Modified Likelihood Ratio Test for Simultaneous Testing of Mean Vectors and Covariance Matrices with Missing Data

Remi Nomura*, Ayaka Yagi* and Takashi Seo*

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

Abstract

In this paper, we consider testing equality of multivariate normal populations when data is incomplete without condition of missing patterns. We obtain the likelihood ratio test (LRT) statistic and maximum likelihood estimators (MLEs) with iterative method. The limit distribution of likelihood ratio test statistics is the chi-square distribution. We also propose the modified likelihood ratio test (MLRT) statistic which improves chi-squared approximation using linear interpolation. Finally, we investigate the accuracy of the approximations by Monte Carlo simulation.

Key Words and Phrases: Likelihood ratio test statistic, Maximum likelihood estimator, Iterative method, Multi-sample problem, Linear interpolation.

1 Introduction

We consider the problem of testing the equality of mean vectors and covariance matrices in a multi-sample problem including missing data which are monotone and general type from multivariate normal population Let $\boldsymbol{x}_1^{(\ell)}, \boldsymbol{x}_2^{(\ell)}, \ldots, \boldsymbol{x}_{n^{(\ell)}}^{(\ell)}$ be independently and identically distributed as the p dimensional normal distribution $N_p(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)}), \ell = 1, 2, \ldots, m$. We consider the following hypothesis,

$$H_0 : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} = \cdots = \boldsymbol{\mu}^{(m)}, \boldsymbol{\Sigma}^{(1)} = \boldsymbol{\Sigma}^{(2)} = \cdots = \boldsymbol{\Sigma}^{(m)} \text{ vs. } H_1 : \text{not } H_0.$$

Under complete data, the LRT statistic and its modified LRT statistic are given (see Muirhead (1982) and Srivastava (2002)). Further, in the case of univariate data, Zhang et al. (2012) discussed and gave the exact likelihood ratio test. On the other hand, for simultaneous testing under missing data, Hao and Krishnamoorthy (2001), Hosoya and Seo (2015, 2016) and others gave the LRT statistics and its approximate upper percentile when the data have monotone pattern that is missing observations. Specifically, Hao and Krishnamoorthy (2001) and Hosoya and Seo (2015) discussed simultaneous testing of the mean vector and the covariance matrix for two-step monotone missing data in a one-sample problem, whereas Hosoya and Seo (2016) discussed that of multi-sample cases for two-step monotone missing data. The multi-population under the condition of general k-step monotone missing data was discussed by Yagi, Yamaguchi and Seo (2016). Thus, simultaneous testing for monotone missing data have been discussed, but simultaneous testing for missing data with general of missing data pattern have not been discussed yet. Therefore, in this paper, we consider simultaneous testing for missing data in a multi-sample problem without of missing pattern, that is, the problem of testing the equality of mean vectors and covariance matrices in a multi-sample problem including missing data from multivariate normal population. We give the likelihood ratio test, introduce its LRT statistics, and propose MLRT statistic.

2 MLE and iterative method

Suppose that there are $K^{(\ell)}$ groups of the incomplete data containing $n_1^{(\ell)}, n_2^{(\ell)} \dots, n_{K^{(\ell)}}^{(\ell)}$ observations from ℓth population for $\ell = 1, 2, \dots, m$. Let $\boldsymbol{x}_{k1}^{(\ell)}, \boldsymbol{x}_{k2}^{(\ell)}, \dots, \boldsymbol{x}_{kn_k^{(\ell)}}^{(\ell)}$ be p_1 dimensional sample vectors in the kth group for $k = 1, 2, \dots, K^{(\ell)}$ and $n^{(\ell)} = \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}$. In order to remove missing data from observation vectors, we transform ones by using incidence matrix defined by Srivastava (1985) and Srivastava and Carter (1986). That is, we multiply $\boldsymbol{x}_{kj}^{(\ell)}$ by incidence matrix $\mathbf{B}_k^{(\ell)}$ which is a $p_k^{(\ell)} \times p_1$ matrix and transform to $\boldsymbol{z}_{kj}^{(\ell)}$. For example,

$$\mathbf{B}_k^{(\ell)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

for $\boldsymbol{x}_{kj}^{(\ell)} = (x_{kj1}^{(\ell)}, x_{kj2}^{(\ell)}, *, x_{kj4}^{(\ell)}, *, x_{kj6}^{(\ell)}, x_{kj7}^{(\ell)})' : 7 \times 1$, and $\boldsymbol{x}_{kj}^{(\ell)}$ would be transformed into $\boldsymbol{z}_{kj}^{(\ell)} = (x_{kj1}^{(\ell)}, x_{kj2}^{(\ell)}, x_{kj4}^{(\ell)}, x_{kj6}^{(\ell)}, x_{kj7}^{(\ell)})' : 5 \times 1$. We assume that $\boldsymbol{z}_{kj}^{(\ell)} \overset{iid}{\sim} N_{p_k^{(\ell)}}(\boldsymbol{\mu}^{(\ell)}, \boldsymbol{\Sigma}^{(\ell)}), j = 1, 2, \dots, n_k^{(\ell)}$. Then we construct the likelihood function based on $\boldsymbol{z}_{kj}^{(\ell)}$ as follows:

$$L(\boldsymbol{\mu}^{(1)}, \boldsymbol{\mu}^{(2)}, \dots, \boldsymbol{\mu}^{(m)}, \boldsymbol{\Sigma}^{(1)}, \boldsymbol{\Sigma}^{(2)}, \dots, \boldsymbol{\Sigma}^{(m)}) = Const. \prod_{\ell=1}^{m} \prod_{k=1}^{K^{(\ell)}} |\boldsymbol{\Lambda}_k^{(\ell)}|^{-\frac{1}{2} n_k^{(\ell)}}$$
$$\times \exp\Big\{ -\frac{1}{2} \sum_{\alpha=1}^{n_k} (\boldsymbol{z}_{k\alpha}^{(\ell)} - \mathbf{B}_k^{(\ell)} \boldsymbol{\mu}^{(\ell)})' \boldsymbol{\Lambda}_k^{(\ell)-1} (\boldsymbol{z}_{k\alpha}^{(\ell)} - \mathbf{B}_k^{(\ell)} \boldsymbol{\mu}^{(\ell)}) \Big\},$$

where

$$\boldsymbol{\Lambda}_k^{(\ell)} = \mathbf{B}_k^{(\ell)} \boldsymbol{\Sigma}^{(\ell)} \mathbf{B}_k^{(\ell)\prime}.$$

