SOME ASPECTS OF DISCRIMINATION FUNCTION COEFFICIENTS

By SOMESH DAS GUPTA*

Indian Statistical Institute

SUMMARY. The exact means and covariances of the sample discrimination function coefficients are computed under the assumption that the parent distribution is multivariate normal. The limiting distribution of these coefficients are obtained under the normality assumption and also for the general case. Various testing problems concerning hypotheses on the population discrimination function coefficients are introduced and some properties of the standard tests are discussed. In particular, it is shown that the power function of the step-down test for testing that all the coefficients are zero has a weak monotonicity property and does not have the usual monotonicity property.

1. Introduction

Fisher introduced the concept of linear discriminant function in order to distinguish between two sets of distributions which are generally taken to be normal distributions with the same covariance matrix. Various uses of this discriminant function are discussed by Rao (1965, 1966) and the relevant references are given therein. For one-population problem the coefficients of the linear discriminant function are usually taken to be proportional to the components of $\Sigma^{-1}\mu$ (with the same constant of proportionality), where μ and Σ are the mean vector and the covariance matrix, respectively, of the distribution concerned. In the sequel we shall take the constant of proportionality to be 1 and define the coefficients of the sample linear discriminant function as the components of $\hat{\Sigma}^{-1}\hat{\mu}$, where $\hat{\Sigma}$ and $\hat{\mu}$ are the standard unbiased estimates of Σ and μ, respectively. Analogous definitions are made for the two-sample case.

The first part of this paper is concerned with the computation of the exact means and covariances of the coefficients of the sample linear discrimination function (abbreviated as d.c.) under the assumption that the parent distribution is normal. Moreover, the large-sample distributions of these coefficients are obtained under the normality assumption and for the general case with some mild assumptions. Analogous results are obtained for the coefficients in the two-sample problem. These results may be used for deriving large sample tests on d.c.'s.

The second part of this paper deals with various testing problems concerning hypotheses on d.c. Some optimum tests are derived for some hypothesis-testing problems and a few properties of some standard tests are discussed. In particular, it is shown that the power function of the step-down test (Roy, 1958) for testing that all the d.c.'s are zero has a 'weak monotonicity' property and does not have the usual monotonicity property.

Families of simultaneous confidence bounds for d.c.'s are given in Rao (1966). This is not considered here. A few problems are also discussed.

---

* Now with the University of Minnesota.

2. The means and the covariances of the sample discrimination function coefficients

In the sequel, whenever we shall mention the distribution of a random symmetric matrix $S = [s_{ij}]$ we would mean the joint distribution of $s_{11}, s_{21}, s_{22}, s_{31}, s_{32}, s_{33}, \ldots, s_{p,p-1}, s_{pp}$. The following lemmas are of interest and they will be used in proving the main results.

Lemma 2.1 : *Let S be a $p \times p$ random matrix distributed according to the Wishart distribution $W(I_p; n, p)$ where I_p is the $p \times p$ identity matrix. For a positive integer k,*

(i) $E(S^k) = c(k, n, p) I_p$,

and, (ii) *if $E(S^{-k})$ exists then it is given by $E(S^{-k}) = d(k, n, p) I_p$; $c(k, n, p)$ and $d(k, n, p)$ are constants depending on k, n, p.*

Proof : It follows from symmetry that the diagonal elements of S^k have the same expectations and the diagonal elements of S^{-k} also have the same expectations. The above two results now follow from the fact that

$$E(S^k) = E(S^{*k}) = L'E(S^k)L,$$

$$E(S^{-k}) = E(S^{*-k}) = L'E(S^{*-k})L,$$

where $S^* = L'SL$ and L is any orthogonal matrix.

Lemma 2.2 : *Let S be a $p \times p$ random matrix distributed according to $W(\Sigma; n, p)$, where Σ is non-singular. Then*

(i) $E(S) = c(1, n, p)\Sigma$

(ii) $E(S^{-1}) = d(1, n, p)\Sigma^{-1}$

(iii) $E(S^{-1}\Sigma S^{-1}) = d(2, n, p)\Sigma^{-1}$

where c and d are defined as in Lemma 2.1.

Proof : Let $\Sigma = MM'$. Define $S^* = M^{-1}S(M')^{-1}$ and note that S^* is distributed according to $W(I_p; n, p)$. Now

$$E(S) = E(MS^*M') = c(1, n, p)MM' = c(1, n, p)\Sigma,$$

$$E(S^{-1}) = E[(M')^{-1}S^{*-1}M^{-1}] = d(1, n, p)(M')^{-1}M^{-1} = d(1, n, p)\Sigma^{-1}$$

$$E(S^{-1}\Sigma S^{-1}) = E[M')^{-1}S^{*-2}M^{-1}] = d(2, n, p)(M')^{-1}M^{-1} = d(2, n, p)\Sigma^{-1}.$$

Lemma 2.3 : *Let the constants $c(k, n, p)$ and $d(k, n, p)$ be defined as in Lemma 2.1. Then*

(i) $c(1, n, p) = n$,

(ii) $c(2, n, p) = n(n+p+1)$,

(iii) $d(1, n, p) = 1/(n-p-1)$, *if* $n-p-1 > 0$,

(iv) $d(2, n, p) = (n-1)/(n-p)(n-p-1)(n-p-3)$, *if* $n-p-3 > 0$.

