ON SEQUENTIAL ESTIMATION OF THE MEAN VECTOR OF A MULTINORMAL POPULATION

V. K. ROHATGI* AND R. T. O'NEILL

(Received Dec. 1, 1970; revised July 16, 1971)

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be a sample from a k-variate normal population $\mathfrak{N}(\mu, \Sigma)$ where $\mu=(\mu_1, \cdots, \mu_k)'$ is the mean vector,

$$\Sigma=\begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_k^2 \end{pmatrix}, \qquad \sigma_i^2<\infty, \ i=1, 2, \cdots, k$$

is the variance covariance matrix and both μ and Σ are unknown. In a recent paper [1] Khan studied the limiting behavior of a stopping rule for the sequential estimation of μ when the elements of Σ become infinite. In this note we show that the regret is bounded in the limit. The results obtained here parallel those of Starr and Woodroofe [4] for the univariate case.

Let

$$\bar{X}_{in}=n^{-1}\sum_{j=1}^{n} X_{ij} , \qquad S_{in}^2=(n-1)^{-1}\sum_{j=1}^{n}(X_{ij}-\bar{X}_{in})^2 ,$$

where $n\geqq 2$, $i=1, 2, \cdots, k$. Let $\bar{X}_n=(\bar{X}_{1n}, \cdots, \bar{X}_{kn})'$. Let the loss incurred in estimating μ by $\bar{X}_n$ be given by

$$L(n)=\sum_{i=1}^{k} \lambda_i |\bar{X}_{in}-\mu_i|^s+n \tag{1}$$

where $s>0$ is a given real number and $\lambda_i>0$, $i=1, 2, \cdots, k$. Following [4], we see that

$$\varphi(n)=\mathcal{E}L(n)=n^{-s/2}C(s)\sum_{1}^{n}\lambda_i\sigma_i^s+n , \qquad C(s)=2^{(s+1)/2}(\sqrt{2\pi})^{-1}\Gamma((s+1)/2) , \tag{2}$$

* This work was supported by the National Science Foundation Grant GP-9396.

which is minimized for $n=n_0$ given by

$$(3)\qquad n_0=\left[\frac{s}{2}C(s)\sum_1^k \lambda_i\sigma_i^s\right]^{2/(s+2)} .$$

The minimum risk, if we use n_0 observations, is

$$(4)\qquad \varphi(n_0)=\left[\frac{2}{s}+1\right]n_0 .$$

Since $\sigma=(\sigma_1,\cdots,\sigma_k)$ is not known we determine a sample of size N by means of the following sequential procedure.

Let

$$(5)\qquad N=\text{smallest integer } n\geqq m \text{ for which } n\geqq\left(\beta\sum_1^k \lambda_i S_{in}^s\right)^{2/(2+s)} ,$$

where $\beta=(s/2)C(s)$ and $m\geqq k$ is the starting sample size.

2. Some preliminary results

In the following we write $\sigma=(\sigma_1,\cdots,\sigma_k)$ and $\sigma\to\infty$ means $\sigma_i\to\infty$, $i=1,2,\cdots,k$. Let us write

$$(6)\qquad \sigma_*=\min(\sigma_1,\sigma_2,\cdots,\sigma_k) , \qquad \sigma^*=\max(\sigma_1,\sigma_2,\cdots,\sigma_k) ,$$

$$(7)\qquad \lambda_*=\min(\lambda_1,\lambda_2,\cdots,\lambda_k) , \qquad \lambda^*=\max(\lambda_1,\lambda_2,\cdots,\lambda_k)$$

and assume that

$$(8)\qquad \sigma^*/\sigma_*\to 1 \qquad \text{as } \sigma\to\infty .$$

Note that

$$(9)\qquad n_*=[\beta k\sigma_*^s\lambda_*]^{2/(2+s)}\leqq n_0\leqq[\beta k\sigma^{*s}\lambda^*]^{2/(2+s)}=n^* ,$$

so that, in view of (8)

$$(10)\qquad n^*/n_*\to(\lambda^*/\lambda_*)^{2/(2+s)} \qquad \text{as } \sigma\to\infty .$$

In the following let c denote a positive generic constant.

LEMMA 1. $P\{N<\infty\}=1$

THEOREM 1. (i) $\lim_{\sigma\to\infty} n_0^{-1}N=1$ *a.s.* (ii) $\lim_{\sigma\to\infty} n_0^{-1}\mathcal{E}N=1$.

