ON A MULTIVARIATE SLIPPAGE PROBLEM I [1)]

M. S. SRIVASTAVA

(Received Nov. 29, 1965)

Summary

In this paper a slippage problem for the covariance matrices of multivariate normal populations is considered and a procedure is given so that the probability of making the correct selection when there has been no slippage exceeds a specified value. The proposed procedure is shown to be admissible. The statistic used in the procedure is a multivariate analogue of Cochran's [1] statistic.

1. Introduction

Slippage problems have been considered many times in statistical literature; for an extensive bibliography the reader is referred to Karlin and Truax [3]. Roughly, the problem (in the controlled case) is as follows: Suppose we are given $k+1$ populations with density functions $f(x, \theta_0)$, $f(x, \theta_1)$, $\cdots$, $f(x, \theta_k)$. Then the problem is to decide on the basis of a sample from each population, if all the θ_i are equal to θ_0, or, if not, which of the k populations has a larger (or smaller) parameter than θ_0, when it is known that *all* but *one* are equal to θ_0. Karlin and Truax, op. cit., considered a slippage problem for the means of multivariate normal populations and gave an admissible procedure in the *restricted* class of *symmetric invariant* procedures; the admissibility of this procedure without these restrictions has recently been proved by the author [5]; the result has been extended for unequal samples also.

In the present investigation, a slippage problem for the covariance matrices of multivariate normal populations is considered and an admissible procedure is given.

2. Problem

Consider $k+1$ multivariate normal populations $\Pi_0, \Pi_1, \cdots, \Pi_k$; Π_i : $N(\mu_i, \Delta_i)$, where $N(\mu_i, \Delta_i)$ denotes a p-variate nonsingular normal distri-

1) Research supported partially by the Office of Naval Research, Contract Nonr 1858(05) at Princeton University.

bution with mean vector μ_i and covariance matrix Δ_i. Suppose we have $k+1$ hypotheses, $H_0: \Delta_1=\cdots=\Delta_k=\Delta_0$, and $H_i: \Delta_j=\Delta_0$ for $j\neq i$, $\Delta_0^{1/2}\Delta_i^{-1}\Delta_0^{1/2}=I_p+\eta\eta'$, $i=1,2,\cdots,k$, where η is a $p\times r$ matrix of rank $r\geqq 1$, and $\Delta_0^{1/2}$ is the unique positive definite symmetric square root of Δ_0, and I_p is a $p\times p$ identity matrix. Suppose it is known that exactly one of these $k+1$ hypotheses is true. We want to decide on the basis of N_j observations from Π_j, $j=0,1,\cdots,k$, which one of these is true subject to the restriction that if H_0 is true, the decision D_0 (D_i is the decision that the hypotheses H_i is true) is to be selected with probability $\geqq 1-\alpha$, where α is a pre-assigned number.

A minimal sufficient set of statistics consists of the sample means $\bar{X}_0, \bar{X}_1, \cdots, \bar{X}_k$ and the sample covariances $S_0, S_1, \cdots, S_k$; S_j is based on N_j-1 degrees of freedom and $E(S_j)=(N_j-1)\Delta_j$.

It is shown that the following procedure for selecting one of the $(k+1)$ decisions $D_0, D_1, \cdots, D_k$ is admissible:

$$\text{Select } D_0 \text{ if } [\min_{1\leq j\leq k}|S_j|/|S_0+S_1+\cdots+S_k|]>R\,,$$

$$\text{(1)}\qquad \text{Select } D_i \text{ if } [\min_{1\leq j\leq k}|S_j|/|S_0+S_1+\cdots+S_k|]\leqq R\,, \text{ and}$$

$$\det(S_i)<\det(S_j) \quad \text{for all} \quad j\neq i,\ j=1,2,\cdots,k\,,$$

where the symbol $\det(A)$ stands for the determinant of a square matrix, i.e., $\det(A)=|A|$, and the constant R depends on α.

