Journal of Statistical Theory and Practice (2024) 18:3
https://doi.org/10.1007/s42519-023-00355-2

ORIGINAL ARTICLE

Simultaneous Tests for Mean Vectors and Covariance Matrices with Three-Step Monotone Missing Data

Remi Sakai[1] · Ayaka Yagi[1] · Takashi Seo[1]

Accepted: 15 November 2023 / Published online: 14 December 2023

Abstract

In this paper, we consider simultaneous tests of the mean vectors and the covariance matrices under three-step monotone missing data for a one-sample and a multi-sample problem. We provide the likelihood ratio test (LRT) statistic and propose statistics for improving the accuracy of the χ^2 approximation. These test statistics are derived by decomposing the likelihood ratio (LR) using the coefficients of the modified LRT statistics with complete data. As an alternative approach, we derive an approximate upper percentile of the LRT statistic with three-step monotone missing data using linear interpolation based on an asymptotic expansion of the LRT statistic with complete data. Finally, we investigate the asymptotic behavior of the upper percentiles of these test statistics and the accuracy of approximate upper percentiles via Monte Carlo simulation. In addition, we give an example of test statistics and approximate upper percentiles proposed in this paper.

Keywords Asymptotic expansion · Likelihood ratio test · Linear interpolation · Maximum likelihood estimator · Modified likelihood ratio test statistic

Mathematics Subject Classification 62E20 · 62H10

1 Introduction

In this paper, we consider simultaneous tests of the mean vectors and the covariance matrices under three-step monotone missing data for a one-sample and a multi-sample problem. Jinadasa and Tracy [4] and Kanda and Fujikoshi [5] discussed MLEs for general k-step monotone missing data. For simultaneous tests, the LRT statistic and the modified LRT statistic with Bartlett correction for the case of complete data were

✉ Ayaka Yagi
yagi@rs.tus.ac.jp

1 Department of Applied Mathematics, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

discussed by Muirhead [6] and Srivastava [7]. Furthermore, the LRT statistic and the test statistics for improving the accuracy of the χ^2 approximation for three-step monotone missing data were proposed by Hao and Krishnamoorthy [1] and Hosoya and Seo [2, 3]. In particular, Hosoya and Seo [2, 3] presented test statistics by decomposing the LR; this paper is an extension of the work presented by Hosoya and Seo [2, 3]. An LRT statistic and test statistics for general k-step monotone missing data, which are obtained by correcting only a part of the missing data, were given by Yagi et al. [9].

The remainder of this paper is organized as follows. Sect. 2 describes the MLEs of the mean vector and covariance matrix and its LRT statistic in the case of three-step monotone missing data for a one-sample problem. Furthermore, we propose three test statistics for improving the accuracy of the χ^2 approximation using the coefficients of the modified LRT statistics with complete data. In addition, we derive an approximate upper percentile of the LRT statistic. Using Monte Carlo simulation, we investigate the χ^2 approximation accuracy of the test statistics and the accuracy of approximate upper percentiles of the LRT statistic. Numerical power comparison of the test statistics for some selected parameters is also presented. Sect. 3 describes the test statistics and approximate upper percentile for a multi-sample problem. Furthermore, via Monte Carlo simulation, we investigate the asymptotic behavior of the upper percentiles of these test statistics and approximate upper percentiles of the LRT statistic. The results are illustrated using an example. Finally, Sect. 4 states our conclusions.

2 One-Sample Problem

In this section, we consider the problem of simultaneous test of the mean vector and the covariance matrix under three-step monotone missing data for a one-sample problem.

2.1 LR with Three-Step Monotone Missing Data

We suppose that the data is normally distributed as follows:

$$\begin{aligned} \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N_1} &\overset{i.i.d.}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \\ \boldsymbol{x}_{(12),N_1+1}, \boldsymbol{x}_{(12),N_1+2}, \ldots, \boldsymbol{x}_{(12),N_1+N_2} &\overset{i.i.d.}{\sim} N_{p_1+p_2}(\boldsymbol{\mu}_{(12)}, \boldsymbol{\Sigma}_{(12)(12)}), \\ \boldsymbol{x}_{1,N_1+N_2+1}, \boldsymbol{x}_{1,N_1+N_2+2}, \ldots, \boldsymbol{x}_{1N} &\overset{i.i.d.}{\sim} N_{p_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}), \end{aligned} \tag{1}$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{(12)} \\ \boldsymbol{\mu}_3 \end{pmatrix},$$

$$\boldsymbol{\Sigma} = \left(\begin{array}{cc|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{13} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \boldsymbol{\Sigma}_{23} \\ \hline \boldsymbol{\Sigma}_{31} & \boldsymbol{\Sigma}_{32} & \boldsymbol{\Sigma}_{33} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{(12)(12)} & \boldsymbol{\Sigma}_{(12)3} \\ \hline \boldsymbol{\Sigma}_{3(12)} & \boldsymbol{\Sigma}_{33} \end{array}\right).$$

We partition $\boldsymbol{x}_j$ into a $p_1 \times 1$ random vector, a $p_2 \times 1$ random vector, and a $p_3 \times 1$ random vector as $\boldsymbol{x}_j = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j}, \boldsymbol{x}'_{3j})'(j = 1, \ldots, N_1)$. In addition, let $\boldsymbol{x}_{(12),j} = (\boldsymbol{x}'_{1j}, \boldsymbol{x}'_{2j})'(j = N_1 + 1, \ldots, N_1 + N_2)$.

Such a dataset has three-step monotone missing data for a one-sample problem:

$$
\begin{pmatrix}
\overbrace{\quad}^{p_1} & \overbrace{\quad}^{p_2} & \overbrace{\quad}^{p_3} \\
\boldsymbol{x}'_{11} & \boldsymbol{x}'_{21} & \boldsymbol{x}'_{31} \\
\vdots & \vdots & \vdots \\
\boldsymbol{x}'_{1N_1} & \boldsymbol{x}'_{2N_1} & \boldsymbol{x}'_{3N_1} \\
\boldsymbol{x}'_{1,N_1+1} & \boldsymbol{x}'_{2,N_1+1} & * \cdots * \\
\vdots & \vdots & \vdots \quad \vdots \\
\boldsymbol{x}'_{1,N_1+N_2} & \boldsymbol{x}'_{2,N_1+N_2} & * \cdots * \\
\boldsymbol{x}'_{1,N_1+N_2+1} & * \cdots * & * \cdots * \\
\vdots & \vdots \quad \vdots & \vdots \quad \vdots \\
\boldsymbol{x}'_{1N} & * \cdots * & * \cdots *
\end{pmatrix},
$$

where $N = N_1 + N_2 + N_3$, $p = p_1 + p_2 + p_3$ and "$*$" indicates a missing observation.

Now, we consider the following hypothesis test when the dataset has a three-step monotone pattern.

$$
H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0 \ \text{vs.} \ H_1 : \text{not} \ H_0, \tag{2}
$$

where $\boldsymbol{\mu}_0$ is a known vector, and $\boldsymbol{\Sigma}_0$ is a known matrix. Without loss of generality, we can assume that $\boldsymbol{\mu} = \boldsymbol{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_p$. Then, we have the following theorem.

Theorem 1 *Suppose the data have a three-step monotone pattern of missing observations and are normally distributed as* (1). *Then, the LR of the hypothesis test* (2) *can be given by*

$$
\begin{aligned}
\lambda_1 = {} & |\widehat{\boldsymbol{\Delta}}_{11}|^{\frac{N}{2}} |\widehat{\boldsymbol{\Delta}}_{22}|^{\frac{N_1+N_2}{2}} |\widehat{\boldsymbol{\Delta}}_{33}|^{\frac{N_1}{2}} \\
& \times \frac{\operatorname{etr}\left(-\frac{1}{2}\sum_{j=1}^{N}\boldsymbol{x}_j\boldsymbol{x}'_j\right)\operatorname{etr}\left(-\frac{1}{2}\sum_{j=1}^{N_1+N_2}\boldsymbol{x}_{2j}\boldsymbol{x}'_{2j}\right)\operatorname{etr}\left(-\frac{1}{2}\sum_{j=1}^{N_1}\boldsymbol{x}_{3j}\boldsymbol{x}'_{3j}\right)}{\exp\left(-\frac{1}{2}Np_1\right)\exp\left(-\frac{1}{2}(N_1+N_2)p_2\right)\exp\left(-\frac{1}{2}N_1p_3\right)},
\end{aligned} \tag{3}
$$

where

$$
\widehat{\boldsymbol{\Delta}}_{11} = \frac{1}{N}\boldsymbol{B}, \ \ \widehat{\boldsymbol{\Delta}}_{22} = \frac{1}{N_1 + N_2}\boldsymbol{A}_{22\cdot 1}, \ \ \widehat{\boldsymbol{\Delta}}_{33} = \frac{1}{N_1}\boldsymbol{W}_{(1)33\cdot 12},
$$

and

$$
\boldsymbol{W}_{(1)} = \sum_{j=1}^{N_1}(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(1)})(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_{(1)})'
$$

$$
\begin{aligned}
&= \left(\begin{array}{cc|c} \boldsymbol{W}_{(1)11} & \boldsymbol{W}_{(1)12} & \boldsymbol{W}_{(1)13} \\ \boldsymbol{W}_{(1)21} & \boldsymbol{W}_{(1)22} & \boldsymbol{W}_{(1)23} \\ \hline \boldsymbol{W}_{(1)31} & \boldsymbol{W}_{(1)32} & \boldsymbol{W}_{(1)33} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{W}_{(1),(12)(12)} & \boldsymbol{W}_{(1),(12)3} \\ \hline \boldsymbol{W}_{(1),3(12)} & \boldsymbol{W}_{(1)33} \end{array}\right), \\
\boldsymbol{W}_{(2)} &= \sum_{j=N_1+1}^{N_1+N_2} \begin{pmatrix} \boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{(2)1} \\ \boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_{(2)2} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{(2)1} \\ \boldsymbol{x}_{2j} - \overline{\boldsymbol{x}}_{(2)2} \end{pmatrix}' \\
&\quad + \frac{N_1 N_2}{N_1 + N_2} \begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1} - \overline{\boldsymbol{x}}_{(2)1} \\ \overline{\boldsymbol{x}}_{(1)2} - \overline{\boldsymbol{x}}_{(2)2} \end{pmatrix} \begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1} - \overline{\boldsymbol{x}}_{(2)1} \\ \overline{\boldsymbol{x}}_{(1)2} - \overline{\boldsymbol{x}}_{(2)2} \end{pmatrix}' \\
&= \begin{pmatrix} \boldsymbol{W}_{(2)11} & \boldsymbol{W}_{(2)12} \\ \boldsymbol{W}_{(2)21} & \boldsymbol{W}_{(2)22} \end{pmatrix}, \\
\boldsymbol{W}_{(3)} &= \sum_{j=N_1+N_2+1}^{N} (\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{(3)})(\boldsymbol{x}_{1j} - \overline{\boldsymbol{x}}_{(3)})' \\
&\quad + \frac{(N_1+N_2)N_3}{N} \left(\overline{\boldsymbol{x}}_{(3)} - \frac{1}{N_1+N_2}(N_1 \overline{\boldsymbol{x}}_{(1)1} + N_2 \overline{\boldsymbol{x}}_{(2)1}) \right) \\
&\quad \times \left(\overline{\boldsymbol{x}}_{(3)} - \frac{1}{N_1+N_2}(N_1 \overline{\boldsymbol{x}}_{(1)1} + N_2 \overline{\boldsymbol{x}}_{(2)1}) \right)', \\
\boldsymbol{W}_{(1)33\cdot 12} &= \boldsymbol{W}_{(1)33} - \boldsymbol{W}_{(1),3(12)} \boldsymbol{W}_{(1),(12)(12)}^{-1} \boldsymbol{W}_{(1),(12)3}, \\
\boldsymbol{A} &= \boldsymbol{W}_{(1),(12)(12)} + \boldsymbol{W}_{(2)}, \quad \boldsymbol{A}_{22\cdot 1} = \boldsymbol{A}_{22} - \boldsymbol{A}_{21} \boldsymbol{A}_{11}^{-1} \boldsymbol{A}_{12}, \\
\boldsymbol{B} &= \boldsymbol{W}_{(1)11} + \boldsymbol{W}_{(2)11} + \boldsymbol{W}_{(3)}, \\
\overline{\boldsymbol{x}}_{(1)} &= \begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1} \\ \overline{\boldsymbol{x}}_{(1)2} \\ \overline{\boldsymbol{x}}_{(1)3} \end{pmatrix}, \; \overline{\boldsymbol{x}}_{(1)1} = \frac{1}{N_1} \sum_{j=1}^{N_1} \boldsymbol{x}_{1j}, \; \overline{\boldsymbol{x}}_{(1)2} = \frac{1}{N_1} \sum_{j=1}^{N_1} \boldsymbol{x}_{2j}, \\
\overline{\boldsymbol{x}}_{(1)3} &= \frac{1}{N_1} \sum_{j=1}^{N_1} \boldsymbol{x}_{3j}, \; \overline{\boldsymbol{x}}_{(2)} = \begin{pmatrix} \overline{\boldsymbol{x}}_{(2)1} \\ \overline{\boldsymbol{x}}_{(2)2} \end{pmatrix}, \; \overline{\boldsymbol{x}}_{(2)1} = \frac{1}{N_2} \sum_{j=N_1+1}^{N_1+N_2} \boldsymbol{x}_{1j}, \\
\overline{\boldsymbol{x}}_{(2)2} &= \frac{1}{N_2} \sum_{j=N_1+1}^{N_1+N_2} \boldsymbol{x}_{2j}, \; \overline{\boldsymbol{x}}_{(3)} = \frac{1}{N_3} \sum_{j=N_1+N_2+1}^{N} \boldsymbol{x}_{1j}.
\end{aligned}
$$

For the derivation of Theorem 1, see the Appendix. After calculations, we get

$$
\begin{aligned}
\lambda_1 &= \left(\frac{e}{N}\right)^{\frac{Np_1}{2}} |\boldsymbol{B}|^{\frac{N}{2}} \left(\frac{e}{N_1+N_2}\right)^{\frac{(N_1+N_2)p_2}{2}} |\boldsymbol{A}_{22\cdot 1}|^{\frac{N_1+N_2}{2}} \left(\frac{e}{N_1}\right)^{\frac{N_1 p_3}{2}} \left|\boldsymbol{W}_{(1)33\cdot 12}\right|^{\frac{N_1}{2}} \\
&\quad \times \operatorname{etr}\left\{ -\frac{1}{2}\left(\boldsymbol{B} + \frac{1}{N}(N_1 \overline{\boldsymbol{x}}_{(1)1} + N_2 \overline{\boldsymbol{x}}_{(2)1} + N_3 \overline{\boldsymbol{x}}_{(3)})(N_1 \overline{\boldsymbol{x}}_{(1)1} + N_2 \overline{\boldsymbol{x}}_{(2)1} + N_3 \overline{\boldsymbol{x}}_{(3)})' \right) \right\} \\
&\quad \times \operatorname{etr}\left\{ -\frac{1}{2}\left(\boldsymbol{A}_{22} + \frac{1}{N_1+N_2}(N_1 \overline{\boldsymbol{x}}_{(1)2} + N_2 \overline{\boldsymbol{x}}_{(2)2})(N_1 \overline{\boldsymbol{x}}_{(1)2} + N_2 \overline{\boldsymbol{x}}_{(2)2})' \right) \right\} \\
&\quad \times \operatorname{etr}\left\{ -\frac{1}{2}(\boldsymbol{W}_{(1)33} + N_1 \overline{\boldsymbol{x}}_{(1)3} \overline{\boldsymbol{x}}_{(1)3}') \right\}.
\end{aligned}
$$

