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

https://doi.org/10.5109/12583

出版情報：Bulletin of informatics and cybernetics. 36, pp.131-136, 2004-12. 統計科学研究会

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

Bulletin of Informatics and Cybernetics, Vol. 36, 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 $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \cdots$ be independent and identically distributed (i.i.d.) random vectors having p-variate normal distribution with mean $\boldsymbol{\mu}_i$ and covariance matrix Σ_i, $N_p(\boldsymbol{\mu}_i, \Sigma_i)$, $(i = 1, 2)$. We assume that the covariance matrices have the structure

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

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 Σ_i are $\tau_{i1} = \sigma_i^2\{1 + (p-1)\rho_i\}$ and $\tau_{i2} = \sigma_i^2(1-\rho_i)$. Here ρ_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 $\boldsymbol{y}_n = \bar{\boldsymbol{x}}_{1,n_1} - \bar{\boldsymbol{x}}_{2,n_2}$ and $\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$, where $\bar{\boldsymbol{x}}_{i,n_i}$ is the usual sample mean based on n_i observations $(i = 1, 2)$.

The problem is to determine the sample sizes satisfying

$$P\{|\boldsymbol{a}'(\boldsymbol{y}_n - \boldsymbol{\mu})| \le d, \text{ for all } \boldsymbol{a} \text{ such that } \boldsymbol{a}'\boldsymbol{a} = 1\} \ge 1 - \alpha, \tag{2}$$

where $d > 0$ and α $(0 < \alpha < 1)$ 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 multinormal 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

$$\begin{aligned} 1-\alpha &= P[(\boldsymbol{y}_n - \boldsymbol{\mu})'(\Sigma_1/n_1 + \Sigma_2/n_2)^{-1}(\boldsymbol{y}_n - \boldsymbol{\mu}) \le \chi_p^2(\alpha)] \\ &= P[\max_{\boldsymbol{b}\neq\boldsymbol{0}} \frac{\{\boldsymbol{b}'(\boldsymbol{y}_n - \boldsymbol{\mu})\}^2}{\boldsymbol{b}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{b}} \le \chi_p^2(\alpha)] \\ &= P[\boldsymbol{b}'\boldsymbol{\mu} \in \boldsymbol{b}'\boldsymbol{y}_n \pm \sqrt{\chi_p^2(\alpha)\boldsymbol{b}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{b}}, \text{ for all } \boldsymbol{b} \neq \boldsymbol{0}], \end{aligned}$$

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

$$\boldsymbol{b}'\boldsymbol{y}_n \pm \sqrt{\chi_p^2(\alpha)\boldsymbol{b}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{b}},$$

which are equivalent to

$$\boldsymbol{a}'\boldsymbol{y}_n \pm \sqrt{\chi_p^2(\alpha)\boldsymbol{a}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{a}}, \tag{3}$$

for all $\boldsymbol{a}$ such that $\boldsymbol{a}'\boldsymbol{a} = 1$. 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

$$\sqrt{\chi_p^2(\alpha)\boldsymbol{a}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{a}} \le d, \quad \text{for all } \boldsymbol{a}\ (\boldsymbol{a}'\boldsymbol{a} = 1). \tag{4}$$

In Section 2, we give the optimal sample sizes, when σ_i and ρ_i are known. When σ_i and ρ_i are unknown, (4) is changed to

$$\sqrt{c_m\boldsymbol{a}'(\hat{\Sigma}_1/n_1 + \hat{\Sigma}_2/n_2)\boldsymbol{a}} \le d, \quad \text{for all } \boldsymbol{a}\ (\boldsymbol{a}'\boldsymbol{a} = 1), \tag{5}$$

where $\hat{\Sigma}_i$ is an estimator of Σ_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 asuume that the covariance matrices are known, that is, σ_i and ρ_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 $\boldsymbol{a}_1 (\in R^p)$ of $\Sigma_1/n_1^* + \Sigma_2/n_2^*$ such that*

$$\chi_p^2(\alpha)\boldsymbol{a}_1'(\Sigma_1/n_1^* + \Sigma_2/n_2^*)\boldsymbol{a}_1 = d^2 \tag{6}$$

and

$$n_i^* = \frac{\chi_p^2(\alpha)}{d^2}\xi_i(\xi_1 + \xi_2), \quad (i = 1, 2) \tag{7}$$

where $\xi_i^2 = \boldsymbol{a}_1\Sigma_i\boldsymbol{a}_1$. Furthermore, when $\boldsymbol{a}_1'\mathbf{1}_p \neq 0$, the corresponding eigen value equals $\tau_{11}/n_1^* + \tau_{21}/n_2^*$. When $\boldsymbol{a}_1'\mathbf{1}_p = 0$, it holds that $\xi_i^2 = \tau_{i2}$ $(i = 1, 2)$.