3. Solution

We use the Bayes technique to prove the admissibility of the proposed rule (1). Hence, it will be shown that the procedure given by (1) is an (a.e.) Unique Bayes procedure. Let

$$\text{(2)}\qquad \theta_j=N_j^{1/2}\mu_j$$

$$\text{(3)}\qquad Y_j=N_j^{1/2}\bar{X}_j$$

$$\text{(4)}\qquad Y=(Y_0, Y_1, \cdots, Y_k)$$

and

$$\text{(5)}\qquad S=(S_0, S_1, \cdots, S_k)\,.$$

The joint density of the sufficient statistics Y and S is

$$\text{(6)}\qquad \text{Const.} \prod_0^k |\Delta_j|^{-N_j/2} \operatorname{etr} -\tfrac{1}{2}[\sum_0^k \Delta_j^{-1}\{S_j+(Y_j-\theta_j)(Y_j-\theta_j)'\}]\,,$$

where the expression $\prod_0^k (\det S_j)^{(N_j-p-2)/2}$ is included in the constant, and where the symbol etr stands for exponential of the trace of a square matrix. The density (6) under H_i, $i=1, 2, \cdots, k$, becomes

$$(7) \qquad \text{Const. } |\Delta_0^{-1}|^{N/2}|I_p+\eta\eta'|^{N_i/2} \operatorname{etr} -\tfrac{1}{2}[\Delta_0^{-1} \sum_0^k \{S_j+(Y_j-\theta_j)(Y_j-\theta_j)'\} + \Delta_0^{-1/2}\eta\eta'\Delta_0^{-1/2}\{S_i+(Y_i-\theta_i)(Y_i-\theta_i)'\}] ,$$

where

$$(8) \qquad N=\sum_0^k N_j .$$

The density (6) under H_0 becomes (7) with $\eta=0$ (zero matrix).

We compute the Bayes procedure relative to the prior distribution $P=\sum_{i=0}^k \xi_i P_i$, $0\leqq\xi_i\leqq 1$, $\sum_0^k \xi_i=1$, on the parameter space $\Omega=\bigcup_{i=0}^k H_i$ with $P(\Omega)<\infty$ and with P_i a finite measure on H_i, $i=0, 1, \cdots, k$. In the present investigation, we consider *simple* loss function, i.e., the loss is assumed to be zero or one, according as a correct or incorrect decision was made. Let φ_i be the probability of accepting the ith decision; $\sum_{i=0}^k \varphi_i=1$. Then, for the simple loss function, a decision rule is a *Bayes rule* relative to the a priori distribution P, if and only if, except on a set of Lebesgue measure zero, $\varphi_i(T)=0$, whenever

$$(9) \qquad \xi_i\int f(T, \gamma)P_i(d\gamma)<\max_{j\neq i} \left\{\xi_j\int f(T, \gamma)P_j(d\gamma)\right\} ,$$

where $f(T, \gamma)$ is the density function with respect to Lebesgue measure of the distribution of $T=(Y, S)$, and $\gamma=(\theta_0, \theta_1, \cdots, \theta_k, \Delta_0, \Delta_1, \cdots, \Delta_k)\in\Omega$.

We now define the prior distribution P. The construction of the prior distribution is similar to that of Kiefer and Schwartz [4]. Let $\xi_0=\xi$, and $\xi_j=(1-\xi)/k$, $j=1, 2, \cdots, k$. Under H_i $(i=1, 2, \cdots, k)$,

$$(10) \qquad \Delta_i^{-1}=\Delta_0^{-1}+\Delta_0^{-1/2}\eta\eta'\Delta_0^{-1/2} .$$

