JOURNAL OF MULTIVARIATE ANALYSIS 8, 134–140 (1978)

A Locally Most Powerful Invariant Test for the Equality of Means Associated with Covariate Discriminant Analysis

TAKEAKI KARIYA

Hitotsubashi University, Tokyo, Japan

AND

MITSUYO KANAZAWA*

Osaka University, Osaka, Japan

Communicated by J. N. Srivastava

In this paper, the authors propose a locally most powerful invariant test for the equality of means in the presence of covariate variables. Also the null and nonnull distributions associated with the above test are developed. This problem arises in covariate discriminant analysis and has been treated by various authors, notably Cochran and Bliss (1948, *Ann. Math. Statist.* 19, 151–176) and Rao (1949, Sankhyā 9 343–366; 1966). The test derived here locally dominates in power the tests proposed so far. It is also shown that the Cochran–Bliss test is uniformly most powerful in the class of conditional invariant tests.

1. INTRODUCTION

Let X_i's and Y_j's be random samples from $(p+q)$-variate normal distributions $N_{p+q}(\mu_x, \Sigma)$ and $N_{p+q}(\mu_y, \Sigma)$ respectively ($i = 1,..., n$; $j = 1,..., m$) where X_i's and Y_j's are independent. Define $\bar{X} = \sum X_i/n$, $\bar{Y} = \sum Y_j/m$, $d = \bar{X} - \bar{Y}$, $\delta = \mu_x - \mu_y$ and $S = [\sum (X_i - \bar{X})(X_i - \bar{X})' + \sum (Y_j - \bar{Y})(Y_j - \bar{Y})']/(n + m - 2)$. Suppose that the X_i's and Y_j's consist of the subvectors of p-main variables and q-covariables as $X_i = (X_{1i}', X_{2i}')'$ and $Y_j = (Y_{1j}', Y_{2j}')'$. Hence we partition $\bar{X}$, $\bar{Y}$, d, δ, S and Σ correspondingly as $\bar{X} = (\bar{X}_1', \bar{X}_2')'$, $d = (d_1', d_2')'$, etc. Further we assume that $n + m \geqslant p + q + 1$.

Received March 12, 1976; revised September 23, 1977

AMS 1970 Subject classification: Primary 62H30.

Key words and phrases: Discriminant analysis with covariates, testing the equality of mean vectors, a locally most powerful invariant test.

* Currently at Fukui University, Fukui-ken, Japan.

0047-259X/78/0081-0134$02.00/0

The problem studied here is one of testing H: $\delta_1 = 0$ under the condition $\delta_2 = 0$. This has been discussed by various authors, associated with covariate discriminant analysis. Especially Cochran and Bliss (1948) and Rao (1949, 1966) proposed the following tests:

$$T_1 = [c(D^2_{p+q} - D_q^2)/(n + m - 2 + cD_q^2)] > k_1 \quad \text{(Cochran–Bliss test)}, \tag{1.1}$$

$$T_2 = [c(D^2_{p+q} - D_q^2)/(n + m - 2)] > k_2 \quad \text{(Rao test)}, \tag{1.2}$$

where $c = nm/(n + m)$, $D^2_{p+q} = d'S^{-1}d$ and $D_q^2 = d_2'S_{22}^{-1}d_2$. As will be shown, the Cochran–Bliss test is equivalent to the LRT (likelihood ratio test). Subrahmaniam and Subrahmaniam (1973, 1976) have studied the relative merits of the procedures.

In this paper, as an alternative test, an LMPI (locally most powerful invariant) test is derived. Hence in a neighborhood of the hypothesis H: $\delta_1 = 0$, this test dominates in power the above tests which are both invariant. The null and nonnull distributions of the test are also obtained, and the LRT is shown to be UMP (uniformly most powerful) in the class of conditional invariant tests.