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 $\boldsymbol{a}_j \in R^p$ $(1 \leq j \leq \ell)$ such that

$$\frac{\partial L}{\partial n_i}(n_1^*, n_2^*) = 0, \quad (i = 1, 2) \tag{8}$$

and

$$\chi_p^2(\alpha)\boldsymbol{a}_j'\Big(\frac{\Sigma_1}{n_1^*} + \frac{\Sigma_2}{n_2^*}\Big)\boldsymbol{a}_j = d^2, \quad (1 \leq j \leq \ell), \tag{9}$$

where L is the Lagrange function

$$L(n_1, n_2) := n_1 + n_2 + \sum_{j=1}^{\ell}\lambda_j\Big\{\chi_p^2(\alpha)\boldsymbol{a}_j'\Big(\frac{\Sigma_1}{n_1^*} + \frac{\Sigma_2}{n_2^*}\Big)\boldsymbol{a}_j - d^2\Big\}, \tag{10}$$

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

$$\lambda_1\chi_p^2(\alpha)\xi_i^2/n_i^{*2} = 1 \quad (i = 1, 2) \tag{11}$$

and

$$\chi_p^2(\alpha)(\xi_1^2/n_1^* + \xi_2^2/n_2^*) = d^2, \tag{12}$$

respectively. Solving (11) and (12) with respect to $(n_1^*, n_2^*, \lambda_1)$, we get (7). On the other hand, $\boldsymbol{a}_1$ is a maximum of $\chi_p^2(\alpha)\boldsymbol{a}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{a}$ subject to $\boldsymbol{a}'\boldsymbol{a} = 1$. Hence there exists a Lagrange multiplier $\eta \geq 0$ such that $\partial M/\partial\boldsymbol{a} = \mathbf{0}$, where

$$M(\boldsymbol{a}) = \chi_p^2(\alpha)\boldsymbol{a}'(\Sigma_1/n_1 + \Sigma_2/n_2)\boldsymbol{a} - \eta(\boldsymbol{a}'\boldsymbol{a} - 1),$$

that is,

$$\chi_p^2(\alpha)(\Sigma_1/n_1^* + \Sigma_2/n_2^*)\boldsymbol{a}_1 - \eta\boldsymbol{a}_1 = \mathbf{0}. \tag{13}$$

Hence $\boldsymbol{a}_1$ 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 $\boldsymbol{a}_1$, we see that the eigen value equals

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

Multiplying (13) by $\mathbf{1}_p$, we have

$$\chi_p^2(\alpha)(\mathbf{1}_p'\Sigma_1\boldsymbol{a}_1/n_1^* + \mathbf{1}_p'\Sigma_2\boldsymbol{a}_1/n_2^*) - \eta\mathbf{1}_p'\boldsymbol{a}_1 = 0. \tag{15}$$

Since $\mathbf{1}_p'\Sigma_i = \tau_{i1}\mathbf{1}_p'$, we get

$$\{\chi_p^2(\alpha)(\tau_{11}/n_1^* + \tau_{21}/n_2^*) - \eta\}\mathbf{1}_p'\boldsymbol{a}_1 = 0. \tag{16}$$

Hence, when $\mathbf{1}_p'\boldsymbol{a}_1 \neq 0$, the eigen value equals $\tau_{11}/n_1^* + \tau_{21}/n_2^*$. When $\mathbf{1}_p'\boldsymbol{a}_1 = 0$, we get $\xi_i^2 = \tau_{i2}$ $(i = 1, 2)$ 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 miniimum.