Let P_i $(i=1, 2, \cdots, k)$ assign all its measure to Δ_0's of the form

$$(11) \qquad \Delta_0^{-1}=I_p+\delta\delta'$$

where δ is a $p\times q$ matrix of rank q, $1\leqq q\leqq p$, and to θ_j's of the form $\Delta_0^{-1}\theta_j=\delta\gamma_j$ for $j\neq i$, $j=0, 1, \cdots, k$, and to θ_i's $(i=1, 2, \cdots, k)$ of the form $\Delta_i^{-1}\theta_i=(\delta, \Delta_0^{-1/2}\eta)\beta_i$, where γ_j's are q-vectors and β_i is a $(q+r)$-vector; Δ_0 and Δ_i are given by the expressions (10) and (11). Under P_i, the density (7) becomes

$$\text{(12)}\qquad \text{Const.}\ |\Delta_0^{-1}|^{N/2}|I_p+\eta\eta'|^{N_i/2}\ \text{etr}\ -\tfrac{1}{2}[\delta\delta' A+\Delta_0^{-1/2}\eta\eta'\Delta_0^{-1/2}(S_i+Y_iY_i')$$
$$-2\delta\textstyle\sum' \gamma_j Y_j' -2(\delta,\ \Delta_0^{-1/2}\eta)\beta_i Y_i' +(\sum' \gamma_j\gamma_j')\delta'\Delta_0\delta$$
$$+\beta_i\beta_i'(\delta,\ \Delta_0^{-1/2}\eta)'\Delta_i(\delta,\ \Delta_0^{-1/2}\eta)]\ ,$$

where the symbol $\sum'$ denotes the summation over all j from 0 to k except for $j=i$, and where the expression (etr $-\frac{1}{2}A$) is included in the constant (we need only to verify that the same expression comes out under H_0 also); A is given by

$$\text{(13)}\qquad A=\sum_0^k (S_j+Y_jY_j')\ .$$

Note that in the above expression (12) and in all the expressions below, Δ_0 and Δ_i are given by (10) and (11), respectively. We now define the prior distributions of γ_j's, $j\neq i$, $j=0, 1, \cdots, k$, β_i and δ. Let the conditional prior density of β_i given η and δ be normally distributed with mean vector 0 and covariance matrix

$$\text{(14)}\qquad I_{q+r}+(\delta,\ \Delta_0^{-1/2}\eta)'(\delta,\ \Delta_0^{-1/2}\eta)\ ,\quad \eta: p\times r\ ,$$

and; the conditional prior distributions of γ_j's given δ be independently and identically normally distributed with mean vector 0 and covariance matrix

$$\text{(15)}\qquad I_q+\delta'\delta\ ,\quad \delta: p\times q\ ,$$

and, is independent of β_i. Using the identities

$$\text{(16)}\qquad (I_q+u'u)^{-1}=I_q-u'(I_p+uu')^{-1}u$$

for $u: p\times q$, and

$$\text{(17)}\qquad |I_q+u'u|=|I_p+uu'|\ ,$$

and taking the expectation of (12) with respect to the prior distributions of γ_j's and β_i, we find that under H_i, the joint density of Y and S given η and δ is

$$\text{Const.}\ |\Delta_0^{-1}|^{(N-k-1)/2}|I_p+\eta\eta'|^{(N_i-1)/2}\iint \text{etr}\ -\tfrac{1}{2}[\delta\delta' A+\Delta_0^{-1/2}\eta\eta'\Delta_0^{-1/2}(S_i+Y_iY_i')$$
$$\text{(18)}\qquad -2\delta\textstyle\sum' \gamma_j\gamma_j' -2(\delta,\ \Delta_0^{-1/2}\eta)\beta_i Y_i' +\sum' \gamma_j\gamma_j' +\beta_i\beta_i']\, d\beta_i \prod_{\substack{j\neq i\\ j=0}}^{k} d\gamma_j$$
$$=\text{Const.}\ |\Delta_0^{-1}|^{(N-k-1)/2}|I_p+\eta\eta'|^{(N_i-1)/2}\text{etr}\ -\tfrac{1}{2}[\delta\delta' A+\Delta_0^{-1/2}\eta\eta'\Delta_0^{-1/2}(S_i+Y_iY_i')$$
$$-\delta\delta'\textstyle\sum' Y_jY_j' -(\delta,\ \Delta_0^{-1/2}\eta)(\delta,\ \Delta_0^{-1/2}\eta)'Y_iY_i']$$
$$=\text{Const.}\ |\Delta_0^{-1}|^{(N-k-1)/2}|I_p+\eta\eta'|^{(N_i-1)/2}\text{etr}\ -\tfrac{1}{2}[\delta\delta'\bar{S}+\Delta_0^{-1/2}\eta\eta'\Delta_0^{-1/2}S_i]\ ,$$

