Jour. Ind. Soc. Ag. Statistics
Vol. XL, No 1 (1988), pp. 51-59

FIXED-SIZE CONFIDENCE REGION ESTIMATION OF THE MEAN VECTOR

R. KARAN SINGH and AJIT CHATURVEDI
Lucknow University

(Received : September, 1987)

SUMMARY

The problem of fixed-size confidence region (ellipsoid and sphere) estimation of the mean vector of a multinormal population is considered. The model considered is same as that of Wang [14]. Two-stage, three-stage and purely sequential procedures are proposed and their properties are studied.

Keywords : Fixed-size confidence region; multinormal population; two-stage, three.stage and sequential estimation procedures; second-order approximations.

Introduction

Let us consider a sequence $\{\underset{\sim}{X}_i\}$, $i = 1, 2, \ldots$ of i. i. d. r. v.'s from a p-variate normal population $N_p(\underset{\sim}{\mu}, \sigma^2\Sigma)$, where $\underset{\sim}{\mu}$ is the $p \times l$ unknown mean vector, σ^2 is an unknown scalar, and Σ is a known $p \times p$ positive definite matrix. For other references where similar models have been adopted for different estimation problems, one may refer to Wang [14] and Chaturvedi [2], [3], [4]. For specified $d \varepsilon (0, \infty)$ and $\alpha \varepsilon (0, 1)$, suppose one wishes to construct an ellipsoid R_n for $\underset{\sim}{\mu}$ in the p-dimensional Euclidean space such that the diameter of R_n is $2d$ and $P(\underset{\sim}{\mu} \varepsilon R_n) > \alpha$. Following Srivastava [11], we define

$$R_n = \{\underset{\sim}{Z} : (\overline{\underset{\sim}{X}}_n - \underset{\sim}{Z})' \, \Sigma^{-1} (\overline{\underset{\sim}{X}}_n - \underset{\sim}{Z}) < d^2\}, \tag{1.1}$$

where $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$ is the sample mean based on a random sample $X_1, \ldots, X_n$ of size n. It can be seen that

$$P(\mu \varepsilon R_n) = \psi\,(nd^2/\sigma^2), \tag{1.2}$$

where $\psi(\cdot)$ stands for the c.d.f. of a χ^2_p r.v. Let 'a' be the upper $100 \cdot \alpha\%$ point of χ^2_p distribution, i.e.,

$$\psi\,(a) = \alpha. \tag{1.3}$$

It is clear from (1.2) and (1.3) that for known σ, in order to achieve $P(\mu \varepsilon R_n) \geqslant \alpha$, the required sample size n is the smallest positive integer $n \geqslant n^*$, where

$$n^* = a\sigma^2/d^2 \tag{1.4}$$

But, when σ is unknown, no fixed sample size procedure can achieve the goals of 'fixed-size' and 'pre-assigned coverage probability' simultaneously for all values of σ. In such a situation, we adopt different estimation procedures based on samples of flexible sizes. These procedures and their relative properties are described below.

In Section 2, along the lines of Stein (12), a two-stage procedure is developed. It is proved that the procedure achieves the required goals. However, this procedure is not 'asymptotically efficient' in Chow-Robbins [6] sense. In order to ensure this property one can, of course, 'modify' the procedure along the lines ot Mukhopadhyay [9]. But, there is a serious drawback in Mukhopadhyay's procedure which is that even the first stage sample size tends to $+\infty$ as $d \to 0$. Such procedures are not feasible on practical grounds. In this connection, Stein [12] has already mentioned that in univariate normal case if one takes the initial sample size sufficiently large (>30), then the ratio of expected sample size to the 'optimal' fixed sample size (e.g., n^* in our case) approaches unity.

In Section 3, a purely sequential procedure is proposed for estimating μ. Such a procedure has advantage over the two-stage procedure since it uses a sufficiently smaller sample size to achieve a confidence interval with very nearly the same coverage. It is also 'asymptotically efficient'. Second-order approximations are achieved for the expected sample size and coverage probability associated with the procedure. For this procedure, along the lines of Simons [10], Srivastava and Bhargava [13] proved that if after stopping, we take, say K, 'additional' observations then, for sufficiently small values of d, $P(\mu \varepsilon R_{N+K}) \geqslant \alpha$ for all μ and σ^2. However, the question, "What should be the value of K?" is un-answered upto now. An attempt is made to obtain a lower bound for K.