In Lemma 2.1, it is easy to see that $\min(\tau_{i1}, \tau_{i2}) \leq \xi_i^2 \leq \max(\tau_{i1}, \tau_{i2})$ by $\boldsymbol{a}_1'\boldsymbol{a}_1 = 1$. The vector $\boldsymbol{a}_1$ would depend on the parameters σ_i and ρ_i, so we write $\boldsymbol{a}_1 = \boldsymbol{a}_1(\sigma_1, \sigma_2, \rho_1, \rho_2)$, $(i = 1, 2)$, which implies that $\boldsymbol{a}_1$ is a function of τ_{ij}.

3. Two-stage procedure

When σ_i and ρ_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 $\boldsymbol{z}_{ir} = (z_{ir,1}, \cdots, z_{ir,p})' = Q(\boldsymbol{x}_{ir} - \boldsymbol{\mu}_i)$, $r = 1, 2, \cdots$ and $i = 1, 2$. Then $\boldsymbol{z}_{ir}$'s are i.i.d. according to $N_p(\mathbf{0}, D)$, where $D_i = \mathrm{diag}(\tau_{i1}, \tau_{i2}, \cdots, \tau_{i2})$.

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

$$\hat{\tau}_{i1} = \frac{1}{m-1}\sum_{r=1}^{m}(z_{ir,1} - \bar{z}_{i,1})^2 \quad \text{and} \quad \hat{\tau}_{i2} = \frac{1}{(p-1)(m-1)}\sum_{j=2}^{p}\sum_{r=1}^{m}(z_{ir,j} - \bar{z}_{i,j})^2, \tag{17}$$

where $(\bar{z}_{i,1}, \cdots, \bar{z}_{i,p})' = \sum_{r=1}^{m}\boldsymbol{z}_{ir}/m$. Then $\hat{\tau}_{i1}$ and $\hat{\tau}_{i2}$ are independent and are unbiased estimators of τ_{i1} and τ_{i2}, respectively, see e.g., Hyakutake, Takada and Aoshima (1995). The estimator of Σ_i is $S_i = Q'\hat{D}_iQ$, where $\hat{D}_i = \mathrm{diag}(\hat{\tau}_{i1}, \hat{\tau}_{i2}, \cdots, \hat{\tau}_{i2})$, say S_i is used in $\hat{\Sigma}_i$ of (5). Hence $\hat{\xi}_i^2 = \hat{\boldsymbol{a}}_1S_i\hat{\boldsymbol{a}}_1$ is an estimator of ξ_i^2, where $\hat{\boldsymbol{a}}_1 = \boldsymbol{a}_1(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\rho}_1, \hat{\rho}_2)$, which is expressed by $\hat{\tau}_{i1} = \hat{\sigma}_i^2\{1 + (p-1)\hat{\rho}_i\}$ and $\hat{\tau}_{i2} = \hat{\sigma}_i^2(1 - \hat{\rho}_i)$. It would hold that $\min(\hat{\tau}_{i1}, \hat{\tau}_{i2}) \leq \hat{\xi}_i^2 \leq \max(\hat{\tau}_{i1}, \hat{\tau}_{i2})$ as in Section 2.

The total sample sizes are defined by

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

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 cummulative distribution function (c.d.f.) of

$$\nu_1 v_{01}/\min(v_{11}, v_{21}) + \nu_2 v_{02}/\min(v_{12}, v_{22}), \tag{19}$$

where v_{1i} 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 conditonal distributions of v_{01} and v_{02} given $\hat{\xi}_1, \hat{\xi}_2$ are χ^2_1 and χ^2_{p-1}, respectively.

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

$$P[(\boldsymbol{y}_N - \boldsymbol{\mu})'(S_1/n_1 + S_2/n_2)^{-1}(\boldsymbol{y}_N - \boldsymbol{\mu}) \le c_m] \ge 1 - \alpha, \tag{20}$$