where

$$\bar{S}=\sum_{0}^{k} S_j \,. \tag{19}$$

We have yet to define the prior distribution of η and δ. Let the conditional prior (Lebesgue) density of η given δ be given by

$$\text{Const.}\ |I_p+\eta\eta'|^{-(N_i-1)/2}\,, \quad N_i>p+r\,, \tag{20}$$

and let the prior density of δ be given by

$$\text{Const.}\ |I_p+\delta\delta'|^{-(N-r-k-1)/2}\,, \quad N>p+q+r+k\,. \tag{21}$$

The integrability of these densities follows from equation (3.7) of Kiefer and Schwartz [4].

Taking the expectation of (18) first with respect to the prior distribution of η and then with respect to the prior distribution of δ, we find that the unconditional density of Y and S under H_i $(i=1, 2, \cdots, k)$

$$\text{Const.}\ (\det S_i)^{-r/2}(\det \bar{S})^{-q/2}\,. \tag{22}$$

We now define the prior distribution P_0 on H_0. Under H_0, $\Delta_1=\Delta_2=\cdots=\Delta_k=\Delta_0$. Let P_0 assign all its measure to Δ_0's of the form

$$\Delta_0^{-1}=I_p+\zeta\zeta'\,, \tag{23}$$

where ζ is a $p\times s$ matrix of rank s, $1\leqq s\leqq p$. Also, under H_0, all measure is assigned to θ_j's of the form $(I_p+\zeta\zeta')\theta_j=\zeta\gamma_j^*$, $j=0, 1, 2, \cdots, k$, where γ_j^*'s are s-vectors. The density (7) with $\eta=0$, then becomes

$$\text{Const.}\ |\Delta_0^{-1}|^{N/2}\,\text{etr}\ -\tfrac{1}{2}[\zeta\zeta' A-2\zeta\sum_{0}^{k}\gamma_j^* Y_j'+(\sum_{0}^{k}\gamma_j^*\gamma_j^{*\prime})\zeta'\Delta_0\zeta] \tag{24}$$

where (as under H_i) the expression (etr $-\frac{1}{2}A$) is included in the constant, and where N and A have been defined in (8) and (13) respectively. The prior distributions of ζ and γ_j^*'s are yet to be defined. Let the conditional prior distributions of γ_j^*'s given ζ be independently and identically normally distributed with mean 0 and covariance matrix

$$I_q+\zeta\zeta'\,, \quad \zeta: p\times q \tag{25}$$

and; let the conditional prior (Lebesgue) density of ζ be given by

$$\text{Const.}\ |I_p+\zeta\zeta'|^{(N-k-1)/2}\,, \quad N>p+s+k\,. \tag{26}$$

The integrability of the above density function follows from equation (3.7) of Kiefer and Schwartz [4].

Taking the expectation of (24) first with respect to the prior distributions of γ_j^*'s and then with respect to the prior distribution of ζ

[using identities (16) and (17)], we find that under H_0, the unconditional density of Y and S is

$$\text{Const.}\ |\bar{S}|^{-s/2}\ . \tag{27}$$

Choose $s=q+r$. [This implies that $2\leqq s\leqq p$, and $1\leqq q$, $r\leqq p-1$. However, the restriction that $r\leqq p-1$ instead of $r\leqq p$ is not a serious restriction since we have an identity matrix attached with $\eta\eta'$ in the representation]. Then for the simple loss function, the form of the Bayes solutions is:

$$\begin{aligned} &\varphi_0(Y,S)=1 \quad \text{if} \quad [\min_{1\leqq j\leqq k} |S_j|\,/\,|S_0+S_1+\cdots+S_k|]\geqq R\ ; \\ (1)\qquad &\varphi_i(Y,S)=1 \quad \text{if} \quad [\min_{1\leqq j\leqq k} |S_j|\,/\,|S_0+S_1+\cdots+S_k|]<R\ , \qquad \text{and} \\ &\qquad\qquad \det(S_i)<\det(S_j) \quad \text{for all} \quad j\neq i,\ j=1,2,\cdots,k\ , \end{aligned}$$

