UNIVERSITY OF NORTH CAROLINA

Department of Statistics

Chapel Hill, N. C.

ON THE MONOTONICITY PROPERTY OF THE THREE MAIN TESTS FOR MULTIVARIATE ANALYSIS OF VARIANCE

by

J. N. Srivastava

December 1962

Contract No. AF 49(638)-213

This paper presents a unified proof of the monotonicity property of each of the three main tests used in multi-response analysis of variance, and also some other interesting side-results.

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

Institute of Statistics
Mimeo Series No. 342

ON THE MONOTONICITY PROPERTY OF THE THREE MAIN TESTS FOR MULTIVARIATE ANALYSIS OF VARIANCE[1]

by

J. N. Srivastava
University of North Carolina

1. This paper presents a unified and relatively simple proof of the monotonicity property of the three well known tests for the multivariate analysis of variance (MANOVA) viz, (i) Hotelling - Lawley trace criterion (ii) Wilk's likelihood ratio criterion and (iii) Roy's largest root criterion. For the last one, the property has already been proved by Roy and Mikhail [3].

2. To provide a background, we shall start with the multivariate linear model, present the three criteria in explicit form and discuss the general nature of the power function. To this end we shall state several results, for the proof of which the reader is referred to [2].

The usual multivariate linear model is:

$$(1)\qquad \underset{n\times p}{E(Y)} = \underset{n\times m}{A}\ \underset{m\times p}{\eta}, \qquad \mathrm{Var}(Y) = I_n \otimes \Sigma ,$$

where Σ is an unknown p.d. matrix, $\otimes$ denotes Kronecker or direct product of matrices, and

$$\underset{n\times p}{Y} = (\underline{y}_1, \ \underline{y}_2, \ldots, \ \underline{y}_p) = \begin{bmatrix} \underline{y}_{(1)} \\ \vdots \\ \underline{y}_{(n)} \end{bmatrix} \begin{matrix} 1\times p \\ \\ 1\times p \end{matrix}$$

is a matrix of observations, in which the j-th column $\underline{y}_j$ corresponds to the j-th variate, and the i-th row to the i-th experimental unit. As indicated in (1), $\underline{y}_{(r)}$ and $\underline{y}_{(r')}$ are assumed to be independent if $r \neq r'$. Further each

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

vector $\underline{y}_{(r)}$ $(r = 1, 2, \ldots, n)$ is assumed to have a multivariate normal distribution. In fact

(2) $\text{Prob}(Y) = \text{const.} \exp \left[-\frac{1}{2} \text{ tr } \Sigma^{-1} (Y' - \eta' A')(Y - A\eta) \right] dY$.

where dY stands for $\prod_{i,j} dy_{ij}$.

Let $A_1(n \times r)$ be a basis of A, where $r = \text{Rank}(A)$, and let

(3) $$\underset{n\times r}{A_1} = \underset{n\times r}{L'} \; \underset{r\times r}{T} , \qquad \underset{n\times m}{A} = (\underset{n\times r}{A_1} \;\vdots\; \underset{n\times \overline{m-r}}{A_2})$$

where T is triangular, and L is orthonormal. Also let $L_1(n \times \overline{n-r})$ be an arbitrary orthogonal completion of L'. Consider the (testable) null hypothesis

(4) $$\underset{s\times m}{C} \; \underset{m\times p}{\eta} = 0_{mp} \iff C_1 \eta = 0$$

where $C_1(s' \times m)$ is the matrix obtained by taking any independent s' rows of C, where $s' = \text{Rank } C$. Now in the model (1), the partitioning (3) of A induces a partitioning of η as below

(5) $$\underset{m\times p}{\eta} = \begin{bmatrix} \eta_1 \\ \\ \eta_2 \end{bmatrix} \begin{matrix} r\times p \\ \\ \overline{m-r}\times p \end{matrix}$$

This is turn induces, through (4), a partitioning of

(6) $$\underset{s'\times m}{C_1} = \left[\underset{s'\times r}{C_{11}} \quad \underset{s'\times \overline{m-r}}{C_{12}} \right] .$$

