Jour. Ind. Soc. Ag. Statistics
Vol. XXXVIII, No. 3 (1986), pp. 395-402

SEQUENTIAL ESTIMATION OF A LINEAR FUNCTION OF k-NORMAL MEANS

RAMKARAN, AJIT CHATURVEDI and S. A. AKBAR
Lucknow University, Lucknow

(Received : September, 1984)

SUMMARY

Sequential procedures are proposed (i) to construct fixed width confidence interval for a linear function of means of k normal populations, and (ii) to estimate this linear combination pointwise (the loss being quadratic). Asymptotic properties of the procedures are studied.

Keywords : Sequential estimation, fixed width confidence interval, point estimation, squared error loss, cost, stopping times, asymptotic efficiency and consistency, risk-efficiency.

Introduction

Robbins, Simons, and Starr [7] developed sequential procedures to construct fixed-width confidence interval for the difference of two normal means. Sequential procedures for estimating this difference under quadratic loss functions have been discussed by Mukhopadhyay [3] and Ghosh and Mukhopadhyay [2]. Sequential procedures for estimating a linear function of two and three normal means have been proposed by Mukhopadhyay [5] and Mukhopadhyay [4], respectively. In the present note, sequential interval and point estimation procedures are derived for estimating a linear function of the means of k normal populations.

Let $\{X_{ij}\}$, $j = 1, 2, \ldots$ be a sequence of random observations from the ith ($i = 1, 2, \ldots, k$) normal population, with mean μ_i and variance σ_i^2. All the $2k$ parameters $\mu_i \varepsilon (-\infty, \infty)$ and $\sigma_i \varepsilon (0, \infty)$ are assumed to be unknown. For given non-zero constants $\lambda_1, \lambda_2, \ldots, \lambda_k$, suppose one wishes to estimate $\mu = \sum_{i=1}^{k} \lambda_i \mu_i$. Without loss of generality, we assume $\lambda_i = 1$ for all $i = 1, 2, \ldots, k$.

Having observed a sample $(X_{i1}, X_{i2}, \ldots, X_{in_i})$ of size $n_i (\geqslant 2)$ from the ith population, let us define $\bar{X}_{n_i} = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}$ and $\sigma_{n_i}^{(i)^2} = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{n_i})^2$, i.e., $\sigma_{n_i}^{(i)^2}$ denotes the sample variance based on a sample of size n_i from the ith population. Moreover, for $n = \sum_{i=1}^{k} n_i$, we propose the estimator $\bar{W}_n = \sum_{i=1}^{k} \bar{X}_{n_i}$ for μ.

In section 2, a confidence interval of prescribed width and coverage probability is constructed for μ. Section 3 is devoted to the point estimation of μ under quadratic loss structures.

Fixed Width Confidence Interval for μ

Given the constants $d, \alpha (d > 0, 0, < \alpha < 1)$, suppose one wishes to construct a confidence interval I_n of width $2d$ and confidence coefficient $1 - \alpha$ for μ such that $P(\mu \varepsilon I_n) \geqslant 1 - \alpha$. We propose $I_n = [\bar{W}_n - d, \bar{W}_n + d]$. Now using the fact that $\bar{W}_n$ is normally distributed with mean μ and variance $\sum_{i=1}^{k} \sigma_i^2 / n_i$, we obtain

$$P(\mu \varepsilon I_n) = 2\varphi \left[d \left(\frac{\sigma_1^2}{n_1} + \ldots + \frac{\sigma_k^2}{n_k} \right)^{-1/2} \right] - 1 \tag{2.1}$$

where $\varphi(\cdot)$ stands for the *cdf* of a standard normal variate.

