NOTES ON OPTIMAL ALLOCATION FOR FIXED SIZE CONFIDENCE REGIONS OF THE DIFFERENCE OF TWO MULTINORMAL MEANS

Hyakutake, Hiroto
Faculty of Mathematics, Kyushu University

Kawasaki, Hidefumi
Faculty of Mathematics, Kyushu University

http://hdl.handle.net/2324/12583

バージョン：published
権利関係：
NOTES ON OPTIMAL ALLOCATION FOR FIXED SIZE CONFIDENCE REGIONS OF THE DIFFERENCE OF TWO MULTINORMAL MEANS

by

Hiroto Hyakutake and Hidefumi Kawasaki

Reprinted from the Bulletin of Informatics and Cybernetics Research Association of Statistical Sciences, Vol.36

FUKUOKA, JAPAN
2004
NOTES ON OPTIMAL ALLOCATION FOR FIXED SIZE CONFIDENCE REGIONS OF THE DIFFERENCE OF TWO MULTINORMAL MEANS

By

Hiroto Hyakutake∗ and Hidefumi Kawasaki†

Abstract

We consider the problem of constructing a fixed-size confidence region of the difference of two multinormal means when the covariance matrices have intraclass correlation structure. When the covariance matrices are known, we derive an optimal allocation. A two-stage procedure is given for the problem with unknown covariance matrices.

Key Words and Phrases: fixed-size confidence interval, intraclass correlation, semi-infinite programming problem, two-stage procedure.

1. Introduction

Let \( \mathbf{x}_{i1}, \mathbf{x}_{i2}, \cdots \) be independent and identically distributed (i.i.d.) random vectors having \( p \)-variate normal distribution with mean \( \mathbf{\mu}_i \) and covariance matrix \( \Sigma_i \), \( N_p(\mathbf{\mu}_i, \Sigma_i) \), \( (i = 1, 2) \). We assume that the covariance matrices have the structure

\[
\Sigma_i = \sigma_i^2 \{(1 - \rho_i)I_p + \rho_i \mathbf{1}_p \mathbf{1}_p'\}, \quad (i = 1, 2),
\]

where \( \sigma_i > 0 \), \( 1 > \rho_i > -1/(p - 1) \), \( I_p \) is the \( p \times p \) identity matrix, and \( \mathbf{1}_p : p \times 1 = (1, \cdots, 1)' \). The eigen values of \( \Sigma_i \) are \( \tau_{i1} = \sigma_i^2 \{1 + (p - 1)\rho_i\} \) and \( \tau_{i2} = \sigma_i^2 (1 - \rho_i) \). Here \( \rho_i \) is called the intraclass correlation coefficient. This structure, which is called an intraclass correlation model or equi-variance and equi-correlation model, is applied to MANOVA for repeated measurements, see e.g. Vonesh and Chinchilli (1997). Let \( \mathbf{y}_n = \bar{\mathbf{x}}_{1,n1} - \bar{\mathbf{x}}_{2,n2} \) and \( \mathbf{\mu} = \mathbf{\mu}_1 - \mathbf{\mu}_2 \), where \( \bar{\mathbf{x}}_{i,n_i} \) is the usual sample mean based on \( n_i \) observations \( (i = 1, 2) \).

The problem is to determine the sample sizes satisfying

\[
P\{\|\mathbf{a}'(\mathbf{y}_n - \mathbf{\mu})\| \leq d, \text{ for all } \mathbf{a} \text{ such that } \mathbf{a}'\mathbf{a} = 1\} \geq 1 - \alpha,
\]

where \( d > 0 \) and \( \alpha \) \( (0 < \alpha < 1) \) are given. For one sample problem, Hyakutake, Takada and Aoshima (1995) solved the problem by a two-stage procedure and a purely sequential procedure. Aoshima (1997) and Hyakutake (1998) considered the problem

