An asymptotic approximation for EPMC in linear discriminant analysis based on three-step monotone missing data

Kazumasa Kurihara[1,*] , Nobumichi Shutoh[1,2], Takashi Seo[3]

[1] *Department of Mathematical Information Science, Graduate School of Science, Tokyo University of Science 1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, JAPAN*

[2] *Research Fellow of the Japan Society for the Promotion of Science*

[3] *Department of Mathematical Information Science, Faculty of Science, Tokyo University of Science 1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, JAPAN*

ABSTRACT

In this paper, we consider the expected probabilities of misclassification (EPMC) in linear discriminant function based on 3-step monotone missing data and derive an asymptotic approximation for EPMC. On condition that the parameters are unknown, we derive the maximum likelihood estimators (MLEs) and the unbiased estimators based on 3-step monotone missing data. Finally, we perform the Monte Carlo simulation for evaluating our result.

1 Introduction

Linear discriminant analysis is well known as one of statistical procedures to assign p dimensional observation vector $\boldsymbol{x}$ which arises from one of some groups into one of them. In particular, we consider that for the case that $\boldsymbol{x}$ comes from one of two groups, i.e., $\Pi^{(1)} : N_p(\boldsymbol{\mu}^{(1)}, \Sigma)$ and $\Pi^{(2)} : N_p(\boldsymbol{\mu}^{(2)}, \Sigma)$. If the parameters are known, then linear discriminant function is constructed as

$$W = (\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})'\Sigma^{-1}\Big[\boldsymbol{x} - \frac{1}{2}(\boldsymbol{\mu}^{(1)} + \boldsymbol{\mu}^{(2)})\Big]. \tag{1}$$

If $W > 0$, $\boldsymbol{x}$ may be assigned to $\Pi^{(1)}$, otherwise it may be assigned to $\Pi^{(2)}$. Then the exact probabilities of misclassification in (1) are

$$\Pr[W \le 0|\boldsymbol{x} \in \Pi^{(1)}] = \Pr[W > 0|\boldsymbol{x} \in \Pi^{(2)}] = \Phi\Big(-\frac{1}{2}\Delta\Big),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution and $\Delta^2 \equiv (\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})'\Sigma^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$ is Mahalanobis squared distance between $\Pi^{(1)}$ and $\Pi^{(2)}$.

However, in general, both $\boldsymbol{\mu}^{(g)}$ and Σ are unknown. Therefore their own estimators are substituted for $\boldsymbol{\mu}^{(g)}$ and Σ in (1), respectively. If we can obtain the following data set:

*Corresponding author, *E-mail address* : j1410610@ed.kagu.tus.ac.jp

$$\begin{pmatrix} x_{11}^{(g)} & x_{12}^{(g)} & \cdots & x_{1p}^{(g)} \\ x_{21}^{(g)} & x_{22}^{(g)} & \cdots & x_{2p}^{(g)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N_1^{(g)}\ 1}^{(g)} & x_{N_1^{(g)}\ 2}^{(g)} & \cdots & x_{N_1^{(g)}\ p}^{(g)} \end{pmatrix}$$

from $\Pi^{(g)}$, in other words, $N_1^{(g)}$ observation vectors $\boldsymbol{x}_j^{(g)}$ $(j = 1, \ldots, N_1^{(g)},\ g = 1, 2)$ can be observed completely from $\Pi^{(g)}$, where $\boldsymbol{x}_j^{(g)} = (x_{j1}^{(g)}, x_{j2}^{(g)}, \ldots, x_{jp}^{(g)})'$ are distributed as $N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$,then linear discriminant function can be constructed as follows:

$$W_1 = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})' S^{-1} \Big[\boldsymbol{x} - \frac{1}{2}(\overline{\boldsymbol{x}}^{(1)} + \overline{\boldsymbol{x}}^{(2)}) \Big], \tag{2}$$

where $\overline{\boldsymbol{x}}^{(g)}$ is sample mean vector from $\Pi^{(g)}$ and S is pooled sample covariance matrix on the basis of $\boldsymbol{x}_j^{(g)}$.

Since it is difficult to have the exact distribution of W_1, we cannot obtain exact EPMC, i.e.,

$$e_1(2|1) = \Pr[W_1 \le 0 | \boldsymbol{x} \in \Pi^{(1)}], \tag{3}$$

$$e_1(1|2) = \Pr[W_1 > 0 | \boldsymbol{x} \in \Pi^{(2)}]. \tag{4}$$

Alternatively, the asymptotic distribution of W_1 is well known. Several authors considered the asymptotic approximations for EPMC using the asymptotic properties.

For instance, Okamoto [8] and Lachenbruch [7] derived the asymptotic approximations under an asymptotic framework:

$$N_1^{(1)} \to \infty,\ N_1^{(2)} \to \infty.$$

On the other hand, Fujikoshi and Seo [2] and Lachenbruch [7] derived the same under another asymptotic framework:

$$N_1^{(1)} \to \infty,\ N_1^{(2)} \to \infty,\ p \to \infty,\ n_1 - p \to \infty,$$
$$\frac{N_1^{(1)}}{N_1^{(2)}} \to \text{positive const. and } \Delta^2 = \mathrm{O}(1),$$

In this paper, we primarily deal with the asymptotic approximation for EPMC proposed by Lachenbruch [7]. Using the property of asymptotic normality of (2), he proposed an asymptotic approximation for EPMC of (3) as

$$e_1(2|1) \simeq \Phi(\mathrm{E}(U)\{\mathrm{E}(V)\}^{-\frac{1}{2}}), \tag{5}$$

where

$$U = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})' S^{-1} (\overline{\boldsymbol{x}}^{(1)} - \boldsymbol{\mu}^{(1)}) - \frac{1}{2} D^2,$$
$$D^2 = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})' S^{-1} (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)}),$$
$$V = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})' S^{-1} \Sigma S^{-1} (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)}).$$

In order to the asymptotic aproximation for EPMC, Lachenbruch [7] derived the following expectations:

$$\mathrm{E}(U) = -\frac{n_1}{2(n_1 - p - 1)} \Big\{ \Delta^2 + \frac{p(N_1^{(1)} - N_1^{(2)})}{N_1^{(1)} N_1^{(2)}} \Big\},\ n_1 - p - 1 > 0, \tag{6}$$

$$\mathrm{E}(V) = \frac{n_1^2 (n_1 - 1)}{(n_1 - p)(n_1 - p - 1)(n_1 - p - 3)} \Big\{ \Delta^2 + \frac{p(N_1^{(1)} + N_1^{(2)})}{N_1^{(1)} N_1^{(2)}} \Big\},\ n_1 - p - 3 > 0. \tag{7}$$

The same for (4) can be also expressed by interchanging $N_1^{(1)}$ with $N_1^{(2)}$ in (7).

In addition, several authors discussed a general missing data or monotone missing data. For a general missing pattern, many statistical methods have been developed by Srivastava [13], Srivastava and Carter [14] and Shutoh et al. [12]. When the missing pattern is monotone, Seo and Srivastava [9] consider the test of equality of means and simultaneous confidence intervals in one sample problem, and Koizumi and Seo [5,6] considered them in K samples problem for k-step monotone missing data. Anderson and Olkin [1] obtained the MLEs of mean vector and covariance matrix for 2-step monotone missing data in one sample problem, and Kanda and Fujikoshi [4] consider the MLE's and the asymptotic expansions of the distributions concerning k-step monotone missing data in one sample problem. Hao and Krishnamoorthy [3] stated tests for the mean vector and a covariance matrix for k-step monotone missing data in one sample problem. Shutoh, Hyodo and Seo [11] stated the MLEs for two sample problem based on 2-step monotone missing data in order to derive an asymptotic approximation for EPMC in linear discriminant function in this case.

In this paper, our purpose is to obtain the asymptotic approximation for EPMC similar to Lachenbruch [7] with 3-step monotone missing data:

$$\begin{pmatrix}
x_{11}^{(g)} & \cdots & x_{1\ p_1}^{(g)} & x_{1,\ p_1+1}^{(g)} & \cdots & x_{1,\ p_1+p_2}^{(g)} & x_{1,\ p_1+p_2+1}^{(g)} & \cdots & x_{1p}^{(g)} \\
x_{21}^{(g)} & \cdots & x_{2\ p_1}^{(g)} & x_{2,\ p_1+1}^{(g)} & \cdots & x_{2,\ p_1+p_2}^{(g)} & x_{2,\ p_1+p_2+1}^{(g)} & \cdots & x_{2p}^{(g)} \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
x_{N_1^{(g)},\ 1}^{(g)} & \cdots & x_{N_1^{(g)},\ p_1}^{(g)} & x_{N_1^{(g)},\ p_1+1}^{(g)} & \cdots & x_{N_1^{(g)},\ p_1+p_2}^{(g)} & x_{N_1^{(g)},\ p_1+p_2+1}^{(g)} & \cdots & x_{N_1^{(g)}p}^{(g)} \\
x_{N_1^{(g)}+1,\ 1}^{(g)} & \cdots & x_{N_1^{(g)}+1,\ p_1}^{(g)} & x_{N_1^{(g)}+1,\ p_1+1}^{(g)} & \cdots & x_{N_1^{(g)}+1,\ p_1+p_2}^{(g)} & * & \cdots & * \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
x_{N_{(12)}^{(g)},\ 1}^{(g)} & \cdots & x_{N_{(12)}^{(g)},\ p_1} & x_{N_{(12)}^{(g)},\ p_1+1} & \cdots & x_{N_{(12)}^{(g)},\ p_1+p_2} & * & \cdots & * \\
x_{N_{(12)}^{(g)}+1,\ 1} & \cdots & x_{N_{(12)}^{(g)}+1,\ p_1} & * & \cdots & * & * & \cdots & * \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
x_{N^{(g)},\ 1} & \cdots & x_{N^{(g)},\ p_1} & * & \cdots & * & * & \cdots & *
\end{pmatrix}, \tag{8}$$

where $N_{(12)}^{(g)} \equiv N_1^{(g)} + N_2^{(g)},\ N^{(g)} \equiv N_1^{(g)} + N_2^{(g)} + N_3^{(g)}$ and the statement “$*$“ denotes missing observation. $N_{(12)}^{(g)}$ samples lose common $p - p_1$ observations and $N_3^{(g)}$ samples lose the common $p - (p_1 + p_2)$ observations, respectively. In other words, we assume that

$$\boldsymbol{x}_j^{(g)} = \left(\boldsymbol{x}_{1j}^{(g)\prime}, \boldsymbol{x}_{2j}^{(g)\prime}, \boldsymbol{x}_{3j}^{(g)\prime}\right)' \sim N_p(\boldsymbol{\mu}^{(g)}, \Sigma) \quad (j = 1, \ldots, N_1^{(g)},\ g = 1, 2), \tag{9}$$

$$\left(\boldsymbol{x}_{1j}^{(g)\prime}, \boldsymbol{x}_{2j}^{(g)\prime}\right)' \sim N_{p_1+p_2}(\boldsymbol{\mu}_{[2]}^{(g)}, \Sigma_{(12)(12)}) \quad (j = N_1^{(g)} + 1, \ldots, N_{(12)}^{(g)},\ g = 1, 2), \tag{10}$$

$$\boldsymbol{x}_{1j}^{(g)} \sim N_{p_1}(\boldsymbol{\mu}_1^{(g)}, \Sigma_{11}) \quad (j = N_{(12)}^{(g)} + 1, \ldots, N^{(g)},\ g = 1, 2), \tag{11}$$

where

$$\boldsymbol{x}_{1j}^{(g)} = (x_{j1}^{(g)}, \ldots, x_{j,p_1}^{(g)})'\ (j = 1, \ldots, N^{(g)},\ g = 1, 2),$$
$$\boldsymbol{x}_{2j}^{(g)} = (x_{j,p_1+1}^{(g)}, \ldots, x_{j,p_1+p_2}^{(g)})'\ (j = 1, \ldots, N_{(12)}^{(g)},\ g = 1, 2),$$
$$\boldsymbol{x}_{3j}^{(g)} = (x_{j,p_1+p_2+1}, \ldots, x_{j,p}^{(g)})'\ (j = 1, \ldots, N_1^{(g)},\ g = 1, 2),$$
$$\boldsymbol{\mu}^{(g)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_2^{(g)} \\ \boldsymbol{\mu}_3^{(g)} \end{pmatrix}, \quad \boldsymbol{\mu}_{[2]}^{(g)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(g)} \\ \boldsymbol{\mu}_2^{(g)} \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix}, \quad \Sigma_{(12)(12)} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