Let 'a' be any constant such that

$$2\varphi(a) - 1 = 1 - \alpha \tag{2.2}$$

It is clear from (2.1) and (2.2) that in order to achieve $P(\mu \varepsilon I_n) \geqslant 1 - \alpha$, $n_1, n_2, \ldots, n_k$ must satisfy the condition

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + \ldots + \frac{\sigma_k^2}{n_k} \leqslant \frac{1}{b} \tag{2.3}$$

where $b = (a/d)^2$. Moreover, using Lagrange's method of undetermined multipliers, the values $n_1^*, n_2^*, \ldots, n_k^*$ of $n_1, n_2, \ldots, n_k$, respectively, for which (2.3) holds and n is minimum, are given by

$$n_1^* = b\sigma_1(\sigma_1 + \sigma_2 + \ldots + \sigma_k)$$
$$n_2^* = b\sigma_2(\sigma_1 + \sigma_2 + \ldots + \sigma_k)$$
$$\vdots$$
$$n_k^* = b\sigma_k(\sigma_1 + \sigma_2 + \ldots + \sigma_k)$$

However, in the absence of any knowledge about σ_i's, no fixed sample size procedure meets the requirements. In such a situation, we adopt a sequential procedure which is described as follows.

Start by taking at least $m(\geqslant 2)$ observations from each of the k populations. If, upto any stage, $N_1 = n_1$ observations from $\{X_{1j}\}$, $N_2 = n_2$ observations from $\{X_{2j}\}, \ldots, N_k = n_k$ observations from $\{X_{kj}\}$ have been taken, the next observation is taken

(A_1) from $\{X_{1j}\}$, if

$$n_1/n_2 \leqslant \sigma_{n_1}^{(1)}/\sigma_{n_2}^{(2)},\ n_1/n_3 \leqslant \sigma_{n_1}^{(1)}/\sigma_{n_3}^{(3)}, \ldots, n_1/n_k \leqslant \sigma_{n_1}^{(1)}/\sigma_{n_k}^{(k)}$$

(A_2) from $\{X_{2j}\}$, if

$$n_2/n_1 \leqslant \sigma_{n_2}^{(2)}/\sigma_{n_1}^{(1)},\ n_2/n_3 \leqslant \sigma_{n_2}^{(2)}/\sigma_{n_3}^{(3)}, \ldots, n_2/n_k \leqslant \sigma_{n_2}^{(2)}/\sigma_{n_k}^{(k)}$$

$$\vdots$$

(A_k) from $\{X_{kj}\}$, if

$$n_k/n_1 \leqslant \sigma_{n_k}^{(k)}/\sigma_{n_1}^{(1)},\ n_k/n_2 \leqslant \sigma_{n_k}^{(k)}/\sigma_{n_2}^{(2)}, \ldots, n_k/n_{k-1} \leqslant \sigma_{n_k}^{(k)}/\sigma_{n_{k-1}}^{(k-1)} \quad (2.4)$$

and the stopping time $N \equiv N(d)$ is the smallest positive integer $n \geqslant km$ such that

$$n_1 \geqslant b\sigma_{n_1}^{(1)}\left(\sigma_{n_1}^{(1)} + \sigma_{n_2}^{(2)} + \ldots + \sigma_{n_k}^{(k)}\right)$$
$$n_2 \geqslant b\sigma_{n_2}^{(2)}\left(\sigma_{n_1}^{(1)} + \sigma_{n_2}^{(2)} + \ldots + \sigma_{n_k}^{(k)}\right)$$
$$\vdots$$
$$n_k \geqslant b\sigma_{n_k}^{(k)}\left(\sigma_{n_1}^{(1)} + \sigma_{n_2}^{(2)} + \ldots + \sigma_{n_k}^{(k)}\right) \quad (2.5)$$

where $N = \sum_{i=1}^{k} N_i$. When stop, construct I_N for μ, The following theorem establishes the results that our sequential procedure is "asymptotically efficient" and "asymptotically consistent" in Chow-Robbins [1] sense.

THEOREM 1. *For all* $0 < \sigma_i < \infty$ $(i = 1, 2, \ldots, k)$ *and* $n^* = \sum_{i=1}^{k} n_i^*$,

$$\lim_{d \to 0} \frac{N}{n^*} = 1 \text{ a.s.} \tag{2.6}$$

$$\lim_{d \to 0} E\left(\frac{N}{n^*}\right) = 1 \tag{2.7}$$

$$\lim_{d \to 0} P(\mu \varepsilon I_N) = 1 - \alpha \tag{2.8}$$

Proof. Note the basic inequality

$$b\left(\sigma_{N_1}^{(1)} + \sigma_{N_2}^{(2)} + \ldots + \sigma_{N_k}^{(k)}\right)^2 \leqslant N \leqslant b\left(\sigma_{N_1'}^{(1)} + \sigma_{N_2'}^{(2)} + \ldots + \sigma_{N_k'}^{(k)}\right)^2 + km \tag{2.9}$$

