_{11} & M_{12} & 0 \\ M_{21} & M_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} .$$

The problem is to test H_0: $M_{21} = 0$ against $M_{21} \neq 0$.

Let us partition Σ as in the above.

$$(5.3)\qquad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix}$$

A class of tests invariant under a certain group of transformations which keeps the problem invariant is obtained by Gleser and Olkin [18]. However, this problem is generally viewed in the conditional set-up described below.

The column vectors of

$$(5.4)\qquad \tilde{Y} = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \end{bmatrix}$$

are conditionally independently distributed as $N_q(\cdot,\tilde{\Sigma})$, given Y_{31}, Y_{32} and Y_{33}; $\tilde{\Sigma}$ is the covariance matrix of the first q components given the last $p-q$ components derived from Σ. The conditional expectation of $\tilde{Y}$ is

$$(5.5)\qquad \mathcal{E}^*(\tilde{Y}) = \begin{bmatrix} M_{11} & M_{12} & 0 \\ M_{21} & M_{22} & 0 \end{bmatrix} + \beta[Y_{31} \quad Y_{32} \quad Y_{33}] ,$$

where β is the matrix of regression coefficients. In this conditional set-up the s.p. matrices due to error and the hypothesis H_0 are respectively defined by (assuming $n-r \geq p-q$)

$$(5.6)\qquad S_e = Y_{23}Y_{23}' - Y_{23}Y_{33}'(Y_{33}Y_{33}')^{-1}Y_{33}Y_{23}'$$

$$(5.7)\qquad S_0 = \hat{M}_{21}(I_s + Y_{31}'(Y_{33}Y_{33}')^{-1}Y_{31})^{-1}\hat{M}_{21}',$$

where

$$(5.8)\qquad \hat{M}_{21} = Y_{21} - Y_{23}Y_{33}'(Y_{33}Y_{33}')^{-1}Y_{31} .$$

In the conditional situations S_e and S_0 are independently distributed as the Wishart distributions $\mathfrak{W}_v(n-r-p+q,\ \Sigma_{22\cdot3})$ and $\mathfrak{W}_v(s,\Sigma_{22\cdot3};\tilde{\Delta})$, respectively, where $\Sigma_{22\cdot3}$ is the covariance matrix of the second set (of v) components given the third set of $(p-q)$ components, and

(5.9) $\tilde{\Delta} = M_{21}(I_s + Y_{31}'(Y_{33}Y_{33}')^{-1}Y_{31})^{-1}M_{21}'$.

As in the MANOVA one might consider those tests which depend only on the characteristic roots of $S_0S_e^{-1}$. In particular, the acceptance region of the likelihood-ratio test is given by $|S_e|/|S_0 + S_e| \geq k$.

The column vectors of $(Y_{31}\ Y_{33})$ are independently distributed as $N_{p-q}(0,\Sigma_{33})$. It is clear that the distribution of $Y_{31}'(Y_{33}Y_{33}')^{-1}Y_{31}$ does not depend on Σ_{33} and we shall assume it to be I_{p-q} . Also for considering the distribution of the roots of $S_0S_e^{-1}$ we might take $\Sigma_{22\cdot3} = I_v$ and replace M_{21} by $\Sigma_{22\cdot3}^{-\frac{1}{2}}M_{21}$. As in the MANOVA case, we can replace $\Sigma_{22\cdot3}^{-\frac{1}{2}}M_{21}$ by a matrix Δ: v x s such that

(5.10) $$\Delta = \left[\begin{array}{c|c} \mathrm{diag}(\tau_1, \ldots, \tau_\ell) & 0 \\ \hline 0 & 0 \end{array}\right]$$

where $\ell=\min(v,s)$ and $\tau_i^2\ (\tau_i > 0)$ are the characteristic roots of $M_{21}'\Sigma_{22\cdot3}^{-1}M_{21}$. This discussion leads us to take $\tilde{\Delta}$ as

(5.11) $\tilde{\Delta} = \Delta(I_s + Y_{31}'(Y_{33}Y_{33}')^{-1}Y_{31})^{-1}\Delta'$.

Arguing as in Anderson and Das Gupta [3] we see that the characteristic roots of $\tilde{\Delta}$ increase if any τ_i is increased. Thus all the results in the MANOVA case can be applied now.

6. Bibliographical Notes.

On Section 1. For a general discussion of MANOVA see Anderson [2], Roy [33], and Lehmann [23].

On Section 2. See Roy [33].

On Section 3. A proof of Theorem 3.1 is given in Bonneson and Fenchel [4]. For Lusternik's generalization of Theorem 3.1 see Hadwiger and Ohman [19] or Henstock and Macbeath [20].

Theorem 3.2 was proved by Anderson [1] when G is the group of sign transformations. Essentially the same proof also holds for any G defined in Theorem 3.2; the general statement is due to Mudholkar [28]. For further generalizations of this theorem see Das Gupta [11].

Theorem 3.3 was proved by Prekopa [31] and Leindler [24] (for n = 1); however, their proofs are quite obscure and somehwat incomplete. The present proof uses essentially the ideas given by Henstock and Macbeath [20]; see Das Gupta [13] for more general results. Theorem 3.4 was proved by Ibragimov [21] and Schoenberg [35] when n = 1; the general case was proved by Davidovic, Korenbljum and Hacet [14]. For a discussion of these results see Das Gupta [13].

<u>On Section 4</u>. Theorem 4.1 is due to Das Gupta, Anderson and Mudholkar [6] where the monotonicity property of the power functions of tests (a), (b), and (c) are established. Roy and Mikhail [34] also proved the monotonicity property of the maximum root test. Srivastava [37] derived the result for tests (a)-(c) although his proofs are incomplete. Theorem 4.2 and its present proof are due to Das Gupta [12]; an alternative proof using Theorem 3.2 is given by Eaton and Perlman [15].