By differentiating it, the MLEs of $\boldsymbol{\mu}^{(\ell)}$ are given by

$$\widehat{\boldsymbol{\mu}}^{(\ell)} = \left[\sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)} \mathbf{B}_k^{(\ell)\prime} \widehat{\boldsymbol{\Lambda}}_k^{(\ell)-1} \mathbf{B}_k^{(\ell)}\right]^{-1} \left[\sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)} \mathbf{B}_k^{(\ell)\prime} \widehat{\boldsymbol{\Lambda}}_k^{(\ell)-1} \bar{\boldsymbol{z}}_k^{(\ell)}\right], \ell = 1, 2, \ldots, m,$$

and $\widehat{\boldsymbol{\Sigma}}^{(\ell)}$, the MLEs of $\boldsymbol{\Sigma}^{(\ell)}$ satisfy

$$\mathbf{H}^{(\ell)} = \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)} \mathbf{B}_k^{(\ell)\prime} \widehat{\boldsymbol{\Lambda}}_k^{(\ell)-1} \mathbf{B}_k^{(\ell)} - \sum_{k=1}^{K^{(\ell)}} \mathbf{B}_k^{(\ell)\prime} \widehat{\boldsymbol{\Lambda}}_k^{(\ell)-1} \widehat{\mathbf{U}}_k^{(\ell)} \widehat{\mathbf{U}}_k^{(\ell)\prime} \widehat{\boldsymbol{\Lambda}}_k^{(\ell)-1} \mathbf{B}_k^{(\ell)} = \boldsymbol{O}, \ell = 1, 2, \ldots, m, \tag{2.1}$$

where

$$\begin{aligned}
\widehat{\boldsymbol{\Lambda}}_k^{(\ell)} &= \mathbf{B}_k^{(\ell)} \widehat{\boldsymbol{\Sigma}}^{(\ell)} \mathbf{B}_k^{(\ell)\prime}, \quad \widehat{\mathbf{U}}_k^{(\ell)} = \mathbf{Z}_k^{(\ell)} - \mathbf{B}_k^{(\ell)} \widehat{\boldsymbol{\mu}}^{(\ell)} \boldsymbol{e}'_{n_k^{(\ell)}}, \quad \bar{\boldsymbol{z}}_k^{(\ell)} = \frac{1}{n_k^{(\ell)}} \sum_{\alpha=1}^{n_k^{(\ell)}} \boldsymbol{z}_{k\alpha}^{(\ell)}, \\
\mathbf{Z}_k^{(\ell)} &= (\boldsymbol{z}_{k1}^{(\ell)}, \boldsymbol{z}_{k2}^{(\ell)}, \ldots, \boldsymbol{z}_{kn_k^{(\ell)}}^{(\ell)}), \quad \boldsymbol{e}_{n_k^{(\ell)}} = (1, 1, \ldots, 1)'.
\end{aligned}$$

Likewise, under H_0, the MLE of $\boldsymbol{\mu}(= \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} = \cdots = \boldsymbol{\mu}^{(m)})$ is given by

$$\widetilde{\boldsymbol{\mu}} = \left[\sum_{\ell=1}^{m} \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)} \mathbf{B}_k^{(\ell)\prime} \widetilde{\boldsymbol{\Lambda}}_k^{(\ell)-1} \mathbf{B}_k^{(\ell)}\right]^{-1} \left[\sum_{\ell=1}^{m} \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)} \mathbf{B}_k^{(\ell)\prime} \widetilde{\boldsymbol{\Lambda}}_k^{(\ell)-1} \bar{\boldsymbol{z}}_k^{(\ell)}\right], \tag{2.2}$$

and $\widetilde{\boldsymbol{\Sigma}}$, the MLE of $\boldsymbol{\Sigma}(= \boldsymbol{\Sigma}^{(1)} = \boldsymbol{\Sigma}^{(2)} = \cdots = \boldsymbol{\Sigma}^{(m)})$, satisfies

$$\sum_{\ell=1}^{m} \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)} \mathbf{B}_k^{(\ell)\prime} \widetilde{\boldsymbol{\Lambda}}_k^{(\ell)-1} \mathbf{B}_k^{(\ell)} = \sum_{\ell=1}^{m} \sum_{k=1}^{K^{(\ell)}} \mathbf{B}_k^{(\ell)\prime} \widetilde{\boldsymbol{\Lambda}}_k^{(\ell)-1} \widetilde{\mathbf{U}}_k^{(\ell)} \widetilde{\mathbf{U}}_k^{(\ell)\prime} \widetilde{\boldsymbol{\Lambda}}_k^{(\ell)-1} \mathbf{B}_k^{(\ell)},$$

where

$$\widetilde{\boldsymbol{\Lambda}}_k^{(\ell)} = \mathbf{B}_k^{(\ell)} \widetilde{\boldsymbol{\Sigma}} \mathbf{B}_k^{(\ell)\prime}, \quad \widetilde{\mathbf{U}}_k^{(\ell)} = \mathbf{Z}_k^{(\ell)} - \mathbf{B}_k^{(\ell)} \widetilde{\boldsymbol{\mu}} \boldsymbol{e}'_{n_k^{(\ell)}}.$$

We want to solve these likelihood equations to obtain $\widehat{\boldsymbol{\Sigma}}^{(\ell)}$ and $\widetilde{\boldsymbol{\Sigma}}$, but it is difficult to obtain the exact solution of them. Therefore, the solution is given by using Newton-Raphson method; see Srivastava and Carter (1986). That is, suppose that $\boldsymbol{\Delta}$ is the update matrix and i is the number of iterations, we substitute

$$\boldsymbol{\Sigma}_{i+1} = \boldsymbol{\Sigma}_i + \boldsymbol{\Delta}_{i+1}$$

into (2.1). When introducing $\widehat{\boldsymbol{\Sigma}}^{(\ell)}$, define

$$\widehat{\mathbf{Q}}_i^{(\ell)} = \sum_{k=1}^{K^{(\ell)}} (n_k^{(\ell)} \widehat{\mathbf{D}}_{ki}^{(\ell)} \otimes \widehat{\mathbf{D}}_{ki}^{(\ell)} - \widehat{\mathbf{D}}_{ki}^{(\ell)} \otimes \widehat{\mathbf{F}}_{ki}^{(\ell)} - \widehat{\mathbf{F}}_{ki}^{(\ell)} \otimes \widehat{\mathbf{D}}_{ki}^{(\ell)}),$$

