

ON THE MONOTONIC CHARACTER OF THE POWER FUNCTIONS OF TWO MULTIVARIATE TESTS¹

BY S. N. ROY AND W. F. MIKHAIL

University of North Carolina

1. Summary and introduction. The largest characteristic root has been proposed in [2] as a test statistic in (i) the multivariate analysis of variance test, and (ii) testing that two sets of variates are independent. In this paper it is shown that, in each case, the power function is a monotonically increasing function of each non-centrality parameter, separately. This property was stated in [2] without proof. This provides a stronger result than would be obtained by any direct use of Anderson's Theorem [1] which implies that the power function increases when all the roots are simultaneously increased in the same ratio. The proof of the monotonicity property for the multivariate analysis of variance is given in Section 3, and in Section 4 it is shown how the proof is modified for testing independence between two sets of variates.

2. Preliminaries for the multivariate analysis of variance situation. Let u , s and n denote respectively the "effective" number of variates, the degrees of freedom of the hypothesis and the degrees of freedom for the error and let $t = \min(u, s)$. Then, with $\mathbf{X} \equiv [x_{ij}]: u \times s$ and $\mathbf{Y} \equiv [y_{ij}]: u \times n$, the canonical form for the d.f. in the multivariate analysis of variance model was obtained in [2, p. 86] to be

$$(2.1) \quad \begin{aligned} & [1/(2\pi)^{\frac{1}{2}u(s+n)}] \\ & \cdot \exp \left[-\frac{1}{2} \left\{ \sum_{i=1}^u \sum_{j=1}^n y_{ij}^2 + \sum_{i=1}^t (x_{ii} - \theta_i)^2 + \sum_{i=t+1}^u x_{ii}^2 + \sum_{i=1}^u \sum_{j \neq i-1}^s x_{ij}^2 \right\} \right] \\ & \cdot \prod_{i=1}^u \prod_{j=1}^s dx_{ij} \prod_{i=1}^u \prod_{j=1}^n dy_{ij} \equiv Q d\mathbf{X} d\mathbf{Y}, \end{aligned}$$

and the acceptance region of size $1 - \alpha$ for the linear hypothesis H_0 of analysis of variance, i.e., for the case $\theta_1 = \theta_2 = \dots = \theta_t = 0$, can be expressed as

$$(2.2) \quad \mathfrak{D} = \{ \mathbf{X}, \mathbf{Y} : c_M[(\mathbf{X}\mathbf{X}')(\mathbf{Y}\mathbf{Y}')^{-1}] \leq \mu \},$$

where μ is given by

$$(2.3) \quad P[\mathbf{X}, \mathbf{Y} \in \mathfrak{D} \mid \theta'_i s = 0] = 1 - \alpha,$$

and where the θ_i^2 's are the non-centrality parameters defined in [2, pp. 85-86],

Received September 1, 1959; revised April 5, 1961.

¹ This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under Contract No. AF 49(638)-213. Reproduction in whole or in part is permitted for any purpose of the United States Government.

and $c_M(\mathbf{A})$ denotes the largest characteristic root of the (square) matrix \mathbf{A} . The problem now is to prove that the integral, $P[\mathbf{X}, \mathbf{Y} \in \mathfrak{D}] = \int_{\mathfrak{D}} Q d\mathbf{X} d\mathbf{Y}$ is a monotonically decreasing function of each θ_i^2 , separately. If we regard the domain \mathfrak{D} as one of dimensionality $u s + u n$, where $u s$ dimensions are associated with \mathbf{X} and $u n$ with \mathbf{Y} , then it is clear that we can rewrite this integral as

$$(2.4) \quad \int_{\mathfrak{D}^*} \text{const. exp} \left[-\frac{1}{2} \left(\sum_{i=1}^u \sum_{j=1}^n y_{ij}^2 + \sum_{i=1}^u \sum_{j=1}^s x_{ij}^2 \right) \right] d\mathbf{X} d\mathbf{Y} \\ = \int_{\mathfrak{D}^*} Q^* d\mathbf{X} d\mathbf{Y}, \quad (\text{say}),$$

