

A Multivariate Analogue of Pooling of Data

Sato, Sokuro
Saga University

<https://doi.org/10.5109/13007>

出版情報：統計数理研究. 10 (3/4), pp.61-76, 1962-10. Research Association of Statistical Sciences

バージョン：

権利関係：



A MULTIVARIATE ANALOGUE OF POOLING OF DATA

By

Sōkuro SATO

(Received August 16th, 1961)

§ 1. Summary and introduction

In this paper the author attempts to extend the inferences of a k -dimensional mean vector on the basis of pooling data in a multivariate normal case and to give concretely certain formulae and properties of them.

The principle and some statistical methods of pooling data have been discussed by Bancroft [1], Kitagawa [1], [2], Bennet [1], Asano [1] and various authors and developed mainly in case when the observations were obtained from the univariate populations. Recently Asano and the author of this paper [1] dealt the inference of a mean vector and a dispersion matrix on the basis of pooling data in the bivariate case. And in their Introduction it has been noted that the inference of a mean vector may be also expressed by the similar formulae and properties in the general k -dimensional multivariate case.

Hence this paper may be considered to be partially an extension of the previous paper of Asano and Sato [1]. Type 1 of this paper gives us the inference of population mean vector with known population dispersion matrix and Type 2 with unknown population dispersion matrix.

In conclusion the author wishes his hearty thanks to Prof. T. Kitagawa and Mr. C. Asano for their kind suggestions and encouragement.

§ 2. Pooling of sample mean vector

2.1. Type 1. (The inference of population mean vector with "known" population dispersion matrix)

Let $O_{N_1} : (\mathbf{x}^{(1)}, \mathbf{x}^{(1)}, \dots, \mathbf{x}_{N_1}^{(1)})$ be a random sample of N_1 vector observations from a k -variate non-degenerate normal population $N[\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}]$ and let $O_{N_2} : (\mathbf{x}^{(2)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}_{N_2}^{(2)})$ be another random sample of N_2 from some k -variate normal population $N[\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}]$. The values of these two common population dispersion matrices are known to us, but the populations have not necessarily the same population mean vector. The distinction between these populations may be regarded, however, as hypothetical. Let us suppose that we may pool the two sample mean vectors and form an estimate vector

of the assumed same population mean vectors simultaneously in case when testing a hypothesis that $\mu^{(1)} = \mu^{(2)}$ shows that the hypothesis cannot be rejected.

Our rule of inference procedure is as follows:

- (i) Let $\bar{\mathbf{x}}^{(i)}$ be sample mean vector defined by O_{Ni} , ($i=1, 2$)
- (ii) Let the statistic $\mathbf{U}'\mathbf{U}$ be defined by

$$\mathbf{U}'\mathbf{U} = (\bar{\mathbf{x}}^{(2)} - \bar{\mathbf{x}}^{(1)})' \left(\frac{1}{N_1} + \frac{1}{N_2} \right)^{-1} \Sigma^{-1} (\bar{\mathbf{x}}^{(2)} - \bar{\mathbf{x}}^{(1)}),$$

where $\mathbf{x}^{(i)} = \sum_{j=1}^{N_i} \mathbf{x}_j^{(i)} / N_i$.

- (iii) Then let us define the statistic $\bar{\mathbf{x}}$ in the following way.

$$(a) \quad \bar{\mathbf{x}} = \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2}, \quad \text{if} \quad \mathbf{U}'\mathbf{U} \leq \chi_k^2(\alpha)$$

$$(b) \quad \bar{\mathbf{x}} = \bar{\mathbf{x}}^{(1)}, \quad \text{if} \quad \mathbf{U}'\mathbf{U} > \chi_k^2(\alpha),$$

where $\chi_k^2(\alpha)$ means the significance value of χ^2 distribution with a significance level α in case when the degrees of freedom is equal to k .

For the sake of convenience, throughout this paper, the notation of $Pr.\{\bar{\mathbf{x}} < \mathbf{u}\}$ shows at the same time the probability that each element of the vector $\bar{\mathbf{x}}$ does not over the corresponding element of the vector \mathbf{u} namely, the distribution functions of the respective elements of $\bar{\mathbf{x}}$ simultaneously.

Theorem 1.1. *The distribution function of $\bar{\mathbf{x}}$ is given by*

$$(1.1) \quad Pr.\{\bar{\mathbf{x}} < \mathbf{u}\}$$

$$\begin{aligned} &= \int \cdots \int \left(\frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \right) \int \cdots \int \left(\frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \prod_{i=1}^k ds_i \right) \\ &\quad \begin{array}{l} r_k < m_k / l_{kk} \\ r_{k-1} < (m_{k-1} - l_{k-1,k} r_k) / l_{k-1,k-1} \\ \dots \dots \dots \\ r_1 < (m_1 - \sum_{j=2}^k l_{1j} r_j) / l_{11} \end{array} \quad \begin{array}{l} (\bar{\mathbf{s}} + \mathbf{L}^{-1}\mathbf{a})'(\bar{\mathbf{s}} + \mathbf{L}^{-1}\mathbf{a}) \leq \chi_k^2(\alpha) \end{array} \\ &+ \int \cdots \int \left(\frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \right) \cdot \int \cdots \int \left(\frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \prod_{i=1}^k ds_i \right) \\ &\quad \begin{array}{l} (\bar{\mathbf{s}} + \mathbf{L}^{-1}\mathbf{a})'(\bar{\mathbf{s}} + \mathbf{L}^{-1}\mathbf{a}) < \chi_k^2(\alpha) \\ r_k < n_k / l_{kk} \\ r_{k-1} < (n_{k-1} - l_{k-1,k} r_k) / l_{k-1,k-1} \\ \dots \dots \dots \\ r_1 < (n_1 - \sum_{j=2}^k l_{1j} r_j) / l_{11} \end{array} \end{aligned}$$

where $m_i \equiv \frac{1}{\sqrt{\sigma_{ii}}} \frac{N_1 + N_2}{N_1 + N_2} u_i - \frac{N_1 \mu_i^{(1)} + N_2 \mu_i^{(2)}}{N_1 + N_2}$ ($i=1, 2, \dots, k$), $n_i \equiv \frac{1}{\sqrt{\sigma_{ii}}} \frac{N_1 + N_2}{N_1 + N_2} (u_i - \mu_i^{(1)})$

+ $\frac{1}{\sqrt{N_2}} \sum_{j=1}^k l_{ij} s_j$, and l_{ij} 's are elements of a matrix \mathbf{L} such that

$$\mathbf{L} = \begin{pmatrix} l_{11} & l_{12} & l_{13} & \cdots & l_{1k} \\ 0 & l_{22} & l_{23} & \cdots & l_{2k} \\ 0 & 0 & l_{33} & \cdots & l_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & l_{kk} \end{pmatrix}, \quad l_{ii} > 0 \quad (i=1, 2, \dots, k),$$

$$L'R^{-1}L = \text{unit matrix}$$

where

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{12} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1k} & \rho_{2k} & \cdots & 1 \end{pmatrix}$$