∗ Faculty of Mathematics, Kyushu University, Ropponmatsu, Fukuoka 810-8560, Japan. hyakutak@math.kyushu-u.ac.jp
† Faculty of Mathematics, Kyushu University 33, Hakozaki, Fukuoka 812-8581, Japan. kawasaki@math.kyushu-u.ac.jp
of constructing the fixed-size spherical confidence region of the difference of two multivariate normal means by a two-stage procedure. However their procedures may not be optimal as stated in Hyakutake (1998), when the covariance matrices are known. For example, when $\sigma_1^2 = 1.0$, $\rho_1 = 0$, $\sigma_2^2 = 2.5$, $\rho_2 = 0.6$, $d = 1.0$, and $\alpha = 0.05$, the required sample sizes are $n_1 = 17.97$ and $n_2 = 35.95$ by Hyakutake (1998), which improves Aoshima (1997). Based on these sample sizes, the coverage probability is 0.971. This suggests that the procedure would be improved. We determine a pair of the sample sizes $n_1^*$ and $n_2^*$ that minimizes $n_1 + n_2$ under the constraint (2).

If the covariance matrices are known, it is easy to see that

$$1 - \alpha = P[(y_n - \mu)'(\Sigma_1/n_1 + \Sigma_2/n_2)^{-1}(y_n - \mu) \leq \chi^2_p(\alpha)]$$

$$= P[\max_{b \neq 0} \frac{b'(y_n - \mu)^2}{(\Sigma_1/n_1 + \Sigma_2/n_2)b} \leq \chi^2_p(\alpha)]$$

$$= P[b'\mu \in b'y_n \pm \sqrt{\chi^2_p(\alpha)b'(\Sigma_1/n_1 + \Sigma_2/n_2)b}, \text{ for all } b \neq 0],$$

where $\chi^2_p(\alpha)$ is the upper 100$(1 - \alpha)$% point of $\chi^2_p$, which is a chi-square distribution with $p$ degrees of freedom. Hence the confidence intervals of $b'\mu$ for all $b \neq 0$ are

$$b'y_n \pm \sqrt{\chi^2_p(\alpha)b'(\Sigma_1/n_1 + \Sigma_2/n_2)b},$$

which are equivalent to

$$a'y_n \pm \sqrt{\chi^2_p(\alpha)a'(\Sigma_1/n_1 + \Sigma_2/n_2)a},$$

for all $a$ such that $a'a = 1$. It is not easy to derive the optimal sample sizes directly by (2). We consider the confidence intervals (3), say the problem is to determine $n_1^*$ and $n_2^*$ such that

$$\sqrt{\chi^2_p(\alpha)a'(\Sigma_1/n_1 + \Sigma_2/n_2)a} \leq d, \text{ for all } a (a'a = 1).$$

(4)

In Section 2, we give the optimal sample sizes, when $\sigma_i$ and $\rho_i$ are known. When $\sigma_i$ and $\rho_i$ are unknown, (4) is changed to

$$\sqrt{c_m a'(\hat{\Sigma}_1/n_1 + \hat{\Sigma}_2/n_2)a} \leq d, \text{ for all } a (a'a = 1),$$

(5)

where $\hat{\Sigma}_i$ is an estimator of $\Sigma_i$ and $c_m$ is a 100$(1 - \alpha)$% point of a distribution which is discussed in Section 3. We propose a two-stage procedure satisfying (5) and investigate its property in Section 3.

2. Optimal sample sizes

In this section, we assume that the covariance matrices are known, that is, $\sigma_i$ and $\rho_i$ are known. The following lemma gives the optimal sample sizes $n_i^*$ ($i = 1, 2$) that minimize $n_1 + n_2$ under the constraint (4).
Lemma 2.1. A sample size \((n_1^*, n_2^*)\) is a minimum if and only if there exists a unit eigen vector \(a_1(\in \mathbb{R}^p)\) of \(\Sigma_1/n_1^* + \Sigma_2/n_2^*\) such that
\[
\chi_p^2(\alpha) a_1'(\Sigma_1/n_1^* + \Sigma_2/n_2^*) a_1 = d^2
\]
and
\[
n_i^* = \frac{\chi_p^2(\alpha)}{d^2} \xi_i(\xi_1 + \xi_2), \quad (i = 1, 2)
\]
where \(\xi_i = a_1 \Sigma_i a_1\). Furthermore, when \(a_1'1_p \neq 0\), the corresponding eigen value equals \(\tau_{11}/n_1^* + \tau_{21}/n_2^*\). When \(a_1'1_p = 0\), it holds that \(\xi_i^2 = \tau_{2i} (i = 1, 2)\).