Proof: Let S be distributed according to $W(I_p; n, p)$ and let $S = [s_{ij}]$.

(i) $c(1, n, p) = E(s_{11}) = n$,

(ii) $c(2, n, p) = E\left(\sum_{i=1}^{p} s_{1i}^2\right) = (n^2+2n)+n(p-1) = n(n+p+1).$

(iii) Let S and S^{-1} be partitioned as follows:

$$S = S_{[p]} = \begin{bmatrix} S_{[p-1]} & S_{p-1,p} \\ S'_{p-1,p} & s_{pp} \end{bmatrix}$$

$$S^{-1} = [s^{ij}] = \begin{bmatrix} S^{[p-1]} & S^{p-1,p} \\ (S^{p-1,p})' & s^{pp} \end{bmatrix}$$

It is known (Anderson, 1958) that

$$1/s^{pp} = s_{pp} - S'_{p-1,p} S_{[p-1]}^{-1} S_{p-1,p}$$

is distributed according to the χ^2 distribution with $n-p+1$ d.f.

In this connection, note the following result. If U is distributed according to the χ^2 distribution with r d.f., then for positive k, $E(U^{-k})$ exists when $2k < r$, and in that case

$$E(U^{-k}) = \Gamma\left(\frac{r}{2}-k\right) \Big/ \Gamma\left(\frac{r}{2}\right). 2^k.$$

Using this result, we get

$$d(1, n, p) = E(s^{pp}) = 1/(n-p-1), \text{ if } n-p-1 > 0.$$

(iv) Note that

$$d(2, n, p) = E\left[\sum_{i=1}^{p} (s^{pi})^2\right],$$

and

$$\sum_{i=1}^{p} (s^{pi})^2 = \frac{b'_{p.(p-1)} b_{p.(p-1)}+1}{(s^{pp})^{-2}},$$

where

$$b_{p.(p-1)} = S_{[p-1]}^{-1} S_{p-1,p}.$$

Hence

$$d(2, n, p) = E[b'_{p.(p-1)} b_{p.(p-1)}+1] E(s^{pp})^2$$

$$= E[\text{tr}(S_{[p-1]}^{-1})+1] E(s^{pp})^2$$

$$= \left[\frac{p-1}{n-p}+1\right] \frac{1}{(n-p-1)(n-p-3)},$$

if $n-p-3 > 0$. Finally

$$d(2, n, p) = (n-1)/(n-p)(n-p-1)(n-p-3)$$

if $n-p-3 > 0$.

Lemma 2.4 : *Let S be distributed according to $W(I_p; n, p)$, and $\mathbf{v}$ be a $p\times 1$ vector. Then*

(i) $$E(S\mathbf{v}\mathbf{v}'S) = n\lambda I_p+(n^2+n)\mathbf{v}\mathbf{v}'$$

(ii) $$E(S^{-1}\mathbf{v}\mathbf{v}'S^{-1}) = \frac{\lambda I_p+(n-p-1)\mathbf{v}\mathbf{v}'}{(n-p)(n-p-1)(n-p-3)}$$

if $n-p-3 > 0$, $\lambda = \mathbf{v}'\mathbf{v}$.

Proof : Let L be an orthogonal matrix such that

$$L\mathbf{v}\mathbf{v}'L' = \lambda\Lambda$$

where Λ is a diagonal matrix whose all the diagonal elements are zero except the last one which is equal to 1.

(i) It is easy to see that

$$E(S\mathbf{v}\mathbf{v}'S) = \lambda L'E(S\Lambda S)L,$$

$$E(S\Lambda S) = E\left[\begin{pmatrix} s_{1p} \\ \vdots \\ s_{pp}\end{pmatrix} (s_{1p} \ldots s_{pp})\right]$$

$$= nI_p+(n^2+n)\Lambda.$$

Hence $$E(S\mathbf{v}\mathbf{v}'S) = n\lambda L'L+(n^2+n)\lambda L'\Lambda L$$

$$= n\lambda I_p+(n^2+n)\mathbf{v}\mathbf{v}'.$$

(ii) Note that

$$E(S^{-1}\mathbf{v}\mathbf{v}'S^{-1}) = \lambda L'E(S^{-1}\Lambda S^{-1})L,$$

$$E(S^{-1}\Lambda S^{-1}) = E\left[\begin{pmatrix} s^{1p} \\ \vdots \\ s^{pp}\end{pmatrix} (s^{1p} \ldots s^{pp})\right]$$

$$= E\left[(s^{pp})^2\begin{pmatrix} b_{p.(p-1)}\, b'_{p.(p-1)} & -b_{p.(p-1)} \\ -b'_{p.(p-1)} & 1\end{pmatrix}\right]$$

$$= E[(s^{pp})^2]E\left[\begin{pmatrix} S^{-1}_{(p-1)} & 0 \\ 0 & 1\end{pmatrix}\right]$$

$$= \frac{1}{(n-p-1)(n-p-3)(n-p)}[I_p+(n-p-1)\Lambda].$$

Hence $$E(S^{-1}\mathbf{v}\mathbf{v}'S^{-1}) = \frac{\lambda I_p+\mathbf{v}\mathbf{v}'(n-p-1)}{(n-p)(n-p-1)(n-p-3)}.$$

Theorem 2.1 : *Let Y and S be distributed independently according to $N_p(\nu, I_p)$ and $W(I_p, n, p)$, respectively. Let $d = S^{-1}Y$. Then*

$$E(d) = \nu/(n-p-1), \quad \text{if} \quad n-p-1 > 0$$

and

$$\operatorname{cov}(d) = E[d-E(d)][d-E(d)]',$$

$$= \frac{(n-1)I_p+\nu\nu' I_p+\nu\nu'(n-p+1)/(n-p-1)}{(n-p)(n-p-1)(n-p-3)}$$

provided $n-p-3 > 0$.