2. An LMPI (Locally Most Powerful Invariant) Test

Without loss of generality, we consider the problem in terms of a sufficient statistic $(\bar{X}, \bar{Y}, S)$ and in this section we study it through invariance. Consider the group $G = \mathscr{A} \times R^{p+q}$, where

$$\mathscr{A} = \left\{ A : (p+q) \times (p+q) \,\middle|\, A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \right.$$
$$\left. A_{12} : p \times q, \; |A_{ii}| \neq 0 \; (i = 1, 2) \right\},$$

which leaves the problem invariant by the actions; for $(A, b) \in G$, $(\bar{X}, \bar{Y}, S) \to (A\bar{X} + b, A\bar{Y} + b, ASA')$, and $(\mu_x, \mu_y, \Sigma) \to (A\mu_x + b, A\mu_y + b, A\Sigma A')$. Here it is well to refer to Giri (1964a) in which in our terms the problem of testing $\tau_1 = 0$ under $\tau_2 = 0$ is treated through invariance where $\tau = \Sigma^{-1}\delta$ and $\tau = (\tau_1', \tau_2')'$ and the LRT is shown to be UMP similar invariant. In our problem such a stronger result cannot be expected as will be shown and an LMPI test we derive below is not the LRT. But the proof of the next lemma is omitted since it is similar to that in Giri (1964b).

Lemma 1. *Under the group G a maximal invariant is (D^2_{p+q}, D_q^2) and a maximal invariant parameter is $\lambda = \delta_1'\Sigma_{11.2}^{-1}\delta_1$ where $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.*

A more convenient choice of a maximal invariant is $(W_1, W_2) \equiv (T_1, cD_q^2/(n + m - 2))$ which is in one–one correspondence with (D^2_{p+q}, D_q^2), where

T_1 is given in (1, 1). Now by invariance our problem is to test $H: \lambda = 0$ versus $K: \lambda > 0$ and the class of level α tests based on (W_1, W_2), denoted by D^I, is only considered $(0 < \alpha < 1)$. Let $F'(\gamma: a, b)$ denote a noncentral F-distribution with degrees of freedom (a, b) and noncentrality parameter γ, and let $F(a, b) = F'(0: a, b)$, $r = n + m - p - q - 1$ and $s = n + m - q - 1$. Then, conditional on $W_2 = w_2$, $U_1 = rW_1/p \sim F'(\zeta(w_2): p, r)$, and unconditionally $U_2 = sW_2/q \sim F(q, s)$ where $\zeta(w_2) = c\lambda/(1 + w_2)$ (see Rao (1966) and Subrahmaniam and Subrahmanian (1973)). Hence under $H: \lambda = 0$ or $\delta_1 = 0$ given $\delta_2 = 0$, W_1 and W_2 are independent and clearly $U_1 \sim F(p, r)$. Here we prove

THEOREM 1. *A unique* LMPI *test is given by the critical region*

$$T_3 = aW_1(1 + W_1)^{-1}(1 + W_2)^{-1} - (1 + W_2)^{-1} > k_3 \quad \textit{where } a = s/p. \tag{2.1}$$

Proof. Let $f(u_1, u_2: \lambda) = g(u_1: u_2, \lambda)\, h(u_2)$ be the joint density of $U_1 = rW_1/p$ and $U_2 = sW_2/q$, where $g(u_1: u_2, \lambda)$ is the conditional density of $F'(\zeta(w_2): p, r)$ and $h(u_2)$ the density of $F(q, s)$ (see, e.g., Johnson and Kotz (1970) for the explicit forms of these densities). Then as in Ferguson (1967, pp. 235–239), an LMP test is given by $[\partial \log f(u_1, u_2: \lambda)/\partial\lambda] |_{\lambda=0} > k$. Evaluating this yields (2.1). The uniqueness follows from Lehmann (1959, Theorem 5, p. 84).

Hence the test (2.1) locally dominates in power the Cochran–Bliss test and the Rao test since both are invariant. In terms of (W_1, W_2) the Rao test is

$$T_2 = W_1(1 + W_2) > k_2. \tag{2.2}$$

3. SOME PROPERTIES OF THE THREE TESTS

PROPOSITION 1. *The Cochran–Bliss test* (1.1) *is equivalent to the* LRT *and the cut-off point is determined from* $F(p, r)$.

