Discussiones Mathematicae
Probability and Statistics 33 (2013) 171–190
doi:10.7151/dmps.1153

TESTS FOR PROFILE ANALYSIS BASED ON TWO-STEP MONOTONE MISSING DATA

Mizuki Onozawa

Department of Mathematical Information Science, Graduate School of Science
Tokyo University of Science, Tokyo 162-8601, Japan

e-mail: j1412701@ed.tus.ac.jp

Sho Takahashi

Chiba University Hospital Clinical Research Center
Chiba 260-8677, Japan

e-mail: sho@chiba-u.jp

and

Takashi Seo

Department of Mathematical Information Science, Faculty of Science
Tokyo University of Science, Tokyo 162-8601, Japan

e-mail: seo@rs.kagu.tus.ac.jp

Abstract

In this paper, we consider profile analysis for the observations with two-step monotone missing data. There exist three interesting hypotheses – the parallelism hypothesis, level hypothesis, and flatness hypothesis – when comparing the profiles of some groups. The T^2-type statistics and their asymptotic null distributions for the three hypotheses are given for two-sample profile analysis. We propose the approximate upper percentiles of these test statistics. When the data do not have missing observations, the test statistics perform lower than the usual test statistics, for example, as in [8]. Further, we consider a parallel profile model for several groups when the data have two-step monotone missing observations. Under the assumption of non-missing data, the likelihood ratio test procedure is derived by [16]. We derive the test statistic based on the likelihood ratio. Finally, in order to investigate the accuracy for the null distributions of the proposed statistics, we perform a Monte Carlo simulation for some selected parameters values.

Keywords: Hotelling's T^2-type statistic, likelihood ratio, profile analysis, two-step monotone missing data.
2010 Mathematics Subject Classification: 62H15.

1. Introduction

Profile analysis is a statistical method used to compare the profiles of several groups. In a normal population, the profile analysis for a two-sample problem has been discussed using Hotelling's T^2-type statistic (see, e.g., [8]). Further, [16] gave a profile analysis of several groups based on the likelihood ratio. For the assumption of nonnormality, [9] discussed profile analysis in elliptical populations. Further, [7] obtained asymptotic expansions of the null distributions of some test statistics for general distributions.

At the same time, we often encounter the problem of missing data in many practical situations. For samples with observations missing at random, many statistical methods have been developed by [3, 14, 15], and [12] among others. Moreover, when the missing observations are of the monotone-type, the test for the equality of means and simultaneous confidence intervals in repeated measures with an intraclass correlation model was discussed by [11] in a one-sample problem, [5] in a two-sample problem, and [6] in a k-sample problem. For two-step monotone missing data, [2] and [10] considered tests for the mean vector in a one-sample problem. [1] obtained the maximum likelihood estimators (MLEs) of the mean vector and covariance matrix in a one-sample problem for two-step monotone missing data, and [4] discussed the distribution of these MLEs and expanded for K-step monotone missing data. In the same way as [1], the MLEs in two-sample problem have been obtained (e.g., [13]).

In this paper, we consider a profile analysis for a two-sample problem comprising several groups and two-step monotone missing observations. In particular, for several groups, we consider the parallelism hypothesis.

The organization of this paper is as follows. In Section 2, we consider a profile analysis for complete data. In Section 3, we derive the MLEs of $\boldsymbol{\mu}^{(i)}$ and Σ when the missing observations are of the two-step monotone-type. In Section 4, we give the T^2-type statistics for profile analysis. In Section 5, we give the likelihood ratio test statistic for the parallelism hypothesis. In Section 6, we perform a Monte Carlo simulation to investigate the accuracy for the null distributions of these statistics. Finally, in Section 7, we conclude this study.

2. Profile analysis for complete data

In this section, we consider the test statistics when the data have non-missing observations. Let the p-dimensional random vector $\boldsymbol{x}_j^{(i)}$ be independently distributed as $N_p(\boldsymbol{\mu}^{(i)}, \Sigma)$ $(j = 1, \ldots, N_1^{(i)},\ i = 1, 2)$, where $\boldsymbol{\mu}^{(i)} = (\mu_1^{(i)}, \ldots, \mu_p^{(i)})'$. Let the i-th sample mean vector, the i-th sample covariance matrix, and the pooled sample covariance matrix be

$$\begin{aligned}\overline{\boldsymbol{x}}^{(i)} &= \frac{1}{N_1^{(i)}} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{x}_j^{(i)},\ S_i = \frac{1}{N_1^{(i)} - 1} \sum_{j=1}^{N_1^{(i)}} (\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}}^{(i)})(\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}}^{(i)})', \\ S &= \frac{(N_1^{(1)} - 1)S_1 + (N_1^{(2)} - 1)S_2}{N_1^{(1)} + N_1^{(2)} - 2},\end{aligned}$$

respectively. When carrying out a profile analysis for two samples, we first consider the parallelism hypothesis that is expressed as

$$H_{P_2} : C\boldsymbol{\mu}^{(1)} = C\boldsymbol{\mu}^{(2)} \quad \text{vs.} \quad A_{P_2} \neq H_{P_2},$$

where C is a $(p-1) \times p$ matrix of rank $p - 1$ such that $C\mathbf{1}_p = \mathbf{0}$ and $\mathbf{1}_p$ is a p-vector of ones. The test statistic for testing hypothesis H_{P_2} can be written as

$$T_{Pc}^2 = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})'C' \left\{ \frac{N_1^{(1)} + N_1^{(2)}}{N_1^{(1)} N_1^{(2)}} (CSC') \right\}^{-1} C(\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)}).$$

In normal populations,

$$T_{Pc}^2 \sim \frac{(N_1^{(1)} + N_1^{(2)} - 2)(p-1)}{N_1^{(1)} + N_1^{(2)} - p} F_{p-1, N_1^{(1)} + N_1^{(2)} - p}.$$

If the parallelism hypothesis is true, we test the level hypothesis or the flatness hypothesis. The level hypothesis is expressed as

$$H_{L_2} : \mathbf{1}_p'\boldsymbol{\mu}^{(1)} = \mathbf{1}_p'\boldsymbol{\mu}^{(2)} \quad \text{vs.} \quad A_{L_2} \neq H_{L_2}.$$

The test statistic for testing hypothesis H_{L_2} can be written as

$$T_{Lc}^2 = (\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)})'\mathbf{1}_p \left\{ \frac{N_1^{(1)} + N_1^{(2)}}{N_1^{(1)} N_1^{(2)}} (\mathbf{1}_p' S \mathbf{1}_p) \right\}^{-1} \mathbf{1}_p'(\overline{\boldsymbol{x}}^{(1)} - \overline{\boldsymbol{x}}^{(2)}).$$

In normal populations,

$$T_{Lc}^2 \sim F_{1, N_1^{(1)} + N_1^{(2)} - 2}.$$

Further, the flatness hypothesis is expressed as

$$H_{F_2} : C(\boldsymbol{\mu}^{(1)} + \boldsymbol{\mu}^{(2)}) = \mathbf{0} \quad \text{vs.} \quad A_{F_2} \neq H_{F_2}.$$

