Fixed Width Confidence Region for The Mean of A Multivariate Normal Distribution

by

Hisao Nagao
Department of Mathematical Sciences
Osaka Prefecture University

and

M.S. Srivastava
Department of Statistics
University of Toronto

Technical Report No. 9713, July 25, (1997)

TECHNICAL REPORT SERIES

University of Toronto

Department of Statistics

FIXED WIDTH CONFIDENCE REGION FOR THE MEAN OF A MULTIVARIATE NORMAL DISTRIBUTION

Hisao Nagao
Department of Mathematical Sciences
Osaka Prefecture University
Sakai, Osaka 593, Japan
e-mail: nagao@ms.osakafu-u.ac.jp

and

M.S.Srivastava
Department of Statistics
University of Toronto
Toronto, Canada
e-mail: srivasta@utstat.toronto.edu

ABSTRACT

Srivastava gave an asymptotically efficient and consistent sequential procedure to obtain a fixed-width confidence region for the mean vector of any p-dimensional random vector with finite second moments. For normally distributed random vectors, Srivastava and Bhargava showed that the specified coverage probability is attained independently of the width, the mean vector and the covariance matrix by taking an additional finite number of observations than T prescribed by the sequential rule. However, the problem of showing that $\mathrm{E}(T - n_0)$ is bounded, where n_0 is the sample size required if the covariance matrix were known, has not been available. In this paper, we not only show that it is bounded but obtain sharper estimates of $\mathrm{E}(T)$ on the lines of Woodroofe. We also give an asymptotic expansion of the coverage probability. Similar results have recently been obtained by Nagao under the assumption that the covariance matrix $\Sigma = \sum_{i=1}^{k} \sigma_i A_i$, where A_i's are known matrices. We obtain the results of this paper under the sole assumption that the largest latent root of Σ is simple.

1. INTRODUCTION

Let $X_1, X_2, \cdots, X_n, \cdots$ be iid p-dimensional random vectors which are normally distributed with unknown mean vector μ and unknown positive definite covariance matrix Σ. We wish to find a spherical confidence region of fixed diameter $2d$ and confidence coefficient $1-\alpha$. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ be the ordered latent roots of Σ, and u be such that

$$\Pr(\chi^2_{[p]} \leq u) = 1 - \alpha,$$

where $\chi^2_{[p]}$ denotes a chi-square random variable with p degrees of freedom. Then if Σ were known, we could take a sample of size

$$n_0 \geq (u\lambda_1/d^2)$$

and obtain a confidence region R_{n_0} of diameter $2d$ for μ as

$$R_{n_0} = \{z : (z - \bar{X}_{n_0})'(z - \bar{X}_{n_0}) \leq d^2\},$$

with the required confidence coefficient $1-\alpha$, namely, $\Pr(\mu \in R_{n_0}) \geq 1 - \alpha$, where $\bar{X}_{n_0} = n_0^{-1} \sum_{\alpha=1}^{n_0} X_\alpha$. For unknown covariance matrix Σ, Srivastava (1967) proposed a sequential procedure in which Σ is estimated at each stage and the sampling is stopped at

$$T = \inf\{n \geq m \mid \hat{\lambda}_{1,n} \leq d^2 n/u_n\}, \tag{1}$$

where $u_n \to u$ and $\hat{\lambda}_{1,n}$ is the largest latent root of

$$A_n = \sum_{\alpha=1}^{n} (X_\alpha - \bar{X}_n)(X_\alpha - \bar{X}_n)'/(n-1)$$

AMS 1991 subject classifications. Primary 62L12; secondary 62H12.
Key words and Phrases. Spherical confidence region; asymptotic expansion; coverage probability; largest latent root of covariance matrix; theorem on implicit function; stopping variable; reverse martingale; contrasts of mean.