Table 1 The upper percentile of $-2\log\lambda_1$ and type I error rates when $(p_1, p_2, p_3) = (3, 3, 3)$

Sample size			Upper percentile	Type I error rate
N_1	N_2	N_3	$-2\log\lambda_1$	α_1
$\alpha = 0.05$				
10	10	10	155.129	0.907
20	10	10	87.775	0.282
40	10	10	78.955	0.127
80	10	10	75.419	0.082
200	10	10	73.431	0.061
$\alpha = 0.01$				
10	10	10	187.825	0.804
20	10	10	99.009	0.115
40	10	10	88.653	0.036
80	10	10	84.693	0.019
200	10	10	82.514	0.013

$\chi^2_{f_1;0.95} = 72.15$, $\chi^2_{f_1;0.99} = 81.07$, $f_1 = 54$

This LR λ_1 is essentially the same as that obtained by Yagi, Yamaguchi, and Seo [9]. Thus, we obtain the LRT statistic $-2\log\lambda_1$. In the complete data case (Sect. 2.2), the LRT statistic for (2) is asymptotically distributed as χ^2 distribution with f_1 degrees of freedom where $f_1 = p(p+3)/2$ (see Muirhead [6, p. 370]). For example, Table 1 presents the simulated values of the upper 100α percentiles of $-2\log\lambda_1$ and Type I error rate, $\alpha_1 = \Pr\{-2\log\lambda_1 > \chi^2_{f_1;1-\alpha}\}$ for the three-step monotone missing data case, where $\chi^2_{f_1;1-\alpha}$ is the upper 100α percentile of the χ^2 distribution with f_1 degrees of freedom.

As demonstrated in Table 1, the accuracy of the χ^2 approximation in this case is not desirable when the sample size is not large; therefore, a test statistic is needed to improve the accuracy of the χ^2 approximation. We propose test statistics that improve the χ^2 approximation using the modified likelihood ratio test statistic of simultaneous test and test of variance for the complete data case described in Sect. 2.2.

2.2 Complete Data

We consider the LRT statistic and modified LRT statistics with Bartlett correction in the case of complete data for a one-sample problem. These results are used in the next subsection. We first consider a simultaneous test for complete data as follows:

$$H_{01} : \boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = \boldsymbol{I} \ \text{ vs. } \ H_{11} : \text{not } \ H_{01}$$

In this case, the LR can be expressed as follows. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ be independently distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and let λ_{S_1} be the LR for the complete data. Then, the LR is given by

$$\lambda_{S_1} = e^{\frac{Np}{2}} \left| \frac{1}{N} \boldsymbol{W} \right|^{\frac{N}{2}} \operatorname{etr}\left(-\frac{1}{2} \boldsymbol{W} \right) \exp\left(-\frac{1}{2} N \overline{\boldsymbol{x}}' \overline{\boldsymbol{x}} \right),$$

where

$$\boldsymbol{W} = \sum_{j=1}^{N} (\boldsymbol{x}_j - \overline{\boldsymbol{x}})(\boldsymbol{x}_j - \overline{\boldsymbol{x}})', \ \overline{\boldsymbol{x}} = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{x}_j.$$

It is known that $-2\log\lambda_{S_1}$ is asymptotically distributed as χ^2 distribution with $f_1(= p(p+3)/2)$ degrees of freedom. Furthermore, the modified LRT statistic with Bartlett correction can be given by $-2\rho_1 \log\lambda_{S_1}$ (Muirhead [6, p. 370]), where

$$\rho_1 = 1 - \frac{2p^2 + 9p + 11}{6N(p+3)}.$$

Next, we consider a covariance test for complete data as follows:

$$H_{02} : \boldsymbol{\Sigma} = \boldsymbol{I} \ \text{ vs. } \ H_{12} : \text{not } \ H_{02}$$

In this case, the LR, which is an unbiased test, can be expressed as follows:

$$\lambda_{V_1} = e^{\frac{(N-1)p}{2}} \left| \frac{1}{N-1} \boldsymbol{W} \right|^{\frac{N-1}{2}} \operatorname{etr}\left(-\frac{1}{2} \boldsymbol{W} \right).$$

The modified LRT statistic with Bartlett correction $-2\rho_2 \log\lambda_{V_1}$ was provided by Muirhead [6, p. 359], where

$$\rho_2 = 1 - \frac{2p^2 + 3p - 1}{6(N-1)(p+1)}.$$

2.3 Test Statistics

We now decompose the LR to derive the test statistic for improving the accuracy of the χ^2 approximation. Let

$$\omega_1 = \exp\left(-\frac{1}{2N} (N_1 \overline{\boldsymbol{x}}_{(1)1} + N_2 \overline{\boldsymbol{x}}_{(2)1} + N_3 \overline{\boldsymbol{x}}_{(3)})' (N_1 \overline{\boldsymbol{x}}_{(1)1} + N_2 \overline{\boldsymbol{x}}_{(2)1} + N_3 \overline{\boldsymbol{x}}_{(3)}) \right),$$

$$\omega_2 = \exp\left(-\frac{1}{2(N_1 + N_2)} (N_1 \overline{\boldsymbol{x}}_{(1)2} + N_2 \overline{\boldsymbol{x}}_{(2)2})' (N_1 \overline{\boldsymbol{x}}_{(1)2} + N_2 \overline{\boldsymbol{x}}_{(2)2}) \right),$$

$$\omega_3 = \exp\left(-\frac{N_1}{2} \overline{\boldsymbol{x}}'_{(1)3} \overline{\boldsymbol{x}}_{(1)3} \right),$$

$$\omega_4 = \left(\frac{e}{N} \right)^{\frac{1}{2} N p_1} |\boldsymbol{B}|^{\frac{N}{2}} \operatorname{etr}\left(-\frac{1}{2} \boldsymbol{B} \right),$$

$$\omega_5 = \left(\frac{e}{N_1 + N_2}\right)^{\frac{1}{2}(N_1+N_2)p_2} |\boldsymbol{A}_{22\cdot1}|^{\frac{N_1+N_2}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{A}_{22\cdot1}\right),$$
$$\omega_6 = \left(\frac{e}{N_1}\right)^{\frac{1}{2}N_1 p_3} \left|\boldsymbol{W}_{(1)33\cdot12}\right|^{\frac{N_1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{W}_{(1)33\cdot12}\right),$$
$$\omega_7 = \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\right) \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{W}_{(1),3(12)}\boldsymbol{W}_{(1),(12)(12)}^{-1}\boldsymbol{W}_{(1),(12)3}\right).$$

Therefore,

$$\lambda_1 = \prod_{i=1}^{7} \omega_i.$$

Then, $\omega_1\omega_4, \omega_2\omega_5, \omega_3\omega_6$ are of the form of LR for H_{01} under non-missing normality. Hence, we can obtain the modified LRT statistics, $-2\rho_{14}\log\omega_1\omega_4$, $-2\rho_{25}\log\omega_2\omega_5$, $-2\rho_{36}\log\omega_3\omega_6$, where

$$\rho_{14} = 1 - \frac{2p_1^2 + 9p_1 + 11}{6N(p_1+3)},$$
$$\rho_{25} = 1 - \frac{2p_2^2 + 9p_2 + 11}{6(N_1+N_2)(p_2+3)},$$
$$\rho_{36} = 1 - \frac{2p_3^2 + 9p_3 + 11}{6N_1(p_3+3)}.$$

Thus, we propose a new test statistic given by $-2\log\tau_1$, where

$$\tau_1 = (\omega_1\omega_4)^{\rho_{14}}(\omega_2\omega_5)^{\rho_{25}}(\omega_3\omega_6)^{\rho_{36}}\omega_7.$$

In addition, we denote

$$\omega_4^* = \left(\frac{e}{n}\right)^{\frac{1}{2}np_1} |\boldsymbol{B}|^{\frac{n}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{B}\right),$$
$$\omega_5^* = \left(\frac{e}{n_1 + n_2}\right)^{\frac{1}{2}(n_1+n_2)p_2} |\boldsymbol{A}_{22\cdot1}|^{\frac{n_1+n_2}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{A}_{22\cdot1}\right),$$
$$\omega_6^* = \left(\frac{e}{n_1}\right)^{\frac{1}{2}n_1 p_3} \left|\boldsymbol{W}_{(1)33\cdot12}\right|^{\frac{n_1}{2}} \operatorname{etr}\left(-\frac{1}{2}\boldsymbol{W}_{(1)33\cdot12}\right),$$

where

$$n = N - 1,\ n_1 = N_1 - (p_1 + p_2) - 1,\ n_1 + n_2 = N_1 + N_2 - p_1 - 1.$$

Subsequently, since $\omega_4^*, \omega_5^*, \omega_6^*$ are of the form of LR for H_{02} under non-missing normality, we can propose the test statistic as $-2\log\phi_1$, where

$$\phi_1 = \omega_1\omega_2\omega_3(\omega_4^*)^{\rho_4^*}(\omega_5^*)^{\rho_5^*}(\omega_6^*)^{\rho_6^*}\omega_7$$

and

$$\rho_4^* = 1 - \frac{2p_1^2 + 3p_1 - 1}{6n(p_1+1)}, \quad \rho_5^* = 1 - \frac{2p_2^2 + 3p_2 - 1}{6(n_1+n_2)(p_2+1)},$$
$$\rho_6^* = 1 - \frac{2p_3^2 + 3p_3 - 1}{6n_1(p_3+1)}.$$

Now, we propose the modified LRT statistic $-2\rho_{L_1}\log\lambda_1$ via linear interpolation, where

$$\rho_{L_1} = \left\{1 - \frac{(p_1+p_2)N_2 + p_1N_3}{p(N_2+N_3)}\right\}\rho_{N_1,1} + \frac{(p_1+p_2)N_2 + p_1N_3}{p(N_2+N_3)}\rho_{N,1},$$

and

$$\rho_{N_1,1} = 1 - \frac{2p^2 + 9p + 11}{6N_1(p+3)}, \quad \rho_{N,1} = 1 - \frac{2p^2 + 9p + 11}{6N(p+3)}.$$

2.4 Asymptotic Expansion Approximation

In this subsection, we give an approximate upper percentile of $-2\log\lambda_1$ when the data have a three-step monotone pattern for a one-sample problem. The upper 100α percentile of $-2\log\lambda_{S_1}$ can be expanded as

$$q_{c_1}^*(\alpha) = \chi^2_{f_1;1-\alpha} + \frac{\nu}{N}\chi^2_{f_1;1-\alpha} + \frac{1}{N^2}\chi^2_{f_1;1-\alpha}\left\{\nu^2 + \frac{2\nu}{f_1} + \frac{2\nu}{f_1(f_1+2)}\chi^2_{f_1;1-\alpha}\right\}$$
$$+ O(N^{-2}),$$

where $\nu = (2p^2 + 9p + 11)/\{6(p+3)\}$ (Hosoya and Seo [2]) Based on linear interpolation and letting $q_1^*(\alpha)$ be the upper 100α percentile of $-2\log\lambda_1$, the following can be obtained:

$$q_1^*(\alpha) = \left\{1 - \frac{(p_1+p_2)N_2 + p_1N_3}{p(N_2+N_3)}\right\}q_{N_1,1}^*(\alpha) + \frac{(p_1+p_2)N_2 + p_1N_3}{p(N_2+N_3)}q_{N,1}^*(\alpha),$$

where

$$q_{N_1,1}^*(\alpha) = \chi^2_{f_1;1-\alpha} + \frac{\nu}{N_1}\chi^2_{f_1;1-\alpha} + \frac{1}{N_1^2}\chi^2_{f_1;1-\alpha}\left\{\nu^2 + \frac{2\nu}{f_1} + \frac{2\nu}{f_1(f_1+2)}\chi^2_{f_1;1-\alpha}\right\},$$
$$q_{N,1}^*(\alpha) = \chi^2_{f_1;1-\alpha} + \frac{\nu}{N}\chi^2_{f_1;1-\alpha} + \frac{1}{N^2}\chi^2_{f_1;1-\alpha}\left\{\nu^2 + \frac{2\nu}{f_1} + \frac{2\nu}{f_1(f_1+2)}\chi^2_{f_1;1-\alpha}\right\}.$$

2.5 Simulation Studies

We evaluate the accuracy and the asymptotic behaviors of the χ^2 approximations via Monte Carlo simulation (10^6 runs). Let

$$\begin{aligned}
\alpha_1 &= \Pr\{-2\log\lambda_1 > \chi^2_{f_1;1-\alpha}\},\\
\alpha_{\rho_{L_1}} &= \Pr\{-2\rho_{L_1}\log\lambda_1 > \chi^2_{f_1;1-\alpha}\},\\
\alpha_{\tau_1} &= \Pr\{-2\log\tau_1 > \chi^2_{f_1;1-\alpha}\},\\
\alpha_{\phi_1} &= \Pr\{-2\log\phi_1 > \chi^2_{f_1;1-\alpha}\},\\
\alpha_{q_1^*} &= \Pr\{-2\log\lambda_1 > q_1^*(\alpha)\}.
\end{aligned}$$

In Tables 2 and 3, we provide the simulated upper 100α percentiles of $-2\log\lambda_1$, $-2\rho_{L_1}\log\lambda_1$, $-2\log\tau_1$, and $-2\log\phi_1$ and the approximate upper percentiles of $-2\log\lambda_1$ ($q_1^*(\alpha)$) and the actual type I error rates α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$; $\alpha = 0.05$; and for the following cases (Case I),

$$(N_1, N_2, N_3) = \begin{cases} (t, 10, 10), & \\ (t, 20, 20), & t = 10, 20, 30, 40, 80, 200, 400, \\ (t, 50, 50), & \end{cases}$$

where (p_1, p_2, p_3) in Tables 2 and 3 are (3, 3, 3), (6, 6, 6), respectively.

Similarly, Tables 4 and 5 exhibit the results for the following cases (Case II),

$$(N_1, N_2, N_3) = \begin{cases} (t, t/2, t/2), & \\ (t, t, t), & t = 10, 20, 30, 40, 80, 200, 400, \\ (t, 2t, 2t), & \end{cases}$$

where (p_1, p_2, p_3) in Tables 4 and 5 are (3, 3, 3), (6, 6, 6) respectively.