The test statistic for testing hypothesis H_{F_2} can be written as

$$T_{Fc}^2 = \overline{\boldsymbol{x}}_{12}' C' \left\{ \frac{1}{N_1^{(1)} + N_1^{(2)}} CSC' \right\}^{-1} C\overline{\boldsymbol{x}}_{12},$$

where

$$\overline{\boldsymbol{x}}_{12} = \frac{N_1^{(1)}}{N_1^{(1)} + N_1^{(2)}} \overline{\boldsymbol{x}}^{(1)} + \frac{N_1^{(2)}}{N_1^{(1)} + N_1^{(2)}} \overline{\boldsymbol{x}}^{(2)}.$$

In normal populations,

$$T_{Fc}^2 \sim \frac{(N_1^{(1)} + N_1^{(2)} - 2)(p-1)}{N_1^{(1)} + N_1^{(2)} - p} F_{p-1, N_1^{(1)} + N_1^{(2)} - p}.$$

In addition, we consider a parallelism hypothesis of several groups when the data have non-missing observations. Let $\boldsymbol{x}_1^{(i)}, \ldots, \boldsymbol{x}_{N_1^{(i)}}^{(i)}$ be $N_1^{(i)}$ independent observations from $N_p(\boldsymbol{\mu}^{(i)}, \Sigma)$ $(i = 1, \ldots, k)$. Then we consider the primarily testing the parallelism hypothesis as follows:

$$H_{P_k} : C\boldsymbol{\mu}^{(1)} = \cdots = C\boldsymbol{\mu}^{(k)} \quad \text{vs.} \quad A_{P_k} \neq H_{P_k}.$$

The MLEs of $\boldsymbol{\mu}^{(i)}$ and Σ under A_{P_k} are

$$\overline{\boldsymbol{x}}^{(i)} = \frac{1}{N_1^{(i)}} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{x}_j^{(i)}, \quad \widehat{\Sigma}_c = \frac{1}{N_1} \sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} (\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}}^{(i)})(\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}}^{(i)})',$$

respectively, where $N_1 = \sum_{i=1}^{k} N_1^{(i)}$. In contrast, the MLEs of $\boldsymbol{\mu}$ and Σ under H_{P_k} are

$$\overline{\boldsymbol{x}} = \frac{1}{N_1} \sum_{i=1}^{k} \sum_{j=1}^{N^{(i)}} \boldsymbol{x}_j^{(i)}, \quad \widetilde{\Sigma}_c = \frac{1}{N_1} \sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} (\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}})(\boldsymbol{x}_j^{(i)} - \overline{\boldsymbol{x}})',$$

respectively. For complete data, using these MLEs, we can construct the following likelihood ratio:

$$\Lambda_c = \frac{|C\widehat{\Sigma}_c C'|^{\frac{1}{2}N_1}}{|C\widetilde{\Sigma}_c C'|^{\frac{1}{2}N_1}}.$$

The likelihood ratio test statistic, $-2\log\Lambda_c$, is asymptotically distributed as a χ^2 distribution with $(p-1)(k-1)$ degrees of freedom as $N_1^{(i)}$s tend to infinity (see [16]). Hence, we reject H_{P_k} when $-2\log\Lambda_c > \chi^2_{(p-1)(k-1),\alpha}$, where $\chi^2_{(p-1)(k-1),\alpha}$

is the upper 100α percentile of a χ^2 distribution with $(p-1)(k-1)$ degrees of freedom. However, convergence to the asymptotic χ^2 distribution can be improved by considering an asymptotic expansion for the likelihood ratio statistic and deriving the modified likelihood ratio statistic as $-2\rho_{c_1}\log\Lambda_c$, where

$$\rho_{c_1} = 1 - \frac{1}{2N_1}(p+k+1).$$

3. MLEs

We consider the case when the missing observations are of the two-step monotone-type. Observations $\{x_{\ell j}^{(i)}\}$ can be written in the following form:

$$\begin{pmatrix} x_{11}^{(i)} & \cdots & x_{1p_1}^{(i)} & x_{1,p_1+1}^{(i)} & \cdots & x_{1p}^{(i)} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{N_1^{(i)}1}^{(i)} & \cdots & x_{N_1^{(i)}p_1}^{(i)} & x_{N_1^{(i)},p_1+1}^{(i)} & \cdots & x_{N_1^{(i)}p}^{(i)} \\ x_{N_1^{(i)}+1,1}^{(i)} & \cdots & x_{N_1^{(i)}+1,p_1}^{(i)} & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{N^{(i)}1}^{(i)} & \cdots & x_{N^{(i)}p_1}^{(i)} & * & \cdots & * \end{pmatrix},$$

where $*$ denotes missing component. Let $\boldsymbol{x}_j^{(i)} \equiv (\boldsymbol{x}_{1j}^{(i)\prime}, \boldsymbol{x}_{2j}^{(i)\prime})'$ $(j = 1,\ldots,N_1^{(i)},\ i = 1,\ldots,k)$ be a p-dimensional observation vector from the i-th group with complete data. Let $\boldsymbol{x}_{1j}^{(i)}$ $(j = N_1^{(i)}+1,\ldots,N^{(i)})$ be p_1-dimensional vectors based on $N_2^{(i)}$ $(= N^{(i)} - N_1^{(i)})$ observations. Now, we assume the distribution of observation vectors:

$$\boldsymbol{x}_j^{(i)} \sim N_p(\boldsymbol{\mu}^{(i)}, \Sigma)\ (j = 1,\ldots,N_1^{(i)},\ i = 1,\ldots,k),$$
$$\boldsymbol{x}_{1j}^{(i)} \sim N_{p_1}(\boldsymbol{\mu}_1^{(i)}, \Sigma_{11})\ (j = N_1^{(i)}+1,\ldots,N^{(i)},\ i = 1,\ldots,k),$$

respectively, where

$$\boldsymbol{\mu}^{(i)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(i)} \\ \boldsymbol{\mu}_2^{(i)} \end{pmatrix},\ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

and $\boldsymbol{\mu}^{(i)}$ and Σ are partitioned according to the blocks of the data set. Therefore, $\boldsymbol{\mu}_\ell^{(i)}$ $(\ell = 1,2)$ is a p_ℓ-dimensional vector and $\Sigma_{\ell m}$ $(\ell, m = 1,2)$ is a $p_\ell \times p_m$ matrix.