Make the factorization $$\underset{s'\times r}{C_{11}} \; \underset{r\times r}{T'^{-1}} = \underset{s'\times s'}{\Gamma} \; \underset{s'\times r}{N} ,$$

where Γ is triangular and N is orthonormal, and let N_1 $(r-s', r)$ be an orthogonal completion of N. Consider first the transformation from Y to W_1, W_2 given by

$$\underset{r\times p}{W_1} = L \; Y , \qquad \underset{\overline{n-r}\times p}{W_2} = L_1 \; Y .$$

One can interpret this by observing (as can be mathematically proved) that W_1 relates to the set of best linear unbiased estimates of estimable linear functions, and W_2 relates to the error. Next make the decomposition:

$$(8) \qquad \underset{s'\times p}{Z_1} = \underset{s'\times r}{N} \; \underset{r\times p}{W_1} \;, \qquad \underset{\overline{r-s'}\times p}{Z_2} = \underset{\overline{r-s'}\times r}{N_1} \; \underset{r\times p}{W_1} \;.$$

It can be shown that Z_1 relates to the set of linear functions which are being tested under H_o, and Z_2 corresponds to the rest of the estimable linear functions. Indeed it can be easily seen [2] that the sum of products matrix due to the hypothesis H_o is $Y'S_H Y$ where

$$(9) \qquad Y' S_H Y = Z_1' Z_1 \;,$$

and the sum of products matrix due to error is

$$(10) \qquad Y' S_E Y = W_2' W_2 \;.$$

Since the three tests of MANOVA are in terms of the roots of $(Y' S_H Y)(Y' S_E Y)^{-1}$ i.e., of $(Z_1' Z_1)(W_2' W_2)^{-1}$, we restrict ourselves to Z_1 and W_2. Their joint distribution is given by

$$(11) \quad \text{Prob}\,(Z_1, W_2) = \text{const.}\; \exp\; -\tfrac{1}{2}\,\text{tr}\,\Sigma^{-1}\left[(Z_1' - \Theta_1')(Z_1 - \Theta_1) + W_2'W_2\right]\; dZ_1 dW_2,$$

where

$$(12) \qquad \Theta_1 = \Gamma^{-1}\, C\, \eta \;.$$

The null hypothesis H_o reduces to

$$(13) \qquad H_o:\; \Theta_1 = 0_{s'p}$$

The following results will now be needed. For its proof, one may see [2].

Theorem 2.1. Let c_j $(j - 1, 2, \ldots, p)$ be the roots of $(Z_1' Z_1)(W_2' W_2)^{-1}$, with $c_1 \geq c_2 \geq \ldots \geq c_p$. Then the joint distribution of these roots involves as parameters just the following quantities:

(14) $$p,\ s',\ n-r,\ \gamma_1,\ \gamma_2,\ \ldots,\ \gamma_p\ ,$$

where $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_p$ are the p unknown latent roots of the matrix $(\Theta_1' \ \Theta_1)\Sigma^{-1}$.

Theorem 2.2. There exists a matrix V_1 $(p \times \overline{n-r})$ whose elements are functions of the elements of Σ and of W_2, and a matrix $V_2(p \times s')$ whose elements are functions of the elements of Σ and of Z_1, such that

(15) $$\text{ch}\,[(Z_1' Z_1)(W_2' W_2)^{-1}] = \text{ch}\,[(V_2 V_2')(V_1 V_1')^{-1}]\ ,$$

so that c_j $(j = 1, 2, \ldots, p)$ are the p roots of $(V_2 V_2')(V_1 V_1')^{-1}$ also. Also the distribution of V_1 and V_2 is given by

$$(2\pi)^{-\frac{p}{2}(n-r+s')} \exp\left[-\frac{1}{2}\quad \text{tr}(V_1V_1' + V_2V_2') + \sum_{i=1}^{t} \gamma_i - 2\sum_{i=1}^{t} (V_2)_{ii}\gamma_i^{\frac{1}{2}}\right]$$

$$dV_1 \ \ dV_2\ ,$$

where $(V_2)_{ii}$ represents the (i,i) element of V_2, and $t = \text{Rank}\,(\Theta_1' \Theta_1)$.