It may be noted from the above-mentioned Tables that the simulated values are closer to the upper percentile of the χ^2 distribution when the sample size increases. In addition, it can be seen that the upper percentile of $-2\log\phi_1$ is considerably better than that of $-2\log\lambda_1$ even for small sample sizes, while the upper percentile of $-2\rho_{L_1}\log\lambda_1$ or $q_1^*(\alpha)$ is not as good as $-2\log\phi_1$.

2.6 Numerical Power

We conduct the power comparison of (I) the LR test using $-2\log\lambda_1$ given in Sect. 2.1, (II) the test using statistic $-2\log\tau_1$ given in Sect. 2.3, and (III) the test using statistic $-2\log\phi_1$ given in Sect. 2.3. Under some parameter settings, the powers of (I), (II), and (III) are compared using corresponding simulated upper 100α percentiles under the null distribution, where $\alpha = 0.05$. The simulation was executed 10^6 times using normal random vectors. When $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, the powers are computed with various values of $\delta_i = \boldsymbol{\mu}_i'\boldsymbol{\mu}_i$, $i = 1, 2, 3$. This follows the power computation for the test of a mean vector in Krishnamoorthy and Pannala [10]. On the other hand, when $\boldsymbol{\mu} = \boldsymbol{0}$,

Table 2 Simulated values for $-2\log\lambda_1$, $-2\rho_{L_1}\log\lambda_1$, $-2\log\tau_1$, and $-2\log\phi_1$ and the approximate values for $-2\log\lambda_1$ and the actual type I error rates α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ when $(p_1, p_2, p_3) = (3, 3, 3)$ for Case I

Sample size		Upper percentile					Type I error				
N_1	$N_2 = N_3$	$-2\log\lambda_1$	$-2\rho_{L_1}\log\lambda_1$	$-2\log\tau_1$	$-2\log\phi_1$	$q_1^*(\alpha)$	α_1	$\alpha_{\rho_{L_1}}$	α_{τ_1}	α_{ϕ_1}	$\alpha_{q_1^*}$
$\alpha = 0.05$											
10	10	155.00	118.54	137.73	74.01	94.23	0.907	0.598	0.834	**0.066**	0.600
20	10	87.82	76.20	84.51	72.20	83.14	0.282	0.090	0.220	**0.050**	0.090
40	10	78.92	73.12	77.49	72.14	77.87	0.127	0.058	0.107	**0.050**	0.058
80	10	75.39	72.40	74.70	72.15	75.13	0.082	0.052	0.074	**0.050**	0.052
200	10	73.43	72.19	73.15	72.15	73.39	0.061	**0.050**	0.059	**0.050**	**0.050**
400	10	72.79	72.17	72.65	72.17	72.78	0.055	**0.050**	0.054	**0.050**	**0.050**
10	20	153.15	120.74	136.58	74.02	92.21	0.895	0.619	0.822	**0.067**	0.608
20	20	86.82	76.61	83.89	72.13	81.92	0.263	0.095	0.209	**0.050**	0.094
40	20	78.56	73.37	77.29	72.15	77.29	0.122	0.061	0.105	**0.050**	0.060
80	20	75.28	72.51	74.64	72.17	74.91	0.080	0.053	0.074	**0.050**	0.053
200	20	73.47	72.28	73.20	72.20	73.34	0.062	0.051	0.059	**0.050**	0.051
400	20	72.77	72.16	72.63	72.13	72.77	0.055	**0.050**	0.054	**0.050**	**0.050**
10	50	151.79	122.58	135.71	74.04	90.67	0.883	0.634	0.810	**0.066**	0.612
20	50	85.81	76.98	83.28	72.14	80.76	0.243	0.099	0.198	**0.050**	0.095
40	50	78.03	73.60	76.98	72.17	76.55	0.115	0.063	0.101	**0.050**	0.062
80	50	75.06	72.67	74.52	72.16	74.54	0.078	0.054	0.072	**0.050**	0.054
200	50	73.41	72.33	73.17	72.20	73.23	0.061	0.051	0.059	**0.050**	0.051
400	50	72.74	72.17	72.62	72.12	72.73	0.055	**0.050**	0.054	**0.050**	**0.050**

The closest value to α from among α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ of each row is in bold. $\chi^2_{f_1;0.95} = 72.15, f_1 = 54$

Table 3 The simulated values for $-2\log\lambda_1$, $-2\rho_{L_1}\log\lambda_1$, $-2\log\tau_1$, and $-2\log\phi_1$ and the approximate values for $-2\log\lambda_1$ and the actual type I error rates α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ when $(p_1, p_2, p_3) = (6, 6, 6)$ for Case I

Sample size		Upper percentile					Type I error				
N_1	$N_2 = N_3$	$-2\log\lambda_1$	$-2\rho_{L_1}\log\lambda_1$	$-2\log\tau_1$	$-2\log\phi_1$	$q_1^*(\alpha)$	α_1	$\alpha_{\rho_{L_1}}$	α_{τ_1}	α_{ϕ_1}	$\alpha_{q_1^*}$
$\alpha = 0.05$											
20	10	405.30	306.26	372.59	226.25	291.12	0.999	0.811	0.995	**0.073**	0.830
30	10	291.64	240.96	279.34	222.52	267.81	0.829	0.198	0.717	**0.052**	0.208
40	10	267.69	231.35	259.75	222.23	256.49	0.575	0.109	0.465	**0.051**	0.112
80	10	241.96	224.22	238.59	222.12	239.57	0.217	0.061	0.179	**0.050**	0.061
200	10	229.52	222.39	228.27	222.08	229.20	0.095	0.051	0.086	**0.050**	0.052
400	10	225.74	222.15	225.12	222.08	225.66	0.070	**0.050**	0.066	**0.050**	**0.050**
20	20	399.33	312.59	368.66	226.25	283.45	0.998	0.846	0.993	**0.072**	0.847
30	20	288.20	243.49	277.07	222.42	262.75	0.799	0.228	0.692	**0.052**	0.229
40	20	265.49	233.05	258.33	222.13	252.90	0.544	0.123	0.443	**0.050**	0.124
80	20	241.18	224.81	238.07	222.02	238.22	0.209	0.064	0.174	**0.050**	0.064
200	20	229.51	222.66	228.31	222.14	228.91	0.095	0.053	0.086	**0.050**	0.053
400	20	225.77	222.26	225.16	222.14	225.58	0.070	0.051	0.066	**0.050**	0.051
20	50	393.38	318.62	364.58	226.12	276.44	0.997	0.872	0.990	**0.072**	0.858
30	50	283.88	245.94	274.18	222.46	257.29	0.755	0.256	0.655	**0.052**	0.245
40	50	262.41	234.93	256.35	222.17	248.53	0.498	0.139	0.412	**0.050**	0.135
80	50	240.07	225.95	237.42	222.14	236.02	0.195	0.071	0.166	**0.050**	0.071
200	50	229.13	222.91	228.05	222.06	228.28	0.092	0.054	0.084	**0.050**	0.054
400	50	225.63	222.32	225.07	222.12	225.38	0.069	0.051	0.066	**0.050**	0.051

The closest to α from among α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ of each row is in bold. $\chi^2_{f_1;0.95} = 222.08$, $f_1 = 189$

Table 4 The simulated values for $-2\log\lambda_1$, $-2\rho_{L_1}\log\lambda_1$, $-2\log\tau_1$, and $-2\log\phi_1$ and the approximate values for $-2\log\lambda_1$ and the actual type I error rates α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ when $(p_1, p_2, p_3) = (3, 3, 3)$ for Case II

Sample size		Upper percentile					Type I error				
N_1	$N_2 = N_3$	$-2\log\lambda_1$	$-2\rho_{L_1}\log\lambda_1$	$-2\log\tau_1$	$-2\log\phi_1$	$q_1^*(\alpha)$	α_1	$\alpha_{\rho_{L_1}}$	α_{τ_1}	α_{ϕ_1}	$\alpha_{q_1^*}$
$\alpha = 0.05$											
10	5	157.03	115.48	138.95	74.07	96.99	0.922	0.569	0.849	**0.067**	0.587
20	10	87.78	76.17	84.48	72.16	83.14	0.282	0.089	0.220	**0.050**	0.090
40	20	78.58	73.38	77.31	72.17	77.29	0.122	0.061	0.105	**0.050**	0.061
80	40	75.13	72.64	74.56	72.17	74.63	0.079	0.054	0.072	**0.050**	0.054
200	100	73.32	72.35	73.11	72.18	73.12	0.060	0.052	0.058	**0.050**	0.052
400	200	72.71	72.23	72.60	72.14	72.63	0.055	0.051	0.054	**0.050**	0.051
10	10	155.06	118.59	137.81	74.03	94.23	0.908	0.598	0.834	**0.066**	0.600
20	20	86.86	76.65	83.94	72.19	81.92	0.263	0.095	0.209	**0.050**	0.094
40	40	78.14	73.55	77.04	72.15	76.71	0.116	0.062	0.101	**0.050**	0.062
80	80	74.90	72.70	74.41	72.13	74.35	0.076	0.055	0.071	**0.050**	0.055
200	200	73.22	72.36	73.04	72.14	73.01	0.059	0.052	0.058	**0.050**	0.052
400	400	72.71	72.29	72.63	72.21	72.58	0.055	0.051	0.054	**0.050**	0.051
10	20	153.27	120.83	136.66	74.03	92.21	0.894	0.619	0.821	**0.066**	0.608
20	40	86.08	76.97	83.45	72.24	80.98	0.247	0.099	0.200	**0.051**	0.095
40	80	77.75	73.64	76.80	72.18	76.27	0.111	0.063	0.099	**0.050**	0.062
80	160	74.72	72.75	74.31	72.14	74.14	0.074	0.055	0.070	**0.050**	0.055
200	400	73.17	72.40	73.02	72.19	72.93	0.059	0.052	0.057	**0.050**	0.052
400	800	72.66	72.27	72.58	72.18	72.54	0.054	0.051	0.054	**0.050**	0.051

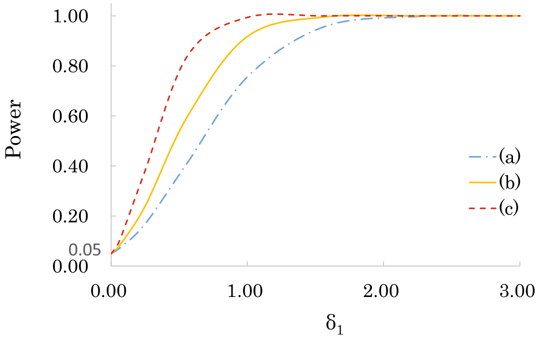
The closest to α from among α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ of each row is in bold. $\chi^2_{f_1;0.95} = 72.15$, $f_1 = 54$

Table 5 The simulated values for $-2\log\lambda_1$, $-2\rho_{L_1}\log\lambda_1$, $-2\log\tau_1$, and $-2\log\phi_1$ and the approximate values for $-2\log\lambda_1$ and the actual type I error rates α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ when $(p_1, p_2, p_3) = (6, 6, 6)$ for Case II

Sample size		Upper percentile					Type I error				
N_1	$N_2 = N_3$	$-2\log\lambda_1$	$-2\rho_{L_1}\log\lambda_1$	$-2\log\tau_1$	$-2\log\phi_1$	$q_1^*(\alpha)$	α_1	$\alpha_{\rho_{L_1}}$	α_{τ_1}	α_{ϕ_1}	$\alpha_{q_1^*}$
$\alpha = 0.05$											
20	10	405.45	306.38	372.68	226.18	291.12	0.999	0.810	0.995	**0.072**	0.830
30	15	289.65	242.47	278.02	222.42	264.82	0.812	0.216	0.704	**0.052**	0.221
40	20	265.59	233.15	258.43	222.24	252.90	0.545	0.124	0.445	**0.051**	0.124
80	40	240.33	225.64	237.56	222.16	236.57	0.199	0.069	0.168	**0.050**	0.069
200	100	228.84	223.25	227.88	222.14	227.65	0.090	0.056	0.084	**0.050**	0.056
400	200	225.40	222.65	224.94	222.10	224.83	0.068	0.053	0.065	**0.050**	0.053
20	20	399.21	312.50	368.57	226.07	283.45	0.998	0.846	0.993	**0.072**	0.848
30	30	286.15	244.72	275.73	222.36	260.07	0.780	0.243	0.675	**0.051**	0.239
40	40	263.08	234.51	256.72	222.12	249.48	0.510	0.136	0.421	**0.050**	0.133
80	80	239.23	226.24	236.84	222.10	234.96	0.186	0.073	0.160	**0.050**	0.072
200	200	228.33	223.37	227.51	222.04	227.03	0.087	0.056	0.081	**0.050**	0.056
400	400	225.12	222.67	224.73	222.03	224.52	0.066	0.053	0.064	**0.050**	0.053
20	40	394.60	317.47	365.43	226.21	277.79	0.998	0.868	0.991	**0.072**	0.857
30	60	283.19	246.28	273.71	222.39	256.48	0.747	0.260	0.647	**0.052**	0.247
40	80	261.00	235.49	255.38	222.25	246.86	0.477	0.144	0.398	**0.051**	0.138
80	160	238.19	226.55	236.12	222.14	233.70	0.175	0.075	0.153	**0.050**	0.073
200	400	228.01	223.55	227.30	222.12	226.54	0.084	0.057	0.080	**0.050**	0.057
400	800	224.99	222.79	224.66	222.15	224.28	0.065	0.053	0.063	**0.050**	0.053

The closest to α from among α_1, $\alpha_{\rho_{L_1}}$, α_{τ_1}, α_{ϕ_1}, and $\alpha_{q_1^*}$ of each row is in bold. $\chi^2_{f_1;0.95} = 222.08$, $f_1 = 189$

Fig. 1 The power plots of the test using statistic $-2\log\phi_1$: (**a**) $(N_1, N_2, N_3) = (10, 10, 10)$, (**b**) $(N_1, N_2, N_3) = (20, 10, 10)$, (**c**) $(N_1, N_2, N_3) = (40, 10, 10)$



we put $\boldsymbol{\Sigma} = \boldsymbol{I}_p + (1/\sqrt{N_1})\boldsymbol{\Omega}$, where $\boldsymbol{\Omega} = \mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_p)$, the powers are computed with various values of $\omega_j = \omega$, $j = 1, 2, \ldots, p$. Table 6 shows the power of three tests where $(p_1, p_2, p_3) = (3, 3, 3)$ and $(N_1, N_2, N_3) = (20, 10, 10)$. We note from Table 6 that the power of three tests have natural power properties. In addition, comparing the power of tests (I), (II), and (III), it can be seen that test (III) has the highest power, while tests (I) and (II) have almost the same power.