<u>On Section 5</u>. See Fujikoshi [16] and Khatri [22].

References

[1] Anderson, T. W. (1955) The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proc. Amer. Math. Soc. 6, 170-176.

[2] Anderson, T. W. (1958) An Introduction to Multivariate Statistical Analysis. Wiley, New York.

[3] Anderson, T. W. and Das Gupta, S. (1964) Monotonicity of the power functions of some tests of independence between two sets of variates. Ann. Math. Statist. 35, 206-208.

[4] Bonneson, T. and Fenchel, W. (1948) Konvexe Körper. Chelsea, New York.

[5] Brunn, H. (1887) Über Ovale und Eiflachen, Inaugural dissertation, München.

[6] Das Gupta, S., Anderson, T. W., and Mudholkar, G. S. (1964) Monotonicity of the power functions of some tests of the multivaraite linear hypothesis. Ann. Math. Statist., 35, 200-205.

[7] Das Gupta, S. (1970) Step-down multiple-decision rules. Essays in Probability and Statistics., Univ. of North Carolina Press, Chapel Hill.

[8] Das Gupta, S. (1972) Noncentral matrix-variate beta distribution and Wilks' U-distribution. Sankhyā, Series A, 34, 357-362.

[9] Das Gupta, S. and Perlman, M. D. (1973) On the power of Wilks' U-test for MANOVA. Jour. Multiv. Anal., 3, 220-225.

[10] Das Gupta, S. and Perlman, M. D. (1974) Power of the noncentral F-test: Effect of additional variates on Hotelling's T^2-test Jour. Amer. Statist. Assoc., 69, 174-180.

[11] Das Gupta, S. (1976) A generalization of Anderson's theorem on unimodal functions. Proc. Amer. Math. Soc., 60, 85-91.

[12] Das Gupta, S. (1977) s-unimodal functions: related inequalities and statistical applications. Sankhyā, Series B.

[13] Das Gupta, S. (1978) Brunn-Minkowski inequality and its aftermath. Tech. Report #310, School of Statistics, Univ. of Minnesota.

[14] Davidovic, Ju. S., Korenbljum, B. I. and Hacet, B. I. (1962) A property of logarithmically concave functions. Soviet. Math. Dokl., 10, #2, 477-480.

[15] Eaton, M. and Perlman, M. D. (1974) A monotonicity property of the power functions of some invariant tests. Ann. Statist., 2, 1022-1028.

[16] Fujikoshi, Y. (1973) Monotonicity of the power functions of some tests in general MANOVA models. Ann. Statist., 1, 388-391.

[17] Ghosh, M. N. (1964) On the admissibility of some tests of MANOVA. Ann. Math. Statist., 35, 789-794.

[18] Gleser, L. and Olkin, I. (1970). Linear model in multivariate analysis. Essays in Probability and Statistics. Univ. of North Carolina Press, Chapel Hill.

[19] Hadwiger, H. and Ohman, D. (1956) Brunn-Minkowskischer Satz und Isoperimetrie. Math. Zeit., 66, 1-8.

[20] Henstock, R. and Macbeath, A. M. (1953) On the measure of sum sets (I). The theorem of Brunn, Minkowski and Lusternik. Proc. Lond. Math. Soc., 3, 182-194.

[21] Ibragimov, I. A. (1956) On the composition of unimodal distributions. Theor. Prob. Appl. (Translation) 1, 255-266.

[22] Khatri, C. G. (1966) A note on MANOVA model applied to problems in growth curves. Ann. Inst. Math. Statist., 18, 75-86.

[23] Lehmann, E. L. (1959) Testing Statistical Hypotheses, Wiley, New York.

[24] Leindler, L. (1972) On a certain converse of Holder's inequality II. Acta Scient. Mat., 33, 217-223.

[25] Lusternik, L. (1925) Die Brunn-Minkowskische Ungleischung fur Beliebge Nessabare Mengen. Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS, Vol. III, (viii) 55-58.

[26] Minkowski, H. (1910) Geometrie der Zahlen, Keipzig and Berlin.

[27] Mudholkar, G. S. (1965) A class of tests with monotone power functions for two problems in multivaraite statistical analysis Ann. Math. Statist., 36, 1794-1801.

[28] Mudholkar, G. S. (1966) The integral of an invariant unimodal function over an invariant convex set--an inequality and applications. Proc. Amer. Math. Soc., 17, 1327-1333.

[29] Perlman, M. D. (1974) Monotonicity of the power function of Pillai's trace test. Jour. Multiv. Anal., 4, 22-30.

[30] Potthoff, R. F. and Roy, S. N. (1964) A generalized multivariate analysis of variance model useful especially for growth curve problems. Biometrika, 51, 313-326.

[31] Prekopa, A. (1973) On logarithmic concave measures and functions. Acta Scient. Mat., 34, 335-343.

[32] Roy, J. (1958) Step-down procedure in multivatiate analysis. Ann. Math. Statist., 29, 1177-1187.

[33] Roy, S. N. (1957). Some aspects of Multivariate Analysis, Wiley, New York.

[34] Roy, S. N. and Mikhail, W. F. (1961) On the monotonic character of the power functions of two multivariate tests. Ann. Math. Statist., 32, 1145-1151.

[35] Schoenberg, I. J. (1951). On Polya frequency functions I: the totally positive functions and their Laplace transforms.

[36] Schwartz, R. E. (1967) Admissible tests in multivariate analysis of variance, Ann. Math. Statist., 38, 698-710.

[37] Srivastava, J. N. (1964) On the monotonicity property of the three multivariate tests for Multivariate Analysis of Variance. J. Roy. Statist. Ser. B., 26, 77-81.