where $\mathbf{A} \otimes \mathbf{B}$ denotes the kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ defined by $\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B})$,

$$\widehat{\mathbf{D}}_{ki}^{(\ell)} = \mathbf{B}_k^{(\ell)\prime} (\mathbf{B}_k^{(\ell)} \widehat{\boldsymbol{\Sigma}}_i^{(\ell)} \mathbf{B}_k^{(\ell)\prime})^{-1} \mathbf{B}_k^{(\ell)},$$

and

$$\widehat{\mathbf{F}}_{ki}^{(\ell)} = \mathbf{B}_k^{(\ell)\prime}(\mathbf{B}_k^{(\ell)}\widehat{\mathbf{\Sigma}}_i^{(\ell)}\mathbf{B}_k^{(\ell)\prime})^{-1}\widehat{\mathbf{U}}_k^{(\ell)}\widehat{\mathbf{U}}_k^{(\ell)\prime}(\mathbf{B}_k^{(\ell)}\widehat{\mathbf{\Sigma}}_i^{(\ell)}\mathbf{B}_k^{(\ell)\prime})^{-1}\mathbf{B}_k^{(\ell)}.$$

For any matrix $\mathbf{A} = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_q)'$, we define $\text{vec}(\mathbf{A}) = (\boldsymbol{a}_1', \dots, \boldsymbol{a}_q')'$. Then (2.1) can be approximately written as

$$\widehat{\mathbf{Q}}_i^{(\ell)}\text{vec}(\widehat{\mathbf{\Delta}}_{i+1}^{(\ell)}) = \text{vec}(\widehat{\mathbf{E}}_i^{(\ell)}), \tag{2.3}$$

where

$$\widehat{\mathbf{E}}_i^{(\ell)} = \sum_{k=1}^{K^{(\ell)}}(n_k^{(\ell)}\widehat{\mathbf{D}}_{ki}^{(\ell)} - \widehat{\mathbf{F}}_{ki}^{(\ell)}).$$

(2.3) can be introduced the following method. Because $\mathbf{H}^{(\ell)} = \boldsymbol{O}$, we consider $\text{vec}(\mathbf{H}^{(\ell)}) = \mathbf{0}$. Substituting $\mathbf{\Sigma}_{i+1} = \mathbf{\Sigma}_i + \mathbf{\Delta}_{i+1}$ into $\mathbf{\Lambda}_{ki}^{(\ell)}(= \mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_i^{(\ell)}\mathbf{B}_k^{(\ell)\prime})$,

$$\mathbf{\Lambda}_{ki}^{(\ell)} = \mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)\prime}[\boldsymbol{I}_{p_k^{(\ell)}} + (\mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)\prime})^{-1}\mathbf{B}_k^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{B}_k^{(\ell)\prime}].$$

Thus,

$$\mathbf{\Lambda}_{ki}^{(\ell)-1} \fallingdotseq [\boldsymbol{I}^{(\ell)} - (\mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)\prime})^{-1}\mathbf{B}_k^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{B}_k^{(\ell)\prime}](\mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)\prime})^{-1},$$

because $(\boldsymbol{I}+\boldsymbol{A})^{-1}$ is approximately equal to $\boldsymbol{I}-\boldsymbol{A}$. Therefore, the first term of $\mathbf{H}^{(\ell)}$except $\sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}$can be written as

$$\mathbf{B}_k^{(\ell)\prime}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{B}_k^{(\ell)} = \mathbf{D}_{ki-1}^{(\ell)} - \mathbf{D}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)},$$

where

$$\mathbf{D}_{ki-1}^{(\ell)} = \mathbf{B}_k^{(\ell)\prime}(\mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)\prime})^{-1}\mathbf{B}_k^{(\ell)}.$$

Using the fact that $\text{vec}(\boldsymbol{AXC}) = (\boldsymbol{C}' \otimes \boldsymbol{A})\text{vec}(\boldsymbol{X})$ holds, we can write

$$\text{vec}[\mathbf{B}_k^{(\ell)\prime}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{B}_k^{(\ell)}] = \text{vec}(\mathbf{D}_{ki-1}^{(\ell)}) - (\mathbf{D}_{ki-1}^{(\ell)} \otimes \mathbf{D}_{ki-1}^{(\ell)})\text{vec}(\mathbf{\Delta}_i^{(\ell)}).$$

Hence,

$$\sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}\text{vec}[\mathbf{B}_k^{(\ell)\prime}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{B}_k^{(\ell)}] = \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}\text{vec}(\mathbf{D}_{ki-1}^{(\ell)}) - \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}(\mathbf{D}_{ki-1}^{(\ell)} \otimes \mathbf{D}_{ki-1}^{(\ell)})\text{vec}(\mathbf{\Delta}_i^{(\ell)}).$$

The second term of $\mathbf{H}^{(\ell)}$ can be organized

$$\begin{aligned} &-\sum_{k=1}^{K^{(\ell)}}\mathbf{B}_k^{(\ell)\prime}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{U}_k^{(\ell)}\mathbf{U}_k^{(\ell)\prime}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{B}_k^{(\ell)} \\ &\quad = -\sum_{k=1}^{K^{(\ell)}}\mathbf{F}_{ki-1}^{(\ell)} + \sum_{k=1}^{K^{(\ell)}}\mathbf{F}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)} \\ &\qquad + \sum_{k=1}^{K^{(\ell)}}\mathbf{D}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{F}_{ki-1}^{(\ell)} - \sum_{k=1}^{K^{(\ell)}}\mathbf{D}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{F}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)}, \end{aligned}$$

where

$$\mathbf{F}_{ki-1}^{(\ell)} = \mathbf{B}_k^{(\ell)'}(\mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)'})^{-1}\mathbf{U}_k^{(\ell)}\mathbf{U}_k^{(\ell)'}(\mathbf{B}_k^{(\ell)}\mathbf{\Sigma}_{i-1}^{(\ell)}\mathbf{B}_k^{(\ell)'})^{-1}\mathbf{B}_k^{(\ell)}.$$

For the fact that $\mathbf{\Delta}^{(\ell)} \otimes \mathbf{\Delta}^{(\ell)} \fallingdotseq \boldsymbol{O}$, note that

$$\begin{aligned}\mathrm{vec}(\mathbf{D}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{F}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)}) &= \mathbf{D}_{ki-1}^{(\ell)} \otimes \mathbf{D}_{ki-1}^{(\ell)}(\mathbf{\Delta}_i^{(\ell)} \otimes \mathbf{\Delta}_i^{(\ell)}\mathrm{vec}(\mathbf{F}_{ki-1}^{(\ell)})) \\ &\fallingdotseq \mathbf{0}.\end{aligned}$$

