Biometrika (1971), 58, 1, *p.* 123
Printed in Great Britain

Some optimal multivariate tests

By S. JOHN
Australian National University

Summary

Tests that are best for detecting small deviations from the null hypothesis are derived for a number of hypotheses concerning multivariate normal populations. Both one-sided and two-sided tests are considered.

1. Introduction

The majority of test criteria used in multivariate analysis are functions of the roots of certain determinantal equations constructed from sample moments of order not greater than two. In general there exists no best test. Pillai (1966), Pillai & Jayachandran (1967, 1968) and Pillai & Dotson (1969) have compared the powers of a few among many possible competing test criteria by evaluating their powers numerically in small dimensional cases. The numerical approach, besides being costly and laborious, has a number of other obvious limitations. That locally best tests can be derived mathematically from known results seems not to have been noted.

2. Test of the hypothesis that the covariance matrix of a normal population is as specified

Only the canonical form of this hypothesis where the specified matrix is the identity matrix need be considered. Let $\mathbf{X}$ be the $k \times N$ matrix whose jth column is the jth of N independent observations from the population. The testing problem is invariant under the following transformations: (i) the addition of an arbitrary vector to each column of $\mathbf{X}$; (ii) the premultiplication of $\mathbf{X}$ by an orthogonal matrix $\mathbf{G}$; (iii) the postmultiplication of $\mathbf{X}$ by an orthogonal matrix $\mathbf{H}$. A test function ϕ that is invariant under these transformations can depend on the observations only through the latent roots $t_1, \ldots, t_k$ of $\mathbf{S} = N\mathbf{V}$, where $\mathbf{V}$ is the sample covariance matrix, and its power π_ϕ will depend on the population parameters only through the latent roots $\omega_1, \ldots, \omega_k$ of the population covariance matrix by a theorem of Lehmann (1950) in the theory of invariant tests.

The first directional derivative of π_ϕ with respect to the ω's in the direction $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k)'$ at $(1, \ldots, 1)'$ is the integral of the product of ϕ and the directional derivative of the joint density of the t's, which can be obtained from James [1960, expression (28)] with $n = N-1$, $\sigma^2 = 1$, $\sigma_i^2 = \omega_i$ and $l_i = t_i/n$ $(i = 1, \ldots, k)$ and equals the product of the null density of the t's and $\frac{1}{2}\Sigma\gamma_i\{\Sigma t_j/k - (N-1)\}$. Hence, by the generalized Neyman–Pearson lemma, the test that maximizes the directional derivative of the power function with respect to the ω's in the direction $\boldsymbol{\gamma}$ at $(1, \ldots, 1)'$ is that which rejects the hypothesis when $T = \Sigma t_i = \operatorname{tr} \mathbf{S} > c$ or when $T < c$, according as $\Sigma\gamma_i > 0$ or $\Sigma\gamma_i < 0$, c being so chosen that the test has the required size; the directional derivative of π_ϕ at $(1, \ldots, 1)'$ in the direction $\boldsymbol{\gamma}$ is zero for any ϕ, if $\Sigma\gamma_i = 0$. Recall that the likelihood ratio criterion is equivalent to $|\mathbf{S}|^N \exp\{-\mathrm{tr}\,(\mathbf{S})\}$. The null distribution of T is chi-squared with $(N-1)k$ degrees of freedom.

If we attempt similarly to obtain the test that maximizes the second directional derivative of π_ϕ at $(1, \ldots, 1)'$ among tests for which the first derivative is zero, we shall find that it depends on the direction inextricably. However, the two sided test that rejects the hypothesis when T falls outside the interval $[c_1, c_2]$, where c_1 and c_2 are so determined that the integrals of $\delta(t)$, the null density of T, and $t\delta(t)$ between c_1 and c_2 equal $(1-\alpha)$ and $(N-1)k(1-\alpha)$ respectively is a size α test whose power function has zero directional derivative with respect to the ω's in any direction. Values of c_1 and c_2 satisfying these requirements are given by a_n and b_n of Tate & Klett (1959, Table 2), with $n = (N-1)k$ and $\epsilon = 1-\alpha$. The above table was constructed for a different purpose and is not sufficiently extensive; it is hoped to publish later a more extensive table.