Proof: $E(d) = E(S^{-1}Y) = E(S^{-1})E(Y) = \frac{1}{n-p-1}\cdot\nu$, provided $n-p-1 > 0$.

Moreover, from Lemma 2.4

$$E(dd') = E(S^{-1}YY'S^{-1}) = \frac{E(Y'Y)I_p+(n-p-1)\,E(YY')}{(n-p)(n-p-1)(n-p-3)}.$$

But

$$E(Y'Y) = p+\nu'\nu$$

$$E(YY') = \nu\nu'+I_p.$$

Hence

$$\operatorname{cov}(d) = E(dd')-E(d)E(d')$$

$$= \frac{(p+\nu'\nu)I_p+(n-p-1)(\nu\nu'+I_p)}{(n-p)(n-p-1)(n-p-3)} - \frac{\nu\nu'}{(n-p-1)^2}$$

$$= \frac{(n-1)I_p+\nu'\nu I_p+\nu\nu'(n-p+1)/(n-p-1)}{(n-p)(n-p-1)(n-p-3)}.$$

Corollary 2.1.1 : *Let X and S be distributed independently according to $N_p(\mu, \Sigma)$ and $W(\Sigma; n, p)$, respectively; Σ being assumed to be non-singular. Let $d^* = (S^*)^{-1}X$. Then*

(i) $E(d^*) = \Sigma^{-1}\mu/(n-p-1)$, *if* $n-p-1 > 0$

and

$$\text{(ii)} \quad \operatorname{cov}(d^*) = \frac{(\mu'\Sigma^{-1}\mu)\Sigma^{-1}+(n-1)\Sigma^{-1}+\dfrac{\Sigma^{-1}\mu\mu'\Sigma^{-1}(n-p+1)}{(n-p-1)}}{(n-p)(n-p-1)(n-p-3)}$$

provided $n-p-3 > 0$.

Proof: Let $\Sigma = \tau\tau'$. Define $Y = \tau^{-1}X$, $S = \tau^{-1}S^*(\tau')^{-1}$, $d = S^{-1}Y = \tau' d^*$ and $\nu = \tau^{-1}\mu$. From Theorem 2.1 we get

$$E(d^*) = E[(\tau')^{-1}d] = \frac{(\tau')^{-1}\nu}{n-p-1} = \frac{\Sigma^{-1}\mu}{n-p-1}$$

provided $n-p-1 > 0$. Moreover, from Theorem 2.1, we get

$$\text{cov}(d^*) = (\tau')^{-1} \text{cov}(d)\tau^{-1}$$

$$= \frac{(n-1)(\tau')^{-1}\tau^{-1}+(\nu'\nu)(\tau')^{-1}\tau^{-1}+\dfrac{(\tau')^{-1}\nu\nu'\tau^{-1}(n-p+1)}{(n-p-1)}}{(n-p)(n-p-1)(n-p-3)}$$

$$= \frac{(n-1)\Sigma^{-1}+(\mu'\Sigma^{-1}\mu)\Sigma^{-1}+\Sigma^{-1}\mu\mu'\Sigma^{-1}(n-p+1)/(n-p-1)}{(n-p)(n-p-1)(n-p-3)}$$

Corollary 2.1.2 : *Let* $X_1 ..., X_N$ *be a randm sample from* $N_p(\mu, \Sigma)$, Σ *being assumed to be non-singular. Let*

$$\bar{X} = \frac{1}{N}\sum_{\alpha=1}^{N} X_\alpha, \quad S = \sum_{\alpha=1}^{N}(X_\alpha-\bar{X})(X_\alpha-\bar{X})'$$

and the vector of sample discrimination coefficients be

$$d = nS^{-1}\bar{X},$$

$n = N-1$. *Then*

$$E(d) = \frac{n}{(n-p-1)}\Sigma^{-1}\mu, \quad \textit{if} \quad n-p-1 > 0,$$

and

$$cov(d) = \frac{n^2}{(n-p)(n-p-1)(n-p-3)}\left[\Sigma^{-1}(\mu'\Sigma^{-1}\mu)+\frac{n-1}{n+1}\Sigma^{-1}+\frac{\Sigma^{-1}\mu\mu'\Sigma^{-1}(n-p+1)}{n-p-1}\right.$$

provided $n-p-3 > 0$.