Then, we can write

$$\begin{aligned}&\mathrm{vec}(-\sum_{k=1}^{K^{(\ell)}}\mathbf{B}_k^{(\ell)'}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{U}_k^{(\ell)}\mathbf{U}_k^{(\ell)'}\mathbf{\Lambda}_{ki}^{(\ell)-1}\mathbf{B}_k^{(\ell)}) \\ &\fallingdotseq -\mathrm{vec}(\sum_{k=1}^{K^{(\ell)}}\mathbf{F}_{ki-1}^{(\ell)}) + \mathrm{vec}(\sum_{k=1}^{K^{(\ell)}}\mathbf{F}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)}) + \mathrm{vec}(\sum_{k=1}^{K^{(\ell)}}\mathbf{D}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{F}_{ki-1}^{(\ell)}).\end{aligned}$$

Since $\mathbf{H}^{(\ell)} = \boldsymbol{O}$,

$$\begin{aligned}\mathbf{0} &= \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}\mathrm{vec}(\mathbf{D}_{ki-1}^{(\ell)}) - \sum_{k=1}^{K^{(\ell)}} n_k^{(\ell)}(\mathbf{D}_{ki-1}^{(\ell)} \otimes \mathbf{D}_{ki-1}^{(\ell)})\mathrm{vec}(\mathbf{\Delta}_i^{(\ell)}) \\ &\quad - \mathrm{vec}(\sum_{k=1}^{K^{(\ell)}}\mathbf{F}_{ki-1}^{(\ell)}) + \mathrm{vec}(\sum_{k=1}^{K^{(\ell)}}\mathbf{F}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)}) + \mathrm{vec}(\sum_{k=1}^{K^{(\ell)}}\mathbf{D}_{ki-1}^{(\ell)}\mathbf{\Delta}_i^{(\ell)}\mathbf{F}_{ki-1}^{(\ell)}).\end{aligned}$$

We note that

$$\begin{aligned}\mathbf{Q}_{i-1}^{(\ell)} &= \sum_{k=1}^{K^{(\ell)}}(n_k^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)} \otimes \mathbf{D}_{ki-1}^{(\ell)} - \mathbf{D}_{ki-1}^{(\ell)} \otimes \mathbf{F}_{ki-1}^{(\ell)} - \mathbf{F}_{ki-1}^{(\ell)} \otimes \mathbf{D}_{ki-1}^{(\ell)}), \\ \mathbf{E}_{i-1}^{(\ell)} &= \sum_{k=1}^{K^{(\ell)}}(n_k^{(\ell)}\mathbf{D}_{ki-1}^{(\ell)} - \mathbf{F}_{ki-1}^{(\ell)}).\end{aligned}$$

Therefore, $\boldsymbol{Q}_{i-1}^{(\ell)}\mathrm{vec}(\mathbf{\Delta}_i^{(\ell)}) = \mathrm{vec}(\boldsymbol{E}_{i-1}^{(\ell)})$.

Similarly, the MLEs under H_0 can be also obtained using (2.2) and

$$\widetilde{\mathbf{Q}}_i\mathrm{vec}(\widetilde{\mathbf{\Delta}}_{i+1}) = \mathrm{vec}(\widetilde{\mathbf{E}}_i),$$

where

$$\begin{aligned}\widetilde{\mathbf{Q}}_i &= \sum_{\ell=1}^{m}\sum_{k=1}^{K^{(\ell)}}(n_k^{(\ell)}\widetilde{\mathbf{D}}_{ki}^{(\ell)} \otimes \widetilde{\mathbf{D}}_{ki}^{(\ell)} - \widetilde{\mathbf{D}}_{ki}^{(\ell)} \otimes \widetilde{\mathbf{F}}_{ki}^{(\ell)} - \widetilde{\mathbf{F}}_{ki}^{(\ell)} \otimes \widetilde{\mathbf{D}}_{ki}^{(\ell)}), \\ \widetilde{\mathbf{D}}_{ki}^{(\ell)} &= \mathbf{B}_k^{(\ell)'}(\mathbf{B}_k^{(\ell)}\widetilde{\mathbf{\Sigma}}_i\mathbf{B}_k^{(\ell)'})^{-1}\mathbf{B}_k^{(\ell)}, \\ \widetilde{\mathbf{F}}_{ki}^{(\ell)} &= \mathbf{B}_k^{(\ell)'}(\mathbf{B}_k^{(\ell)}\widetilde{\mathbf{\Sigma}}_i\mathbf{B}_k^{(\ell)'})^{-1}\widetilde{\mathbf{U}}_k^{(\ell)}\widetilde{\mathbf{U}}_k^{(\ell)'}(\mathbf{B}_k^{(\ell)}\widetilde{\mathbf{\Sigma}}_i\mathbf{B}_k^{(\ell)'})^{-1}\mathbf{B}_k^{(\ell)}, \\ \widetilde{\mathbf{E}}_i &= \sum_{\ell=1}^{m}\sum_{k=1}^{K^{(\ell)}}(n_k^{(\ell)}\widetilde{\mathbf{D}}_{ki}^{(\ell)} - \widetilde{\mathbf{F}}_{ki}^{(\ell)}).\end{aligned}$$

3 LRT and MLRT statistics

We give the likelihood ratio

$$\lambda = \prod_{\ell=1}^{m}\prod_{k=1}^{K^{(\ell)}} \left(\frac{|\widehat{\mathbf{\Lambda}}_k^{(\ell)}|}{|\widetilde{\mathbf{\Lambda}}_k^{(\ell)}|} \right)^{\frac{1}{2}n_k^{(\ell)}}.$$

The limit distribution of likelihood ratio test statistics $-2\log\lambda$ is χ^2 distribution with $f(= p_1(p_1+3)(m-1)/2)$ degrees of freedom. Therefore, if $-2\log\lambda > \chi_f^2(\alpha)$, the hypothesis H_0 is rejected, where $\chi_f^2(\alpha)$ is the upper $100\alpha\%$ percentiles of χ^2 distribution with f degrees of freedom. For complete data, the MLRT statistics $-2\rho_n\log\lambda$ is given by Muirhead (1982, p.308), where

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

With missing data, we consider an approximation to the correction factor ρ for the MLRT statistic $(-2\rho\log\lambda)$ instead of ρ_n. From the linear interpolation, the MLRT statistic $-2\rho^*\log\lambda$, can be calculated where