This paper is organized as follows. Section 2 states the outline for derivation of the MLEs for two groups based on 3-step monotone missing data. Section 2 also presents some Lemmas for our purpose in this paper. In Section 3, we state the main results in this paper, i.e., the asymptotic approximation for EPMC based on 3-step monotone missing data and the estimators of the Mahalanobis squared distances. Section 4 conducts Monte Carlo simulations in order to evaluate our result. Section 5 concluded this paper. In Appendix, we present the proof as to several theorems.

2 MLEs based on three-step monotone missing data

In this section, we derive the MLEs for two groups based on 3-step monotone missing data defined in (9)–(11). First we consider the transformation of the observation vector $\boldsymbol{x}_j^{(g)}$ by multiplying

$$\begin{pmatrix} I_{p_1} & O & O \\ -\Sigma_{21}\Sigma_{11}^{-1} & I_{p_2} & O \\ -\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1} & & I_{p_3} \end{pmatrix}$$

on the left-hand for obtaining MLEs, where O is a matrix with 0's. The transformed observation vectors have the following distributions:

$$\begin{aligned} \boldsymbol{y}_{1j}^{(g)} &= \boldsymbol{x}_{1j}^{(g)} \sim N_{p_1}(\boldsymbol{\mu}_1^{(g)}, \Psi_{11}), \\ \boldsymbol{y}_{2j}^{(g)} &= \boldsymbol{x}_{2j}^{(g)} - \Psi_{21}\boldsymbol{x}_{1j}^{(g)} \sim N_{p_2}(\boldsymbol{\mu}_2^{(g)} - \Psi_{21}\boldsymbol{\mu}_1^{(g)}, \Psi_{22}), \\ \boldsymbol{y}_{3j}^{(g)} &= \boldsymbol{x}_{3j}^{(g)} - \Psi_{3(12)}\begin{pmatrix} \boldsymbol{x}_{1j}^{(g)} \\ \boldsymbol{x}_{2j}^{(g)} \end{pmatrix} \sim N_{p_3}(\boldsymbol{\mu}_3^{(g)} - \Psi_{3(12)}\boldsymbol{\mu}_{[2]}^{(g)}, \Psi_{33}), \end{aligned}$$

where

$$\begin{aligned} \Psi &= \begin{pmatrix} \Psi_{(12)(12)} & \Psi_{(12)3} \\ \Psi_{3(12)} & \Psi_{33} \end{pmatrix} \\ &= \begin{pmatrix} \Psi_{(12)(12)} & \Sigma_{(12)(12)}^{-1}\Sigma_{(12)3} \\ \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1} & \Sigma_{33\cdot 12} \end{pmatrix}, \\ \Psi_{(12)(12)} &= \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{11}^{-1}\Sigma_{12} \\ \Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22\cdot 1} \end{pmatrix}, \end{aligned}$$

$\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$, $\Sigma_{33\cdot 12} = \Sigma_{33} - \Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}$ and Σ_{11} has nonsingularity. Thus we consider the new parameters $\{\boldsymbol{\mu}_1^{(g)}, \boldsymbol{\mu}_2^{(g)}, \boldsymbol{\mu}_3^{(g)}, \Psi\}$ instead of $\{\boldsymbol{\mu}_1^{(g)}, \boldsymbol{\mu}_2^{(g)}, \boldsymbol{\mu}_3^{(g)}, \Sigma\}$.

It should be noted that $\boldsymbol{y}_{1j}^{(g)}$, $\boldsymbol{y}_{2j}^{(g)}$ and $\boldsymbol{y}_{3j}^{(g)}$ are mutually independent and the likelihood function to obtain the MLEs is constructed as

$$\begin{aligned} L(\boldsymbol{\mu}_1^{(g)}, \boldsymbol{\mu}_2^{(g)}, \boldsymbol{\mu}_3^{(g)}, \Psi) = \text{Const.} \times\, &|\Psi_{11}|^{-\frac{1}{2}(N^{(1)}+N^{(2)})}|\Psi_{22}|^{-\frac{1}{2}(N_{(12)}^{(1)}+N_{(12)}^{(2)})}|\Psi_{33}|^{-\frac{1}{2}(N_1^{(1)}+N_1^{(2)})} \\ &\times \operatorname{etr}\left(-\frac{1}{2}\sum_{g=1}^{2}\sum_{j=1}^{N^{(g)}} (\boldsymbol{x}_{1j}^{(g)} - \boldsymbol{\mu}_1^{(g)})'\Psi_{11}^{-1}(\boldsymbol{x}_{1j}^{(g)} - \boldsymbol{\mu}_1^{(g)}) \right) \end{aligned}$$

$$\times \operatorname{etr}\Bigg(-\frac{1}{2}\sum_{g=1}^{2}\sum_{j=1}^{N_{(12)}^{(g)}} \Big\{ (\boldsymbol{x}_{2j}^{(g)} - \Psi_{21}\boldsymbol{x}_{1j}^{(g)}) - (\boldsymbol{\mu}_2^{(g)} - \Psi_{21}\boldsymbol{\mu}_1^{(g)}) \Big\}' \Psi_{22}^{-1}$$

$$\times \Big\{ (\boldsymbol{x}_{2j}^{(g)} - \Psi_{21}\boldsymbol{x}_{1j}^{(g)}) - (\boldsymbol{\mu}_2^{(g)} - \Psi_{21}\boldsymbol{\mu}_1^{(g)}) \Big\} \Bigg)$$

$$\times \operatorname{etr}\Bigg(-\frac{1}{2}\sum_{g=1}^{2}\sum_{j=1}^{N_1{}^{(g)}} \Big\{ (\boldsymbol{x}_{3j}^{(g)} - \Psi_{3(12)}\begin{pmatrix}\boldsymbol{x}_{1j}^{(g)}\\ \boldsymbol{x}_{2j}^{(g)}\end{pmatrix}) - (\boldsymbol{\mu}_3^{(g)} - \Psi_{3(12)}\boldsymbol{\mu}_{[2]}^{(g)}) \Big\}' \Psi_{33}^{-1}$$

$$\times \Big\{ (\boldsymbol{x}_{3j}^{(g)} - \Psi_{3(12)}\begin{pmatrix}\boldsymbol{x}_{1j}^{(g)}\\ \boldsymbol{x}_{2j}^{(g)}\end{pmatrix}) - (\boldsymbol{\mu}_3^{(g)} - \Psi_{3(12)}\boldsymbol{\mu}_{[2]}^{(g)}) \Big\} \Bigg). \tag{12}$$

We define the sample mean vectors

$$\begin{aligned}
\overline{\boldsymbol{x}}_i^{(g,1)} &= \frac{1}{N_1^{(g)}}\sum_{j=1}^{N_1^{(g)}} \boldsymbol{x}_{ij}^{(g)} \quad (i = 1, 2, 3),\\
\overline{\boldsymbol{x}}_i^{(g,2)} &= \frac{1}{N_2^{(g)}}\sum_{j=N_1^{(g)}+1}^{N_1^{(g)}+N_2^{(g)}} \boldsymbol{x}_{ij}^{(g)} \quad (i = 1, 2),\\
\overline{\boldsymbol{x}}_1^{(g,3)} &= \frac{1}{N_3^{(g)}}\sum_{j=N_{(12)}^{(g)}+1}^{N_1^{(g)}+N_2^{(g)}+N_3^{(g)}} \boldsymbol{x}_{1j}^{(g)}
\end{aligned}$$

and

$$\begin{aligned}
\overline{\boldsymbol{x}}_1^{[g,3]} &= \frac{1}{N_1^{(g)}+N_2^{(g)}+N_3^{(g)}}\Big(N_1^{(g)}\overline{\boldsymbol{x}}_1^{(g,1)} + N_2^{(g)}\overline{\boldsymbol{x}}_1^{(g,2)} + N_3^{(g)}\overline{\boldsymbol{x}}_1^{(g,3)}\Big),\\
\overline{\boldsymbol{x}}_i^{[g,2]} &= \frac{1}{N_1^{(g)}+N_2^{(g)}}\Big(N_1^{(g)}\overline{\boldsymbol{x}}_i^{(g,1)} + N_2^{(g)}\overline{\boldsymbol{x}}_i^{(g,2)}\Big) \quad (i = 1, 2),\\
\overline{\boldsymbol{x}}_i^{[g,1]} &= \frac{1}{N_1^{(g)}}\Big(N_1^{(g)}\overline{\boldsymbol{x}}_i^{(g,1)}\Big) \quad (i = 1, 2, 3).
\end{aligned}$$