3. Test of the hypothesis that the covariance matrix of a normal population is proportional to a given matrix

Here again we need consider only the canonical form of the hypothesis where the given matrix is the identity matrix. Let $\mathbf{X}$ have the same meaning as in § 2. The testing problem is invariant under the following operations: (i) addition of the same vector to each column of $\mathbf{X}$; (ii) premultiplication of $\mathbf{X}$ by a scalar multiple of an orthogonal matrix; (iii) the post multiplication of $\mathbf{X}$ by an orthogonal matrix. A test that is invariant under these operations can depend on the observations only through $z_i = t_i/\Sigma t_j$ $(i = 1, \ldots, k-1)$, where the t's have the same meaning as in § 2. The power of any such test depends on the population parameters only through $\omega_i/\Sigma\omega_j$ $(i = 1, \ldots, k-1)$, the ω's being the latent roots of the population covariance matrix. We may therefore write without loss of generality $\omega_i = \exp\{(\gamma_i - \bar{\gamma})\lambda\}$ $(i = 1, \ldots, k)$, where $\bar{\gamma} = \Sigma\gamma_i/k$.

The joint density of the z's can be obtained from that of the t's. The jth derivative with respect to λ at $\lambda = 0$ of the power of any test function ϕ is the integral of the product of ϕ and the similar derivative of the joint density of the z's. For $j = 1$, it is zero for arbitrary ϕ. For $j = 2$, the second derivative at $\lambda = 0$ of the joint density of the z's with respect to λ is equal to the product of the null density of the z's and a linear function of

$$U = \sum_{i=1}^{k-1} z_i^2 + \left(1 - \sum_{i=1}^{k-1} z_i\right)^2 = \Sigma t_i^2/(\Sigma t_i)^2,$$

in which the coefficient of U is positive. It now follows from the generalized Neyman–Pearson lemma that the test that maximizes the second derivative with respect to λ of the power function at $\lambda = 0$ is the one that rejects the null hypothesis when $U > c$, where c is so determined that the test has the required size. This test seems not to have been proposed previously. An equivalent criterion is

$$V = \sum_{i<j} t_i t_j/(\Sigma t_i)^2 = \tfrac{1}{2}(1-U).$$

Recall that the likelihood ratio criterion is $(\Pi t_i)/(\Sigma t_i)^k$ and note that it equals V if and only if $k = 2$, in which case the null distribution of $(N-2)(\frac{1}{2}V^{-\frac{1}{2}} - 1)$ can be found from Mauchly (1940) and is the F distribution with degrees of freedom 2 and $2N-4$. If $k > 2$, the null distribution of V cannot be given in a simple form and is postponed.

4. Test of independence of two sets of variables

Let $\mathbf{X}$ be the $(p+q)\times N$ matrix whose jth column is the jth of N independent observations from a $(p+q)$-variate normal distribution. The problem of testing that the first p variates are independent of the other q variates is invariant under the following transformations: (i) addition of the same vector to each column of $\mathbf{X}$; and (ii) premultiplication of $\mathbf{X}$ by a $(p+q)\times(p+q)$ matrix of the form

$$\begin{bmatrix}\mathbf{C}_1 & \mathbf{O}\\ \mathbf{0} & \mathbf{C}_2\end{bmatrix},$$

where $\mathbf{C}_1$ and $\mathbf{C}_2$ are arbitrary $p\times p$ and $q\times q$ nonsingular matrices; (iii) postmultiplication of $\mathbf{X}$ by an orthogonal matrix. We may assume without loss of generality that $p \leqslant q$. A test that is invariant under the above transformations can depend on the observations only through the squared canonical correlations r_i^2 $(i = 1, \ldots, p)$, and conversely. The power π_ϕ of any such test function ϕ depends on the population parameters only through ρ_i^2 $(i = 1, \ldots, p)$, the squared population canonical correlations. Exactly as in § 2, we can prove by making use of Constantine [1963, expression (59)] that the invariant test that maximizes the directional derivative of π_ϕ with respect to the ρ^2's in the direction $\boldsymbol{\gamma}$ at $(0, \ldots, 0)'$ is that which rejects the hypothesis when $\Sigma r_i^2 > c$, where c is such that the test has the required size. Since the ρ^2's cannot be negative, we may assume that $\gamma_i \geqslant 0$ $(i = 1, \ldots, p)$. Recall that the likelihood ratio criterion is $\Pi(1-r_i^2)$. The null distribution of Σr_i^2 is that of $V^{(s)}$, tabulated by Pillai (1960). Pillai's s, m and n should be equated to p, $\frac{1}{2}(q-p-1)$ and $\frac{1}{2}(N-p-q-2)$, respectively.