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

Since $\boldsymbol{u} = (u_1, \boldsymbol{u}_2')' = Q(\boldsymbol{y}_N - \boldsymbol{\mu})$ is distributed as $N(\mathbf{0}, D_1/N_1 + D_2/N_2)$ given (N_1, N_2), the conditional distributions of $v_{01} = u_1^2/(\tau_{11}/N_1 + \tau_{21}/N_2)$ and $v_{02} = \boldsymbol{u}_2'\boldsymbol{u}_2/(\tau_{12}/N_1 + \tau_{22}/N_2)$ are χ^2_1 and χ^2_{p-1}, respectively. Hence we have

$$\begin{aligned}
& P\{(\boldsymbol{y}_N - \boldsymbol{\mu})'(S_1/N_1 + S_2/N_2)^{-1}(\boldsymbol{y}_N - \boldsymbol{\mu}) \le c_m\} \\
= \ & P\Big\{\frac{u_1^2}{\tau_{11}/N_1 + \tau_{21}/N_2}\frac{\tau_{11}/N_1 + \tau_{21}/N_2}{\hat{\tau}_{11}/N_1 + \hat{\tau}_{21}/N_2} + \frac{\boldsymbol{u}_2'\boldsymbol{u}_2'}{\tau_{12}/N_1 + \tau_{22}/N_2}\frac{\tau_{12}/N_1 + \tau_{22}/N_2}{\hat{\tau}_{12}/N_1 + \hat{\tau}_{22}/N_2} \le c_m\Big\} \\
= \ & P\Big\{\frac{v_{01}}{q_1\hat{\tau}_{11}/\tau_{11} + (1 - q_1)\hat{\tau}_{21}/\tau_{21}} + \frac{v_{02}}{q_2\hat{\tau}_{12}/\tau_{12} + (1 - q_2)\hat{\tau}_{22}/\tau_{22}} \le c_m\Big\} \\
= \ & P\{\nu_1 v_{01}/(q_1 v_{11} + (1 - q_1)v_{21}) + \nu_2 v_{02}/(q_2 v_{12} + (1 - q_2)v_{22}) \le c_m\},
\end{aligned}$$

where $q_j = (\tau_{1j}/N_1)(\tau_{1j}/N_1 + \tau_{2j}/N_2)$ $(j = 1, 2)$. Since $q_j v_{1j} + (1 - q_1)v_{2j} \ge \min(v_{1j}, v_{2j})$, we have

$$\begin{aligned}
& P\{\nu_1 v_{01}/(q_1 v_{11} + (1 - q_1)v_{21}) + \nu_2 v_{02}/(q_2 v_{12} + (1 - q_2)v_{22}) \le c_m\} \\
\ge \ & P\{\nu_1 v_{01}/\min(v_{11}, v_{21}) + \nu_2 v_{02}/\min(v_{12}, v_{22}) \le c_m\} \\
= \ & 1 - \alpha,
\end{aligned}$$

which completes the proof.

Next we discuss an asymptotic property of the procedure. It is easy to see that $v_{ij}/\nu_j \to 1$ $(i, j = 1, 2)$ almost surely as $m \to \infty$ by $\hat{\tau}_{ij} \to \tau_{ij}$ almost surely as $m \to \infty$, see e.g., Hyakutake, Takada and Aoshima (1995). Then the limiting distribution of (19) is χ^2_p, say $c_m \to \chi^2_p(\alpha)$ as $m \to \infty$. Under the assumption that $m \to \infty$ and $d^2 m \to 0$ as $d \to 0$, we have

$$\lim_{d \to 0} \frac{E(N_1 + N_2)}{n_1^* + n_2^*} = 1,$$

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

Aoshima, M. (1997). A two-stage procedure for fixed-size confidence region when covariance matrices have some structures, *J. Japan Statist. Soc.*, 27, 59-67.

Ben-Tal, A., Teboulle, M. and Zowe, J. (1979). Second order necessary optimality conditions for semi-infinite programming problems, Lecture Notes in Control and Information Sciences 15, Semi-Infinite Programming (ed. R. Hettich), Springer.

Hyakutake, H. (1998). Fixed-size confidence region of the difference of two multinormal means when the covariance matrices have structure, *J. Japan Statist. Soc.*, 28, 175-180.

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

Kawasaki, H. (2004). Extremal Problems (Japanese), Yokohama Publishers.

Vonesh, E.F. and Chinchilli, V.M. (1997). Linear and Nonlinear Models for the Analysis of Repeated Measurements, Dekker.

Takada, Y. (1988). Two-stage procedures for a multivariate normal distribution, *Kumamoto J. Math.*, 1, 1-8.

Received October 16, 2003
Revised July 2, 2004