Corollary 2.1.3 : *Let* $X_1, ..., X_{N_1}$ *be a random sample from* $N_p(\mu_1, \Sigma)$ *and* $X_{N_1+1}, ..., X_{N_1+N_2}$ *be a random sample from* $N_p(\mu_2, \Sigma)$, Σ *being assumed to be non-singular. Let*

$$\bar{X}_{(1)} = \frac{1}{N_1}\sum_{\alpha=1}^{N_1} X_\alpha, \quad \bar{X}_{(2)} = \frac{1}{N_2}\sum_{\alpha=N_1+1}^{N_1+N_2} X_\alpha$$

$$S = \sum_{\alpha=1}^{N_1}(X_\alpha-\bar{X}_{(1)})(X_\alpha-\bar{X}_{(1)})'+\sum_{\alpha=N_1+1}^{N_1+N_2}(X_\alpha-\bar{X}_{(2)})(X_\alpha-\bar{X}_{(2)})'$$

and the vector of sample discrimination coefficients be

$$d = nS^{-1}(\bar{X}_{(1)}-\bar{X}_{(2)}),$$

$n = N_1+N_2-2$. *Then*

$$E(d) = \frac{n}{n-p-1}\Sigma^{-1}\mu, \quad \textit{if} \quad n-p-1 > 0$$

and

$$cov\,(d) = \frac{n^2[(\mu'\Sigma^{-1}\mu)\Sigma^{-1}+(n-1)m\Sigma^{-1}+\Sigma^{-1}\mu\mu'\Sigma^{-1}(n-p+1)/(n-p-1)]}{(n-p)(n-p-1)(n-p-3)}$$

provided $n-p-3 > 0$; $m = 1/N_1+1/N_2$, $\mu = \mu_1-\mu_2$.

The above two corollaries follow easily from Corollary 2.1.1 and their proofs are omitted.

3. Asymptotic distribution of the sample discrimination function coefficients

Theorem 3.1 : *Let $X_1, ..., X_N$ be a random sample from $N_p(\nu, I_p)$. Let*

$$\bar{X}_{(N)} = \frac{1}{N}\sum_{\alpha=1}^{N} X_\alpha, \quad S_{(N)} = \sum_{\alpha=1}^{N}(X_\alpha-\bar{X}_{(N)})(X_\alpha-\bar{X}_{(N)})',$$

$$d_N = nS_{(N)}^{-1}\bar{X}_{(N)}, \quad n = N-1.$$

Then the distribution of $\sqrt{N}(d_N-\nu)$ tends to $N_p[0, \nu\nu'+(1+\nu'\nu)I_p]$ as $N\to\infty$.

Proof: It can be seen that

$$\sqrt{N}(d_N-\nu) = nS_{(N)}^{-1}\sqrt{N}(\bar{X}_{(N)}-\nu)-(N/n)^{1/2}\sqrt{n}(n^{-1}S_{(N)}-I_p)(nS_{(N)}^{-1})\nu.$$

Note the following facts : As $N\to\infty$,

$$\mathcal{L}[\sqrt{N}(\bar{X}_{(N)}-\nu)]\to N_p(0, I_p), \qquad nS_{(N)}^{-1}\to I_p \text{ in probability,}$$

and

$$\mathcal{L}[\sqrt{n}(n^{-1}S_{(N)}-I_p)\nu]\to N_p(0, \nu\nu'+\nu'\nu I_p).$$

The last result follows from Anderson (1958, Theorem 4.2.5) and Lemma 2.4.

Corollary 3.1.1 : *Let $X_1, ..., X_N$ be a random sample from $N_p(\mu, \Sigma)$, Σ being assumed to be non-singular. Let d_N be defined as in Theorem 3.1. Then the distribution of $\sqrt{N}(d_N-\Sigma^{-1}\mu)$ tends to $N_p(0, \Gamma)$ as $N\to\infty$, where*

$$\Gamma = \Sigma^{-1}\mu\mu'\Sigma^{-1}+(1+\mu'\Sigma^{-1}\mu)\Sigma^{-1}.$$

Proof: Let $\Sigma = \tau\tau'$.

Define

$$X_i^* = \tau^{-1}X_i, \quad i = 1, ..., N,$$

$$S_{(N)}^* = \tau^{-1}S_{(N)}(\tau')^{-1},$$

$$\bar{X}_{(N)}^* = \tau^{-1}\bar{X}_{(N)}, \quad \nu = \tau^{-1}\mu$$

$$d_N^* = n(S_{(N)}^*)^{-1}\bar{X}_{(N)}^* = \tau' d_N.$$

It follows from Theorem 3.1 that the limiting distribution of $\sqrt{N}(d_N^*-\nu)$ is $N_p(0, \nu\nu'+(1+\nu'\nu)I_p)$. Hence the limiting distribution of

$$\sqrt{N}(d_N-\Sigma^{-1}\mu) = (\tau')^{-1}\sqrt{N}(d_N^*-\nu)$$

is N (0, Γ), where

$$\Gamma = (\tau')^{-1}[\nu\nu'+(1+\nu'\nu)I_p]\tau^{-1}$$
$$= \Sigma^{-1}\mu\mu'\Sigma^{-1}+(1+\mu'\Sigma^{-1}\mu)\Sigma^{-1}.$$

Now we shall derive the asymptotic distribution of d_N as defined in Theorem 3.1, without making the assumption that the underlying distribution is multivariate normal.