5. Test of equality of covariance matrices of two normal populations

Let $\mathbf{X}^{(i)}$ $(i = 1, 2)$ be the $m\times N^{(i)}$ matrix whose jth column is the jth of $N^{(i)}$ independent observations from the ith population. The testing problem is invariant under the following transformations: (i) addition of an arbitrary vector $\mathbf{a}^{(i)}$ to each column of $\mathbf{X}^{(i)}$ $(i = 1, 2)$; (ii) premultiplication of $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ by an arbitrary nonsingular matrix $\mathbf{C}$; (iii) post multiplication of $\mathbf{X}^{(i)}$ by an orthogonal matrix $\mathbf{H}^{(i)}$ $(i = 1, 2)$. A test that is invariant under the above transformations can depend on the observations only through the roots f_i $(i = 1, \ldots, m)$ of the equation $|\mathbf{S}^{(1)}-f\mathbf{S}^{(2)}| = 0$, where $\mathbf{S}^{(i)} = N^{(i)}\mathbf{V}^{(i)}$, $\mathbf{V}^{(i)}$ being the ith sample covariance matrix. The power of any such test will depend on the population parameters only through the roots ω_i $(i = 1, \ldots, m)$, of the equation $|\boldsymbol{\Sigma}^{(1)}-\omega\boldsymbol{\Sigma}^{(2)}| = 0$.

The first directional derivative of the joint density of the f's with respect to the ω's in the direction $\boldsymbol{\gamma}$ at $(1, \ldots, 1)'$ can be obtained from the expression (65) of James (1964) with $p = N^{(1)}-1$ and $n = N^{(2)}-1$ and is the product of the null density of the f's and $\frac{1}{2}(\Sigma\gamma_i)\{N^{(2)}-1-(N^{(1)}+N^{(2)}-2)\Sigma(1+f_i)^{-1}/m\}$. Together with the generalized Neyman–Pearson lemma this implies that the test that maximizes the directional derivative of the power function with respect to the ω's in the direction $\boldsymbol{\gamma}$ at $(1, \ldots, 1)'$ is the one that rejects the hypothesis when $T = \Sigma(1+f_i)^{-1} = \operatorname{tr}\{(\mathbf{S}^{(1)}+\mathbf{S}^{(2)})^{-1}\mathbf{S}^{(2)}\} < c$ or when $T > c$, according as $\Sigma\gamma_i > 0$ or < 0, c being so chosen that the test has the required size; the directional derivative of the power function in the direction $\boldsymbol{\gamma}$ at $(1, \ldots, 1)'$ is zero for arbitrary $\boldsymbol{\gamma}$, if $\Sigma\gamma_i = 0$. The null distribution of T is the same as that of Pillai's (1960) $V^{(s)}$, with Pillai's s, m and n equal to m, $\frac{1}{2}(N^{(2)}-m-2)$ and $\frac{1}{2}(N^{(1)}-m-2)$ in our notation, assuming that $N^{(i)} \geqslant m+1$ $(i = 1, 2)$.

If one attempts to find the test that maximizes the second directional derivative of the power function at $(1, \ldots, 1)'$ among tests for which the first derivative is zero, one would find that it depends on the direction inextricably. A two-sided size-α test for which the first directional derivative of the power function with respect to the ω's at $(1, \ldots, 1)'$ is zero for any direction is the one that rejects the hypothesis when T falls outside the interval $[c_1, c_2]$, where c_1 and c_2 are determined so that the integrals of $\delta(t)$, the null density function of T, and $t\delta(t)$ between c_1 and c_2 equal $1-\alpha$ and $m\{N^{(2)}-1\}(1-\alpha)/\{N^{(1)}+N^{(2)}-2\}$. It is hoped to publish later tables of c_1 and c_2.

6. Test of a linear hypothesis

In the canonical form the multivariate linear hypothesis is to test that $E(\mathbf{X}) = \mathbf{0}$, using observations $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$, where $\mathbf{X}$ is an $m \times p$ matrix, $\mathbf{Y}$ is an $m \times n$ matrix and $\mathbf{Z}$ is an $m \times q$ matrix, the columns of $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ being mutually independent normal variables with common covariance matrix $\mathbf{\Sigma}$ and with $E(\mathbf{Y}) \equiv \mathbf{0}$ and $E(\mathbf{Z})$ arbitrary whether the hypothesis is true or not; see Hsu (1941). The problem is invariant under the following transformations: (i) addition of an arbitrary matrix to $\mathbf{Z}$; (ii) premultiplication of $\mathbf{X}$ and $\mathbf{Y}$ by an arbitrary nonsingular matrix $\mathbf{C}$; (iii) post multiplication of $\mathbf{X}$ by an orthogonal matrix $\mathbf{G}$ and of $\mathbf{Y}$ by an orthogonal matrix $\mathbf{H}$. The test that is invariant under the above transformations can depend on the observations only through the roots f_i $(i = 1, \ldots, m)$, of the equation

$$|\mathbf{XX}' - f\mathbf{YY}'| = 0$$

and will have power depending on the population parameters only through the roots ω_i $(i = 1, \ldots, m)$ of the equation $|\mathbf{MM}' - \omega\mathbf{\Sigma}| = 0$, where $\mathbf{M} = E(\mathbf{X})$.