Further, Fig. 1 shows the power plots of (III) the test using statistic $-2\log\phi_1$ when (a) $(N_1, N_2, N_3) = (10, 10, 10)$, (b) $(N_1, N_2, N_3) = (20, 10, 10)$ and (c) $(N_1, N_2, N_3) = (40, 10, 10)$ with $(p_1, p_2, p_3) = (3, 3, 3)$ and $\delta_2 = \delta_3 = \omega = 0$. Fig. 1 illustrates that the power is an increasing function of the sample size. The power studies are performed for other sample sizes and dimensions, and similar trends are observed. Therefore, the results are not listed here.

3 Multi-Sample Problem

In this section, we will consider simultaneous tests of the mean vector and the covariance matrix under three-step monotone missing data for a multi-sample problem.

3.1 LR with Three-Step Monotone Missing Data

Let $\boldsymbol{x}_1^{(\ell)}, \boldsymbol{x}_2^{(\ell)}, \ldots, \boldsymbol{x}_{N_1^{(\ell)}}^{(\ell)}$ be independent p-dimensional sample vectors, $\boldsymbol{x}_{(12),N_1^{(\ell)}+1}^{(\ell)}$, $\boldsymbol{x}_{(12),N_1^{(\ell)}+2}^{(\ell)}, \ldots, \boldsymbol{x}_{(12),N_1^{(\ell)}+N_2^{(\ell)}}^{(\ell)}$ be independent (p_1+p_2)-dimensional sample vectors and $\boldsymbol{x}_{1,N_1^{(\ell)}+N_2^{(\ell)}+1}^{(\ell)}, \boldsymbol{x}_{1,N_1^{(\ell)}+N_2^{(\ell)}+2}^{(\ell)}, \ldots, \boldsymbol{x}_{1N^{(\ell)}}^{(\ell)}$ be independent p_1-dimensional sample vectors from the ℓth population ($\ell = 1, \ldots, m$). We suppose that the data is normally distributed as follows:

$$\boldsymbol{x}_1^{(\ell)}, \boldsymbol{x}_2^{(\ell)}, \ldots, \boldsymbol{x}_{N_1^{(\ell)}}^{(\ell)} \overset{i.i.d.}{\sim} N_p(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)}),$$

$$\boldsymbol{x}_{(12),N_1^{(\ell)}+1}^{(\ell)}, \boldsymbol{x}_{(12),N_1^{(\ell)}+2}^{(\ell)}, \ldots, \boldsymbol{x}_{(12),N_1^{(\ell)}+N_2^{(\ell)}}^{(\ell)} \overset{i.i.d.}{\sim} N_{p_1+p_2}(\boldsymbol{\mu}_{(12)}^{(\ell)}, \boldsymbol{\Sigma}_{(12)(12)}^{(\ell)}),$$

Table 6 The power comparison of (I), (II), and (III)

δ_1	δ_2	δ_3	ω	(I)	(II)	(III)
0	0	0	0	0.050	0.050	0.050
0.0100	0	0	0	0.053	0.053	0.054
0.0625	0	0	0	0.074	0.074	0.082
0.2500	0	0	0	0.187	0.188	0.240
0.5625	0	0	0	0.480	0.484	0.600
1.0000	0	0	0	0.837	0.839	0.918
1.5625	0	0	0	0.984	0.984	0.996
2.2500	0	0	0	1.000	1.000	1.000
0	0.0100	0	0	0.053	0.053	0.053
0	0.0625	0	0	0.068	0.068	0.074
0	0.2500	0	0	0.145	0.144	0.180
0	0.5625	0	0	0.342	0.338	0.439
0	1.0000	0	0	0.662	0.656	0.779
0	1.5625	0	0	0.913	0.909	0.963
0	2.2500	0	0	0.992	0.991	0.998
0	3.0625	0	0	1.000	1.000	1.000
0	0	0.0100	0	0.052	0.052	0.052
0	0	0.0625	0	0.061	0.060	0.064
0	0	0.2500	0	0.105	0.102	0.125
0	0	0.5625	0	0.212	0.203	0.273
0	0	1.0000	0	0.416	0.398	0.528
0	0	1.5625	0	0.685	0.662	0.800
0	0	2.2500	0	0.894	0.879	0.953
0	0	3.0625	0	0.981	0.977	0.995
0	0	4.0000	0	0.998	0.998	1.000
0	0	5.0625	0	1.000	1.000	1.000
0	0	0	0.10	0.051	0.051	0.060
0	0	0	0.25	0.053	0.054	0.081
0	0	0	0.50	0.065	0.067	0.133
0	0	0	0.75	0.086	0.090	0.214
0	0	0	1.00	0.119	0.125	0.323
0	0	0	1.25	0.170	0.180	0.455
0	0	0	1.50	0.242	0.255	0.593
0	0	0	1.75	0.333	0.348	0.720
0	0	0	2.00	0.440	0.457	0.823
0	0	0	2.25	0.554	0.570	0.896
0	0	0	2.50	0.666	0.680	0.945
0	0	0	3.00	0.843	0.852	0.987
0	0	0	4.00	0.982	0.984	1.000
0	0	0	5.00	0.999	0.999	1.000
0	0	0	6.00	1.000	1.000	1.000

$$\boldsymbol{x}^{(\ell)}_{1,N_1^{(\ell)}+N_2^{(\ell)}+1}, \boldsymbol{x}^{(\ell)}_{1,N_1^{(\ell)}+N_2^{(\ell)}+2}, \ldots, \boldsymbol{x}^{(\ell)}_{1N^{(\ell)}} \overset{i.i.d.}{\sim} N_{p_1}(\boldsymbol{\mu}_1^{(\ell)}, \boldsymbol{\Sigma}_{11}^{(\ell)}), \tag{4}$$

where

$$\boldsymbol{\mu}^{(\ell)} = \begin{pmatrix} \boldsymbol{\mu}_1^{(\ell)} \\ \boldsymbol{\mu}_2^{(\ell)} \\ \boldsymbol{\mu}_3^{(\ell)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{(12)}^{(\ell)} \\ \boldsymbol{\mu}_3^{(\ell)} \end{pmatrix},$$
$$\boldsymbol{\Sigma}^{(\ell)} = \left(\begin{array}{cc|c} \boldsymbol{\Sigma}_{11}^{(\ell)} & \boldsymbol{\Sigma}_{12}^{(\ell)} & \boldsymbol{\Sigma}_{13}^{(\ell)} \\ \boldsymbol{\Sigma}_{21}^{(\ell)} & \boldsymbol{\Sigma}_{22}^{(\ell)} & \boldsymbol{\Sigma}_{23}^{(\ell)} \\ \hline \boldsymbol{\Sigma}_{31}^{(\ell)} & \boldsymbol{\Sigma}_{32}^{(\ell)} & \boldsymbol{\Sigma}_{33}^{(\ell)} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{(12)(12)}^{(\ell)} & \boldsymbol{\Sigma}_{(12)3}^{(\ell)} \\ \hline \boldsymbol{\Sigma}_{3(12)}^{(\ell)} & \boldsymbol{\Sigma}_{33}^{(\ell)} \end{array}\right),$$

and

$$\boldsymbol{x}_j^{(\ell)} = (\boldsymbol{x}_{1j}^{(\ell)\prime}, \boldsymbol{x}_{2j}^{(\ell)\prime}, \boldsymbol{x}_{3j}^{(\ell)\prime})', \ j = 1, \ldots, N_1^{(\ell)},$$
$$\boldsymbol{x}_{(12),j}^{(\ell)} = (\boldsymbol{x}_{1j}^{(\ell)\prime}, \boldsymbol{x}_{2j}^{(\ell)\prime})', \ j = N_1^{(\ell)} + 1, \ldots, N_1^{(\ell)} + N_2^{(\ell)},$$
$$N^{(\ell)} = N_1^{(\ell)} + N_2^{(\ell)} + N_3^{(\ell)}, \ \ p = p_1 + p_2 + p_3.$$

Such a dataset has three-step monotone missing data for a multi-sample problem for the ℓth population:

$$\begin{pmatrix} \overbrace{\qquad\qquad}^{p_1} & \overbrace{\qquad\qquad}^{p_2} & \overbrace{\qquad}^{p_3} \\ \boldsymbol{x}_{11}^{(\ell)\prime} & \boldsymbol{x}_{21}^{(\ell)\prime} & \boldsymbol{x}_{31}^{(\ell)\prime} \\ \vdots & \vdots & \vdots \\ \boldsymbol{x}_{1N_1^{(\ell)}}^{(\ell)\prime} & \boldsymbol{x}_{2N_1^{(\ell)}}^{(\ell)\prime} & \boldsymbol{x}_{3N_1^{(\ell)}}^{(\ell)\prime} \\ \boldsymbol{x}_{1,N_1^{(\ell)}+1}^{(\ell)\prime} & \boldsymbol{x}_{2,N_1^{(\ell)}+1}^{(\ell)\prime} & * \cdots * \\ \vdots & \vdots & \vdots \quad \vdots \\ \boldsymbol{x}_{1,N_1^{(\ell)}+N_2^{(\ell)}}^{(\ell)\prime} & \boldsymbol{x}_{2,N_1^{(\ell)}+N_2^{(\ell)}}^{(\ell)\prime} & * \cdots * \\ \boldsymbol{x}_{1,N_1^{(\ell)}+N_2^{(\ell)}+1}^{(\ell)\prime} & * \cdots * & * \cdots * \\ \vdots & \vdots \quad \vdots & \vdots \quad \vdots \\ \boldsymbol{x}_{1N^{(\ell)}}^{(\ell)\prime} & * \cdots * & * \cdots * \end{pmatrix},$$

where "$*$" indicates a missing observation.

We consider the following hypothesis:

$$H_{m0} : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} = \cdots = \boldsymbol{\mu}^{(m)}, \ \boldsymbol{\Sigma}^{(1)} = \boldsymbol{\Sigma}^{(2)} = \cdots = \boldsymbol{\Sigma}^{(m)} \ \text{vs.} \ H_{m1} : \text{not} \ H_{m0} \tag{5}$$

To derive the MLEs of the mean vectors and the covariance matrices, we consider the following transformation matrix $\boldsymbol{Z}^{(\ell)}$:

$$\boldsymbol{Z}^{(\ell)} = \left(\begin{array}{cc|c} \boldsymbol{I}_{p_1} & \boldsymbol{O} & \boldsymbol{O} \\ -\boldsymbol{\Sigma}_{21}^{(\ell)}\boldsymbol{\Sigma}_{11}^{(\ell)-1} & \boldsymbol{I}_{p_2} & \\ \hline \multicolumn{2}{c|}{-\boldsymbol{\Sigma}_{3(12)}^{(\ell)}\boldsymbol{\Sigma}_{(12)(12)}^{(\ell)-1}} & \boldsymbol{I}_{p_3} \end{array}\right).$$

The transformed vector $\boldsymbol{y}_j^{(\ell)} = (\boldsymbol{y}_{1j}^{(\ell)'}, \boldsymbol{y}_{2j}^{(\ell)'}, \boldsymbol{y}_{3j}^{(\ell)'})'$ is

$$\begin{aligned} \boldsymbol{y}_j^{(\ell)} &= \boldsymbol{Z}^{(\ell)}\boldsymbol{x}_j^{(\ell)} \\ &= \begin{pmatrix} \boldsymbol{x}_{1j}^{(\ell)} \\ -\boldsymbol{\Sigma}_{21}^{(\ell)}\boldsymbol{\Sigma}_{11}^{(\ell)-1}\boldsymbol{x}_{1j}^{(\ell)} + \boldsymbol{x}_{2j}^{(\ell)} \\ -\boldsymbol{\Sigma}_{3(12)}^{(\ell)}\boldsymbol{\Sigma}_{(12)(12)}^{(\ell)-1}\begin{pmatrix} \boldsymbol{x}_{1j}^{(\ell)} \\ \boldsymbol{x}_{2j}^{(\ell)} \end{pmatrix} + \boldsymbol{x}_{3j}^{(\ell)} \end{pmatrix}. \end{aligned}$$

The transformed parameters $(\boldsymbol{\eta}^{(\ell)}, \boldsymbol{\Delta}^{(\ell)})$ are defined as

$$\boldsymbol{\eta}^{(\ell)} = \begin{pmatrix} \boldsymbol{\eta}_1^{(\ell)} \\ \boldsymbol{\eta}_2^{(\ell)} \\ \boldsymbol{\eta}_3^{(\ell)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1^{(\ell)} \\ -\boldsymbol{\Sigma}_{21}^{(\ell)}\boldsymbol{\Sigma}_{11}^{(\ell)-1}\boldsymbol{\mu}_1^{(\ell)} + \boldsymbol{\mu}_2^{(\ell)} \\ -\boldsymbol{\Sigma}_{3(12)}^{(\ell)}\boldsymbol{\Sigma}_{(12)(12)}^{(\ell)-1}\begin{pmatrix} \boldsymbol{\mu}_1^{(\ell)} \\ \boldsymbol{\mu}_2^{(\ell)} \end{pmatrix} + \boldsymbol{\mu}_3^{(\ell)} \end{pmatrix},$$

$$\boldsymbol{\Delta}^{(\ell)} = \left(\begin{array}{cc|c} \boldsymbol{\Delta}_{11}^{(\ell)} & \boldsymbol{\Delta}_{12}^{(\ell)} & \boldsymbol{\Delta}_{13}^{(\ell)} \\ \boldsymbol{\Delta}_{21}^{(\ell)} & \boldsymbol{\Delta}_{22}^{(\ell)} & \boldsymbol{\Delta}_{23}^{(\ell)} \\ \hline \boldsymbol{\Delta}_{31}^{(\ell)} & \boldsymbol{\Delta}_{32}^{(\ell)} & \boldsymbol{\Delta}_{33}^{(\ell)} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Delta}_{(12)(12)}^{(\ell)} & \boldsymbol{\Delta}_{(12)3}^{(\ell)} \\ \hline \boldsymbol{\Delta}_{3(12)}^{(\ell)} & \boldsymbol{\Delta}_{33}^{(\ell)} \end{array}\right),$$

where

$$\begin{aligned} \boldsymbol{\Delta}_{11}^{(\ell)} &= \boldsymbol{\Sigma}_{11}^{(\ell)}, \\ \boldsymbol{\Delta}_{12}^{(\ell)} &= \boldsymbol{\Delta}_{21}^{(\ell)'} = \boldsymbol{\Sigma}_{11}^{(\ell)-1}\boldsymbol{\Sigma}_{12}^{(\ell)}, \\ \boldsymbol{\Delta}_{22}^{(\ell)} &= \boldsymbol{\Sigma}_{22\cdot 1}^{(\ell)} = \boldsymbol{\Sigma}_{22}^{(\ell)} - \boldsymbol{\Sigma}_{21}^{(\ell)}\boldsymbol{\Sigma}_{11}^{(\ell)-1}\boldsymbol{\Sigma}_{12}^{(\ell)}, \\ \boldsymbol{\Delta}_{(12)3}^{(\ell)} &= \boldsymbol{\Delta}_{3(12)}^{(\ell)'} = \boldsymbol{\Sigma}_{(12)(12)}^{(\ell)-1}\boldsymbol{\Sigma}_{(12)3}^{(\ell)}, \\ \boldsymbol{\Delta}_{33}^{(\ell)} &= \boldsymbol{\Sigma}_{33\cdot 12}^{(\ell)} = \boldsymbol{\Sigma}_{33}^{(\ell)} - \boldsymbol{\Sigma}_{3(12)}^{(\ell)}\boldsymbol{\Sigma}_{(12)(12)}^{(\ell)-1}\boldsymbol{\Sigma}_{(12)3}^{(\ell)}. \end{aligned}$$