Proof. It follows from the necessary optimality condition for a semi-infinite programming problem that there exist a number \(1 \leq \ell \leq 2\) (2 is the number of the variables \(n_1\) and \(n_2\)), multipliers \(\lambda_j \geq 0\), and vectors \(a_j \in \mathbb{R}^p (1 \leq j \leq \ell)\) such that
\[
\frac{\partial L}{\partial n_i}(n_1^*, n_2^*) = 0, \quad (i = 1, 2)
\]
and
\[
\chi_p^2(\alpha) a_j'\left(\frac{\Sigma_1}{n_1^*} + \frac{\Sigma_2}{n_2^*}\right) a_j = d^2, \quad (1 \leq j \leq \ell),
\]
where \(L\) is the Lagrange function
\[
L(n_1, n_2) := n_1 + n_2 + \sum_{j=1}^\ell \lambda_j \left\{\chi_p^2(\alpha) a_j'\left(\frac{\Sigma_1}{n_1^*} + \frac{\Sigma_2}{n_2^*}\right) a_j - d^2\right\},
\]
see, e.g., Theorem 3.2 in Ben-Tal et al (1979) or Theorem 10.13.1 in Kawasaki (2004).

It is not hard to show that either \(n_1^*\) or \(n_2^*\) is negative when \(\ell = 2\), so \(\ell = 1\). Then (8) and (9) reduce to
\[
\lambda_1 \chi_p^2(\alpha) \xi_i^2/n_1^{*2} = 1 \quad (i = 1, 2)
\]
and
\[
\chi_p^2(\alpha) (\xi_1^2/n_1^* + \xi_2^2/n_2^*) = d^2,
\]
respectively. Solving (11) and (12) with respect to \((n_1^*, n_2^*, \lambda_1)\), we get (7). On the other hand, \(a_1\) is a maximum of \(\chi_p^2(\alpha) a'(\Sigma_1/n_1 + \Sigma_2/n_2) a\) subject to \(a'a = 1\). Hence there exists a Lagrange multiplier \(\eta \geq 0\) such that \(\partial M/\partial a = 0\), where
\[
M(a) = \chi_p^2(\alpha) a'\left(\Sigma_1/n_1 + \Sigma_2/n_2\right) a - \eta(a'a - 1),
\]
that is,
\[
\chi_p^2(\alpha) (\Sigma_1/n_1^* + \Sigma_2/n_2^*) a_1 - \eta a_1 = 0.
\]
Hence $a_1$ is an eigen vector of $\Sigma_1/n_1^* + \Sigma_2/n_2^*$ and its eigen value is equal to $\eta/\chi_p^2(\alpha)$. Multiplying (13) by $a_1$, we see that the eigen value equals

$$\frac{\xi_1^2}{n_1^*} + \frac{\xi_2^2}{n_2^*}. \tag{14}$$

Multiplying (13) by $1_p$, we have

$$\chi_p^2(\alpha)(1_p'\Sigma_1 a_1/n_1^* + 1_p'\Sigma_2 a_1/n_2^*) - \eta 1_p' a_1 = 0. \tag{15}$$

Since $1_p'\Sigma_i = \tau_i 1_p'$, we get

$$\{\chi_p^2(\alpha)(\tau_{11}/n_1^* + \tau_{21}/n_2^*) - \eta\}1_p' a_1 = 0. \tag{16}$$

Hence, when $1_p' a_1 \neq 0$, the eigen value equals $\tau_{11}/n_1^* + \tau_{21}/n_2^*$. When $1_p' a_1 = 0$, we get $\xi_i^2 = \tau_i (i = 1, 2)$ from the form of (1).

