Ann. Inst. Statist. Math.
Vol. 44, No. 2, 357–362 (1992)

ON AN OPTIMUM TEST OF THE EQUALITY OF TWO COVARIANCE MATRICES*

N. GIRI

Department of Mathematics and Statistics, University of Montreal, P.O. Box 6128, Station A, Montreal, Quebec, Canada H3C 3J7

(Received September 13, 1990; revised February 21, 1991)

Abstract. Let $X : p \times 1, Y : p \times 1$ be independently and normally distributed p-vectors with unknown means ξ_1, ξ_2 and unknown covariance matrices Σ_1, Σ_2 (> 0) respectively. We shall show that Pillai's test, which is locally best invariant, is locally minimax for testing $H_0 : \Sigma_1 = \Sigma_2$ against the alternative $H_1 : \mathrm{tr}(\Sigma_2^{-1}\Sigma_1 - I) = \sigma > 0$ as $\sigma \to 0$. However this test is not of type D among G-invariant tests.

Key words and phrases: Locally best invariant tests, locally minimax tests, type D critical region.

1. Introduction

Let $X : p \times 1$, $Y : p \times 1$ be independently distributed normal p-vectors with unknown means ξ_1, ξ_2 and unknown covariance matrices Σ_1, Σ_2 respectively. We are interested to test the null hypothesis $H_0 : \Sigma_1 = \Sigma_2$. This problem remains invariant under the group of affine transformations $G_l(p) \times R^p$ transforming $X \to gX + b_1$, $Y \to gY + b_2$, $g \in G_l(p)$—the multiplicative group of $p \times p$ nonsingular matrices and b_1, $b_2 \in R^p$. Let $X_1, \dots, X_{N_1}$ and $Y_1, \dots, Y_{N_2}$ be the samples of sizes N_1, N_2 from X, Y respectively. Let $\bar{X} = N_1^{-1} \sum_1^{N_1} X_i$, $\bar{Y} = N_2^{-1} \sum_1^{N_2} Y_j$, $S_1 = \sum_1^{N_1} (X_i - \bar{X})(X_i - \bar{X})'$, $S_2 = \sum_1^{N_2} (Y_j - \bar{Y})(Y_j - \bar{Y})'$, $N = N_1 + N_2$. Several invariant test criteria known for this problem are (i) a test based on $|S_2|/|S_1|$, (ii) a test based on $\mathrm{tr}\, S_1 S_2^{-1}$, (iii) a test based on the largest and the smallest characteristic roots of $S_2 S_1^{-1}$ (Roy (1953)), (iv) a test based on $|S_1 + S_2|/|S_2|$ (Kiefer and Schwartz (1965)) and (v) a test based on $\mathrm{tr}\, S_2(S_1 + S_2)^{-1}$ (Pillai (1955)). Giri (1968) has shown that the test which rejects H_0 for small values of $\mathrm{tr}\, S_2(S_1 + S_2)^{-1}$ (or large values) (Pillai's test) is locally best invariant (LBI) for testing H_0 against $H_1 : \sigma = \mathrm{tr}(\Sigma_1 \Sigma_2^{-1} - I) > 0$ (or < 0) as $\sigma \to 0$. From Anderson and Das Gupta (1964) it follows that the power function of each of the above tests is a monotonically increasing function in each of the characteristic roots of $\Sigma_1 \Sigma_2^{-1}$.

* Research supported by the Canadian N.S.E.R.C. Grant.

Cohen and Marden (1988) have obtained the minimal complete class of invariant tests and it follows from their results or from the essential uniqueness of its local properties that Pillai's test is also admissible for our problem. We shall show here that the LBI test is locally minimax in the sense of Giri and Kiefer (1964) but it is not of type D (Isaacson (1951)).

The property of being an LBI test under the affine group $G_l(p) \times R^p$, which is possessed by Pillai's test, is unsatisfactory because this group, with $p \geq 2$, does not satisfy the conditions of the Hunt-Stein Theorem. The only satisfactory property known to us at this writing is the admissibility of the Kiefer-Schwartz test, the LBI property and the admissibility of Pillai's test.