We give some notations for the sample mean vectors. Let $\overline{\boldsymbol{x}}_{1T}^{(i)}$ be the sample mean vector of $\boldsymbol{x}_{11}^{(i)},\ldots,\boldsymbol{x}_{1N^{(i)}}^{(i)}$. Let $(\overline{\boldsymbol{x}}_{1F}^{(i)\prime},\ \overline{\boldsymbol{x}}_{2F}^{(i)\prime})'$ be the sample mean vector of

$\boldsymbol{x}_1^{(i)}, \ldots, \boldsymbol{x}_{N_1^{(i)}}^{(i)}$, where $\overline{\boldsymbol{x}}_{\ell F}^{(i)\prime} : p_\ell \times 1$ $(\ell = 1, 2)$. That is,

$$\overline{\boldsymbol{x}}_{1T}^{(i)} = \frac{1}{N^{(i)}} \sum_{j=1}^{N^{(i)}} \boldsymbol{x}_{1j}^{(i)}, \quad \overline{\boldsymbol{x}}_{1F}^{(i)} = \frac{1}{N_1^{(i)}} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{x}_{1j}^{(i)}, \quad \overline{\boldsymbol{x}}_{2F}^{(i)} = \frac{1}{N_1^{(i)}} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{x}_{2j}^{(i)}.$$

Since the MLEs based on the complete data case cannot be used, we have to estimate $\boldsymbol{\mu}^{(i)}$ and Σ under two-step monotone missing data. Let $\widehat{\boldsymbol{\mu}}^{(i)}$ and $\widehat{\Sigma}$ be the MLEs of $\boldsymbol{\mu}$ and Σ. These have the same patterns of partition as $\boldsymbol{\mu}^{(i)}$ and Σ. The likelihood function is

$$\begin{aligned} & L(\boldsymbol{\mu}^{(i)}, \Sigma) \\ = & \prod_{i=1}^{k} \left[\prod_{j=1}^{N_1^{(i)}} \frac{1}{(2\pi)^{\frac{p}{2}} \mid \Sigma \mid^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (\boldsymbol{x}_j^{(i)} - \boldsymbol{\mu}^{(i)})' \Sigma^{-1} (\boldsymbol{x}_j^{(i)} - \boldsymbol{\mu}^{(i)}) \right\} \right. \\ & \left. \times \prod_{j=N_1^{(i)}+1}^{N^{(i)}} \frac{1}{(2\pi)^{\frac{p_1}{2}} \mid \Sigma_{11} \mid^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (\boldsymbol{x}_{1j}^{(i)} - \boldsymbol{\mu}_1^{(i)})' \Sigma_{11}^{-1} (\boldsymbol{x}_{1j}^{(i)} - \boldsymbol{\mu}_1^{(i)}) \right\} \right]. \end{aligned}$$

Let A be a $p \times p$ transformation matrix:

$$A = \begin{pmatrix} I_{p_1} & O \\ -\Sigma_{21}\Sigma_{11}^{-1} & I_{p_2} \end{pmatrix}.$$

Then we have

$$A\boldsymbol{x}_j^{(i)} = \begin{pmatrix} \boldsymbol{x}_{1j}^{(i)} \\ \boldsymbol{x}_{2j}^{(i)} - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{x}_{1j}^{(i)} \end{pmatrix} \sim N_p(A\boldsymbol{\mu}^{(i)}, A\Sigma A'),$$

where the mean vector and the covariance matrix of transformed observation vectors are

$$\begin{aligned} A\boldsymbol{\mu}^{(i)} = \boldsymbol{\eta}^{(i)} = \begin{pmatrix} \boldsymbol{\eta}_1^{(i)} \\ \boldsymbol{\eta}_2^{(i)} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\mu}_1^{(i)} \\ \boldsymbol{\mu}_2^{(i)} - \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1^{(i)} \end{pmatrix}, \\ A\Sigma A' &= \begin{pmatrix} \Sigma_{11} & O \\ O & \Sigma_{22\cdot 1} \end{pmatrix}, \end{aligned}$$

and $\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. It should be noted that $\boldsymbol{\mu}^{(i)}$ and Σ have one-to-one correspondence with $\boldsymbol{\eta}^{(i)}$ and Ψ, where

$$\Psi = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{11}^{-1}\Sigma_{12} \\ \Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22\cdot 1} \end{pmatrix}.$$

For parameters $\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(k)}$ and Ψ, the likelihood function is

$$
\begin{aligned}
& L(\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(k)}, \Psi) \\
= \;& \text{Const.} \times \mid \Psi_{11} \mid^{-\frac{1}{2}N} \mid \Psi_{22} \mid^{-\frac{1}{2}N_1} \\
& \times \exp\left\{ -\frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{N^{(i)}} (\boldsymbol{x}_{1j}^{(i)} - \boldsymbol{\eta}_1^{(i)})' \Psi_{11}^{-1} (\boldsymbol{x}_{1j}^{(i)} - \boldsymbol{\eta}_1^{(i)}) \right\} \\
& \times \exp\left\{ -\frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} (\boldsymbol{x}_{2j}^{(i)} - \Psi_{21}\boldsymbol{x}_{1j}^{(i)} - \boldsymbol{\eta}_2^{(i)})' \Psi_{22}^{-1} (\boldsymbol{x}_{2j}^{(i)} - \Psi_{21}\boldsymbol{x}_{1j}^{(i)} - \boldsymbol{\eta}_2^{(i)}) \right\},
\end{aligned}
$$

where $N = \sum_{i=1}^{k} N^{(i)}$.

Differentiating the log likelihood function, we get that

$$
\begin{aligned}
\widehat{\boldsymbol{\eta}}_1^{(i)} &= \overline{\boldsymbol{x}}_{1T}^{(i)}, \\
\widehat{\boldsymbol{\eta}}_2^{(i)} &= \overline{\boldsymbol{x}}_{2F}^{(i)} - \widehat{\Psi}_{21}\overline{\boldsymbol{x}}_{1F}^{(i)},
\end{aligned}
$$

and that

$$
\begin{aligned}
\widehat{\Psi}_{11} &= \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{N^{(i)}} (\boldsymbol{x}_{1j}^{(i)} - \overline{\boldsymbol{x}}_{1T}^{(i)})(\boldsymbol{x}_{1j}^{(i)} - \overline{\boldsymbol{x}}_{1T}^{(i)})', \\
\widehat{\Psi}_{21} &= \left[\sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{z}_{2j}^{(i)} \boldsymbol{z}_{1j}'^{(i)} \right] \left[\sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{z}_{1j}^{(i)} \boldsymbol{z}_{1j}'^{(i)} \right]^{-1}, \\
\widehat{\Psi}_{22} &= \frac{1}{N_1} \left\{ \sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{z}_{2j}^{(i)} \boldsymbol{z}_{2j}'^{(i)} \right. \\
& \quad \left. - \left[\sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{z}_{2j}^{(i)} \boldsymbol{z}_{1j}'^{(i)} \right] \left[\sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{z}_{1j}^{(i)} \boldsymbol{z}_{1j}'^{(i)} \right]^{-1} \left[\sum_{i=1}^{k} \sum_{j=1}^{N_1^{(i)}} \boldsymbol{z}_{1j}^{(i)} \boldsymbol{z}_{2j}'^{(i)} \right] \right\}, \\
& \boldsymbol{z}_{1j}^{(i)} = \boldsymbol{x}_{1j}^{(i)} - \overline{\boldsymbol{x}}_{1F}^{(i)}, \quad \boldsymbol{z}_{2j}^{(i)} = \boldsymbol{x}_{2j}^{(i)} - \overline{\boldsymbol{x}}_{2F}^{(i)}.
\end{aligned}
$$