Proof. It is easy to show that the LRT is equivalent to

$$| cdd' + (n + m - 2)S |/| cd_2d_2' + (n + m - 2)\, S_{22} |\, | S_{11.2} | > k \quad \text{or}$$

$$\{(n + m - 2) + cD_{p+q}^2\}/\{(n + m - 2) + cD_q^2\} > k,$$

from which $T_1 > k_1$ follows where $S_{11.2} = S_{11} - S_{12}S_{22}^{-1}S_{21}$. The second part is clear from $U_1 \sim F(p, r)$ under H.

As studied in Subrahmaniam and Subrahmaniam (1973), the Rao test is in general superior to the Cochran–Bliss test or the LRT. However this does not mean that the LRT is uniformly dominated in power by the Rao test. On the other hand, the LRT has the following optimality.

THEOREM 2. *The* LRT *is a* UMP *test in the class of conditional invariant tests*

$$D^I(w_1 \mid w_2) = \left\{ \phi \in D^I \;\middle|\; \int \phi(w_1, w_2) f_1(w_1)\, dw_1 \leqslant \alpha \text{ a.e. } (w_2) \right\} \tag{3.1}$$

where $f_1(w_1)$ *is the density of* W_1 *under* $H: \lambda = 0$. *Especially it is a* UMP *test in the class of level* α *tests based on* W_1 *alone.*

Proof. Recall that, conditional on $W_2 = w_2$, $U_1 = rT_1/p \sim F'(\zeta(w_2): p, r)$. Since $F'(\zeta(w_2): p, r)$ has a monotone likelihood ratio property in U_1, the critical region $W_1 > k(w_2)$ is UMP in $D^I(w_1 \mid w_2)$ from Lehmann (1959, p. 68, Theorem 2). But the cut-off point can be determined independently of w_2 since W_1 and W_2 are independent under H. This completes the proof.

We remark that there exists no unconditional UMP test in D^I since a most powerful test for testing $\lambda = 0$ versus $\lambda = \lambda_0$ (fixed) cannot be free from the fixed λ_0. Further we remark that the Rao test and the LMPI test do not belong to the class $D^I(w_1 \mid w_2)$. But the three tests have the following property in common.

PROPOSITION 2. *The power functions of the* LRT, *the Rao test and the* LMPI *test are strictly increasing.*

Proof. Although the Rao test and the LMPI test are not conditional tests of level α, conditional on $W_2 = w_2$, these are of the form $W_1 > k(w_2)$. On the other hand, the conditional distribution of $U_1 = rW_1/p$, i.e., $F'(\zeta(w_2): p, r)$, has the monotone likelihood ratio in U_1, hence the power function of any of the three tests is strictly increasing in $\zeta(w_2) = c\lambda/(1 + w_2)$ or in λ for each given w_2. Since the distribution of W_2 gives a positive mass to any nonempty open set, the power functions of the three tests are unconditionally strictly increasing in λ. (See Lehmann (1959, p. 68, Theorem 2 and p. 312, Problem 4).)

As mentioned above, a close numerical comparison between the LRT and the Rao test has shown the superiority of the Rao test. But the Rao test is locally dominated by the LMPI test. Hence it is necessary to compare these two tests globally. However, this problem is hard to analytically treat and is left here.

4. THE NULL AND NONNULL DISTRIBUTIONS OF THE LMPI TEST

The null and nonnull distributions of the LRT and the Rao test are given in Rao (1966) and Subrahmaniam and Subrahmaniam (1973, 1976). Here we consider those of the LMPI test. Let $V_1 = W_1/(1 + W_1)$ and $V_2 = 1/(1 + W_2)$. Then the LMPI test can be written as $T_3 = (aV_1 - 1)V_2 > k_3$, and given $V_2 = v_2$, $V_1 \sim Be'(c\lambda v_2\,; p/2, r/2)$ and $V_2 \sim Be(s/2, q/2)$ where $Be'(\gamma: \alpha, \beta)$ denotes a noncentral beta distribution with noncentrality parameter γ and