We note that the pair $(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)})$ is in one-to-one correspondence with $(\boldsymbol{\eta}^{(\ell)}, \boldsymbol{\Delta}^{(\ell)})$. Under H_1, we define the MLEs of $(\boldsymbol{\eta}^{(\ell)}, \boldsymbol{\Delta}^{(\ell)})$ as $(\widehat{\boldsymbol{\eta}}^{(\ell)}, \widehat{\boldsymbol{\Delta}}^{(\ell)})$,

$$\widehat{\boldsymbol{\eta}}_1^{(\ell)} = \frac{1}{N^{(\ell)}}(N_1^{(\ell)}\overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)}\overline{\boldsymbol{x}}_{(2)1}^{(\ell)} + N_3^{(\ell)}\overline{\boldsymbol{x}}_{(3)}^{(\ell)}),$$

$$
\begin{aligned}
\widehat{\boldsymbol{\eta}}_2^{(\ell)} &= \frac{1}{N_1^{(\ell)} + N_2^{(\ell)}}\{N_1^{(\ell)}\overline{\boldsymbol{x}}_{(1)2}^{(\ell)} + N_2^{(\ell)}\overline{\boldsymbol{x}}_{(2)2}^{(\ell)} - \widehat{\boldsymbol{\Delta}}_{21}^{(\ell)}(N_1^{(\ell)}\overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)}\overline{\boldsymbol{x}}_{(2)1}^{(\ell)})\},\\
\widehat{\boldsymbol{\eta}}_3^{(\ell)} &= \overline{\boldsymbol{x}}_{(1)3}^{(\ell)} - \widehat{\boldsymbol{\Delta}}_{3(12)}^{(\ell)}\overline{\boldsymbol{x}}_{(1)(12)}^{(\ell)},\\
\widehat{\boldsymbol{\Delta}}_{11}^{(\ell)} &= \frac{1}{N^{(\ell)}}(\boldsymbol{W}_{(1)11}^{(\ell)} + \boldsymbol{W}_{(2)11}^{(\ell)} + \boldsymbol{W}_{(3)}^{(\ell)}), \qquad (6)\\
\widehat{\boldsymbol{\Delta}}_{22}^{(\ell)} &= \frac{1}{N_1^{(\ell)} + N_2^{(\ell)}}(\boldsymbol{W}_{(1),(12)(12)}^{(\ell)} + \boldsymbol{W}_{(2)}^{(\ell)})_{22\cdot 1},\ \widehat{\boldsymbol{\Delta}}_{33}^{(\ell)} = \frac{1}{N_1^{(\ell)}}(\boldsymbol{W}_{(1)33\cdot 12}^{(\ell)}),\\
\widehat{\boldsymbol{\Delta}}_{12}^{(\ell)} &= \widehat{\boldsymbol{\Delta}}_{21}^{(\ell)\prime} = (\boldsymbol{W}_{(1)11}^{(\ell)} + \boldsymbol{W}_{(2)11}^{(\ell)})^{-1}(\boldsymbol{W}_{(1)12}^{(\ell)} + \boldsymbol{W}_{(2)12}^{(\ell)}),\\
\widehat{\boldsymbol{\Delta}}_{(12)3}^{(\ell)} &= \widehat{\boldsymbol{\Delta}}_{3(12)}^{(\ell)\prime} = (\boldsymbol{W}_{(1),(12)(12)}^{(\ell)})^{-1}\boldsymbol{W}_{(1),(12)3}^{(\ell)}, \qquad (7)
\end{aligned}
$$

where

$$
\begin{aligned}
\overline{\boldsymbol{x}}_{(1)}^{(\ell)} &= \begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} \\ \overline{\boldsymbol{x}}_{(1)2}^{(\ell)} \\ \overline{\boldsymbol{x}}_{(1)3}^{(\ell)} \end{pmatrix} = \begin{pmatrix} \overline{\boldsymbol{x}}_{(1)(12)}^{(\ell)} \\ \overline{\boldsymbol{x}}_{(1)3}^{(\ell)} \end{pmatrix},\\
\overline{\boldsymbol{x}}_{(1)1}^{(\ell)} &= \frac{1}{N_1^{(\ell)}}\sum_{j=1}^{N_1^{(\ell)}}\boldsymbol{x}_{1j}^{(\ell)},\ \overline{\boldsymbol{x}}_{(1)2}^{(\ell)} = \frac{1}{N_1^{(\ell)}}\sum_{j=1}^{N_1^{(\ell)}}\boldsymbol{x}_{2j}^{(\ell)},\ \overline{\boldsymbol{x}}_{(1)3}^{(\ell)} = \frac{1}{N_1^{(\ell)}}\sum_{j=1}^{N_1^{(\ell)}}\boldsymbol{x}_{3j}^{(\ell)},\\
\overline{\boldsymbol{x}}_{(2)}^{(\ell)} &= \begin{pmatrix} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} \\ \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} \end{pmatrix},\ \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} = \frac{1}{N_2^{(\ell)}}\sum_{j=N_1^{(\ell)}+1}^{N_1^{(\ell)}+N_2^{(\ell)}}\boldsymbol{x}_{1j}^{(\ell)},\ \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} = \frac{1}{N_2^{(\ell)}}\sum_{j=N_1^{(\ell)}+1}^{N_1^{(\ell)}+N_2^{(\ell)}}\boldsymbol{x}_{2j}^{(\ell)},\\
\overline{\boldsymbol{x}}_{(3)}^{(\ell)} &= \frac{1}{N_3^{(\ell)}}\sum_{j=N_1^{(\ell)}+N_2^{(\ell)}+1}^{N^{(\ell)}}\boldsymbol{x}_{1j}^{(\ell)},\\
\boldsymbol{W}_{(1)}^{(\ell)} &= \sum_{j=1}^{N_1}(\boldsymbol{x}_j^{(\ell)} - \overline{\boldsymbol{x}}_{(1)}^{(\ell)})(\boldsymbol{x}_j^{(\ell)} - \overline{\boldsymbol{x}}_{(1)}^{(\ell)})'\\
&= \left(\begin{array}{cc|c} \boldsymbol{W}_{(1)11}^{(\ell)} & \boldsymbol{W}_{(1)12}^{(\ell)} & \boldsymbol{W}_{(1)13}^{(\ell)} \\ \boldsymbol{W}_{(1)21}^{(\ell)} & \boldsymbol{W}_{(1)22}^{(\ell)} & \boldsymbol{W}_{(1)23}^{(\ell)} \\ \hline \boldsymbol{W}_{(1)31}^{(\ell)} & \boldsymbol{W}_{(1)32}^{(\ell)} & \boldsymbol{W}_{(1)33}^{(\ell)} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{W}_{(1),(12)(12)}^{(\ell)} & \boldsymbol{W}_{(1),(12)3}^{(\ell)} \\ \hline \boldsymbol{W}_{(1),3(12)}^{(\ell)} & \boldsymbol{W}_{(1)33}^{(\ell)} \end{array}\right),\\
\boldsymbol{W}_{(2)}^{(\ell)} &= \sum_{j=N_1^{(\ell)}+1}^{N_1^{(\ell)}+N_2^{(\ell)}}\begin{pmatrix} \boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} \\ \boldsymbol{x}_{2j}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} \end{pmatrix}\begin{pmatrix} \boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} \\ \boldsymbol{x}_{2j}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} \end{pmatrix}'\\
&\quad + \frac{N_1^{(\ell)}N_2^{(\ell)}}{N_1^{(\ell)} + N_2^{(\ell)}}\begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} \\ \overline{\boldsymbol{x}}_{(1)2}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} \end{pmatrix}\begin{pmatrix} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} \\ \overline{\boldsymbol{x}}_{(1)2}^{(\ell)} - \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} \end{pmatrix}'\\
&= \begin{pmatrix} \boldsymbol{W}_{(2)11}^{(\ell)} & \boldsymbol{W}_{(2)12}^{(\ell)} \\ \boldsymbol{W}_{(2)21}^{(\ell)} & \boldsymbol{W}_{(2)22}^{(\ell)} \end{pmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{W}_{(3)}^{(\ell)} = &\sum_{j=N_1^{(\ell)}+N_2^{(\ell)}+1}^{N^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(3)}^{(\ell)})(\boldsymbol{x}_{1j}^{(\ell)} - \overline{\boldsymbol{x}}_{(3)}^{(\ell)})' \\
&+ \frac{(N_1^{(\ell)} + N_2^{(\ell)})N_3^{(\ell)}}{N^{(\ell)}} \left(\overline{\boldsymbol{x}}_{(3)}^{(\ell)} - \frac{1}{N_1^{(\ell)} + N_2^{(\ell)}} (N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)}) \right) \\
&\times \left(\overline{\boldsymbol{x}}_{(3)}^{(\ell)} - \frac{1}{N_1^{(\ell)} + N_2^{(\ell)}} (N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)}) \right)'.
\end{aligned}
$$

Conversely, under H_{m0}, we define MLEs of $\boldsymbol{\eta}(= \boldsymbol{\eta}^{(1)} = \cdots = \boldsymbol{\eta}^{(m)})$, $\boldsymbol{\Delta}(= \boldsymbol{\Delta}^{(1)} = \cdots = \boldsymbol{\Delta}^{(m)})$ as $(\widetilde{\boldsymbol{\eta}}, \widetilde{\boldsymbol{\Delta}})$. Subsequently, we obtain

$$
\begin{aligned}
\widetilde{\boldsymbol{\eta}}_1 &= \frac{1}{N} \sum_{\ell=1}^{m} (N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)} + N_3^{(\ell)} \overline{\boldsymbol{x}}_{(3)}^{(\ell)}), \\
\widetilde{\boldsymbol{\eta}}_2 &= \frac{1}{N_1 + N_2} \sum_{\ell=1}^{m} \{ N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)2}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)2}^{(\ell)} - \widetilde{\boldsymbol{\Delta}}_{21} (N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)}) \}, \\
\widetilde{\boldsymbol{\eta}}_3 &= \frac{1}{N_1} \sum_{\ell=1}^{m} N_1^{(\ell)} \{ \overline{\boldsymbol{x}}_{(1)3}^{(\ell)} - \widetilde{\boldsymbol{\Delta}}_{3(12)} \overline{\boldsymbol{x}}_{(1)(12)}^{(\ell)} \}, \\
\widetilde{\boldsymbol{\Delta}}_{11} &= \frac{1}{N} \sum_{\ell=1}^{m} \sum_{j=1}^{N^{(\ell)}} (\boldsymbol{x}_{1j}^{(\ell)} - \widetilde{\boldsymbol{\eta}}_1)(\boldsymbol{x}_{1j}^{(\ell)} - \widetilde{\boldsymbol{\eta}}_1)', \qquad (8) \\
\widetilde{\boldsymbol{\Delta}}_{22} &= \frac{1}{N_1 + N_2} \sum_{\ell=1}^{m} \sum_{j=1}^{N_1^{(\ell)}+N_2^{(\ell)}} (-\widetilde{\boldsymbol{\Delta}}_{21} \boldsymbol{x}_{1j}^{(\ell)} + \boldsymbol{x}_{2j}^{(\ell)} - \widetilde{\boldsymbol{\eta}}_2)(-\widetilde{\boldsymbol{\Delta}}_{21} \boldsymbol{x}_{1j}^{(\ell)} + \boldsymbol{x}_{2j}^{(\ell)} - \widetilde{\boldsymbol{\eta}}_2)', \qquad (9) \\
\widetilde{\boldsymbol{\Delta}}_{33} &= \frac{1}{N_1} \sum_{\ell=1}^{m} \sum_{j=1}^{N_1^{(\ell)}} (-\widetilde{\boldsymbol{\Delta}}_{3(12)} \boldsymbol{x}_{(12)j}^{(\ell)} + \boldsymbol{x}_{3j}^{(\ell)} - \widetilde{\boldsymbol{\eta}}_3)(-\widetilde{\boldsymbol{\Delta}}_{3(12)} \boldsymbol{x}_{(12)j}^{(\ell)} + \boldsymbol{x}_{3j}^{(\ell)} - \widetilde{\boldsymbol{\eta}}_3)', \\
\widetilde{\boldsymbol{\Delta}}_{21} &= \widetilde{\boldsymbol{\Delta}}_{12}' \\
&= \sum_{\ell=1}^{m} \left[\sum_{j=1}^{N_1^{(\ell)}+N_2^{(\ell)}} \boldsymbol{x}_{2j}^{(\ell)} \boldsymbol{x}_{1j}^{(\ell)'} - \frac{1}{N_1 + N_2} \left\{ \sum_{k=1}^{m} (N_1^{(k)} \overline{\boldsymbol{x}}_{(1)2}^{(k)} + N_2^{(k)} \overline{\boldsymbol{x}}_{(2)2}^{(k)}) \right\} \right. \\
&\quad \left. \times (N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)})' \right] \sum_{\ell=1}^{m} \left[\sum_{j=1}^{N_1^{(\ell)}+N_2^{(\ell)}} \boldsymbol{x}_{1j}^{(\ell)} \boldsymbol{x}_{1j}^{(\ell)'} \right. \\
&\quad \left. - \frac{1}{N_1 + N_2} \left\{ \sum_{k=1}^{m} (N_1^{(k)} \overline{\boldsymbol{x}}_{(1)1}^{(k)} + N_2^{(k)} \overline{\boldsymbol{x}}_{(2)1}^{(k)}) \right\} (N_1^{(\ell)} \overline{\boldsymbol{x}}_{(1)1}^{(\ell)} + N_2^{(\ell)} \overline{\boldsymbol{x}}_{(2)1}^{(\ell)})' \right]^{-1},
\end{aligned}
$$

$$\begin{aligned}\widetilde{\boldsymbol{\Delta}}_{3(12)} &= \widetilde{\boldsymbol{\Delta}}'_{(12)3} \\ &= \sum_{\ell=1}^{m} \left\{ \sum_{j=1}^{N_1^{(\ell)}} \boldsymbol{x}_{3j}^{(\ell)} \boldsymbol{x}_{(12),j}^{(\ell)\prime} - N_1^{(\ell)} \left(\frac{1}{N_1} \sum_{k=1}^{m} N_1^{(k)} \overline{\boldsymbol{x}}_{(1)3}^{(\ell)} \right) \overline{\boldsymbol{x}}_{(1),(12)}^{(\ell)} \right\} \\ &\quad \times \sum_{\ell=1}^{m} \left\{ \sum_{j=1}^{N_1^{(\ell)}} \boldsymbol{x}_{(12),j}^{(\ell)} \boldsymbol{x}_{(12),j}^{(\ell)\prime} - N_1^{(\ell)} \left(\frac{1}{N_1} \sum_{k=1}^{m} N_1^{(k)} \overline{\boldsymbol{x}}_{(1),(12)}^{(\ell)} \right) \overline{\boldsymbol{x}}_{(1),(12)}^{(\ell)} \right\}^{-1}, \end{aligned} \tag{10}$$

where $N = \sum_{\ell=1}^{m} N^{(\ell)}$, $N_1 = \sum_{\ell=1}^{m} N_1^{(\ell)}$, $N_2 = \sum_{\ell=1}^{m} N_2^{(\ell)}$. From the preceding MLEs, we get the following theorem.