We remark that both Lemma 1 and Theorem 1 hold if we replace the loss function (1) by

$$(11)\qquad L^*(n)=\sum_1^k \lambda_i|\bar{X}_{in}-\mu_i|^s+\log n .$$

Lemmas 2 and 3 below are of independent interest. The method of proof adopted here parallels closely the work of Simons [3].

LEMMA 2. $P\{N=m\}=O(\sigma_*^{-k(m-1)})$ *as* $\sigma\to\infty$ *in such a way that* (8) *holds.*

PROOF.

$$\begin{aligned}P\{N=m\}&=P\left\{\left(\beta\sum_1^k\lambda_i S_{im}^s\right)^{2/(2+s)}\leqq m\right\}\\&\leqq P\left\{\sum_1^k\frac{S_{im}^s}{\sigma_i^s}\leqq m^{(s+2)/2}(\beta\lambda_*\sigma_*^s)^{-1}\right\}\\&\leqq P\left\{\sum_1^k\frac{S_{im}^2}{\sigma_i^2}\leqq p^{2/s}m^{(s+2)/s}(\beta\lambda_*\sigma_*^s)^{-2/s}\right\}\end{aligned}$$

where we have used the elementary inequality (see [2], p. 264)

$$\left(\sum_1^k a_i^2\right)^{s/2}\leqq p\sum_1^k|a_i|^s,$$

$p=k^{s/2-1}$ if $s\geqq 2$, and $=1$ if $s\leqq 2$. Thus

$$P\{N=m\}\leqq P\{\chi^2_{k(m-1)}\leqq p^{2/s}(m-1)m^{(s+2)/s}(\beta\lambda_*\sigma_*^s)^{-2/s}\}=O(\sigma_*^{-k(m-1)})\,. \tag{12}$$

On the other hand

$$\begin{aligned}P\{N=m\}&\geqq P\left\{\bigcap_1^k[\lambda_i S_{im}^s\leqq(\beta k)^{-1}m^{(2+s)/2}]\right\}\\&\leqq\prod_1^k P\left\{\frac{S_{im}^s}{\sigma_i^s}\leqq(\beta k\lambda_*\sigma_*^s)^{-1}m^{(2+s)/2}\right\}\\&=[P\{\chi^2_{m-1}\leqq(\beta k\lambda_*\sigma_*^s)^{-2/s}m^{(2+s)/s}(m-1)\}]^k\\&=O(\sigma_*^{-k(m-1)})\,.\end{aligned}\tag{13}$$

LEMMA 3. *For fixed* θ, $0<\theta<1$

$$P\{N\leqq\theta n_0\}=O(\sigma_*^{-k(m-1)})\qquad as\ \sigma\to\infty$$

and (8) *holds.*

The methods used in [3] can be similarly modified to yield a proof of Lemma 3. We omit the details.

3. The main result

In this section we return to the quadratic loss function used by Khan [1]

$$L(n)=\sum_1^k\lambda_i|\bar{X}_{in}-\mu_i|^2+n\,, \tag{14}$$

so that

$$n_0=(\sum \lambda_i\sigma_i^2)^{1/2}, \tag{15}$$

$$v(\sigma)=\varphi(n_0)=2n_0, \tag{16}$$

and

$$N=\text{smallest integer } n\geqq m \text{ for which } n\geqq\left(\sum_1^k \lambda_i S_{in}^2\right)^{1/2}. \tag{17}$$

Then

$$\bar{v}(\sigma)=\mathcal{E}L(N)=n_0^2\mathcal{E}(N^{-1})+\mathcal{E}(N). \tag{18}$$

Let us write $w(\sigma)=\bar{v}(\sigma)-v(\sigma)$ for the regret. Clearly

$$w(\sigma)=n_0^2\mathcal{E}(N^{-1}-n_0^{-1})+\mathcal{E}(N-n_0). \tag{19}$$

THEOREM 2. *As $\sigma\rightarrow\infty$ such that* (8) *holds*

$$w(\sigma)=O(1) \tag{20}$$

if and only if $m\geqq 2/k+1$.

PROOF. The proof follows closely the development in [4]. We only indicate here the modifications necessary and omit the details.

For the necessity part of the proof we follow the analysis on page 287 of [4] and obtain

$$\begin{aligned} w(\sigma) &\geqq c\sum_1^k \lambda_i\sigma_i^2 n_0^{-2}(m-n_0)^2\,\mathrm{P}\{N=m\} \\ &\geqq ck\lambda_*\sigma_*^2\left(\frac{m}{n_0}-1\right)^2 O(\sigma_*^{-k(m-1)})=O(\sigma_*^{2-k(m-1)}) \end{aligned}$$

and the necessity of $m\geqq 2/k+1$ for $w(\sigma)=O(1)$ follows.