Proof. Now let us put for a moment

$$(1.2) \quad \mathbf{y}^{(i)'} = (y_1^{(i)}, y_2^{(i)}, \dots, y_k^{(i)}) = \left(\frac{\bar{x}_1^{(i)} - \mu_1^{(i)}}{1/\sigma_{11}/\sqrt{N_i}}, \frac{\bar{x}_2^{(i)} - \mu_2^{(i)}}{1/\sigma_{22}/\sqrt{N_i}}, \dots, \frac{\bar{x}_k^{(i)} - \mu_k^{(i)}}{1/\sigma_{kk}/\sqrt{N_i}} \right) \quad (i=1, 2),$$

then the joint elementary probability of $(y_1^{(1)}, y_2^{(1)}, \dots, y_k^{(1)}, y_1^{(2)}, y_2^{(2)}, \dots, y_k^{(2)})$ may be given by

$$(1.3) \quad \frac{1}{(2\pi)^k |\mathbf{R}|} \exp \left[-\frac{1}{2} \sum_{i=1}^2 \mathbf{y}^{(i)'} \mathbf{R}^{-1} \mathbf{y}^{(i)} \right] \prod_{i=1}^2 \prod_{j=1}^k dy_j^{(i)}.$$

Further, the following orthogonal transformation

$$(1.4) \quad p = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \left(-\frac{\mathbf{y}^{(1)}}{\sqrt{N_1}} + \frac{\mathbf{y}^{(2)}}{\sqrt{N_2}} \right), \quad q = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \left(\frac{\mathbf{y}^{(1)}}{\sqrt{N_2}} + \frac{\mathbf{y}^{(2)}}{\sqrt{N_1}} \right)$$

gives us the elementary probability of $p_1, p_2, \dots, p_k, q_1, q_2, \dots, q_k$

$$(1.5) \quad h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) = \frac{1}{(2\pi)^k |\mathbf{R}|} \exp \left[-\frac{1}{2} (\mathbf{p}' \mathbf{R}^{-1} \mathbf{p} + \mathbf{q}' \mathbf{R}^{-1} \mathbf{q}) \right] \prod_{i=1}^k (dp_i dq_i)$$

Now let the sample space be divided into two distinct sets D_1 and D_2 which are defined as a domain $\mathbf{U}'\mathbf{U} = (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) \leq \chi_k^2(\alpha)$,

where $\mathbf{a} = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \mathbf{c}^{-1} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})$

$$\mathbf{c} = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{kk}} \end{pmatrix},$$

and the complement respectively. Then we may and shall decompose the

probability (1.1) into two parts.

$$(1.6) \quad Pr.\{\bar{\mathbf{x}} < \mathbf{u}\} = Pr.\left\{\frac{N_1\bar{\mathbf{x}}^{(1)} + N_2\bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1\right\} + Pr.\{\mathbf{x}^{(1)} < \mathbf{u}, D_2\},$$

where $Pr.\{\bar{\mathbf{x}} < \mathbf{u}, D_1\}$ means the probability that two events happen simultaneously and similar for other cases.

On the other hand in view of (1.2) and (1.4), we may obtain that

$$(1.7) \quad \frac{N_1\bar{\mathbf{x}}^{(1)} + N_2\bar{\mathbf{x}}^{(2)}}{N_1 + N_2} = \mathbf{E} + \mathbf{F}\mathbf{q} = \boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} + \frac{N_2}{N_1 + N_2}(\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})$$

$$\bar{\mathbf{x}} = \boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} - \mathbf{G}\mathbf{F}\mathbf{p},$$

where

$$\mathbf{E} = \frac{N_1\boldsymbol{\mu}^{(1)} + N_2\boldsymbol{\mu}^{(2)}}{N_1 + N_2}, \quad \mathbf{F} = \frac{1}{\sqrt{N_1 + N_2}} \begin{pmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{kk}} \end{pmatrix}, \quad \text{and } \mathbf{G} = \sqrt{\frac{N_2}{N_1}}.$$

Then the first term of (1.6) may be given by

$$(1.8) \quad Pr.\left\{\frac{N_1\bar{\mathbf{x}}^{(1)} + N_2\bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1\right\} = \int \cdots \int h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i).$$

$$\begin{aligned} & \mathbf{q} < \mathbf{F}^{-1}(\mathbf{u} - \mathbf{E}) \\ & (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1}(\mathbf{p} + \mathbf{a}) \leq \chi_k^2(\alpha) \end{aligned}$$

Now let $F_{\alpha\beta}(t)$ be a matrix whose diagonal elements are all one, (α, β) element ($\alpha \neq \beta$) is t , and another elements are all zero. Then

$$\mathbf{R}_1 = \mathbf{R}^{-1} \mathbf{E}_{12} \left(-\frac{\rho^{12}}{\rho^{11}} \right) \mathbf{E}_{13} \left(-\frac{\rho^{13}}{\rho^{11}} \right) \cdots \mathbf{E}_{1k} \left(-\frac{\rho^{1k}}{\rho^{11}} \right)$$

$$= \begin{pmatrix} \rho^{11} & \rho^{12} & \cdots & \rho^{1k} \\ \rho^{21} & \rho^{22} & \cdots & \rho^{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{k1} & \rho^{k2} & \cdots & \rho^{kk} \end{pmatrix} \begin{pmatrix} 1 & -\frac{\rho^{12}}{\rho^{11}} & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 & \cdots & -\frac{\rho^{1k}}{\rho^{11}} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

becomes a matrix whose elements of first row are all zero except (1.1) element. Similarly, all elements of the second row of the matrix

$$\mathbf{R}_2 = \mathbf{R}_1 \mathbf{E}_{23} \left(-\frac{\rho^{23}}{\rho^{22}} \right) \mathbf{E}_{24} \left(-\frac{\rho^{24}}{\rho^{22}} \right) \cdots \mathbf{E}_{2k} \left(-\frac{\rho^{2k}}{\rho^{22}} \right)$$

where \mathbf{I} is the unit matrix of k degree, and $l_{ii} > 0$ ($i=1, 2, \dots, k$). By using the following transformations

$$(1.8) \quad \mathbf{p} = \mathbf{L}\mathbf{s}, \quad \mathbf{q} = \mathbf{L}\mathbf{r},$$

where $\mathbf{s}' = (s_1, s_2, \dots, s_k)$, $\mathbf{r}' = (r_1, r_2, \dots, r_k)$, the probability (1.8) may be written in the following way :

$$(1.9) \quad \text{Pr.} \left\{ \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1 \right\} = \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \\ \mathbf{L}\mathbf{r} < \mathbf{F}^{-1}(\mathbf{u} - \mathbf{E}) \\ \cdot \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \prod_{i=1}^k ds_i \\ (\mathbf{L}\mathbf{s} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{L}\mathbf{s} + \mathbf{a}) \leq \chi_k^2(\alpha)$$