(It is clear that the truth of Theorem 2.2 implies the proof of Theorem 2.1).

We now come to the three tests for the hypothesis $H_0 : \Theta_1 = 0_{s'p}$. These are

(i) Hotelling-Lawley trace criterion, say ψ_1 where

(16) $$\psi_1 = \sum_{j=1}^{p} c_j$$

(ii) Wilks' likelihood ratio criterion:

$$\psi_2 = (1 + c_1)(1 + c_2) \dots (1 + c_p) \tag{17}$$

(iii) Roy's largest root criterion:

$$\psi_3 = c_1 \quad . \tag{18}$$

The distribution of ψ_1 does not involve as an unknown parameter just $\sum_{j=1}^{p} \gamma_j$, but all the roots $\gamma_1, \gamma_2, \dots, \gamma_p$. Similarly the distribution of ψ_2 and ψ_3 does not involve just $\prod_{j=1}^{p} (1 + \gamma_j)$ or γ_1, but all the roots separately. The problem before us is to prove that the power of each of these three criteria is a monotonically increasing function of each of the p parameters $\gamma_1, \gamma_2, \dots, \gamma_p$.

Write

$$\underset{p \times s'}{V_2} = \begin{bmatrix} v_{11} & \dots\dots & v_{1s'} \\ \\ v_{p1} & \dots\dots & v_{ps'} \end{bmatrix} \qquad \underset{p \times n-r}{V_1} = \begin{bmatrix} v'_{11} & \dots\dots & v'_{1,n-r} \\ \\ v'_{p1} & \dots\dots & v'_{p,n-r} \end{bmatrix} \tag{19}$$

Let ψ be any of the three criteria, and β_ψ the corresponding power. Then the acceptance region is of the form

$$\psi \leq \mu \quad ,$$

where μ is a constant, and consequently we have by (16):

$$(20)\quad 1-\beta_\psi = \int_{\psi \leq \mu} (2\pi)^{-\frac{p}{2}(n-r+s')} \exp\left[-\frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{n-r} v'^2_{ij} + \right.$$

$$\left. + \sum_{i=1}^{p} \sum_{j=1}^{s'} v^2_{ij} + \sum_{i=1}^{p} \gamma_i - 2 \sum_{i=1}^{t} v_{ii} \sqrt{\gamma_i} \right] d\,V_1 \; dV_2$$

$$= \int_{\psi \leq \mu} (2\pi)^{-\frac{p}{2}(n-r+s')} \exp\left[-\frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{n-r} v'^2_{ij} + \sum_{i=1}^{p} \sum_{j=2}^{s'} v^2_{ij} + \right.$$

$$\left. + \sum_{i=1}^{p} (v_{ii} - \sqrt{\gamma_i}\,)^2 \right] dV_1 \; dV_2 \;,$$

where we define $\sqrt{\gamma_i} = 0, \;\; p \geq i > t$. Now write

$$(21)\qquad \underset{p\times s'}{V_2} = \left[\underset{p\times 1}{\underline{v}_1} \quad \underset{p\times s'-1}{V_{21}} \right] ,$$

so that $\underline{v}_1$ is the first column of V_2. Then the last integral can be written

$$(22)\quad \int_{V_1, V_{21}} \left\{ \int_{R_{1\psi}} (2\pi)^{-\frac{p}{2}} \exp \; -\frac{1}{2} \left[\sum_{i=2}^{p} v^2_{i1} + (v_{11} - \sqrt{\gamma_1}\,)^2 \right] \right.$$

$$\left. dv_{11}\; dv_{21} \ldots dv_{p1} \right\} \times \; \phi_o \; dV_1 \;\; dV_{21} \;\; ,$$

where ϕ_o represents the part of the integrand (say at (20)) which is not inside the big curly brackets, and where $R_{1\psi}$ represents the section of the region

$$\psi \leq \mu$$

obtained by taking any fixed value of V_1 and V_{21}. For any given ψ, the proof of the monotonicity with respect to $\sqrt{\gamma_1}$ will be completed, if we can show that as $\sqrt{\gamma_1}$ increases the quantity in the big curly brackets decreases, for any fixed value of V_1 and V_{21}. From the symmetry of the proof it will then follow that β_ψ is an increasing function of $\sqrt{\gamma_i}$ or γ_i for $i = 1, 2, \ldots p$.