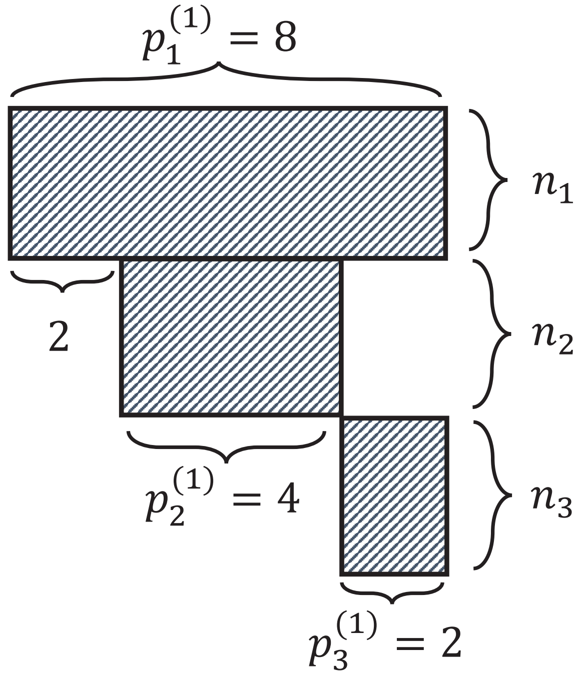
$$\begin{aligned}
\rho^* &= w\rho_n - (1-w)\rho_{n_1}, w = \frac{S^* - S_{n_1}}{S_n - S_{n_1}},\\
\rho_{n_1} &= 1 - \frac{2p_1^2+9p_1+11}{6n_1(m-1)(p_1+3)}\left(\sum_{\ell=1}^{m}\frac{n_1}{n_1^{(\ell)}} - 1\right),\ n_1 = \sum_{\ell=1}^{m} n_1^{(\ell)},\\
S^* &= \sum_{\ell=1}^{m}\sum_{k=1}^{K^{(\ell)}} p_k^{(\ell)} n_k^{(\ell)},\ S_n = p_1 n,\ S_{n_1} = p_1 n_1.
\end{aligned}$$

If the data sets have the same two-step monotone missing data, then ρ^* coincides with the correction factor ρ_L of Hosoya and Seo (2016).

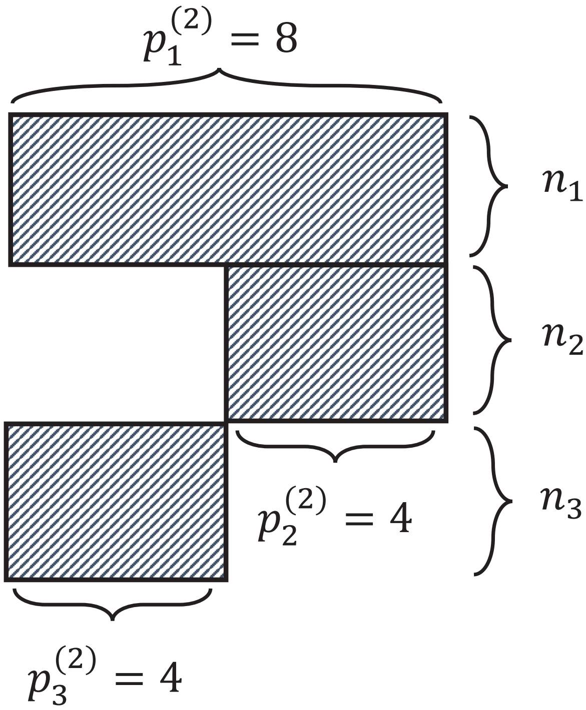
4 Simulation studies

In this section, we will evaluate the numerical accuracy of the χ^2 approximations by Monte Carlo simulation. We compute the upper $100\alpha\%$ percentiles of $-2\log\lambda$, $-2\rho_n\log\lambda$ and $-2\rho^*\log\lambda$. For each parameter, the LRT and the MLRT statistics are computed 10^5 times based on the normal random vectors generated from $N_{p_k^{(\ell)}}(\mathbf{0}, \boldsymbol{I}_{p_k^{(\ell)}}), k = 1, 2, \ldots, K^{(\ell)}, \ell = 1, 2, \ldots, m$. The data patterns are of the form:

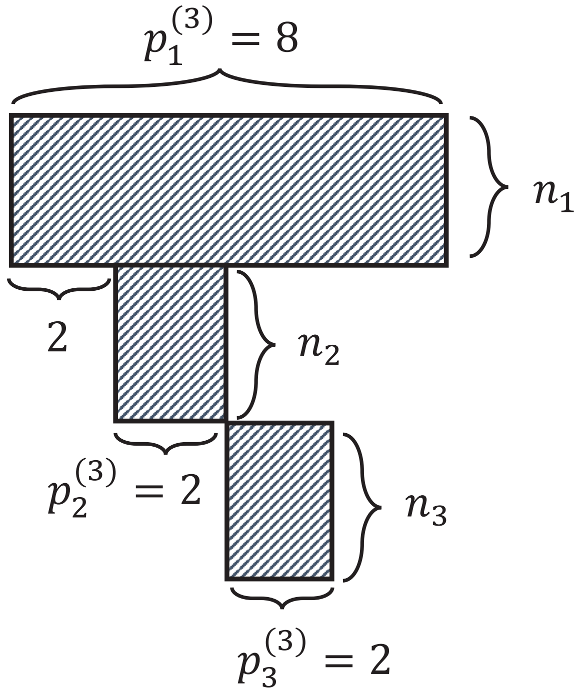
(i)



(ii)



(iii)



We compute the following cases (with significant level α=0.05, 0.01):

(I) Equal Case : $m = 2$, Pattern (i)
$(p_1^{(\ell)}, p_2^{(\ell)}, p_3^{(\ell)}) = (8, 4, 2)$,
$n_1^{(\ell)} = 30, 40, 50, 100, 200, 400, 800,\ n_2^{(\ell)} = 10, 20, 50,\ n_3^{(\ell)} = n_2^{(\ell)},\ \ell = 1, 2.$

(II) Unequal Case : $m = 2$, Pattern (i), (ii)
$(p_1^{(1)}, p_2^{(1)}, p_3^{(1)}) = (8, 4, 2),\ (p_1^{(2)}, p_2^{(2)}, p_3^{(2)}) = (8, 4, 4)$,
$n_1^{(\ell)} = 30, 40, 50, 100, 200, 400, 800,\ n_2^{(\ell)} = 10, 20, 50,\ n_3^{(\ell)} = n_2^{(\ell)},\ \ell = 1, 2.$