Let us put $(\mathbf{F}^{-1}(\mathbf{u} - \mathbf{E}))' = \mathbf{m}' = (m_1, m_2, \dots, m_k)$ where $m_i = \frac{\sqrt{N_1 + N_2}}{\sqrt{\sigma_{ii}}}$

$$\left(u_i - \frac{N_1 \mu_i^{(1)} + N_2 \mu_i^{(2)}}{\sqrt{N_1 + N_2}} \right),$$

then

$$(1.10) \quad \text{Pr.} \left\{ \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1 \right\} \\ = \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \prod_{i=1}^k ds_i \\ r_k < m_k / l_{kk} \quad (\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})' (\mathbf{s} + \mathbf{L}^{-1}\mathbf{a}) \leq \chi_k^2(\alpha) \\ r_{k-1} < (m_{k-1} - l_{k-1, k} r_k) / l_{k-1, k-1} \\ \dots \dots \dots \\ r_1 < (m_1 - \sum_{j=2}^k l_{1j} r_j) / l_{11}$$

The second term of (1.1) may be similarly given as follows :

$$(1.11) \quad \text{Pr.} \{ \bar{\mathbf{x}} < \mathbf{u}, D_2 \} = \int \cdots \int h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) \\ \mathbf{q} < \mathbf{F}^{-1}(\mathbf{u} - \mu^{(1)}) + \mathbf{G}\mathbf{p} \\ (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) > \chi_k^2(\alpha) \\ = \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \cdot \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \prod_{i=1}^k dr_i \prod_{i=1}^k ds_i \\ \mathbf{L}\mathbf{r} < \mathbf{F}^{-1}(\mathbf{u} - \mu^{(1)}) + \mathbf{G}\mathbf{L}\mathbf{s} \\ (\mathbf{L}\mathbf{s} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{L}\mathbf{s} + \mathbf{a}) > \chi_k^2(\alpha) \\ = \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \cdot \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \prod_{i=1}^k ds_i \\ (\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})' (\mathbf{s} + \mathbf{L}^{-1}\mathbf{a}) > \chi_k^2(\alpha) \quad r_k < n_k / l_{kk} \\ r_{k-1} < (n_{k-1} - l_{k-1, k} r_k) / l_{k-1, k-1} \\ \dots \dots \dots \\ r_1 < (n_1 - \sum_{j=2}^k l_{1j} r_j) / l_{11}$$

where $\mathbf{n}' = (n_1, n_2, \dots, n_k) \equiv (\mathbf{F}^{-1}(\mathbf{u} - \boldsymbol{\mu}^{(1)}) + \mathbf{GLs})'$.

Thus we obtain the second term of the right-hand side of (1.1). In combination of (1.10) and (1.11), we may obtain (1.1) to be proved.

Theorem 1.2. *The mean vector $E\{\bar{\mathbf{x}}\}$ and the mean square error $M.S.E.\{\bar{\mathbf{x}}\}$ of the estimate $\bar{\mathbf{x}}$ are given by*

$$(1.12) \quad E\{\bar{\mathbf{x}}\} = \boldsymbol{\mu}^{(1)} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \text{Pr.}\{D_1\} - \sqrt{\frac{N_2}{(N_1 + N_2)N_1}} \mathbf{c} I_{D_2}\{\mathbf{p}\}$$

and

$$(1.13) \quad M.S.E.\{\bar{\mathbf{x}}\} = \frac{1}{N_1 + N_2} \boldsymbol{\Sigma} + \left(\frac{N_2}{N_1 + N_2} \right)^2 (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' \text{Pr.}\{D_1\} \\ + \frac{N_2}{(N_1 + N_2)N_1} I_{D_2}\{\mathbf{c} \mathbf{p} \mathbf{p}' \mathbf{c}'\}$$

where we put

$$(1.14) \quad \text{Pr.}\{D_1\} = \int \cdots \int \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{p}' \mathbf{R}^{-1} \mathbf{p} \right] \prod_{i=1}^k dp_i \\ (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) \leq \chi_k^2(\alpha)$$

$$(1.15) \quad I_{D_2}\{\mathbf{p}\} = \int \cdots \int \mathbf{p} \cdot \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{p}' \mathbf{R}^{-1} \mathbf{p} \right] \prod_{i=1}^k dp_i \\ (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) > \chi_k^2(\alpha)$$

Proof. Making use of the notation in the enunciation and proof of theorem 1.1., we may write

$$(1.16) \quad E\{\bar{\mathbf{x}}\} = \int \cdots \int_{D_1} \boldsymbol{\mu}^{(1)} + \mathbf{F} \mathbf{q} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) \\ + \int \cdots \int_{D_2} (\boldsymbol{\mu}^{(1)} + \mathbf{F} \mathbf{q} - \mathbf{G} \mathbf{F} \mathbf{p}) h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) \\ = \boldsymbol{\mu}^{(1)} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \text{Pr.}\{D_1\} - \sqrt{\frac{N_2}{(N_1 + N_2)N_1}} \mathbf{c} I_{D_2}\{\mathbf{p}\},$$

because, $E\{\mathbf{q}\} = 0$.

Thus the mean vector (1.12) is obtained. The mean square error matrix may be obtained from a following relation:

$$(1.17) \quad E\{\bar{\mathbf{x}} \bar{\mathbf{x}}'\} = \int \cdots \int_{D_1} \left(\boldsymbol{\mu}^{(1)} + \mathbf{F} \mathbf{q} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \right) \left(\boldsymbol{\mu}^{(1)} + \mathbf{F} \mathbf{q} \right. \\ \left. + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \right)' h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i)$$

$$\begin{aligned}
& + \int \cdots \int_{D_2} (\boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} - G\mathbf{F}\mathbf{p}) (\boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} - G\mathbf{F}\mathbf{p})' h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) \\
= & \int \cdots \int_{D_1} \left\{ \boldsymbol{\mu}^{(1)} \boldsymbol{\mu}^{(1)'} + \mathbf{F}\mathbf{q}\mathbf{q}'\mathbf{F}' + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(1)} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \boldsymbol{\mu}^{(1)'}) \right. \\
& + \left(\frac{N_2}{N_1 + N_2} \right)^2 (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(1)} \mathbf{q}'\mathbf{F}' + \mathbf{F}\mathbf{q}\boldsymbol{\mu}^{(1)'}) \\
& + \frac{N_2}{N_1 + N_2} (\mathbf{F}\mathbf{q} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \mathbf{q}'\mathbf{F}') h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) \\
& + \int \cdots \int_{D_2} \left\{ \boldsymbol{\mu}^{(1)} \boldsymbol{\mu}^{(1)'} + \mathbf{F}\mathbf{q}\mathbf{q}'\mathbf{F}' + G^2 \mathbf{F}\mathbf{p}\mathbf{p}'\mathbf{F}' + (\boldsymbol{\mu}^{(1)} \mathbf{q}'\mathbf{F}' + \mathbf{F}\mathbf{q}\boldsymbol{\mu}^{(1)'}) - G(\boldsymbol{\mu}^{(1)} \mathbf{p}'\mathbf{F}' + \mathbf{F}\mathbf{p}\boldsymbol{\mu}^{(1)'}) \right. \\
& \left. \left. - G(\mathbf{F}\mathbf{q}\mathbf{p}'\mathbf{F}' + \mathbf{F}\mathbf{p}\mathbf{q}'\mathbf{F}') h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) \right\} \right.
\end{aligned}$$

