

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

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

where  $\sigma_i > 0$ ,  $1 > \rho_i > -1/(p-1)$ ,  $I_p$  is the  $p \times p$  identity matrix, and  $\mathbf{1}_p : p \times 1 = (1, \dots, 1)'$ . 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  $\mathbf{y}_n = \bar{\mathbf{x}}_{1,n_1} - \bar{\mathbf{x}}_{2,n_2}$  and  $\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ , 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

$$P\{|\mathbf{a}'(\mathbf{y}_n - \boldsymbol{\mu})| \leq d, \text{ for all } \mathbf{a} \text{ such that } \mathbf{a}'\mathbf{a} = 1\} \geq 1 - \alpha, \quad (2)$$

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 multi-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

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

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

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

which are equivalent to

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

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

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

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

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

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  $\mathbf{a}_1 (\in R^p)$  of  $\Sigma_1/n_1^* + \Sigma_2/n_2^*$  such that*

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

and

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

where  $\xi_i^2 = \mathbf{a}_1 \Sigma_i \mathbf{a}_1$ . Furthermore, when  $\mathbf{a}_1' \mathbf{1}_p \neq 0$ , the corresponding eigen value equals  $\tau_{11}/n_1^* + \tau_{21}/n_2^*$ . When  $\mathbf{a}_1' \mathbf{1}_p = 0$ , 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  $\mathbf{a}_j \in R^p$  ( $1 \leq j \leq \ell$ ) such that

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

and

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

where  $L$  is the Lagrange function

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

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

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

and

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

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

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

that is,

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

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

$$\xi_1^2/n_1^* + \xi_2^2/n_2^*. \quad (14)$$

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

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

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

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

Hence, when  $\mathbf{1}_p' \mathbf{a}_1 \neq 0$ , the eigen value equals  $\tau_{11}/n_1^* + \tau_{21}/n_2^*$ . When  $\mathbf{1}_p' \mathbf{a}_1 = 0$ , 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 optimality condition turns out to be a sufficient condition for a minimum.

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  $\mathbf{a}_1' \mathbf{a}_1 = 1$ . The vector  $\mathbf{a}_1$  would depend on the parameters  $\sigma_i$  and  $\rho_i$ , so we write  $\mathbf{a}_1 = \mathbf{a}_1(\sigma_1, \sigma_2, \rho_1, \rho_2)$ , ( $i = 1, 2$ ), which implies that  $\mathbf{a}_1$  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  $(1/\sqrt{p}, \dots, 1/\sqrt{p})$ , and define  $\mathbf{z}_{ir} = (z_{ir,1}, \dots, z_{ir,p})' = Q(\mathbf{x}_{ir} - \boldsymbol{\mu}_i)$ ,  $r = 1, 2, \dots$  and  $i = 1, 2$ . Then  $\mathbf{z}_{ir}$ 's are i.i.d. according to  $N_p(\mathbf{0}, D)$ , where  $D_i = \text{diag}(\tau_{i1}, \tau_{i2}, \dots, \tau_{ip})$ .

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

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

where  $(\bar{z}_{i,1}, \dots, \bar{z}_{i,p})' = \sum_{r=1}^m \mathbf{z}_{ir}/m$ . Then  $\hat{\tau}_{i1}$  and  $\hat{\tau}_{i2}$  are independent and are unbiased estimators of  $\tau_{i1}$  and  $\tau_{i2}$ , respectively, see e.g., Hyakutake, Takada and Aoshima (1995). The estimator of  $\Sigma_i$  is  $S_i = Q' \hat{D}_i Q$ , where  $\hat{D}_i = \text{diag}(\hat{\tau}_{i1}, \hat{\tau}_{i2}, \dots, \hat{\tau}_{ip})$ , say  $S_i$  is used in  $\hat{\Sigma}_i$  of (5). Hence  $\hat{\xi}_i^2 = \hat{\mathbf{a}}_1' S_i \hat{\mathbf{a}}_1$  is an estimator of  $\xi_i^2$ , where  $\hat{\mathbf{a}}_1 = \mathbf{a}_1(\hat{\sigma}_1, \hat{\sigma}_2, \hat{\rho}_1, \hat{\rho}_2)$ , 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

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

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

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

where  $v_{1i}$  and  $v_{2i}$  are independently distributed as  $\chi_{\nu_i}^2$  with  $\nu_1 = m-1$  and  $\nu_2 = (p-1)\nu_1$ , and the conditional distributions of  $v_{01}$  and  $v_{02}$  given  $\hat{\xi}_1, \hat{\xi}_2$  are  $\chi_1^2$  and  $\chi_{p-1}^2$ , 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

$$P[(\mathbf{y}_N - \boldsymbol{\mu})'(S_1/n_1 + S_2/n_2)^{-1}(\mathbf{y}_N - \boldsymbol{\mu}) \leq c_m] \geq 1 - \alpha, \quad (20)$$

where  $\mathbf{y}_N = \bar{\mathbf{x}}_{1,N_1} - \bar{\mathbf{x}}_{2,N_2}$  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  $\mathbf{u} = (u_1, \mathbf{u}_2')' = Q(\mathbf{y}_N - \boldsymbol{\mu})$  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  $v_{02} = \mathbf{u}_2' \mathbf{u}_2/(\tau_{12}/N_1 + \tau_{22}/N_2)$  are  $\chi_1^2$  and  $\chi_{p-1}^2$ , respectively. Hence we have

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

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

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

which completes the proof.

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

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

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*