3. Some simple general results on monotonicity.

Lemma 3.1. Let $\psi(x_1, x_2, \ldots, x_p)$ be a positive, real-valued monotonically decreasing function of $x_1, x_2, \ldots, x_p$ (separately) in a rectangle R in the p-dimensional space of $x_1, x_2, \ldots, x_p$. Let $\Omega \subset R$, and suppose (for any i) that $\Omega_i(\theta_i) \subset R$ is the region Ω shifted by a distance θ_i (> 0) along the axis X_i. Then for all i :

$$(23) \quad \int_{\Omega} \psi(x_1, x_2, \ldots, x_p)\, d\,\underline{x} > \int_{\Omega_i(\theta_i)} \psi(x_1, x_2, \ldots, x_p)\, d\,\underline{x}$$

Proof: Take for example $i = 1$. We have

$$\int_{\Omega_1(\theta_1)} \psi(x_1, x_2, \ldots, x_p)\, d\,x_1\, dx_2 \ldots dx_p$$

$$= \int_{\Omega} \psi(x_1 + \theta_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_p$$

$$< \int_{\Omega} \psi(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_p \; ,$$

since ψ is a decreasing function of θ_1 .

Lemma 3.2.

In the above, suppose further that

$$(24) \qquad \psi(x_1, x_2, \ldots, x_p) = \psi(g_1x_1, g_2x_2, \ldots, g_px_p) \; ,$$

where $g_i = +1$ or -1, $i = 1, 2, \ldots p$.

Suppose also that Ω ($\subset R$) is such that

(25) $$(x_1, x_2, \ldots, x_p) \in \Omega \implies (c_1x_1, c_2x_2, \ldots, c_px_p) \in \Omega ,$$

for $-1 \leq c_i \leq 1$, $i = 1, 2, \ldots, p$. Then again, for any i,

(26) $$\int_{\Omega} \psi(x_1, x_2, \ldots, x_p)\, d\underline{x} > \int_{\Omega_i(\theta_i)} \psi(x_1, x_2, \ldots, x_p)\, d\underline{x} .$$

Proof: Consider for example the case $i = 1$. We have

$$\int_{\Omega_1(\theta_1)} \psi(x_1, x_2, \ldots, x_p)\, d\underline{x}$$

$$= \int_{x_2, x_3, \ldots, x_p} \left\{ \int_{[x_1 | \underline{x} \in \Omega_1(\theta_1)]} \psi(x_1, x_2, \ldots, x_p)\, dx_1 \right\} dx_2 \ldots dx_p .$$

$$= \int_{x_2, x_3, \ldots, x_p} \left\{ \int_{[x_1 | \underline{x} \in \Omega]} \psi(x_1 + \theta_1, x_2, \ldots, x_p)\, dx_1 \right\} dx_2 \ldots dx_p .$$

Now take a fixed value of $x_2, x_3, \ldots, x_p$ for which there exists an x_1 such that $\underline{x} \in \Omega$. From (25), it follows that the range of possible values of x_1 in Ω is an interval $(-R_1, R_1)$, where $R_1 \geq 0$, and $R_1 = R_1(x_2, x_3, \ldots, x_p)$ depends on the chosen value of $x_2, x_3, \ldots, x_p$. Hence the last integral can be written as

$$\int\limits_{x_2, x_3, \ldots, x_p} \left\{ \int_{-R_1(x_2, x_3, \ldots, x_p)}^{R_1(x_2, x_3, \ldots, x_p)} \psi(x_1 + \theta_1, x_2, \ldots, x_p)\, dx_1 \right\} dx_2\, dx_3 \ldots dx_p$$

$$= \int\limits_{x_2, x_3, \ldots, x_p} \left\{ \int_{-R_1 + \theta_1}^{R_1 + \theta_1} \psi(x_1, x_2, \ldots, x_p) dx_1 \right\} dx_2 \ldots dx_p \ .$$