We thus obtain the MLEs of $\boldsymbol{\mu}^{(i)}$ and Σ in general:

$$\widehat{\boldsymbol{\mu}}^{(i)} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1^{(i)} \\ \widehat{\boldsymbol{\mu}}_2^{(i)} \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{x}}_{1T}^{(i)} \\ \overline{\boldsymbol{x}}_{2F}^{(i)} - \widehat{\Psi}_{21}(\overline{\boldsymbol{x}}_{1F}^{(i)} - \overline{\boldsymbol{x}}_{1T}^{(i)}) \end{pmatrix},$$

$$\widehat{\Sigma} = \begin{pmatrix} \widehat{\Sigma}_{11} & \widehat{\Sigma}_{12} \\ \widehat{\Sigma}_{21} & \widehat{\Sigma}_{22} \end{pmatrix} = \begin{pmatrix} \widehat{\Psi}_{11} & \widehat{\Psi}_{11}\widehat{\Psi}_{12} \\ \widehat{\Psi}_{21}\widehat{\Psi}_{11} & \widehat{\Psi}_{22} + \widehat{\Psi}_{21}\widehat{\Psi}_{11}\widehat{\Psi}_{12} \end{pmatrix}.$$

4. TWO-SAMPLE PROFILE ANALYSIS WITH TWO-STEP MONOTONE MISSING DATA

By using the MLEs given in Section 3, we obtain the T^2-type statistics. In this section, let $k = 2$. The T^2-type statistic under H_{P_2} can be written as

$$T_{Pm}^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' C' \{C \widehat{\Xi} C'\}^{-1} C (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}),$$

where $\widehat{\Xi}$ is the MLE of $\Xi = \{\mathrm{Cov}[\widehat{\boldsymbol{\mu}}^{(1)}] + \mathrm{Cov}[\widehat{\boldsymbol{\mu}}^{(2)}]\}$,

$$\widehat{\Xi} = \begin{pmatrix} \dfrac{N}{N^{(1)}N^{(2)}}\widehat{\Sigma}_{11} & \dfrac{N}{N^{(1)}N^{(2)}}\widehat{\Sigma}_{12} \\ \dfrac{N}{N^{(1)}N^{(2)}}\widehat{\Sigma}_{21} & \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2^{(1)}] + \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2^{(2)}] \end{pmatrix}$$

and

$$\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2^{(1)}] + \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2^{(2)}]$$
$$= \sum_{i=1}^{2} \left\{ \frac{1}{N_1^{(i)}} \left(\widehat{\Sigma}_{22} - \frac{N_2^{(i)}}{N^{(i)}} \widehat{\Sigma}_{21} \widehat{\Sigma}_{11}^{-1} \widehat{\Sigma}_{12} \right) + \frac{N_2^{(i)} p_1}{N^{(i)} N_1^{(i)} (N_1^{(i)} - p_1 - 2)} \widehat{\Sigma}_{22\cdot 1} \right\}.$$

For details of the MLEs, see [4]. T_{Pm}^2 is asymptotically distributed as a χ^2 distribution with $p-1$ degrees of freedom when $N_1^{(i)}$s are large.

The T^2-type statistic under H_{L_2} can be written as

$$T_{Lm}^2 = (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)})' \mathbf{1}_p \{\mathbf{1}_p' \widehat{\Xi} \mathbf{1}_p\}^{-1} \mathbf{1}_p' (\widehat{\boldsymbol{\mu}}^{(1)} - \widehat{\boldsymbol{\mu}}^{(2)}).$$

T_{Lm}^2 is asymptotically distributed as a χ^2 distribution with 1 degree of freedom when $N_1^{(i)}$s are large.

When we consider the case under H_{F_2}, we can join the two samples and regard it as a one-sample problem. The T^2-type statistic under H_{F_2} can be written as

$$T_{Fm}^2 = (C\widehat{\boldsymbol{\mu}})' \{C \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}] C'\}^{-1} (C\widehat{\boldsymbol{\mu}}),$$

where

$$\widehat{\boldsymbol{\mu}} = \begin{pmatrix} \widehat{\boldsymbol{\mu}}_1 \\ \widehat{\boldsymbol{\mu}}_2 \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{x}}_{1T} \\ \overline{\boldsymbol{x}}_{2F} - \widehat{\Sigma}_{21}\widehat{\Sigma}_{11}^{-1}(\overline{\boldsymbol{x}}_{1F} - \overline{\boldsymbol{x}}_{1T}) \end{pmatrix},$$

$$\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}] = \begin{pmatrix} \frac{1}{N}\widehat{\Sigma}_{11} & \frac{1}{N}\widehat{\Sigma}_{12} \\ \frac{1}{N}\widehat{\Sigma}_{21} & \widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2] \end{pmatrix},$$

$$\widehat{\mathrm{Cov}}[\widehat{\boldsymbol{\mu}}_2] = \frac{1}{N_1}\left(\widehat{\Sigma}_{22} - \frac{N_2}{N}\widehat{\Sigma}_{21}\widehat{\Sigma}_{11}^{-1}\widehat{\Sigma}_{12}\right) + \frac{N_2 p_1}{NN_1(N_1 - p_1 - 2)}\widehat{\Sigma}_{22\cdot 1}$$

and

$$\overline{\boldsymbol{x}}_{1T} = \frac{1}{N}\sum_{i=1}^{2}\sum_{j=1}^{N^{(i)}}\boldsymbol{x}_{1j}^{(i)}, \quad \overline{\boldsymbol{x}}_{1F} = \frac{1}{N_1}\sum_{i=1}^{2}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{x}_{1j}^{(i)}, \quad \overline{\boldsymbol{x}}_{2F} = \frac{1}{N_1}\sum_{i=1}^{2}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{x}_{2j}^{(i)},$$

$$N_2 = \sum_{i=1}^{k} N_2^{(i)}.$$

These estimators are extended for the MLEs obtained by [4]. T_{Fm}^2 is asymptotically distributed as a χ^2 distribution with $p-1$ degrees of freedom when $N_1^{(i)}$s are large.

However, the upper percentiles of the χ^2 distribution are not a good approximation for the T^2-type statistic when the sample size is small, and it is difficult to obtain the exact upper percentiles of these statistics when the data have missing observations. Hence, we give the approximate upper percentiles based on the idea of [10] where it is assumed that the true upper percentiles exist between $T^2_{p-1,N_1-p,\alpha}$ and $T^2_{p-1,N-p,\alpha}$. $F^*_{1,\alpha}$ can give the approximate upper percentiles of T_{Pm} and T_{Fm}.

$$F^*_{1,\alpha} = T^2_{p-1,N_1-p,\alpha} - \frac{Np - N_2 p_2}{Np}\left(T^2_{p-1,N_1-p,\alpha} - T^2_{p-1,N-p,\alpha}\right),$$

where

$$T^2_{p-1,N-p,\alpha} = \frac{(N-2)(p-1)}{N-p}F_{p-1,N-p,\alpha},$$

$$T^2_{p-1,N_1-p,\alpha} = \frac{(N_1-2)(p-1)}{N_1-p}F_{p-1,N_1-p,\alpha},$$

and $F_{p,q,\alpha}$ is the upper 100α percentile of F distribution with p and q degrees of freedom. Further, $F^*_{2,\alpha}$ can give the approximate upper percentiles of T_{Lm}.