In order to construct a fixed-width confidence interval for the mean of a univariate normal population, Hall [7] adopted a three-stage procedure. The logic behind the use of this procedure was that in many situations significant savings in time and money may be achieved. This procedure is strongly competetive with the purely sequential procedure from the point of view of efficiency. In Section 4, a multivariate extension of Hall's three-stage procedure is provided.

2. The Two-Stage Procedure

Let us define [see, Wang (1980)] $\sigma_n^2 = [p(n-1)]^{-1} \sum_{i=1}^{n} (\underset{\sim}{X}_i - \overline{\underset{\sim}{X}}_n)' \Sigma^{-1}(\underset{\sim}{X}_i - \overline{\underset{\sim}{X}}_n)$ as the estimator of σ^2. Let us start with a sample of size m ($\geqslant 2$). Then, the second stage sample size being given by

$$N = \max \{m, [pb_m \, \sigma_m^2/d^2] + 1\} \tag{2.1}$$

where $[y]$ denotes the integer part of y and b_m is the upper $100 \cdot \alpha\%$ point of Snedeckor's F-distribution with p, $p(m-1)$ degrees of freedom. Construct R_N for $\underset{\sim}{\mu}$.

The properties of the two-stage procedure (2.1) are mentioned in the following theorem.

THEOREM 1 : $$\lim_{d \to 0} E(N/n^*) = pb_m/a > 1 \tag{2 2}$$

$$P(\underset{\sim}{\mu} \varepsilon R_N) \geqslant \alpha \tag{2.3}$$

Proof. We note the inequality

$$(pb_m \, \sigma_m^2/d^2) \leqslant N \leqslant (pb_m \, \sigma_m^2/d^2) + m \tag{2.4}$$

or,

$$(pb_m \, \sigma_m^2/a\sigma^2) \leqslant (N/n^*) \leqslant (pb_m \, \sigma_m^2/a\sigma^2) + m/n^*). \tag{2.5}$$

Result (2.2) now follows on taking expectation throughout (2.5) and using the results $E(\sigma_m^2) = \sigma^2$, $\lim_{d \to 0} n^* = \infty$.

Since the events '$N = n$' and '$\bar{X}_n$' are stochastically independent (see, Wang [14]), we obtain

$$P(\mu \epsilon R_N) = E(\psi (Nd^2/\sigma^2)].$$

Using the basic inequality (2.4), one gets

$$P(\mu \epsilon R_N) > E[\psi (pb_m \ \sigma_m^2/\sigma^2)]$$

$$= P\{F(p, p(m-1)) \leqslant b_m\}$$

and (2.3) follows.

Remark 1. For $p = 1, \mu = \mu, \Sigma = I = 1$, (1.1) reduces to an interval of width $2d$ and coverage probability α for the mean μ of a univariate normal population. In this case, since $b_m = F(1, m-1) = t^2_{(m-1)}$, where $t_{(m-1)}$ is the upper $100 \cdot \alpha\%$ point of Student's t-distribution with $(m-1)$ degrees of freedom, the procedure (2.1) reduces to the Stein's two-stage procedure.

Remark 2 : It is straightforward from (2.2) that the procedure (2.1) is not 'asymptotically efficient' in view of Chow and Robbins [6]. As in Mukhopadhyay [9], one can modify the procedure to ensure this property by taking $m = \max\{2, [(a/d^2)]^{1/(1+\gamma)} + 1\}$, where $\gamma(> 0)$ is some known constant. This is due to the fact that $\lim_{d \to 0} m = \infty$ and $\lim_{m \to \infty} F(p, p(m-1)) = p^{-1} \chi^2_p$. But, as we have mentioned earlier, even $m \to \infty$ as $d \to 0$.

3. The Purely Sequential Procedure

Let us start with a sample of size $m (> 2)$. Then, the stopping time $N = N(d)$ is defined by

$$N = \inf\{n > m : n \geqslant a\sigma_n^2/d^2\}$$

$$= \inf\{n > m : S_n \leqslant (pn^2/n^*) L(n)\} \quad (3.1)$$

where $L(n) = 1 - n^{-1}$ and $S_n = \sum_{j=1}^{p(n-1)} Z_j^2$, with $Z_j \sim N(0, 1)$.

Let us consider the difference

$$R_d = (pN^2/n^*) L(N) - S_N. \quad (3.2)$$

The mean, say ν, of the asymptotic (as $d \to 0$) distribution of R_d can be obtained from Theorem 2.2 of Woodroofe [15].