Then, by (12) and the correction of their own coefficients, we can express the following estimators

$$\widehat{\boldsymbol{\mu}}^{(g)} = \begin{pmatrix}\widehat{\boldsymbol{\mu}}_1^{(g)}\\ \widehat{\boldsymbol{\mu}}_2^{(g)}\\ \widehat{\boldsymbol{\mu}}_3^{(g)}\end{pmatrix} \equiv \begin{pmatrix} \overline{\boldsymbol{x}}_1^{[g,3]}\\ \overline{\boldsymbol{x}}_2^{[g,2]} - \widehat{\Psi}_{21}(\overline{\boldsymbol{x}}_1^{[g,2]} - \overline{\boldsymbol{x}}_1^{[g,3]})\\ \overline{\boldsymbol{x}}_3^{[g,1]} - \widehat{\Psi}_{3(12)}\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[g,1]} - \overline{\boldsymbol{x}}_1^{[g,3]}\\ (\overline{\boldsymbol{x}}_2^{[g,1]} - \overline{\boldsymbol{x}}_2^{[g,2]}) + \widehat{\Psi}_{21}(\overline{\boldsymbol{x}}_1^{[g,2]} - \overline{\boldsymbol{x}}_1^{[g,3]})\end{pmatrix}\end{pmatrix},$$

and

$$\widehat{\Psi}_{11} = \frac{1}{n}\Big(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)}\Big),\ \widehat{\Psi}_{22} = \frac{1}{n_{(12)}}\Big(\Gamma_{(12)(12)}^{(1)} + \Gamma^{(2)}\Big)_{22\cdot 1},\ \widehat{\Psi}_{33} = \frac{1}{n_1}\Gamma_{33\cdot 12}^{(1)}, \tag{13}$$

$$\widehat{\Psi}_{21} = \Big(\Gamma_{21}^{(1)} + \Gamma_{21}^{(2)}\Big)\Big(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)}\Big)^{-1},\ \widehat{\Psi}_{3(12)} = \Gamma_{3(12)}^{(1)}\Big(\Gamma_{(12)(12)}^{(1)}\Big)^{-1}, \tag{14}$$

$$n = N^{(1)} + N^{(2)} - 2,\ n_{(12)} = N_{(12)}^{(1)} + N_{(12)}^{(2)} - 2 \text{ and } n_1 = N_1^{(1)} + N_2^{(2)} - 2.$$

Therefore $\widehat{\Sigma}$ has the following forms:

$$\widehat{\Sigma} = \begin{pmatrix}\widehat{\Sigma}_{(12)(12)} & \widehat{\Sigma}_{(12)3}\\ \widehat{\Sigma}_{3(12)} & \widehat{\Sigma}_{33}\end{pmatrix},$$

where

$$\widehat{\Sigma}_{(12)(12)} = \begin{pmatrix} \widehat{\Psi}_{11} & \widehat{\Psi}_{11}\widehat{\Psi}_{12} \\ \widehat{\Psi}_{21}\widehat{\Psi}_{11} & \widehat{\Psi}_{11} + \widehat{\Psi}_{21}\widehat{\Psi}_{11}\widehat{\Psi}_{12} \end{pmatrix},$$

$$\widehat{\Sigma}_{3(12)} = \widehat{\Sigma}'_{(12)3} = \widehat{\Psi}_{3(12)}\widehat{\Sigma}_{(12)(12)}, \ \widehat{\Sigma}_{33} = \widehat{\Psi}_{33} + \widehat{\Psi}_{3(12)}\widehat{\Sigma}_{(12)(12)}\widehat{\Sigma}_{(12)3}.$$

In addition, $\Gamma^{(1)}$, $\Gamma^{(2)}$ and $\Gamma^{(3)}$ are defined as

$$\Gamma^{(1)} = \begin{pmatrix} \Gamma^{(1)}_{11} & \Gamma^{(1)}_{12} & \Gamma^{(1)}_{13} \\ \Gamma^{(1)}_{21} & \Gamma^{(1)}_{22} & \Gamma^{(1)}_{23} \\ \Gamma^{(1)}_{31} & \Gamma^{(1)}_{32} & \Gamma^{(1)}_{33} \end{pmatrix} = n_1 S^{(1)},$$

$$\Gamma^{(2)} = \begin{pmatrix} \Gamma^{(2)}_{11} & \Gamma^{(2)}_{12} \\ \Gamma^{(2)}_{21} & \Gamma^{(2)}_{22} \end{pmatrix} = n_2 S^{(2)} + \sum_{g=1}^{2} \frac{N_1^{(g)} N_2^{(g)}}{N_1^{(g)} + N_2^{(g)}} \begin{pmatrix} \overline{\boldsymbol{x}}_1^{(g,1)} - \overline{\boldsymbol{x}}_1^{(g,2)} \\ \overline{\boldsymbol{x}}_2^{(g,1)} - \overline{\boldsymbol{x}}_2^{(g,2)} \end{pmatrix} \begin{pmatrix} \overline{\boldsymbol{x}}_1^{(g,1)} - \overline{\boldsymbol{x}}_1^{(g,2)} \\ \overline{\boldsymbol{x}}_2^{(g,1)} - \overline{\boldsymbol{x}}_2^{(g,2)} \end{pmatrix}'$$

and

$$\Gamma^{(3)} = n_3 S^{(3)} + \sum_{g=1}^{2} \frac{(N_1^{(g)} + N_2^{(g)}) N_3^{(g)}}{N^{(g)}} \left(\overline{\boldsymbol{x}}_1^{(g,3)} - \frac{1}{N_1^{(g)} + N_2^{(g)}} (N_1^{(g)} \overline{\boldsymbol{x}}_1^{(g,1)} + N_2^{(g)} \overline{\boldsymbol{x}}_1^{(g,2)}) \right)$$
$$\times \left(\overline{\boldsymbol{x}}_1^{(g,3)} - \frac{1}{N_1^{(g)} + N_2^{(g)}} (N_1^{(g)} \overline{\boldsymbol{x}}_1^{(g,1)} + N_2^{(g)} \overline{\boldsymbol{x}}_1^{(g,2)}) \right)',$$

where

$$S^{(1)} = \frac{1}{n_1} \sum_{g=1}^{2} \sum_{j=1}^{N_1^{(g)}} \left(\boldsymbol{x}_j^{(g)} - \overline{\boldsymbol{x}}_j^{(g,1)} \right) \left(\boldsymbol{x}_j^{(g)} - \overline{\boldsymbol{x}}_j^{(g,1)} \right)',$$

$$S^{(2)} = \frac{1}{n_2} \sum_{g=1}^{2} \sum_{j=N_1^{(g)}+1}^{N_{(12)}^{(g)}} \begin{pmatrix} \boldsymbol{x}_{1j}^{(g)} - \overline{\boldsymbol{x}}_1^{(g,2)} \\ \boldsymbol{x}_{2j}^{(g)} - \overline{\boldsymbol{x}}_2^{(g,2)} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{1j}^{(g)} - \overline{\boldsymbol{x}}_1^{(g,2)} \\ \boldsymbol{x}_{2j}^{(g)} - \overline{\boldsymbol{x}}_2^{(g,2)} \end{pmatrix}',$$

$$S^{(3)} = \frac{1}{n_3} \sum_{g=1}^{2} \sum_{j=N_{(12)}^{(g)}+1}^{N^{(g)}} \left(\boldsymbol{x}_{1j}^{(g)} - \overline{\boldsymbol{x}}_1^{(g,3)} \right) \left(\boldsymbol{x}_{1j}^{(g)} - \overline{\boldsymbol{x}}_1^{(g,3)} \right)',$$

$$n_2 = N_2^{(1)} + N_2^{(2)} - 2,$$

$$n_3 = N_3^{(1)} + N_3^{(3)} - 2,$$

$$\Gamma^{(1)}_{(12)(12)} = \begin{pmatrix} \Gamma^{(1)}_{11} & \Gamma^{(1)}_{12} \\ \Gamma^{(1)}_{21} & \Gamma^{(1)}_{22} \end{pmatrix}, \ \Gamma^{(1)}_{3(12)} = \Gamma^{(1)\prime}_{(12)3} = \begin{pmatrix} \Gamma^{(1)}_{31} & \Gamma^{(1)}_{32} \end{pmatrix}.$$

Then $\Gamma^{(1)}_{ij}$ and $\Gamma^{(2)}_{ij}$ are $p_i \times p_j$ partitioned matrices of $\Gamma^{(1)}$ and $\Gamma^{(2)}$. At the end of this section, we have several Lemmas for the distributions and expectations of (13) and (14).

Lemma 1. $\Gamma^{(1)}_{11}$, $\Gamma^{(1)}_{11} + \Gamma^{(2)}_{11} + \Gamma^{(3)}$, $(\Gamma^{(1)}_{(12)(12)} + \Gamma^{(2)})_{22\cdot 1}$ *and* $\Gamma^{(3)}_{33\cdot 12}$ *have the following distributions receptively*:

$$\Gamma^{(1)}_{11} \sim W_{p_1}(n_1, \Sigma_{11}),$$

$$\Gamma^{(1)}_{11} + \Gamma^{(2)}_{11} + \Gamma^{(3)} = n\widehat{\Psi}_{11} \sim W_{p_1}(n, \Sigma_{11}),$$

$$(\Gamma^{(1)}_{(12)(12)} + \Gamma^{(2)})_{22\cdot 1} = n_{(12)}\widehat{\Psi}_{22} \sim W_{p_2}(n_{(12)} - p_1, \Sigma_{22\cdot 1}),$$

$$\Gamma^{(1)}_{33\cdot 12} = n_1\widehat{\Psi}_{33} \sim W_{p_3}(n_1 - p_1 - p_2, \Sigma_{33\cdot 12}),$$

where $W_d(m,\Omega)$ denotes Wishart distribution with the parameters m and Ω.

Lemma 2. *Suppose that G has $W_d(m,\Omega)$. Let C be a $d \times d$ constant matrix or a random matrix which is independent of G. Then the following expectations can be obtained*:

$$\begin{aligned}
\mathrm{E}(G^{-1}) &= \frac{1}{m-d-1}\Omega^{-1},\ m-d-1>0,\\
\mathrm{E}(G^{-1}CG^{-1}) &= \frac{m-d-2}{(m-d)(m-d-1)(m-d-3)}\Omega^{-1}C\Omega^{-1}\\
&\quad + \frac{1}{(m-d)(m-d-1)(m-d-3)}\Big\{\Omega^{-1}C'\Omega^{-1}+\mathrm{tr}[\mathrm{C}\Omega^{-1}]\Omega^{-1}\Big\},\\
&\quad m-d-3>0,
\end{aligned}$$

respectively.

Lemma 3. *The following conditional expectations given $\Gamma_{11}^{(1)}$ and $\Gamma_{11}^{(2)}$ can be obtained as follows*:

$$\begin{aligned}
\mathrm{E}\big(\widehat{\Psi}_{12}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)}\big) &= \Sigma_{11}^{-1}\Sigma_{12},\\
\mathrm{E}\big(\widehat{\Psi}_{12}C_{22}\widehat{\Psi}_{21}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)}\big) &= \mathrm{tr}[C_{22}\Sigma_{22\cdot1}](\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}+\Sigma_{11}^{-1}\Sigma_{12}C_{22}\Sigma_{21}\Sigma_{11}^{-1},
\end{aligned}$$

where C_{ij} is $p_i \times p_j$ constant matrix.

Lemma 4. *The following conditional expectations given $\Gamma_{(12)(12)}^{(1)}$ can be obtained as follows*:

$$\begin{aligned}
\mathrm{E}\big(\widehat{\Psi}_{3(12)}|\Gamma_{(12)(12)}^{(1)}\big) &= \Sigma_{3(12)}\Sigma_{(12)(12)},\\
\mathrm{E}\big(\widehat{\Psi}_{(12)3}C_{33}\widehat{\Psi}_{3(12)}|\Gamma_{(12)(12)}^{(1)}\big) &= \mathrm{tr}[C_{33}\Sigma_{33\cdot12}](\Gamma_{(12)(12)}^{(1)})^{-1}+\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}C_{33}\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1},
\end{aligned}$$

where C_{ij} is $p_i \times p_j$ constant matrix.