On the other hand, since the present semi-infinite programming problem is a convex programming problem, the necessary optimility condition turns out to be a sufficient condition for a minimum.

In Lemma 2.1, it is easy to see that $\min(\tau_{11}, \tau_{12}) \leq \xi_i^2 \leq \max(\tau_{11}, \tau_{12})$ by $a_i' a_1 = 1$. The vector $a_1$ would depend on the parameters $\sigma_i$ and $\rho_i$, so we write $a_1 = a_1(\sigma_1, \sigma_2, \rho_1, \rho_2), (i = 1, 2)$, which implies that $a_1$ is a function of $\tau_{ij}$.

3. Two-stage procedure

When $\sigma_i$ and $\rho_i$ are unknown, there is no fixed sample size procedure. We give a two-stage procedure satisfying (5). Let the first low of a $p \times p$ orthogonal matrix $Q$ defined by $(1/\sqrt{p}, \cdots, 1/\sqrt{p})$, and define $z_{ir} = (z_{ir,1}, \cdots, z_{ir,p})' = Q(x_{ir} - \mu_i), r = 1, 2, \cdots$ and $i = 1, 2$. Then $z_{ir}$'s are i.i.d. according to $N_p(0, D_i)$, where $D_i = \text{diag}(\tau_{11}, \tau_{12}, \cdots, \tau_{12})$.

First take the initial sample size $m(> p)$ from each population and compute

$$\tau_{11} = \frac{1}{m - 1} \sum_{r = 1}^{m} (z_{ir,1} - \bar{z}_{i,1})^2 \quad \text{and} \quad \tau_{12} = \frac{1}{(p - 1)(m - 1)} \sum_{r = 1}^{p} \sum_{j = 1}^{m} (z_{ir,j} - \bar{z}_{i,j})^2, \tag{17}$$

where $(\bar{z}_{i,1}, \cdots, \bar{z}_{i,p})' = \sum_{r = 1}^{m} z_{ir}/m$. Then $\hat{\tau}_{11}$ and $\hat{\tau}_{12}$ are independent and are unbiased estimators of $\tau_{11}$ and $\tau_{12}$, respectively, see e.g., Hyakutake, Takada and Aoshima (1995). The estimator of $\Sigma_i$ is $S_i = Q'D_iQ$, where $D_i = \text{diag}(\hat{\tau}_{11}, \hat{\tau}_{12}, \cdots, \hat{\tau}_{12})$, say $S_i$ is used in $\Sigma_i$ of (5). Hence $\hat{\xi}_i^2 = \bar{a}_1 S_i \bar{a}_1$ is an estimator of $\xi_i^2$, where $\bar{a}_1 = a_1(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\rho}_1, \hat{\rho}_2)$, which is expressed by $\hat{\tau}_{11} = \hat{\sigma}_1^2 (1 + (p - 1)\hat{\rho}_1)$ and $\hat{\tau}_{12} = \hat{\sigma}_1^2 (1 - \hat{\rho}_1)$. It would hold that $\min(\hat{\tau}_{11}, \hat{\tau}_{12}) \leq \hat{\xi}_i^2 \leq \max(\hat{\tau}_{11}, \hat{\tau}_{12})$ as in Section 2.

The total sample sizes are defined by

$$N_i = \max\left\{m, \left[\frac{c_m \hat{\xi}_i (\hat{\xi}_1 + \hat{\xi}_2)}{d^2}\right] + 1\right\}, \quad (i = 1, 2), \tag{18}$$

where $[q]$ denotes the greatest integer less than $q$ and $c_m$ is a solution of an equation $H(c_m) = 1 - \alpha$. $H(c_m)$ is a cumulative distribution function (c.d.f. of
\[ \nu_1 v_{10} / \min(v_{11}, v_{21}) + \nu_2 v_{20} / \min(v_{12}, v_{22}), \]

where \( v_{ij} \) and \( v_{2i} \) are independently distributed as \( \chi^2_{\nu_i} \) with \( \nu_1 = m-1 \) and \( \nu_2 = (p-1)\nu_1 \), and the conditional distributions of \( v_{01} \) and \( v_{02} \) given \( \xi_1, \xi_2 \) are \( \chi^2_{\nu_i} \) and \( \chi^2_{\nu - p-1} \), respectively.