2. Locally minimax test

We use the theory of local minimax results contained in Giri and Kiefer (1964). In deriving the local minimax results of $H_0 : \Sigma_1 = \Sigma_2$ we first observe that the full linear group $G_l(p)$, which leaves the problem invariant, does not satisfy the conditions of the Hunt-Stein Theorem. Since the subgroup $G_T(p)$ of $p \times p$ nonsingular lower triangular matrices satisfies the conditions of the theorem, we shall use the group $G = G_T(p) \times R^p$ for the invariance of the problem and find the probability ratio R of the distributions of a maximal invariant under G for testing H_0 against H_1, by using a result of Wijsman (1967). The techniques developed so far, to establish the local minimax property, are mainly for one population problems and depend on the derivation of the exact distribution of the maximal invariant relevent to the problem. For two populations testing problems concerning covariance matrices the distribution of the maximal invariant under $G_T(p)$ is much more involved. Using Wijsman (1967) and integrations over groups we are able to establish here the local minimax property of Pillai's test.

Let $Z = T(X)$ be a maximal invariant in the space of $(\bar{X}, \bar{Y}, S_1, S_2)$ under G transforming $(\bar{X}, \bar{Y}, S_1, S_2) \to (g\bar{X} + b_1, g\bar{Y} + b_2, gS_1g', gS_2g')$ with $g = (g_{ij}) \in G_T(p)$ and $b_1, b_2 \in R^p$. Writing $G_T(p) = G_T$, the probability ratio R of the distribution of $T(X)$ can be written as

$$(2.1) \quad R = \frac{\int_{G_T} |\Sigma_1^{-1}|^{\frac{N_1-1}{2}} |\Sigma_2^{-1}|^{\frac{N_2-1}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}(\Sigma_1^{-1} g s_1 g' + \Sigma_2^{-1} g s_2 g')\right\} \prod_1^P (g_{ii}^2)^{\frac{N-2-i}{2}} dg}{\int_{G_T} |\Sigma_1^{-1}|^{\frac{N-2}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\Sigma_1^{-1}(g s_1 g' + g s_2 g')\right\} \prod_1^p (g_{ii}^2)^{\frac{N-2-i}{2}} dg}$$

where $dg = \prod_{i \leq j} dg_{ij}$. It may be remarked that a left invariant Haar measure on $G_T(p)$ is $\prod_1^p (g_{ii}^2)^{-i/2} dg$ (see Giri (1977)). In what follows we write $A^{1/2}$ as a lower triangular nonsingular matrix such that $A = A^{1/2}A^{1/2'} > 0$, A symmetric and $A^{-1/2} = (A^{1/2})^{-1}$. Let $\Delta = \Sigma_2^{-1/2}\Sigma_1\Sigma_2^{-1/2'}$, $\theta = \Delta - I$ and $\sigma = \operatorname{tr}\theta$. In terms of σ our testing problem reduces to testing $H_0 : \sigma = 0$ against $H_1 : \sigma > 0$. The dual alternative $\sigma < 0$ can be reduced to H_1 by interchanging the roles of X's and Y's. We shall assume throughout that $N_i > p$, $i = 1, 2$, so that S_1, S_2 are positive definite with probability one.