Lemma 5. *Suppose that G_1 has $W_d(m_1,\Omega)$, G_2 has $W_d(m_2,\Omega)$ and they are independent. Let C_1 and C_2 be constant matrices respectively. If $m-d-3>0$ and $m_1-d-1>0$, then*

$$\begin{aligned}
\mathrm{E}(\mathrm{tr}[C_1G_1^{-1}C_2(G_1+G_2)^{-1}]) &= c_1(m-d-2)\mathrm{tr}[C_1\Omega^{-1}C_2\Omega^{-1}]\\
&\quad + c_1\{\mathrm{tr}[C_1\Omega^{-1}C_2'\Omega^{-1}]+\mathrm{tr}[C_1\Omega^{-1}]\mathrm{tr}[C_2\Omega^{-1}],
\end{aligned}$$

where $m=m_1+m_2$, $c_1=1/\{(m-d)(m-d-3)(m_1-d-1)\}$.

As to proof, refer Shutoh, Hyodo and Seo [11].

3 Asymptotic approximation for EPMC and estimators of Mahalanobis distances

Now we consider the asymptotic approximation for EPMC in the case of 3-step monotone missing data. If $\boldsymbol{\mu}^{(g)}$ and Σ are unknown, then the linear discriminant function is constructed as

$$W_3 = (\widehat{\boldsymbol{\mu}}^{(1)}-\widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\Sigma}^{-1}\big[\boldsymbol{x}-\frac{1}{2}(\widehat{\boldsymbol{\mu}}^{(1)}+\widehat{\boldsymbol{\mu}}^{(2)})\big], \tag{15}$$

where $\widehat{\boldsymbol{\mu}}^{(g)}$ $(g=1,2)$ and $\widehat{\Sigma}$ are estimators obtained in Section 2. Under $\boldsymbol{x} \in \Pi^{(1)}$, the EPMC such that $\boldsymbol{x}$ is assigned to $\Pi^{(2)}$ is expressed as

$$e_3(2|1) = \Pr[W_3 \le 0|\boldsymbol{x} \in \Pi^{(1)}]. \tag{16}$$

Moreover, another EPMC is also expressed as

$$e_3(1|2) = \Pr[W_3 > 0|\boldsymbol{x} \in \Pi^{(2)}]. \tag{17}$$

We consider (16) and define

$$U_3 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\Sigma}^{-1}(\widehat{\boldsymbol{\mu}}^{(1)} - \boldsymbol{\mu}^{(1)}) - \frac{1}{2}D_3^2, \tag{18}$$

$$D_3^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\Sigma}^{-1}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}), \tag{19}$$

$$V_3 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\Sigma}^{-1}\Sigma\widehat{\Sigma}^{-1}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}). \tag{20}$$

then $Z_3 = V_3^{-\frac{1}{2}}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\Sigma}^{-1}(\boldsymbol{x} - \widehat{\boldsymbol{\mu}}^{(1)})$ is independent of (U_3,V_3) and is distributed as $N(0,1)$ given $\widehat{\boldsymbol{\mu}}^{(1)}$, $\widehat{\boldsymbol{\mu}}^{(2)}$, $\widehat{\Sigma}$. By using (18)–(20), (15) can be also expressed as

$$W_3 = V_3^{\frac{1}{2}}Z_3 - U_3. \tag{21}$$

Consequently, W_3 can be expressed in the same way as Lachenbruch [7] in case of 3-step monotone missing data. As to EPMC of (16), it can be expressed by using (21) as

$$e_3(2|1) = \Pr\big[W_3 \le 0|\boldsymbol{x} \in \Pi^{(1)}\big] = \Pr\Big[Z_3 \le U_3V_3^{-\frac{1}{2}}|\boldsymbol{x} \in \Pi^{(1)}\Big]. \tag{22}$$

Since Z_3 is distributed as $N(0,1)$, (22) can be expressed as

$$e_3(2|1) = \mathrm{E}\Big[\Phi(U_3V_3^{-\frac{1}{2}})\Big]. \tag{23}$$

We consider the asymptotic approximation for EPMC stated in (16). W_3 is not distributed as normal distribution exactly, but is closely normal asymptotically. thus we propose

$$e_3(2|1) \simeq \Phi\Big[\mathrm{E}(U_3)\{\mathrm{E}(V_3)\}^{-\frac{1}{2}}\Big]. \tag{24}$$

It follows from

$$\widehat{\Sigma}^{-1} = \begin{pmatrix} \widehat{\Psi}_{11}^{-1} + \widehat{\Psi}_{12}\widehat{\Psi}_{22}^{-1}\widehat{\Psi}_{21} + \widehat{\Psi}_{13}\widehat{\Psi}_{33}^{-1}\widehat{\Psi}_{31} & -\widehat{\Psi}_{12}\widehat{\Psi}_{22}^{-1} + \widehat{\Psi}_{13}\widehat{\Psi}_{33}^{-1}\widehat{\Psi}_{32} & -\widehat{\Psi}_{13}\widehat{\Psi}_{33}^{-1} \\ -\widehat{\Psi}_{22}^{-1}\widehat{\Psi}_{21} + \widehat{\Psi}_{23}\widehat{\Psi}_{33}^{-1}\widehat{\Psi}_{31} & \widehat{\Psi}_{22}^{-1} + \widehat{\Psi}_{23}\widehat{\Psi}_{33}^{-1}\widehat{\Psi}_{32} & -\widehat{\Psi}_{23}\widehat{\Psi}_{33}^{-1} \\ -\widehat{\Psi}_{33}^{-1}\widehat{\Psi}_{31} & -\widehat{\Psi}_{33}^{-1}\widehat{\Psi}_{32} & \widehat{\Psi}_{33}^{-1} \end{pmatrix},$$

and Lemmas 1–4 that Theorem 6 holds.

Theorem 6. *If $n_1 - p - 3 > 0$, then the exact expectation of U_3 and the asymptotic expectation of V_3 for large $N_{(12)}^{(g)}$ $(g=1,2)$ can be obtained as*

$$\mathrm{E}(U_3) = \mathrm{E}(U_2) - u_1\left\{\delta_{11}^2 + \frac{p_1(N^{(1)} - N^{(2)})}{N^{(1)}N^{(2)}}\right\} + u_2\left\{\delta_{11}^2 + \frac{p_1(N_{(12)}^{(1)} - N_{(12)}^{(2)})}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\right\},$$

$$\mathrm{E}(V_3) \simeq \mathrm{E}(V_2) + v_1 \left\{ \delta_{11}^2 + \frac{p_1(N^{(1)} + N^{(2)})}{N^{(1)} N^{(2)}} \right\}$$
$$- v_2 \left\{ \delta_{11}^2 + \frac{p_1(N_{(12)}^{(1)} + N_{(12)}^{(2)})}{N_{(12)}^{(1)} N_{(12)}^{(12)}} \right\},$$

where

$$u_1 = \frac{n}{2(n - p_1 - 1)}, \; u_2 = \frac{n_{(12)}}{2(n_{(12)} - p_1 - 1)},$$
$$v_1 = \frac{n^2(n-1)}{(n-p_1)(n-p_1-1)(n-p_1-3)}$$
$$+ \frac{2nn_{(12)}p_2(n-1)}{(n-p_1)(n-p_1-3)(n_{(12)}-p_1-1)(n_{(12)}-p_1-p_2-1)}$$
$$+ \frac{2nn_1p_3(n-1)}{(n-p_1)(n-p_1-3)(n_1-p-1)(n_1-p_1-p_2-1)},$$
$$v_2 = \frac{n_{(12)}^2(n_{(12)}-1)}{(n_{(12)}-p_1)(n_{(12)}-p_1-1)(n_{(12)}-p_1-3)}$$
$$+ \frac{2n_{(12)}^2p_2(n_{(12)}-1)}{(n_{(12)}-p_1)(n_{(12)}-p_1-3)(n_{(12)}-p_1-1)(n_{(12)}-p_1-p_2-1)}$$
$$+ \frac{2n_{(12)}n_1p_3(n_{(12)}-1)}{(n_{(12)}-p_1)(n_{(12)}-p_1-3)(n_1-p-1)(n_1-p_1-p_2-1)}$$

and δ_{11}^2 *denotes* $(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})'\Sigma_{11}^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)})$*. In addition,* U_2 *is* U_m *stated in Shutoh, Hyodo and Seo [11].*

For details of proof, refer A.1. Therefore the asymptotic approximation stated in (16) can be obtained by substituting the result of Theorem 5 into (24). Moreover, the asymptotic approximation stated in (17) can be also obtained by interchanging $N_i^{(1)}$ and $N_i^{(2)}$ for $i = 1, 2, 3$.

Consequently, we have the asymptotic approximation for EPMC under the respective frameworks

$$N_1^{(g)} \to \infty, \; N_2^{(g)} \to \infty, \; N_3^{(g)} \to \infty.$$

and

$$N_1^{(g)} \to \infty, \; N_2^{(g)} \to \infty, \; N_3^{(g)} \to \infty, \; p_1 \to \infty, \; p_1 + p_2 \to \infty, \; p \to \infty,$$
$$\frac{N_1^{(1)}}{N_1^{(2)}} \to \text{positive const.}, \; \frac{N_{(12)}^{(1)}}{N_{(12)}^{(2)}} \to \text{positive const.}, \; \frac{N^{(1)}}{N^{(2)}} \to \text{positive const.},$$
$$n - p_1 \to \infty, \; n_{(12)} - p_1 - p_2 \to \infty, \; n_1 - p \to \infty, \; \Delta^2 = \mathrm{O}(1).$$

Besides, to check on the results in Theorem 6, we can obtain the following corollary.

Corollary 7. *If we put* $N_3^{(g)} = 0$ $(g = 1, 2)$*, then the results given by Theorem 5 coincides with the result derived by Shutoh, Hyodo and Seo [11] respectively.*

Next, we can obtain the asymptotic approximation for EPMC in the previous section, but $\boldsymbol{\mu}^{(g)}$ and Σ are unknown. Then we derive the unbiased estimators of the Mahalanobis distances Δ^2, δ_{11}^2 and δ_{12}^2, where $\delta_{11}^2 = (\boldsymbol{\mu}_1^{(1)} - \boldsymbol{\mu}_1^{(2)})'\Sigma_{11}^{-1}(\boldsymbol{\mu}_1^{(1)} - \boldsymbol{\mu}_1^{(2)})$ and $\delta_{12}^2 = (\boldsymbol{\mu}_{[2]}^{(1)} - \boldsymbol{\mu}_{[2]}^{(2)})'\Sigma_{(12)(12)}^{-1}(\boldsymbol{\mu}_{[2]}^{(1)} - \boldsymbol{\mu}_{[2]}^{(2)})$.