However we can write

$$(27) \quad \int_{-R_1 + \theta_1}^{R_1 + \theta_1} \psi(x_1, x_2, \ldots, x_p)\, dx_1$$

$$= \int_{-R_1}^{R_1} \psi(\underline{x})\, dx_1 - \int_{-R_1}^{-R_1 + \theta_1} \psi(\underline{x})\, dx_1 + \int_{R_1}^{R_1 + \theta_1} \psi(\underline{x})\, dx_1$$

But
$$\int_{R_1}^{R_1 + \theta_1} \psi(\underline{x})\, dx_1 - \int_{-R_1}^{-R_1 + \theta_1} \psi(\underline{x})\, dx_1$$

$$= \int_{R_1}^{R_1 + \theta_1} \psi(\underline{x})\, dx_1 - \int_{R_1 - \theta_1}^{R_1} \psi(\underline{x})\, d\underline{x} \ , \quad \text{using (24)} \ .$$

This last expression is negative for all real θ_1, since $\psi(x_1, x_2, \ldots, x_p)$ is a decreasing function of x_1 for $x_1 \geq 0$. Hence the expression (27) is less than

$$\int\limits_{x_2, x_3, \ldots, x_p} \left\{ \int_{-R_1}^{R_1} \psi(\underline{x})\, dx_1 \right\} dx_2 \ldots dx_p$$

$$= \int_{\Omega} \psi(x_1, x_2, \ldots, x_p)\, d\underline{x} \ ,$$

which completes the proof.

4. The monotonicity property of the trace and the likelihood ratio criteria.

Theorem 4.1. The power of the trace criterion is a monotonically increasing function of $|\gamma_i|$ for each $i = 1, 2, \ldots, p$.

Proof: We have

$$\psi_2 = c_1 + c_2 + \ldots + c_p = \mathrm{tr}\, (V_2 V_2')(V_1 V_1')^{-1} .$$

Let

$$(28) \qquad (V_1 V_1')^{-1} = V' \; V ,$$

where V (pxp) is triangular. Then

$$\begin{aligned} \psi_2 &= \mathrm{tr}\, (V_2 V_2')\, (V' V) \\ &= \mathrm{tr}\, V\, (\underline{v}_1 \;\; V_{21}) \begin{pmatrix} \underline{v}_1' \\ V_{21}' \end{pmatrix} V' \\ &= \mathrm{tr}\, V\, (\underline{v}_1\, \underline{v}_1')\, V' + \mathrm{tr}\, V\, V_{21}\, V_{21}'\, V' . \\ &= \mathrm{tr}\, (\underline{v}_1'\, V'\, V\, \underline{v}_1) + \mathrm{tr}\, V\, V_{21}\, V_{21}'\, V' \\ &= \underline{v}_1'\; V'\, V\; \underline{v}_1 + \mathrm{tr}\, V\, V_{21}\, V_{21}'\, V' . \end{aligned}$$

Hence the region $\psi_2 \leq \mu$ can be written

$$(29) \qquad \underline{v}_1'\, (V_1 V_1')^{-1}\, \underline{v}_1 \leq \mu - \mathrm{tr}\, (V_{21} V_{21}')\, (V_1 V_1')^{-1} .$$

Since $V_1 V_1'$ is clearly positive definite, the region (29) is the interior of an ellipsoid (centered at the origin) in the p-dimensional space of $v_{11}, v_{12}, \ldots, v_{1p}$. Also then there exists an orthogonal matrix P^* such that

$$P^{*\prime}\, (V_1 V_1')^{-1}\, P^* = D_1 ,$$

where D_1 is the diagonal matrix of the roots of $(V_1 V_1')^{-1}$.