and

$$\bar{X}_n = n^{-1} \sum_{\alpha=1}^{n} X_\alpha.$$

The confidence region for μ is given by

$$R_T = \{z : (z - \bar{X}_T)'(z - \bar{X}_T) \leq d^2\}.$$

For $u_n \equiv u$, Srivastava and Bhargava (1979) showed that there exists a finite k independent of d, μ and Σ such that

$$\Pr(\mu \in R_{T+k}) \geq 1 - \alpha,$$

where k may depend on α. However, the problem of showing that $\mathrm{E}(T - n_0)$ is bounded remained open. Recently, Nagao (1996) obtained sharper bounds on $\mathrm{E}(T)$ under the assumption that the covariance matrix $\Sigma = \sum_{i=1}^{k} \sigma_i A_i$, where A_i's are known symmetric matrices. This structure contains some interesting models. Also see Hyakutake, Takada and Aoshima (1995). The objective of this paper is to provide a sharper bound for $\mathrm{E}(T)$ for general Σ. We also give an asymptotic expansion of the coverage probability. This problem remained open for long time. The sharper bound on $\mathrm{E}(T)$ is obtained under the assumption that λ_1, the largest latent root of Σ is simple, that is, $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_p > 0$. When $p = 1$, this problem has been considered many times in the statistical literature, see Anscombe(1953), Chow and Robbins(1965), Starr(1966), Simons(1968), Woodroofe(1977), Srivastava and Bhargava(1969, 1979) and Nagao and Takada (1980). The last reference proposes new stopping rules for this problem and compares them with other rules. Also the last second reference gives confidence interval using t-statistic and width smaller than $2d$. Similar results on ellipsoidal confidence regions, except for asymptotic results given in Srivastava and Bhargava(1979), are not yet available.

2. BOUNDS ON THE AVERAGE SAMPLE SIZE

We shall consider the case when the sequence u_n is of the form

$$u_n = u\ell_n = u[1 + \frac{1}{n}\ell_0 + o(n^{-1})],\ \ell_0 > 0,$$

and λ_1, the largest latent root of Σ, is simple. Thus our stopping rule (1) can be written as

$$T_a = \inf\{n \geq m \mid n \geq \frac{u\ell_n\hat{\lambda}_{1,n}}{d^2}\}, \tag{2.1}$$

where $m \geq p$, and $\hat{\lambda}_{1,n}$ is the largest latent root of A_n. Since the latent roots are invariant under orthogonal transformations, we shall assume without any loss of generality that the covariance matrix Σ is a diagonal matrix Λ with its diagonal elements $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_p > 0$. Thus, it is assumed that λ_1 is simple. By Helmert's transformation, we can write $A_{n+1} = \sum_{\alpha=1}^{n} y_\alpha y_\alpha'/n$, where y_α's are iid N(0,Λ). At first, we show that

$$\begin{aligned} \hat{\lambda}_{1,n+1} &= \lambda_1 + (a_{11}^{(n)} - \lambda_1) + v_n \\ &= a_{11}^{(n)} + v_n, \end{aligned} \tag{2.2}$$

where $A_{n+1} = (a_{ij}^{(n)})$ and v_n is a slowly changing random variable. We apply a theorem on implicit function. Let $f(A_{n+1}, \ell) = |A_{n+1} - \ell I| = 0$. Regarding $\hat{\lambda}_{1,n+1}$ as a function of $A_{n+1} = (a_{ij}^{(n)})$, we expand it around Λ. Letting $a_{ij}^{(n)} = a_{ij}$, we get

$$\frac{\partial f}{\partial a_{11}}|_{A_{n+1}=\Lambda} = \begin{vmatrix} 1 & & & 0 \\ & \lambda_2 - \lambda_1 & & \\ & & \ddots & \\ 0 & & & \lambda_p - \lambda_1 \end{vmatrix} = \prod_{k=2}^{p}(\lambda_k - \lambda_1)$$

and

$$\frac{\partial f}{\partial \ell}|_{A_{n+1}=\Lambda} = -\prod_{k=2}^{p}(\lambda_k - \lambda_1).$$

The other values $\partial f/\partial a_{ij}$ on Λ are zero. We put $f_{ij}(A_{n+1},\ell) = \dfrac{\partial}{\partial a_{ij}} f(A_{n+1},\ell)$, $f_{ij:kl}(A_{n+1},\ell) = \dfrac{\partial^2}{\partial a_{ij}\partial a_{kl}} f(A_{n+1},\ell)$, $f_\ell(A_{n+1},\ell) = \dfrac{\partial}{\partial \ell} f(A_{n+1},\ell)$ and etc. Then

$$f_{ij}(A_{n+1},\ell) + f_\ell(A_{n+1},\ell)\frac{\partial \ell}{\partial a_{ij}} = 0 \tag{2.3}$$

and

$$\begin{aligned} &f_{ij:kl}(A_{n+1},\ell) + f_{ij:\ell}(A_{n+1},\ell) + f_{kl:\ell}(A_{n+1},\ell) \\ &+ f_{\ell:\ell}(A_{n+1},\ell)\frac{\partial \ell \partial \ell}{\partial a_{ij}\partial a_{kl}} + f_\ell(A_{n+1},\ell)\frac{\partial^2 \ell}{\partial a_{ij}\partial a_{kl}} = 0. \end{aligned} \tag{2.4}$$