For each (σ, η) in the parametric space of the distribution of Z let $p(Z; \sigma, \eta)$ be the probability density function of Z with respect to some σ-finite measure u. For each α, $0 < \alpha < 1$ consider a critical region of the form $\tilde{R} = \{Z : U(Z) \leq C_\alpha\}$ where U is bounded, positive and has a continuous distribution function for each (σ, η), equicontinuous for some $\sigma < \sigma_0$ (fixed) and

$$P_{0,\eta}(\tilde{R}) = \alpha, \qquad P_{\lambda,\eta}(\tilde{R}) = \alpha + h(\lambda) + q(\lambda, \eta) \tag{2.2}$$

where $q(\lambda, \eta) = o(h(\lambda))$ uniformly in η with $h(\lambda) > 0$ for $\lambda > 0$ and $h(\lambda) = o(1)$. Throughout this section notations like $o(1)$, $o(h(\lambda))$ are to be interpreted as $\lambda \to 0$. Let ξ_0, ξ_λ denote the a priori probability measure on the set $\sigma = 0$ and $\sigma = \lambda$ respectively. Let us assume that there exist ξ_0 and ξ_λ such that

$$\frac{\int p(z; \lambda, \eta) d\xi_\lambda(d\eta)}{\int p(z; 0, \eta) d\xi_0(d\eta)} = 1 + h(\lambda)(g(\lambda) + r(\lambda)U(z)) + B(z, \lambda) \tag{2.3}$$

where $0 < c_i < r(\lambda) < c_2 < \infty$ for λ sufficiently small and $g(\lambda) = O(1)$ and $B(z, \lambda) = o(h(\lambda))$. Then $\tilde{R}$ is locally minimax in the sense of Giri and Kiefer (1964) for testing $H_0 : \sigma = 0$ against $H_i : \sigma = \lambda$ as $\lambda \to 0$. If the set $\sigma = 0$ is a single point, by letting ξ_0 give measure one to the single point we can rewrite the left-hand side of (2.3) as

$$\int \frac{p(z; \lambda, \eta)}{p(z; 0, \eta)} \xi_\lambda(d\eta). \tag{2.4}$$

In our search for a locally minimax test of $H_0 : \sigma = 0$ against $H_i : \sigma = \lambda > 0$ through the Hunt-Stein theorem we consider the subgroup $G = G_T(p) \times R^p$ for the invariance of the problem. Since G satisfies the conditions of the theorem we conclude that a test which is locally minimax among G-invariant level α tests is locally minimax among all level α tests.

Let $V = (V_{ij}) = S_2^{-1/2} S_1 S_2^{-1/2}$. Then R can be written as

$$R = \beta^{-1} \int_{G_T} |\Delta|^{(N_2-1)/2} \exp\left\{-\frac{1}{2}\operatorname{tr}(gvg' + \Delta gg')\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg \tag{2.5}$$

where

$$\beta = \int_{G_T} \exp\left\{-\frac{1}{2}\operatorname{tr}(g(v + I)g')\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg.$$

Since $|\Delta|^{(N_2-1)/2} = 1 + (N_2 - 1)\sigma/2 + O(\sigma)$ and the set $\sigma = 0$ is a single point in this case, we can rewrite the left-hand side of (2.3) as $\int R\xi_\lambda(d\eta)$. By Taylor expansion of $\exp\{-\operatorname{tr}(gvg' + \Delta gg')/2\}$ around $\sigma = 0$ $(\Delta = I)$ we obtain

$$\begin{aligned} R = {} & 1 + \frac{1}{2}\sigma(N_2 - 1) \\ & - \frac{1}{2}\beta^{-1} \int \operatorname{tr}(\Delta gg') \exp\left\{-\frac{1}{2}\operatorname{tr}(g(v + I)g')\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg \\ & + O(\sigma). \end{aligned} \tag{2.6}$$

To simplify further we need the following lemma whose proof is straightforward and hence is omitted.

LEMMA 2.1.

(a) $$\int_{G_T} g_{ij} \exp\left\{-\frac{1}{2}\operatorname{tr} gg'\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg = 0,$$

(b) $$\beta^{-1}\int_{G_T} g_{ij}g_{i'j'} \exp\left\{-\frac{1}{2}\operatorname{tr} gg'\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg$$
$$= \begin{cases} 1, & \textit{if } (i,j) = (i'j')(i \neq j), \\ 0, & \textit{if } (i,j) \neq (i',j'), \end{cases}$$

(c) $$\beta^{-1}\int_{G_T} g_{ii}^2 \exp\left\{-\frac{1}{2}\operatorname{tr} gg'\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg = (N-i-1),$$

(d) $$\beta^{-1}\int gg' \exp\left\{-\frac{1}{2}\operatorname{tr} gg'\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg = D$$

where $D = (d_{ij})$ is a diagonal matrix with diagonal elements

$$d_{ii} = (N-i-1)w_{ii} + \sum_{j<i} w_{ij}$$

where $W = (w_{ij}) = (v+I)^{-1}$.