Theorem 8. *The unbiased estimators for* δ_{11}^2*,* δ_{12}^2 *and* Δ^2 *can be obtained as*

$$\widehat{\delta}_{11}^2 = \frac{1}{c_1}d_{11}^2 - \frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}}, \tag{25}$$

$$\widehat{\delta}_{12}^2 = \frac{1}{c_2}d_{12}^2 - \frac{1}{c_2}(c_1-c_3)\widehat{\delta}_{11}^2 - \frac{c_1}{c_2}\frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}} - \frac{p_2(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{N_{(12)}^{(1)}N_{(12)}^{(2)}} - c_3\frac{p_1p_2(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{n_{(12)}N_{(12)}^{(1)}N_{(12)}^{(2)}}, \tag{26}$$

$$\widehat{\Delta}^2 = \frac{1}{c_4}D^2 - \frac{1}{c_4}(c_1-c_3)\widehat{\delta}_{11}^2 - \frac{1}{c_4}(c_2-c_5)\widehat{\delta}_{12}^2 - \frac{c_1}{c_4}\frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}} - \frac{c_2}{c_4}\frac{p_2(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{N_{(12)}^{(1)}N_{(12)}^{(2)}} - \frac{c_2c_3}{c_4}\frac{p_1p_2(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{n_{(12)}N_{(12)}^{(1)}N_{(12)}^{(2)}} - \frac{p_3(N_1^{(1)}+N_1^{(2)})}{N_1^{(1)}N_1^{(2)}} - c_5\frac{p_3(p_1+p_2)(N_1^{(1)}+N_1^{(2)})}{n_1N_1^{(1)}N_1^{(2)}}, \tag{27}$$

respectively, where $c_1 = n/(n-p_1-1)$, $c_2 = n_{(12)}/(n_{(12)}-p_1-p_2-1)$, $c_3 = n_{(12)}/(n_{(12)}-p_1-1)$, $c_4 = n_1/(n_1-p-1)$, $c_5 = n_1/(n_1-p_1-p_2-1)$,

$$d_{11}^2 = (\widehat{\boldsymbol{\mu}}_1^{(1)} - \widehat{\boldsymbol{\mu}}_1^{(2)})'\widehat{\Sigma}_{11}^{-1}(\widehat{\boldsymbol{\mu}}_1^{(1)} - \widehat{\boldsymbol{\mu}}_1^{(2)}),$$

$$d_{12}^2 = \begin{pmatrix}\widehat{\boldsymbol{\mu}}_1^{(1)} - \widehat{\boldsymbol{\mu}}_1^{(2)} \\ \widehat{\boldsymbol{\mu}}_2^{(1)} - \widehat{\boldsymbol{\mu}}_2^{(2)}\end{pmatrix}' \widehat{\Sigma}_{(12)(12)}^{-1} \begin{pmatrix}\widehat{\boldsymbol{\mu}}_1^{(1)} - \widehat{\boldsymbol{\mu}}_1^{(2)} \\ \widehat{\boldsymbol{\mu}}_2^{(1)} - \widehat{\boldsymbol{\mu}}_2^{(2)}\end{pmatrix},$$

$$D^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})'\widehat{\Sigma}^{-1}(\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}).$$

As to a proof, refer to A.2.

4 Simulation studies

In this section, we compare the accuracy of the result which is proposed in (24) as $\widehat{e}_K(2|1)$ with other approximations in the case of complete data, i.e., the result of Lachenbruch [7] denoted by $\widehat{e}_L(2|1)$. As Monte Carlo simulation for $\widehat{e}_K(2|1)$ and $\widehat{e}_L(2|1)$, we carry out 1,000,000 replications. Moreover, we can acquire the exact probability of misclassification by Monte Carlo simulation. Therefore, we evaluate the difference between the asymptotic approximation for EPMC and the exact probability of misclassification.

First, in the case of $\Delta = 1.05$, we select the parameters as

$$(p_1,p_2,p_3) = (2,1,1),\ (2,2,1),\ (3,2,1),\ (4,3,2),\ (5,4,3),$$
$$(M_1,M_2,M_3) = (10,10,10),\ (20,20,20),\ (50,50,50),\ (100,100,100),$$

where $M_i = N_i^{(1)} = N_i^{(2)}$ $(i=1,2,3)$.
Then, the simulation results are presented in Table 1. As for $e_1(2|1)$ and $\widehat{e}_L(2|1)$, we put $M_2 = M_3 = 0$. Moreover, the higher the value of Δ is, the superior it is that both the approximation value and the difference between the asymptotic approximation for EPMC and the exact probability of misclassification. In this section, we set that $\Delta = 1.05$.

Next, we performed Monte Carlo simulation under unequal sample sizes. We assume that the sample size from $\Pi^{(1)}$ is two times the same from $\Pi^{(2)}$, i.e, $N_i^{(1)} = 2N_i^{(2)}$ $(i=1,2,3)$. Then, we select the parameters as

$$(p_1,p_2,p_3) = (2,1,1),\ (2,2,1),\ (3,2,1),\ (4,3,2),$$
$$(L_1,L_2) = (10,5),\ (20,10),\ (50,25),\ (100,50),$$

and $\Delta = 1.05$,where $L_g = N_1^{(g)} = N_2^{(g)} = N_3^{(g)}$ $(g = 1, 2)$. Then, the simulation results are presented in Table 2. As for $e_1(2|1)$ and $\widehat{e}_L(2|1)$, we put $L_i = N_1^{(i)}$ and $N_2^{(g)} = N_3^{(g)} = 0$ $(g = 1, 2)$.

Finally, we consider the value of Δ, δ_{12} and δ_{11} so that we perform the simulation peculiar to 3-step monotone missing data. Then, we select the parameters as

$$(\Delta, \delta_{12}, \delta_{11}) =(1.05, 0.70, 0.35),\ (1.05, 0.60, 0.15),$$
$$(M_1, M_2, M_3) =(10, 10, 10),\ (20, 20, 20),\ (50, 50, 50),\ (100, 100, 100).$$

As for M_i $(i = 1, 2, 3)$, they are similar to first simulation case. This simulation results are presented in Table 3 and 4.

5 Conclusion

In this paper, we derived the MLEs of 3-step monotone missing data in Section 2. By using the MLEs, we considered the linear discriminant function based on 3-step monotone missing data. And then, by using existing results and several Lemmas, we proposed the approximation for EPMC as to Lachenbruch's [7] type. In addition, we proposed not only the asymptotic approximation for the EPMC but also the unbiased estimators. Finally, we compared our result with Lachenbruch's [7] approximation in Section 5. These results implied that the method we proposed in this paper could be useful. But, when the difference between the number of the dimension and the number of the sample is small, it could be obtained that the Lachenbruch's [7] approximation is more accurate than the proposed result.

A.1 Proof of Theorem 6

First, we consider the expectations of sample mean vectors. If we put $\boldsymbol{\delta}_\ell = \boldsymbol{\mu}_\ell^{(1)} - \boldsymbol{\mu}_\ell^{(2)} (\ell = 1, 2, 3)$ and C_{ij} is $p_i \times p_j$ matrix and independent of sample mean vectors, then we can obtain the expectations of sample mean vectors as

$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]})' C_{11} (\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]})\Big] = \boldsymbol{\delta}_1' C_{11} \boldsymbol{\delta}_1 + \frac{N^{(1)} + N^{(2)}}{N^{(1)} N^{(2)}} \mathrm{tr}(\Sigma_{11} C_{11}'),$$
$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]})' C_{11} (\overline{\boldsymbol{x}}_1^{[1,3]} - \boldsymbol{\mu}_1^{(1)})\Big] = \frac{1}{N^{(1)}} \mathrm{tr}(\Sigma_{11} C_{11}'),$$
$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]})' C_{11} (\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]})\Big] = \boldsymbol{\delta}_1' C_{11} \boldsymbol{\delta}_1 + \frac{N_{(12)}^{(1)} + N_{(12)}^{(1)}}{N_{(12)}^{(1)} N_{(12)}^{(2)}} \mathrm{tr}(\Sigma_{11} C_{11}'),$$
$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]})' C_{11} (\overline{\boldsymbol{x}}_1^{[1,2]} - \boldsymbol{\mu}_1^{(1)})\Big] = \frac{1}{N_{(12)}^{(1)}} \mathrm{tr}(\Sigma_{11} C_{11}'),$$
$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]})' C_{12} (\overline{\boldsymbol{x}}_2^{[1,2]} - \overline{\boldsymbol{x}}_2^{[2,2]})\Big] = \boldsymbol{\delta}_1' C_{12} \boldsymbol{\delta}_2 + \frac{N_{(12)}^{(1)} + N_{(12)}^{(1)}}{N_{(12)}^{(1)} N_{(12)}^{(2)}} \mathrm{tr}(\Sigma_{12} C_{12}'),$$
$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]})' C_{12} (\overline{\boldsymbol{x}}_2^{[1,2]} - \boldsymbol{\mu}_2^{(1)})\Big] = \frac{1}{N_{(12)}^{(1)}} \mathrm{tr}(\Sigma_{12} C_{12}'),$$
$$\mathrm{E}\left[\begin{pmatrix} \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_2^{[1,1]} - \overline{\boldsymbol{x}}_2^{[2,1]} \end{pmatrix}' C_{(12)(12)} \begin{pmatrix} \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_2^{[1,1]} - \overline{\boldsymbol{x}}_2^{[2,1]} \end{pmatrix}\right] = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \end{pmatrix}' C_{(12)(12)} \begin{pmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \end{pmatrix} + \frac{N_1^{(1)} + N_1^{(2)}}{N_1^{(1)} N_1^{(2)}} \mathrm{tr}(\Sigma_{(12)(12)} C_{(12)(12)}'),$$
$$\mathrm{E}\left[\begin{pmatrix} \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_2^{[1,1]} - \overline{\boldsymbol{x}}_2^{[2,1]} \end{pmatrix}' C_{(12)(12)} \begin{pmatrix} \overline{\boldsymbol{x}}_1^{[1,1]} - \boldsymbol{\mu}_1^{(1)} \\ \overline{\boldsymbol{x}}_2^{[1,1]} - \boldsymbol{\mu}_2^{(1)} \end{pmatrix}\right] = \frac{N_1^{(1)} + N_1^{(2)}}{N_1^{(1)} N_1^{(2)}} \mathrm{tr}(\Sigma_{(12)(12)} C_{(12)(12)}'),$$