where \mathfrak{D}^* is merely the domain \mathfrak{D} translated by θ_i along x_{ii} , that is, along the i th axis (with $i = 1, 2, \dots, t$). Notice that if, in the integral (2.4), we replace the domain \mathfrak{D}^* by \mathfrak{D} , the integral over the new domain becomes equal to $1 - \alpha$, where α is the probability of the first kind of error. Let

$$(2.5) \quad \mathbf{Y}\mathbf{Y}' = (\tilde{\mathbf{V}}'\tilde{\mathbf{V}})^{-1},$$

where $\tilde{\mathbf{V}}$ is a $u \times u$ triangular matrix with zeros above the main diagonal. Observe that

$$c_M[(\mathbf{X}\mathbf{X}')(\mathbf{Y}\mathbf{Y}')^{-1}] = c_M[(\mathbf{X}\mathbf{X}')(\tilde{\mathbf{V}}'\tilde{\mathbf{V}})] = c_M[(\tilde{\mathbf{V}}\mathbf{X})(\tilde{\mathbf{V}}\mathbf{X})'],$$

and rewrite (2.2), that is, the domain \mathfrak{D} as

$$(2.6) \quad \mathfrak{D} = \{\mathbf{Y}, \mathbf{X}: c_M[\tilde{\mathbf{V}}\mathbf{X}(\tilde{\mathbf{V}}\mathbf{X})'] \leq \mu\}.$$

Notice that $\tilde{\mathbf{V}}$ is a function of \mathbf{Y} given by (2.5).

The problem now can be rephrased in the following way. How does the integral of Q^* given by (2.4) over the domain \mathfrak{D} given by (2.6) change under successive translations of θ_1 along x_{11} , of θ_2 along x_{22} , \dots , θ_t along x_{tt} ? It is clear that the successive changes are cumulative. It will be also seen from the mechanics of the demonstration that if we can prove that the integral decreases for the first shift of \mathfrak{D} , namely, by θ_1 along x_{11} , then the general theorem itself will be proved.

3. Proof of the monotonicity property for the multivariate analysis of variance situation. The proof is developed in three main steps discussed in the following subsections.

3.1 *The proof for the univariate case.* In this case, $u = 1$ and we can drop the first subscript in \mathbf{X} , \mathbf{Y} . The domain \mathfrak{D} of (2.6) now takes on the form

$$(3.1.1) \quad \mathfrak{D} = \left\{ \mathbf{X}, \mathbf{Y}: \sum_{j=1}^s x_j^2 \leq \mu \sum_{j=1}^n y_j^2 \right\},$$

and the integral of Q over this domain takes the final form

$$(3.1.2) \quad \int_{\mathfrak{D}^*} \exp \left[-\frac{1}{2} \left(\sum_{j=1}^n y_j^2 + \sum_{j=1}^s x_j^2 \right) \right] \prod_{j=1}^n dy_j \prod_{j=1}^s dx_j.$$

Notice that now \mathfrak{D}^* is just a shift of \mathfrak{D} along x_1 by θ . It is evident from the form of (3.1.2) that the integral (3.1.2) decreases under this shift if, for any given set of y_j 's for $j = 1, 2, \dots, n$, and x_j 's, for $j = 2, 3, \dots, s$,

$$(3.1.3) \quad \int_{-a+\theta}^{a+\theta} \exp \left[-\frac{1}{2} x_1^2 \right] dx_1 < \int_{-a}^a \exp \left[-\frac{1}{2} x_1^2 \right] dx_1,$$

where $a = +[\mu \sum_{j=1}^n y_j^2 - \sum_{j=2}^s x_j^2]^{\frac{1}{2}}$; it is clear that it doesn't matter whether we take θ to be positive or negative. It is easy to verify this and also an even more general result, namely that

$$(3.1.4) \quad \int_{-a+\lambda}^{a+\lambda} \phi(x) dx \leq \int_{-a}^a \phi(x) dx,$$

for all real λ and $a > 0$, where $\phi(x)$ is a continuous function of x , symmetric about 0 and monotonically decreasing with $|x|$. It is also clear that the left side of (3.1.3) monotonically decreases with $|\theta|$.