Theorem 2 *Suppose that the datasets have a three-step monotone pattern of missing observations and are normally distributed as* (4)*. Then, the LR for* (5) *can be given by*

$$\lambda_m = \frac{\prod_{\ell=1}^{m} |\widehat{\boldsymbol{\Delta}}_{11}^{(\ell)}|^{\frac{N^{(\ell)}}{2}} |\widehat{\boldsymbol{\Delta}}_{22}^{(\ell)}|^{\frac{N_1^{(\ell)}+N_2^{(\ell)}}{2}} |\widehat{\boldsymbol{\Delta}}_{33}^{(\ell)}|^{\frac{N_1^{(\ell)}}{2}}}{|\widetilde{\boldsymbol{\Delta}}_{11}|^{\frac{N}{2}} |\widetilde{\boldsymbol{\Delta}}_{22}|^{\frac{N_1+N_2}{2}} |\widetilde{\boldsymbol{\Delta}}_{33}|^{\frac{N_1}{2}}},$$

where $\widehat{\boldsymbol{\Delta}}_{ii}$ *and* $\widetilde{\boldsymbol{\Delta}}_{ii}$ $(i = 1, 2, 3)$ *are given in* (6)–(10).

Thus, we obtain LRT statistic $-2\log\lambda_m$. $-2\log\lambda_m$ is asymptotically distributed as a χ^2 distribution with $f_m = p(p+3)(m-1)/2$ degrees of freedom. However, it is known that the accuracy of this approximation is not good for small samples. Therefore, we propose the test statistics that are a good approximation to χ^2 distribution using several methods based on the complete data case in Sect. 3.2.

3.2 Complete Data

In this subsection, we discuss the LRT statistic in the case of complete data and the modified LRT statistics with Bartlett correction. The results will be used to propose the test statistics in the next subsection. First, we consider a simultaneous test with complete data as follows:

$$H_{03} : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} = \cdots = \boldsymbol{\mu}^{(m)},\ \boldsymbol{\Sigma}^{(1)} = \boldsymbol{\Sigma}^{(2)} = \cdots = \boldsymbol{\Sigma}^{(m)} \ \text{ vs. } \ H_{13} : \text{not } H_{03}$$

$\boldsymbol{x}_1^{(\ell)}, \boldsymbol{x}_2^{(\ell)}, \ldots, \boldsymbol{x}_{N^{(\ell)}}^{(\ell)}$ be independently distributed as $N_p(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)})$, and let λ_{S_m} be the LR for the complete data. Then, the LR is given by

$$\lambda_{S_m} = \frac{\prod_{\ell=1}^{m} \left| \frac{1}{N^{(\ell)}} \boldsymbol{V}^{(\ell)} \right|^{\frac{1}{2}N^{(\ell)}}}{\left| \frac{1}{N}(\boldsymbol{V} + \boldsymbol{B}) \right|^{\frac{1}{2}N}},$$

where

$$\boldsymbol{V}^{(\ell)} = \sum_{j=1}^{N^{(\ell)}} (\boldsymbol{x}_j^{(\ell)} - \overline{\boldsymbol{x}}^{(\ell)})(\boldsymbol{x}_j^{(\ell)} - \overline{\boldsymbol{x}}^{(\ell)})', \quad \boldsymbol{V} = \sum_{\ell=1}^{m} \boldsymbol{V}^{(\ell)},$$

$$\boldsymbol{B} = \sum_{\ell=1}^{m} N^{(\ell)} (\overline{\boldsymbol{x}}^{(\ell)} - \overline{\boldsymbol{x}})(\overline{\boldsymbol{x}}^{(\ell)} - \overline{\boldsymbol{x}})',$$

$$\overline{\boldsymbol{x}}^{(\ell)} = \frac{1}{N^{(\ell)}} \sum_{j=1}^{N^{(\ell)}} \boldsymbol{x}_j^{(\ell)}, \quad \overline{\boldsymbol{x}} = \frac{1}{N} \sum_{\ell=1}^{m} N^{(\ell)} \overline{\boldsymbol{x}}^{(\ell)}, \quad N = \sum_{\ell=1}^{m} N^{(\ell)}.$$

Furthermore, the modified LRT statistic with Bartlett correction can be given by $-2\rho_3 \log \lambda_{S_m}$ (Muirhead [6, p. 513]), where

$$\rho_3 = 1 - \frac{2p^2 + 9p + 11}{6N(p+3)(m-1)} \left(\sum_{\ell=1}^{m} \frac{N}{N^{(\ell)}} - 1 \right).$$

Next, we consider the covariance test in the case of complete data as follows:

$$H_{04} : \boldsymbol{\Sigma}^{(1)} = \boldsymbol{\Sigma}^{(2)} = \cdots = \boldsymbol{\Sigma}^{(m)} \text{ vs. } H_{14} : \text{not } H_{04}$$

The modified LRT statistic $-2\rho_4 \log \lambda_{V_m}$ was provided by Muirhead [6, p. 308], where

$$\rho_4 = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(m-1)} \left(\sum_{\ell=1}^{m} \frac{1}{n^{(\ell)}} - \frac{1}{n} \right), \quad \lambda_{V_m} = \frac{\prod_{\ell=1}^{m} \left| \frac{1}{n^{(\ell)}} \boldsymbol{V}^{(\ell)} \right|^{\frac{n^{(\ell)}}{2}}}{\left| \frac{1}{n} \boldsymbol{V} \right|^{\frac{n}{2}}},$$

and

$$n^{(\ell)} = N^{(\ell)} - 1, \quad n = \sum_{\ell-1}^{m} n^{(\ell)}.$$

3.3 Test Statistics

Using LR of simultaneous test with complete data from the previous subsection, we propose test statistics by decomposing the LR λ_m with three-step monotone missing data. First, the LR can be decomposed as $\lambda_m = \xi_1 \xi_2 \xi_3$, where

$$\xi_1 = \frac{\prod_{\ell=1}^{m} \left| \widehat{\boldsymbol{\Delta}}_{11}^{(\ell)} \right|^{\frac{N^{(\ell)}}{2}}}{\left| \widetilde{\boldsymbol{\Delta}}_{11} \right|^{\frac{N}{2}}}, \quad \xi_2 = \frac{\prod_{\ell=1}^{m} \left| \widehat{\boldsymbol{\Delta}}_{22}^{(\ell)} \right|^{\frac{N_1^{(\ell)} + N_2^{(\ell)}}{2}}}{\left| \widetilde{\boldsymbol{\Delta}}_{22} \right|^{\frac{N_1 + N_2}{2}}}, \quad \xi_3 = \frac{\prod_{\ell=1}^{m} \left| \widehat{\boldsymbol{\Delta}}_{33}^{(\ell)} \right|^{\frac{N_1^{(\ell)}}{2}}}{\left| \widetilde{\boldsymbol{\Delta}}_{33} \right|^{\frac{N_1}{2}}}.$$

Because ξ_1 is of the form of LR for H_{03} in the case of without missing data, the modified LRT statistic $-2\rho_{\xi_1}\log\xi_1$ is given, where

$$\rho_{\xi_1}=1-\frac{2p_1^2+9p_1+11}{6N(p_1+3)(m-1)}\left(\sum_{\ell=1}^{m}\frac{N}{N^{(\ell)}}-1\right).$$

Next, ξ_2 can be decomposed as $\xi_2=\xi_2^{\dagger}\xi_2^{\ddagger}$, where

$$\xi_2^{\dagger}=\frac{\prod_{\ell=1}^{m}\left|\widehat{\boldsymbol{\Delta}}_{22}^{(\ell)}\right|^{\frac{N_1^{(\ell)}+N_2^{(\ell)}}{2}}}{\left|\frac{1}{N_1+N_2}(\boldsymbol{V}_{p_2}+\boldsymbol{B}_{p_2})\right|^{\frac{N_1+N_2}{2}}},\quad \xi_2^{\ddagger}=\frac{\left|\frac{1}{N_1+N_2}(\boldsymbol{V}_{p_2}+\boldsymbol{B}_{p_2})\right|^{\frac{N_1+N_2}{2}}}{\left|\widetilde{\boldsymbol{\Delta}}_{22}\right|^{\frac{N_1+N_2}{2}}},$$

and

$$\boldsymbol{V}_{p_2}^{(\ell)}=\sum_{j=1}^{N_1^{(\ell)}+N_2^{(\ell)}}(\boldsymbol{x}_{2j}^{(\ell)}-\widehat{\boldsymbol{\Delta}}_{21}^{(\ell)}\boldsymbol{x}_{1j}^{(\ell)}-\widehat{\boldsymbol{\eta}}_2^{(\ell)})(\boldsymbol{x}_{2j}^{(\ell)}-\widehat{\boldsymbol{\Delta}}_{21}^{(\ell)}\boldsymbol{x}_{1j}^{(\ell)}-\widehat{\boldsymbol{\eta}}_2^{(\ell)})',$$

$$\boldsymbol{V}_{p_2}=\sum_{\ell=1}^{m}\boldsymbol{V}_{p_2}^{(\ell)},\ \ \boldsymbol{B}_{p_2}=\sum_{\ell=1}^{m}(N_1^{(\ell)}+N_2^{(\ell)})(\widehat{\boldsymbol{\eta}}_2^{(\ell)}-\widetilde{\boldsymbol{\eta}}_2)(\widehat{\boldsymbol{\eta}}_2^{(\ell)}-\widetilde{\boldsymbol{\eta}}_2)'.$$

Since $\xi_2^{\dagger}$ is of the form of LR for H_{03} in the case of complete data, the modified LRT statistic $-2\rho_{\xi_2}\log\xi_2^{\dagger}$ is given, where

$$\rho_{\xi_2}=1-\frac{2p_2^2+9p_2+11}{6(N_1+N_2)(p_2+3)(m-1)}\left(\sum_{\ell=1}^{m}\frac{N_1+N_2}{N_1^{(\ell)}+N_2^{(\ell)}}-1\right).$$

Similarly, ξ_3 can be decomposed as $\xi_3=\xi_3^{\dagger}\xi_3^{\ddagger}$, where

$$\xi_3^{\dagger}=\frac{\prod_{\ell=1}^{m}\left|\widehat{\boldsymbol{\Delta}}_{33}^{(\ell)}\right|^{\frac{N_1^{(\ell)}}{2}}}{\left|\frac{1}{N_1}(\boldsymbol{V}_{p_3}+\boldsymbol{B}_{p_3})\right|^{\frac{N_1}{2}}},\quad \xi_3^{\ddagger}=\frac{\left|\frac{1}{N_1}(\boldsymbol{V}_{p_3}+\boldsymbol{B}_{p_3})\right|^{\frac{N_1}{2}}}{\left|\widetilde{\boldsymbol{\Delta}}_{33}\right|^{\frac{N_1}{2}}},$$

and

$$\boldsymbol{V}_{p_3}^{(\ell)}=\sum_{j=1}^{N_1^{(\ell)}}(\boldsymbol{x}_{3j}^{(\ell)}-\widehat{\boldsymbol{\Delta}}_{3(12)}^{(\ell)}\boldsymbol{x}_{(12)j}^{(\ell)}-\widehat{\boldsymbol{\eta}}_3^{(\ell)})(\boldsymbol{x}_{3j}^{(\ell)}-\widehat{\boldsymbol{\Delta}}_{3(12)}^{(\ell)}\boldsymbol{x}_{(12)j}^{(\ell)}-\widehat{\boldsymbol{\eta}}_3^{(\ell)})',$$

$$\boldsymbol{V}_{p_3}=\sum_{\ell=1}^{m}\boldsymbol{V}_{p_3}^{(\ell)},\ \ \boldsymbol{B}_{p_3}=\sum_{\ell=1}^{m}N_1^{(\ell)}(\widehat{\boldsymbol{\eta}}_3^{(\ell)}-\widetilde{\boldsymbol{\eta}}_3)(\widehat{\boldsymbol{\eta}}_3^{(\ell)}-\widetilde{\boldsymbol{\eta}}_3)'.$$

Since $\xi_3^\dagger$ is of the form of LR for H_{03} in the case of complete data, the modified LRT statistic $-2\rho_{\xi_3}\log\xi_3^\dagger$ is given, where

$$\rho_{\xi_3} = 1 - \frac{2p_3^2 + 9p_3 + 11}{6N_1(p_3+3)(m-1)}\left(\sum_{\ell=1}^{m}\frac{N_1}{N_1^{(\ell)}} - 1\right).$$

Therefore, the decomposed form of $\lambda_m = \xi_1\xi_2\xi_3$ is $\lambda_m = \xi_1\xi_2^\dagger\xi_2^\ddagger\xi_3^\dagger\xi_3^\ddagger$. We give a correction only for $\xi_1, \xi_2^\dagger$, and $\xi_3^\dagger$. Thus, we give the test statistic $-2\log\tau_m$ for improving the accuracy of the χ^2 approximation, where

$$\tau_m = (\xi_1)^{\rho_{\xi_1}}(\xi_2^\dagger)^{\rho_{\xi_2}}\xi_2^\ddagger(\xi_3^\dagger)^{\rho_{\xi_3}}\xi_3^\ddagger.$$