$Be(\alpha, \beta)$ denotes a beta distribution. For $rW_1/p \sim F'(\zeta(w_2): p, r)$ for a given $W_2 = w_2$, and $sW_2/q \sim F(q, s)$. Let $b(z: \alpha, \beta) = [B(\alpha, \beta)]^{-1} z^{\alpha-1}(1-z)^{\beta-1}$, $Po(j : \gamma) = \gamma^j e^{-\gamma}/j!$ and $\tau = c\lambda/2$ where $B(\alpha, \beta)$ denotes a beta function. Then the joint density of V_1 and V_2 is

$$h(v_1, v_2 : \tau) = \sum_{k=0}^{\infty} Po(k : \tau v_2)\, b\left(v_1 : \frac{p}{2} + k, \frac{r}{2}\right) b\left(v_2 : \frac{s}{2}, \frac{q}{2}\right) \quad (0 < v_1, v_2 < 1) \tag{4.1}$$

(see Johnson and Kotz (1970)). To evaluate $H(x : \tau) \equiv P((aV_1 - 1)V_2 \leqslant x : \tau)$, we distinguish the cases: (1) $x \geqslant a - 1$, (2) $0 \leqslant x < a - 1$, (3) $-1 \leqslant x < 0$ and (4) $x < -1$. Note that $a = s/p = (n + m - q - 1)/p > 1$ and $Po(k : \tau v_2) = \sum_{j=0}^{\infty} e^{2\tau} Po(k : \tau)\, Po(j : \tau)(-1)^j v_2^{k+j}$. The following results are easily obtained by integrating (4.1) over each region.

Case (1) $x \geqslant a - 1$: $H(x : \tau) = 1$ and so $H(x : 0) = 1$.

Case (2) $0 \leqslant x < a - 1$: $H(x : \tau) = K_1(x : \tau) + K_2(x : \tau)$ where with $I(z: \alpha, \beta) = \int_0^z b(t: \alpha, \beta)\, dt$

$$K_1(x : \tau) = \sum_k \sum_j e^{2\tau} Po(k : \tau)\, Po(j : \tau)(-1)^j \times \left[B\left(\frac{s}{2} + k + j, \frac{q}{2}\right) \Big/ B\left(\frac{s}{2}, \frac{q}{2}\right)\right] I\left(\frac{x+1}{a} : \frac{p}{2} + k, \frac{r}{2}\right) \tag{4.2}$$

and with $J(z: \alpha, \beta) = \int_1^z t^{\alpha-1}(t-1)^{\beta-1}\, dt$ and $l = s/2 + k + j$

$$K_2(z : \tau) = \sum_k \sum_j \sum_\alpha \sum_\beta e^{2\tau} Po(k : \tau)\, Po(j : \tau)(-1)^{j+\alpha} (\alpha + l)^{-1} x^{\alpha+l} a^{-(l+\alpha+\beta)} \binom{q/2 - 1}{\alpha} \times \binom{\alpha + \beta + l - 1}{\beta} \left[B\left(l, \frac{q}{2}\right) B\left(\frac{p}{2} + k, \frac{r}{2}\right)\right]^{-1} \times J\left(\frac{a}{x+1} : j + \alpha + \beta + 1, \frac{r}{2}\right). \tag{4.3}$$

Hence the null distribution is given by

$$H(x : 0) = I\left(\frac{x+1}{a} : \frac{p}{2}, \frac{r}{2}\right) + \sum_\alpha \sum_\beta \binom{q/2 - 1}{\alpha} \binom{\alpha + \beta + s/2 - 1}{\beta} (-1)^\alpha \left(\alpha + \frac{s}{2}\right)^{-1} x^{\alpha + s/2} \times a^{-s/2-\alpha-\beta} \left[B\left(\frac{s}{2}, \frac{q}{2}\right) B\left(\frac{p}{2}, \frac{r}{2}\right)\right]^{-1} J\left(\frac{a}{x+1} : \alpha + \beta + 1, \frac{q}{2}\right). \tag{4.4}$$

