MONOTONICITY OF THE POWER FUNCTIONS OF SOME TESTS OF INDEPENDENCE BETWEEN TWO SETS OF VARIATES[1]

By T. W. Anderson and S. Das Gupta

Columbia University

1. Summary. For testing independence between two sets of normally distributed variates we consider the class of test procedures which are invariant under certain groups of transformations and depend only on the sample canonical correlation coefficients [1]. The power function of such a test depends only on the population canonical correlation coefficients as parameters, which may be regarded as measures of deviation from the hypothesis. In this paper sufficient conditions on the invariant procedure for the power function to be a monotonically increasing function of each of the parameters are obtained. The likelihood-ratio test [1] and Roy's maximum root test [5] satisfy these conditions. In [4] only the unbiasedness of the maximum root test was proved, although the authors claimed to prove the monotonicity property.

2. Introduction. Consider a $(p + q) \times (n + 1)$ random matrix $\mathbf{Z}$ whose column vectors $\mathbf{z}_\alpha$'s are independently distributed according to a $(p + q)$-variate normal distribution $N(\boldsymbol{\xi}, \boldsymbol{\Sigma})$, where the positive definite matrix $\boldsymbol{\Sigma}$ is partitioned into p and q rows and columns as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}.$$

Consider the problem of testing the hypothesis

$$\mathcal{H}_0 : \boldsymbol{\Sigma}_{12} = \mathbf{0}\ (p \times q)$$

against all alternatives. Let the sample covariance matrix be $\mathbf{S}$ which is similarly partitioned as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix},$$

where $n\mathbf{S} = \mathbf{Z}\mathbf{Z}' - (n + 1)\bar{\mathbf{z}}\bar{\mathbf{z}}'$ and $\bar{\mathbf{z}} = [1/(n + 1)]\sum_{\alpha=1}^{n+1} \mathbf{z}_\alpha$. This problem is invariant under transformations

$$\begin{pmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{pmatrix} \mathbf{Z}\mathbf{F}, \qquad \mathbf{z}_\alpha + \mathbf{b}, \qquad \alpha = 1, \cdots, n + 1,$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are nonsingular matrices of order p and q, respectively, and $\mathbf{F}$ is orthogonal. A test procedure which is invariant under these transformations depends only on the characteristic roots $c_1 \geqq c_2 \geqq \cdots \geqq c_p$ of $\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\mathbf{S}_{21}$,

Received 9 August 1963.

[1] This research was sponsored by the Office of Naval Research under Contract Number Nonr-266(33), Project Number NR 042-034. Reproduction in whole or in part is permitted for any purpose of the United States Government.

which are the squares of the nonzero sample canonical correlation coefficients. (We assume $p \leqq q$.) The power function of such a test depends only on the characteristic roots $\rho_1 \geqq \rho_2 \geqq \cdots \geqq \rho_p$ of $\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$, which are the squares of the possibly nonzero population canonical correlation coefficients. The distribution of the characteristic roots of $\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\mathbf{S}_{21}$ is the same as the distribution of the roots of $(\mathbf{XX}')^{-1}(\mathbf{XY}')(\mathbf{YY}')^{-1}(\mathbf{YX}')$, where the density of the matrices $\mathbf{X} = [x_{i\alpha}]: p \times n$ and $\mathbf{Y} = [y_{i\alpha}]: q \times n$ is given by

$$(2.1)\quad (2\pi)^{-\frac{1}{2}(p+q)n} \prod_{i=1}^{p} (1-\rho_i^2)^{-\frac{1}{2}n} \exp\left[-\frac{1}{2}\left\{\sum_{i=1}^{p} (1-\rho_i^2)^{-1} \cdot \sum_{\alpha=1}^{n} (x_{i\alpha}^2 + y_{i\alpha}^2 - 2\rho_i x_{i\alpha} y_{i\alpha}) + \sum_{i=p+1}^{q} \sum_{\alpha=1}^{n} y_{i\alpha}^2\right\}\right],$$

and $\mathcal{H}_0$ holds if and only if $\rho_1 = \rho_2 = \cdots = \rho_p = 0$.