Thus we have, as $n \to \infty$, from (2.4)

$$\frac{\partial^2 \ell}{\partial a_{ij}\partial a_{kl}} \to \begin{cases} \dfrac{2}{\lambda_j - \lambda_1} & i = k = 1, \quad l = j \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

Therefore we have by (2.3) and (2.4)

$$\begin{aligned} \hat{\lambda}_{1,n+1} &= \lambda_1 + (a_{11}^{(n)} - \lambda_1) + v_n \\ &= a_{11}^{(n)} + v_n, \end{aligned} \tag{2.6}$$

where

$$v_n = \sum_{i \le j}\sum_{k \le l} \triangle_{ij,kl}(\Lambda^*)(a_{ij} - \lambda_i\delta_{ij})(a_{kl} - \lambda_k\delta_{kl}) \tag{2.7}$$

and $\triangle_{ij,kl} = \dfrac{\partial^2 \ell}{2\partial a_{ij}\partial a_{kl}}$, δ_{ij} is the Kronecker product, and Λ^* is some point between Λ and A_{n+1}. Since $\sqrt{n}(A_{n+1} - n\Lambda)$ has a limiting normal distribution, it follows that v_n is a slowly changing random variable. For asymptotic expansions, see Nagao (1970) and Sugiura (1973).

We shall now introduce another stopping variable N_a defined by

$$N_a = \inf\{n \ge m-1 \mid Z_n > a\}, \tag{2.8}$$

where

$$Z_n = \frac{(n+1)\lambda_1}{\hat{\lambda}_{1,n+1}\ell_{n+1}} \text{ and } a \equiv n_0 = \frac{u\lambda_1}{d^2}.$$

Then, $T_a = N_a + 1$. From (2.6) we can write Z_n as

$$Z_n = S_n + \xi_n, \tag{2.9}$$

where $S_n = \sum_{\alpha=1}^{n}(2 - \frac{1}{\lambda_1}y_\alpha^{(1)^2})$ and $y_\alpha^{(1)}$ is the first component of vector Y_α and

$$\begin{aligned} \xi_n &= -2(\ell_0 - 1) - \frac{n}{\lambda_1}v_n + \frac{\lambda_1}{\lambda_1^{*3}}(n - (\ell_0 - 1 + b_n))(a_{11}^{(n)} - \lambda_1 + v_n)^2 \\ &+ \frac{(\ell_0 - 1 + b_n)}{\lambda_1}(a_{11}^{(n)} + v_n) = -2(\ell_0 - 1) + V_n \end{aligned} \tag{2.10}$$

and λ_1^* is some point between λ_1 and $\hat{\lambda}_{1,n+1}$ and $b_n = o(1)$. Here we prove that ξ_n is a slowly changing. We find that $\xi_n/n \to 0$ a.s., since $v_n \to 0$ and $a_{11}^{(n)} - \lambda_1 \to 0$ a.s. Next we show that ξ_n is uniformly continuous in probability (u.c.i.p.). Since $|\Delta_{ij,kl}(\Lambda^*)| \leq M$ and $\sqrt{n}(a_{ij}^{(n)} - \lambda_i\delta_{ij})$ are u.c.i.p. and converges in law, v_n is also u.c.i.p. Similarly we can show that $\frac{\lambda_1}{\lambda_1^{*3}}\{\sqrt{n}(a_{11}^{(n)} - \lambda_1)\}^2$ is also u.c.i.p. For other terms, here we use the following simple lemma.

Lemma 2.1. Let U_n be a u.c.i.p. and converge in law. If a real sequence a_n converges to zero, then a_nU_n is also a u.c.i.p.