Next we take \( N_i - m \) additional observations from each population and compute the sample mean \( \bar{x}_{i,N_i} \), \( i = 1, 2 \). Then we have the following theorem.

**Theorem 3.1.** If \( N_1 \) and \( N_2 \) are determined by (17), then (5) is satisfied.

**Proof.** If it is shown that

\[ P((y_N - \mu)'(S_1/n_1 + S_2/n_2)^{-1}(y_N - \mu) \leq c_m) \geq 1 - \alpha, \]

where \( y_N = \bar{x}_{1,N_1} - \bar{x}_{2,N_2} \) and \( \Sigma_i = \hat{\sigma}^2_i \{(1 - \hat{\rho}_i)I_p + \hat{\rho}_i 1_p 1'_p \} \), then (5) is satisfied by Lemma 1.

Since \( u = (u_1, u_2)' = Q(y_N - \mu) \) is distributed as \( N(0, D_1/N_1 + D_2/N_2) \) given \( (N_1, N_2) \), the conditional distributions of \( v_{01} = u_1^2 / (\tau_{11}/N_1 + \tau_{21}/N_2) \) and \( v_{02} = u_2' u_2 / (\tau_{12}/N_1 + \tau_{22}/N_2) \) are \( \chi^2_{\nu} \) and \( \chi^2_{\nu - p-1} \), respectively. Hence we have

\[
P\{ (y_N - \mu)'(S_1/n_1 + S_2/n_2)^{-1}(y_N - \mu) \leq c_m \}
= \frac{u_1^2}{\tau_{11}/N_1 + \tau_{21}/N_2} + \frac{u_2' u_2}{\tau_{12}/N_1 + \tau_{22}/N_2} \leq c_m
\]

where \( q_j = (\tau_{1j}/N_1)(\tau_{1j}/N_1 + \tau_{2j}/N_2) \) \( j = 1, 2 \). Since \( q_j v_{1j} + (1 - q_j) v_{2j} \geq \min(v_{1j}, v_{2j}) \), we have

\[
P\{ \nu_1 v_{10} / (q_1 v_{11} + (1 - q_1) v_{21}) + \nu_2 v_{20} / (q_2 v_{12} + (1 - q_2) v_{22}) \leq c_m \}
\geq P\{ \nu_1 v_{10} / v_{11} + \nu_2 v_{20} / v_{22} \leq c_m \}
= 1 - \alpha,
\]

which completes the proof.

Next we discuss an asymptotic property of the procedure. It is easy to see that \( v_{ij}/v_j \to 1 \) \( i, j = 1, 2 \) almost surely as \( m \to \infty \) by \( \tau_{ij} \to \tau_{ij} \) almost surely as \( m \to \infty \), see e.g., Hyakutake, Takada and Aoshima (1995). Then the limiting distribution of (19) is \( \chi^2_{p} \), say \( c_m \to \chi^2_{p}(\alpha) \) as \( m \to \infty \). Under the assumption that \( m \to \infty \) and \( d^2m \to 0 \) as \( d \to 0 \), we have
\[
\lim_{d \to 0} E\left(\frac{N_1 + N_2}{n_1^* + n_2^*}\right) = 1,
\]
that is the two-stage procedure based on (18) is asymptotic efficient. This can be shown by the same method as in Takada (1988), so the proof is omitted.

**Acknowledgement**

Authors would like to thank Professor Yanagawa for his comments and encouragement. Thanks also due to the referee for his valuable comments.

**References**


*Received October 16, 2003*

*Revised July 2, 2004*