$$F^*_{2,\alpha} = T^2_{1,N_1-2,\alpha} - \frac{Np - N_2 p_2}{Np}(T^2_{1,N_1-2,\alpha} - T^2_{1,N-2,\alpha}),$$

where

$$\begin{aligned} T^2_{1,N-2,\alpha} &= F_{1,N-2,\alpha}, \\ T^2_{1,N_1-2,\alpha} &= F_{1,N_1-2,\alpha}. \end{aligned}$$

5. Parallelism hypothesis for several groups with two-step monotone missing data

We have two-step monotone missing data when $k \geq 3$, as in Section 3. First, we transform the observation vectors using C. Then we have

$$\begin{aligned} \boldsymbol{u}^{(i)}_j &= C\boldsymbol{x}^{(i)}_j \sim N_{p-1}(\boldsymbol{\theta}^{(i)}, \Gamma), \\ \boldsymbol{u}^{(i)}_{1j} &= C_1\boldsymbol{x}^{(i)}_{1j} \sim N_{p_1-1}(\boldsymbol{\theta}^{(i)}_1, \Gamma_{11}), \end{aligned}$$

where $\boldsymbol{\theta}^{(i)} = C\boldsymbol{\mu}^{(i)},\ \Gamma = C\Sigma C'$, and C_1 is a $(p_1-1)\times p_1$ matrix of rank (p_1-1) such that $C_1\mathbf{1}_{p_1} = \mathbf{0}$ and $\mathbf{1}_{p_1}$ is a p_1-vector of ones.

$$\boldsymbol{\theta}^{(i)} = \begin{pmatrix} \boldsymbol{\theta}^{(i)}_1 \\ \boldsymbol{\theta}^{(i)}_2 \end{pmatrix},\ \Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}.$$

$\boldsymbol{\theta}^{(i)}$ and Γ are partitioned according to the blocks of the data set. It should be noted that $\boldsymbol{\theta}_1 : (p_1-1)\times 1,\ \boldsymbol{\theta}_2 : p_2\times 1,\ \Gamma_{11} : (p_1-1)\times(p_1-1),\ \Gamma_{12} = \Gamma'_{21} : (p_1-1)\times p_2$, and $\Gamma_{22} : p_2\times p_2$. To construct a likelihood ratio, we obtain the MLEs of $\boldsymbol{\theta}^{(i)}$ and Γ in general and under the hypothesis H_{P_k}. These can be obtained in the same way as earlier:

$$\begin{aligned} \widehat{\boldsymbol{\theta}}^{(i)} &= \begin{pmatrix} \widehat{\boldsymbol{\theta}}^{(i)}_1 \\ \widehat{\boldsymbol{\theta}}^{(i)}_2 \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{u}}^{(i)}_{1T} \\ \overline{\boldsymbol{u}}^{(i)}_{2F} - \widehat{\Phi}_{21}(\overline{\boldsymbol{u}}^{(i)}_{1F} - \overline{\boldsymbol{u}}^{(i)}_{1T}) \end{pmatrix}, \\ \widehat{\Gamma} &= \begin{pmatrix} \widehat{\Gamma}_{11} & \widehat{\Gamma}_{12} \\ \widehat{\Gamma}_{21} & \widehat{\Gamma}_{22} \end{pmatrix} = \begin{pmatrix} \widehat{\Phi}_{11} & \widehat{\Phi}_{11}\widehat{\Phi}_{12} \\ \widehat{\Phi}_{21}\widehat{\Phi}_{11} & \widehat{\Phi}_{22} + \widehat{\Phi}_{21}\widehat{\Phi}_{11}\widehat{\Phi}_{12} \end{pmatrix}, \end{aligned}$$

where

$$\overline{\boldsymbol{u}}_{1T}^{(i)} = \frac{1}{N^{(i)}}\sum_{j=1}^{N^{(i)}}\boldsymbol{u}_{1j}^{(i)}, \quad \overline{\boldsymbol{u}}_{1F}^{(i)} = \frac{1}{N_1^{(i)}}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{u}_{1j}^{(i)}, \quad \overline{\boldsymbol{u}}_{2F}^{(i)} = \frac{1}{N_1^{(i)}}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{u}_{2j}^{(i)},$$

and

$$\begin{aligned}
\widehat{\Phi}_{11} &= \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N^{(i)}}(\boldsymbol{u}_{1j}^{(i)} - \overline{\boldsymbol{u}}_{1T}^{(i)})(\boldsymbol{u}_{1j}^{(i)} - \overline{\boldsymbol{u}}_{1T}^{(i)})', \\
\widehat{\Phi}_{21} &= \left[\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{y}_{2j}^{(i)}\boldsymbol{y}_{1j}'^{(i)}\right]\left[\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{y}_{1j}^{(i)}\boldsymbol{y}_{1j}'^{(i)}\right]^{-1}, \\
\widehat{\Phi}_{22} &= \frac{1}{N_1}\left\{\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{y}_{2j}^{(i)}\boldsymbol{y}_{2j}'^{(i)}\right. \\
&\quad \left. - \left[\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{y}_{2j}^{(i)}\boldsymbol{y}_{1j}'^{(i)}\right]\left[\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{y}_{1j}^{(i)}\boldsymbol{y}_{1j}'^{(i)}\right]^{-1}\left[\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{y}_{1j}^{(i)}\boldsymbol{y}_{2j}'^{(i)}\right]\right\}, \\
&\quad \boldsymbol{y}_{1j}^{(i)} = \boldsymbol{u}_{1j}^{(i)} - \overline{\boldsymbol{u}}_{1F}^{(i)}, \quad \boldsymbol{y}_{2j}^{(i)} = \boldsymbol{u}_{2j}^{(i)} - \overline{\boldsymbol{u}}_{2F}^{(i)}.
\end{aligned}$$

Similarly, the MLEs of $\boldsymbol{\theta}$ and Γ under H_{P_k} are as follows:

$$\begin{aligned}
\widetilde{\boldsymbol{\theta}} &= \begin{pmatrix} \widetilde{\boldsymbol{\theta}}_1 \\ \widetilde{\boldsymbol{\theta}}_2 \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{u}}_{1T} \\ \overline{\boldsymbol{u}}_{2F} - \widetilde{\Phi}_{21}(\overline{\boldsymbol{u}}_{1F} - \overline{\boldsymbol{u}}_{1T}) \end{pmatrix}, \\
\widetilde{\Gamma} &= \begin{pmatrix} \widetilde{\Gamma}_{11} & \widetilde{\Gamma}_{12} \\ \widetilde{\Gamma}_{21} & \widetilde{\Gamma}_{22} \end{pmatrix} = \begin{pmatrix} \widetilde{\Phi}_{11} & \widetilde{\Phi}_{11}\widetilde{\Phi}_{12} \\ \widetilde{\Phi}_{21}\widetilde{\Phi}_{11} & \widetilde{\Phi}_{22} + \widetilde{\Phi}_{21}\widetilde{\Phi}_{11}\widetilde{\Phi}_{12} \end{pmatrix},
\end{aligned}$$