Now in (22) consider the integral

$$(\ 0)\qquad \int_{R_{1\psi_1}} (2\pi)^{-\frac{p}{2}} \exp\left\{ -\frac{1}{2}\left[\sum_{i=2}^{p} v_{i1}^2 + (v_{11} - \sqrt{\gamma_1}\,)^2 \right] \right\} d\,\underline{v}_1 \quad ,$$

and make the transformation

$$\underline{v}_1 = P^* \underline{u}$$

to get

$$(\ 1)\qquad \int_{\underline{u}' D_1 \underline{u} \le \mu - \mathrm{tr}(V_{21}V_{21}')(V_1V_1')^{-1}} (2\pi)^{-\frac{p}{2}} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{p} (u_i - p_{i1}^* \sqrt{\gamma_1})^2 \right\} d\underline{u} \quad ,$$

where p_{i1}^* form the first column of P^*, and $\underline{u} = (u_1, \ldots, u_p)$. If now in lemma 3.2, we take

$$: \ \underline{u}' D_1 \underline{u} \le \text{const.}$$

$$\theta_i = p_{i1}^* \sqrt{\gamma_1}\ , \quad i = 1, 2, \ldots p, \text{ and}$$

$$\psi(u_1, u_2, \ldots, u_p) = \text{const. } \exp\left[-\frac{1}{2} \sum_{1}^{p} u_i^2 \right] ,$$

then the conditions of the lemma are satisfied. From this it follows that (30) is a decreasing function of $\sqrt{\gamma_1}$. This completes the proof.

<u>Theorem 4.2.</u> The power of the likelihood ratio test

$$(32)\qquad \psi_2 = (1 + c_1)(1 + c_2) \ldots (1 + c_p) \le \mu$$

is a monotonically increasing function of $\sqrt{\gamma_i}$ for each $i = 1, 2, \ldots, p$.

<u>Proof:</u> From the mechanics of the proof of the last theorem, it is clear that it is enough to show that the region $R_{1\psi_2}$ in (22) is an ellipsoid with center at the origin. Now

$$\psi_2 = \text{Det.}\left[I_p + (V_2V_2')\,(V_1V_1')^{-1} \right]$$

$$= \left| V_1V_1' + V_2V_2' \right| \big/ \left| V_1V_1' \right| \quad .$$

Hence (32) reduces to

$$\left| V_1 V_1' + V_2 V_2' \right| \leq \mu \quad \left| V_1 V_1' \right|$$

But

$$\left| V_1 V_1' + V_2 V_2' \right|$$

$$= \left| V_1 V_1' + V_{21} V_{21}' + \underline{v}_1 \underline{v}_1' \right|$$

$$= \left| \underline{v}_1 \underline{v}_1' + V_1^* \right| , \quad \text{say} .$$

Clearly V_1^* is nonsingular p.d. almost everywhere. Hence

$$\underline{v}_1 \underline{v}_1' + V_1^* \quad = \det \left[(\underline{v}_1 \underline{v}_1' V_1^{*-1} + I_p) V_1^* \right]$$

$$= \left| V_1^* \right| . \det (I_p + \underline{v}_1 \underline{v}_1' V_1^{*-1})$$

Let ω be the nonzero root of $\underline{v}_1 \underline{v}_1' V_1^{*-1}$, then $\left| I_p + \underline{v}_1 \underline{v}_1' V_1^{*-1} \right| = (\omega+1)(o+1) \ldots (o+1) = \omega + 1$. But ω is then also the root of $\underline{v}_1' V_1^{*-1} \underline{v}_1$, which is a scalar. Thus

$$\left| \underline{v}_1 \underline{v}_1' + V_1^* \right| = \left| V_1^* \right| \cdot (1 + \underline{v}_1' V_1^{*-1} \underline{v}_1) ,$$

so that (32) becomes

$$(33) \qquad \underline{v}_1' V_1^{*-1} \underline{v}_1 \leq \mu \cdot \frac{V_1 V_1'}{V_1 V_1' + V_{21} V_{21}'} - 1$$

Clearly V_1^* depends just on V_1 and V_{21} and is p.d. Hence by (33) $R_{1\psi_2}$ is an ellipsoid with center at the origin. This completes the proof.