In the following theorem, we shall obtain second-order approximations for the expected sample size and coverage probability associated with the sequential procedure (3.1).

THEOREM 2 : *For all* $m > 1 + 2p^{-1}$, *as* $d \to 0$,

$$E(N) = n^* + p^{-1} \nu - 2 + 0(1) \quad (3.3)$$

$$P(\underset{\sim}{\mu} \varepsilon R_N) = \alpha + (a/n^*) \{p^{-1} \nu - 4 - a + p\} \xi(a) + 0(d^2), \quad (3.4)$$

where $\xi(\cdot)$ stands for the p.d.f. of a χ^2_p r.v.

Proof. From (3.2) and Wald's equation,

$$pE(N - 1) = (p/n^*) E\{N^2 - N\} - \nu$$
$$= (p/n^*) E\{n^{*2} + 2n^* (N - n^*) + (N - n^*)^2 - (N - n^*) - n^*\} - \nu$$

or,

$$p(1 - 1/n^*) E(N - n^*) = \nu - pE\{(N - n^*)^2/n^*\} \quad (3.5)$$

It follows from a result of Bhattacharya and Mallik [1] that $(N - n^*)/(n^*)^{1/2} \xrightarrow{\mathcal{L}} N(0, 2)$ as $d \to 0$. Moreover, following the proof of Lemma 4 in Chaturvedi (to appear) it can be shown that $(N - n^*)^2/n^*$ is uniformly integrable for all $m > 1 + 2p^{-1}$. Thus, we obtain from (3.5) for all $m > 1 + 2p^{-1}$, as $d \to 0$,

$$pE(N - n^*) = \nu - p\{2 + 0(1)\}$$

and (3.3) holds.

Denoting by $\psi'(\cdot)$ and $\psi''(\cdot)$, the first and second derivatives of $\psi(\cdot)$ and applying Taylor series expansion, we obtain for $| a - W | \leqslant | (N/n^*) - 1 |$.

$$P(\underset{\sim}{\mu} \varepsilon R_N) = E[\psi(aN/n^*)]$$
$$= \psi(a) + (a/n^*) E(N - n^*) \psi'(a) + (a^2/n^*) \cdot E\{(N - n^*)^2/n^* \psi''(W)\} \quad (3.6)$$

Using the results $(N - n^*)/(n^*)^{1/2} \xrightarrow{\mathcal{L}} N(0, 2)$, $(N - n^*)^2/n^*$ is uniformly integrable for all $m > 1 + 2p^{-1}$, $\psi''(W) \to \psi''(a)$ as $d \to 0$, and making substitution from (3.3), equation (3.6) leads us to

$$P(\underset{\sim}{\mu}\varepsilon R_N) = \alpha + (a/n^*)\{p^{-1}\nu - 2 + 0(1)\}\,\psi'(a) + (a^2/n^*)$$
$$\{2 + 0(1)\}\,\psi''(a),$$

Result (3.4) now follows on using $\psi'(a) = \xi(a)$ and $\psi''(a) = \{(-1/2) + (p/2 - 1)\,a^{-1}\}\,\xi(a)$, where $\xi(\cdot)$ denotes the p.d.f. of a χ^2_p r.v.

In the next theorem, we shall obtain a lower bound for the 'additional' observations—'K', which ensures $P(\underset{\sim}{\mu}\varepsilon R_N) \geqslant \alpha$ for all $\underset{\sim}{\mu}$, σ^2 and sufficiently small d. We first establish a basic lemma.

LEMMA 1 : *Let* $M = [a\sigma^2_N/d^2]^+ + K$, *where* N *is determined by the rule* (3.1) *and* $[y]^+$ *stands for the least positive integer* $\geqslant y$. *Then, for all* $m > 1 + 2p^{-1}$, *as* $d \to 0$,

$$E(M) = (a\sigma^2/d^2) + K - 2 + 0(1) \tag{3.7}$$

$$\text{Var}\,(M) = (2a\sigma^2/d^2) + 0\,(d^{-2}) \tag{3.8}$$

and, for $r > 0$,

$$E\,|\,M - E(M)|^r = 0\,(d^{-r}) \tag{3.9}$$

Proof: For R_d defined at (3.2), let

$$R^*_d = \{a\sigma^2/d^2\}\,p^{-1}\,(N-1)^{-1}R_d$$

Since $\sigma^2_{N-1} \geqslant (N-1)\,d^2/a$, we have

$$R^*_d \leqslant N - (a/d^2)\,(N-1)^{-1}p^{-1}\cdot \sum_{i=1}^{N-1} (\underset{\sim}{X}_i - \underset{\sim}{\overline{X}}_{N-1})'\,\Sigma^{-1}(\underset{\sim}{X}_i - \underset{\sim}{\overline{X}}_{N-1})$$