where

$$\overline{\boldsymbol{u}}_{1T} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N^{(i)}}\boldsymbol{u}_{1j}^{(i)}, \quad \overline{\boldsymbol{u}}_{1F} = \frac{1}{N_1}\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{u}_{1j}^{(i)}, \quad \overline{\boldsymbol{u}}_{2F} = \frac{1}{N_1}\sum_{i=1}^{k}\sum_{j=1}^{N_1^{(i)}}\boldsymbol{u}_{2j}^{(i)},$$

and

$$\widetilde{\Phi}_{11} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N^{(i)}}(\boldsymbol{u}_{1j}^{(i)} - \overline{\boldsymbol{u}}_{1T})(\boldsymbol{u}_{1j}^{(i)} - \overline{\boldsymbol{u}}_{1T})',$$

$$
\begin{aligned}
\widetilde{\Phi}_{21} &= \left[\sum_{i=1}^{k}\sum_{j=1}^{N^{(i)}} \boldsymbol{w}_{2j}^{(i)}\boldsymbol{w}_{1j}^{\prime(i)}\right]\left[\sum_{i=1}^{k}\sum_{j=1}^{N^{(i)}} \boldsymbol{w}_{1j}^{(i)}\boldsymbol{w}_{1j}^{\prime(i)}\right]^{-1}, \\
\widetilde{\Phi}_{22} &= \frac{1}{N_1}\sum_{i=1}^{k}\left\{\sum_{j=1}^{N^{(i)}} \boldsymbol{w}_{2j}^{(i)}\boldsymbol{w}_{2j}^{\prime(i)}\right. \\
&\quad \left. - \left[\sum_{j=1}^{N^{(i)}} \boldsymbol{w}_{2j}^{(i)}\boldsymbol{w}_{1j}^{\prime(i)}\right]\left[\sum_{j=1}^{N^{(i)}} \boldsymbol{w}_{1j}^{(i)}\boldsymbol{w}_{1j}^{\prime(i)}\right]^{-1}\left[\sum_{j=1}^{N^{(i)}} \boldsymbol{w}_{1j}^{(i)}\boldsymbol{w}_{2j}^{\prime(i)}\right]\right\}, \\
&\boldsymbol{w}_{1j}^{(i)} = \boldsymbol{u}_{1j}^{(i)} - \overline{\boldsymbol{u}}_{1F}, \quad \boldsymbol{w}_{2j}^{(i)} = \boldsymbol{u}_{2j}^{(i)} - \overline{\boldsymbol{u}}_{2F}.
\end{aligned}
$$

We have a likelihood ratio for the parallelism hypothesis as follows:

$$
\Lambda_m = \prod_{i=1}^{k} \frac{L(\widetilde{\boldsymbol{\theta}}_1^{(i)}, \widetilde{\boldsymbol{\theta}}_2^{(i)}, \widetilde{\Gamma})}{L(\widehat{\boldsymbol{\theta}}_1^{(i)}, \widehat{\boldsymbol{\theta}}_2^{(i)}, \widehat{\Gamma})} = \frac{|\widehat{\Gamma}^*|^{\frac{1}{2}N_1}}{|\widetilde{\Gamma}^*|^{\frac{1}{2}N_1}} \times \frac{|\widehat{\Gamma}_{11}|^{\frac{1}{2}N_2}}{|\widetilde{\Gamma}_{11}|^{\frac{1}{2}N_2}},
$$

where

$$
\widehat{\Gamma}^* = \begin{pmatrix} \widehat{\Gamma}_{11} & O \\ O & \widehat{\Gamma}_{22} - \widehat{\Gamma}_{21}\widehat{\Gamma}_{11}^{-1}\widehat{\Gamma}_{12} \end{pmatrix}, \quad \widetilde{\Gamma}^* = \begin{pmatrix} \widetilde{\Gamma}_{11} & O \\ O & \widetilde{\Gamma}_{22} - \widetilde{\Gamma}_{21}\widetilde{\Gamma}_{11}^{-1}\widetilde{\Gamma}_{12} \end{pmatrix}.
$$

Then the likelihood ratio statistic $-2\log\Lambda_m$ is asymptotically distributed as a χ^2 distribution with $(p-1)(k-1)$ degrees of freedom as $N_1^{(i)}$'s tend to infinity. Hence, we reject H_{P_k} when $-2\log\Lambda_m > \chi^2_{(p-1)(k-1),\alpha}$. However, it is difficult to obtain the modified likelihood ratio statistic directly when the data have missing observations. As such, much like in the two-sample case, we use ρ_m that improves convergence to a χ^2 distribution, and put it into the test statistic:

$$
\rho_m = \left\{\frac{1}{\rho_{c_1}} - \frac{Np - N_2p_2}{Np}\left(\frac{1}{\rho_{c_1}} - \frac{1}{\rho_{c_2}}\right)\right\}^{-1},
$$

where

$$
\begin{aligned}
\rho_{c_1} &= 1 - \frac{1}{2N_1}(p+k+1), \\
\rho_{c_2} &= 1 - \frac{1}{2N}(p+k+1)
\end{aligned}
$$

and $\rho_{c_1}, \rho_{c_2} \neq 0$. Then we reject H_{P_k} when $-2\rho_m\log\Lambda_m > \chi^2_{(p-1)(k-1),\alpha}$.

6. Simulation studies

In this section, we examine the accuracy of the approximations of the proposed test statistics. The Monte Carlo simulation for the upper percentiles of the T^2-type statistics and the likelihood ratio test statistic is implemented for selected values of the parameters. The settings of the parameters α, p $(= p_1 + p_2)$, and M $(= M_1 + M_2)$ for the simulation are as follows:

$$\begin{aligned} k &= 2, 3, 6, \\ \alpha &= 0.05, \\ (p_1, p_2) &= (2,2),\ (3,1),\ (2,6),\ (6,2), \\ (M_1, M_2) &= (10,10),\ (20,10),\ (50,10),\ (100,10), \\ &\quad\ (10,100),\ (20,100),\ (50,100),\ (100,100), \end{aligned}$$

where $M_j = N_j^{(i)}$ $(j = 1, 2)$. Further, we compare their type I error rates. As a numerical experiment, we carry out 1,000,000 replications. It should be noted that our results may be applicable to the case where the sample size differs for each population. However, for simplicity, we show the results under the same sample size.

Tables 1–3 list the percentiles of the T^2-type statistics and the values of F_1^* and F_2^*. They also list the results for the comparison of the type I error rates under the T^2-type statistics when the null hypothesis is rejected, using F_1^*, F_2^*, and a χ^2 distribution. The T^2-type statistics are closer to the χ^2 distribution when the sample size is large. Comparing the type I error rates, we have that $F_{1,\alpha}^*$ and $F_{2,\alpha}^*$ seem to be closer to 0.05 than the percentiles of the χ^2 distribution especially when the sample size is small. The value tends to be closer to 0.05 under the level hypothesis than under the parallelism hypothesis and the flatness hypothesis.

Tables 4 and 5, which are compare $-2\log\Lambda_m$ and $-2\rho_m\log\Lambda_m$, list the percentiles and type I error rates using a χ^2 distribution. $-2\log\Lambda_m$ and $-2\rho_m\log\Lambda_m$ are close to the χ^2 distribution when the sample size is large. Furthermore, $-2\rho_m\log\Lambda_m$ is closer to the χ^2 distribution than $-2\log\Lambda_m$.