where $\varphi_i(Y,S)$ is the probability of taking the decision D_i, $i=0,1,\cdots;$ k, and R depends on ξ. Since the set of (Y,S)'s which yield ties among the minimum of the statistics [$\det(S_j)/\det(S_0+S_1+\cdots+S_k)$] has Lebesgue measure zero, it is an a.e. unique Bayes procedure. Hence the above procedure is admissible.

We know that the constant R depends on ξ. A simple continuity argument show that when ξ varies between 0 and 1, the constant R varies continuously from its largest possible value to its smallest possible value. In particular, when $\xi=0$, $\varphi_0(Y,S)\equiv 0$, and when $\xi=1$, $\varphi_0(Y,S)\equiv 1$. For any prescribed R (chosen so as to have the probability of accepting D_0 when H_0 is true $\geqq\alpha$), by continuity, we obtain the existence of ξ^* such that the given procedure of (1) defined by the constant R is Bayes against ξ^*.

Remark 1. The integrability of the prior distributions [equations (20), and (26)] requires that N_i be $>p+r$. However, if we choose $r=1$, we need $N_i>p+1$, $i=1,2,\cdots,k$. Hence, the admissibility for the minimum sample size p from populations $\Pi_1,\cdots,\Pi_k$ is not proved.

Remark 2. For the special case $\eta\eta'=aI_p$, $a>0$ and known, an admissible procedure can easily be obtained from section 3. The procedure will be minimax also. This problem was considered by the author in [6] and [7], where the admissibility was proved only in the restricted class of procedures—procedures invariant under the additive group of transformations as well as under the group of all nonsingular triangular group of transformations. (The procedure in [6] and [7] needs a correction: All the inequalities in the procedure should be in the reverse direction. The author is indebted to Professor R. A. Wijsman for pointing out this slip.)

Remark 3. If the hypothesis H_i of the problem is changed to: $\Delta_j = \Delta_0$ for $j \neq i$, and $\Delta_i^{1/2}\Delta_0^{-1}\Delta_i^{1/2} = I_p + \eta\eta'$, the statistic used in the procedure (1) will be a multivariate analogue of Cochran's statistic (if there is no control population).

Remark 4. When equal numbers of observations are taken from each population, the method of finding the percentage points is similar to that of Cochran [1]. This problem, however, will be considered in a later paper along with multivariate generalizations of some other statistics (e.g., the statistic $\max_{1 \leq j \leq k} |S_j| / \min_{1 \leq j \leq k} |S_j|$, a multivariate analogue of Hartley's $F_{\max}$ test statistic, [2]).

Acknowledgement

I wish to express my sincere thanks to Professor R. A. Wijsman for his helpful suggestions and criticisms.

PRINCETON UNIVERSITY AND UNIVERSITY OF TORONTO[2)]

REFERENCES

[1] W. G. Cochran, "The distribution of the largest of a set of estimated variances as a fraction of their total," *Ann. Eugenics*, 11 (1941), 47-52.

[2] H. O. Hartley, "The maximum *F*-ratio as a short-cut test for heterogeneity of variance," *Biometrika*, 37 (1950), 308-312.

[3] S. Karlin and D. Truax, "Slippage problems," *Ann. Math. Statist.*, 31 (1960), 296-324.

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

[5] M. S. Srivastava, "On a multivariate slippage problem II," to be submitted for publication, 1965.

[6] M. S. Srivastava, "On a multivariate slippage problem," (Abstract), *Ann. Math. Statist.*, 35 (1964), 1848.

[7] M. S. Srivastava, "Optimum procedures for classification and related problems," Technical report No. 11 under NSG Grant GP-40, Stanford University, 1964.

2) On leave of absence from the University of Toronto.