3.2. *The nature of the multivariate domain* (2.6). We now characterize \mathfrak{D} as a domain in $(x_{11}, x_{21}, \dots, x_{u1})$ for fixed values of $\tilde{\mathbf{V}}, \mu$ and \mathbf{X} (excluding the first column). Toward this end, put $\mathbf{X}^* = \tilde{\mathbf{V}}\mathbf{X}$ and observe that, if ν is any characteristic root of $\mathbf{X}^*\mathbf{X}^{*'} \equiv \mathbf{S}^*$, then

$$(3.2.1) \quad |\mathbf{S}^* - \nu\mathbf{I}| = 0 = |\mathbf{S}^*| - \nu \text{tr}_{p-1} + \nu^2 \text{tr}_{p-2} - \dots - +(-1)^u \nu^u,$$

where tr_j is the trace of the j th order, or in other words, the sum of the j th rowed principal minor determinants of \mathbf{S}^* . But, given x_{ij}^* (for $i = 1, 2, \dots, u; j = 2, \dots, s$), $|\mathbf{S}^*| \equiv |\mathbf{X}^*\mathbf{X}^{*'}|$ is a homogeneous quadratic function of $(x_{11}^*, \dots, x_{u1}^*)$ + a constant which is really a function of the other x_{ij}^* 's just mentioned. The coefficients of the quadratic function are also each a polynomial function of x_{ij}^* 's (for $i = 1, 2, \dots, u; j = 2, 3, \dots, s$). Likewise, if we take any q -rowed principal minor determinant of $[s_{ij}^*]$, say the one with rows and columns numbered, $1, 2, \dots, q$, then that determinant is

$$\left| \begin{bmatrix} x_{11}^* & \dots & x_{1s}^* \\ \cdot & \dots & \cdot \\ x_{q1}^* & \dots & x_{qs}^* \end{bmatrix} \begin{bmatrix} x_{11}^* & \dots & x_{q1}^* \\ \cdot & \dots & \cdot \\ x_{1s}^* & \dots & x_{qs}^* \end{bmatrix} \right|,$$

which, given the other x_{ij}^* 's, is a homogeneous quadratic function of $(x_{11}^*, \dots, x_{q1}^*)$ (in which the coefficients are polynomials in the other x_{ij}^* 's) + a constant which is really a function of the other x_{ij}^* 's. Thus, given ν and the other x_{ij}^* 's, the equation (3.2.1) in ν yields a homogeneous quadric surface in $x_{11}^*, \dots, x_{u1}^*$. Now recall from (2.5) that, given y_{ij} 's, that is $\tilde{\mathbf{V}}$, the $(x_{11}^*, \dots, x_{u1}^*)$ are linear functions of (x_{11}, \dots, x_{u1}) and likewise $(x_{ij}^*, \dots, x_{uj}^*)$ are linear functions of (x_{1j}, \dots, x_{uj}) , for $j = 2, \dots, s$. Thus, given ν and

$$x_{ij} \text{ (for } i = 1, 2, \dots, u; j = 2, \dots, s),$$

the equation (3.2.1) yields a homogeneous quadric surface in (x_{11}, \dots, x_{u1}) in which the coefficients and the constant term are all functions of ν, \mathbf{Y} and the other x_{ij} 's already referred to. This is for any characteristic root ν .

Let us now rewrite (2.2) in the (almost everywhere) equivalent form

$$(3.2.2) \quad \mathfrak{D} = \left\{ \mathbf{X}, \mathbf{Y}: \sup_{\mathbf{a}} \frac{(x_{11} + a_2 x_{21} + \cdots + a_u x_{u1})^2 + \cdots + (x_{1s} + a_2 x_{2s} + \cdots + a_u x_{us})^2}{(y_{11} + a_2 y_{21} + \cdots + a_u y_{u1})^2 + \cdots + (y_{1n} + a_2 y_{2n} + \cdots + a_u y_{un})^2} \leq \mu \right\},$$

where $\mathbf{a}' = (a_2, \dots, a_u)$. Now, given μ, \mathbf{Y} and x_{ij} 's (for $i = 1, \dots, u; j = 2, \dots, s$), (3.2.2) represents the domain of (x_{11}, \dots, x_{u1}) in an u -dimensional Euclidean space, the boundary being given by the surface defined by the equality sign. An equivalent form of the same surface is the homogeneous quadric associated with (3.2.1) after ν is replaced by μ . Next, it is easy to check from the definition of \mathfrak{D} and the manner in which the vector (x_{11}, \dots, x_{u1}) occurs in it that (3.2.2) implies the following:

(i) $(\mathbf{X}, \mathbf{Y}) \in \mathfrak{D}$ implies that $((c\mathbf{x}_1, \mathbf{X}_2), \mathbf{Y}) \in \mathfrak{D}$ for $0 \leq c \leq 1$, where \mathbf{X} is decomposed into $(\mathbf{x}_1, \mathbf{X}_2)$ such that $\mathbf{x}'_i = (x_{11}, \dots, x_{u1})$ and \mathbf{X}_2 is a matrix with the other elements of \mathbf{X} .

(ii) Any straight line passing through the origin $(0, \dots, 0)$ has an intersection with the domain, of finite length.

Thus, given μ, \mathbf{Y} and the other x_{ij} 's (already described), (2.2) or (2.6) can be regarded as a domain of (x_{11}, \dots, x_{u1}) which is the interior of a u -dimensional ellipsoid whose boundary is given by (3.2.1) after μ is substituted for ν . It is well known that there is an orthogonal transformation by which the ellipsoid can be referred to principal axes, or in other words, the transformed equation to the surface becomes free from the product terms in the transformed variables and involves only the square terms with positive coefficients. Let $\mathbf{x}'_1 = [x_{11}, \dots, x_{u1}]$ and

$$(3.2.3) \quad \mathbf{z} = \mathbf{L}\mathbf{x}_1,$$

where $\mathbf{L}: u \times u$ is an orthogonal matrix that transforms the ellipsoid into principal axes. This \mathbf{L} can be determined and the rows of \mathbf{L} , say $(l_{i1}, \dots, l_{iu}), i = 1, 2, \dots, u$, are the direction cosines of the different principal axes. Note that $\mathbf{z}'\mathbf{z} = \mathbf{x}'_1\mathbf{x}_1$. It would be useful to rewrite (2.4), after substitution of \mathfrak{D} for \mathfrak{D}^* and omission of the constant, in the form

$$(3.2.4) \quad \int_{\mathfrak{D}} \exp \left[-\frac{1}{2} \left\{ \sum_{i=1}^u \sum_{j=1}^n y_{ij}^2 + \sum_{i=1}^u \sum_{j=2}^s x_{ij}^2 + \sum_{i=1}^u z_i^2 \right\} \right] d\mathbf{Y} \prod_{i=1}^u \prod_{j=2}^s dx_{ij} \prod_{i=1}^u dz_i,$$

where, given μ, \mathbf{Y} and the x_{ij} 's, the domain \mathfrak{D} , as a domain in (z_1, \dots, z_u) , forms the interior of an ellipsoid referred to principal axes (that is, in a form which is free from the product terms of z 's and involves only the square terms with positive coefficients). In other words, \mathfrak{D} is symmetric about the origin in each z_i separately. A displacement θ_1 along the direction of x_{11} might be regarded as the resultant of a displacement $l_{11}\theta_1$ along z_1 , that is, along the direction with cosines $(l_{11}, l_{12}, \dots, l_{1u})$, a displacement $l_{21}\theta_1$ along z_2 , that is, along the direc-

tion with cosines $(l_{21}, l_{22}, \dots, l_{2u})$, and so on, and finally a displacement $l_{u1}\theta_1$ along z_u , that is, along the direction with cosines (l_{u1}, \dots, l_{uu}) . It should be remembered that these l_{ij} 's are functions of μ, \mathbf{Y} and the x_{ij} 's of (3.2.4).