7. Conclusions

We discussed profile analysis when the observations have two-step monotone missing data. In Section 3, we first derived the MLEs of several groups. In Section 4, we constructed the T^2-type statistics under the three hypotheses for a two-sample problem using the MLEs given in Section 3. We gave the likelihood ratio test statistic under the parallelism hypothesis for several groups in Section 5. Finally, we performed a Monte Carlo simulation for the type I error rates in Section 6. As a result, we confirmed that $F_{1,\alpha}^*$ and $F_{2,\alpha}^*$ are better approximations than the upper percentiles of a χ^2 distribution. We confirm that both $-2\log\Lambda_m$ and

$-2\rho_m \log \Lambda_m$ are closer to the χ^2 distribution as the sample size becomes large. We can also see that $-2\rho_m \log \Lambda_m$ is always closer to the χ^2 distribution than $-2 \log \Lambda_m$ for any sample size. Therefore, we confirm that convergence to the asymptotic χ^2 distribution is improved by inputting ρ_m into the likelihood ratio statistic $-2 \log \Lambda_m$.

Table 1. Upper percentiles and type I error rates of T^2_{Pm} and F^*_1 values.

						percentile		type I error rate	
p	p_1	p_2	M	M_1	M_2	T^2_{Pm}	F^*_1	T^2_{Pm}	χ^2
4	2	2	20	10	10	9.671	9.540	0.052	0.089
$\chi^2_{3,0.05} = 7.815$			30	20	10	8.750	8.684	0.051	0.071
			60	50	10	8.212	8.194	0.050	0.059
			110	100	10	8.001	8.014	0.050	0.054
			110	10	100	9.176	9.339	0.047	0.078
			120	20	100	7.996	8.446	0.050	0.064
			150	50	100	8.075	8.061	0.050	0.056
			200	100	100	7.974	7.950	0.051	0.054
	3	1	20	10	10	9.198	9.308	0.048	0.080
			30	20	10	8.664	8.644	0.050	0.069
			60	50	10	8.182	8.191	0.050	0.058
			110	100	10	8.020	8.013	0.050	0.055
			110	10	100	8.261	8.676	0.042	0.060
			120	20	100	8.137	8.221	0.048	0.057
			150	50	100	7.987	8.010	0.050	0.054
			200	100	100	7.953	7.936	0.050	0.053
8	6	2	20	10	10	18.184	20.645	0.030	0.120
$\chi^2_{7,0.05} = 14.067$			30	20	10	17.200	17.288	0.049	0.108
			60	50	10	15.465	15.444	0.050	0.076
			110	100	10	14.787	14.779	0.050	0.063
			110	10	100	14.195	18.371	0.014	0.052
			120	20	100	15.011	15.655	0.041	0.067
			150	50	100	14.774	14.685	0.049	0.061
			200	100	100	14.498	14.499	0.050	0.058
	2	6	20	10	10	26.607	23.487	0.073	0.251
			30	20	10	18.428	17.640	0.060	0.131
			60	50	10	15.624	15.470	0.052	0.078
			110	100	10	14.811	14.783	0.050	0.063
			110	10	100	25.615	25.559	0.050	0.234
			120	20	100	17.715	17.534	0.052	0.117
			150	50	100	15.306	15.160	0.052	0.072
			200	100	100	14.695	14.600	0.052	0.061

Table 2. Upper percentiles and type I error rates of T^2_{Lm} and F^*_2 values.

						percentile		type I error rate	
p	p_1	p_2	M	M_1	M_2	T^2_{Lm}	F^*_1	T^2_{Lm}	χ^2
4	2	2	20	10	10	4.048	7.322	7.322	0.048
$\chi^2_{1,0.05} = 3.841$			30	20	10	3.999	4.014	7.094	0.055
			60	50	10	3.925	3.922	6.857	0.050
			110	100	10	3.880	3.885	6.733	0.050
			110	10	100	3.990	4.005	7.399	0.050
			120	20	100	3.950	3.926	6.973	0.051
			150	50	100	3.883	3.884	6.749	0.050
			200	100	100	3.871	3.868	6.686	0.050
	3	1	20	10	10	4.175	4.177	7.699	0.050
			30	20	10	4.023	4.022	7.178	0.050
			60	50	10	3.916	3.923	6.845	0.050
			110	100	10	3.880	3.885	6.745	0.050
			110	10	100	4.218	4.125	7.923	0.052
			120	20	100	4.025	3.971	7.198	0.051
			150	50	100	3.898	3.895	6.802	0.050
			200	100	100	3.879	3.871	6.728	0.050
8	6	2	20	10	10	3.433	4.138	6.242	0.032
$\chi^2_{1,0.05} = 3.841$			30	20	10	3.976	4.014	7.075	0.049
			60	50	10	3.928	3.922	6.847	0.050
			110	100	10	3.886	3.885	6.729	0.050
			110	10	100	2.840	4.005	5.510	0.024
			120	20	100	3.859	3.926	6.863	0.048
			150	50	100	3.890	3.884	6.760	0.050
			200	100	100	3.867	3.868	6.698	0.050
	2	6	20	10	10	4.239	4.217	7.860	0.051
			30	20	10	4.065	4.030	7.258	0.051
			60	50	10	3.942	3.924	6.893	0.051
			110	100	10	3.884	3.885	6.758	0.050
			110	10	100	4.264	4.245	8.113	0.050
			120	20	100	4.070	4.017	7.291	0.051
			150	50	100	3.917	3.905	6.877	0.050
			200	100	100	3.865	3.874	6.738	0.050

Table 3. Upper percentiles and type I error rates of T^2_{Fm} and F^*_1 values.

						percentile		type I error rate	
p	p_1	p_2	M	M_1	M_2	T^2_{Fm}	F^*_1	T^2_{Fm}	χ^2
4	2	2	20	10	10	10.699	9.540	0.069	0.112
$\chi^2_{3,0.05} = 7.815$			30	20	10	9.072	8.684	0.057	0.078
			60	50	10	8.301	8.194	0.052	0.061
			110	100	10	8.065	8.014	0.051	0.056
			110	10	100	10.672	9.339	0.072	0.112
			120	20	100	8.898	8.446	0.059	0.074
			150	50	100	8.212	8.061	0.053	0.059
			200	100	100	8.017	7.950	0.051	0.055
	3	1	20	10	10	10.071	9.308	0.100	0.100
			30	20	10	8.913	8.644	0.055	0.075
			60	50	10	8.294	8.191	0.052	0.061
			110	100	10	8.060	8.013	0.051	0.055
			110	10	100	9.414	8.676	0.085	0.085
			120	20	100	8.479	8.221	0.055	0.065
			150	50	100	8.106	8.010	0.052	0.057
			200	100	100	7.980	7.936	0.051	0.054
8	6	2	20	10	10	24.303	20.645	0.084	0.230
$\chi^2_{7,0.05} = 14.067$			30	20	10	18.111	17.288	0.061	0.127
			60	50	10	15.663	15.444	0.053	0.080
			110	100	10	14.838	14.779	0.051	0.064
			110	10	100	21.011	18.371	0.077	0.168
			120	20	100	16.274	15.655	0.059	0.091
			150	50	100	14.982	14.774	0.053	0.067
			200	100	100	14.606	14.499	0.052	0.060
	2	6	20	10	10	30.222	23.487	0.103	0.314
			30	20	10	30.222	19.236	0.070	0.148
			60	50	10	15.834	15.470	0.055	0.083
			110	100	10	14.904	14.783	0.052	0.065
			110	10	100	30.757	25.559	0.086	0.324
			120	20	100	18.966	17.534	0.068	0.144
			150	50	100	15.630	15.160	0.057	0.079
			200	100	100	14.842	14.600	0.054	0.064