Proof. $\Pr(\max_{0\leq k\leq \delta n}|a_{n+k}U_{n+k} - a_nU_n| \geq \epsilon) = \Pr((\max_{0\leq k\leq \delta n}|a_{n+k}(U_{n+k} - U_n) + (a_{n+k} - a_n)U_n| \geq \epsilon) \leq \Pr(\max_{0\leq k\leq \delta n} a_{n+k}|U_{n+k} - U_n| \geq \epsilon/2) + \Pr(\epsilon'|U_n| \geq \epsilon/2) \leq \epsilon$, for enough small $\epsilon' > 0$.

Thus ξ_n can be shown to be a u.c.i.p. Since nv_n and $\sqrt{n}(a_{11} - \lambda_1)$ converge in law, we have $\xi_n/\sqrt{n} \to 0$ in P. By Woodroofe (1982), we have

Lemma 2.2. Let $n_0 = \frac{u\lambda_1}{d^2}$, then we have

$$\frac{N_a - n_0}{\sqrt{n_0}} \to \mathrm{N}(0, 2). \tag{2.11}$$

Next we have

Lemma 2.3.

(1) $\quad \dfrac{N_a}{n_0} \longrightarrow 1$

(2) $\quad \mathrm{E}(\sup_{a\geq 1}(\dfrac{N_a}{n_0})^q) < \infty \qquad \text{for all} \quad q > 0$

(3) $\quad \int_{N_a>2a} N_a^2 \mathrm{dP} \longrightarrow 0 \quad \text{as} \quad a \longrightarrow \infty.$

We give an outline of the proof. The term (1) follows from $\xi_n/n \longrightarrow 0$ a.s., since v_n and $(a_{11}^{(n)} - \lambda_1)$ converge to 0 a.s. For (2), we note that $(\dfrac{N_a}{a})^q = (\dfrac{T_a - 1}{a})^q \leq (\dfrac{\ell_n p \mathrm{tr} A_n}{p\lambda_1})^q$ and $\dfrac{\mathrm{tr} A_n}{p\lambda_1}$ is a reverse martingale, and hence the result follows. For (3), we need to show according to Woodroofe (1991), that

$$\sum_{n=1}^{\infty} n\Pr(\xi_n \leq -n\epsilon) < \infty \quad \text{for some } 0 < \epsilon < 1. \tag{2.12}$$

We first note that,

$$\Pr(\xi_n \leq -n\epsilon) \leq \Pr(\frac{1}{\lambda_1}|v_n| \geq \epsilon/3) + \Pr(\frac{\lambda_1}{\lambda_1^{*3}}(n - (\ell_0 - 1 + b_n))$$

$$\times (a_{11}^{(n)} - \lambda_1 + v_n)^2 \geq \epsilon/3) + \Pr(\frac{(\ell_0 - 1 + b_n)}{\lambda_1}(a_{11}^{(n)} + |v_n|) \geq \epsilon/3).$$

By considering higher order moments of $(a_{ij}^{(n)} - \lambda_i\delta_{ij})^2$, we can show that $\Pr((a_{ij}^{(n)} - \lambda_i\delta_{ij})^2 \geq c\epsilon) \leq (\mathrm{const})n^{-\alpha}$ $(\alpha \geq 3)$, and hence the first term is finite. The other terms can be proved by similar consideration. Thus we have (3).

Lemma 2.4.

$$\Pr(N_a = m - 1) \sim \Pr(N_a \leq \frac{a}{2}) \sim ca^{-\frac{1}{2}p(m-1)}. \tag{2.13}$$

Proof. We have

$$\Pr(N_a = m-1) = \Pr(\frac{m\lambda_1}{\hat{\lambda}_{1,m}\ell_m} > a) \le \Pr(\frac{mp\lambda_1}{a\ell_m} \ge \mathrm{tr}A_m)$$

$$\le \Pr(\frac{m(m-1)p\lambda_1}{a\lambda_p\ell_m} > \chi^2_{[p(m-1)]}) \sim c'a^{-p(m-1)/2}.$$

and similarly,

$$\Pr(N_a = m-1) = \Pr(\frac{m\lambda_1}{a\ell_m} > \hat{\lambda}_{1,m}) \ge \Pr(\frac{m\lambda_1}{a\ell_m} > \mathrm{tr}A_m) \sim c''a^{-p(m-1)/2}.$$

Then, the proof follows as in Simons(1968), Srivastava(1973), Srivastava and Bhargava(1979) or Woodroofe(1982).