Then substituting the following relations

$$I_{D_1+D_2}\{\mathbf{q}\mathbf{q}'\} = \int \cdots \int_{D_1+D_2} \mathbf{q}\mathbf{q}' h(\mathbf{p}', \mathbf{q}') \prod_{i=1}^k (dp_i dq_i) = \mathbf{R},$$

$$I_{D_1+D_2}\{\mathbf{q}\} = I_{D_1}\{\mathbf{q}\} = 0, \quad I_{D_2}\{\mathbf{q}\mathbf{p}'\} = I_{D_2}\{\mathbf{p}\mathbf{q}'\} = 0$$

for (1.17), we may obtain that

$$\begin{aligned}
E\{\bar{\mathbf{x}}\mathbf{x}'\} &= \boldsymbol{\mu}^{(1)} \boldsymbol{\mu}^{(1)'} + \frac{1}{N_1 + N_2} \boldsymbol{\Sigma} + \frac{N_2}{N_1 + N_2} \{ \boldsymbol{\mu}^{(1)} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \boldsymbol{\mu}^{(1)'} \} Pr.\{D_1\} \\
&+ \left(\frac{N_2}{N_1 + N_2} \right)^2 (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' Pr.\{D_1\} + \frac{N_2}{N_1} I_{D_2}\{\mathbf{F}\mathbf{p}\mathbf{p}'\mathbf{F}'\} \\
&- \sqrt{\frac{N_2}{N_1}} I_{D_2}\{ \boldsymbol{\mu}^{(1)} \mathbf{p}'\mathbf{F}' + \mathbf{F}\mathbf{p}\boldsymbol{\mu}^{(1)'} \}.
\end{aligned}$$

Thus the mean square error matrix $M.S.E.\{\bar{\mathbf{x}}\}$ of the estimate $\bar{\mathbf{x}}$ may be obtained in the following manner.

$$\begin{aligned}
M.S.E.\{\bar{\mathbf{x}}\} &= E\{(\bar{\mathbf{x}} - \boldsymbol{\mu}^{(1)}) (\bar{\mathbf{x}} - \boldsymbol{\mu}^{(1)})'\} \\
&= E\{\bar{\mathbf{x}}\mathbf{x}'\} - \boldsymbol{\mu}^{(1)} E\{\mathbf{x}'\} - E\{\bar{\mathbf{x}}\} \boldsymbol{\mu}^{(1)'} + \boldsymbol{\mu}^{(1)} \boldsymbol{\mu}^{(1)'} \\
&= \frac{1}{N_1 + N_2} \boldsymbol{\Sigma} + \left(\frac{N_2}{N_1 + N_2} \right)^2 (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' Pr.\{D_1\} \\
&+ \frac{N_2}{(N_1 + N_2)N_1} I_{D_2}\{\mathbf{c}\mathbf{p}\mathbf{p}'\mathbf{c}'\}.
\end{aligned}$$

Corollary 1.1. *Specially when $\boldsymbol{\mu}^{(2)} = \boldsymbol{\mu}^{(1)}$, $E\{\bar{\mathbf{x}}\} = \boldsymbol{\mu}^{(1)}$, that is, $\bar{\mathbf{x}}$ is an unbiased estimate of $\boldsymbol{\mu}^{(1)}$, but the variance components and the absolute value of covariance components of the mean square error matrix are greater*

than $\sigma_{ii}/(N_1+N_2)$ and $\sigma_{ij}/(N_1+N_2)$ and less than σ_{ii}/N_1 and σ_{ij}/N_1 respectively, $(i, j=1, 2)$.

Corollary 1.2. *Specially when N_1 is very large compared with N_2 , that is $N_1 \gg N_2$, the bias of the estimate $\bar{\mathbf{x}}$ becomes small, and the mean square error matrix becomes $\Sigma/(N_1+N_2)$ nearly.*

Corollary 1.3. *Specially when $\mu^{(2)} = \mu^{(1)}$, the variance component of the M.S.E. $\{\bar{\mathbf{x}}\}$ in case of multivariate pooling of data becomes smaller than the variance of the corresponding estimate in case of univariate pooling of data. The difference is given by*

$$\frac{N_2 \sigma_{ii}}{(N_1+N_2) N_1} [I_{D_2} \{p_i^2\} - I_{D_2^*} \{p_i^2\}] < 0 \quad (i=1, 2, \dots, k)$$

where

$$I_{D_2} \{p_i^2\} = \int \dots \int p_i^2 \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{p}' \mathbf{R}^{-1} \mathbf{p} \right] \prod_{i=1}^k dp_i$$

$$(\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) > \chi_{k^2}^2(\alpha)$$

$$I_{D_2^*} \{p_i^2\} = \int p_i^2 \frac{1}{(2\pi)^{1/2}} \exp \left[-\frac{1}{2} p_i^2 \right] dp_i$$

$$p_i + a_i > (\chi_1^2(\alpha))^{1/2}$$

2.2. Type 2. (The case with "unknown population dispersion matrix" in Type 1)

Let $O_{N_1} : (\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \dots, \mathbf{x}_{N_1}^{(1)})$ be a random sample of N_1 vector observations from a k -variate non-degenerate normal population $N[\mu^{(1)}, \Sigma]$ and let $O_{N_2} : (\mathbf{x}_1^{(2)}, \mathbf{x}_2^{(2)}, \dots, \mathbf{x}_{N_2}^{(2)})$ be another random sample of N_2 from some k -variate normal population $N[\mu^{(2)}, \Sigma]$. In this section, the values of these two common population dispersion matrices are assumed to be unknown to us, but the populations have not necessarily the same population mean vector. Our attempt of the present inference procedure is the same in Type 1.