or,

$$\frac{\left(\sigma_{N_1}^{(1)} + \sigma_{N_2}^{(2)} + \ldots + \sigma_{N_k}^{(k)}\right)^2}{(\sigma_1 + \sigma_2 + \ldots + \sigma_k)^2} \leqslant \frac{N}{n^*} \leqslant \frac{\left(\sigma_{N_1'}^{(1)} + \sigma_{N_2'}^{(2)} + \ldots + \sigma_{N_k'}^{(k)}\right)^2}{(\sigma_1 + \sigma_2 + \ldots + \sigma_k)^2} + \frac{km}{n^*}$$

where $N_i' = N_i$ or N_{i-1} (for $i = 1, 2, \ldots, k$) depending on which population is sampled at the final stage. If i_0th population is sampled at the final stage then $N_i' = N_{i_0} - 1$ and $N_i' = N_i$ $i \neq i_0$, $i = 1, 2, \ldots, k$. Now, using the fact that for all $i = 1, 2, \ldots, k$ $\lim_{d \to 0} N_i = \infty$ a.s., (2.9) gives (2.6).

It follows from Wiener ergodic theorem (see, Wiener [10]) that for all $i = 1, 2, \ldots, k$, $\sup_{n_i \geqslant 2} \{\sigma_{n_i}^{(i)}\}^2$ has its second moment finite. Thus, the expression on the right hand side of N/n^* in (2.9) is integrable and (2.6), together with dominated convergence theorem provides (2.7).

Finally, note that

$$P(\mu \varepsilon I_N) = 2E\left[\varphi d\left(\frac{\sigma_1^2}{N_1} + \ldots + \frac{\sigma_k^2}{N_k}\right)^{-1/2}\right] - 1$$

Now, using the technique of Robbins, Simons, and Starr [7], (2.6) and dominated convergence theorem lead us to (2.7)

Point Estimation of μ

Let the total loss incurred in estimating μ by $\overline{W}_n$ be squared-error plus the linear cost of sampling, i.e.,

$$L_n(c) = A(\overline{W}_n - \mu)^2 + c(n_1 + n_2 + \ldots + n_k) \tag{3.1}$$

where $A(> 0)$ is known weight, and $c(> 0)$ is the known cost of sampling per unit observation.

For the loss $L_n(c)$, the risk comes out to be

$$\nu_n(c) = A\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + \ldots + \frac{\sigma_k^2}{n_k}\right) + c(n_1 + n_2 + \ldots + n_k) \tag{3.2}$$

Treating n_i's as continuous variables, the values $n_1^*, n_2^*, \ldots, n_k^*$ of $n_1, n_2, \ldots, n_k$, respectively, which minimize $\nu_n(c)$ are

$$n_1^* = (A/c)^{1/2}\,\sigma_1,\ n_2^* = (A/c)^{1/2}\,\sigma_2, \ldots, n_k^* = (A/c)^{1/2}\,\sigma_k \tag{3.3}$$

and, for $n^* = \sum_{i=1}^{k} n_1^*$, the corresponding minimum risk is

$$\nu_n^*(c) = 2cn^* \tag{3.4}$$

However, when all or some of the σ_i's are unknown, no fixed sample size procedure serves the purpose. In this situation, adopt the following sequential procedure.

The sampling scheme is same as that defined at (2.4). Motivated from (3.3), the stopping time $N \equiv N(c)$ is the smallest positive integer $n \geqslant km$ such that

$$n_1 \geqslant (A/c)^{1/2}\,\sigma_{n_1}^{(1)},\ n_2 \geqslant (A/c)^{1/2}\,\sigma_{n_2}^{(2)}, \ldots, n_k \geqslant (A/c)^{1/2}\,\sigma_{n_k}^{(k)} \tag{3.5}$$

When stop, estimate μ by $\overline{W}_N$.