Lemma 2.5. If $p(m-1) > 2$, $N_a^* = \dfrac{N_a - n_0}{\sqrt{n_0}}$ is uniformly integrable for $a > 0$.

Proof. At first we show that

$$\int_{N_a \le a/2} N_a^{*2} \mathrm{dP} + \int_{N_a > 2a} N_a^{*2} \mathrm{dP} \to 0. \tag{2.14}$$

Since $\int_{N_a \le a/2} N_a^{*2} \mathrm{dP} \le a^{-1}(m-1-a)^2 \Pr(N_a \le a/2)$ and $p(m-1) > 2$, we get the first part of (2.14) from Lemma 2.4, the second part follows from Lemma 2.3 (3). Next we consider the probability $\Pr(N_a > a/2, N_a^* < -x)$. When $\sqrt{a}/2 \le x$, the event is empty, so we consider $0 \le x \le \sqrt{a}/2$. Thus we have $((k+1)\lambda_1)/(\hat{\lambda}_{1,k+1}\ell_{k+1}) > a$ for some $k \in (a/2, a - \sqrt{a}x]$. Since $(k+1)/(a\ell_{k+1}) \le 1/(a\ell_{k+1})(a - \sqrt{x} + 1) = 1 - \dfrac{x}{\sqrt{a}} + \dfrac{\epsilon}{a}$ for some $\epsilon > 0$. We have

$$\Pr(N_a > a/2, N_a^* < -x) \le \Pr(1 - \frac{x}{\sqrt{a}} + \frac{\epsilon}{a} > \frac{\hat{\lambda}_{1,k+1}}{\lambda_1} \; \exists k \in (a/2, a - \sqrt{a}x])$$

$$\le \Pr(\max_{k \le a} k|\frac{\hat{\lambda}_{1,k+1}}{\lambda_1} - 1| \ge \frac{x\sqrt{a}}{2} + \epsilon') \le \Pr(\max_{k \le a} k|\frac{a_{11}^{(k)}}{\lambda_1} - 1| \ge \frac{x\sqrt{a}}{4})$$

$$+\Pr(\max_{k \le a} k|\frac{v_k}{\lambda_1}| \ge \frac{x\sqrt{a}}{4}).$$

By martingale inequality, the last two formulas are less than cx^{-4}. Next let us consider $\Pr(N_a < 2a, N_a^* > x)$. When $0 \le x \le \sqrt{a}$, we have

$$\Pr(N_a < 2a, N_a^* > x) \le \Pr(N_a \ge a + \sqrt{a}x) \le \Pr(\frac{(a+\sqrt{a}x)\lambda_1}{\hat{\lambda}_{1,a+\sqrt{a}x}\ell_{a+\sqrt{a}x}} \le a)$$
$$= \Pr((1+\frac{x}{\sqrt{a}})\frac{1}{\ell_{a+\sqrt{a}x}} < \frac{\hat{\lambda}_{1,a+\sqrt{a}x}}{\lambda_1}) \le \Pr(\frac{x}{\sqrt{a}} + o(a^{-1/2}) < |\frac{\hat{\lambda}_{1,a+\sqrt{a}x}}{\lambda_1} - 1|)$$

Since $\hat{\lambda}_{1,n+1} = a_{11}^{(n)} + v_n$ in (2.6), Markov inequality yields $\Pr(N_a < 2a, N_a^* > x) \le cx^{-4}$. Thus N_a^* is uniformly integrable.

3. ASYMPTOTIC EXPANSION

In this section, we give asymptotic expansions of the mean of the time T_a and $\Pr((\bar{X}_{T_a} - \mu)'(\bar{X}_{T_a} - \mu) \le d^2)$. Since the random walk is given by $\sum_{\alpha=1}^{n}(2 - \frac{1}{\lambda_1}y_\alpha^{(1)^2})$, we have $E(2 - \frac{1}{\lambda_1}y_\alpha^{(1)^2}) = 1$ and $\mathrm{Var}(2 - \frac{1}{\lambda_1}y_\alpha^{(1)^2}) = 2$. Also by (2.5) as $n \to \infty$, we have $V_n \to -\sum_{j=2}^{p}\frac{1}{\lambda_j - \lambda_1}W_{1j}^2 + \lambda_1^{-2}W^2 + (\ell_0 - 1)$, where W_{1j} are normally distributed random variables with mean 0 and variance $\lambda_1\lambda_j$ and W is normally distributed random variable with mean 0 and variance $2\lambda_1^2$. Also we show that ξ_{n+k} for $0 \le k \le n$ is uniformly integrable. We have