3.3 *The final step in the proof of the monotonicity property.* Looking at (3.2.4) and using (3.1.4) we observe that a displacement of \mathfrak{D} by $l_{11}\theta_1$ along z_1 will decrease the integral under (3.2.4), because, for any given set μ, \mathbf{Y}, x_{ij} 's and z_2, z_3, \dots, z_u ,

$$(3.3.1) \quad \int_{-a+l_{11}\theta_1}^{a+l_{11}\theta_1} \exp\left[-\frac{1}{2}z_1^2\right] dz_1 < \int_a^a \exp\left[-\frac{1}{2}z_1^2\right] dz_1,$$

where a and $l_{11}\theta_1$, without any loss of generality, can be assumed to be positive. Recall that a is a function of μ, \mathbf{Y}, x_{ij} 's and z_2, \dots, z_u . Using the same argument for successive displacements by $l_{21}\theta_1$ along z_2 , by $l_{31}\theta_1$ along z_3 , and so on, and finally by $l_{u1}\theta_1$ along z_u we have successive decreases of the integral. In other words, the resultant displacement which is along x_{11} and by θ_1 decreases the integral. At this point we go back to the integral over \mathfrak{D} of Q^* , forget about the z_i 's, use the result just stated about a displacement by θ_1 along x_{11} , apply successive displacements by θ_2 along x_{22} , θ_3 along x_{33} and so on, and finally θ_t along x_{tt} and eventually obtain an integral over the displaced domain \mathfrak{D}^* which is less than the one over the original domain \mathfrak{D} . It is also clear from the mechanics of the proof that the integral over \mathfrak{D}^* decreases as each $|\theta_i|, i = 1, 2, \dots, t$, increases separately. This proves the monotonicity property.

4. **The case of the test for independence between two sets of variates.** With a $(p + q)$ set ($p \leq q$) of variables let us assume, for a sample of size $n + 1$ ($> p + q$), the canonical distribution law ([2], p. 68)

$$(4.1) \quad \left[1 / (2\pi)^{\frac{1}{2}(p+q)n} \prod_{i=1}^p (1 - \rho_i^2)^{n/2} \right] \cdot \exp \left[-\frac{1}{2} \left\{ \sum_{i=1}^p \frac{1}{1 - \rho_i^2} (x_{ij}^2 + y_{ij}^2 - 2\rho_i x_{ij} y_{ij}) + \sum_{i=p+1}^q \sum_{j=1}^n y_{ij}^2 \right\} \right] \cdot \prod_{i=1}^p \prod_{j=1}^n dx_{ij} \prod_{i=1}^q \prod_{j=1}^n dy_{ij},$$

where ρ_i 's are the population canonical correlation coefficients. The hypothesis of independence H_o is equivalent to the hypothesis that ρ_i 's = 0; the acceptance region (of size $1 - \alpha$) for H_o is

$$(4.2) \quad \mathfrak{D} = \{ \mathbf{X}, \mathbf{Y} : c_M[(\mathbf{X}\mathbf{X}')^{-1}(\mathbf{X}\mathbf{Y}')(\mathbf{Y}\mathbf{Y}')^{-1}(\mathbf{Y}\mathbf{X}')] \leq \mu \},$$

where μ is given by

$$P[\mathbf{X}, \mathbf{Y} \in \mathfrak{D} | H_o] = 1 - \alpha.$$

The monotonicity in this case is proved in exactly the same way as in the previous case. For this purpose we rewrite the \mathfrak{D} of (4.2) as

$$(4.3) \quad \mathfrak{D} = \left\{ \mathbf{X}, \mathbf{Y}: c_M[(\mathbf{U}\mathbf{U}')(\mathbf{V}\mathbf{V}')^{-1}] \leq \frac{\mu}{1-\mu}, \text{ i.e., } \leq \mu^* \text{ (say)} \right\},$$

and the d.f. of (4.1) as

$$(4.4) \quad \text{const. exp} \left[-\frac{1}{2} \left(\sum_{i=1}^p \sum_{j=1}^q (u_{ij} - \gamma_{ij} t_{ij})^2 + \sum_{i=1}^q \sum_{j=1}^i t_{ij}^2 + \sum_{i=1}^p \sum_{j=q+1}^n v_{ij}^2 \right) \right] \\ \cdot d\mathbf{U} d\mathbf{V} \prod_{i=1}^q t_{ii}^{n-i} d\tilde{\mathbf{T}},$$

where

$$(4.5) \quad \gamma_{ij} = \rho_i / (1 - \rho_i^2)^{\frac{1}{2}} = \theta_i \text{ (say)},$$