As in Starr [9], define the risk-efficiency of the sequential procedure to be

$$\eta(c) = \bar{\nu}(c)/\nu_n^*(c) \tag{3.6}$$

where $\bar{\nu}(c)$ is the expected-loss of the sequential procedure, i.e.,

$$\bar{\nu}(c) = AE\left(\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2} + \ldots + \frac{\sigma_k^2}{N_k}\right) + cE(N) \tag{3.7}$$

Now prove the following theorem, which establishes the results that our sequential procedure is asymptotically as efficient as the optimal fixed sample size procedure.

THEOREM 2: *For all* $0 < \sigma_i < \infty$ $(i = 1, 2, \ldots, k)$ *and* $n^* = \sum_{i=1}^{k} n_i^*$,

$$\lim_{c \to 0} \frac{N}{n^*} = 1 \text{ a.s.} \tag{3.8}$$

$$\lim_{c \to 0} E\left(\frac{N}{n^*}\right) = 1 \tag{3.9}$$

$$\lim_{c \to 0} \eta(c) = 1 \tag{3.10}$$

Proof. It is easy to verify that for all $i = 1, 2, \ldots, k$, $\lim_{c \to 0} N_i = \infty$ *a.s.* Now the proofs of (3.8) and (3.9) are similar to that of (2.6) and (2.7), respectively.

From (3.4) and (3.7), substituting the values of $\nu_{n^*}(c)$ and $\bar{\nu}(c)$, we obtain after little algebra

$$\eta(c) = \frac{n_1^* E(n_1^*/N_1) + n_2^* E(n_2^*/N_2) + \ldots + n_k^* E(n_k^*/N_k)}{2(n_1^* + n_2^* + \ldots + n_k^*)} + \tfrac{1}{2} E\left(\frac{N}{n^*}\right)$$

Since $\lim_{c \to 0} E(N/n^*) = 1$, it suffices to prove that for all $i = 1, 2, \ldots k$,

$$\lim_{c \to 0} E\left(\frac{n_i^*}{N_i}\right) = 1.$$

We shall establish the proof of $\lim_{c \to 0} E(n_1^*/N_1) = 1$ only, and the remaining proofs are routine.

From $\lim_{c \to 0} N_1/n_1^* = 1$ *a.s.* and Fatou's lemma,

$$\lim_{c \to 0} \inf E\left(\frac{n_1^*}{N_1}\right) \geqslant 1 \tag{3.11}$$

To prove "lim sup" part, we proceed as follows.

Define the following quantities

$$\theta = (1 + \varepsilon) n_1^* \qquad (0 < \varepsilon < 1)$$

$$a_{n_1} = \frac{1}{2}(n_1 - 1)\left(\frac{n_1}{n_1^*}\right)^2$$

$$l(n_1, n_1^*) = \Gamma^{-1}\left(1 + \frac{(n_1 - 1)}{2}\right) a_{n_1}^{(n_1 - 1)/2}$$

$$L(n_1) = (n_1^*)^{(n_1 - 1)} \cdot l(n_1, n_1^*)$$

It is to be noted that $L(n_1)$ is independent of n_1^*.

Write

$$E\left(N_1^{-1}\right) = \sum_{n_1 \geq m} n_1^{-1} P(N_1 = n_1)$$

$$= \pi_1 + \pi_2 + \pi_3 \text{ (say)}$$

where

$$\pi_1 = m^{-1} P(N_1 = m),\ \pi_2 = \sum_{m+1 \leqslant n_1 \leqslant \theta} n_1^{-1} P(N_1 = n_1),$$

$$\pi_3 = \sum_{n_1 > \theta} n_1^{-1} P(N_1 = n_1)$$

First of all

$$\pi_1 = m^{-1} P\left\{ m \geqslant (A/c)^{1/2} \sigma_m^{(1)} \right\}$$

$$= m^{-1} P\left\{ (m-1) \frac{\sigma_m^{(1)2}}{\sigma_1^2} \leqslant (m-1)\left(\frac{m}{n_1^*}\right)^2 \right\}$$

$$= m^{-1} P\left\{ \chi^2_{(m-1)} \leqslant (m-1)\left(\frac{m}{n_1^*}\right)^2 \right\}$$

$$= m^{-1} \Gamma^{-1}\left(\frac{m-1}{2}\right) \int_0^{a_m} e^{-z} z^{(m-1)/2-1} dz$$