Theorem 3.2 : *Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. random vectors* ($p\times 1$) *with the common mean* ν *and the common covariance matrix* I_p. *Moreover, it is assumed that*

$$\tau_4(i, j, k, l) = E[X_{i\alpha}X_{j\alpha}X_{k\alpha}X_{l\alpha}]$$

exists for any i, j, k, l, where $X'_\alpha = (X_{1\alpha}, \ldots, X_{p\alpha})$. Let d_N be defined as in Theorem 3.1. Then the limiting distribution of $\sqrt{N}(d_N-\nu)$ is $N_p(0, \Gamma)$ as $N\to\infty$, where

$$\Gamma = (r_{ij}),$$

$$r_{ii} = 1+\sum_{j=1}^{p}\sum_{j'=1}^{p}\nu_j\nu_{j'}\tau_4(i, i, j, j')-2\sum_{j=1}^{p}\nu_j\tau_3(i, i, j),$$

$$r_{ii'} = \sum_{j=1}^{p}\sum_{j'=1}^{p}\nu_j\nu_{j'}\tau_4(i, i', j, j')-\nu_i\nu_{i'}-2\sum_{j=1}^{p}\nu_j\tau_3(i, i', j), \quad i\neq i',$$

$$\tau_3(i, j, k) = E[X_{i\alpha}X_{j\alpha}X_{k\alpha}], \quad \nu' = (\nu_1, \ldots, \nu_p).$$

Proof : We define $\bar{X}_{(N)}$, $S_{(N)}$ and d_N as in Theorem 3.1 and use the following relation :

$$\sqrt{N}(d_N-\nu) = nS_{(N)}^{-1}\sqrt{N}(\bar{X}_{(N)}-\nu)-\left(\frac{N}{n}\right)^{\frac{1}{2}}\sqrt{n}(n^{-1}S_{(N)}-I_p)(nS_{(N)}^{-1})\nu$$

Note that

$$n^{-1}S_{(N)} = \frac{1}{n}\sum_{\alpha=1}^{N}(X_\alpha-\nu)(X_\alpha-\nu)'-\frac{N}{n}(\bar{X}_{(N)}-\nu)(\bar{X}_{(N)}-\nu)'.$$

Since

$$(\bar{X}_{(N)}-\nu)(\bar{X}_{(N)}-\nu)'\to 0(p\times p) \text{ in probability,}$$

we have $\qquad n^{-1}S_{(N)}\to I_p$ in probability.

Hence $\qquad nS_{(N)}^{-1}\to I_p$ in probability.

Let $B_{(N)} = \sum_{\alpha=1}^{N}(X_\alpha-\nu)(X_\alpha-\nu)'$. Then $\sqrt{n}(n^{-1}S_{(N)}-I_p)$ and $\sqrt{n}(n^{-1}B_{(N)}-I_p)$ are asymptotically equivalent, since $\bar{X}_{(N)}-\nu$ is $0_p\left(\frac{1}{\sqrt{N}}\right)$ The required result is now obtained by computing the cov $[(n^{-1}S_{(N)}-I_p)\nu]$ and cov$[\bar{X}_{(N)}-\nu, (n^{-1}S_{(N)}-I_p)\nu]$.

Corollary 3.2.1 : *Let* $\{X_N\}$ *be a sequence of i.i.d. random vectors* $(p\times 1)$ *with the common mean* μ *and the common covariance matrix* Σ; Σ *being assumed to be nonsingular. Let* d_N *be defined as in Theorem 3.1. Then the distribution of* $\sqrt{N}(d_N-\Sigma^{-1}\mu)$ *tends to* $N_p(0, \Gamma^*)$ *as* $N\to\infty$, *where*

$$\Gamma^* = (\tau')^{-1}\Gamma\tau^{-1},$$

$\Sigma = \tau\tau'$, *and the matrix* Γ *be defined as in Theorem 3.2 with* $\nu = \tau^{-1}\mu$ *and* τ_4 *and* τ_3 *are replaced respectively by* τ_4^* *and* τ_3^*, *where*

$$\tau_4^*(i, j, k, l) = E[X^*_{i\alpha}X^*_{j\alpha}X^*_{k\alpha}X^*_{l\alpha}],$$

$$\tau_3^*(j, j, k) = E[X^*_{i\alpha}X^*_{j\alpha}X^*_{k\alpha}],$$

$$X^*_\alpha = \tau^{-1}X_\alpha.$$

The proof of Corollary 3.2.1 is the same as the proof of Corollary 3.1.1; use the results of Theorem 3.2 instead of those of Theorem 3.1.