Table 4. Upper percentiles and type I error rates using $-2\log\Lambda_m$ and $-2\rho_m\log\Lambda_m$values for $k=3$.

						percentile		type I error rate	
p	p_1	p_2	M	M_1	M_2	LRT	modified LRT	LRT	modified LRT
4	2	2	20	10	10	14.314	13.108	0.086	0.060
$\chi^2_{6,0.05}=12.592$			30	20	10	13.437	12.789	0.066	0.054
			60	50	10	12.923	12.631	0.056	0.051
			110	100	10	12.771	12.615	0.053	0.050
			110	10	100	14.132	13.126	0.082	0.060
			120	20	100	13.306	12.840	0.064	0.055
			150	50	100	12.894	12.702	0.056	0.052
			200	100	100	12.726	12.620	0.053	0.051
	3	1	20	10	10	13.961	12.906	0.078	0.056
			30	20	10	13.287	12.671	0.064	0.051
			60	50	10	12.893	12.604	0.056	0.050
			110	100	10	12.757	12.602	0.053	0.050
			110	10	100	13.544	12.967	0.069	0.057
			120	20	100	13.051	12.747	0.059	0.053
			150	50	100	12.782	12.630	0.054	0.051
			200	100	100	12.718	12.623	0.052	0.051
8	6	2	20	10	10	28.011	24.822	0.128	0.067
$\chi^2_{14,0.05}=23.685$			30	20	10	25.836	24.039	0.084	0.055
			60	50	10	24.586	23.760	0.064	0.051
			110	100	10	24.166	23.726	0.057	0.051
			110	10	100	26.788	25.009	0.102	0.070
			120	20	100	25.089	24.204	0.071	0.057
			150	50	100	24.285	23.851	0.059	0.052
			200	100	100	24.033	23.762	0.055	0.051
	2	6	20	10	10	29.686	25.521	0.168	0.079
			30	20	10	26.312	24.333	0.094	0.059
			60	50	10	24.669	23.826	0.064	0.052
			110	100	10	24.201	23.758	0.057	0.051
			110	10	100	29.538	25.110	0.164	0.072
			120	20	100	26.219	24.371	0.092	0.060
			150	50	100	24.613	23.952	0.064	0.054
			200	100	100	24.165	23.832	0.057	0.052

Table 5. Upper percentiles and type I error rates using $-2\log\Lambda_m$ and $-2\rho_m \log\Lambda_m$values for $k = 6$.

						percentile		type I error rate	
p	p_1	p_2	M	M_1	M_2	LRT	modified LRT	LRT	modified LRT
4	2	2	20	10	10	27.213	25.642	0.085	0.059
$\chi^2_{15,0.05} = 24.996$			30	20	10	26.079	25.215	0.066	0.053
			60	50	10	25.462	25.066	0.057	0.051
			110	100	10	25.243	25.031	0.053	0.050
			110	10	100	27.213	25.642	0.085	0.059
			120	20	100	25.918	25.298	0.063	0.054
			150	50	100	25.395	25.135	0.055	0.052
			200	100	100	25.189	25.044	0.053	0.051
	3	1	20	10	10	26.759	25.372	0.077	0.055
			30	20	10	25.951	25.125	0.064	0.052
			60	50	10	25.407	25.016	0.056	0.050
			110	100	10	25.211	25.000	0.053	0.050
			110	10	100	26.158	25.410	0.067	0.056
			120	20	100	25.580	25.175	0.058	0.052
			150	50	100	25.301	25.094	0.054	0.051
			200	100	100	25.157	25.028	0.052	0.050
8	6	2	20	10	10	54.759	50.882	0.114	0.061
$\chi^2_{35,0.05} = 49.802$			30	20	10	52.432	50.154	0.080	0.053
			60	50	10	50.959	49.888	0.062	0.051
			110	100	10	50.382	49.808	0.056	0.050
			110	10	100	53.064	50.957	0.088	0.062
			120	20	100	51.423	50.305	0.068	0.055
			150	50	100	50.581	49.802	0.058	0.052
			200	100	100	50.248	49.895	0.054	0.051
	2	6	20	10	10	56.757	51.821	0.148	0.072
			30	20	10	53.077	50.585	0.088	0.058
			60	50	10	51.083	49.992	0.064	0.052
			110	100	10	50.431	49.854	0.056	0.051
			110	10	100	56.585	51.391	0.145	0.067
			120	20	100	52.916	50.608	0.086	0.058
			150	50	100	51.004	50.150	0.062	0.053
			200	100	100	50.427	49.993	0.056	0.052

Acknowledgements

The research of the third author was supported in part by Grant-in-Aid for Scientific Research (C) (23500360).

References

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

[2] W.-Y. Chang and D.St.P. Richards, *Finite-sample inference with monotone incomplete multivariate normal data. I*, J. Multivariate Anal. **100** (2009) 1883–1899.

[3] A.P. Dempster, N.M. Laird and D.B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, J.R. Stat. Soc. Ser. B Stat. Methodol. **39** (1977) 1–38.

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

[5] K. Koizumi and T. Seo, *Testing equality of two mean vectors and simultaneous confidence intervals in repeated measures with missing data*, J. Japanese Soc. Comput. Statist. **22** (2009) 33–41.

[6] K. Koizumi and T. Seo, *Simultaneous confidence intervals among k mean vectors in repeated measures with missing data*, Amer. J. Math. Management Sci. **29** (2009) 263–275.

[7] Y. Maruyama, *Asymptotic expansions of the null distributions of some test statistics for profile analysis in general distributions*, J. Statist. Plann. Inference **137** (2007) 506–526.

[8] D.F. Morrison, Multivariate Statistical Methods, 4th ed (Duxbury, 2005).

[9] N. Okamoto, N. Miura and T. Seo, *On the distributions of some test statistics for profile analysis in elliptical populations*, Amer. J. Math. Management Sci. **26** (2006) 1–31.

[10] N. Seko, A. Yamazaki and T. Seo, *Tests for mean vector and simultaneous confidence intervals with two-step monotone missing data*, SUT J. Math. **48** (2012) 13–36.

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

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

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

[14] M.S. Srivastava, *Multivariate data with missing observations*, Comm. Statist. Theory Methods **14** (1985) 775–792.

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

[16] M.S. Srivastava, *Profile analysis of several groups*, Comm. Statist. Theory Methods **16** (1987) 909–926.

Received 14 August 2013
Revised 18 November 2013