$$= N - (a/d^2)\,(N-1)^{-1}\,(N-2)\,\sigma^2_{N-1}$$
$$\leqslant 2.$$

Moreover, from the definition of N, $R^*_d \geqslant 0$. Thus, $0 \leqslant R^*_d \leqslant 2$ and hence, $E(R^*_d) \to p^{-1}\nu$ as $d \to 0$. Now, using (3.3) we obtain for all $m > 1 + 2p^{-1}$, as $d \to 0$,

$$E(R^*_d) = p^{-1}\nu$$
$$= E(N) - E\,\{a\sigma^2_N/d^2\}$$

or,

$$E\{a\sigma^2_N/d^2\} = (a\sigma^2/d^2) - 2 + 0(1).$$

giving (3.7).

We note that $\text{Var}(M) = \text{Var}(N)$. Let $h(N) = | \{a\sigma^2/d^2\}^{-1/2} \{N - a\sigma^2/d^2\} |$. Since $h(N) \xrightarrow{\mathcal{L}} N(0, 2)$ as $d \to 0$ and for all $m > 1 + 2p^{-1}$, $h^2(N)$ is uniformly integrable, we get

$$\text{Var}(M) = (2a\sigma^2/d^2)\{1 + 0(1)\}$$

and (3.8) follows.

For the proof of (3.9), we refer to Hall [8].

Now we establish the main theorem.

THEOREM 3 : *For all* $(\underset{\sim}{\mu}, \sigma^2)$ *and sufficiently small* d,

$$P(\underset{\sim}{\mu} \varepsilon R_{N+K}) \geqslant \alpha \quad \text{if } K \geqslant (5/2) - (p-2)/2a - a.$$

Proof: Since $M = N + K$,

$$P(\underset{\sim}{\mu} \varepsilon R_{N+K}) = E\{\psi(Md^2/\sigma^2)\}$$

$$= \psi(a) + E(Md^2/\sigma^2 - a)\,\psi'(a) + (1/2)\,E(Md^2/\sigma^2 - a)^2\,\psi''(a) + r(d)$$

where $r(d) = 0\,(d^6 E\,|\,M - E(M)|^3) = 0\,(d^2)$ on using (3.9). Hence,

$$P(\underset{\sim}{\mu} \varepsilon R_{N+K}) = \alpha + (d^2/\sigma^2)\,E(M - a\sigma^2/d^2)\,\xi(a) + (d^4/2\sigma^4)\{-1/2 + (p/2 - 1)\,a^{-1}\}\{\text{Var}(M) + (EM - a)^2\}\,\xi(a) + 0\,(d^2)$$

which, on using (3.7) and (3.8), gives

$$P(\underset{\sim}{\mu} \varepsilon R_{N+K}) = \alpha + (d^2/\sigma^2)\,[K - 2 + \{(-1/2) + (p/2 - 1)\,a^{-1} + a\}]\,\xi(a) + 0\,(d^2),$$

and the theorem follows.

4. The Three-Stage Procedure

Let $\eta \varepsilon (0, 1)$ be specified. Start with a sample $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_k$ of size $k (\geqslant 2)$. Define $M = \max\{k, [\eta a\sigma_k^2/d^2] + 1\}$. If $M > k$, take the additional observations to complete the sample $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_M$. Let

$$N = \max\{M, [a\sigma_M^2/d^2] + 1\} \tag{4.1}$$

and if $N > M$, sample the difference to obtain $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_N$.

Now we state the following lemma, the proof of which can be obtained exactly along the lines of that of Theorem 1 in Hall (1981).

LEMMA 2 : *As* $d \to 0$,

$$E(N) = n^* + 1/2 - 2\eta^{-1} + 0\,(1) \tag{4.2}$$

$$E\mid N - n^*|^2 = 2\eta^{-1} n^* + 0\,(1) \tag{4.3}$$

$$E\mid N - n^*|^3 = 0\,(d^{-4}) \tag{4.4}$$

The main result is now stated in the following theorem.