Some invariant tests are:

(i) The likelihood-ratio test [1] whose acceptance region is $\prod_{i=1}^{p} (1 - c_i) \geqq \lambda_1$, a constant.

(ii) Roy's maximum root test [5] whose acceptance region is $c_1 \leqq \lambda_2$, a constant.

3. Tests of independence between two sets of variates. Starting from the canonical form (2.1) of the density of **X** and **Y**, we find that given **Y**, the column vectors $\mathbf{x}_j$'s of **X** are independently distributed, each according to a p-variate normal distribution with covariance matrix **D** which is diagonal with diagonal elements $1 - \rho_1^2, 1 - \rho_2^2, \cdots, 1 - \rho_p^2$. The marginal distribution of **Y** does not depend on ρ_i's. Moreover, the conditional expectation of **X**, given **Y**, is $\mathcal{E}(\mathbf{X} \mid \mathbf{Y}) = \mathbf{AY}$, where $\mathbf{A} = [\mathbf{R}\ \mathbf{0}]: p \times q$, and **R** is the diagonal matrix with diagonal elements $\rho_1, \rho_2, \cdots, \rho_p$. Define

$$\mathbf{S}_h = (\mathbf{XY}')(\mathbf{YY}')^{-1}(\mathbf{YX}'), \qquad \mathbf{S}_e = \mathbf{XX}' - (\mathbf{XY}')(\mathbf{YY}')^{-1}(\mathbf{YX}').$$

If c_i is the ith largest root of $(\mathbf{XX}')^{-1}(\mathbf{XY}')(\mathbf{YY}')^{-1}(\mathbf{YX}')$, then $c_i/(1 - c_i)$ is the ith largest root of $\mathbf{S}_h\mathbf{S}_e^{-1}$. Thus the class of test procedures based on the roots of $(\mathbf{XX}')^{-1}(\mathbf{XY}')(\mathbf{YY}')^{-1}(\mathbf{YX}')$ is the same as the class of test procedures based on the roots of $\mathbf{S}_h\mathbf{S}_e^{-1}$. Let $\mathbf{U} = \mathbf{BXF}$, $\mathbf{V} = \mathbf{BXG}$, where **B** is nonsingular, and $\mathbf{F}: n \times q$ and $\mathbf{G}: n \times (n - q)$ are such that

$$\mathbf{FF}' = \mathbf{Y}'(\mathbf{YY}')^{-1}\mathbf{Y}, \qquad \mathbf{GG}' = \mathbf{I}_n - \mathbf{Y}'(\mathbf{YY}')^{-1}\mathbf{Y}.$$

Then the roots of $\mathbf{S}_h\mathbf{S}_e^{-1}$ are the same as the roots of $(\mathbf{UU}')(\mathbf{VV}')^{-1}$. The matrices **B**, **F** and **G** can be found (see [5], p. 86) such that the conditional density of $\mathbf{U} = [u_{ij}]: p \times q$ and $\mathbf{V} = [v_{ij}]: p \times (n - q)$, given **Y**, is

$$(2\pi)^{-\frac{1}{2}pn} \exp\left[-\frac{1}{2}\left\{\mathrm{tr}\,(\mathbf{VV}') + \sum_{i=1}^{p} (u_{ii} - \tau_i)^2 + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{q} u_{ij}^2\right\}\right],$$

where $\tau_1^2 \geqq \tau_2^2 \geqq \cdots \geqq \tau_p^2$ are the characteristic roots of $\mathbf{AYY}'\mathbf{A}'\mathbf{D}^{-1}$.

THEOREM 1. *An invariant test for which the acceptance region is convex in each* *column vector of* $\mathbf{U}$ *for each set of fixed* $\mathbf{V}$ *and fixed values of the other column vectors of* $\mathbf{U}$ *has a power function which is monotonically increasing in each* ρ_i.