For the sufficiency part we obtain, as in [4],

$$w(\sigma)\leqq O(\sigma^*)^2\left\{O(\sigma_*^{-k(m-1)})+O(\sigma_*^{-2})\mathcal{E}\left[\frac{(N-n_0)^2}{n_0}\right]\right\}.$$

It only remains to show that

$$\mathcal{E}\left[\frac{(N-n_0)^2}{n_0}\right]=O(1) \tag{21}$$

as $\sigma\rightarrow\infty$ and (8) holds. On integration by parts we get

$$\begin{aligned} \mathcal{E}\left[\frac{(N-n_0)^2}{n_0}\right] &\leqq 1+2\int_1^{\sqrt{n_0}} \lambda\,\mathrm{P}\{N-n_0<-\lambda n_0^{1/2}\}d\lambda \\ &\quad +2\int_1^{\infty} \lambda\,\mathrm{P}\{N-n_0>\lambda n_0^{1/2}\}d\lambda, \end{aligned} \tag{22}$$

which is inequality (11) in [4]. We have

$$2\int_1^{\sqrt{n_0}} \lambda \, \mathrm{P}\{N-n_0<-\lambda\sqrt{n_0}\}\,d\lambda$$
$$\leqq n_0 \, \mathrm{P}\{N\leqq n_0/2\}+2\int_1^{\sqrt{n_0}/2} \lambda \mathrm{P}\{N-n_0<-\lambda\sqrt{n_0}\}\,d\lambda$$
$$=O(1)+2\int_1^{\sqrt{n_0}/2} \lambda \, \mathrm{P}\left\{\sum_1^k \lambda_i S_{il}^2<(n_0-\lambda n_0^{1/2});\ l\geqq n_0/2\right\}d\lambda\,.$$

Thus

$$2\int_1^{\sqrt{n_0}/2} \lambda \, \mathrm{P}\left\{\sum_1^k \lambda_i S_{il}^2<(n_0-\lambda n_0^{1/2})^2;\ l\geqq n_0/2\right\}d\lambda$$
$$=2\int_1^{\sqrt{n_0}/2} \lambda \, \mathrm{P}\left\{\sum_1^k \lambda_i (S_{il}^2-\sigma_i^2)<[(n_0-\lambda n_0^{1/2})^2-n_0^2];\ l\geqq n_0/2\right\}d\lambda$$
$$\leqq 2\int_1^{\infty} \lambda \, \mathrm{P}\left\{\sum_1^k \lambda_i (S_{il}^2-\sigma_i^2)<-\frac{3}{4}\lambda n_0^{3/2};\ l\geqq n_0/2\right\}d\lambda$$
$$\leqq 2\int_1^{\infty} \lambda \sum_1^k \mathrm{P}\left\{S_{il}^2-\sigma_i^2<-\frac{3}{4\lambda_i k}\lambda n_0^{3/2};\ l\geqq n_0/2\right\}d\lambda$$
$$\leqq 2\sum_1^k \int_1^{\infty} \lambda\left(k\frac{4\lambda_i}{3\lambda}\right)^4 n_0^{-6}\mathcal{E}\,|S_{in_2}^2-\sigma_i^2|^4 d\lambda\,,$$

where n_2 is the largest integer $\leqq n_0/2$ (see, for example, [4]). But

$$n_0^{-6}\mathcal{E}\,|S_{in_2}^2-\sigma_i^2|^4\leqq c n_0^{-6} n_2^{-2}\sigma_i^8\leqq c n_0^{-8}\sigma^{*8}\leqq c$$

is bounded. It follows that the first integral in (22) is bounded if $m\geqq 2/k+1$. A similar argument applies to the second integral in (22), proving (21).

BOWLING GREEN STATE UNIVERSITY
CHALLENGER RESEARCH

REFERENCES

[1] Khan, R. A. (1968). Sequential estimation of the mean vector of a multivariate normal distribution, *Sankhyā*, **30**, Ser. A, 331-334.

[2] Loève, M. (1963). *Probability Theory*, D. Van Nostrand, Princeton.

[3] Simons, G. (1968). On the cost of not knowing the variance when making a fixed width confidence interval for the mean, *Ann. Math. Statist.*, **39**, 1946-1952.

[4] Starr, N. and Woodrofe, M. B. (1969). Remarks on sequential point estimation, *Proc. Nat. Acad. Sci.*, **63**, 285-288.