Let $\Delta - I = \theta = (\theta_{ij})$. Using Lemma 2.1 we get,

LEMMA 2.2.

$$\beta^{-1}\int_{G_T} \operatorname{tr}(\theta gg') \exp\left\{-\frac{1}{2}\operatorname{tr} g(v+I)^{-1}g'\right\} \prod_1^p (g_{ii}^2)^{(N-i)/2-1} dg$$
$$= \sum_1^p w_{ii}\left[(N-i-1)\theta_{ii} + \sum_{j>i}\theta_{jj}\right]$$
$$= \sigma\sum_1^p w_{ii}\left[(N-i-1)\eta_{ii} + \sum_{j>i}\eta_{jj}\right].$$

Let $\eta = (\eta_{11}, \ldots, \eta_{pp})$. From (2.5) using the above lemmas we obtain

$$\int R\xi_\lambda(d\eta) = 1 + \frac{\lambda}{2}\left(N_2 - 1 - \sum_1^p w_{ii}\left[(N-i-1)\eta_{ii}^0 + \sum_{j>i}\eta_{jj}^0\right]\right) + o(\sigma)$$

where ξ_λ assigns measure one to the single point $\eta^0(\text{say}) = (\eta_{11}^0, \ldots, \eta_{pp}^0)$ whose j-th coordinate is

$$\eta_{jj}^0 = (N-2-j)^{-1}(N-2-j+1)^{-1}(N-2)(N-2-p)$$

so that

$$\sum_{j>i} \eta_{jj}^0 + (N-1-i)\eta_{ii}^0 = \frac{N-2}{p}.$$

Hence we get

$$\int R\xi_\lambda(dn) = 1 + \frac{\lambda}{2}\left[N_2 - 1 - \frac{(N-2)}{p}\operatorname{tr} S_2(S_1S_2)^{-1}\right] + B(w,\lambda,\eta) \tag{2.7}$$

where $B(w,\lambda,\eta) = o(\lambda)$ uniformly in w, η. Giri (1968) has shown that the test which rejects H_0 for small values of $U(Z) = \operatorname{tr} S_2(S_2+S_1)^{-1}$ is LBI for testing H_0 against $H_i : \sigma = \lambda$ as $\lambda \to 0$ and its power function is given by $\alpha + h(\lambda) + o(\lambda)$ where $h(\lambda) = b\lambda$ with b a positive constant. Thus from (2.7) we obtain

THEOREM 2.1. *For every p, N_1, N_2 and α Pillai's test which rejects H_0: $\theta = 0$ for small values of $\operatorname{tr} S_2(S_2+S_1)^{-1}$ is locally minimax against H_i: $\sigma = \lambda > 0$ as $\lambda \to 0$.*

3. Type D and Type E regions

The notion of a type D or E region is due to Isaacson (1951). Lehmann (1959) showed that in finding a type D region invariance could be invoked in the manner of the Hunt-Stein theorem and this can be done with type E region provided that one works with the group of transformations which operates as the identity on the nuisance parameter space. We refer to Giri and Kiefer (1964) for details about these regions and the theory of associated D_A and D_M regions. We will only show here that Pillai's test is not type D or E among G-invariant level α tests and hence it is not of type D_A or D_M among all level α tests. In the notation of Giri and Kiefer (1964) we write $\theta = \Delta - I$, $\eta = \Sigma_2$. From (2.6) the power function of a G-invariant level α test ϕ can be written as

$$\beta_\phi(\theta,\eta) = \alpha\Big(1 + \sum_1^p \theta_{ii}\Big) - \sum \theta_{ii} \sum_{j\le i} a_{ij} \int \phi(w) w_{jj} p(w;0,\eta) u(dw) + o(\delta) \tag{3.1}$$

where $R = p(w;\lambda,\eta)/p(w;0,\eta)$ and $a_{ij} = 1$ if $i > j$, $= 0$ if $i < j$, $= (N-j-1)$ if $i = j$ respectively.