4. Tests of hypotheses concerning discrimination coefficients

Let $X_1, \ldots, X_N$ be a random sample from the nonsingular $N_p(\mu, \Sigma)$. Define

$$\bar{X} = \frac{1}{N}\sum_{\alpha=1}^{N} X_\alpha,\quad S = \sum_{\alpha=1}^{N}(X_\alpha-\bar{X})(X_\alpha-\bar{X})' = [s_{ij}].$$

The joint density of $X_1, \ldots, X_N$ is

$$\frac{1}{(2\pi)^{Np/2}|\Sigma|^{N/2}}\exp\left[-\frac{N}{2}\operatorname{tr}\Sigma^{-1}\mu\mu'-\frac{1}{2}\operatorname{tr}\Sigma^{-1}\left(\sum_{\alpha=1}^{N}X_\alpha X_\alpha'\right)+N\gamma'\bar{X}\right],$$

where $(\gamma_1, \ldots, \gamma_p) = \gamma' = \mu'\Sigma^{-1}$. Let

$$\Sigma_{(t)} = \begin{bmatrix}\sigma_{11}\cdots\sigma_{1t}\\ \cdots\cdots\cdots\\ \sigma_{t1}\cdots\sigma_{tt}\end{bmatrix},\quad \mu_{(t)} = \begin{bmatrix}\mu_1\\ \vdots\\ \mu_t\end{bmatrix},\quad \Delta_t^2 = \mu'_{(t)}\Sigma_{(t)}^{-1}\mu_{(t)},$$

$$S_{(t)} = \begin{bmatrix}s_{11}\cdots s_{1t}\\ \cdots\cdots\\ s_{t1}\cdots s_{tt}\end{bmatrix},\quad \bar{X}_{(t)} = \begin{bmatrix}\bar{X}_1\\ \vdots\\ \bar{X}_t\end{bmatrix},\quad D_t^2 = \bar{X}'_{(t)}S_{(t)}^{-1}\bar{X}_{(t)}.$$

Let H denote the hypothesis to be tested and K the alternative hypothesis. We shall consider the following testing problems.

(a) $H : \gamma_i^{(1)} \leqslant \gamma_i \leqslant \gamma_i^{(2)}$; $K : \gamma_i > \gamma_i^{(2)}$ or $\gamma_i < \gamma_i^{(1)}$

$\gamma_i^{(1)}$ and $\gamma_i^{(2)}$ are specified constants and i is a fixed integer lying between 1 to p.

(b) $H : \gamma_i = \gamma_i^{(0)}$; $K : \gamma_i \neq \gamma_i^{(0)}$.

$\gamma_i^{(0)}$ is a specified constant and i is a fixed integer lying between 1 and p.

(c) $H : \sum_{i=1}^{p} b_i \gamma_i = b_0$; $K : \sum_{i=1}^{p} b_i\gamma_i \neq b_0$.

b_i's are specified constants.

(d) $H : \gamma_p = \gamma_{p-1} = \ldots = \gamma_{p-q+1} = 0$; K : not $H\,(1 \leqslant q \leqslant p)$.

It can be seen that the hypothesis H is equivalent to the hypothesis :

$$\Delta_p^2 = \Delta_{p-q}^2 (\Delta_0^2 = 0)$$

(e) $H : B\gamma = 0$; $K : B\gamma \neq 0$.

B is a known $q \times p$ matrix of rank q. The problem (d) is a special case of this problem.

(f) H : $0 \leqslant \Delta_p^2 - \Delta_{p-q}^2 \leqslant c_0$; K: $\Delta_p^2 - \Delta_{p-q}^2 > c_0$.

c_0 is a specified constant.

Without going into details we shall make comments on the existence of 'good' tests for the above problems. Note that the distributions of $X_1, \ldots, X_N$ constitute an exponential family. It now follows from Lehmann (1959) that the UMP unbiased test exists for each of the problems (a), (b) and (c). However, these tests are conditional tests except when $\gamma_i^{(0)} = 0$ in (b) and $b_0 = 0$ in (c). Giri (1964, 1965) considered a suitable invariant class of tests for the problem (d) and showed that the likelihood-ratio-test is UMP in the class of all similar and invariant tests. It is easily seen that the form of the problem (d) is obtained by the canonical reduction of the problem (e). There exists a matrix C such that the hypothesis in (e) is equivalent to $\gamma_p^* = \ldots = \gamma_{p-q+1}^* = 0$ where γ^* is the vector of discrimination function coefficients for the distribution $N_p(C\mu,\ C\Sigma C')$. The likelihood-ratio-test for the problem (e) is given in Rao (1965). For the problem (f) if we restrict our attention to the class of invariant tests as considered by Giri (1964) for the problem (d), then there exists a UMP unbiased test in this invariant class; however, this test is a conditional test, unless $c_0 = 0$.

4.1. *Properties of the power functions of certain tests.* First, we consider the problem (d). The likelihood-ratio-test for this problem rejects H iff

$$\left[(D_p^2 - D_{p-q}^2)\Big/\left(\frac{1}{N} + D_{p-q}^2\right)\right] > \frac{q}{N-p}\ F(\alpha; q, N-p) = A^*$$

where $F(\alpha; q, N-p)$ is the upper 100α percent point of the F-distribution with d.f. q and $N-p$. The conditional probability of the acceptance region, given D_{p-q}^2, is

$$I(D_{p-q}^2; \Delta_p^2 - \Delta_{p-q}^2) = \int_0^{A^*} f^*\left(y; q, N-p; \frac{\Delta_p^2 - \Delta_{p-q}^2}{\frac{1}{N} + D_{p-q}^2}\right) dy$$

where $f^*(\cdot\,; a, b; \tau^2)$ is the density of noncentral F-distribution with d.f. a and b and noncentrality parameter τ^2. Since the density of the noncentral F-distribution has the 'monotone likelihood ratio' property (Lehmann, 1959), I is a decreasing function of $\Delta_p^2-\Delta_{p-q}^2$ but an increasing function of D_{p-q}^2. It can be seen that the probability of the acceptance region is

$$I^*(\Delta_p^2-\Delta_{p-q}^2,\ \Delta_{p-q}^2) = \int_0^\infty I(D_{p-q}^2;\ \Delta_p^2-\Delta_{p-q}^2) f^*\left(D_{p-q}^2\,\frac{N-p+q}{p-q};\ p-q,\ N-p+q;\ \Delta_{p-q}^2\right) d\left(D_{p-q}^2\,\frac{N-p+q}{p-q}\right).$$