$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]})'C_{33}(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]})\Big]=\boldsymbol{\delta}_3'C_{33}\boldsymbol{\delta}_3+\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\mathrm{tr}(\Sigma_{33}C_{33}'),$$

$$\mathrm{E}\Big[(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]})'C_{33}(\overline{\boldsymbol{x}}_3^{[1,1]}-\boldsymbol{\mu}_3^{(1)})\Big]=\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\mathrm{tr}(\Sigma_{33}C_{33}'),$$

$$\mathrm{E}\left[\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]}-\overline{\boldsymbol{x}}_1^{[2,1]}\\ \overline{\boldsymbol{x}}_2^{[1,1]}-\overline{\boldsymbol{x}}_2^{[2,1]}\end{pmatrix}'C_{(12)3}(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]})\right]=\begin{pmatrix}\boldsymbol{\delta}_1\\ \boldsymbol{\delta}_2\end{pmatrix}'C_{(12)3}\boldsymbol{\delta}_3^{(1)}+\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\mathrm{tr}(\Sigma_{(12)3}C_{(12)3}'),$$

$$\mathrm{E}\left[\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]}-\overline{\boldsymbol{x}}_1^{[2,1]}\\ \overline{\boldsymbol{x}}_2^{[1,1]}-\overline{\boldsymbol{x}}_2^{[2,1]}\end{pmatrix}'C_{(12)3}(\overline{\boldsymbol{x}}_3^{[1,1]}-\boldsymbol{\mu}_3^{(1)})\right]=\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\mathrm{tr}(\Sigma_{(12)3}C_{(12)3}').$$

So we have the expectation of U_3 including U_2:

$$\begin{aligned}U_3=\;&U_2\\&-n_{(12)}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\boldsymbol{\mu}_1^{(1)})\\&+\frac{n_{(12)}}{2}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})\\&+n(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\boldsymbol{\mu}_1^{(1)})\\&-\frac{n}{2}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}).\end{aligned}$$

The expectation of U_2 has already derived by Shutoh, Hyodo and Seo [11]. By using Lemma 2, after deriving the expectations of sample mean vectors, we can have the expectations of the random matrices. Therefore we have the exact expectation of U_3 since $\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}$ is independent of $(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]},\overline{\boldsymbol{x}}_1^{[1,2]}-\boldsymbol{\mu}_1^{(1)})$ and $\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)}$ is independent of $(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]},\overline{\boldsymbol{x}}_1^{[1,3]}-\boldsymbol{\mu}_1^{(1)})$, respectively.

Next we consider the expectation of V_3. In a similar manner to U_3, we express V_3 which depends on V_2 as

$$\begin{aligned}V_3=\;&V_2\\&+n^2(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}) &(28)\\&-n_{(12)}^2(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]}) &(29)\\&+2nn_{(12)}(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]})'(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}) &(30)\\&-2nn_{12}(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]})'(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}(\Gamma_{21}^{(1)}+\Gamma_{21}^{(2)})(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\\&\qquad\times\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}) &(31)\\&-2nn_{(12)}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\\&\qquad\times\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}) &(32)\\&+2nn_{(12)}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\\&\qquad\times(\Gamma_{21}^{(1)}+\Gamma_{21}^{(2)})(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}) &(33)\\&-2n_{(12)}^2(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]})'(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]}) &(34)\\&+2n_{(12)}^2(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]})'(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}(\Gamma_{21}^{(1)}+\Gamma_{21}^{(2)})(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\\&\qquad\times\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]}) &(35)\\&+2n_{(12)}^2(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\\&\qquad\times\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]}) &(36)\\&-2n_{(12)}^2(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\\&\qquad\times(\Gamma_{21}^{(1)}+\Gamma_{21}^{(2)})(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]}) &(37)\end{aligned}$$

$$+ 2nn_1(\overline{\boldsymbol{x}}_3^{[1,1]} - \overline{\boldsymbol{x}}_3^{[2,1]})'(\Gamma_{33\cdot 12}^{(1)})^{-1}\Sigma_{31}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]}) \tag{38}$$

$$- 2nn_1(\overline{\boldsymbol{x}}_3^{[1,1]} - \overline{\boldsymbol{x}}_3^{[2,1]})'(\Gamma_{33\cdot 12}^{(1)})^{-1}\Gamma_{3(12)}^{(1)}(\Gamma_{(12)(12)}^{(1)})^{-1} \times \Sigma_{(12)1}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]}) \tag{39}$$

$$- 2nn_1\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]}\end{pmatrix}'(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}^{(1)}(\Gamma_{33\cdot 12}^{(1)})^{-1} \times \Sigma_{31}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]}) \tag{40}$$

$$+ 2nn_1\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]}\end{pmatrix}'(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}^{(1)}(\Gamma_{33\cdot 12}^{(1)})^{-1}\Gamma_{3(12)}^{(1)}(\Gamma_{(12)(12)}^{(1)})^{-1} \times \Sigma_{(12)1}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,3]} - \overline{\boldsymbol{x}}_1^{[2,3]}) \tag{41}$$

$$- 2n_{(12)}n_1(\overline{\boldsymbol{x}}_3^{[1,1]} - \overline{\boldsymbol{x}}_3^{[2,1]})'(\Gamma_{33\cdot 12}^{(1)})^{-1}\Sigma_{31}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]}) \tag{42}$$

$$+ 2n_{(12)}n_1(\overline{\boldsymbol{x}}_3^{[1,1]} - \overline{\boldsymbol{x}}_3^{[2,1]})'(\Gamma_{33\cdot 12}^{(1)})^{-1}\Gamma_{3(12)}^{(1)}(\Gamma_{(12)(12)}^{(1)})^{-1} \times \Sigma_{(12)1}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]}) \tag{43}$$

$$+ 2n_{(12)}n_1\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]}\end{pmatrix}'(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}^{(1)}(\Gamma_{33\cdot 12}^{(1)})^{-1} \times \Sigma_{31}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]}) \tag{44}$$

$$- 2n_{(12)}n_1\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]} \\ \overline{\boldsymbol{x}}_1^{[1,1]} - \overline{\boldsymbol{x}}_1^{[2,1]}\end{pmatrix}'(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}^{(1)}(\Gamma_{33\cdot 12}^{(1)})^{-1}\Gamma_{3(12)}^{(1)}(\Gamma_{(12)(12)}^{(1)})^{-1} \times \Sigma_{(12)1}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]} - \overline{\boldsymbol{x}}_1^{[2,2]}). \tag{45}$$

By using Lemmas 1–5, we have the following results:

$$\mathrm{E}\Big[(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}\Big] = \frac{n-1}{(n-p_1)(n-p_1-1)(n-p_1-3)}\Sigma_{11}^{-1},$$

$$\mathrm{E}\Big[(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} +)^{-1}\Big] = \frac{n_{(12)}-1}{(n_{(12)}-p_1)(n_{(12)}-p_1-1)(n_{(12)}-p_1-3)}\Sigma_{11}^{-1},$$

$$\mathrm{E}\Big[\Sigma_{11}^{-1}\Sigma_{12}(\Gamma_{(12)(12)}^{(1)} + \Gamma^{(2)})_{22\cdot 1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}\Big] = \frac{1}{n_{(12)}-p_1-p_2-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot 1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1},$$

$$\mathrm{E}\Big[\Sigma_{11}^{-1}\Sigma_{12}(\Gamma_{(12)(12)}^{(1)} + \Gamma^{(2)})_{22\cdot 1}^{-1}(\Gamma_{21}^{1} + \Gamma_{21}^{(2)})(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}|\Gamma_{11}^{(1)}, \Gamma_{11}^{(2)}, \Gamma^{(3)}\Big] = \frac{1}{n_{(12)}-p_1-p_2-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot 1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1},$$

$$\mathrm{E}\Big[(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)} + \Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)} + \Gamma^{(2)})_{22\cdot 1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1}|\Gamma_{11}^{(1)}, \Gamma_{11}^{(2)}, \Gamma^{(3)}\Big] = \frac{1}{n_{(12)}-p_1-p_2-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot 1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)} + \Gamma^{(3)})^{-1},$$

$$\mathrm{E}\Big[(\Gamma_{(12)(12)}^{(1)} + \Gamma^{(2)})_{22\cdot 1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}\Big] = \frac{1}{(n_{(12)}-p_1-1)(n_{(12)}-p_1-p_2-1)}\Sigma_{22\cdot 1}^{-1}\Sigma_{21}\Sigma_{11}^{-1},$$

$$\mathrm{E}\Big[(\Gamma_{(12)(12)}^{(1)} + \Gamma^{(2)})_{22\cdot 1}^{-1}(\Gamma_{21}^{1} + \Gamma_{21}^{(2)})(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}\Sigma_{11}^{-1}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1}|\Gamma_{11}^{(1)}, \Gamma_{11}^{(2)}\Big] = \frac{1}{n_{(12)}-p_1-p_2-1}\Sigma_{22\cdot 1}\Sigma_{21}(\Gamma_{11}^{(1)} + \Gamma_{11}^{(2)})^{-1},$$

$$\mathrm{E}\Big[(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)}\Big]$$
$$=\frac{1}{n_{(12)}-p_1-p_2-1}\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot1}\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1},$$
$$\mathrm{E}\Big[(\Gamma_{33\cdot12}^{(1)})^{-1}\Sigma_{31}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}\Big]=\frac{1}{(n-p_1-1)(n_1-p-1)}\Sigma_{33\cdot12}^{-1}\Sigma_{31}\Sigma_{11}^{-1},$$
$$\mathrm{E}\Big[(\Gamma_{33\cdot12}^{(1)})^{-1}\Gamma_{3(12)}^{(1)}(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)1}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)},\Gamma^{(3)}\Big]$$
$$=\frac{1}{n_1-p-1}\Sigma_{33\cdot12}^{-1}\Sigma_{31}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1},$$
$$\mathrm{E}\Big[(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}^{(1)}\Gamma_{33\cdot12}^{(1)}\Sigma_{31}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)},\Gamma^{(3)}\Big]$$
$$=\frac{1}{n_1-p-1}\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}\Sigma_{33\cdot12}^{-1}\Sigma_{31}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1},$$
$$\mathrm{E}\Big[(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}^{(1)}\Gamma_{33\cdot12}^{(1)}\Gamma_{3(12)}^{(1)}(\Gamma_{(12)(12)}^{(1)})^{-1}\Sigma_{(12)1}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)},\Gamma^{(3)}\Big]$$
$$=\frac{1}{n_1-p-1}\Big[p_3(\Gamma_{(12)(12)}^{(1)})^{-1}\Sigma_{(12)1}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}$$
$$+\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}\Sigma_{33\cdot12}\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\Sigma_{(12)1}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}\Big],$$
$$\mathrm{E}\Big[(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}(\Gamma_{21}^{(1)}+\Gamma_{21}^{(2)})(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\Sigma_{11}$$
$$\times(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}|\Gamma_{11}^{(1)},\Gamma_{11}^{(2)},\Gamma^{(3)}\Big]$$
$$=\frac{1}{n_{(12)}-p_1-p_2-1}\Big[p_2(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}\Sigma_{11}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}$$
$$+\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot1}^{-1}\Sigma_{21}(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)})^{-1}\Big].$$

To obtain the above expectations of random matrices, we should note the dependence of each random matrix. First, since $\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)}+\Gamma^{(3)}$ is independent of $\overline{\boldsymbol{x}}_1^{[1,3]}-\overline{\boldsymbol{x}}_1^{[2,3]}$, the expectation stated in (28) can be obtained:

$$\frac{n^2(n-1)}{(n-p_1)(n-p_1-1)(n-p_1-3)}\left(\delta_{11}^2+\frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}}\right).$$

