ON SEQUENTIAL ESTIMATION OF A CERTAIN ESTIMABLE FUNCTION OF THE MEAN VECTOR OF A MULTIVARIATE NORMAL DISTRIBUTION

V. K. ROHATGI* and SURESH C. RASTOGI

(Received 7th July 1971)

Communicated by P. D. Finch.

Summary. Consider a k-variable normal distribution $\mathcal{N}(\mu,\Sigma)$ where $\mu = (\mu_1,\mu_2,\cdots,\mu_k)'$ and Σ is a diagonal matrix of unknown elements $\sigma_i^2 > 0, i = 1, 2, \cdots, k$. The problem of sequential estimation of $\Sigma_{i=1}^k \alpha_i\mu_i$ is considered. The stopping rule used is shown to have some interesting limiting properties when the σ_i's become infinite.

1. Introduction

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

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

is the dispersion matrix and both μ and Σ are unknown. Sequential estimation of μ when Σ is unknown has been considered by Khan [1] where references to previous work on the subject (for the case $k = 1$) may also be found. We consider here the problem of estimation of

$$\nu = \sum_{i=1}^{k} \alpha_i\mu_i,$$

where α_i, $i = 1,2,\cdots,k$ are given real numbers. We assume, without loss of generality, that $\alpha_i \neq 0$, $i = 1,2,\cdots,k$. Let

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

* Supported by the NSF Grant Number GP-22585.

where $n \geqq 2$ and $X_i = (X_{i1}, \cdots, X_{in})'$, $i = 1, 2, \cdots, \kappa$. Let $\bar{X}_n = (\bar{X}_{1n}, \cdots, \bar{X}_{kn})'$ Following [5] we use $Y = \Sigma_{i=1}^{k} \alpha_i \bar{X}_{in}$ as an unbiased estimate of v and measure the loss incurred by

$$L(n) = |Y - v|^s + n \tag{1}$$

where $s > 0$ is a given real number. Clearly

$$\phi(n) = EL(n) = a(s)(\Sigma_1^{\kappa} \alpha_i^2 \sigma_i^2)^{s/2} n^{-s/2} + n, \tag{2}$$

where $a(s) = (\sqrt{2\pi})^{-1} 2^{(s+1)/2} \Gamma((s+1)/2)$. The risk $\phi(n)$ is minimized for $n = n_0$ given by

$$n_0 = \beta^{2/(s+2)} (\Sigma \alpha_i^2 \sigma_i^2)^{s/(s+2)} \tag{3}$$

where $\beta = sa(s)/2$. If σ_i's are known we take n_0 observations and estimate v by Y. The risk in doing so is given by

$$v(\sigma) = \phi(n_0) = (1 + 2/s) n_0. \tag{4}$$

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

Let

$$\begin{cases} N = \text{smallest integer } n \geqq m \text{ for which} \\ n > (\beta^{2/s} \Sigma_1^k \alpha_i^2 S_{in}^2)^{s/(s+2)}, \end{cases} \tag{5}$$

where m is the starting sample size.

2. Results

In what follows we write $\sigma \to \infty$ to mean that $\sigma_i \to \infty$, $i = 1, 2, \cdots, k$. c will denote a generic positive constant.

Write

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

$$\alpha_*^2 = \min(\alpha_1^2, \alpha_2^2, \cdots, \alpha_k^2), \qquad \alpha^{*2} = \max(\alpha_2^2, \alpha_2^2, \cdots, \alpha_k^2), \tag{7}$$

and assume that

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

Note that

$$\beta^{2/(2+s)} (k\alpha_*^2 \sigma_*^2)^{s/(s+2)} \leqq n_0 \leqq \beta^{2/(s+2)} (k\alpha^{*2} \sigma^{*2})^{s/(s+2)}. \tag{9}$$

Some immediate results follow along the lines of [1].

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

https://doi.org/10.1017/S1446788700013203 Published online by Cambridge University Press

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

(ii) $\lim_{\sigma\to\infty} \mathscr{E}(n_0^{-1}N) = 1$

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

$$L^*(n) = |Y - v|^s + \log n. \tag{10}$$

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

LEMMA 3. *For fixed θ, $0 < \theta < 1$*

$$P\{N \leqq \theta n_0\} = O(\sigma^{*-k(m-1)}) \text{ as } \sigma \to \infty \textit{ and } (8) \textit{ holds.}$$

The proofs of Lemmas 2 and 3 can easily be constructed by minor modification of the methods of Simons [4]. For example, in case of Lemma 2, we have

$$\begin{aligned} P\{N = m\} &= P\{\Sigma_1^k \alpha_i^2 S_{im}^2 \leqq \beta^{-2/s} m^{(s+2)/s}\} \\ &\leqq P\left\{\Sigma_1^k \frac{(m-1)S_{im}^2}{\sigma_i^2} \leqq \beta^{-2/s}\sigma_*^{-2}\alpha_*^{-2}(m-1)m^{(s+2)/s}\right\} \\ &= P\{\chi^2_{k(m-1)} \leqq \beta^{-2/s}\sigma_*^{-2}\alpha_*^{-2}(m-1)m^{(s+2)/s}\} \\ &= O(\sigma_*^{-k(m-1)}) \text{ as } \sigma \to \infty. \end{aligned}$$

Again

$$\begin{aligned} P\{N = m\} &\geqq P\left\{\bigcap_{i=1}^k [\alpha_i^2 S_{im}^2 \leqq \beta^{-2/s}k^{-1}m^{(s+2)/s}]\right\} \\ &\geqq [P\{\chi^2_{m-1} \leqq \beta^{-2/s}k^{-1}\sigma^{*-2}\alpha^{*-2}(m-1)m^{(s+2)/s}\}]^k \\ &= O(\sigma^{*-k(m-1)}) \text{ as } \sigma \to \infty. \end{aligned}$$

This completes the proof of Lemma 2. The proof of Lemma 3 can be constructed in a similar manner.