PROOF. It follows from Theorem 3 of [3] that, for given $\mathbf{Y}$, the conditional probability of the acceptance region monotonically decreases in each r_i. Note that $r_1^2, \cdots, r_p^2$ are the possibly nonzero roots of $\mathbf{YY}'\mathbf{A}'\mathbf{D}^{-1}\mathbf{A} = \mathbf{TT}'\mathbf{\Gamma}$, where $\mathbf{YY}' = \mathbf{TT}'$, $\mathbf{T}$ being a nonsingular $q \times q$ matrix, and

$$\mathbf{\Gamma} = \mathbf{A}'\mathbf{D}^{-1}\mathbf{A} = \begin{bmatrix} \mathbf{D}_\theta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

and $\mathbf{D}_\theta$ is a diagonal matrix with diagonal elements $\theta_1, \cdots, \theta_p$, where $\theta_i = \rho_i^2/(1 - \rho_i^2)$. We denote the ith largest characteristic root of a matrix by $\text{ch}_i(\cdot)$. Then

$$\text{ch}_i(\mathbf{TT}'\mathbf{\Gamma}) = \text{ch}_i(\mathbf{T}'\mathbf{\Gamma}\mathbf{T}), \qquad i = 1, \cdots, p.$$

Let $\mathbf{\Gamma}^*$ be the matrix obtained by changing the nonzero elements of $\mathbf{\Gamma}$ from θ_i to θ_i^*, where $\theta_i^* \geqq \theta_i (i = 1, \cdots, p)$. Then $\mathbf{T}'\mathbf{\Gamma}^*\mathbf{T} - \mathbf{T}'\mathbf{\Gamma}\mathbf{T}$ is a positive semidefinite matrix. It follows from ([2], pp. 33–34) that

$$\text{ch}_i(\mathbf{T}'\mathbf{\Gamma}^*\mathbf{T}) \geqq \text{ch}_i(\mathbf{T}'\mathbf{\Gamma}\mathbf{T}), \qquad i = 1, \cdots, p.$$

Thus, for any given $\mathbf{Y}$, the conditional probability of the acceptance region decreases monotonically in each ρ_i. The theorem now follows from the fact that the marginal distribution of $\mathbf{Y}$ does not involve any ρ_i.

Let $e_1 \geqq e_2 \geqq \cdots \geqq e_p$ be the roots of $(\mathbf{UU}')(\mathbf{VV}')^{-1}$. Then $e_i = c_i/(1 - c_i)$. Thus the relation $c_1 \leqq \mu$ is equivalent to the relation $e_1 \leqq \mu/(1 - \mu) = \mu^*$, say. Let $d_i = 1 + e_i (i = 1, \cdots, p)$, and let W_k be the sum of all different products of $d_1, \cdots, d_p$ taken k at a time $(k = 1, \cdots, p)$. In particular,

$$W_p = \prod_{i=1}^{p} d_i = \left[\prod_{i=1}^{p} (1 - c_i)\right]^{-1}.$$

The following results are obtained from Section 3 of [3] and Theorem 1:

COROLLARY 1.1. *The maximum root test of Roy has a power function which is monotonically increasing in each* ρ_i.

THEOREM 2. *A test having the acceptance region* $\sum_{k=1}^{p} a_k W_k \leqq \mu$, $a_k \geqq 0$, *has a power function which is monotonically increasing in each* ρ_i.

COROLLARY 2.1. *The power function of the likelihood-ratio test monotonically increases as each* ρ_i *increases.*

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[2] COURANT, R. and HILBERT, D. (1953). *Methods of Mathematical Physics,* **1**. Interscience, New York.

[3] DAS GUPTA, S., ANDERSON, T. W., and MUDHOLKAR, G. S. (1964). Monotonicity of the power functions of some tests of the multivariate linear hypothesis. *Ann. Math. Statist.* **35** 200–205.

[4] 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.

[5] ROY, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.