Similarly to the above, (29) can be also obtained:

$$-\frac{n_{(12)}^2(n_{(12)}-1)}{(n_{(12)}-p_1)(n_{(12)}-p_1-1)(n_{(12)}-p_1-3)}\left(\delta_{11}^2+\frac{p_1(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\right).$$

Next, by the exact and conditional expectations, both (30) and (31) are canceled. In addition, (34) and (35), (38) and (39), (42) and (43) are canceled similarly. Therefore, we consider the expectations of (32)–(33), (36)–(37), (40)–(41) and (44)–(45). As to the expectations of (32)–(33), using Lemma 5, we can obtain as follows:

$$\frac{2nn_{(12)}p_2(n-1)}{(n-p_1)(n-p_1-3)(n_{(12)}-p_1-1)(n_{(12)}-p_1-p_2-1)}\left(\delta_{11}^2+\frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}}\right).$$

Similarly, we can express the expectations of (37)–(38), (41)–(42) and (45)–(46) as follows:

$$-\frac{2n_{(12)}^2p_2(n_{(12)}-1)}{(n_{(12)}-p_1)(n_{(12)}-p_1-3)(n_{(12)}-p_1-1)(n_{(12)}-p_1-p_2-1)}\left(\delta_{11}^2+\frac{p_1(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\right),$$
$$\frac{2nn_1p_3(n-1)}{(n-p_1)(n-p_1-3)(n_1-p-1)(n_1-p_1-p_2-1)}\left(\delta_{11}^2+\frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}}\right),$$

$$-\frac{2n_{(12)}n_1p_3(n_{(12)}-1)}{(n_{(12)}-p_1)(n_{(12)}-p_1-3)(n_1-p-1)(n_1-p_1-p_2-1)}\left(\delta_{11}^2+\frac{p_1(N_{(12)}^{(1)}+N_{(12)}^{(2)})}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\right),$$

respectively. We can obtain the result by substituting all the expectations in order. Thus, the proof is completed.

A.2 Proof of Theorem 8

First, as to d_{11}^2, using Lemma 2, we can obtain

$$\mathrm{E}(d_{11}^2)=\frac{n}{n-p_1-1}\left(\delta_{11}^2+\frac{p_1(N^{(1)}+N^{(2)})}{N^{(1)}N^{(2)}}\right). \tag{46}$$

Next, we consider the expectation of d_{12}^2. Then it can be expressed as follows:

$$\begin{aligned}
d_{12}^2 = d_{11}^2 &\\
&+n_{(12)}(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]})'(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]}) &(47)\\
&+n_{(12)}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}\\
&\qquad\times(\Gamma_{21}^{(1)}+\Gamma_{21}^{(2)})(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]}) &(48)\\
&-2n_{(12)}(\overline{\boldsymbol{x}}_1^{[1,2]}-\overline{\boldsymbol{x}}_1^{[2,2]})'(\Gamma_{11}^{(1)}+\Gamma_{11}^{(2)})^{-1}(\Gamma_{12}^{(1)}+\Gamma_{12}^{(2)})(\Gamma_{(12)(12)}^{(1)}+\Gamma^{(2)})_{22\cdot1}^{-1}(\overline{\boldsymbol{x}}_2^{[1,2]}-\overline{\boldsymbol{x}}_2^{[2,2]}). &(49)
\end{aligned}$$

By using Lemmas 1 and 2, the expectation of (47) can be expressed as

$$\frac{n_{(12)}}{n_{(12)}-p_1-p_2-1}\left((\boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)})'\Sigma_{22\cdot1}^{-1}(\boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)})+\frac{N_{(12)}^{(1)}+N_{(12)}^{(2)}}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\mathrm{tr}(\Sigma_{22}\Sigma_{22\cdot1}^{-1})\right). \tag{50}$$

By using Lemmas 1–4, the expectation of (48) can be expressed as

$$\begin{aligned}
&\frac{N_{(12)}^{(1)}+N_{(12)}^{(2)}}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\cdot\frac{n_{(12)}}{n_{(12)}-p_1-p_2-1}\left(\frac{p_1p_2}{n_{(12)}-p_1-1}+\mathrm{tr}(\Sigma_{12}\Sigma_{22\cdot1}^{-1}\Sigma_{21}\Sigma_{11}^{-1})\right)\\
&+\frac{n_{(12)}}{n_{(12)}-p_1-p_2-1}\left(\frac{p_2}{n_{(12)}-p_1-1}(\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)})'\Sigma_{11}^{-1}(\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)})\right.\\
&\qquad\left.+(\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)})'\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot1}^{-1}\Sigma_{21}\Sigma_{11}^{-1}(\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)})\right).
\end{aligned} \tag{51}$$

Moreover, by using Lemmas 1–3, the expectation of (49) can be expressed as

$$-\frac{2n_{(12)}}{n_{(12)}-p_1-p_2-1}\left((\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)})'\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22\cdot1}^{-1}(\boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)})+\frac{N_{(12)}^{(1)}+N_{(12)}^{(2)}}{N_{(12)}^{(1)}N_{(12)}^{(2)}}\mathrm{tr}(\Sigma_{12}\Sigma_{22\cdot1}^{-1}\Sigma_{21}\Sigma_{11}^{-1})\right). \tag{52}$$

Therefore, we can obtain (26).

Similarly, D^2 can be written as

$$\begin{aligned}
D^2 = d_{12}^2 &\\
&+n_1(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]})'(\Gamma_{33\cdot12}^{(1)})^{-1}(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]}) &(53)\\
&+n_1\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]}-\overline{\boldsymbol{x}}_1^{[2,1]}\\ \overline{\boldsymbol{x}}_2^{[1,1]}-\overline{\boldsymbol{x}}_2^{[2,1]}\end{pmatrix}'(\Gamma_{(12)(12)}^{(1)})^{-1}\Gamma_{(12)3}(\Gamma_{33\cdot12}^{(1)})^{-1}\Gamma_{3(12)}(\Gamma_{(12)(12)}^{(1)})^{-1}\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]}-\overline{\boldsymbol{x}}_1^{[2,1]}\\ \overline{\boldsymbol{x}}_2^{[1,1]}-\overline{\boldsymbol{x}}_2^{[2,1]}\end{pmatrix} &(54)\\
&-2n_1(\overline{\boldsymbol{x}}_3^{[1,1]}-\overline{\boldsymbol{x}}_3^{[2,1]})'(\Gamma_{33\cdot12}^{(1)})^{-1}\Gamma_{3(12)}(\Gamma_{(12)(12)}^{(1)})^{-1}\begin{pmatrix}\overline{\boldsymbol{x}}_1^{[1,1]}-\overline{\boldsymbol{x}}_1^{[2,1]}\\ \overline{\boldsymbol{x}}_2^{[1,1]}-\overline{\boldsymbol{x}}_2^{[2,1]}\end{pmatrix}. &(55)
\end{aligned}$$

As to the expectation of d_{12}^2, we use the formula stated in (26). Then, we have only the expectations of (53)–(55). By using Lemmas 1 and 2, the expectation of (53) can be expressed as follows:

$$\frac{n_1}{n_1-p-1}\left((\boldsymbol{\mu}_3^{(1)}-\boldsymbol{\mu}_3^{(2)})'\Sigma_{33\cdot 12}^{-1}(\boldsymbol{\mu}_3^{(1)}-\boldsymbol{\mu}_3^{(2)})+\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\mathrm{tr}(\Sigma_{33}\Sigma_{33\cdot 12})^{-1}\right). \tag{57}$$

By using Lemmas 1–5, the expectation of (48) can be obtained as

$$\begin{aligned}&\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\cdot\frac{n_1}{n_1-p-1}\left(\frac{(p_1+p_2)p_3}{n_1-p_1-p_2-1}+\mathrm{tr}(\Sigma_{(12)3}\Sigma_{33\cdot 12}^{-1}\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1})\right)\\ &+\frac{n_1}{n_1-p-1}\left(\frac{p_3}{n_1-p_1-p_2-1}\begin{pmatrix}\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)}\\ \boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)}\end{pmatrix}'\Sigma_{(12)(12)}^{-1}\begin{pmatrix}\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)}\\ \boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)}\end{pmatrix}\right.\\ &\qquad\left.+\begin{pmatrix}\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)}\\ \boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)}\end{pmatrix}'\Sigma_{(12)(12)}^{-1}\Sigma_{(12)3}\Sigma_{33\cdot 12}^{-1}\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\begin{pmatrix}\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)}\\ \boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)}\end{pmatrix}\right).\end{aligned} \tag{57}$$

Moreover, using Lemmas 1–2 and 4, the expectation of (55) can be expressed as follows:

$$\begin{aligned}&-\frac{2n_1}{n_1-p-1}\left((\boldsymbol{\mu}_3^{(1)}-\boldsymbol{\mu}_3^{(2)})'\Sigma_{33\cdot 12}^{-1}\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1}\begin{pmatrix}\boldsymbol{\mu}_1^{(1)}-\boldsymbol{\mu}_1^{(2)}\\ \boldsymbol{\mu}_2^{(1)}-\boldsymbol{\mu}_2^{(2)}\end{pmatrix}\right.\\ &\qquad\left.+\frac{N_1^{(1)}+N_1^{(2)}}{N_1^{(1)}N_1^{(2)}}\mathrm{tr}(\Sigma_{(12)3}\Sigma_{33\cdot 12}^{-1}\Sigma_{3(12)}\Sigma_{(12)(12)}^{-1})\right).\end{aligned} \tag{58}$$

Thus, the proof is completed.

References

[1] Anderson, T. W. and Olkin, I. (1985). Maximum-lilelihood estimation of the parameters of a multivariate normal distribution, *Linear Algebra and its Applications*, **70**, 147–171.

[2] Fujikoshi, Y. and Seo, T. (1998). Asymptotic approximations for EPMC's of the linear and quadratic discriminant functions when the sample sizes and the dimension are large. *Random Operators and Stochastic Equations*, **6**, 269–280.

[3] Hao, J. and Krishnamoorthy, K. (2001). Inferences on a normal covariance matrix and generalized variance with monotone missing data. *Journal of Multivariate Analysis*, **78**, 62–82.

[4] Kanda, T. and Fujikoshi, Y. (1998). Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data. *American Journal of Mathematical and Management Science*, **18**, 161–190.

[5] Koizumi, K. and Seo, T. (2009a). Testing equality of two mean vectors and simultaneous confidence intervals in repeated measures with missing data. *Journal of tha Japanese Society of Computational Statistics*, **22**, 33–41.

[6] Koizumi, K. and Seo, T. (2009b). Simultaneous confidence intervals among k mean vectors in repeated measures with missing data. *American Journal of Mathematical and Management Sciences*, **29**, 263–275.