$$\Pr(\max_{0\le k\le n}|\xi_{n+k}| \ge y) \le \Pr(\max_{0\le k\le n}|\frac{(n+k)}{\lambda_1}v_{n+k}| \ge y/3)$$
$$+\Pr(\max_{0\le k\le n}\frac{\lambda_1}{\lambda_1^{*3}}(n+k-(\ell_0 - 1 + b_{n+k}))(a_{11}^{(n+k)} - \lambda_1 + v_{n+k})^2 \ge y/3)$$
$$+\Pr(\max_{0\le k\le n}\frac{(\ell_0 - 1 + b_{n+k})}{\lambda_1}(a_{11}^{(n+k)} - \lambda_1 + v_{n+k})^2 \ge y/3).$$

For example, by martingale inequality, we can choose α large enough such that $\Pr(\max_{0\le k\le n}|(n+k)(a_{ij}^{(n+k)} - \lambda_i\delta_{ij})^2| \ge cy) \le \Pr(\max_{0\le k\le n}|(n+k)^2(a_{ij}^{(n+k)}\lambda_i\delta_{ij})^2| \ge cny) \le cy^{-\alpha}$. Similarly we can get the desired inequality by considering the same way. The formula $\sum_{n=1}^{\infty}\Pr(\xi_n \le -n\epsilon) < \infty$ for some $0 < \epsilon < \infty$, can be obtained by (2.11). Thus we have

Theorem 3.1. Let T_a be a stopping time defined by (2.1). Then $a \to \infty$,

$$E(T_a) = a + \rho + (\ell_0 - 2) + \sum_{j=2}^{p} \frac{\lambda_j}{\lambda_j - \lambda_1} + o(1), \tag{3.1}$$

where $\rho = \frac{3}{2} - \sum_{k=1}^{\infty} k^{-1} E(S_k^-)$.

We note that $\sum_{j=2}^{p} \frac{\lambda_j}{\lambda_j - \lambda_1}$ vanishes when $p = 1$ and this summation is negative. Also when $p = 1$, this result reduces to the case in Woodroofe(1977). Finally ,we evaluate the probability $\Pr((\bar{X}_{T_a} - \mu)'(\bar{X}_{T_a} - \mu) \le d^2) \ge \Pr(\chi^2_{[p]} \le \frac{T_a d^2}{\lambda_1}) = E\psi_p(\frac{T_a d^2}{\lambda_1}) = E\psi_p(u\frac{N_a + 1}{n_0})$, where $\psi_p(x) = \Pr(\chi^2_{[p]} \le x)$. By Lemma 2.3, $\frac{N_a}{n_0} \to 1$. We give Taylor expansion $\psi_p(ux + h)$ about $x = 1$. After some calculation, we have

$$E\psi_p(u\frac{N_a + 1}{n_0}) = \psi_p(u) + u\psi_p'(u)E(\frac{T_a}{n_0} - 1) + \frac{u^2}{2}E\psi_p''(*)E(\frac{N_a}{n_0} - 1)^2 + o(a^{-1}),$$

where $*$ is some point between u and $u\frac{N_a}{n_0}$. Since $\psi_p''(x) = \psi_p'(x)\{(p-2)x^{-1} - 1\}/2, \psi_p''(x)$ is bounded when $p - 2 \ge 2$ or $p - 2 = 0$. In these cases, by Lemma 2.5, $E\psi''(*)N_a^{*2} \to 2\psi_p''(u)$. When $p = 3$, we have $|\psi_3''(x)| \le c_1 + c_2 x^{-1/2}$ for some constant numbers c_1 and c_2. When $N_a > \frac{a}{2}$, we have $* > \frac{u}{2}$. Thus $|\psi_3''(*)| \le c_1 + c_2(u/2)^{-1/2}$. Then on $N_a > \frac{a}{2}$, $\psi_3''(*)N_a^{*2}$ is uniformly integrable. The last work is to show that $\int_{N_a \le a/2} *^{-1/2} N_a^{*2} dP \to 0$ as $a \to \infty$. Since $N_a \le a/2$, we have $* \ge \frac{N_a u}{a}$. Therefore

$$\int_{N_a \le a/2} *^{-1/2} N_a^{*2} dP \le u^{-1/2} \int_{N_a \le a/2} (\frac{a}{N_a})^{1/2} N_a^{*2} dP \le (mu)^{-1/2} a^{3/2} \Pr(N_a \le a/2)$$