$$\leqslant m^{-1} l(m, n_1^*)$$

$$= m^{-1} (n_1^*)^{-(m-1)} L(m)$$

Secondly,

$$\pi_2 \leqslant (m+1)^{-1} \sum_{m+1 \leqslant n_1 \leqslant \theta} P(N_1 = n_1)$$

$$< (m+1)^{-1} \sum_{m+1 < n_1 < \theta} P\left\{ n_1 \geqslant (A/c)^{1/2} \sigma_{n_1}^{(1)} \right\}$$

$$= (m+1)^{-1} \sum_{m+1 < n_1 \leqslant \theta} P\left\{ \chi^2_{(n_1-1)} \leqslant (n_1 - 1)\left(\frac{n_1}{n_1^*}\right)^2 \right\}$$

$$= (m+1)^{-1} \sum_{m+1 < n_1 < \theta} l(n_1, n_1^*)$$

$$= (m+1)^{-1} \sum_{m+1 < n_1 \leqslant \theta} (n_1^*)^{-(n_1-1)} L(n_1)$$

$$\leqslant (m+1)^{-1} (n_1^*)^{-m} \sum_{m+1 \leqslant n_1 < \theta} L(n_1)$$

Using the fact that $L(n_1)$ is monotonic increasing in n_1, we get

$$\pi_3 \leqslant (m+1)^{-1} (\theta - m) L(\theta) \tag{3.13}$$

Finally,

$$\pi_3 \leqslant \theta^{-1} P(N_1 > \theta) \tag{3.14}$$

Combining (3.12), (3.13) and (3.14), we obtain

$$\lim_{c\to 0} \sup E\left(\frac{n_1^*}{N_1}\right) \leqslant 1 - \delta(\varepsilon) \tag{3.15}$$

where $0 < \delta(\varepsilon) < 1$. Since ε is arbitrary, inequalities (3.11) and (3.15) jointly provide $\lim_{c\to 0} E(n_1^*/N_1) = 1$.

Remark : It is to be noted that not only the sequential procedures developed by Robbins, Simons and Starr [7], Mukhopadhyay [4], [5] and Ghosh and Mukhopadhyay [2] are particular cases of our procedures, the sequential point and interval estimation procedures derived by Robbins [6] and Starr [8], respectively follow immediately just by taking $k = 1$.

ACKNOWLEDGEMENT

The authors are thankful to the referee for his many valuable comments, and for pointing out a mistake in the original manuscript.

REFERENCES

[1] Chow, Y. S. and Robbins, H. (1965) : On the asymptotic theory of fixed width sequential confidence intervals for the mean, *Ann. Math. Stat.*, **36** : 457-462.

[2] Ghosh, M. and Mukhopadhyay, N. (1980) : Sequential point estimation of the difference of two normal means, *Ann. Math. Stat.*, **8** : 221-225.

[3] Mukhopadhyay, N. (1975) : Sequential methods in estimation and prediction, *Ph.D. dissertation*, ISI, Calcutta.

[4] Mukhopadhyay, N. (1976) : Sequential estimation of a linear function of means of three normal populations, *Jour. Amer. Statist. Assoc.*, **71** : 149-153.

[5] Mukhopadhyay, N. (1977) : Remarks on sequential estimation of a linear function of two means : The normal case, *Metrika*, **24** : 197-201.

[6] Robbins, H. (1959) : Sequential estimation of the mean of a normal population, *Prob. and Stat.* (H. Cramer's Vol.), Uppsala, 235-245.

[7] Robbins, H. Simons, G. and Starr, N. (1967) : A sequential analogue of the Behrens—Fisher problem, *Ann. Math. Stat.*, **38** : 1384-1391.

[8] Starr, N. (1966a) : The performance of a sequential procedure for the fixed-width interval estimation of the mean, *Ann. Math. Stat.*, **37** : 36-50.

[9] Starr, N. (1966b) : On the asymptotic efficiency of a sequential procedure for estimating the mean, *Ann. Math. Stat.*, **37** : 1173-1185.

[10] Wiener, N. (1939) : The ergodic theorem, *Duke Math. Jour.*, **5** : 1-18.