The first directional derivative of the joint density of the f's with respect to the ω's in the direction $\boldsymbol{\gamma}$ at $(0, \ldots, 0)'$, which can be obtained from James [1964, expression (73)] derived from Constantine [1963, expression (41)], is the product of the null density of the f's and $\frac{1}{2}(\Sigma\gamma_i)\{(p+n)\Sigma(1+f_i)^{-1}f_i/(mp)-1\}$. Together with the generalized Neyman–Pearson lemma this implies that the invariant test with the maximum value of the directional derivative of the power function with respect to the ω's in the direction $\boldsymbol{\gamma}$ at $(0, \ldots, 0)'$ is that which rejects the hypothesis when $T = \Sigma(1+f_i)^{-1}f_i = \operatorname{tr}\{(\mathbf{XX}'+\mathbf{YY}')^{-1}\mathbf{XX}'\} > c$ or when $T < c$, according as $\Sigma\gamma_i > 0$ or < 0, c being so chosen that the test has the required size; the directional derivative of the power function in the direction $\boldsymbol{\gamma}$ is zero for any test, if $\Sigma\gamma_i = 0$. The null distribution of T is that of Pillai's (1960) $V^{(s)}$, Pillai's s, m and n being in our notation $\min(m,p)$, $\frac{1}{2}(|m-p|-1)$ and $\frac{1}{2}(n-m-1)$, assuming that $n \geqslant m$. Recall that the likelihood ratio criterion is $\Pi(1+f_i)^{-1}$.

As in the case of the hypotheses considered earlier, the test that maximizes the second directional derivative of the power function at $(0, \ldots, 0)'$ among those for which the first derivative is zero depends on the direction inextricably. However, the test that rejects the hypothesis when T lies outside the interval $[c_1, c_2]$, where c_1 and c_2 are so determined that the integrals of $\delta(t)$, the null density function of T, and $t\delta(t)$ from c_1 to c_2 equal $(1-\alpha)$ and $mp(1-\alpha)/(p+n)$ has size α and a zero first directional derivative of the power function with respect to the ω's at $(0, \ldots, 0)'$ in any direction. It is hoped to publish later tables of c_1 and c_2 satisfying the above conditions. Note that $\mathbf{XX}'$ is the sum of products matrix 'due to the hypothesis' and $\mathbf{YY}'$ is the 'error' sum of products matrix.

References

Constantine, A. G. (1963). Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist.* **34**, 1270–85.

Hsu, P. L. (1941). Canonical reduction of the general regression problem. *Ann. Eugenics* **11**, 42–6.

James, A. T. (1960). The distribution of the latent roots of the covariance matrix. *Ann. Math. Statist.* **31**, 151–8.

James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35**, 475–501.

Lehmann, E. L. (1950). Some principles of the theory of testing hypotheses. *Ann. Math. Statist.* **21**, 1–26.

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Statist.* **11**, 204–9.

Pillai, K. C. S. (1960). *Statistical Tables for Tests of Multivariate Hypotheses.* The Statistical Center, University of the Phillippines, Manila.

Pillai, K. C. S. (1966). Noncentral multivariate beta distribution and the moments of traces of some matrices. In *Multivariate Analysis*, ed. P. R. Krishnaiah, pp. 237–51. Academic Press: New York.

Pillai, K. C. S. & Dotson, C. O. (1969). Power comparisons of tests of two multivariate hypotheses based on individual characteristic roots. *Ann. Inst. Statist. Math.* **21**, 49–66.

Pillai, K. C. S. & Jayachandran, K. (1967). Power comparisons of tests of two multivariate hypotheses based on four criteria. *Biometrika* **54**, 195–210.

Pillai, K. C. S. & Jayachandran, K. (1968). Power comparisons of tests of equality of two covariance matrices based on four criteria. *Biometrika* **55**, 335–42.

Tate, R. F. & Klett, G. W. (1959). Optimal confidence intervals for the variance of a normal distribution. *J. Am. Statist. Ass.* **54**, 674–82.

[*Received May* 1970]

NOTE ADDED IN PROOF: The paper of N. Giri (*Ann. Math. Statist.* **39** (1968), 275–7) should be referred to in connexion with §5.

Some key words: Hypothesis testing for the multivariate normal distribution; Tests about covariance matrices; Test of sphericity; Test of independence of two sets of normal variables; Multivariate dispersion analysis; Invariant tests; Locally most powerful tests; Unbiased tests; Two-sided multivariate tests.