(III) Equal Case : $m = 3$, Pattern (i)
$(p_1^{(\ell)}, p_2^{(\ell)}, p_3^{(\ell)}) = (8, 4, 2)$,
$n_1^{(\ell)} = 30, 40, 50, 100, 200, 400, 800,\ n_2^{(\ell)} = 10, 20, 50,\ n_3^{(\ell)} = n_2^{(\ell)},\ \ell = 1, 2, 3.$

(IV) Unequal Case : $m = 3$, Pattern (i), (ii), (iii)
$(p_1^{(1)}, p_2^{(1)}, p_3^{(1)}) = (8, 4, 2),\ (p_1^{(2)}, p_2^{(2)}, p_3^{(2)}) = (8, 4, 4),\ (p_1^{(3)}, p_2^{(3)}, p_3^{(3)}) = (8, 2, 2)$,
$n_1^{(\ell)} = 30, 40, 50, 100, 200, 400, 800,\ n_2^{(\ell)} = 10, 20, 50,\ n_3^{(\ell)} = n_2^{(\ell)},\ \ell = 1, 2, 3.$

In addition, the type I error rates for the upper percentiles of $-2\log\lambda$, $-2\rho_n\log\lambda$ and

$-2\rho^* \log \lambda$ are denoted by

$$\begin{aligned}\alpha_0 &= \Pr\{-2\log\lambda > \chi^2_f(\alpha)\},\\ \alpha_n &= \Pr\{-2\rho_n\log\lambda > \chi^2_f(\alpha)\},\\ \alpha^* &= \Pr\{-2\rho^*\log\lambda > \chi^2_f(\alpha)\},\end{aligned}$$

respectively, where $\chi^2_f(\alpha)$ is the upper $100\alpha\%$ percentiles of χ^2 distribution with f degrees of freedom. Table 1 shows the simulated results of Case (I). It may be noted from Table 1 that the values are closer to the upper percentiles of the χ^2 distribution when the $n_k^{(\ell)}$, especially $n_1^{(\ell)}$ becomes large. When $n_1^{(\ell)}$ is small, type I error rates of $-2\log\lambda$ is not good, but that of $-2\rho^*\log\lambda$ can improve the accuracy relatively. In Table 2, we present the simulated results of Case (II). Even though data pattern is not the same between the samples, the similar trends can be obtained from the results. Tables 3 and 4 show the results of three sample cases (Case (III), (IV)). When $n_1^{(\ell)}$ is large, the number of samples doesn't affect the approximation accuracy. When $n_1^{(\ell)}$ is small, type I error rate of $-2\log\lambda$ is not good compared to the two sample case. Using $-2\rho^*\log\lambda$, the closer approximation accuracy can be given not depending on the number of samples. Furthermore, we tried the additional simulations changing the dimension. In perspective, for the bigger p_1, the more n_1 is needed to obtain the approximation accuracy to the same degree. Regardless of the dimension, MLRT statistics is effective.

Table 1 : The simulated values for $-2\log\lambda$, $-2\rho_n\log\lambda$ and $-2\rho^*\log\lambda$ and the type I error rates when $m=2$ and $(p_1^{(\ell)},p_2^{(\ell)},p_3^{(\ell)})=(8,4,2)$.

Sample Size			Upper Percentile			Type I error rate		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$-2\log\lambda$	$-2\rho_n\log\lambda$	$-2\rho^*\log\lambda$	α_0	α_n	α^*
				α=0.05				
30	10	10	71.533	64.672	61.813	0.215	0.097	0.063
40	10	10	68.575	63.094	61.382	0.157	0.077	0.059
50	10	10	66.859	62.279	61.134	0.127	0.067	0.056
100	10	10	63.421	60.887	60.570	0.081	0.053	0.051
200	10	10	61.993	60.641	60.557	0.064	0.051	0.050
400	10	10	61.336	60.636	60.614	0.058	0.051	0.051
800	10	10	60.832	60.476	60.471	0.053	0.050	0.050
30	20	20	71.556	66.654	62.569	0.211	0.125	0.071
40	20	20	68.329	64.233	61.673	0.153	0.092	0.061
50	20	20	66.609	63.060	61.285	0.124	0.076	0.057
100	20	20	63.421	61.249	60.706	0.081	0.057	0.052
200	20	20	61.923	60.686	60.531	0.064	0.052	0.051
400	20	20	61.256	60.588	60.547	0.057	0.051	0.051
800	20	20	60.833	60.486	60.475	0.053	0.050	0.050
40	50	50	67.828	65.505	61.875	0.145	0.108	0.064
50	50	50	66.291	64.172	61.523	0.120	0.090	0.060
100	50	50	63.354	61.835	60.886	0.079	0.063	0.053
200	50	50	61.850	60.861	60.552	0.064	0.054	0.051
400	50	50	61.202	60.615	60.524	0.057	0.051	0.050
800	50	50	60.789	60.465	60.440	0.053	0.050	0.050
				α=0.01				
30	10	10	81.137	73.355	70.113	0.076	0.025	0.013
40	10	10	77.677	71.469	69.529	0.049	0.018	0.012
50	10	10	75.912	70.711	69.411	0.037	0.015	0.012
100	10	10	72.172	69.288	68.927	0.019	0.011	0.011
200	10	10	70.467	68.931	68.835	0.014	0.011	0.010
400	10	10	69.553	68.759	68.734	0.012	0.010	0.010
800	10	10	69.137	68.733	68.727	0.011	0.010	0.010
30	20	20	81.314	75.743	71.101	0.076	0.035	0.016
40	20	20	77.439	72.797	69.896	0.047	0.022	0.013
50	20	20	75.827	71.787	69.767	0.036	0.018	0.012
100	20	20	72.171	69.699	69.081	0.020	0.012	0.011
200	20	20	70.308	68.903	68.727	0.014	0.010	0.010
400	20	20	69.493	68.736	68.689	0.012	0.010	0.010
800	20	20	68.996	68.602	68.589	0.011	0.010	0.010
40	50	50	77.026	74.387	70.265	0.043	0.028	0.014
50	50	50	75.234	72.829	69.822	0.034	0.023	0.013
100	50	50	71.908	70.184	69.107	0.019	0.014	0.011
200	50	50	70.316	69.192	68.841	0.014	0.011	0.010
400	50	50	69.645	68.977	68.872	0.012	0.011	0.010
800	50	50	69.165	68.796	68.768	0.011	0.010	0.010

Note. $\chi^2_f(0.05)=60.481, \chi^2_f(0.01)=68.710, f=44$

Table 2 : The simulated values for $-2\log\lambda$, $-2\rho_n\log\lambda$ and $-2\rho^*\log\lambda$ and the type I error rates when $m=2$ and $(p_1^{(1)}, p_2^{(1)}, p_3^{(1)}) = (8,4,2), (p_1^{(2)}, p_2^{(2)}, p_3^{(2)}) = (8,4,4)$.