Now our rule of inference procedure is as follows:

(i) Let $\bar{\mathbf{x}}^{(i)}$ be sample mean vector defined by O_{N_i} , $i=1, 2$. and let S be the unbiased estimate of common population dispersion given by

$$S = (N_1 + N_2 - 2)^{-1} \left\{ \sum_{i=1}^{N_1} (\mathbf{x}_i^{(1)} - \bar{\mathbf{x}}^{(1)}) (\mathbf{x}_i^{(1)} - \bar{\mathbf{x}}^{(1)})' + \sum_{j=1}^{N_2} (\mathbf{x}_j^{(2)} - \bar{\mathbf{x}}^{(2)}) (\mathbf{x}_j^{(2)} - \bar{\mathbf{x}}^{(2)})' \right\}$$

(ii) Let the statistic T^2 be defined by

$$T^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{\mathbf{x}}^{(2)} - \bar{\mathbf{x}}^{(1)})' S^{-1} (\bar{\mathbf{x}}^{(2)} - \bar{\mathbf{x}}^{(1)})$$

(iii) Then let us define the statistic $\bar{\mathbf{x}}$ in the following way:

$$(a) \quad \bar{\mathbf{x}} = \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2}, \text{ if } \frac{N_1 + N_2 - k - 1}{k} \cdot \frac{T^2}{N_1 + N_2 - 2} \leq F_{k, N_1 + N_2 - k - 1}(\alpha)$$

$$(b) \cdot \bar{\mathbf{x}} = \bar{\mathbf{x}}^{(1)} \quad , \quad \text{if } \frac{N_1 + N_2 - k - 1}{k} \cdot \frac{T^2}{N_1 + N_2 - 2} > F_{k, N_1 + N_2 - k - 1}(\alpha)$$

where $F_{k, N_1 + N_2 - k - 1}(\alpha)$ denotes α -percent point of F -distribution with the pair of degrees of freedom $(k, N_1 + N_2 - k - 1)$.

Theorem 2.1. *The distribution function of $\bar{\mathbf{x}}$ is given by*

(2.1) $Pr.\{\bar{\mathbf{x}} < \mathbf{u}\}$

$$\begin{aligned} &= \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp\left[-\frac{1}{2} \sum_{i=1}^k r_i^2\right] \prod_{i=1}^k dr_i \\ &\quad \begin{array}{l} r_k < m_k / l_{kk} \\ r_{k-1} < (m_{k-1} - l_{k-1, k} r_k) / l_{k-1, k-1} \\ \cdots \\ r_1 < (m_1 - k \sum_{j=2}^k l_{1j} r_j) / l_{11} \end{array} \\ &\cdot \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp\left[-\frac{1}{2} \sum_{i=1}^k s_i^2\right] \frac{(\chi^2/2)^{(N_1 + N_2 - k - 1)/2 - 1} \exp[-\chi^2/2]}{2\Gamma\{(N_1 + N_2 - k - 1)/2\}} \left(\prod_{i=1}^k ds_i\right) d\chi^2 \\ &\quad \frac{(s + L^{-1}a)'(s + L^{-1}a)}{\chi^2} \leq \frac{k}{N_1 + N_2 - k - 1} F_{k, N_1 + N_2 - k - 1}(\alpha) \\ &+ \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp\left[-\frac{1}{2} \sum_{i=1}^k s_i^2\right] \frac{(\chi^2/2)^{(N_1 + N_2 - k - 1)/2 - 1} \exp[-\chi^2/2]}{2\Gamma\{(N_1 + N_2 - k - 1)/2\}} \\ &\quad \frac{(s + L^{-1}a)'(s + L^{-1}a)}{\chi^2} > \frac{k}{N_1 + N_2 - k - 1} F_{k, N_1 + N_2 - k - 1}(\alpha) \\ &\cdot \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp\left[-\frac{1}{2} \sum_{i=1}^k r_i^2\right] \left(\prod_{i=1}^k dr_i\right) \left(\prod_{i=1}^k ds_i\right) d\chi^2 \\ &\quad \begin{array}{l} r_k < n_k / l_{kk} \\ r_{k-1} < (n_{k-1} - l_{k-1, k} r_k) / l_{k-1, k-1} \\ \cdots \\ r_1 < (n_1 - k \sum_{i=2}^k l_{1j} r_j) / l_{11} \end{array} \end{aligned}$$

Proof. Let us use of the notation in the enunciation of this section and proof of Theorem 1.1.

Let \mathbf{D} be a nonsingular matrix such that $\mathbf{D}\Sigma\mathbf{D}' = \mathbf{I}$ and define

$$\mathbf{Y} = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} (\bar{\mathbf{x}}^{(2)} - \bar{\mathbf{x}}^{(1)}), \quad \mathbf{Y}^* = \begin{pmatrix} Y_1^* \\ Y_2^* \\ \vdots \\ Y_k^* \end{pmatrix} = \mathbf{D}\mathbf{Y}, \quad \mathbf{S}^* = \mathbf{D}\mathbf{S}\mathbf{D}'.$$

Then \mathbf{Y} is distributed according to $N\left[\sqrt{\frac{N_1 N_2}{N_1 + N_2}} (\mathbf{u}^{(2)} - \mathbf{u}^{(1)}), \Sigma\right]$, and $(N_1 + N_2 - 2)\mathbf{S}$ is distributed as $\sum_{\beta=1}^{N_1 + N_2 - 2} \mathbf{Z}_\beta \mathbf{Z}_\beta'$ with the \mathbf{Z}_β independent, each with distribution $N[\mathbf{O}, \Sigma]$, and \mathbf{Z}_β is independent of $\bar{\mathbf{x}}^{(1)}$, $\bar{\mathbf{x}}^{(2)}$, thereby is independent of \mathbf{p} and $\mathbf{q} \cdot \mathbf{Y}^*$ is distributed according to $N\left[\sqrt{\frac{N_1 N_2}{N_1 + N_2}} \mathbf{D}(\mathbf{u}^{(2)} - \mathbf{u}^{(1)}), \Sigma\right]$, and $(N_1 + N_2 - 2)\mathbf{S}^*$ is distributed as $\sum_{\beta=1}^{N_1 + N_2 - 2} \mathbf{Z}_\beta^* \mathbf{Z}_\beta^{*'} = \sum_{\beta=1}^{N_1 + N_2 - 2} \mathbf{D}\mathbf{Z}_\beta (\mathbf{D}\mathbf{Z}_\beta)'$ with

the Z_β^* independent, each with distribution $N[O, I]$, and Z_β^* is independent of Y^* , p and q .

Let the first row of a $k \times k$ orthogonal matrix Q be defined by

$$q_{1i} = \frac{Y_i^*}{\sqrt{Y^{*'} Y^*}} \quad i=1, 2, \dots, k$$

this is permissible since $\sum_{i=1}^k q_{1i}^2 = 1$. Since Q depends on Y^* it is a random matrix. Now let

$$V = QY^*, \quad B = Q(N_1 + N_2 - 2)S^*Q'$$

Then the conditional distribution of B given Q is that of $\sum_{\beta=1}^{N_1+N_2-2} W_\beta W_\beta'$ where conditionally the W_β defined by QZ_β^* are independent, each with distribution $N[O, I]$.

If we put

$$B = (b_{ij}) = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{(1)} \\ b_{(1)}' & B_{22} \end{pmatrix}$$

then $b_{11 \cdot 2 \dots k} = b_{11} - b_{(1)} B_{22}^{-1} b_{(1)}'$ is conditionally distributed as $\sum_{\beta=1}^{N_1+N_2-2-(k-1)} X_\beta^2$ where conditionally the X_β are independent, each with the distribution $N(0, 1)$; that is, $b_{11 \cdot 2 \dots k}$ is conditionally distributed as χ^2 with $N_1 + N_2 - k - 1$ degrees of freedom. Since the conditional distribution of $b_{11 \cdot 2 \dots k}$ does not depend on Q , it is unconditionally distributed as χ^2 , and is independent of p, q .