THEOREM 3 : $\lim\limits_{d \to 0} E(N/n^*) = 1$ (4.5)

$$\lim_{d \to 0} P(\underset{\sim}{\mu} \varepsilon R_N) = \alpha + (a/n^*)\,[1/2 - 2\eta^{-1} + 2a\eta^{-1}/n^*$$

$$\{-1/2 + (p/2 - 1)\,a^{-1}\}]\,\xi\,(a) + 0\,(d^2) \tag{4.6}$$

Proof. Result (4.5) is a direct consequence of the result (4.2).

For the procedure (4.1), we obtain, as $d \to 0$,

$$\begin{aligned} P(\underset{\sim}{\mu} \varepsilon R_N) &= E[\psi\,(aN/n^*)] \\ &= \psi(a) + (a/n^*)\,E(N - n^*)\,\xi\,(a) + (a^2/2n^{*2}) \\ &\quad \cdot E\{(N - n^*)^2\}\,\{-1/2 + (p/2 - 1)\,a^{-1}\}\,\xi\,(a) \\ &\quad + (1/n^{*3})\,0\,(E\mid N - n^*|^3). \end{aligned}$$

The theorem now follows on making use of Lemma 2.

Remark 3 . One can also construct a spherical confidence region R_N^* for $\underset{\sim}{\mu}$ of diameter $2d$, i.e., $R_N^* = \{\underset{\sim}{Z} : (\underset{\sim}{\bar{X}}_N - \underset{\sim}{Z})'\,(\underset{\sim}{\bar{X}}_N - \underset{\sim}{Z}) \leqslant d^2\}$. It is easy to verify that $R_N^* \supset \{\underset{\sim}{Z} : \lambda\,(\underset{\sim}{\bar{X}}_N - \underset{\sim}{Z})'\,\Sigma^{-1}\,(\underset{\sim}{\bar{X}}_N - \underset{\sim}{Z}) \leqslant d^2\}$, where λ = maximum characteristic root of Σ. Thus, $P(\underset{\sim}{\mu} \varepsilon R_N^*) \geqslant P\,(\chi_p^2 < Nd^2/\lambda\sigma^2)$. Obviously, the 'optimal' fixed sample size needed to achieve the coverage probability α is $n \geqslant n_0$, where $n_0 = (a\lambda\sigma^2/d^2)$. The corresponding estimation procedures can be easily defined.

REFERENCES

[1] Bhattacharya, P. K. and Mallik, A. (1973) : Asymptotic normality of the stopping times of some sequential procedures, *Ann. Statist.*, 1 : 1203-1211.

[2] Chaturvedi, A. (1985) : On the multivariate analogue of sequential simultaneous estimation problem, *Jour. Indian Soc. Ag. Statist.*, 37 : 271-275.

[3] Chaturvedi, A. (1986a) : Sequential estimation of the difference of two multinormal means, *Sankhya*, A48 : 331-338.

[4] Chaturvedi. A. (1986b) : Further remarks on sequential point estimation of the mean of a multinormal population, *Sequen. Analy.*, 5(3) : 263-275.

[5] Chaturvedi A. (to appear) : On sequential procedures for the point estimation of the mean of a normal population. *Ann. Inst, Statist. Math.*

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

[7] Hall, P. (1981) : Asymptotic theory of triple sampling for sequential estimation of a mean, *Ann. Statist.*, 9 : 1229-1238.

[8] Hall. P. (1983) : Sequential estimation saving sampling operations, *Jour. Roy. Statist. Soc.*, B45 : 219-223.

[9] Mukhopadhyay, N. (1980) : A consistent and asymptotically efficient two-stage procedure to construct fixed width confidence intervals for the mean, *Metrika*, 27 : 281-284.

[10] 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.

[11] Srivastava, M. S. (1967) : On fixed-width confidence bounds for regression parameters and mean vector, *Jour. Roy. Statist. Soc.*, B29 : 132-140.

[12] Stein, C. (1945) : A two sample test for a linear hypothesis whose power is independent of the variance, *Ann. Math. Statist.*, 16 : 243-258.

[13] Srivastava, M. S. and Bhargava, R. P. (1979) : On fixed-width confidence region for the mean, *Metron*, 37 : 163-174.

[14] Wang, Y. H. (1980) : Sequential estimation of the mean of a multinormal population, *Jour. Amer. Statist. Assoc.*, 75 : 977-983.

[15] Woodroofe, M. (1977) : Second order approximations for sequential point and interval estimation, *Ann. Statist.*, 5 : 984-995,