Sample Size			Upper Percentile			Type I error rate		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$-2\log\lambda$	$-2\rho_n\log\lambda$	$-2\rho^*\log\lambda$	α_0	α_n	α^*
				α=0.05				
30	10	10	71.516	64.657	62.084	0.212	0.096	0.065
40	10	10	68.343	62.880	61.344	0.155	0.074	0.058
50	10	10	66.693	62.124	61.096	0.126	0.066	0.056
100	10	10	63.516	60.978	60.693	0.082	0.055	0.052
200	10	10	62.031	60.679	60.603	0.065	0.052	0.051
400	10	10	61.235	60.536	60.517	0.057	0.050	0.050
800	10	10	60.838	60.482	60.477	0.053	0.050	0.050
30	20	20	71.080	66.210	62.558	0.205	0.119	0.071
40	20	20	68.031	63.953	61.659	0.150	0.089	0.061
50	20	20	66.499	62.955	61.361	0.124	0.076	0.059
100	20	20	63.544	61.368	60.878	0.082	0.058	0.054
200	20	20	61.840	60.604	60.465	0.063	0.051	0.050
400	20	20	61.264	60.596	60.559	0.058	0.051	0.051
800	20	20	60.683	60.337	60.327	0.052	0.049	0.049
40	50	50	67.559	65.245	61.991	0.143	0.106	0.065
50	50	50	66.031	63.920	61.545	0.117	0.088	0.061
100	50	50	63.315	61.797	60.943	0.080	0.062	0.054
200	50	50	61.879	60.890	60.612	0.063	0.054	0.051
400	50	50	61.288	60.701	60.618	0.057	0.052	0.051
800	50	50	61.003	60.678	60.655	0.054	0.052	0.051
				α=0.01				
30	10	10	81.411	73.603	70.675	0.075	0.026	0.015
40	10	10	77.870	71.647	69.896	0.047	0.018	0.013
50	10	10	75.861	70.664	69.495	0.036	0.015	0.012
100	10	10	72.049	69.170	68.846	0.019	0.011	0.010
200	10	10	70.421	68.886	68.799	0.014	0.010	0.010
400	10	10	69.383	68.591	68.568	0.012	0.010	0.010
800	10	10	69.422	69.016	69.010	0.011	0.011	0.011
30	20	20	80.738	75.207	71.059	0.071	0.033	0.016
40	20	20	77.405	72.765	70.155	0.045	0.022	0.013
50	20	20	75.635	71.605	69.791	0.035	0.017	0.012
100	20	20	72.206	69.732	69.176	0.020	0.012	0.011
200	20	20	70.353	68.947	68.789	0.014	0.011	0.010
400	20	20	69.452	68.695	68.653	0.012	0.010	0.010
800	20	20	69.047	68.653	68.642	0.011	0.010	0.010
40	50	50	76.916	74.281	70.576	0.042	0.029	0.014
50	50	50	74.851	72.458	69.766	0.032	0.022	0.013
100	50	50	72.022	70.295	69.324	0.019	0.014	0.011
200	50	50	70.434	69.308	68.992	0.014	0.011	0.011
400	50	50	69.554	68.887	68.793	0.012	0.010	0.010
800	50	50	69.480	69.110	69.084	0.012	0.011	0.011

Note. $\chi_f^2(0.05) = 60.481, \chi_f^2(0.01) = 68.710, f = 44$

Table 3 : The simulated values for $-2\log\lambda$, $-2\rho_n\log\lambda$ and $-2\rho^*\log\lambda$ and the type I error rates when $m=3$ and $(p_1^{(\ell)}, p_2^{(\ell)}, p_3^{(\ell)}) = (8, 4, 2)$.

Sample Size			Upper Percentile			Type I error rate		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$-2\log\lambda$	$-2\rho_n\log\lambda$	$-2\rho^*\log\lambda$	α_0	α_n	α^*
				α=0.05				
30	10	10	128.928	117.936	113.356	0.276	0.113	0.069
40	10	10	123.945	115.139	112.387	0.195	0.084	0.061
50	10	10	121.044	113.673	111.830	0.153	0.070	0.056
100	10	10	115.744	111.632	111.118	0.090	0.055	0.051
200	10	10	113.378	111.181	110.962	0.068	0.052	0.051
400	10	10	112.175	111.037	110.980	0.059	0.051	0.051
800	10	10	111.588	111.008	110.994	0.055	0.051	0.051
30	20	20	128.348	120.532	114.019	0.266	0.146	0.074
40	20	20	123.551	116.968	112.854	0.187	0.103	0.065
50	20	20	120.746	115.027	112.167	0.147	0.083	0.059
100	20	20	115.611	112.091	111.211	0.090	0.059	0.052
200	20	20	113.442	111.427	111.024	0.069	0.054	0.051
400	20	20	112.133	111.046	110.938	0.059	0.051	0.050
800	20	20	111.521	110.955	110.927	0.054	0.050	0.050
40	50	50	122.682	118.947	113.110	0.175	0.124	0.066
50	50	50	120.143	116.729	112.461	0.141	0.101	0.061
100	50	50	115.419	112.959	111.422	0.087	0.065	0.054
200	50	50	113.335	111.724	110.919	0.067	0.055	0.050
400	50	50	112.135	111.180	110.941	0.059	0.052	0.050
800	50	50	111.611	111.083	111.017	0.055	0.051	0.051
				α=0.01				
30	10	10	141.563	129.494	124.466	0.108	0.031	0.015
40	10	10	136.208	126.531	123.507	0.064	0.020	0.013
50	10	10	132.739	124.656	122.636	0.046	0.016	0.011
100	10	10	126.912	122.404	121.840	0.022	0.011	0.010
200	10	10	124.571	122.157	121.916	0.016	0.011	0.010
400	10	10	123.205	121.954	121.892	0.013	0.010	0.010
800	10	10	122.428	121.792	121.776	0.011	0.010	0.010
30	20	20	141.027	132.439	125.282	0.101	0.043	0.009
40	20	20	135.293	128.084	123.579	0.062	0.026	0.013
50	20	20	132.372	126.103	122.968	0.044	0.020	0.012
100	20	20	126.721	122.862	121.898	0.022	0.012	0.010
200	20	20	124.498	122.287	121.845	0.015	0.011	0.010
400	20	20	123.392	122.197	122.077	0.013	0.011	0.011
800	20	20	122.679	122.056	122.025	0.012	0.011	0.010
40	50	50	134.806	130.702	124.288	0.056	0.035	0.015
50	50	50	131.925	128.176	123.490	0.041	0.026	0.013
100	50	50	126.937	124.232	122.541	0.022	0.015	0.011
200	50	50	124.506	122.737	121.852	0.015	0.012	0.010
400	50	50	123.566	122.512	122.249	0.013	0.011	0.011
800	50	50	122.796	122.214	122.141	0.012	0.011	0.011