On the other hand, if we put

$$Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1k} \\ q_{21} & q_{22} & \cdots & q_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ q_{k1} & q_{k2} & \cdots & q_{kk} \end{pmatrix}$$

then $V_1 = \sum_{i=1}^k q_{1i} Y_i^* = \sqrt{Y^{*'} Y^*}$, $V_j = \sum_{i=1}^k q_{ji} Y_i^* = \sqrt{Y^{*'} Y^*} \sum_{i=1}^k q_{ji} q_{1i} = 0$ and

$$\begin{aligned} V_1^2 &= Y^{*'} Y^* = Y' D' D Y = Y' (D \Sigma)^{-1} (D \Sigma) D' D Y \\ &= Y' \Sigma^{-1} D^{-1} (D \Sigma D') D Y = Y' \Sigma^{-1} Y \\ &= \sqrt{\frac{N_1 N_2}{N_1 + N_2}} (\bar{x}^{(2)} - \bar{x}^{(1)})' \Sigma^{-1} \sqrt{\frac{N_1 N_2}{N_1 + N_2}} (\bar{x}^{(2)} - \bar{x}^{(1)}) \\ &= (c(p + a))' \Sigma^{-1} (c(p + a)) \\ &= (p + a)' c' \Sigma^{-1} c (p + a) \end{aligned}$$

$$\begin{aligned}
 &= (\mathbf{p} + \mathbf{a})' (\mathbf{c}^{-1} \Sigma (\mathbf{c}')^{-1})^{-1} (\mathbf{p} + \mathbf{a}) \\
 &= (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a})
 \end{aligned}$$

has a noncentral χ^2 -distribution with k degrees of freedom. Then $T^2/(N_1 + N_2 - 2) = \mathbf{Y}' \mathbf{S}^{-1} \mathbf{Y} / (N_1 + N_2 - 2) = \mathbf{Y}^*{}' \mathbf{S}^{*-1} \mathbf{Y}^* / (N_1 + N_2 - 2)$

$$\begin{aligned}
 &= \mathbf{V}' \mathbf{B}^{-1} \mathbf{V} = (\mathbf{V}_1 0 \dots 0) \begin{pmatrix} b^{11} & b^{12} & \dots & b^{1k} \\ b^{21} & b^{22} & \dots & b^{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b^{k1} & b^{k2} & \dots & b^{kk} \end{pmatrix} \begin{pmatrix} V_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = V_1^2 b^{11} = \frac{V_1^2}{b_{11 \cdot 2}},
 \end{aligned}$$

where $(b_{ij}) = \mathbf{B}^{-1}$, is distributed as the ratio of a noncentral χ^2 and an independent χ^2 . Therefore if $\boldsymbol{\mu}^{(2)} = \boldsymbol{\mu}^{(1)}$, $[T^2/(N_1 + N_2 - 2)] \cdot [(N_1 + N_2 - k - 1)/k]$ is distributed as a F with k and $N_1 + N_2 - k - 1$ degrees of freedom. From above relation the inference rule of this theorem follows.

The joint probability density of $(\mathbf{p}, \mathbf{q}, \chi^2)$ where $\chi^2 = b_{11 \cdot 2 \dots k}$ may be given by

$$\begin{aligned}
 h(\mathbf{p}', \mathbf{q}') f_{N_1 + N_2 - k - 1}(\chi^2) &= \frac{1}{(2\pi)^{k/2}} \mathbf{R} \exp \left[-\frac{1}{2} (\mathbf{p}' \mathbf{R}^{-1} \mathbf{p} + \mathbf{q}' \mathbf{R}^{-1} \mathbf{q}) \right] \\
 &\quad \frac{(\chi^2/2)^{(N_1 + N_2 - k - 1)/2 - 1} \exp[-\chi^2/2]}{2\Gamma[(N_1 + N_2 - k - 1)/2]}
 \end{aligned}$$

Now let the sample space be divided into two mutually distinct sets D_1 and D_2 which are defined as domains $\{(\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a})\} / \chi^2 \mid [(N_1 + N_2 - k - 1)/k] \leq F_{k, N_1 + N_2 - k - 1}(\alpha)$ and the complement respectively. Then

$$(2.2) \quad Pr. \{ \bar{\mathbf{x}} < \mathbf{u} \} = Pr. \left\{ \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1 \right\} + Pr. \{ \bar{\mathbf{x}}^{(1)} < \mathbf{u}, D_2 \}$$

The first term of the right-hand side of (2.2) may be

$$\begin{aligned}
 Pr. \left\{ \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1 \right\} &= \int \dots \int h(\mathbf{p}', \mathbf{q}') f_{N_1 + N_2 - k - 1}(\chi^2) \prod_{i=1}^k (dp_i dq_i) \cdot d\chi^2, \\
 &\quad \begin{matrix} q < F^{-1}(\mathbf{u} - \mathbf{E}) \\ \{(\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a})\} / \chi^2 \mid [(N_1 + N_2 - k - 1)/k] \leq F_{k, N_1 + N_2 - k - 1}(\alpha) \end{matrix}
 \end{aligned}$$

and by using the transformation (1.8), the probability may be given by

$$\begin{aligned}
 Pr. \left\{ \frac{N_1 \bar{\mathbf{x}}^{(1)} + N_2 \bar{\mathbf{x}}^{(2)}}{N_1 + N_2} < \mathbf{u}, D_1 \right\} &= \int \dots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \mathbf{r}' \mathbf{r} \right] \prod_{i=1}^k dr_i \\
 &\quad \begin{matrix} L\mathbf{r} < F^{-1}(\mathbf{u} - \mathbf{E}) \\ \{L\mathbf{s} + \mathbf{a}\}' \mathbf{R}^{-1} (L\mathbf{s} + \mathbf{a}) / \chi^2 \mid [(N_1 + N_2 - k - 1)/k] \leq F_{k, N_1 + N_2 - k - 1}(\alpha) \end{matrix} \\
 &\quad \cdot \int \dots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \mathbf{s}' \mathbf{s} \right] \frac{(\chi^2/2)^{(N_1 + N_2 - k - 1)/2 - 1} \exp[-\chi^2/2]}{2\Gamma[(N_1 + N_2 - k - 1)/2]} \prod_{i=1}^k ds_i d\chi^2
 \end{aligned}$$