5. Some interesting sidelights on the proof of the monotonicity of the largest root criterion.

As stated earlier, a proof of this property has already been given by Roy and Mikhail. However, if we attempt to carry out the proof along the lines used for the other two tests we pass through certain results which appear to be interesting on their own. We shall state these and present the proof only in case the result is new or the proof is of use otherwise.

Lemma 5.1: If c is any latent root of $(V_2V_2')(V_1V_1')^{-1}$, then the equation

$$(34) \qquad \left| (V_2V_2')(V_1V_1')^{-1} - c\, I_p \right| = 0$$

yields a homogeneous second degree surface in $\underline{v}_1$, provided V_1 and V_{21} are held fixed.

Proof: The equation (34) is equivalent to

$$\left| \underline{v}_1\, \underline{v}_1' + V_{21}\, V_{21}' - c\, V_1V_1' \right| = 0$$

(35) or

$$\left| \underline{v}_1\, \underline{v}_1' + V_{1c}^{*} \right| = 0 \text{ , say .}$$

In case V_{1c}^{*} is nonsingular, this is equivalent to (as in the last theorem):

$$1 + \underline{v}_1'\, V_{1c}^{*-1}\, \underline{v}_1 = 0 \text{ ,} \qquad \text{or}$$

(36)

$$\underline{v}_1'\, (-V_{1c}^{*})^{-1}\, \underline{v}_1 = 1 \text{ ,}$$

which completes the proof.

Lemma 5.2. (Random Separation of the largest root).

There exists a real number λ_1^{*} which depends on V_1 and V_{21} but not on $\underline{v}_1$, such that

$$(37) \qquad c_1 \geq \lambda_1^{*} \geq c_2 \text{ .}$$

It is well known that both equalities can hold only at a set of measure zero.

<u>Proof.</u> Assume $c_1 > c_2$, which is true a.e. Let $\underline{b}_1$ and $\underline{b}_2$ be the normalised characteristic vectors of $(V_2V_2')(V_1V_1')^{-1} = \Gamma$, say. Then it is well known that

$$\underline{b}_1' \underline{b}_2 = 0 ,$$

$$\underline{b}_j' \Gamma \underline{b}_j = c_j, \; j = 1, 2 .$$

Let

$$(V_1V_1')^{-1} = V' V , \quad \text{where } V \text{ is triangular, and}$$

$$\underline{d}_j = \underline{b}_j' V , \quad j = 1, 2 .$$

Then

$$c_j = \underline{d}_j' \underline{v}_1 \underline{v}_1' \underline{d}_j + \underline{d}_j' V_{21} V_{21}' \underline{d}_j , \quad j = 1, 2 .$$

Define

$$\lambda_j = \underline{d}_j' V_{21} V_{21}' \underline{d}_j , \quad j = 1, 2, \ldots,$$

and suppose that the largest root of $V V_{21} V_{21}' V'$ is λ_1^*. Then

$$\lambda_1^* = \sup_{\underline{b}} (\underline{b}' V V_{21} V_{21}' V' \underline{b})$$

$$= \underline{b}_1^{*\prime} V V_{21} V_{21}' V' \underline{b}_1^* , \text{ say} .$$

Hence

(38) $$\lambda_1^* \leq \underline{b}_1^{*\prime} V (\underline{v}_1 \underline{v}_1' + V_{21} V_{21}') V' \underline{b}_1^*$$

$$= \underline{b}_1^{*\prime} \Gamma \underline{b}_1^* \leq c_1 .$$

For any fixed $V \underline{v}_1$, choose a vector $\underline{g}$ such that

(39) $$\underline{g}' V \underline{v}_1 = 0, \quad \underline{g}' \underline{g} = 1, \text{ and}$$

$$\underline{g} = \theta_1 \underline{b}_1 + \theta_2 \underline{b}_2 ,$$

where θ_1, θ_2 are real numbers. Notice that this can always be done, and that $0 \leq \theta_1^2 , \theta_2^2 \leq 1$. But

$$\begin{aligned} \underline{g}' \Gamma \underline{g} &= (\theta_1 \underline{b}_1' + \theta_2 \underline{b}_2') \Gamma (\theta_1 \underline{b}_1 + \theta_2 \underline{b}_2) \\ &= \theta_1^2 \underline{b}_1' \Gamma \underline{b}_1 + \theta_2^2 \underline{b}_2' \Gamma \underline{b}_2 \\ &= c_1 \theta_1^2 + c_2 \theta_2^2 \\ &= c_1 \theta_1^2 + (1 - \theta_1^2) c_2 \ . \end{aligned}$$