Case (3) $-1 \leqslant x < 0$: $H(x : \tau) = K_1(x : \tau) + K_3(x : \tau)$, where

$$K_3(x : \tau) = \sum_k \sum_j \sum_\alpha \sum_\beta e^{2\tau} Po(k : \tau)\, Po(j : \tau)(-1)^{j+\alpha} (\alpha + l)^{-1} (-x)^{\alpha+l}$$
$$\times a^\beta \binom{q/2 - 1}{\alpha}$$
$$\times \binom{\alpha + \beta + l - 1}{\beta} \left[B\left(\frac{p}{2} + k + \beta, \frac{r}{2}\right) \Big/ B\left(\frac{s}{2}, \frac{q}{2}\right) B\left(\frac{p}{2} + k, \frac{r}{2}\right)\right]$$
$$\times I\left(\frac{x+1}{a} : \frac{p}{2} + k + \beta, \frac{r}{2}\right). \tag{4.5}$$

Hence the null distribution is given by

$$H(x : 0) = I\left(\frac{x+1}{a} : \frac{p}{2}, \frac{r}{2}\right)$$
$$+ \sum_\alpha \sum_\beta (-1)^\alpha \left(\alpha + \frac{s}{2}\right)^{-1} (-x)^{\alpha + s/2} \binom{q/2 - 1}{\alpha} \binom{\alpha + \beta + s/2 - 1}{\beta}$$
$$\times \left[B\left(\frac{p}{2} + \beta, \frac{r}{2}\right) \Big/ B\left(\frac{s}{2}, \frac{q}{2}\right) B\left(\frac{p}{2}, \frac{r}{2}\right)\right] I\left(\frac{x+1}{a} : \frac{p}{2} + \beta, \frac{r}{2}\right) \tag{4.6}$$

Case (4) $x < -1$: $H(x : \tau) = 0$ and so $H(x : 0) = 0$.

For a given level α, the cut-off point k_3 is determined by $H(k_3 : 0) = \alpha$ and the power function of the LMPI test is given by $1 - H(x : \tau)$ with $\tau = c\lambda$. However, for the determination of k_3 and the evaluation of the power we have to wait for a numerical analysis.

Acknowledgment

The authors wish to thank the referee for his valuable comments.

References

Cochran, W. G. (1964). Comparison of two methods of handling covariates in discriminatory analysis. *Ann. Inst. Statist. Math.* **16** 43–55.

Cochran, W. G., and Bliss, C. I. (1948). Discriminant functions with covariance. *Ann. Math. Statist.* **19** 151–176.

Eaton, M. L. (1972). *Multivariate Statistical Analysis*. Institute of Math. Statist. University of Copenhagen, Denmark.

Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York.

GIRI, N. (1964a). On the likelihood ratio test of a normal multivariate testing problem. *Ann. Math. Statist.* **35** 181–190. Correction (1964).

GIRI, N. (1964b). On the likelihood ratio test of a normal multivariate testing problem II. *Ann. Math. Statist.* **36** 1061–1065.

JOHNSON, N. L., AND KOTZ, S. (1970). *Continuous Univariate Distributions* 2. Wiley, New York.

KSHIRSARGAR, A. M. (1972). *Multivariate Analysis*, pp. 200–203. Dekker, New York.

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

MEMON, A. Z., AND OKAMOTO, M. (1970). The classification statistic W^* in covariate discriminant analysis. *Ann. Math. Statist.* **41** 1491–1499.

RAO, C. R. (1949). On some problems out of discrimination with multiple character. *Sankhya* **9** 343–366.

RAO, C. R. (1950). A note on the distance of $D_{p+q}^2 - D_q^2$ and some computational aspects of D^2 statistic and discriminant functions. *Sankhya* **10** 257–268.

SUBRAHMANIAM, K., AND SUBRAHMANIAM, K. (1973). On the distribution of $(D_{p+q}^2 - D_q^2)$ statistic: percentage points and the power of the test. *Sankhya* **B 53** 51–78.

SUBRAHMANIAM, K. AND SUBRAHMANIAM, K. (1976). On the performance of some statistics in discriminant analysis based on covariates. *J. Multivar. Anal.* **6** 330–337.