[7] Lachenbruch, P. A. (1968). On expected probabilities of misclassification in discriminant analysis, necessary sample size, and a relation with the multiple correlation coefficient. *Biometrics,* **24**, 823–834.

[8] Okamoto, M. (1963). An asymptotic expansion for the distribution of the linear discriminant function. *Annals of Mathematical Statistics*, **34**, 1286–1301.

[9] Seo, T. and Srivastava, M. S. (2000). Testing equality of means and simultaneous confidence intervals in repeated measures with missing data, *Biometrical Journal*, **42**, 981–993.

[10] Shutoh, N. (2011). Asymptotic expansion relating to discrimination based on two-step monotone missing samples. *Journal of Statistical Planning and Inference*, **141**, 1297–1306.

[11] Shutoh, N., Hyodo, M. and Seo, T. (2011). An asymptotic approximation for EPMC in linear discriminant analysis based on two-step monotone missing samples. *Journal of Multivariate Analysis*, **102**, 252–263.

[12] Shutoh, N., Kusumi, M., Morinaga, W., Yamada, S. and Seo, T. (2010). Testing equality of mean vector in two sample problem with missing data, *Communications in Statistics – Simulation and Computation*, **39**, 487–500.

[13] Srivastava, M. S. (1985). Multivariate data with missing observations, *Communications in Statistics – Theory and Methods*, **14**, 775–792.

[14] Srivastava, M. S. and Carter, E. M. (1986). The maximum likelihood method for non–response in sample survey, *Survey Methodology*, **12**, 61–72.

Table 1

The accuracy of asymptotic approximations for EPMC $\Delta = 1.05$

(p_1, p_2, p_3)	(M_1, M_2, M_3)	$e_1(2\|1)$	$\widehat{e}_L(2\|1)$	$\widehat{e}_L(2\|1) - e_1(2\|1)$	$e_3(2\|1)$	$\widehat{e}_K(2\|1)$	$\widehat{e}_K(2\|1) - e_3(2\|1)$
(2,1,1)	(10,10,10)	0.359936	0.408087	0.048151	0.339725	0.368785	0.029060
	(20,20,20)	0.331155	0.356067	0.024912	0.319168	0.330313	0.011145
	(50,50,50)	0.313104	0.322412	0.009308	0.307796	0.311244	0.003448
	(100,100,100)	0.306688	0.311126	0.004438	0.304240	0.305427	0.001187
(2,2,1)	(10,10,10)	0.370428	0.418821	0.048393	0.348130	0.380535	0.032405
	(20,20,20)	0.339964	0.363719	0.023755	0.324471	0.335633	0.011162
	(50,50,50)	0.316683	0.326269	0.009586	0.309150	0.313409	0.004259
	(100,100,100)	0.308855	0.313122	0.004267	0.304839	0.306461	0.001622
(3,2,1)	(10,10,10)	0.380722	0.428697	0.047975	0.355718	0.390997	0.035279
	(20,20,20)	0.347301	0.370401	0.023100	0.328517	0.339423	0.010906
	(50,50,50)	0.321139	0.329747	0.008608	0.311504	0.314788	0.003284
	(100,100,100)	0.311202	0.315056	0.003854	0.305850	0.307199	0.001349
(4,3,2)	(10,10,10)	0.406218	0.456319	0.050101	0.390769	0.443115	0.052346
	(20,20,20)	0.365220	0.388379	0.023159	0.345841	0.360694	0.014853
	(50,50,50)	0.329934	0.339762	0.009828	0.318157	0.322726	0.004569
	(100,100,100)	0.316118	0.320669	0.004551	0.309331	0.311087	0.001756
(5,4,3)	(10,10,10)	0.428524	0.483275	0.054751	0.424156	0.484035	0.059879
	(20,20,20)	0.380594	0.403624	0.023030	0.363805	0.381424	0.017619
	(50,50,50)	0.339997	0.348760	0.008763	0.326214	0.330501	0.004287
	(100,100,100)	0.321302	0.325947	0.004645	0.312498	0.314917	0.002419

Table 2

The accuracy of asymptotic approximations for EPMC $N_i^{(1)} = 2N_i^{(2)}$ $(i = 1, 2, 3)$

(p_1, p_2, p_3)	(L_1, L_2)	$e_1(2\|1)$	$\widehat{e}_L(2\|1)$	$\widehat{e}_L(2\|1) - e_1(2\|1)$	$e_3(2\|1)$	$\widehat{e}_K(2\|1)$	$\widehat{e}_K(2\|1) - e_3(2\|1)$
(2,1,1)	(10,5)	0.341016	0.400051	0.059035	0.329661	0.373823	0.044162
	(20,10)	0.322140	0.349749	0.027609	0.313375	0.327650	0.014275
	(50,25)	0.309707	0.319765	0.010058	0.305294	0.309526	0.004232
	(100,50)	0.304642	0.309789	0.005147	0.302817	0.304500	0.001683
(2,2,1)	(10,5)	0.347668	0.408917	0.061249	0.337201	0.388290	0.051089
	(20,10)	0.326629	0.353791	0.027162	0.317423	0.331659	0.014236
	(50,25)	0.312397	0.322064	0.009667	0.307898	0.311015	0.003117
	(100,50)	0.305818	0.311036	0.005218	0.302890	0.305209	0.002319
(3,2,1)	(10,5)	0.354980	0.419486	0.064506	0.344932	0.405943	0.061011
	(20,10)	0.331159	0.357808	0.026649	0.320938	0.335230	0.014292
	(50,25)	0.314004	0.324277	0.010273	0.307139	0.312074	0.004935
	(100,50)	0.307927	0.312253	0.004326	0.304330	0.305680	0.001350
(4,3,2)	(10,5)	0.377852	0.466836	0.088984	0.380846	0.474031	0.093185
	(20,10)	0.341389	0.369563	0.028174	0.334242	0.355649	0.021407
	(50,25)	0.319890	0.330032	0.010142	0.312986	0.317951	0.004965
	(100,50)	0.311208	0.315772	0.004564	0.307150	0.308469	0.001319

Table 3

The accuracy of asymptotic approximations for EPMC $\Delta = 1.05$, $\delta_{12} = 0.70$ and $\delta_{11} = 0.35$

(p_1, p_2, p_3)	(M_1, M_2, M_3)	$e_1(2\|1)$	$\widehat{e}_L(2\|1)$	$\widehat{e}_L(2\|1) - e_1(2\|1)$	$e_3(2\|1)$	$\widehat{e}_K(2\|1)$	$\widehat{e}_K(2\|1) - e_3(2\|1)$
(2,1,1)	(10,10,10)	0.359864	0.408075	0.048211	0.340666	0.382973	0.042307
	(20,20,20)	0.331345	0.356096	0.024751	0.319122	0.338599	0.019477
	(50,50,50)	0.312972	0.322477	0.009505	0.306782	0.314754	0.007972
	(100,100,100)	0.306497	0.311132	0.004635	0.303539	0.307168	0.003629
(2,2,1)	(10,10,10)	0.369826	0.418719	0.048893	0.349179	0.393023	0.043844
	(20,20,20)	0.339666	0.363593	0.023927	0.324339	0.343898	0.019559
	(50,50,50)	0.316627	0.326276	0.009649	0.309321	0.317064	0.007743
	(100,100,100)	0.307680	0.313155	0.005475	0.304547	0.308346	0.003799
(3,2,1)	(10,10,10)	0.380918	0.428740	0.047822	0.356425	0.402307	0.045882
	(20,20,20)	0.348265	0.370548	0.022283	0.329676	0.347964	0.018288
	(50,50,50)	0.320395	0.329813	0.009418	0.310769	0.318728	0.007959
	(100,100,100)	0.311000	0.315080	0.004080	0.306089	0.309215	0.003126
(4,3,2)	(10,10,10)	0.406507	0.456331	0.049824	0.388825	0.444124	0.055299
	(20,20,20)	0.365315	0.388557	0.023242	0.346084	0.367291	0.021207
	(50,50,50)	0.329834	0.339811	0.009977	0.317700	0.326661	0.008961
	(100,100,100)	0.315484	0.320666	0.005182	0.308911	0.313250	0.004339
(5,4,3)	(10,10,10)	0.427818	0.483223	0.055405	0.418758	0.480375	0.061617
	(20,20,20)	0.380761	0.403758	0.022997	0.362696	0.385061	0.022365
	(50,50,50)	0.339419	0.348717	0.009298	0.326680	0.334229	0.007549
	(100,100,100)	0.321206	0.325955	0.004749	0.313356	0.317207	0.003851

Table 4

The accuracy of asymptotic approximations for EPMC $\Delta = 1.05$, $\delta_{12} = 0.60$ and $\delta_{11} = 0.15$

(p_1, p_2, p_3)	(M_1, M_2, M_3)	$e_1(2\|1)$	$\widehat{e}_L(2\|1)$	$\widehat{e}_L(2\|1) - e_1(2\|1)$	$e_3(2\|1)$	$\widehat{e}_K(2\|1)$	$\widehat{e}_K(2\|1) - e_3(2\|1)$
(2,1,1)	(10,10,10)	0.359774	0.408057	0.048283	0.341192	0.385583	0.044391
	(20,20,20)	0.331665	0.356112	0.024447	0.319137	0.340104	0.020967
	(50,50,50)	0.312801	0.322490	0.009689	0.306608	0.315360	0.008752
	(100,100,100)	0.306588	0.311133	0.004545	0.303752	0.307466	0.003714
(2,2,1)	(10,10,10)	0.369902	0.418726	0.048824	0.349376	0.395237	0.045861
	(20,20,20)	0.339708	0.363586	0.023878	0.324349	0.345369	0.021020
	(50,50,50)	0.316861	0.326270	0.009409	0.309499	0.317686	0.008187
	(100,100,100)	0.308110	0.313158	0.005048	0.304781	0.308674	0.003893
(3,2,1)	(10,10,10)	0.380711	0.428736	0.048025	0.356027	0.404235	0.048208
	(20,20,20)	0.348279	0.370548	0.022269	0.329697	0.349421	0.019724
	(50,50,50)	0.320518	0.329834	0.009316	0.310953	0.319407	0.008454
	(100,100,100)	0.310894	0.315084	0.004190	0.306049	0.309559	0.003510
(4,3,2)	(10,10,10)	0.406227	0.456313	0.050086	0.388386	0.444103	0.055717
	(20,20,20)	0.365185	0.339816	0.023403	0.346026	0.368320	0.022294
	(50,50,50)	0.329916	0.339811	0.009900	0.318256	0.327318	0.009062
	(100,100,100)	0.315206	0.320672	0.005466	0.308670	0.313628	0.004958
(5,4,3)	(10,10,10)	0.427584	0.483261	0.055677	0.417638	0.479598	0.061960
	(20,20,20)	0.381044	0.403787	0.022743	0.362690	0.385542	0.022852
	(50,50,50)	0.339250	0.348711	0.009461	0.326564	0.334834	0.008270
	(100,100,100)	0.321289	0.325954	0.004665	0.313569	0.317592	0.004023