Next, using the LR of the covariance test with complete data from the previous subsection, we propose test statistics by decomposing the LR λ_m with three-step monotone missing data. Let

$$\xi_{11}^* = \frac{\prod_{\ell=1}^m \left|\frac{\boldsymbol{V}_{p_1}^{(\ell)}}{n^{(\ell)}}\right|^{\frac{n^{(\ell)}}{2}}}{\left|\frac{\boldsymbol{V}_{p_1}}{n}\right|^{\frac{n}{2}}},\quad \xi_{21}^* = \frac{\prod_{\ell=1}^m \left|\frac{\boldsymbol{V}_{p_2}^{(\ell)}}{n_1^{(\ell)}+n_2^{(\ell)}}\right|^{\frac{n_1^{(\ell)}+n_2^{(\ell)}}{2}}}{\left|\frac{\boldsymbol{V}_{p_2}}{n_1+n_2}\right|^{\frac{n_1+n_2}{2}}},\quad \xi_{31}^* = \frac{\prod_{\ell=1}^m \left|\frac{\boldsymbol{V}_{p_3}^{(\ell)}}{n_1^{(\ell)}}\right|^{\frac{n_1^{(\ell)}}{2}}}{\left|\frac{\boldsymbol{V}_{p_3}}{n_1}\right|^{\frac{n_1}{2}}},$$

where

$$\boldsymbol{V}_{p_1}^{(\ell)} = \sum_{j=1}^{N^{(\ell)}}(\boldsymbol{x}_{1j}^{(\ell)} - \widehat{\boldsymbol{\eta}}_1^{(\ell)})(\boldsymbol{x}_{1j}^{(\ell)} - \widehat{\boldsymbol{\eta}}_1^{(\ell)})', \quad \boldsymbol{V}_{p_1} = \sum_{\ell=1}^m \boldsymbol{V}_{p_1}^{(\ell)},$$

$$n_1^{(\ell)} = N_1^{(\ell)} - (p_1+p_2) - 1, \quad n_1 = \sum_{\ell=1}^m n_1^{(\ell)},$$

$$n_1^{(\ell)} + n_2^{(\ell)} = N_1^{(\ell)} + N_2^{(\ell)} - p_1 - 1, \quad n_1 + n_2 = \sum_{\ell=1}^m (n_1^{(\ell)} + n_2^{(\ell)}).$$

Because ξ_{11}^*, ξ_{21}^*, and ξ_{31}^* are of the form of LR for H_{04} in the case of complete data, the modified LRT statistics $-2\rho_{\xi_{11}^*}\log\xi_{11}^*$, $-2\rho_{\xi_{21}^*}\log\xi_{21}^*$, and $-2\rho_{\xi_{31}^*}\log\xi_{31}^*$ are given, where

$$\rho_{\xi_{11}^*} = 1 - \frac{2p_1^2 + 3p_1 - 1}{6(p_1+1)(m-1)}\left(\sum_{\ell=1}^m \frac{1}{n^{(\ell)}} - \frac{1}{n}\right),$$

$$\rho_{\xi_{21}^*} = 1 - \frac{2p_2^2 + 3p_2 - 1}{6(p_2+1)(m-1)}\left(\sum_{\ell=1}^m \frac{1}{n_1^{(\ell)} + n_2^{(\ell)}} - \frac{1}{n_1+n_2}\right),$$

$$\rho_{\xi_{31}^*} = 1 - \frac{2p_3^2 + 3p_3 - 1}{6(p_3+1)(m-1)} \left(\sum_{\ell=1}^{m} \frac{1}{n_1^{(\ell)}} - \frac{1}{n_1} \right).$$

Therefore, we propose the test statistic $-2\log\phi_m$ to improve the accuracy of the χ^2 approximation, where

$$\phi_m = (\xi_{11}^*)^{\rho_{\xi_{11}^*}} (\xi_{21}^*)^{\rho_{\xi_{21}^*}} (\xi_{31}^*)^{\rho_{\xi_{31}^*}} \frac{\lambda}{\xi_{11}\xi_{21}\xi_{31}},$$

and

$$\xi_{11} = \frac{\prod_{\ell=1}^{m} \left| \frac{\boldsymbol{V}_{p_1}^{(\ell)}}{N^{(\ell)}} \right|^{\frac{N^{(\ell)}}{2}}}{\left| \frac{\boldsymbol{V}_{p_1}}{N} \right|^{\frac{N}{2}}}, \quad \xi_{21} = \frac{\prod_{\ell=1}^{m} \left| \frac{\boldsymbol{V}_{p_2}^{(\ell)}}{N_1^{(\ell)}+N_2^{(\ell)}} \right|^{\frac{N_1^{(\ell)}+N_2^{(\ell)}}{2}}}{\left| \frac{\boldsymbol{V}_{p_2}}{N_1+N_2} \right|^{\frac{N_1+N_2}{2}}}, \quad \xi_{31} = \frac{\prod_{\ell=1}^{m} \left| \frac{\boldsymbol{V}_{p_3}^{(\ell)}}{N_1^{(\ell)}} \right|^{\frac{N_1^{(\ell)}}{2}}}{\left| \frac{\boldsymbol{V}_{p_3}}{N_1} \right|^{\frac{N_1}{2}}}.$$

Next, we propose the test statistic $-2\rho_{L_m}\log\lambda_m$ via linear interpolation, where

$$\rho_{L_m} = \left\{ 1 - \frac{(p_1+p_2)N_2 + p_1 N_3}{p(N_2+N_3)} \right\} \rho_{N_1,m} + \frac{(p_1+p_2)N_2 + p_1 N_3}{p(N_2+N_3)} \rho_{N,m}$$

and

$$\rho_{N,m} = 1 - \frac{2p^2+9p+11}{6N(p+3)(m-1)} \left(\sum_{\ell=1}^{m} \frac{N}{N^{(\ell)}} - 1 \right),$$

$$\rho_{N_1,m} = 1 - \frac{2p^2+9p+11}{6N_1(p+3)(m-1)} \left(\sum_{\ell=1}^{m} \frac{N_1}{N_1^{(\ell)}} - 1 \right).$$

3.4 Asymptotic Expansion Approximation

In this subsection, we give an approximate upper 100α percentile of $-2\log\lambda_m$ with three-step monotone missing data for a multi-sample problem. The upper 100α percentile of $-2\log\lambda_{S_m}$ can be expanded as

$$q_{c_m}^*(\alpha) = \chi^2_{f_m;1-\alpha} + \frac{\nu}{N} \left(\sum_{\ell=1}^{m} \frac{1}{k_1^{(\ell)}} - 1 \right) \chi^2_{f_m;1-\alpha} + \frac{\chi^2_{f_m;1-\alpha}}{N^2} \left\{ \nu^2 \left(\sum_{\ell=1}^{m} \frac{1}{k_1^{(\ell)}} - 1 \right)^2 + \frac{2\gamma_1}{f_m} + \frac{2\gamma_1 \chi^2_{f_m;1-\alpha}}{f_m(f_m+2)} \right\} + O(N^{-3}),$$

where

$$\nu = \frac{2p^2+9p+11}{6(p+3)(m-1)},\ k_1^{(\ell)} = \frac{N^{(\ell)}}{N},$$
$$\gamma_1 = \frac{1}{288}\left[6p(p+1)(p+2)(p+3)\left(\sum_{\ell=1}^{m}\frac{1}{\{k_1^{(\ell)}\}^2}-1\right)\right.$$
$$\left.-\frac{(2p^2+9p+11)^2(2p-1)}{p(p+3)(m-1)}\left(\sum_{\ell=1}^{m}\frac{1}{k_1^{(\ell)}}-1\right)^2\right],$$

where $\chi^2_{f_m;1-\alpha}$ is the upper 100α percentile of the χ^2 distribution with f_m degrees of freedom (Hosoya and Seo [3]). Based on linear interpolation and letting $q_m^*(\alpha)$ be the upper 100α percentile of $-2\log\lambda_m$, the following can be obtained:

$$q_m^*(\alpha) = \left\{1 - \frac{(p_1+p_2)N_2+p_1N_3}{p(N_2+N_3)}\right\}q_{N_1,m}(\alpha) + \frac{(p_1+p_2)N_2+p_1N_3}{p(N_2+N_3)}q_{N,m}(\alpha)$$

where

$$q_{N,m}(\alpha) = \chi^2_{f_m;1-\alpha} + \frac{\nu}{N}\left(\sum_{\ell=1}^{m}\frac{1}{k_1^{(\ell)}}-1\right)\chi^2_{f_m;1-\alpha}$$
$$+\frac{1}{N^2}\chi^2_{f_m;1-\alpha}\left\{\nu^2\left(\sum_{\ell=1}^{m}\frac{1}{k_1^{(\ell)}}-1\right)^2 + \frac{2\gamma_1}{f_m} + \frac{2\gamma_1}{f_m(f_m+2)}\chi^2_{f_m;1-\alpha}\right\},$$
$$q_{N_1,m}(\alpha) = \chi^2_{f_m;1-\alpha} + \frac{\nu}{N_1}\left(\sum_{\ell=1}^{m}\frac{1}{k_2^{(\ell)}}-1\right)\chi^2_{f_m;1-\alpha}$$
$$+\frac{1}{N_1^2}\chi^2_{f_m;1-\alpha}\left\{\nu^2\left(\sum_{\ell=1}^{m}\frac{1}{k_2^{(\ell)}}-1\right)^2 + \frac{2\gamma_2}{f_m} + \frac{2\gamma_2}{f_m(f_m+2)}\chi^2_{f_m;1-\alpha}\right\},$$
$$k_2^{(\ell)} = \frac{N_1^{(\ell)}}{N_1},$$
$$\gamma_2 = \frac{1}{288}\left[6p(p+1)(p+2)(p+3)\left(\sum_{\ell=1}^{m}\frac{1}{\{k_2^{(\ell)}\}^2}-1\right)\right.$$
$$\left.-\frac{(2p^2+9p+11)^2(2p-1)}{p(p+3)(m-1)}\left(\sum_{\ell=1}^{m}\frac{1}{k_2^{(\ell)}}-1\right)^2\right].$$

3.5 Simulation Studies

We evaluate the accuracy and the asymptotic behaviors of the χ^2 approximations via Monte Carlo simulation (10^6 runs). Now let

$$\begin{aligned}
\alpha_m &= \Pr\{-2\log\lambda_m > \chi^2_{f_m;1-\alpha}\},\\
\alpha_{\rho_{L_m}} &= \Pr\{-2\rho_{L_m}\log\lambda_m > \chi^2_{f_m;1-\alpha}\},\\
\alpha_{\tau_m} &= \Pr\{-2\log\tau_m > \chi^2_{f_m;1-\alpha}\},\\
\alpha_{\phi_m} &= \Pr\{-2\log\phi_m > \chi^2_{f_m;1-\alpha}\},\\
\alpha_{q^*_m} &= \Pr\{-2\log\lambda_m > q^*_m(\alpha)\}.
\end{aligned}$$

In Tables 7, 8, and 9, we provide the simulated upper 100α percentiles of $-2\log\lambda_m$, $-2\rho_{L_m}\log\lambda_m$, $-2\log\tau_m$, and $-2\log\phi_m$ and the approximate upper percentiles of $-2\log\lambda_m$ ($q^*_m(\alpha)$) and the actual type I error rates $\alpha_m, \alpha_{\rho_{L_m}}, \alpha_{\tau_m}, \alpha_{\phi_m}$, and $\alpha_{q^*_m}$; $\alpha = 0.05$;

$$(N_1^{(\ell)}, N_2^{(\ell)}, N_3^{(\ell)}) = \begin{cases} (t, t, t), \\ (t, t/2, t/2),\ t = 20, 40, 80, 160, 320, \\ (t, 2t, 2t), \end{cases}$$

where (p_1, p_2, p_3) is $(4, 4, 4)$.

The simulated values are closer to the upper percentile of the χ^2 distribution when the sample size increases. However, the accuracy of the simulated values is not very good compared with one-sample case, even if the sample size is quite large. In addition, by comparing the Type I error rates $\alpha_m, \alpha_{\rho_{L_m}}, \alpha_{\tau_m}, \alpha_{\phi_m}$, the accuracy of the approximate percentile ($q^*_m(\alpha)$) is the best.

3.6 Numerical Example

In this section, we give an example of test statistics and approximate upper percentiles proposed in this paper. The data consisted of cholesterol values measured during treatment at five time points (baseline, 6 months, 12 months, 20 months, and 24 months) of a placebo group and a high dose group (Wei and Lachin [8]). We used data with values available for up to 24 months, data with values available for up to 20 months, and data with values available for up to 12 months to construct the three-step monotone missing data. That is $m = 2$, $p_1 = 3$, $p_2 = 1$, $p_3 = 1$. For the placebo group ($\ell = 1$), $N_1^{(1)} = 31$, $N_2^{(1)} = 4$, $N_3^{(1)} = 3$, and for the high dose group ($\ell = 2$), $N_1^{(2)} = 36$, $N_2^{(2)} = 7$, $N_3^{(2)} = 12$. Then, LRT statistic and test statistics are

$$\begin{aligned}
-2\log\lambda_m &= 82.201,\ \ -2\rho_{L_m}\log\lambda_m = 75.425,\\
-2\log\tau_m &= 66.542,\ \ -2\log\phi_m = 80.527.
\end{aligned}$$

Table 7 Simulated values for $-2\log\lambda_m$, $-2\rho_{L_m}\log\lambda_m$, $-2\log\tau_m$, and $-2\log\phi_m$ and the approximate values for $-2\log\lambda_m$ and the actual type I error rates α_m, $\alpha_{\rho_{L_m}}$, α_{τ_m}, α_{ϕ_m}, and $\alpha_{q_m^*}$ for $(p_1, p_2, p_3) = (4, 4, 4)$, $m = 2$

Sample size		Upper percentile					Type I error				
$N_1^{(\ell)}$	$N_2^{(\ell)} = N_3^{(\ell)}$	$-2\log\lambda_m$	$-2\rho_{L_m}\log\lambda_m$	$-2\log\tau_m$	$-2\log\phi_m$	$q_m^*(\alpha)$	α_m	$\alpha_{\rho_{L_m}}$	α_{τ_m}	α_{ϕ_m}	$\alpha_{q_m^*}$
$\alpha = 0.05$											
20	20	169.48	131.16	159.56	149.79	153.97	0.842	0.260	0.740	0.591	**0.162**
40	40	133.02	117.98	129.97	126.58	129.75	0.304	0.090	0.252	0.197	**0.072**
80	80	121.89	115.00	120.59	119.10	120.49	0.133	0.063	0.117	0.101	**0.059**
160	160	117.27	113.95	116.66	115.96	116.58	0.082	0.055	0.077	0.071	**0.054**
320	320	115.19	113.56	114.89	114.54	114.80	0.065	0.053	0.062	0.060	**0.052**
20	10	172.43	128.56	161.33	151.16	159.08	0.868	0.220	0.765	0.617	**0.138**
40	20	134.31	117.22	130.76	127.22	131.82	0.329	0.082	0.266	0.208	**0.066**
80	40	122.47	114.68	120.96	119.42	121.41	0.140	0.061	0.121	0.104	**0.057**
160	80	117.60	113.86	116.88	116.14	117.01	0.085	0.055	0.079	0.072	**0.054**
320	160	115.29	113.46	114.94	114.57	115.01	0.066	**0.052**	0.063	0.060	**0.052**
20	40	166.93	132.96	157.96	148.44	150.44	0.815	0.288	0.716	0.567	**0.174**
40	80	131.87	118.45	129.21	125.94	128.22	0.284	0.093	0.239	0.188	**0.074**
80	160	121.37	115.20	120.28	118.85	119.79	0.126	0.065	0.113	0.098	**0.060**
160	320	117.06	114.08	116.55	115.87	116.25	0.080	0.057	0.076	0.070	**0.055**
320	640	114.94	113.48	114.70	114.37	114.64	0.063	**0.052**	0.061	0.058	**0.052**