Note. $\chi^2_f(0.05) = 110.898, \chi^2_f(0.01) = 121.767, f = 88$

Table 4 : The simulated values for $-2\log\lambda$, $-2\rho_n\log\lambda$ and $-2\rho^*\log\lambda$ and the type I error rates when $m=3$ and $(p_1^{(1)},p_2^{(1)},p_3^{(1)})=(8,4,2)$, $(p_1^{(2)},p_2^{(2)},p_3^{(2)})=(8,4,4)$, $(p_1^{(3)},p_2^{(3)},p_3^{(3)})=(8,2,2)$.

Sample Size			Upper Percentile			Type I error rate		
$n_1^{(\ell)}$	$n_2^{(\ell)}$	$n_3^{(\ell)}$	$-2\log\lambda$	$-2\rho_n\log\lambda$	$-2\rho^*\log\lambda$	α_0	α_n	α^*
				α=0.05				
30	10	10	129.001	118.003	113.421	0.277	0.115	0.069
40	10	10	123.933	115.129	112.377	0.195	0.085	0.061
50	10	10	121.270	113.885	112.039	0.152	0.072	0.058
100	10	10	115.572	111.467	110.953	0.089	0.054	0.050
200	10	10	113.364	111.168	111.031	0.069	0.052	0.051
400	10	10	111.942	110.806	110.771	0.057	0.049	0.049
800	10	10	111.413	110.834	110.825	0.054	0.050	0.050
30	20	20	128.176	120.371	113.866	0.264	0.144	0.073
40	20	20	123.299	116.729	112.623	0.184	0.100	0.063
50	20	20	120.638	114.924	112.068	0.150	0.083	0.058
100	20	20	115.684	112.162	111.281	0.090	0.059	0.053
200	20	20	113.249	111.238	110.986	0.067	0.052	0.051
400	20	20	112.065	110.980	110.912	0.058	0.051	0.050
800	20	20	111.478	110.913	110.895	0.054	0.050	0.050
40	50	50	122.714	118.977	113.139	0.177	0.127	0.068
50	50	50	120.325	116.905	112.631	0.141	0.101	0.063
100	50	50	115.519	113.056	111.518	0.088	0.066	0.054
200	50	50	113.216	111.608	111.105	0.067	0.054	0.051
400	50	50	112.032	111.077	110.928	0.058	0.051	0.050
800	50	50	111.555	111.027	110.986	0.054	0.051	0.051
				α=0.01				
30	10	10	141.827	129.736	124.698	0.109	0.031	0.016
40	10	10	136.257	126.577	123.552	0.064	0.020	0.013
50	10	10	133.047	124.945	122.919	0.047	0.017	0.012
100	10	10	127.258	122.738	122.173	0.022	0.012	0.011
200	10	10	124.523	122.111	121.960	0.016	0.011	0.010
400	10	10	123.059	121.810	121.771	0.012	0.010	0.010
800	10	10	122.299	121.664	121.654	0.011	0.010	0.010
30	20	20	140.851	132.273	125.126	0.102	0.042	0.017
40	20	20	135.405	128.190	123.681	0.060	0.026	0.013
50	20	20	132.661	126.378	123.236	0.044	0.020	0.012
100	20	20	127.094	123.224	122.257	0.022	0.013	0.011
200	20	20	124.355	122.146	121.870	0.015	0.011	0.010
400	20	20	123.216	122.022	121.947	0.013	0.010	0.010
800	20	20	122.386	121.765	121.746	0.011	0.010	0.010
40	50	50	134.666	130.566	124.159	0.057	0.035	0.015
50	50	50	131.860	128.113	123.429	0.042	0.026	0.013
100	50	50	126.844	124.141	122.451	0.021	0.014	0.011
200	50	50	124.415	122.647	122.095	0.015	0.012	0.011
400	50	50	123.143	122.093	121.929	0.013	0.011	0.010
800	50	50	121.947	121.369	121.324	0.010	0.009	0.009

Note. $\chi_f^2(0.05)=110.898, \chi_f^2(0.01)=121.767, f=88$

5 Conclusion

We developed the approximate upper percentiles of the LRT statistic and the MLRT statistic for tests of equality of multivariate normal populations with general missing data. The null distribution of the MLRT statistic $-2\rho^* \log \lambda$ presented in this paper has considerably good approximation to the χ^2 distribution even when the sample size is small.

Acknowledgments

The second and third authors' research is partly supported by a Grant-in-Aid for Young Scientists (JSPS KAKENHI Grant Number JP19K20225) and a Grant-in-Aid for Scientific Research (C) (JSPS KAKENHI Grant Number JP17K00058), respectively.

References

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

[2] Hosoya, M. and Seo, T. (2015). Simultaneous testing of the mean vector and the covariance matrix with two-step monotone missing data. *SUT Journal of Mathematics*, **51**, 83-98.

[3] Hosoya, M. and Seo, T. (2016). On the likelihood ratio test for the equality of multivariate normal populations with two-step monotone missing data. *Journal of Statistical Theory and Practice*, **10**, 673-692.

[4] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Willey, New York.

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

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

[7] Srivastava, M. S. (2002). *Methods of Multivariate Statistics.* Willey, New York.

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

[9] Yagi, A., Yamaguchi, R. and 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] Zhang, L., Xu, X. and Chen, G. (2012). The exact likelihood ration test for equality of two normal populations. *The American Statistician,* **66**, 180-184.