For testing $H_0 : \theta = 0$ against $H_i : \theta \ne 0$ let $B_\phi(\eta)$ be the second derivative of $B_\phi(\theta,\eta)$ with respect to θ_i, $i = 1,2,\dots,p$ at $\theta = 0$ where $\theta_{ii} = \theta_i^2$. Let $\Delta\phi(\eta) = |B_\phi(\eta)|$. Since $B_\phi(\theta,\eta)$ does not depend on η, $B_\phi(\eta) = B_\phi$ and from (3.1) B_ϕ is a diagonal matrix whose i-th diagonal element is $\alpha + E_0(\sum_{j<i} a_{ij} W_{jj}\phi(W))$, where E_0 denotes the expectation under H_0. By Lemma 2 of Giri and Kiefer (1964) for any arbitrary matrix Q with q_i as its i-th diagonal element, $\operatorname{tr} QB_\phi$ is maximized over such ϕ by a ϕ^* of the form

$$\phi^*(W) = \begin{cases} 1, & \sum_i \sum_{j\le i} a_{ij} q_i W_{jj} \le C, \\ 0, & \text{otherwise}. \end{cases} \tag{3.2}$$

Pillai's test which rejects H_0 for small values of $\operatorname{tr} W$ has power function of the form $\alpha + b\sum_i \theta_{ii} + o(\sigma)$ with $b > 0$. The matrix B_ϕ of this test is I. When all q_i's are equal the critical region (3.2) reduces to $\sum_i \sum_{j\le i} a_{ij} W_{jj} \le C$, which is not the Pillai's critical region. Hence we obtain

THEOREM 3.1. *For $0 < \alpha < p < N_i$, $i = 1, 2$. Pillai's test is not of type D among G-invariant tests and hence it is not of type D_A or D_M among all tests.*

REFERENCES

Anderson, T. W. and Das Gupta, S. (1964). A monotonicity property of the power functions of some tests of the equality of two covariance matrices, *Ann. Math. Statist.*, **35**, 1059–1063.

Cohen, A. and Marden, J. I. (1988). Minimal complete classes of invariant tests for equality of normal covariance matrices and sphericity, *J. Multivariate Anal.*, **27**, 131–150.

Giri, N. (1968). On tests of the equality of two covariance matrices, *Ann. Math. Statist.*, **39**, 275–277.

Giri, N. (1977). *Multivariate Statistical Inference*, Academic Press, New York.

Giri, N. and Kiefer, J. (1964). Local and asymptotic minimax properties of multivariate tests, *Ann. Math. Statist.*, **35**, 21–35.

Isaacson, S. L. (1951). On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters, *Ann. Math. Statist.*, **22**, 217–234.

Kiefer, J. and Schwartz, R. (1965). Admissible Bayes character of T^2, R^2—and other fully invariant tests for classical multivariate normal problems, *Ann. Math. Statist.*, **36**, 747–770.

Lehmann, E. (1959). *Testing Statistical Hypotheses*, Wiley, New York.

Pillai, K. C. S. (1955). Some new test criteria in multivariate analysis, *Ann. Math. Statist.*, **26**, 117–121.

Roy, S. N. (1953). On a heuristic method of test construction and its use in multivariate analysis, *Ann. Math. Statist.*, **24**, 220–238.

Wijsman, R. A. (1967). Cross-Sections of orbits and their application to densities of maximal invariants, *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, Vol. 1, 389–400, University of California Press, Berkeley.