THEOREM 2. *As $\sigma \to \infty$ and (8) holds*

$$\mathscr{E}L(N) - v(\sigma) = O(1) \tag{11}$$

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

PROOF. For the proof we will follow the analysis of [5] closely and indicate only the modifications required. First, using an argument similar to the one used in Lemma 3 of [2] we have $\mathscr{E}\{L(N)|N = n\} = \mathscr{E}L(n)$ so that

$$\bar{v}(\sigma) = \mathscr{E}L(N) = \frac{2}{s} n_0^{(s+2)/2} \mathscr{E}N^{-s/2} + \mathscr{E}N. \tag{12}$$

https://doi.org/10.1017/S1446788700013203 Published online by Cambridge University Press

As in [5] we write $w(\sigma) = \tilde{v}(\sigma) - v(\sigma)$ for the regret and see that

(13) $$w(\sigma) = \frac{2}{s} n_0^{(s+2)/2} \mathscr{E}\{N^{-s/2} - n_0^{-s/2}\} + \mathscr{E}\{N - n_0\}.$$

The necessity part of the proof is obtained by simply replacing $O(\sigma^{-(m+1)})$ by $O(\sigma_*^{-k(m-1)})$ in the computations on page 287 of [5].

For the sufficiency part we obtain, as in [5],

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

and it suffices to show that

$$\mathscr{E}\left\{\frac{(N-n_0)^2}{n_0}\right\} = O(1)$$

as $\sigma \to \infty$ and (8) holds. On integration by parts one obtains

(14) $$\mathscr{E}\left\{\frac{(N-n_0)^2}{n_0}\right\} \leqq 1 + 2\int_1^{\sqrt{n_0}} \lambda P\{N - n_0 < -\lambda\sqrt{n_0}\}d\lambda + 2\int_1^{\infty} \lambda P\{N - n_0 > \sqrt{n_0}\}d\lambda,$$

which is inequality (11) in [5]. We have

$$\begin{aligned}
&2\int_1^{\sqrt{n_0}} \lambda P\{N - n_0 < -\lambda\sqrt{n_0}\}d\lambda \\
&\leqq n_0 P\{N \leqq \tfrac{1}{2}n_0\} + 2\int_1^{\sqrt{n_0}/2} \lambda P\{N - n_0 < -\lambda\sqrt{n_0}\}d\lambda \\
&\leqq O(1) + 2\int_1^{\sqrt{n_0}/2} \lambda P\{\Sigma_1^k \alpha_i^2 S_{il}^2 < \beta^{-2/s}(n_0 - \lambda\sqrt{n_0})^{(s+2)/s};\ l \geqq n_0/2\}d\lambda \\
&= O(1) + 2\int_1^{\infty} \lambda P\{\Sigma_1^k \alpha_i^2 (S_{il}^2 - \sigma_i^2) < \beta^{-2/s} \\
&\qquad [(n_0 - \lambda\sqrt{n_0})^{(s+2)/s} - n_0^{(s+2)/2}];\quad l \geqq n_0/2\}d\lambda \\
&\leqq O(1) + 2\int_1^{\infty} \lambda \Sigma_1^k P\{\alpha_i^2(S_{.l}^2 - \sigma_i^2) \\
&\qquad < -\beta^{-2/s} k^{-1}(s+2)s^{-1}(n_0/2)^{2/s}\lambda\sqrt{n_0};\ l > n_0/2\}d\lambda
\end{aligned}$$

(15) $$= O(1) + 2\Sigma_1^k \int_1^{\infty} \lambda\left[\frac{\alpha_i^2 k\beta^{2/s}(s+2)}{2\lambda}\right]^4 n_0^{-(8+2s)/s}\mathscr{E}\left|S_{in_2}^2 - \sigma_i^2\right|^4 d\lambda$$

where n_2 is the greatest integer $\leqq n_0/2$ and we have used exactly the same argument as in [5] page 288.

Since

$$n_0^{-(8+2s)/s} \mathscr{E} \left| S_{in_2}^2 - \sigma_i^2 \right|^4 \leqq c n_0^{-(8+2s)/s} n_2^{-2} \sigma_i^8$$
$$\leqq c n_0^{-(8+4s)/s} \sigma^{*8} = 0(1),$$

it follows that

$$2 \int_1^{\sqrt{n_0}} \lambda P\{N - n_0 < -\lambda \sqrt{n_0}\} d\lambda = 0(1).$$

A similar argument can be used to show the boundedness of the second integral in (14). This completes the proof of the theorem.

Finally we remark that Rohatgi and O'Neill [3] have established recently that the risk is bounded if one uses the sequential procedure described by Khan [1] to estimate the mean vector of a multivariate normal population.

References

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

[2] H. Robbins, *Sequential estimation of the mean of a normal population. Probability and Statistics*, (John Wiley & Sons, N. Y.) (1959), 235–245.

[3] V. K. Rohatgi, and R. T. O'Neill, 'On the sequential estimation of the mean vector of a multionormal population' (To appear in Ann. Inst. Statist. Maths.)

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

[5] N. Starr, and M. B. Woodroofe, 'Remarks on sequential point estimation', *Proc. National Acad. Sciences*, 63 (1969), 285–288.

Statistical Laboratory
Catholic University of America and University of Maryland
Washington D. C. 20017 U. S. A.

Present address

Bowling Green State University
Bowling Green, Ohio, U.S.A.

https://doi.org/10.1017/S1446788700013203 Published online by Cambridge University Press