It follows now that for fixed $\Delta_p^2-\Delta_{p-q}^2$, I^* is an increasing function of Δ_{p-q}^2. Thus we arrive at the following theorem.

Theorem 4.1 : *The power function of the likelihood-ratio test for the problem (d) is an increasing function of $\Delta_p^2-\Delta_{p-q}^2$ but a decreasing function of Δ_{p-q}^2.*

Note : The UMP unbiased tests for the problem (b) with $\gamma_1^{(0)} = 0$ and for the problem (c) with $b_0 = 0$ coincide with the corresponding likelihood-ratio tests.

Next we consider the problem (d) with $q = p$. The hypothesis H is then equivalent to $\mu = 0$ and the step-down test (Rao, 1965) for this problem accepts H iff

$$F_i \leqslant F_i^0, \quad i = 1, \ldots, p.$$

where

$$F_i = \frac{D_i^2-D_{i-1}^2}{\frac{1}{N}+D_{i-1}^2}, \quad i = 1, \ldots, p \quad (D_0^2 = 0)$$

and F_i^0 are computed from the table of F-distribution and the size condition. The power function of such a test involves

$$\partial_i = \Delta_i^2-\Delta_{i-1}^2, \quad i = 1, \ldots, p,$$

as parameters. We shall investigate whether the power function of the step-down test increases as each ∂_i separately increases. We shall find that the step-down test does not have this strong property; however, it has some weaker properties regarding the monotonicity of the power function with respect to the parameters involved.

Theorem 4.2 : *The power function of the setp-down test for the problem (d) with $q = p$ is an increasing function of ∂_i when $\partial_1, \ldots \partial_{i-1}$ are fixed and*

$$\partial_p = \partial_{p-1} = \ldots = \partial_{i+1} = 0 \quad (i = 1, \ldots, p).$$

Proof : When $\partial_p = \partial_{p-1} = \ldots = \partial_{i+1} = 0$, $F_p, F_{p-1}, \ldots, F_{i+1}$ are distributed independently of $F_i, \ldots, F_1$ and

$$P[F_j \leqslant F_j^0,\ j = i+1, \ldots, p]$$

does not involve any of the parameters. Moreover, the conditional distribution of $(N-i)F_i$ given $F_1, \ldots, F_{i-1}$ is the non-central F-distribution with d.f. 1 and $N-i$ and non-centrality parameter $\partial_i \Big/ \left(\frac{1}{N}+D_{i-1}^2\right)$. Note that D_{i-1}^2 is function of $F_1, \ldots, F_{i-1}$. Since the F-distribution has the 'monotone likelihood ratio' property, we have

$$P[F_i < F_i^0 | F_1, \ldots, F_{i-1}; \partial_i] > P[F_i < F_i^0 | F_1, \ldots, F_{i-1}; \partial_i^*]$$

if $\partial_i^* > \partial_i$. Combining the above results, we get the theorem. The above result also follows from Das Gupta, Anderson and Mudholkar (1964); as a matter of fact, this is mentioned by Khatri (1966) who derived the result following Das Gupta, Anderson and Mudholkar (1964).

We shall now show that the power function of the step-down test is not an increasing function separately with respect to each ∂_i. To illustrate this fact we shall consider, for simplicity, $p = 2$.

Let α be the size of the step-down test. Let $F^0_{1,\alpha_1}(N-1)$ be the upper $100\alpha_1$ percent point of the F-distribution with d.f. 1 and $N-1$ and let $(N-2)F^0_{2,\alpha_2}$ be the upper $100\alpha_2$ percent point of the F-distribution with d.f. 1 and $N-2$; we take α_1 and α_2 such that $(1-\alpha_1)(1-\alpha_2) = 1-\alpha$. Thus, given α we may take any of the different combinations of (F_1^0, F_2^0) subject to the above size condition. If for every α_1 we have

$$P\left(F_1 < F^0_{1,\alpha_1}, F_2 < F^0_{2,\alpha_2}; \partial_1, \partial_2\right) > P\left(F_1 < F^0_{1,\alpha_1}, F_2 < F^0_{2,\alpha_2}; \partial_1^*, \partial_2\right) \quad \ldots \quad (1)$$

whenever $\partial_1^* > \partial_1$, then we shall arrive at a contradiction. When $\alpha_1 = 0$, $\alpha_2 = \alpha$, then

$$P\left(F_1 < F^0_{1,\alpha_1}, F_2 < F^0_{2,\alpha_2}; \partial_1, \partial_2\right)$$

strictly decreases if ∂_1 increases. This follows from Theorem 4.1. If we take the limits of both the sides in (1) as $\alpha_1 \to 0$, we get a contradiction.