When $(m-1)p > 3$, the above converges to zero. Therefore we have

Theorem 3.2. If $(m-1)p > 3$, we have

$$\Pr((\bar{X}_{T_a} - \mu)'(\bar{X}_{T_a} - \mu) \le d^2) \ge (1 - \alpha) + \frac{u}{a}\psi_p'(u)E(T_a - n_0) + \frac{u^2}{a}\psi_p''(u) + o(a^{-1}), \tag{3.2}$$

where u is the upper $100\alpha\%$ point of the chi-square distribution with p degrees of freedom.

When $p = 1$, we get Woodroofe's(1977) result.

Thus $\Pr(\mu \in R_{T_a})$ could be less than $1 - \alpha$. On the other hand, it has been shown by Srivastava and Bhargava(1979) that there exists a k such that $\Pr(\mu \in R_{T+k}) \geq 1 - \alpha$. It would thus be desirable to take some addtional observations than dictated by the stopping rule. Starr(1966) has showed that the initial sample size plays a significant rule. Thus if possible, one should start with a reasonable initial sample size.

4. SAMPLE SIZE FOR ESTIMATING CONTRASTS OF THE MEANS

We consider the following problem. We want the sample size n such that

$$\max_{||a||=1} E((a'\bar{X}_n - a'\mu)^2) + cn \tag{4.1}$$

is minimized. This problem arises when we wish to compare some linear combinations of the components of the mean vector. Then if Σ were known, the minimum sample size minimizing (4.1) is given by $n_0 = \sqrt{\lambda_1/c}$, where λ_1 is a maximum latent root of Σ. When Σ is unknown, we consider the following stopping time with $\ell_n = 1 + \frac{\ell_0}{n} + o(n^{-1}),\ \ell_0 < 0$,

$$T^* = \inf\{n \geq m \mid n \geq \sqrt{\hat{\lambda}_{1,n}\ell_n/c}\}. \tag{4.2}$$

By the similar consideration in the previous section, we define

$$N_a^* = \inf\{n \geq m - 1 \mid Z_n \geq a\},$$

where $Z_n = (n+1)\sqrt{\lambda_1/(\hat{\lambda}_{1,n+1}\ell_{n+1})}$ and $a = \sqrt{\lambda_1/c}$. Then we have $T^* = N_a^* + 1$ and $Z_n = S_n + \xi_n$, where $S_n = \sum_{\alpha=1}^n(\frac{3}{2} - \frac{1}{2\lambda_1}y_\alpha^{(1)^2})$ and

$$\xi_n = \frac{3}{2} - \frac{1}{2n\lambda_1}\sum_{\alpha=1}^n y_\alpha^{(1)^2} - \frac{(n+1)}{2\lambda_1}v_n + \frac{3}{8}(n+1)b_n^{-5/2}(\frac{1}{n\lambda_1}\sum_{\alpha=1}^n y_\alpha^{(1)^2} + \frac{v_n}{\lambda_1} - 1)^2$$

with b_n between 1 and $\hat{\lambda}_{1,n+1}/\lambda_1$. Therefore we have

Theorem 4.1. The mean of T^* in (4.2) is given by

$$\mathrm{E}(T^*) = a + \rho - (\ell_0 + 2) + \sum_{j=2}^{p} \frac{\lambda_1}{\lambda_j - \lambda_1} + o(1),$$

where $\rho = \dfrac{3}{2} - \sum_{k=1}^{\infty} \dfrac{1}{k}\mathrm{E}S_k^-$.