Hence

(40) $$c_2 \leq \underline{g}' \Gamma \underline{g} \leq c_1 \ .$$

However

$$\begin{aligned} \underline{g}' \Gamma \underline{g} &= \underline{g}' V \underline{v}_1 \underline{v}_1' V' \underline{g} + \underline{g}' V V_{21} V_{21}' V' \underline{g} \\ &\leq \underline{g}' V V_{21} V_{21}' V' \underline{g} \\ &\leq \lambda_1^* \ . \end{aligned}$$

Thus

$$c_2 \leq \lambda_1^* \ ,$$

and the lemma is proved.

Corollary. Consider the roots $c_1, c_2, \ldots, c_p$ for any fixed value of V_1 and V_{21}. As $\underline{v}_1$ varies, let $S_{1i}(\mu)$ $(i = 1, 2, \ldots, p)$ be the surface

$$c_i = \mu$$

in the p-dimensional space of $\underline{v}_1$, where μ is a constant. Then for any value of μ, if $S_{11}(\mu)$ exists, then none of $S_{12}(\mu), \ldots, S_{1p}(\mu)$ exist. Similarly if any $S_{1i}(\mu)$ $(i \geq 2)$ exists, then $S_{11}(\mu)$ does not exist. A proof follows from the fact $S_{11}(\mu)$ exists or not according as $\mu \geq \lambda_1^*$ or $\mu < \lambda_1^*$.

Lemma 5.3. (a) Recall (36). Let $S_1(\mu)$ be the surface given by the equation

(41) $$\underline{v}_1' (-V_{1\mu}^*)^{-1} \underline{v}_1 = 1 \ ,$$

where μ is a given real number. Then

$$S_1(\mu) = S_{11}(\mu) \cup S_{12}(\mu) \cup \ldots \cup S_{1p}(\mu)$$

(b) Let μ be a given real number. For fixed V_1 and V_{21}, if the surface $S_{11}(\mu)$ exists (i.e. if $\mu \geq \lambda_1^*$), then its equation is given by (40).

The proof of this lemma is easy. The proof of the monotonicity of Roy's test then follows by showing that (41) is an ellipsoid or that $(-V_{1\mu}^*)$ is positive definite (if $\mu \geq \lambda_1^*$). But this it true since

$$(-V_{1\mu}^*) = \mu(V' V)^{-1} - V_{21} V_{21}'$$

$$= [\mu I_p - (V_{21} V_{21}')(V' V)](V' V)^{-1} ,$$

and λ_1^* is the largest root of $(V_{21} V_{21}')(V' V)$.

The crucial role played by the separation property (of Lemma 5.2) is evident. The author feels that this and other related properties may be useful in other techniques also in normal multivariate analysis.

6. Acknowledgement. Thanks are due to Professor S. N. Roy for going through this paper and for the research opportunity he gave me during 1961-62. Lemma 5.2 was suggested by him, together with a partial proof. Professor Roy also informs me that some of the above results have been independently obtained by G. S. Mudholkar (using a different approach) some time after I obtained them.

REFERENCES

[1] Anderson, R. W., Introduction to Multivariate Analysis, John Wiley and Sons, New York, 1956.

[2] Roy, S. N., Some Aspects of Multivariate Analysis, John Wiley and Sons, New York, 1957.

[3] Roy, S. N. and Mikhail, W., "On the monotonic character of the power functions of two multivariate tests", Annals of Mathematical Statistics, Vol. 32, 1961.