However, we shall show that given $\partial_1^* > \partial_1$ it is possible to choose α_1 so that

$$P\left[F_1 < F^0_{1,\alpha_1}, F_2 < F^0_{2,\alpha_2}; \partial_1, \partial_2\right] > P\left[F_1 < F^0_{1,\alpha_1}, F_2 < F^0_{2,\alpha_2}; \partial_1^*, \partial_2^*\right].$$

Let
$$I(F_1; \partial_2) = P\left[F_2 < F^0_{2,\alpha_2} | F_1; \partial_2\right].$$

Choose F^0_{1,α_1} so that

$$p\left(F^0_{1,\alpha_1}; \partial_1\right) / p\left(F^0_{1,\alpha_1}; \partial_1^*\right) = 1$$

where $p(\cdot\,;\,\partial)$ is the density of the non-central F-distribution with d.f. 1 and $N-1$ and noncentrality parameter ∂. Now

$$P\left[F_1 < F^0_{1,\alpha_1},\ F_2 < F^0_{2,\alpha_2};\ \partial_1, \partial_2\right]$$

$$= \int_0^{F^0_{1,\alpha_1}} I(F_1;\ \partial_2)p(F_1;\ \partial_1)dF_1$$

$$= \int_0^{F^0_{1,\alpha_1}} I(F_1;\ \partial^2)\, p(F_1;\ \partial^*_1)\, \frac{p(F_1;\ \partial_1)}{p(F_1;\ \partial^*_1)}\, dF_1$$

$$\geqslant \int_0^{F^0_{1,\alpha_1}} I(F_1\,;\ \partial_2)\, p(F_1;\ \partial^*_1)dF_1$$

$$= P\left[F_1 < F^0_{1,\alpha_1},\ F_2 \leqslant F^0_{2,\alpha_2};\ \partial^*_1, \partial_2\right]$$

since $p(F_1;\ \partial_1)/p(F_1, \partial^*_1)$ is a decreasing function of F_1.

For the general case, given $\partial^*_i > \partial_i$, the points $F^0_i, F^0_{i-1}, \ldots, F^0_1$ can be chosen suitably so that the power at ∂^*_i is more than the power at ∂_i when the other ∂_j's are fixed.

4.2. Next, we shall consider some testing problems in the two-sample case. We have a random sample of size N_1 from $N_p(\mu_1, \Sigma)$ and a random sample of size N_2 from $N_p(\mu_2, \Sigma)$. Define $\gamma' = (\mu_1-\mu_2)'\Sigma^{-1} = (\gamma_1, \ldots, \gamma_p)$. The following testing problems are of interest.

(A) $H:\ \gamma^{(1)}_i \leqslant \gamma_i \leqslant \gamma^{(2)}_i;\quad K:\gamma_i > \gamma^{(2)}_i$ or $< \gamma^{(1)}_i$

(B) $H:\gamma_i = \gamma^{(0)}_i;\qquad K:\gamma_i \neq \gamma^{(0)}_i$.

(C) $H:\sum_{i=1}^{p} b_i\gamma_i = b_0;\qquad K:\sum_{i=1}^{p} b_i\gamma_i \neq b_0$

(D) $\gamma_p = \gamma_{p-1} = \ldots = \gamma_{p-q+1} = 0;\qquad K:$ not H.

It follows from Lehmann (1959) that UMP unbiased test exists for each of the problems (A), (B) and (C). For the problem (D), the UMP test exists in the class of all similar, invariant tests; however, invariance should be considered with respect to the group of transformations taken in Giri (1964) along with the translation group.

Remark 1 : It would be interesting to study the problem of testing the hypothesis $H:\mu_1\Sigma_1^{-1} = c\mu_2\Sigma_2^{-1}$ when the two normal populations have covariance matrices Σ_1 and Σ_2. The likelihood-ratio test for this problem is quite complicated.

Remark 2 : The problem of testing $\gamma_1 = \gamma_2 = \ldots = \gamma_p = 0$ is formulated in a multiple-decision set-up in Das Gupta (1960) and another property of the step-down test is also studied there.

REFERENCES

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

DAS GUPTA, S. (1966): Step-down multiple-decision rules. To be published in Professor S. N. Roy Memorial Volume, University of North Carolina, Chapel Hill.

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 Stat.*, **35**, 200–205.

GIRI, N. C. (1964): On the likelihood ratio test of a normal multivariate testing problem. *Ann. Math. Stat.*, **35**, 181–189.

——— (1965): On the likelihood ratio test of a normal multivariate testing problem II. *Ann. Math. Stat.*, **36**, 1061-1065.

KHATRI, C. G. (1966): A note on a MANOVA model applied to problems of growth curves. *Ann. Inst. Stat. Math.*, **18**, 75–85.

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

RAO, C. R. (1965): *Linear Statistical Inference and its Applications*, Wiley, New York.

——— (1966): Inference on discrimination function coefficients. Unpublished.

ROY, J. (1958): Step-down procedure in multivariate analysis. *Ann. Math. Stat.*, **29**, 1177–1187.

Paper received : May, 1967.

Revised : January, 1968.