Next we consider the regret

$$R = \mathrm{E}(\frac{\lambda_1}{T^*} + cT^*) - \mathrm{E}(\frac{\lambda_1}{n_0} + cn_0) = cn_0\{\mathrm{E}(\frac{n_0}{T^*} + \frac{T^*}{n_0}) - 2\}.$$

Let $f(x) = x^{-1} + x$ then the regret is written by

$$R = cn_0\mathrm{E}(f(\frac{T^*}{n_0}) - f(1)).$$

Following as in section 2, we set asymptotic results of T^* as T_a as $a \to \infty$. Then $R = cn_0\mathrm{E}(b^{-3}(\dfrac{T^*}{n_0} - 1)^2\}$, where b is some point between 1 and T^*/n_0. As $a \to \infty$,

$$n_0 \int_{T^* \geq n_0/2} b^{-3}(\frac{T^*}{n_0} - 1)^2 \mathrm{dP} = \int_{T^* \geq n_0/2} b^{-3}\frac{(T^* - n_0)^2}{n_0}\mathrm{dP} \to \frac{1}{2}.$$

Also for $T^*/n_0 \leq \frac{1}{2}$, we evaluate

$$0 \leq n_0 \int_{T^*/n_0 \leq 1/2} \{\frac{n_0}{T^*} + \frac{T^*}{n_0} - 2\}\mathrm{dP} \leq n_0 \int_{T^*/n_0 \leq 1/2} \{\frac{n_0}{T^*} - \frac{3}{2}\}\mathrm{dP}$$

$$\leq n_0^2 \int_{T^*/n_0 \leq 1/2} \frac{1}{T^*}\mathrm{dP} \leq \frac{n_0}{m}\mathrm{Pr}(\frac{T^*}{n_0} \leq \frac{1}{2}) = o(1) \text{ if } p(m-1) > 2.$$

Thus we have the following theorem.

Theorem 4.2. If $p(m-1) > 2$, the regret R is given by

$$R = \frac{c}{2} + o(c).$$

As a final problem, we consider the confidence interval of any linear combination of mean vector μ for a fixed confidence coefficient $(1-\alpha)$ and fixed width $d > 0$, that is, $\Pr(|a'\bar{X}_n - a'\mu| \leq d \text{ for all } ||a|| = 1) \geq 1-\alpha$. But, by Schwarz's inequality, we have $\Pr(|a'\bar{X}_n - a'\mu| \leq d \text{ for all } ||a|| = 1) = \Pr((\bar{X}_n - \mu)'(\bar{X}_n - \mu) \leq d^2)$. Thus this problem reduces to one treated in section 2.

ACKNOWLEDGMENT

A part of this paper was completed during Nagao's visit to the Univ. of Toronto. This reseach was partially supported by Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Anscombe, F.J.(1953). Sequential estimation. Jour. Roy. Statist. Soc., B, **15**, 1-29.

Chow, Y.S. and Robbins, H. (1965). Asymptotic theory of fixed width confidence intervals for the mean. Ann. Math. Statist. **36**, 457-462.

Hyakutake, H., Takada, Y. and Aoshima, M.(1995). Fixed-size confidence regions for the multivariate mean in an intraclass correlation model. Amer. Jour. Math. & Manag. Sci. **15**, 291-308.

Nagao, H(1970). Asymptotic expansions of some test criteria for homogeneity of variances and covariance matrices from normal populations. Jour. Sci. Hiroshima Univ. Ser. A-I, **34**, 153-247.

Nagao, H.(1996). On fixed width confidence regions for multivariate normal mean when the covariance matrix has some structure. Sequential Analysis **15**, 37-46.

Nagao, H. and Takada, M.(1980). On sequential point estimation of the mean of a normal distribution. Ann. Inst. Statist. Math. **32**, 201-210.

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

Srivastava, M.S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. Jour. Roy. Statist. Soc., B, **28**, 132-140.

Srivastava, M.S. and Bhargava, R.P.(1979). On fixed width confidence region for the mean. Metron **27**, 163-174.

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

Sugiura, N.(1973). Derivatives of the characteristic root of a symmetric or a hermitian matrix with two applications in multivariate analysis. Commun. in Statist. **1**, 393-417.

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

Woodroofe, M. (1982). "Nonlinear renewal theory in sequential analysis". SIAM, Philadelphia.

Woodroofe, M. (1991). The role of renewal theory in sequential analysis. 145-167. (Handbook of sequential analysis, Dekker, Ed., Ghosh, B.K. and Sen, P.K.)