(for $j = 1, 2, \dots, i; i = 1, 2, \dots, p$), and $\gamma_{ij} = 0$,

otherwise, and where $\tilde{\mathbf{T}}: q \times q$, $\mathbf{U}: p \times q$, and $\mathbf{V}: p \times (n - q)$ are related to $(\mathbf{X}: p \times n, \mathbf{Y}: q \times n)$ in the following way:

$$(4.6) \quad \mathbf{Y} = \tilde{\mathbf{T}}\mathbf{L},$$

where $\mathbf{L}: q \times n$ is orthonormal and $\tilde{\mathbf{T}}$ is lower triangular. $\mathbf{M}: (n - q) \times n$ is an orthogonal completion of \mathbf{L} , $\mathbf{D}_{a_i}; p \times p$ stands for a diagonal matrix with diagonal elements a_1, a_2, \dots, a_p , and \mathbf{U} and \mathbf{V} are given by

$$(4.7) \quad \mathbf{U} = \mathbf{D}_{(1-\rho_i^2)^{-\frac{1}{2}}} \mathbf{X}\mathbf{L}', \quad \mathbf{V} = \mathbf{D}_{(1-\rho_i^2)^{-\frac{1}{2}}} \mathbf{X}\mathbf{M}'.$$

In the transformation from (4.1) to (4.4), \mathbf{M} does not occur explicitly, \mathbf{L} does, but is easily integrated out as in [2, pp. 196-197].

The probability of the second kind of error is given by integrating (4.4) over the domain (4.3). It is easy to see that, aside from the positive constant factor, this is equivalent to

$$(4.8) \quad \int_{\mathfrak{D}^*} \exp \left[-\frac{1}{2} \left(\sum_{i=1}^q \sum_{j=1}^i t_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^q u_{ij}^2 + \sum_{i=1}^p \sum_{j=q+1}^n v_{ij}^2 \right) \right] d\mathbf{U} d\mathbf{V} \prod_{i=1}^q t_{ii}^{n-i} d\tilde{\mathbf{T}}$$

where, for any given set of $\tilde{\mathbf{T}}$ and \mathbf{V} , \mathfrak{D}^* is just \mathfrak{D} displaced by $\theta_1 t_{11}$ along u_{11} , by $\theta_2 t_{21}$ along u_{21} and $\theta_2 t_{22}$ along u_{22} , and so on, and finally by $\theta_p t_{p1}$ along u_{p1} , $\theta_p t_{p2}$ along u_{p2} , \dots , $\theta_p t_{pp}$ along u_{pp} . Notice that when H_0 is true, that is, when $\theta_i = 0$, we should have \mathfrak{D}^* replaced by \mathfrak{D} in the integral (4.8). Using the same kind of argument as in Section 3 it follows that, for any given $\tilde{\mathbf{T}}$, the partial integral over \mathbf{U} and \mathbf{V} decreases as \mathfrak{D} is displaced by $\theta_1 t_{11}$ along u_{11} , where $t_{11} > 0$, almost everywhere, and with this displacement of \mathfrak{D} , it is easy to see that the total integral (if we now integrate over $\tilde{\mathbf{T}}$) will also decrease. From considerations of symmetry, the same result would follow for the other displacements in that the displacement associated with any θ_i^2 could be represented as a $\theta_i t_{i1}$ along u_{i1} under a suitable transformation. Thus (4.15) monotonically decreases as each $|\theta_i|$, that is, each $|\rho_i|$ separately increases.

Concluding remarks. The power functions of the λ -criteria for the multivariate linear hypothesis and for the test of independence between two sets of

variates have also somewhat similar monotonicity properties that will be discussed in a subsequent paper.

REFERENCES

- [1] T. W. ANDERSON, "The integral of a symmetrical unimodal function," *Proc. Amer. Math. Soc.*, Vol. 6 (1955), pp. 170-176.
- [2] S. N. ROY, *Some Aspects of Multivariate Analysis*, John Wiley and Sons, New York, 1957.
- [3] S. N. ROY AND R. GNANADESIKAN, "Further contributions to multivariate confidence bounds," *Biometrika*, Vol. 44 (1957), pp. 339-410.
- [4] S. N. ROY AND R. GNANADESIKAN, "Some contributions to ANOVA in one or more dimensions—Parts I and II," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 304-340.