The closest value to α from among α_m, $\alpha_{\rho_{L_m}}$, α_{τ_m}, α_{ϕ_m}, and $\alpha_{q_m^*}$ of each row is in bold. $\chi^2_{f_m;0.95} = 113.15$, $f_m = 90$

Table 8 Simulated values for $-2\log\lambda_m$, $-2\rho_{L_m}\log\lambda_m$, $-2\log\tau_m$, and $-2\log\phi_m$ and the approximate values for $-2\log\lambda_m$ and the actual type I error rates α_m, $\alpha_{\rho_{L_m}}$, α_{τ_m}, α_{ϕ_m}, and $\alpha_{q_m^*}$ for $(p_1, p_2, p_3) = (4, 4, 4)$, $m = 3$

Sample size		Upper percentile					Type I error				
$N_1^{(\ell)}$	$N_2^{(\ell)} = N_3^{(\ell)}$	$-2\log\lambda_m$	$-2\rho_{L_m}\log\lambda_m$	$-2\log\tau_m$	$-2\log\phi_m$	$q_m^*(\alpha)$	α_m	$\alpha_{\rho_{L_m}}$	α_{τ_m}	α_{ϕ_m}	$\alpha_{q_m^*}$
$\alpha = 0.05$											
20	20	301.38	240.80	285.82	265.97	277.83	0.941	0.325	0.869	0.684	**0.192**
40	40	244.74	220.15	239.73	233.05	239.35	0.397	0.099	0.324	0.233	**0.077**
80	80	226.75	215.36	224.60	221.67	224.40	0.160	0.066	0.139	0.112	**0.062**
160	160	219.18	213.67	218.17	216.79	217.99	0.092	0.057	0.085	0.075	**0.056**
320	320	215.62	212.91	215.13	214.46	215.06	0.068	**0.053**	0.065	0.061	**0.053**
20	10	306.27	237.02	288.81	268.26	286.03	0.957	0.275	0.890	0.712	**0.161**
40	20	246.91	219.00	241.11	234.07	242.74	0.429	0.090	0.343	0.245	**0.070**
80	40	227.78	214.90	225.26	222.16	225.91	0.171	0.064	0.145	0.116	**0.059**
160	80	219.62	213.42	218.44	216.98	218.71	0.096	0.056	0.087	0.077	**0.055**
320	160	215.88	212.82	215.30	214.59	215.40	0.070	0.053	0.066	0.062	**0.052**
20	40	297.14	243.39	283.18	264.08	272.10	0.923	0.358	0.848	0.657	**0.207**
40	80	242.78	220.82	238.46	232.08	236.85	0.367	0.105	0.306	0.221	**0.081**
80	160	225.76	215.55	223.94	221.16	223.24	0.151	0.068	0.133	0.108	**0.063**
160	320	218.69	213.75	217.84	216.53	217.44	0.088	0.057	0.082	0.074	**0.056**
320	640	215.41	212.98	215.00	214.36	214.79	0.067	**0.053**	0.064	0.061	**0.053**

The closest value to α from among α_m, $\alpha_{\rho_{L_m}}$, α_{τ_m}, α_{ϕ_m}, and $\alpha_{q_m^*}$ of each row is in bold. $\chi^2_{f_m;0.95} = 212.30$, $f_m = 180$

Table 9 Simulated values for $-2\log\lambda_m$, $-2\rho_{L_m}\log\lambda_m$, $-2\log\tau_m$, and $-2\log\phi_m$ and the approximate values for $-2\log\lambda_m$ and the actual type I error rates α_m, $\alpha_{\rho_{L_m}}$, α_{τ_m}, α_{ϕ_m}, and $\alpha_{q_m^*}$ for $(p_1, p_2, p_3) = (4, 4, 4)$, $m = 4$

Sample size		Upper percentile					Type I error				
$N_1^{(\ell)}$	$N_2^{(\ell)} = N_3^{(\ell)}$	$-2\log\lambda_m$	$-2\rho_{L_m}\log\lambda_m$	$-2\log\tau_m$	$-2\log\phi_m$	$q_m^*(\alpha)$	α_m	$\alpha_{\rho_{L_m}}$	α_{τ_m}	α_{ϕ_m}	$\alpha_{q_m^*}$
$\alpha = 0.05$											
20	20	428.72	347.94	408.06	378.36	397.27	0.977	0.384	0.933	0.750	**0.222**
40	40	353.35	320.06	346.59	336.56	345.88	0.471	0.108	0.384	0.261	**0.083**
80	80	329.04	313.54	326.12	321.74	325.75	0.183	0.069	0.156	0.121	**0.064**
160	160	318.65	311.14	317.27	315.20	317.07	0.099	0.058	0.091	0.078	**0.056**
320	320	313.89	310.19	313.22	312.20	313.08	0.071	0.054	0.068	0.063	**0.053**
20	10	435.55	343.22	412.30	381.39	408.27	0.985	0.329	0.947	0.777	**0.186**
40	20	356.23	318.48	348.37	337.85	350.45	0.508	0.098	0.407	0.276	**0.074**
80	40	330.33	312.83	326.91	322.27	327.80	0.196	0.066	0.163	0.125	**0.060**
160	80	319.30	310.84	317.68	315.50	318.04	0.104	0.056	0.093	0.080	**0.055**
320	160	314.26	310.10	313.47	312.41	313.55	0.073	**0.053**	0.069	0.064	**0.053**
20	40	422.86	351.15	404.36	375.76	389.55	0.967	0.422	0.918	0.724	**0.241**
40	80	350.62	320.89	344.80	335.29	342.50	0.436	0.114	0.362	0.248	**0.086**
80	160	327.62	313.73	325.17	321.02	324.18	0.170	0.070	0.148	0.115	**0.065**
160	320	317.98	311.24	316.83	314.88	316.32	0.095	0.058	0.088	0.077	**0.057**
320	640	313.58	310.26	313.03	312.08	312.71	0.070	0.054	0.067	0.062	**0.053**

The closest value to α from among α_m, $\alpha_{\rho_{L_m}}$, α_{τ_m}, α_{ϕ_m}, and $\alpha_{q_m^*}$ of each row is in bold. $\chi^2_{f_m;0.95} = 309.33$, $f_m = 270$

And, approximate upper percentile is

$$q_m^*(0.05) = 34.422,\ \ q_m^*(0.01) = 41.196,$$

and $\chi^2_{20;0.95} = 31.410$, $\chi^2_{20;0.99} = 37.566$. Thus, the null hypothesis is rejected for all test statistics and approximate upper percentile.

4 Conclusions

We discussed simultaneous tests for mean vectors and covariance matrices with three-step monotone missing data for a one-sample and a multi-sample problem. For a one-sample problem, we proposed two test statistics ($-2\log\tau_1$, $-2\log\phi_1$) by decomposing the LR and correcting it by extracting the LR of the simultaneous test and the test of the variance in the case of complete data. We also proposed a test statistic ($-2\rho_{L_1}\log\lambda_1$) via linear interpolation. In addition, we provided an approximate upper 100α percentile ($q_1^*(\alpha)$). Based on the simulation results, the test statistic $-2\log\phi_1$, which was modified only for the LR part of the test of the variance for the complete data, gave the most accurate results. Similarly, for a multi-sample problem, we proposed three test statistics ($-2\rho_{L_m}\log\lambda_m$, $-2\log\tau_m$, $-2\log\phi_m$) and an approximate upper percentile ($q_m^*(\alpha)$). Furthermore, based on the simulation results, the approximate upper 100α percentile $q_m^*(\alpha)$ is the most accurate. Finally, we gave an example of the proposed test statistics. The results of this paper can be extended to the k-step monotone missing data. We are currently investigating this problem.

Acknowledgements The authors would like to thank the referee for the helpful comments and suggestions. The second author's research is partly supported by a Grant-in-Aid for Early-Career Scientists (19K20225, 22K13961).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix

Derivation of Theorem 1

Following the derivation of MLEs of the mean vector and the covariance matrix with two-step monotone missing data in Kanda and Fujikoshi [5], we consider the transformation matrix

$$\boldsymbol{Z} = \left(\begin{array}{cc|c} \boldsymbol{I}_{p_1} & \boldsymbol{O} & \multirow{2}{*}{$\boldsymbol{O}$} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{I}_{p_2} & \\ \hline \multicolumn{2}{c|}{-\boldsymbol{\Sigma}_{3(12)}\boldsymbol{\Sigma}_{(12)(12)}^{-1}} & \boldsymbol{I}_{p_3} \end{array}\right)$$

for the three-step monotone missing data case. In this case, the transformed vector $\boldsymbol{y}_j$ is

$$\boldsymbol{y}_j = \boldsymbol{Z}\boldsymbol{x}_j = \begin{pmatrix} \boldsymbol{x}_{1j} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_{1j} + \boldsymbol{x}_{2j} \\ -\boldsymbol{\Sigma}_{3(12)}\boldsymbol{\Sigma}_{(12)(12)}^{-1}\begin{pmatrix} \boldsymbol{x}_{1j} \\ \boldsymbol{x}_{2j} \end{pmatrix} + \boldsymbol{x}_{3j} \end{pmatrix}.$$

The transformed parameters are defined as

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2 \\ -\boldsymbol{\Sigma}_{3(12)}\boldsymbol{\Sigma}_{(12)(12)}^{-1}\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} + \boldsymbol{\mu}_3 \end{pmatrix},$$

$$\boldsymbol{\Delta} = \left(\begin{array}{cc|c} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} & \boldsymbol{\Delta}_{13} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} & \boldsymbol{\Delta}_{23} \\ \hline \boldsymbol{\Delta}_{31} & \boldsymbol{\Delta}_{32} & \boldsymbol{\Delta}_{33} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Delta}_{(12)(12)} & \boldsymbol{\Delta}_{(12)3} \\ \hline \boldsymbol{\Delta}_{3(12)} & \boldsymbol{\Delta}_{33} \end{array}\right),$$

where

$$\begin{aligned} \boldsymbol{\Delta}_{11} &= \boldsymbol{\Sigma}_{11}, \\ \boldsymbol{\Delta}_{12} &= \boldsymbol{\Delta}_{21}' = \boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}, \\ \boldsymbol{\Delta}_{22} &= \boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}, \\ \boldsymbol{\Delta}_{(12)3} &= \boldsymbol{\Delta}_{3(12)}' = \boldsymbol{\Sigma}_{(12)(12)}^{-1}\boldsymbol{\Sigma}_{(12)3}, \\ \boldsymbol{\Delta}_{33} &= \boldsymbol{\Sigma}_{33\cdot 12} = \boldsymbol{\Sigma}_{33} - \boldsymbol{\Sigma}_{3(12)}\boldsymbol{\Sigma}_{(12)(12)}^{-1}\boldsymbol{\Sigma}_{(12)3}. \end{aligned}$$

Then, the likelihood function after the transformation can be written as

$$\begin{aligned} L(\boldsymbol{\eta}, \boldsymbol{\Delta}) = Const.&|\boldsymbol{\Delta}_{11}|^{-\frac{1}{2}N}|\boldsymbol{\Delta}_{22}|^{-\frac{1}{2}(N_1+N_2)}|\boldsymbol{\Delta}_{33}|^{-\frac{1}{2}N_1} \\ &\times \exp\left\{-\frac{1}{2}\sum_{j=1}^{N}(\boldsymbol{y}_{1j} - \boldsymbol{\eta}_1)'\boldsymbol{\Delta}_{11}^{-1}(\boldsymbol{y}_{1j} - \boldsymbol{\eta}_1)\right\} \\ &\times \exp\left\{-\frac{1}{2}\sum_{j=1}^{N_1+N_2}(\boldsymbol{y}_{2j} - \boldsymbol{\eta}_2)'\boldsymbol{\Delta}_{22}^{-1}(\boldsymbol{y}_{2j} - \boldsymbol{\eta}_2)\right\} \\ &\times \exp\left\{-\frac{1}{2}\sum_{j=1}^{N_1}(\boldsymbol{y}_{3j} - \boldsymbol{\eta}_3)'\boldsymbol{\Delta}_{33}^{-1}(\boldsymbol{y}_{3j} - \boldsymbol{\eta}_3)\right\}. \end{aligned}$$

We note that the pair $(\boldsymbol{\eta}, \boldsymbol{\Delta})$ is in one-to-one correspondence with $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The MLEs of $\boldsymbol{\eta}$ and $\boldsymbol{\Delta}$ ($\widehat{\boldsymbol{\eta}}$ and $\widehat{\boldsymbol{\Delta}}$) can be obtained by differentiating the log likelihood function $\log L(\boldsymbol{\eta}, \boldsymbol{\Delta})$ with respect to $\boldsymbol{\eta}$ and $\boldsymbol{\Delta}$, respectively. Calculating

$$\lambda_1 = \frac{L(\mathbf{0}, \boldsymbol{I}_p)}{L(\widehat{\boldsymbol{\eta}}, \widehat{\boldsymbol{\Delta}})}$$

yields (3).

References

1. Hao J, Krishnamoorthy K (2001) Inferences on a normal covariance matrix and generalized variance with monotone missing data. J Multivar Anal 78:62–82
2. Hosoya M, Seo T (2015) Simultaneous testing of the mean vector and the covariance matrix with two-step monotone missing data. SUT J Math 51:83–98
3. Hosoya M, Seo T (2016) On the likelihood ratio test for the equality of multivariate normal populations with two-step monotone missing data. J Stat Theory Pract 10:673–692
4. Jinadasa KG, Tracy DS (1992) Maximum likelihood estimation for multivariate normal distribution with monotone sample. Commun Stat Theory Methods 21:41–50
5. Kanda T, Fujikoshi Y (1998) Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data. Am J Math Manag Sci 18:161–190
6. Muirhead RJ (2005) Aspects of multivariate statistical theory. John Wiley & Sons Inc, Hoboken
7. Srivastava MS (2002) Methods of multivariate statistics. John Wiley & Sons Inc, New York
8. Wei LJ, Lachin JM (1984) Two-sample asymptotically distribution-free tests for incomplete multivariate observations. J Am Stat Assoc 79:653–661
9. Yagi A, Yamaguchi R, Seo T (2016) Simultaneous testing of mean vectors and covariance matrices with monotone missing data, Technical Report No.16-02, Statistical Research Group, Hiroshima University, Hiroshima, Japan
10. K., Krishnamoorthy Maruthy K., Pannala (1999) Confidence estimation of a normal mean vector with incomplete data Abstract Canadian Journal of Statistics 27(2) 395–407 https://doi.org/10.2307/3315648