Using the same notation as that of the theorem 1.1, we may obtain

$$\begin{aligned}
(2.3) \quad Pr. \left\{ \frac{N_1 \mathbf{x}^{(1)} + N_2 \mathbf{x}^{(2)}}{N_1 + N_2} < \mathbf{u}_1, D_1 \right\} \\
= \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \\
\begin{array}{l} r_k < m_k / l_{kk} \\ r_{k-1} < (m_{k-1} - l_{k-1,k} r_k) / l_{k-1,k-1} \\ \dots \\ r_1 < (m_1 - k \sum_{j=2}^k l_{1j} r_j) / l_{11} \end{array} \\
\cdot \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \frac{(\chi^2/2)^{(N_1+N_2-k-1)/2-1} \exp[-\chi^2/2]}{2\Gamma\{(N_1+N_2-k-1)/2\}} \left(\prod_{i=1}^k ds_i \right) d\chi^2 \\
[\{(\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})'(\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})\}/\chi^2] \leq \{k/(N_1+N_2-k-1)\} F_{k, N_1+N_2-k-1}(\alpha)
\end{aligned}$$

The second term of the right-hand side of (2.2) may be similarly

$$\begin{aligned}
(2.4) \quad Pr. \{ \bar{\mathbf{x}}^{(1)} < \mathbf{u}, D_2 \} = \int \cdots \int h(\mathbf{p}', \mathbf{q}') f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k (dp_i dq_i) d\chi^2 \\
\begin{array}{l} q < F^{-1}(\mathbf{u} - \mathbf{E}) + G\mathbf{p} \\ [(\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1}(\mathbf{p} + \mathbf{a})]/\chi^2 > [N_1+N_2-k-1]/k > F_{k, N_1+N_2-k-1}(\alpha) \end{array} \\
= \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} (\mathbf{r}'\mathbf{r} + \mathbf{s}'\mathbf{s}) \right] \frac{(\chi^2/2)^{(N_1+N_2-k-1)/2-1} \exp[-\chi^2/2]}{2\Gamma\{(N_1+N_2-k-1)/2\}} \left(\prod_{i=1}^k dr_i ds_i \right) \cdot d\chi^2 \\
\begin{array}{l} L\mathbf{r} < F^{-1}(\mathbf{u} - \mathbf{u}^{(1)}) + G\mathbf{L}\mathbf{s} \\ [(\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})'(\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})]/\chi^2 > \{k/(N_1+N_2-k-1)\} F_{k, N_1+N_2-k-1}(\alpha) \end{array} \\
= \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k s_i^2 \right] \frac{(\chi^2/2)^{(N_1+N_2-k-1)/2-1} \exp[-\chi^2/2]}{2\Gamma\{(N_1+N_2-k-1)/2\}} \\
[\{(\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})'(\mathbf{s} + \mathbf{L}^{-1}\mathbf{a})\}/\chi^2] > \{k/(N_1+N_2-k-1)\} F_{k, N_1+N_2-k-1}(\alpha) \\
\cdot \int \cdots \int \frac{1}{(2\pi)^{k/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^k r_i^2 \right] \prod_{i=1}^k dr_i \prod_{i=1}^k ds_i \cdot d\chi^2 \\
\begin{array}{l} r_k < n_k / l_{kk} \\ r_{k-1} < (n_{k-1} - l_{k-1,k} r_k) / l_{k-1,k-1} \\ \dots \\ r_1 < (n_1 - k \sum_{j=2}^k l_{1j} r_j) / l_{11} \end{array}
\end{aligned}$$

In combination of (2.3) and (2.4), we shall reach (2.1) which was to be proved.

Theorem 2.2. *The mean vector $E\{\mathbf{x}\}$ and the mean square error matrix $M.S.E.\{\bar{\mathbf{x}}\}$ of the estimate \mathbf{x} are given by*

$$(2.5) \quad E\{\mathbf{x}\} = \mathbf{u}^{(1)} + \frac{N_2}{N_1 + N_2} (\mathbf{u}^{(2)} - \mathbf{u}^{(1)}) Pr.\{D_1\} - \sqrt{\frac{N_2}{N_1(N_1 + N_2)}} \mathbf{c} I_{D_2}\{\mathbf{p}\}$$

and

$$\begin{aligned}
(2.6) \quad M.S.E.\{\mathbf{x}\} = \frac{1}{N_1 + N_2} \Sigma + \left(\frac{N_2}{N_1 + N_2} \right)^2 (\mathbf{u}^{(2)} - \mathbf{u}^{(1)}) (\mathbf{u}^{(2)} - \mathbf{u}^{(1)})' Pr.\{D_1\} \\
+ \frac{N_2}{N_1} I_{D_2}\{\mathbf{F}\mathbf{p}\mathbf{p}'\mathbf{F}'\}
\end{aligned}$$

where we put

$$(2.7) \quad Pr.\{D_1\} = \int \cdots \int \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{p}' \mathbf{R}^{-1} \mathbf{p} f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k dp_i \cdot d\chi^2 \right. \\ \left. [\{ (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) \} / \chi^2] [(N_1 + N_2 - k - 1) / k] \leq F_{k, N_1 + N_2 - k - 1}(\alpha) \right] \\ I_{D_2}\{\mathbf{p}\} = \int \cdots \int \frac{\mathbf{p}}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{p}' \mathbf{R}^{-1} \mathbf{p} \right] f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k dp_i \cdot d\chi^2 \\ [\{ (\mathbf{p} + \mathbf{a})' \mathbf{R}^{-1} (\mathbf{p} + \mathbf{a}) \} / \chi^2] [(N_1 + N_2 - k - 1) / k] > F_{k, N_1 + N_2 - k - 1}(\alpha)$$

Proof. Making use of the notation in the enunciation and proof of Theorem 2.1, we may write

$$E\{\bar{\mathbf{x}}\} = \int \cdots \int_{D_1} \{ \boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \} h(\mathbf{p}', \mathbf{q}') f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2 \\ + \int \cdots \int_{D_2} (\boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} - \mathbf{G}\mathbf{F}\mathbf{p}) h(\mathbf{p}', \mathbf{q}') f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2 \\ = \boldsymbol{\mu}^{(1)} + \mathbf{F} I_{D_1+D_2}\{\mathbf{q}\} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) Pr.\{D_1\} - \mathbf{G}\mathbf{F} I_{D_2} \\ = \boldsymbol{\mu}^{(1)} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) Pr.\{D_1\} - \sqrt{\frac{N_2}{N_1(N_1 + N_2)}} \mathbf{c} I_{D_2}\{\mathbf{p}\},$$

because $I_{D_1+D_2}\{\mathbf{q}\} = E\{\mathbf{q}\} = 0$. Thus the mean vector (2.5) is obtained. The mean square error matrix $M.S.E.\{\bar{\mathbf{x}}\}$ may be obtained from following relation:

$$(2.9) \quad E\{\bar{\mathbf{x}}\bar{\mathbf{x}}'\} = \int \cdots \int_{D_1} \left\{ \boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \right\} (\boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} \\ + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}))' h(\mathbf{p}', \mathbf{q}') f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2 \\ + \int \cdots \int_{D_2} (\boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} - \mathbf{G}\mathbf{F}\mathbf{p}) (\boldsymbol{\mu}^{(1)} + \mathbf{F}\mathbf{q} \\ - \mathbf{G}\mathbf{F}\mathbf{p})' h(\mathbf{p}', \mathbf{q}') f_{N_1+N_2-k-1}(\chi^2) \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2 \\ = \int \cdots \int_{D_1} \{ \boldsymbol{\mu}^{(1)} \boldsymbol{\mu}^{(1)'} + \mathbf{F}\mathbf{q}\mathbf{q}'\mathbf{F}' + \frac{N_2}{N_1 + N_2} (\boldsymbol{\mu}^{(1)} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \boldsymbol{\mu}^{(1)'}) \\ + \left(\frac{N_2}{N_1 + N_2} \right)^2 (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(1)} \mathbf{q}' \mathbf{F}' + \mathbf{F}\mathbf{q} \boldsymbol{\mu}^{(1)'}) \\ + \frac{N_2}{N_1 + N_2} (\mathbf{F}\mathbf{q} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})' + (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \mathbf{q}' \mathbf{F}') h(\mathbf{p}', \mathbf{q}') f_{N_1+N_2-k-1}(\chi^2) \\ \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2$$

$$\begin{aligned}
& + \int \cdots \int_{D_2} \{ \mu^{(1)} \mu^{(1)'} + F q q' F' + G^2 F p p' F' + (\mu^{(1)} q' F' + F q \mu^{(1)'}) \\
& - G(\mu^{(1)} p' F' + F p \mu^{(1)'}) - G(F q p' F' + F p q' F') \} h(p', q') f_{N_1+N_2-k-1}(\chi^2) \\
& \cdot \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2
\end{aligned}$$

substituting the following relations

$$\begin{aligned}
I_{D_1+D_2} \{ q q' \} &= \int \cdots \int_{D_1+D_2} q q' \cdot h(p', q') f_{N_1+N_2-k-1}(\chi^2) \cdot \prod_{i=1}^k dp_i \cdot \prod_{i=1}^k dq_i \cdot d\chi^2 = E \{ q q' \} = R, \\
I_{D_1+D_2} \{ q \} &= I_{D_2} \{ q \} = 0, \\
I_{D_2} \{ q p' \} &= I_{D_2} \{ p q' \} = 0
\end{aligned}$$

for (2.9), we may obtain that

$$\begin{aligned}
E \{ \bar{x} \bar{x}' \} &= \mu^{(1)} \mu^{(1)'} + \frac{1}{N_1 + N_2} \Sigma + \frac{N_2}{N_1 + N_2} \{ \mu^{(1)} (\mu^{(2)} - \mu^{(1)})' + (\mu^{(2)} - \mu^{(1)}) \mu^{(1)'} \} Pr. \{ D_1 \} \\
&+ \left(\frac{N_2}{N_1 + N_2} \right)^2 (\mu^{(2)} - \mu^{(1)}) (\mu^{(2)} - \mu^{(1)})' Pr. \{ D_1 \} + \frac{N_2}{N_1} I_{D_2} \{ F p p' F' \} \\
&- \sqrt{\frac{N_2}{N_1}} I_{D_2} \{ \mu^{(2)} p' F' + F p \mu^{(1)'} \}.
\end{aligned}$$

Thus the mean square error matrix $M.S.E. \{ \bar{x} \}$ of the estimate \bar{x} may be obtained in the following manner.

$$\begin{aligned}
M.S.E. \{ \bar{x} \} &= E \{ (\bar{x} - \mu^{(1)}) (\bar{x} - \mu^{(1)})' \} \\
&= E \{ \bar{x} \bar{x}' \} - \mu^{(1)} E \{ \bar{x}' \} - E \{ \bar{x} \} \mu^{(1)'} + \mu^{(1)} \mu^{(1)'} \\
&= \frac{1}{N_1 + N_2} \Sigma + \left(\frac{N_2}{N_1 + N_2} \right)^2 (\mu^{(2)} \mu^{(1)}) (\mu^{(2)} - \mu^{(1)})' Pr. \{ D_1 \} \\
&+ \frac{N_2}{N_1} I_{D_2} \{ F p p' F' \}.
\end{aligned}$$

Corollary 2.1. *Specially under the same assumption as corollary 1.1 or corollary 1.2, the same results are obtained correspondently.*

Corollary 2.2. *Specially when $\mu^{(2)} = \mu^{(1)}$, the variance component of the $M.S.E. \{ \bar{x} \}$ in case of multivariate pooling of data becomes smaller than the variance of the corresponding estimate in case of univariate pooling of data. The difference is given by*

$$\frac{N_2 \sigma_{ii}}{(N_1 + N_2) N_1} [I_{D_2} \{ p_i^2 \} - I_{D_{2*}} \{ p_i^2 \}] < 0 \quad (i=1, 2, \dots, k)$$

where

$$\begin{aligned}
 I_{D_2}\{p_i^2\} &= \int \cdots \int_{D_2} p_i^2 \cdot \frac{1}{(2\pi)^{k/2}} |\mathbf{R}|^{1/2} \exp\left[-\frac{1}{2} \mathbf{p}' \mathbf{R}^{-1} \mathbf{p}\right] f_{N_1+N_2-k-1}(\chi^2) \cdot \prod_{i=1}^k dp_i \cdot d\chi^2 \\
 I_{D_2^*}\{p_i^2\} &= \int \cdots \int p_i^2 \cdot \frac{1}{(2\pi)^{1/2}} \exp\left[-\frac{1}{2} p_i^2\right] f_{N_1+N_2-k-1}(\chi^2) dp_i d\chi^2 \\
 &\quad \frac{(N_1+N_2-2)^{1/2} p_i + \left(\frac{N_1 N_2}{N_1+N_2}\right)^{1/2} \frac{\mu_i^{(2)} - \mu_i^{(1)}}{\sigma}}{(\chi^2)^{1/2}} \rightarrow t_{N_1+N_2-2}(\alpha)
 \end{aligned}$$

References

- ASANO, C. [1] : *Estimations after preliminary test of significance and their applications to biometrical researches*, Bull. Math. Stat., **9-1, 2** (1960), 1-23.
- ASANO, C. and SATO, S. [1] : *A bivariate analogue of pooling data*, Bull. Math. Stat. **11** (1962), (in preparation).
- BANCROFT, T. A. [1] : *On biases in estimation due to the use of preliminary tests of significance*, Ann. Math. Stat., **15** (1944), 190-204.
- BENNET, C. I. [1] : *Estimation of means on the basis of preliminary tests of significance*, Ann. Inst. Stat. Math., Tokyo, **4** (1952), 31-43.
- KITAGAWA, T. [1] : *Successive process of statistical inferences*, Mem. Facu. of Sci. Kyushu Univ., Ser. A, **5** (1950), 139-180.
- KITAGAWA, T. [2] : *Successive process of statistical inferences*, (in Japanese, *Suisoku katei Ron*), Iwauami, Tokyo, 1958.

SAGA UNIVERSITY